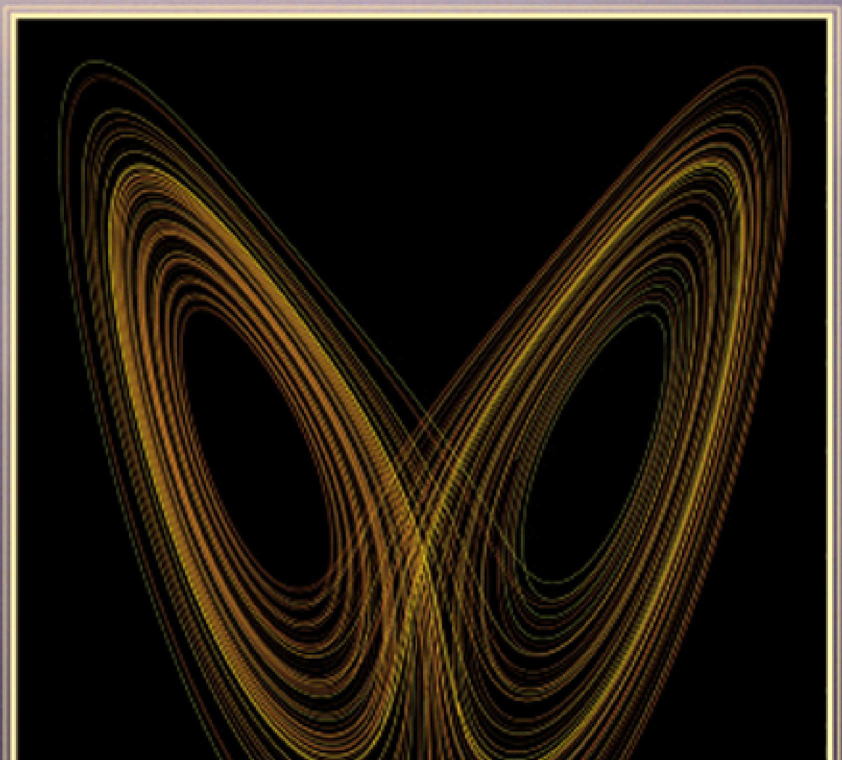
Handbook of
THE PHILOSOPHY OF SCIENCE

General Editors: DOV M. GABBAY, PAUL THAGARD, AND JOHN WOODS

PHILOSOPHY *of* COMPLEX SYSTEMS



Edited by Cliff Hooker

Philosophy of Complex Systems

Handbook of the Philosophy of Science

General Editors

Dov M. Gabbay
Paul Thagard
John Woods

Cover picture shows the Lorenz attractor: Projection of trajectory of Lorenz system in phase space with "canonical" values of parameters r=28, σ = 10, b = 8/3. Reprinted under GNU Free Documentation License.

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
North Holland is an imprint of Elsevier

N·H

Handbook of the Philosophy of Science

Volume 10

Philosophy of Complex Systems

Edited by
Cliff Hooker
Emeritus Professor,
University of Newcastle,
Australia

ELSEVIER
AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
North Holland is an imprint of Elsevier
N·H

North Holland is an imprint of Elsevier
The Boulevard, Langford lane, Kidlington, Oxford, OX5 1GB, UK
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
225 Wyman Street, Waltham, MA 02451, USA

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-444-52076-0
ISSN: 0031-8019

For information on all North Holland publications
visit our web site at *elsevierdirect.com*

Printed and bound in Great Britain

11 11 10 9 8 7 6 5 4 3 2 1

GENERAL PREFACE

Dov Gabbay, Paul Thagard, and John Woods

Whenever science operates at the cutting edge of what is known, it invariably runs into philosophical issues about the nature of knowledge and reality. Scientific controversies raise such questions as the relation of theory and experiment, the nature of explanation, and the extent to which science can approximate to the truth. Within particular sciences, special concerns arise about what exists and how it can be known, for example in physics about the nature of space and time, and in psychology about the nature of consciousness. Hence the philosophy of science is an essential part of the scientific investigation of the world.

In recent decades, philosophy of science has become an increasingly central part of philosophy in general. Although there are still philosophers who think that theories of knowledge and reality can be developed by pure reflection, much current philosophical work finds it necessary and valuable to take into account relevant scientific findings. For example, the philosophy of mind is now closely tied to empirical psychology, and political theory often intersects with economics. Thus philosophy of science provides a valuable bridge between philosophical and scientific inquiry.

More and more, the philosophy of science concerns itself not just with general issues about the nature and validity of science, but especially with particular issues that arise in specific sciences. Accordingly, we have organized this Handbook into many volumes reflecting the full range of current research in the philosophy of science. We invited volume editors who are fully involved in the specific sciences, and are delighted that they have solicited contributions by scientifically-informed philosophers and (in a few cases) philosophically-informed scientists. The result is the most comprehensive review ever provided of the philosophy of science.

Here are the volumes in the Handbook:

Philosophy of Science: Focal Issues, edited by Theo Kuipers.

Philosophy of Physics, edited by Jeremy Butterfield and John Earman.

Philosophy of Biology, edited by Mohan Matthen and Christopher Stephens.

Philosophy of Mathematics, edited by Andrew Irvine.

Philosophy of Logic, edited by Dale Jacquette.

Philosophy of Chemistry and Pharmacology, edited by Andrea Woodsy, Robin Hendry and Paul Needham.

Philosophy of Statistics, edited by Prasanta S. Bandyopadhyay and Malcolm Forster.

Philosophy of Information, edited by Pieter Adriaans and Johan van Benthem.

Philosophy of Technology and Engineering Sciences, edited by Anthonie Meijers.

Philosophy of Complex Systems, edited by Cliff Hooker.

Philosophy of Ecology, edited by Bryson Brown, Kent Peacock and Kevin de Laplante.

Philosophy of Psychology and Cognitive Science, edited by Paul Thagard.

Philosophy of Economics, edited by Uskali Mäki.

Philosophy of Linguistics, edited by Ruth Kempson, Tim Fernando and Nicholas Asher.

Philosophy of Anthropology and Sociology, edited by Stephen Turner and Mark Risjord.

Philosophy of Medicine, edited by Fred Gifford.

Details about the contents and publishing schedule of the volumes can be found at http://www.elsevier.com/wps/find/bookdescription.cws_home/BS_HPHS/description#description

As general editors, we are extremely grateful to the volume editors for arranging such a distinguished array of contributors and for managing their contributions. Production of these volumes has been a huge enterprise, and our warmest thanks go to Jane Spurr and Carol Woods for putting them together. Thanks also to Lauren Schulz and Gavin Becker at Elsevier for their support and direction.

CONTRIBUTORS

Adele Abrahamsen
University of California, San Diego, USA.
aabrahamsen@gmail.com

Peter M. Allen
Cranfield University, UK.
p.m.allen@cranfield.ac.uk

Xabier Barandiaran
University of Sussex, UK.
xabier.academic@barandiaran.net

William Bechtel
University of California, San Diego, USA.
bill@mechanism.ucsd.edu

Mark H. Bickhard
Lehigh University, USA.
mhb0@lehigh.edu

Robert C. Bishop
Wheaton College, USA.
robert.c.bishop@wheaton.edu

Thomas S. Brinsmead
Commonwealth Scientific and Industrial Research Organisation, Australia.
thomas.brinsmead@csiro.au

James A. Coffman
Mount Desert Island Biological Laboratory, USA.
jcoffman@mdibl.org

Sean S. Downey
Stanford University, USA.
sean@codexdata.com

John Foster
University of Queensland, Australia.
j.foster@uq.edu.au

Yin Gao
University of Newcastle, Australia.
yin.gao@newcastle.edu.au

David G. Green
Monash University, Asutralia.
david.green@infotech.monash.edu.au

William Harms
Philosophy Instructor and Independent Scholar, USA.
williamfharms@gmail.com

William Herfel
University of Newcastle, Australia.
wmherfel@yahoo.com.au

Wolfgang Hofkirchner
Vienna University of Technology, Austria.
wolfgang.hofkirchner@tuwien.ac.at

Cliff Hooker
University of Newcastle, Australia.
cliff.hooker@newcastle.edu.au

Heidi Kloos
University of Cincinnati, USA.
heidi.kloos@uc.edu

J. Stephen Lansing
University of Arizona, Stockholm Resilience Centre and Santa Fe Institute, USA.
lansing@santafe.edu

Tania Leishman
Monash University, Australia.
tania.bransden@infotech.monash.edu.au

Michael D. Mastrandrea
Stanford University, USA.
mikemas@stanford.edu

Alvaro Moreno
University of the Basque Country, Spain.
alvaro.moreno@ehu.es

Allan Muir
City University, London, UK.
allan.muir1@virgin.net

Stuart A. Newman
New York Medical College, USA.
newman@nymc.edu

S. Nolfi
Institute of Cognitive Sciences and Technologies, CNR, Italy.
stefano.nolfi@istc.cnr.it

Jay Odenbaugh
Lewis and Clark College, USA.
jay@lclark.edu;odenbaugh@gmail.com

Dean Rickles
University of Sydney, Australia.
d.rickles@usyd.edu.au

D. J. Rodrigues
Newcastle, Australia.
drod3524@uni.sydney.edu.au

Kepa Ruiz-Mirazo
University of the Basque Country, Spain.
kepa.ruiz-mirazo@ehu.es

Alex J. Ryan
School of Advanced Military Studies, Command and General Staff College, USA.
alex.ryan@gmail.com

Matthias Schafranek
University of Salzburg, Austria.
matthias.schafranek@sbg.ac.at

Jan C. Schmidt
Darmstadt University of Applied Sciences, Germany.
jan.schmidt@h-da.de

Stephen H. Schneider
Stanford University, USA.
shs@stanford.edu

Adam Sheya
Indiana University, USA.
aasheya@gmail.com

Linda B. Smith
Indiana University, USA.
smith4@indianda.edu

Carolyn W. Snyder
Stanford University, USA.
carolyn.snyder@stanford.edu

Guy C. Van Orden
University of Cincinnati, USA.
vanordg@ucmail.uc.edu

Sebastian Wallot
University of Cincinnati, USA
wallotsn@mail.uc.edu

Olaf Wolkenhauer
University of Rostock, Germany.
olaf.wolkenhauer@uni-rostock.de

CONTENTS

CONTENTS

Part I

General Foundations

INTRODUCTION TO PHILOSOPHY OF COMPLEX SYSTEMS: A

Cliff Hooker

Part A: Towards a framework for complex systems

1 INTRODUCTION

Every essay in this book is original, often highly original, and they will be of interest to practising scientists as much as they will be to philosophers of science — not least because many of the essays are by leading scientists who are currently creating the emerging new complex systems paradigm. This is no accident. The impact of complex systems on science is a recent, ongoing and profound revolution. But with a few honourable exceptions, it has largely been ignored by scientists and philosophers alike as an object of reflective study.[1] Hence the scientist participants; while the small band of concerned philosophers is well represented in this volume, scientists in the midst of creating the revolution are often better placed to reflect on it than others. (Needless to add, many more of both were invited than ultimately participated.)

In consequence, I have sired a cross-border bastard, properly belonging to neither philosophy or science but an inheritor, and supporter, of both. Being an ex-scientist turned philosopher, it is a bastard of my own kind. No matter its perceived legitimacy, a bastard is always a sign of fruitful productivity. And in this case the offspring is needed, for it is born into a time of revolutionary foment

[1]For a substantial introduction see [Scott, 2004]. Scientists of course just get on with making and using specific innovations as needed, and that suffices. Honorable exceptions who have also reflected on the issues involved include the scientists Kenneth Boulding, Herbert Simon and Richard Levins (see wwweb information) and among philosophers include William Wimsatt and myself, all active across the last 4+ decades that mark the explosion of complex systems into the heartland of science. Wimsatt's work goes back to his early concerns with modelling methodology and explanation when confronted with complex systems and to some extent with systems ontology, with redduction/emergence bridging between them, now all reprinted and summed up in [Wimsatt, 2007]. [Hooker, 1978] showed that when the then-standard philosophical arguments about sense data in perception were embedded in a dynamical systems process setting, as the science already required, it re-cast them so as to make obvious their defects, [Hooker, 1981b, Part III] provided a first analysis of functional reduction in complex systems (subsequently updated in [Hooker, 2004; cf. this volume]) and [Hooker, 1995] a first attempt to recast philosophy of science itself as a dynamical systems process, see section 4.1.3 below.

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

whose constructive resolution is best eased forward by straddling both the philosophy and the science. This bastard will, I hope, be widely adopted and in turn prove fertile.

This present essay and its matching closing essay ([Hooker-b, this volume][2]) are intended to be complementary and between them provide at least a first presentation of an intellectual framework for understanding the foundational and philosophical issues raised by the complex systems revolution. The present essay is designed to introduce and broadly review the domain of complex systems, with an eye to identifying the historical setting (section 2), the key systems properties at issue (section 3) and a collection of sub-domains that do not receive treatment in a dedicated essay (section 4). The closing essay is an attempt to systematically survey the specific components and issues that make up a scientific paradigm (section 5) and philosophy of science (section 6) that together comprise a foundational/philosophical analysis of the role of complex systems in science, as they currently appear.

Readers at least somewhat familiar with complex systems should find the essays to follow reasonably accessible, with references that invite further exploration. Those entering the field for the first time might like to first consult one or more of the books referenced at note 6 below and/or the websites referenced at [Hooker-b, this volume, note 15] (or any of the hundreds of other instructive books and websites available).

Ultimately, the goal is to develop mature foundations/philosophy of complex systems. But attempting this is premature at this time. First, despite enormous progress over the past 30 years, there is no unified science of complex systems. There are increasingly general insights and frameworks, the mathematical and operational mastery of chaos provides an early example, the current emergence of generalised network dynamics provides a contemporary example (cf. [Green and Leishman, this volume]). However, by themselves these do not a completed science make (cf. sections 4.2.6, 5.1.1 below), and at present they remain largely separate (if with multiple specific links — a typically complex situation!). Around them lie a patchwork of specific models and applications that presently remain irreducibly various. One aspect of the variability is the variety of complex systems phenomena engaged: in one application it may be counter-intuitive dynamics — such as the survival of cooperation in a sea of cutthroat competition — in another, self-organisation — such as rhythmic entrainment among food-stressed slime mould amoebae — in still another the onset of chaos — such as in local climate fluctuations — and so on. Another aspect of the variability is that characterising complex system principles is often a 'wicked' problem where the key dynamics generating a phenomenon is itself a function of the application conditions. To take a simple

[2]Other essays in this volume are indicated by '[Name, this volume]' and readers should turn directly to the indicated essay. The essays by Hooker in the volume are indicated by '[Hooker-a; b; c, this volume]', a = present essay, b = closing essay, c = reduction/emergence essay. These references to Hooker in this volume are also entered in the essay bibliography to disambiguate the intended reference.

example, traffic jams on expressways may be caused by any of entry/exit rates, truck/car proportions, flow density, driver pathway correlations, etc. Moreover, the dynamics of jam formation for each of these conditions is significantly different. For instance, truck/car speed differential is important near lane-change originated jams but less important for high density braking-originated jams, and unimportant for pathway convergence jams. So there is no usefully generalisable, detailed dynamical rules for traffic jam formation. In sum, the empirical domain of complex systems is itself complex — at this time irreducibly complex![3]

Irrespective of this developmental complexity, let us be clear about the extent of the complex systems revolution now taking place. When I trained in science (physics — PhD 1968) the contemporary icon for complex systems, chaos, was in process of discovery and few or no courses on complex systems were offered, those few problems considered involved several interacting bodies and were boxed as idiosyncratic special cases of applied mathematics. Today (2009) many of the most prominent scientific disciplines could not exist but for the complex systems models and methods on which they depend, among them synthetic and systems biology, climate science, control engineering, neurophysiology, developmental neuropsychology, astrophysics, geo-dynamics, traffic engineering, ... (cf. [Scott, 2004]). And there cannot be a single scientific discipline that has not now felt the complex systems winds of change blow through it to some extent — as this book testifies this now applies even to anthropology, Chinese medicine and warfare. The very way that science is transforming itself as complex systems penetrates it, is itself an excellent example of complex systems emergence through self-organisation, and one that, like many other instances, is re-defining the relationships between the emergent entity and the encompassing environment from which it emerged (see also section 4.1.3 below).

The scale and sweep of the change is truly vast — entire disciplines or subdisciplines have come into being, departments and institutes of hundreds of scientists now exist that did not exist two decades ago, and Google entries under complex systems headings run into the tens of millions of pages — far too many for any individual to consult, even in a lifetime, thus creating an emergent reliance on systemic institutional structure and processes to generate scientific coherence, yet another facet of complex systems in science. (The same institutional structure needs to construct effective quality controls for generating this information deluge or, speaking impolitely, it is often a major challenge to winnow insight from the false, trivial and groundlessly speculative and the powerful crap detector required by all new, fecund fields is required here too.) Complementing this, policy analysis, industrial development and large scale financial planning all require complex systems modelling while vast enterprises in bio-agriculture, bio-medicine, manufacturing design and so on have arisen and flourished on the backs of these

[3]And here a philosophical issue already raises its head: contrary to common opinion, a general model/theory of traffic jams will evidently be vaguer and less empirically precise and less explanatory than any of its specific sub-cases. So is scientific unification not to be preferred? See further [Hooker-b, this volume, section 6.2.6].

developments and are steadily transforming our lived world. Increasingly, people whose education does not include relevant competency in complex systems are excluded from science, policy and large scale business or find themselves increasingly dependent on those who have it.

Nor should the depth of the complex systems revolution be under-estimated. As the essays in this volume (including this one) attest, complex systems impacts every aspect of a science, from experimental design, what counts as evidence and the treatment of errors, through basic theoretical concepts of component, interaction, individuality, equilibrium (especially of dynamic versus static equilibrium), organisation and self-organisation, dynamic thresholds and irreversible form transitions, to deep limits on prediction and control and the relations of laws, explanation and confirmation to realisation conditions. (These notions will be explained in what follows.) And on a larger scale complex systems dynamical models naturally facilitate new disciplinary foci (such as systems biology) and new interdisciplinary interrelationships (such as synthetic biology and computational anthropology) while at the same time raising foundational issues concerning complexity and the representation of dynamics. In short, the complex systems-driven revolution is as deep as the revolutions in physics a century ago, but much wider in impact, even if they do not disturb our sense of fundamental reality in the same way.

What best to do? We have a revolution occurring in our midst, moreover one too complex and widespread for any one individual (a fortiori for me) to master in any detail across the board, and as yet lacking in any settled, or even well established, philosophical framework. Of course, I have tried to engage the sparse relevant philosophers (inevitably, not always successfully, for the usual practical reasons) and there are excellent essays herein that bear witness to those efforts. However, I had anyway concluded that, in the circumstances, to come to proper grips with the revolution and its challenges it would be necessary to engage the scientists themselves in the enterprise of reflecting on their own activities as they willy nilly develop complex systems based science. They are in the best position to comment on what the complex systems revolution involves for their discipline, and what its prospects are, and will remain so for many decades to come, even while complex systems philosophy of science develops.

Engaging eminent scientists has typically proven still more difficult than it was for philosophers, and understandably so: they are not only busy teaching, publishing and (these days) administering, as philosophers are, but have in addition to run their laboratories and field studies; moreover, they are rewarded for producing science, not for reflecting on that production, and it is often an intellectual wrench to take up the latter approach. Nevertheless, many scientists have willingly given their time and intellectual energy and many of the outstanding essays of this volume — essays that in themselves break new ground — have resulted from their collaboration.

In consequence, I have often played a more active role as editor than would be typical if this were a standard philosophy work, although many of the philosopher authors too have happily engaged in rounds of drafting discussion. Editorial activism was always welcomed (and often at author request) as especially scientists sought (sometimes 'fought' would be more apt) to shift to the reflective mode and bring a complex maze of practical activity into focus. In doing so I have not sought to dictate the content of an essay (this attempt would anyway have been self-defeating) but to assist authors to organise, clarify and enlarge upon what they intuitively want to contribute. That collaboration often ran through many drafts, and it has been one of the most rewarding aspects of editing the volume. I have learned a lot along the way and had the privilege and joy of some of the most stimulating discussions of my career. My deep thanks to all those, scientists and philosophers, who gave so unstintingly to this pioneering volume.

The result is a volume quite different from the bulk of publishing in the area which typically focuses on a range of technical articles by scientists either developing particular techniques or applying them to practical situations, all material that could equally appear in the relevant disciplinary journals. Sometimes there will be added to front or rear a few very general articles about the 'complexity world view' or the like, at home in any general cultural discussion. This is true, e.g., of the recent Handbook on Simulating Social Complexity and the Encyclopedia of Complexity and Systems Science, both 2009 and no criticism of their primary aims and content. However, this volume fits instead exactly into the intermediate space left by such efforts: the detailed reflective discussion of the differences that employing the concepts, principles and methods of complex systems makes to the methodology and theorising of those sciences and the challenges posed to scientific metaphysics, epistemology and methodology arising therefrom. All the while it retains the connection to current examples and practices to vivify and discipline the reflections. At this stage it is premature to attempt a fully developed philosophy of science for the use of complex systems modelling, what might be achieved is briefly and schematically surveyed in [Hooker-b, this volume]. What this book does is provide the proper preliminary foundation of rich, higher order reflection on the changes and challenges as they are currently experienced in the sciences, material from which a more mature view will eventually emerge in the fullness of time — as is only fitting for a complex framework emerging from the complex adaptive process that is science.

Shalizi [2006] distinguishes four components to the content of complex systems work: **patterns** — the classification and study of the characteristic patterns in state space, e.g. period doubling before chaos onset, made by the trajectories of complex systems when they are displaying their distinctive complex dynamical behaviours; **topics** — the array of complex systems features (e.g. chaos) and exemplary cases (e.g. logistic population dynamics) frequently discussed; **tools** — the mathematical methods for data analysis that are appropriate for analysing data pertaining to complex dynamics, e.g. data mining techniques to discover relationships, especially in high dimensional, low density data distributions, and time series analysis (see [Rickles, this volume]); we are some distance yet from general tools for revealing system structure, though particular methods are developing for local conditions; **foundations** — the mathematical foundations of complex systems dynamics, unfortunately currently very patchy, confined to theoretical fragments and particular applied models (see section 5 below). While this book is closest to the tools component, it is not primarily about any of these components — although some of its essays employ examples from patterns and topics and discuss issues that the tools and foundations raise. Rather, it is a reflection on what is distinctive to the conduct of science, especially as it pertains to metaphysics, epistemology and method, what this might amount to foundationally/philosophically, and what challenges are thus posed to science and philosophy.

Needless to say, there is yet much missing from this volume that would legitimately belong to it. Many domains contain numerous sub-domains of application of complex systems models. Biology, e.g., includes at least genome, cellular, multicellular, developmental, ecological and evolutionary modelling with further intra- and inter- sub-domain modelling specialities. But even were their separate inclusion practical (it is not) it would typically not add to the basic reflective issues thrown up. However, because of their prominence in their disciplines, I particularly regret the absence of essays on dynamical models in engineering, chemistry and social group dynamics. These and other disciplinary essays (e.g. on geodynamics, physiology, archaeology, business management and science itself) were vigorously sought but for one reason or another failed to materialise. While we lose the richness of their idiosyncratic responses to the entry of complex systems to their fields, most or all of the more general issues involved will have been well covered by the other essays. Some brief comments are made on most of these missing items in section 4.2 below, plus 4.1.3 (science) and [Hooker-b, this volume, section 5.3] (archaeology). Balancing these absences, on the other hand, are strong essays on Anthropology, Traditional Chinese Medicine, Military Policy and Planning and Public Policy and Planning that represent welcome surprises, areas beyond those normally attended to where the application of complex systems is delivering genuine insight and practical advantage.

Equally regretful is the absence of an essay on the primary mathematical issue raised by complex systems theory, namely the challenge it poses to standard analytical mathematical dynamics and the correlative disarray in unified mathematical underpinnings for complex systems. In my view, the suppression of this issue in many contemporary textbooks on analytical mechanics constitutes something of an intellectual scandal. There is currently no coherent mathematical framework for complex systems theory, as noted earlier there is instead a collection of diverse specific complex systems models and an equally diverse range of at best weakly interrelated mathematical research groups (see also 4.2.1 below). Perhaps this explains why it proved impossible to persuade anyone to write about it (and often even to respond), despite many invitations, especially to leading mathematicians. This is to be particularly regretted since the situation can instead be regarded as a stimulating challenge to future research. At any event the issue is unavoidable when trying to understand the nature of many key complex systems features, such as self-organised emergence. The basic issue is thus presented briefly, and within my limitations, in the essay on reduction and emergence in complex systems ([Hooker-c this volume], cf. [Bickhard, this volume; Bishop, this volume]). But its consequences re-appear throughout the review of our several foundational ignorances in section 5.

This essay plus its complementary closing essay attempts to provide what is possibly the first comprehensive review of the philosophical-cum-foundational issues deriving from the impact of complex systems in science. The review is consciously schematic, deferring to the other essays herein to fill in domain and discipline specifics. My intention is to clarify where possible and start as many hares running as possible. And there are plenty of the latter, the essays are awash with topics worth investigating. Even so, they are also certainly incomplete.

Finally, while there will be plenty of controversial issues arising, I hope for the most part that my authorship per se plays no major intervening role, except perhaps in the reach/lack-of-reach and synthesis of material presented. Here there is a clear bias in these essays toward thinking in physical terms and using physical examples. I am trained as a physicist and there are fewer social examples than another researcher might use. (But natural science examples, being uncluttered with the complications of human capacities, are often also clearer.) The keenest lack is not a matter of domain emphasis, which I think has small effect here, but of tools: there is a relative absence of computational analyses and computational examples. Mea culpa, I can only point out that I do recognise computational analyses on occasion and refer the reader to the review by [Green and Leishman, this volume] for a glimpse of how remaining bias might be corrected. Beyond this, as fair forewarning to readers, my philosophical orientation makes itself felt in two particular ways in what follows. First, there is my preference for dynamical analysis over logical and semantic analysis when it comes to fundamental concepts and principles. I deeply believe that, especially in the domain of complex systems science, this is the only productive way forward. Correlatively, second, there is an occasional willingness to close off philosophical debate where I judge science has provided a clear dynamical indication of how best to proceed (but always flagged, so that debate may ensue).

2 HISTORICAL NOTES ON THE DEVELOPMENT OF COMPLEX SYSTEMS IN SCIENCE[4]

Preparations for the emergence of complex systems in science — both their study and use — have been over 150 years in the making. Even so, the explicit recognition of complex systems concepts, principles and models in science is a recent phenomenon. And the explosion of its application into widespread use, with all the consequences explored in this volume, is a still more recent matter of the last 25 years. Especially because of its sudden, late emergence its history is worth

[4]The following historical discussion is excerpted and adapted with modifications and additions from [Hooker, 2009a, section C.5, 6], where it first appeared.

appreciating. Science has many long-term sub-processes running through it, not all fully coherent with its public consensus at any one time (or even in any one century). Any of these can unexpectedly burst forth to transform the way science is thought about and done. The rise of complex systems has been one of the more important of those processes. This is a reflection of the deep complexity of the science process, a self-reflexive application of complex systems ideas that has yet to reach a mature expression.[5] It is perhaps useful then to briefly rehearse a little of this historical background here. In what follows I stick to a simplified but multi-disciplinary, conceptual focus; there are plenty of other, complementary works that capture particular developments in more detail.[6]

The story is best told by beginning with the prior physics-based framework that essentially excluded complex systems, making the latter's appearance in science all the more surprising. For the roughly 260 years from the publication of Newton's *Principia* to the 1945 close of the Second World War, the defining characteristic of fundamental advance in physics was the understanding of dynamical symmetry and conservation. Symmetries are invariances under operations, the invariant quantity said to be conserved. For physics the fundamental symmetries are the invariances of — hence the conservation of — dynamical quantities under various continuous space-time shifts, for example conservation of linear (respectively, angular) momentum under shift in spatial position (respectively, orientation) or of energy under time shift. Noether gave systematic form to this in 1918 and showed that it was the invariance of the form of the dynamical laws themselves that was expressed. Collections of the same space-time shifts form mathematical groups and the corresponding invariances then form dynamical symmetry groups. For instance, Newton's equations obey the Galilean symmetry group. Symmetry forms the deepest principle for understanding and investigating fundamental dynamical laws [Icke, 1995; van Fraassen, 1989].

In addition to their general dynamical symmetries, many system states have additional symmetries, for example the lattice symmetries of a crystal. Within this framework thermodynamics emerged and thermodynamic equilibrium became the primary dynamical state condition for analytic dynamics of many-body systems because all residual motion is random, hence spatio-temporally stochastically symmetric. Moreover each stable equilibrium state is invariant with respect to transitory pathways leading to it (the outcome is independent of those initial conditions), so its history can be ignored in studying its dynamics. The dynamics itself can then be developed in a simplified form, namely in terms of local, small and reversible — hence linearisable — departures from stable equilibria, yielding classical thermodynamics. The only place for complexity here is simply the large number of particles involved, the least profound dimension of the notion (see

[5]Some preliminary thoughts about it are found in section 4.1.3 below.

[6]These range from popularised expositions that still offer valuable introductions, such as [Gleick, 1988], to elegant mathematical reviews such as [Holmes, 2005]. Betwixt these lie many hundred good works, among them [Gell-mann, 1994; Goodwin, 1994; Kauffman, 1995; Layzer, 1990; Morowitz, 2002; Scott, 2004].

section 3 close).

The study of simple physical systems of a few components and of many-component systems at or near stable equilibrium supported the idea that the paradigm of scientific understanding was linear causal analysis and reduction to linear causal mechanisms, with the real as what was invariant under symmetry groups (a formal stability) or invariant to small perturbations (dynamical stability). Paradigm cases included 2-body solar system dynamics, engineering lever and circuit equations, equilibrium thermodynamics of gases.[7]

The philosophy of science evolved compatibly, focusing on determinism, universal a-temporal (hence condition-independent) causal laws, analysis into fundamental constituents then yielding bottom-up mechanical synthesis. To this was added a simple deductive model of explanation and prediction — deduction from theory plus initial conditions gives explanation after the event and prediction before it. Reduction to fundamental laws and separate contingent initial conditions became the basic explanatory requirement. This supports an ideal of scientific method as logical inference: logical induction from the data to form the most probably correct theory, deduction from theory for prediction and explanation, and deduction from data that conflict with prediction to a falsification of the predicting theory, or other assumptions. (However, it turns out (interestingly!) that neither the logical or the methodological situation is so simple; both scientific practice and rational method are, and must be, much more complex than this.[8])

The scientific paradigm and the philosophy of science together constituted the intellectual framework of scientific orthodoxy for more than a century of scientific understanding. The evident fit between philosophy and paradigm supported the conviction that both were right, the logical clarity and elegance of the philosophy reinforcing that conviction. From within this framework, the greatest challenge is that of quantum theory to determinism and simple causality. But while this is a profound problem, the immediate theoretical challenge is also limited since the fundamental dynamical idea of a universal deterministic flow on a manifold characterised by its symmetries remains at the core (indeed is abstractly strengthened — see e.g. [Brown and Harré, 1988]).

That this orthodoxy could prove unduly confining became obvious first in biology, even if it had made earlier, tacit appearances in physics and engineering. With

[7]See any of the many textbooks on these subjects. If all this seems somewhat impenetrable, it suffices to grasp the idea that symmetry, equilibrium and stability are the central structural features of dynamics in physics. On this framework see, for example, [Brading and Castellani, 2003; Icke, 1995; Stewart and Golubitsky, 1992; Van Fraassen, 1989] and on symmetry disruption by newer systems dynamics ideas see these and, for example, [Mainzer, 2005; Schmidt, this volume] and Brading's *Stanford Encyclopedia of Philosophy* entry at `http://plato.stanford.edu/entries/symmetry-breaking/`. Even with the common simplifying occurrence of symmetry and equilibrium, it is seldom appreciated that learning the dynamics of these systems would still have not been easy but for the fortunate presence of several other simplying conditions, see [Hooker, 1994b].

[8]See classics of the time like [Nagel, 1961] on induction and reduction, and on falsification see [Popper, 1972]. For overview and discussion of the difficulties with this conception of the methodological situation see for example [Hooker, 1987 passim; 1995, ch.2].

respect to biology, Woese [2004] rightly remarks that the nineteenth century had seen the first wave of modern revolution: Pasteur and others had removed spontaneous generation and revealed the huge range of kinds of life, Darwin had brought their evolution within science, the cell had emerged and the idea of genes had begun to form (cf. [Sapp, 2003]). Subsequently, the formation period of modern biology, characterised by the rise of genetics and its incorporation into evolutionary theory, and the subsequent emergence of elementary molecular genetics in its support, was initially all understood within the physics-inspired orthodox framework. The simple fundamental laws of evolutionary population genetics and of molecular genetics that underlay them were held to provide the universal, unchanging causal reality underlying the apparently bewildering diversity of biological phenomena. The observed diversity was to be seen simply as reflecting a diversity of initial conditions independent of these laws, whether generated as exogenous geo-ecological events or as endogenous random mutations. Reduction to molecular genetics thus became a defining strategy. But many biologists resisted this orthodoxy, insisting that the telling characteristics of living organisms — metabolism, regeneration, regulation, growth, replication, evolution and developmental and behavioural teleology — could not informatively be brought within the confines of the prevailing orthodoxy (e.g. [Mayr, 2004]). The consequence of orthodoxy for biology is that either life is radically reduced to simple chemical mechanisms and then to applied traditional physics, or it has to be taken outside the paradigm altogether and asserted as metaphysically sui generis. Both were, and are, implausible positions.

These are undoubtedly the early theory building stages through which any science has to go as it laboriously assembles better understanding. Possibly this was itself intuitively understood by many scientists. Even so, there was enough dogmatic conviction in science, and certainly in philosophy of science, that the results were not pleasant for dissenters, who were denied a hearing and research funding and often ostracised. One might recall, as examples of this, the fates of Baldwin and Lamarck and others in biology, and of Piaget in biology and philosophy — all now being at least partially rehabilitated as the old simple dogmas break down — not to mention those in entire sub-disciplines such as embryology and ecology who were sidelined for many years before they have again returned to the forefront of scientific progress.

Yet all the while scientific work itself was quietly and often unintentionally laying the groundwork for superseding this orthodoxy, both scientifically and philosophically. To understand why this might be so one has only to contemplate what the previous paradigm excludes from consideration, namely all irreversible, far-from-equilibrium thermodynamic phenomena. This comprises the vast majority of the subject matter of interest to science, everything from super-galactic formation in the early cooling of the universe down to planet formation, all or most of our planet's geo-climatic behaviour, all phase change behaviour, natural to the planet or not, and of course all life forms, since these are irreversible far-from-equilibrium systems.[9] What all of these phenomena exploit is dis-equilibrium, spontaneous

[9]Later static and dynamic equilibria will be distinguished, the latter co-existing with static

instability and symmetry-breaking (specifically, non-local, irreversible symmetry-breaking to form increased complexity). This is starkly clear for cosmic condensation: the universe begins as a super-hot super-symmetric expanding point-sphere, but as it expands it cools and differentiates, breaking its natal super-symmetry; the four fundamental forces differentiate out, their non-linearities amplifying the smallest fluctuation differences into ever-increasing structural features. In sum, all of this vast sweep of phenomena are characterised by the opposite of the symmetry/ stability-equilibrium paradigm.[10]

Thus it is not surprising that from early on, even while the elegantly simple mathematics of the symmetry/stability-equilibrium paradigm were being developed and its striking successes explored, scientists sensed the difficulties of remaining within its constraints, albeit in scattered and hesitant forms. Maxwell, who formulated modern electromagnetic theory in the later nineteenth century and sought to unify physics, drew explicit attention to the challenge posed by instability and failure of universality for formulating scientific laws. His young contemporary Poincaré spearheaded an investigation of both non-linear differential equations and instability, especially geometric methods for their characterisation. By the 1920's static, dynamic and structural equilibria and instabilities had been distinguished.[11] In engineering, non-linearity and emergent dynamics appeared in an analytically tractable manner with the entry of feedback and the development of dynamical control theory. Maxwell in 1868 provided the first rigorous mathematical analysis of a feedback control system (Watt's 1788 steam governor). By the early twentieth century General Systems Theory was developed by von Bertalanffy and others, with notions like feedback/feedforward, homing-in and homeostasis at their basis. Later Cybernetics (the term coined by Weiner in [1948]) emerged from control engineering as its applied counterpart at the same time as Walter [1950; 1953] was constructing his Machina speculatrix turtle robots and Ashby [1947] first introduced the term 'self-organisation'.[12] Classical control theory, which became

---

dis-equilibrium to form fluid phenomena like hurricanes and all living systems. Phase changes illustrate structural instabilities, where the dynamical form itself changes during the transient trajectory.

[10]An early mathematical classic on non-linear instabilities referred to the old paradigm as the 'stability dogma' — see [Guckenheimer and Holmes, 1983, pp. 256ff]. See also the deep discussion of the paradigm by a pioneer of irreversible thermodynamics [Prigogine, 1967; 1980; Prigigine, Stengers, 1984] plus note 7 references. I add the phase-shift cosmogony of Daodejing, ch. 42, translated by my colleague Dr. Yin Gao, because the West has been slow to appreciate the deep dynamical systems orientation of this tradition in Chinese metaphysics, for instance in medicine [Herfel, *et al.*, this volume]:

The dao (the great void) gives rise to one (singularity)
Singularity gives rise to two (yin and yang)
Yin and yang give rise to three (yin, yang and the harmonizing force)
Yin yang and the harmonizing force give birth to the ten thousand things/creatures.

[11]Thanks to Birkhoff and Andropov, following Poincaré. Lyapunov's study of the stability of nonlinear differential equations was in 1892, but its significance was not generally realised until the 1960's.

[12]See [Weiner, 1948; Walter, 1950; 1953; Ashby, 1947; 1952], cf. `http://pespmc1.vub.ac.be/DEFAULT.htm` (Principia Cybernetica Web) and [Hofkirchner and Schafranek, this volume]

a disciplinary paradigm by the 1960's, forms the basis of the use of dynamical systems models in many contemporary systems applications.

In 1887 Poincaré had also become the first person to discover a chaotic deterministic system (Newton's 3-body system), later introducing ideas that ultimately led to modern chaos theory. Meanwhile Hadamard 1898 studied a system of idealised 'billiards' and was able to show that all trajectories diverge exponentially from one another (sensitivity to initial conditions), with a positive Lyapunov exponent. However, it was only with the advent of modern computers in the late 1960's that investigation of chaotic dynamics developed, for example for atmospheric dynamics (Lorenz). By the mid-1970's chaos had been found in many diverse places, including physics (both empirical and theoretical work on turbulence), chemistry (the Belousov-Zhabotinskii and like systems) and biology (logistic map population dynamics and Lotka-Volterra equations for 4 or more species) and the mathematical theory behind it was solidly established (Feigenbaum, Mandelbrot, Ruelle, Smale and others).[13]

Since then the story is one of exponential explosion of, and increasing complexity in, content. Bishop [this volume, section 2.4] recounts the complexities slowly uncovered in sensitivity to initial conditions and chaos in systems. Meanwhile, the identification and understanding of self-organised emergence shows no signs of global consensus yet. While many philosophers early on preferred 'change inexplicable from constituents' and the like (see e.g. [McLaughlin, 1992; O'Connor, Wong, 2002]), scientists prefer something less human-dependent, but without much agreement. Commencing from the most general of features and narrowing down there are, first, the common appeals to spontaneous symmetry breaking and failure of superposition (cf. [Landau, 1937; Anderson, 1972], followed by 'expresses logical depth' (see note 43 below) [Bennett 1985; 1992; Kauffman, 1995], 'better global system predictability than do the constituents'- [Crutchfield, 1994; Shalizi, 2001; *et al.*, 2004], 'exhibiting downward causation' — [Sperry, 1969; 1986; 1987; Campbell, 1974; Bickhard, 2000a; Goldstein, 2002], 'relatively macro constraint formation' — [Collier and Hooker, 1999; Hooker-c, this volume], these last three being cousins if the better predictability is because of the emergent constraints on the system, 'optimal prediction is simulation or reproduction' — [Bedau, 1997]

on General Systems Theory. In control engineering Airy (1840) developed a feedback device for pointing a telescope, but it was subject to oscillations; he subsequently became the first to discuss the instability of closed-loop systems, and the first to use differential equations in their analysis. Following Maxwell and others, in 1922 Minorsky became the first to use a proportional-integral-derivative (PID) controller (in his case for steering ships), and considered nonlinear effects in the closed-loop system. By 1932 Nyquist derived a mathematical stability criterion for amplifiers related to Maxwell's analysis and in 1934 Házen published the *Theory of Servomechanisms*, establishing the use of mathematical control theory in such problems as orienting devices (for example naval guns). Later development of the use of transfer functions, block diagrams and frequency-domain methods saw the full development of classical control theory. For more details see, among many, [Bennett, 1986; 1993].

[13]For a reasonable coverage of books on complex systems dynamics over the period 1975-1995 see the *ed works in the bibliography to [Hooker, 1995b] and for contemporary review see [Scott, 2004].

(cf. complexity as the amount of past information required to predict future system states, Shalizi), 'requiring meta-modelling' — [Rosen, 1985; Heylighen, 1991), and 'expresses global semantics' — [Pattee, 1968; 1997; Kubik, 2003] and 'greater scope' [Ryan, 2007]. See further [Bedau and Humphries, 2008; Boschetti, *et al.*, 2009; Holland, 1998]. Similar stories can be told about most of the other key complex systems features (see section 3 below for a list), and where only brief, vaguer stories are available it is because too little is yet known. The best introduction to the modern period is, therefore, the essays herein themselves.

Returning to the larger narrative, this historical account itself is unavoidably selective and sketchy, but it sufficiently indicates the slow build-up of an empirically grounded conceptual break with the simple symmetry/stability-equilibrium orthodoxy, and a corresponding background murmur of doubt and difficulty within the foundational/philosophical tradition. However the new approach often still remained superficial to the cores of the sciences themselves. In physics this is for deep reasons to do with the lack of a way to fully integrate instability processes, especially structural instabilities, into the fundamental dynamical flow framework (at present they remain interruptions of flows), the lack of integration of irreversibility into fundamental dynamics and the related difficulty of dealing with global organisational constraints in flow characterisation (see sections 3, 4 and 5 below). For biology all that had really developed was a partial set of mathematical tools applied to a disparate collection of isolated examples that were largely superficial to the then core principles and developmental dynamics of the field.

Nonetheless, by the late 1970's it is clear in retrospect that science had begun to pull together many of the major ideas and principles that would undermine the hegemony of the simple symmetry/ equilibrium orthodoxy. Instabilities were seen to play crucial roles in many real-life systems — they even conferred sometimes valuable properties on those systems, such as sensitivity to initial conditions and structural lability in response. These instabilities broke symmetries and in doing so produced the only way to achieve more complex dynamical conditions. The phenomenon of deterministic chaos was not only surprising to many, to some extent it pulled apart determinism from analytic solutions, and so also from prediction, and hence also pulled explanation apart from prediction. It also emphasised a principled, as opposed to a merely pragmatic, role for human finitude in understanding the world.[14] The models of phase change especially, but also those of far-from-equilibrium dynamical stability, created models of emergence with causal power ('downward' causality — see above) and hence difficulty for any straightforward idea of reduction to components.[15] And, although not appreciated until recently, they created an alternative paradigm for situation or condition-dependent, rather than universal, laws.[16] Thus a new appreciation for the sciences of complex dy-

[14]The point being that any finite creature can only make finitely accurate measurements, independently of any further constraints arising from specific biology or culture; there is always a residual uncertainty and any sensitivity to initial conditions will amplify that uncertainty over time.

[15]See for example [Boogerd, *et al.*, 2002; 2005; Bruggeman, *et al.*, 2002].

[16]See further sections 3, 6.2.2 below; for the basic idea see [Hooker, 2004].

namical systems began to emerge.

Following this development a little further into contemporary biology as an example, and then briefly to psychology, will complete this invocation of a historical context for the essays to follow. The mid-twentieth-century period quietly set the stage for the undoing of geneticism: the assertion of a simple gene-trait-fitness model that effectively by-passes the cellular phenotype, with its biosynthetic pathways and multi-cellular physiological processes, as unnecessary to evolutionary explanation. The manner of undoing this paradigm paved the way for the more intimate introduction of complex systems methods into the heart of biology.

In mid-twentieth-century physics (1930-1970), Prigogine (following Schrodinger and Turing) worked on irreversible thermodynamics as the foundation for life. This generates a (macroscopic) metabolic picture in which the full internally regulated body, including each of its cells individually, is essential to life [Prigogine, 1967; 1980]. In the same period cellular biology was revived and underwent a rapid development. This was partly driven by new, biochemically based problems (understanding kinds and rates of chemical reactions like electron transport, etc.) and partly by new instrumentation (electron microscope, ultracentrifuge) that allowed much more detailed examination of intracellular structure and behaviour. In consequence, there was increasing molecular understanding of genetic organisation, especially development of RNA roles in relation to DNA, of regulator genes and higher order operon formation, and of the roles of intra-cellular biochemical gradients, inter-cellular signalling and the like in cellular specialisation and multi-cellular development. All this prepared the ground for envisioning the cell as a site of many interacting biochemical processes, in which DNA played complex interactive roles as some chemicals among others, rather than the dynamics being viewed as a consequence of a simple deterministic genetic programme. (The surprising modesty of the number of genes in the human genome, given our phenotypic complexity, emphasises the importance of the gene regulatory networks.) Genetics was replaced by 'omics' (genomics, proteomics, ... metabolomics, etc.).[17] These are the very ideas that, allied to the development of generalised network analysis emerging from circuit theory and chemical process engineering and elsewhere, would later underlie contemporary systems and synthetic biological modelling, now re-casting cellular biology and bio-engineering.[18]

During roughly the same period, Rashevsky and others pioneered the application of mathematics to biology. With the slogan *mathematical biophysics : biology :: mathematical physics : physics*, Rashevsky proposed the creation of a quantitative theoretical biology. He was an important figure in the introduction of

[17]This term also stands for the technologies that provide data on these aspects of cellular function. See, for example, [Bechtel, 1989; Westerhoff and Palsson, 2004] and references, and the Omics symposia at http://www.keystonesymposia.org/Meetings/viewPastMeetings.cfm?MeetingID=980&CFID=376546&CFTOKEN=248205721.

[18]For an introduction to the field of systems and synthetic biology, see e.g. www.systems-biology.org/, http://syntheticbiology.org/ and the IEEE Proceedings, systems biology, at www.iee.org/Publish/Journals/ProfJourn/Proc/SYB/index.cfm. For a sense of the integration required from genes to cognition see Miklos (1993).

quantitative dynamical models and methods into biology, ranging from models of fluid flow in plants to various medical applications. That general tradition was continued by his students, among them Rosen, whose edited volumes on mathematical biology of the 1960's and 1970's did much to establish the approach.[19] In this tradition various physiologists began developing the use of dynamic systems to model various aspects of organism functioning. In 1966, for example, Guyton developed an early computer model which gave the kidney pre-eminence as the long-term regulator of blood pressure and went on to develop increasingly sophisticated dynamical network models of this kind. The next generation expanded these models to include intra-cellular dynamics. Tyson, for example, researched mathematical models of chemical systems like Belousov-Zhabotinskii in the 1970's, passing to cellular aggregation systems like *Dictyostelium* in the 1980's and to intra-cellular network dynamics models in the 1990's and this was a common progression.[20] See also the increasingly sophisticated models of timing (circadian rhythms)[21] and growth (L-systems) that complement more traditional growth models [Green and Leishman, this volume]. In this manner physiology has supported a smooth introduction of increasingly refined dynamical models into biology, providing a direct resource for contemporary systems and synthetic biology.

There has also been a correlative revival of a developmental perspective in biology, in embryology generally and early cellular differentiation in particular. This became linked to evolutionary 'bottlenecks' and evolutionary dynamics generally to form evo-devo as a research focus. Added to this was work on epigenetics and non-nuclear cellular inheritance, especially down the maternal line, and early work on enlarging evolutionary dynamics to include roles for communal (selection bias, group selection) and ecological factors, culminating in the wholistic 'developmental systems' movement.[22]

Ecology too has been studied as dynamic network (Lotka/Volterra, May, Levins and others), as an irreversible far-from-equilibrium dissipative flux network (Ulanowicz) or food-web energetics system (Odum), as a spatio-temporally differentiated

[19] Indeed, as Rosen remarks, "It is no accident that the initiative for System Theory itself came mostly from Biology; of its founders, only Kenneth Boulding came from another realm, and he told me he was widely accused of "selling out" to biologists." On Rashevsky see, for example, [Rashevsky, 1964], `http://www.kli.ac.at/theorylab/AuthPage/R/RashevskyN.html`. For many years (1939-1972), he was editor and publisher of the journal *The Bulletin of Mathematical Biophysics*. For Rosen see `http://www.panmere.com/?page_id=10` and notes 61, 64 and section 4.1.1 below. The quote comes from his *Autobiographical Reminiscence* at `http://www.rosen-enterprises.com/RobertRosen/rrosenautobio.html`.

[20] Among other resources see respectively `http://www.umc.edu/guyton/`, `http://mpf.biol.vt.edu/people/tyson/tyson.html`. Compare the work of Hogeweg, e.g. [2002a], [Hogeweg and Takeuchi, 2003] and see `http://www-binf.bio.uu.nl/master/`. For an eclectic but stimulating introduction to complex modelling of immune systems see `http://jason.brownlee05.googlepages.com/home22`.

[21] [Glass and Mackey, 1988; Kaplan and Glass, 1995; Winfree, 1987].

[22] See, for instance and respectively, [Solé, *et al.*, 1992; Raff, 1996; Raff and Kaufman, 1983; Goodwin and Saunders, 1989; Goodwin 1994; Jablonka and Lamb, 1995; Gray 1992; Oyama *et al.*, 2001], Gray's title itself indicating something of the macro intellectual landscape in which the idea emerged.

energy and matter flow pathway network (Pahl-Wostl) self-organising through inter-organism interaction (Holling, Solé/ Bascompte) and as an organized complex dynamic system employing threshhold (bifurcation) dynamics, spatial organisation and exhibiting adaptive resilience (Holling, Walker and others), responding in complex, often counter-intuitive, ways to policy-motivated inputs.[23] All these features are found within cells, albeit more tightly constrained by cellular regenerative coherence, and fruitful scientific cross-fertilisation should eventually be expected, perhaps particularly with respect to the recent emphasis in both on understanding the coordination of spatial with functional organisation.

All of these scientific developments, still in process, work toward replacing black box geneticism with a larger model of a mutually interacting set of evolutionary/ developmental/ communal/ ecological dynamic processes.[24] Although still a collection of diverse models and methods, dynamical network methods are emerging across these disciplines as a shared methodological toolkit.[25] Combined, these developments present a picture of life as a complex system of dynamic processes running on different groups of timescales at different spatial scales, with longer term, more extended processes setting more local conditions for shorter term, less extensive processes while shorter term, local products accumulate to alter the longer term, more extensive processes, the whole interwoven with self-organised assembly of near-to-chaos criticality and resolutions of it.

A similar overall story could be told for several other subjects. As with biology, in each case the story is about replacing a 'black box' of internal quasi-formal elements, whose interrelations dominate the character of the whole process, with dynamical complex systems models that open up interactive modelling across the domain. For geneticism the elements were genes (both for individuals and populations), their interrelationships determining traits for individuals and gene frequencies for populations. For the behaviourist psychology that dominated the first half last century, the elements were stimuli and responses, their association interrelations determining behaviour. Under criticism this model shifted to a functionally generalised Behaviourism, viz. computational cognitivism, whose elements were symbols and whose relations were logistic or formal symbolic programming. That model is still the orthodoxy today. Geneticism excluded or ignored what did not fit, namely the sciences of what embodied the elements and their interrelations — embryology, developmental and socio-biology and others. Similarly, behaviorist and cognitivist psychology excluded or ignored what embodied its associations or symbolic processes — developmental and neuro psychology and others.

Geneticism was undermined by the increasing emergence of dynamical models

[23]See in rough text order, among many others, the following works and their references: [May, 1976; Ulanowicz 1997; Odum, 1971; Pahl-Wostl, 1995, Holling, 1992; Gunderson *et al.*, 1997; Gunderson and Holling, 1992; Solé and Bascompte, 2006; Dieckmann *et al.*, 2000; Cash *et al.*, 2006; Haila and Levins, 1992; Scheffer *et al.*, 2002] and further [Odenbaugh, this volume].

[24]Cf., for example, [Jablonka and Lamb, 2005] and note 22 references. See further section 6.2.4.

[25]For recent reviews see e.g. [Barabási, 2002; Boccaletti, *et al.*, 2006; Barabási and Bonabeau, 2003; Strogatz, 2001].

of what embodied the elements and their interrelations. Similarly, cognitivism is being undermined by the increasing emergence of dynamical models of what embodied the elements and their interrelations, e.g. by the interactive-constructive tradition in bio-psychology of Piaget [Piaget, 1971a; 1971b; Piaget and Garcia, 1989] (cf. Hooker [1994c]) and Vygotsky [1986] and Bickhard's Interactivism [Bickhard, 2005; 2008; Bickhard and Terveen, 1995]; the cognitive cybernetics tradition of Ashby [1947; 1952], Pask [1960; 1981], Cunninham [1972] and Cariani [1992; 1993]; the introduction and spread of neural network models from McCullogh and Pitts and Rosenblatt to Rummelhart *et al.* and on[26]; embodied agent robotics from Braitenberg [1984], to Brooks [1991] and Beer *et al.* [1993], Smithers [1995], Nolfi [this volume]; the related dynamical psychology tradition (Heath [2000], Ward [2002], Sheya and Smith [this volume], Van Orden *et al.* [this volume]; and through the embodied intentionality movement in philosophy from Merleau-Ponty [1962] to [Dreyfus and Dreyfus, 1986; Dreyfus, 1996] (cf. partial synthetic reviews by [Clark, 1997; Exteberria, 1998]).There is a similar story to tell about the neo-classical economic model of an agent (Homo economicus) as an internal bundle of expressed preferences and a market as an aggregate of these, each operating under a generalised Nash bargaining equilibrium process and being undermined by dynamical models of what embodies these processes.[27]

In both of these cases, as with biology, the entrance of complex dynamical systems models and methods did not simply replace one internal model with another. Instead it opened up the internal processes to much richer modelling of individuals and of their interactive relationships with other individuals, communities and ecologies (human and natural). This again indirectly undermines the hegemony of the simple symmetry/stability-equilibrium orthodoxy by supporting the development of an alternative suited to a complex dynamical world (sections 5, 6 below). And, as with biology, in doing so it opened up the prospect for rich cross-disciplinary integrations, from preference dynamics and decision psychology through collective interactions (for example through multi-agent network dynamics) to industrial ecology and evolution of market structure. We can even contemplate incorporating culture into this picture as a feature of a piece with it rather than as a separate box of memes. (Yet another internal box model like

[26]See respectively [McCullogh and Pitts, 1943; Rosenblatt, 1958; Rummelhart, *et al.*, 1986]. See further history sites like `http://en.wikipedia.org/wiki/History_of_artificial_intelligence` and [Churchland, 1989; 1995; Churchland, 1986; Hooker 1995a], among many, for wider review and references. However, connectionist methodology and rhetoric often retains the general internalist assumption that cognition is basically formal problem solving in a separable input-output processing unit that can be given a largely intrinsic characterisation.

[27]See e.g. econophysics [Rickles, this volume], evolutionary economics [Foster, this volume], while [Anderson, *et al.*, 1988; Arthur, *et al.*, 1997; Blume, Durlauf, 2006] offers a general review. Despite its recent suppression, the dynamical approach has a long history, see e.g. [Roos, 1927], and is now entering orthodox economic discussions from every side, see e.g. the Journal of Economic Dynamics and Control `http://www.elsevier.com/wps/find/journaldescription.cws_home/505547/description#description`, and the Society for Economic Dynamics and their journal at `http://www.economicdynamics.org/index.html`.

those above.)[28] In short, in place of a dis-unified collection of assorted boxes of elements, dynamical complex systems approaches offer a rich, unified conception of the world and its science based on interrelating the dynamical models being employed in each — albeit a complex unity of differing but mutually interacting processes and levels governed by condition-dependent laws (cf. section 3 below).

3 CONCEPTUAL NOTES ON COMPLEXITY

It is now a good idea to obtain some grasp on what complex systems are. Unhappily, as noted earlier, there is currently no unified theory available, nor do we know how to systematically interrelate all the characteristic properties involved. Grimly, this disunity opens the way to an opposite affliction, mis-placed simplicity: too often when discussing complex systems only a select few characteristics are considered (most commonly, just chaos) and the rest ignored. In consequence, discussion is often inadequate because too narrowly based. Similarly, simplistic notions of complexity are too often assumed, e.g. that it is just showing counter-intuitive or unpredictable behaviour, or having many independent components and so requiring a long description. What follows is an informal rather than a technical exposition, although technical remarks appear. It is made with the aim of surveying an adequate range of system features entering complex systems, but also with an eye on the issues they raise for understanding complexity and complexity science.

The complex systems that constitute our life world are characterised by deterministic dynamics that manifest some or all of the properties in the following list. Roughly, properties lower on the list are increasingly richly possessed by living systems and present increasing contemporary challenges to our dynamical understanding. The properties are briefly characterised, in order of appearance.

Non-linear interactions; non-additivity;
Irreversibility;
Constraints — holonomic and non-holonomic;
Equilibria and Stabilities — static and dynamic;
Amplification; sensitivity to initial conditions;
Finite deterministic unpredictability;
Symmetry breaking; bifurcations; self-organisation; emergence;
Constraints — enabling and coordinated;
Intrinsically global coherence;
Order; organisation;
Modularity; hierarchy;

[28] For the internal-box-model tradition in biology, psychology and sociology see [Christensen and Hooker, 1998], for preliminary thoughts about modelling culture dynamically see [Hooker, 2002] and section 4.1.2 below. For a complementary conception of integrated or holistic dynamical agency see note 51 references. Culture has also been introduced to evolutionary processes as simply a selection bias or similar 'external factor'; insofar as these processes are dynamically modelled they also contribute to this picture, albeit less fully.

Path-dependence and historicity;
Constraint duality; super-system formation;
Coordinated spatial and temporal differentiation with functional organization;
Multi-scale and multi-order functional organisation;
Autonomy; adaptation; adaptiveness; learning;
Model specificity/model plurality; model centredness;
Condition-dependent laws.[29]

The diversity and domain-specificity of these properties explains the diversity of notions of complexity, self-organisation, etc. The challenges to understanding that these properties continue to pose reduces hope for any unified account of the complexity domain in the near future, although progress continues to be made (see below).

Just as there is no canonical list of complex systems properties, many of these terms have no canonical expositions — recall views of emergence, above. (This is another of the complexities of the complex systems field.) The proper exposition of many of these terms might each take an essay and it will be a kind of reflexive manifesting of the diversity of the field to review what features, under what descriptions, this volume's authors use. I have no wish to impose expository order here. (Readers should feel free to use the expositions as a base from which to explore alternative conceptions, including ones that contradict opinions expressed below.) Nonetheless, a bat-roost squawk of disparate terms left unattended may confuse readers dipping their toes into new domains. Moreover, while the gists of some of these terms are well enough known, some terms that are in common usage are often wrongly considered well known. Instead, they present ongoing challenges to understanding (e.g. self-organisation, emergence and organisation) and/or are vague or misunderstood (constraint duality, path-dependence, autonomy, model centredness). So, in the following the gists of all these terms will be briefly provided, as a general orientation (only) to the phenomena and issues subsequently discussed in this essay and the essays to follow.

Non-linear interactions and non-additivity. An interaction is *non-linear* in some variable v if the interaction force does not vary proportionately to v (F linearity: F(kx) = kF(x), k a number). Gravitational force, e.g., is spatially non-linear since it varies as the inverse square of the interaction distance (G(kr) = k^{-2}G(r)). Interaction linearity yields linear dynamical equations describing the system, F/m = $d^2x/dt^2 = a^2x$, yielding exponential ($a^2 > 0$) or oscillatory ($a^2 < 0$) motion, respectively x(t) = Ae^{at}, Ae^{iat}. These are characterised by additivity: any numerical combination of solutions (e.g. $Ae^{at} + Be^{at}$) is also a solution. No complex dynamical behaviour is possible. Conversely, interaction non-linearity

[29] This list elaborates on that in [Herfel and Hooker, 1999]. I happily acknowledge Herfel's various contributions to my understanding of them at that time. An abbreviated intermediate version without most commentary appeared in [Hooker, 2009a, section D.7]. Another list appears in [Bishop, this volume]. Bishop's list and mine, together with the conclusions we draw from them, were arrived at mutually independently.

yields non-linear dynamical equations characterised by *non-additivity*: numerical combinations of solutions are in general not solutions. (This is essentially because non-linearities ensure that the consequence of an increase or decrease in some quantity is not proportional to it.) Non-adddativity is a necessary condition for complex dynamical behaviour.[30]

Irreversibility. A process that is reversible can also be run backwards while still satisfying the same laws. Classical dynamics is time-reversible in this sense. Every dynamically possible process running forward is equally possible running backwards. But virtually all real processes are dissipative, the energy they degrade in quality and shed as waste during the process cannot be retrieved (e.g. because it is converted to heat and distributed randomly throughout the universe) so that they cannot be run in reverse. They also will not persist unless a supply of sufficiently high quality energy, typically in requisite material forms, is available to continually renew their dissipative processes. Hence they are inherently open systems. Here only very small, brief changes may be approximately reversible, a condition especially obtaining near a dynamical equilibrium. Many, not all, examples of complex dynamics, but all those concerned with living systems, are of this kind.

Constraints — holonomic and non-holonomic. Constraints on a dynamical process are those limitations on the relationships among its variables that arise from the imposed physical conditions in which the process takes place. A marble rolling in a bowl, e.g., is confined to the surface of the bowl, whereas a small spacecraft has no such constraints, though both move under local gravitational forces. A system's effective degrees of freedom are those provided by its inherent variabilities (its dynamical variables) minus those removed through constraints.[31] A dynamical explanation consists in deriving the behaviour in question from a model of dynamical variable interrelations, constraints and initial conditions.[32]

[30]See also [Bishop, this volume, section 2.2] on failure of superposition (= addativity). But its failure is not a sufficient condition for complexity, as the example of 2-body gravitational motion shows. We currently have no general account of which kinds and combinations of non-linearities will yield which classes of complex systems since interesting complex systems can be governed by relatively structurally simple, 1 or 2 variable, equations (e.g. 1-variable logistic reproductive and 3-body gravitational systems) as well as by more complex equations, while some algorithmically very complex non-linear systems (for example, a gas) generate no interesting complex behaviour.

[31]It is possible in principle to also constrain a system's parameter ranges — in effect, to constrain the range of quantitative forms its dynamics can take — but this is difficult and rare for all but artificial systems set up for this purpose. The net effect of such constraints will still appear as (indirect, often extremely complex) constraints on variables.

[32]The initial conditions fix its particular trajectory from among the possible ones. Constraints are taken to include limiting constraints (e.g. that interaction forces go to zero at infinite separations). These are often labelled the system 'boundary conditions'. However 'constraints' is both more faithful to the mathematical structure of dynamics and more general since it continues to apply even when boundaries are diffuse or uncertain (cf. [Bishop, this volume, section 3.5) and when the constraints penetrate inside accepted boundaries (e.g. an externally applied electric field, the zero limit for distant forces). On dynamics see further [Hooker-c, this volume].

As the marble rolling without and with friction respectively shows, constraints may apply to both reversible and irreversible processes and are typically required to characterise specific processes, e.g. in a cell. Currently, we can only form general analytic dynamics, the Lagrangian/Hamiltonian dynamics, for systems that do no work on their constraints. The core of these in turn is formed by systems also having holonomic constraints (roughly, constraints that are purely a matter of space-time geometry and independent of the system's dynamical states) and are energy conserving. There are various simple extensions of this core theory into the non-holonomic, non-conserving domain but they remain special cases. It is precisely these features that complex non-linear irreversible systems lack. E.g. a river running between sand or gravel banks has non-holonomic constraints (the banks) which it alters through doing work on them and thus dissipating (not conserving) its energy. Similarly, a group of closed self-reproducing processes (e.g. cell metabolism) must do work on their many constraints in order to recreate them, while many dynamical bifurcations (e.g. boiling) are initiated in this way. *Prima facie*, all of these systems fall outside the scope of analytical Lagrangian dynamics.[33]

Equilibria and stabilities — static and dynamic. Qualitatively, some aspect A of a dynamical system is in equilibrium if (and only if) there is no net force acting on the A aspect of the system (its A forces are in balance) and there are thus no net 2^{nd} order rates of change (accelerations) in A. Across the range of possible A choices an important distinction is then drawn between static and dynamic equilibria, that is, between cases where the time invariance concerns state parameters and variables (A = system state) and cases where it concerns process parameters and rate variables (A = system processes). Static equilibria require no energy input or output to persist, e.g. a crystal at rest. Dynamical equilibria typically require an irreversible ordered energy (negentropy) flow to sustain them, e.g. water flow to sustain the wave structure of river rapids, together with appropriate waste (degraded or entropic) outputs, e.g. turbulent water. For living systems there is water, food and hydrogen or oxygen input flow to sustain them and heat and chemicals as waste outputs. For these and other dynamically stable systems energy-material flows act to stabilise the system processes so that, for small disturbances that do not affect the underlying flows, these systems will behave as if they have static equilibria in these flow variables. In other respects however, such as river flows acting on river banks and metabolism producing cellular aging, the system may do work that alters, and eventually undermines, its dynamic equilibria.

Equilibria of any sort are stable, meta-stable or unstable. An equilibrium in some aspect A is stable, with respect to some class of disturbances (perturbations) D, if (and only if) its response to any disturbance from D is to soon return near

[33] See, among many, [Bloch, 2003] and further discussion in [Hooker-c, this volume]. As Bloch demonstrates, it is possible to control non-holonomic systems even where we have no analytical dynamics for them, because we can still locally model the dynamical behaviour.

(including exactly) to its original A condition under its own dynamical processes and remain there. An equilibrium is unstable to a class D of disturbances if it does not return near to its original A condition and it is meta-stable to D if it is stable for some disturbances from D and unstable for others.[34] The closed set of states a system repeatedly traverses when at equilibrium is its attractor (the marble's rest point is a point attractor, if it circled frictionlessly around that point but up the basin wall it would be a cyclic attractor) and the wider set of states it can pass through while still returning to its attractor is its 'attractor basin' (the bowl provides a literal attractor basin).[35] A complex dynamics may generate several different equilibria of various stabilities (attractor basins of various shapes) either directly intersecting or connected by transient paths (where small disturbances may change its destination, cf. rolling on horizontal surfaces) plus other transient paths that do not end. This 'attractor landscape' is a system's dynamical signature, expressing its dynamical form.[36] A system that remains within a single attractor landscape is structurally stable (= autonomous dynamics in mathematical parlance) and otherwise is structurally meta- or un- stable. While mathematical dynamics typically assumes structural stability, many complex systems are structurally unstable (= bifurcate in mathematical parlance), e.g. exhibiting phase changes.

Amplification. Transient paths aside, a disturbance to a system at equilibrium will have one of three consequences: leave it within its current attractor basin, push it into another attractor basin, or transform the attractor landscape, leaving the system in an attractor basin in the new landscape.[37] In the first case, the

[34]Often presentations will use a simplified form by taking A to be all state parameters and variables or all process parameters and rate variables and taking D to be all disturbances, so that stability is global stability and so on, but relativising the notion to an aspect A and a class D of disturbances provides for a more precise and richly structured, and realistic, account. Note that this formulation holds for both static and dynamical equilibria, it is just that in the static case it applies to a set of state parameters and variables while in the dynamical case it will apply to a set of process parameters and rate variables. To obtain a quantitative notion it is necessary to specify how near a return is near enough and how soon the return is soon enough, and perhaps how far it can wander (how much it can change its equilibrium parameter and variable values) before returning and how easy/hard it is to create that wander with a perturbation (how much resistence it has). These are details here, but they can be useful for various purposes and are required in any rigorous analysis — e.g. see [Brinsmead and Hooker, 2006] on ecological resilience.

[35]In 'bowl' terms, for a class D of disturbances of energy less than the bowl height, equilibria can be stable (= simple bowl-like attractor basin, the system response to all perturbations is to return to the attractor), meta-stable (= a bowl much shallower in some directions than others, the system response to some sufficiently energetic perturbations escapes the bowl and does not return to the attractor) and unstable (= atop an inverted bowl, the system response to all perturbations is to not return to the attractor). For an introduction, among many, see [Lorenz, 1997].

[36]A dynamics for which superposition holds [Bishop, this volume], that is, a linear dynamics (see above), has no attractor basins, hence no equilibria, simply a 'flat' landscape filled with transients. (Technically it is everywhere in a degenerate 'neutral' equilibrium, remaining in whatever state it happens to be in initially.)

[37]Including transients complicates the story without adding anything essential.

disturbance will be suppressed — negatively amplified (amplified down) — as the system returns near to its original state. In the other two cases the disturbance will be augmented — positively amplified (amplified up) — as the system departs from its original state. Amplification is the norm in non-linear systems.

Sensitivity to initial conditions. Non-linearities permit small differences in system state to be amplified into large differences in subsequent system trajectory. This is sensitivity to initial conditions. For instance, a system may be poised near the threshold of changing either attractor basin within the same landscape or changing landscape, so that a small disturbance is amplified to produce a larger change. And while systems left by disturbances inside the same attractor basin are insensitive to initial conditions in respect of their ultimate destination, they may still be locally sensitive to the path taken back to the equilibrium attractor.[38] Sensitivity to initial conditions is as common as amplification, but under certain conditions it takes a special form where a 'strange attractor' is formed in which motion is said to be chaotic because it occurs at random (in certain respects). However, the motion remains deterministic and, far from being more disordered than a normal attractor, is best viewed as super-ordered since every point within it may manifest sensitivity to initial conditions. The randomness is confined to trajectory locations sampled across the attractor and the like. Indeed, measures of chaotic system states show statistical distributions with 'fat tails', that is, with events that would be rare and lie far out from the mean were the processes fully random now showing up much more often [Rickles, this volume]; this is to be expected since the trajectories remain confined to the strange attractor, reflecting their internal interdependencies. Such phenomena are characteristic of chaos but are not confined to that condition, they may occur wherever subtle correlation is combined with quasi-randomness.

Finite deterministic unpredictability. Systems manifesting sensitivity to initial conditions present the problem that small uncertainties (including errors) in initial conditions may be amplified into large subsequent uncertainties in system location and trajectory. That is, system predictability is limited by knowledge of initial conditions. How severe a limitation this is in practice, and in what respects, depends on the amplification processes involved. In particular, while prediction that a system's state will remain within a strange attractor is often legitimate, knowledge of location within the attractor can be quickly lost (though not always). Conversely, since systems showing sensitivity to initial conditions can be significantly influenced using only small signals (disturbances), then so long as the relevant state conditions can be distinguished, these conditions can be used to

[38]Think of a marble rolling in a bowl with a pinpoint protuberance on its side surface. As these cases all show, sensitivity is strictly relative to the specified magnitudes of the disturbances. On this and following notions see also [Bishop, this volume].

sensitively guide or control them.[39]

Symmetry breaking. A symmetry is an invariance under an operation, e.g. the molecular structure of a cubic crystal is invariant under spatial shift along any axis, classical dynamics is invariant under reversal of time. Symmetry breaking occurs when an existing symmetry is disrupted. If water or another fluid is heated from below in a pan, e.g., then its previous complete kinetic symmetry (same random motion profile of molecules throughout the pan) is broken vertically as layers nearer the bottom heat up while those at the top remain cooler, passing the applied heat upward by conduction. This already changes the dynamical form of the system, from a stable dynamic equilibrium maintained by internal molecular collisions producing no net macro force to a stabilised dynamic equilibrium maintained by an irreversible vertical transmission of heat (kinetic energy). If the applied heat is increased there comes a point where rolling boiling sets in, conduction is replaced by convection and the fluid breaks up horizontally (and vertically) into convection cells, each matched to its neighbour along its boundary. This change corresponds to the breakdown of previous horizontal symmetry and is again maintained by the (increased) heat flow. In each symmetry breaking the orderedness and complexity of the system behaviour increased, and this is typical. Symmetry breaking may be spontaneous, that is, brought about by the system's own dynamics, or imposed, as in the heating example above. Spontaneous symmetry-breaking transitions are assumed to account for all the emergence of order and complexity in the universe since the super-symmetry of the big bang.[40]

Bifurcation. A bifurcation occurs when a structural instability in a system leads to a change in its dynamical form, that is, a change in the structure of its attractor landscape. There are many dynamically different ways in which this can occur, broadly classified as either local — where the form changes continuously as some dynamical parameter or parameters continuously vary — or global changes that involve more complex shifts.[41] Among the latter are phase transitions (e.g. gas to liquid, liquid to solid, or reverse), including critical point transitions (e.g. si-

[39] Satisfying the qualifying clause is of course the difficulty. In this respect, there has been claimed to be an ideal condition for generating complex behaviour (strictly: computational capacity in a cellular automaton) and that is near to the chaotic condition, where there are multiple sensitivities, but not fully chaotic (in a strange attractor) where there are too many of them. This is edge-of-chaos criticality. Van Orden *et al* [this volume] give it a large role. However, this idea has also been criticised. See `http://en.wikipedia.org/wiki/Edge_of_chaos` for a summary and references.

[40] See [Landau, 1937] and the classic [Anderson, 1972] and note 7 references.

[41] See, e.g., `http://www.scholarpedia.org/article/Bifurcation`, including [Guckenheimer and Holmes, 1983; Humphries, 1994] on the Ising model, and the International Journal of Bifurcation and Chaos. For a local example, imagine how the motion of the marble rolling across a round-bottomed plastic bowl changes when a screw is attached to its base at some point and is either lowered to create a local basin or raised to create a local hill in the bowl contour. The marble dynamics changes as the bolt shifts away from neutral, so the bifurcation parameter is the height of the bolt.

multaneous transitions among gas, liquid and solid states), where changes can be discontinuous and uncomputable, essentially because fluctuations on every scale up to that of the whole system are simultaneously possible. While we can study mathematically the conditions under which a bifurcation occurs, beyond the simplest cases we typically have no dynamical mathematical analysis of the process of the change. Rather, the conditions of occurrence are deduced, where possible, by matching up characterisations of the antecedent and subsequent dynamical states, e.g. in terms of parameter changes across the bifurcation threshold.

Self-organisation. Self-organisation occurs when a system bifurcates, sufficiently under its own dynamics, to a form exhibiting more ordered and/or more complex behaviour. The molecular motion in a heated pan of water shifting from conduction through random collisions to cellular convecting, provides a core intuitive example. By contrast the reverse bifurcations as heat is reduced to the water, equally dynamically transforming, would not normally be considered self-organisations (they might be considered self-disorganisations). Since the condensing of molten iron to form a solid iron crystal is also considered self-organisation it is clear that self-organisation has little to do with organisation proper, since an iron crystal is too ordered to be significantly organised (see below). Many self-organised states could also be brought about through external manipulation; e.g. it is possible to build up an iron crystal lattice by spraying iron ions a few at a time on a template. While the outcome is the same, here the active 'self' is missing. All things considered, it is probably most useful to consider self-organisation to occur where (and only where) a system bifurcates, sufficiently under its own dynamics, so as to bring to bear an additional system-wide constraint (or at any rate an additional multi-component, that is, relatively macro, constraint).[42]

The formation of a new relatively macro constraint, however brought about, creates a new level proper in the system, since the constraint now filters out more microscopic relation detail incompatible with it. The iron crystal lattice, e.g., filters out thermal fluctuations and many external perturbations, dissipating their energy as lattice vibrations. (Otherwise the constraint would not be stable against

[42]See [Hooker-c, this volume]. Some may prefer to refer to the formation of a new 'top-down' constraint or to 'downward' causation [Campbell, 1974; Bickhard, 2000a], to emphasise the new dynamical entity as outcome, its component dynamical interrelations constrained by its macro dynamical form. Others may instead prefer to refer to mutually constrained behaviour in order to emphasise that the emergent dynamical change is produced by the way interactions among constituents issue in mutual constraint. Cf. [Craver and Bechtel, 2007; Emmerche, *et al.*, 2000; O'Connor and Wong, 2002]. But these two are in fact equivalent, since the mutual constraint is macroscopically real, producing new forms of work. However, what counts as "sufficiently under its own dynamics" to justify the 'self' in self-organisation can be a vexed matter. The pan of heated fluid counts as self-organised because, while the applied heat is key to 'forcing' the dynamical changes, it is applied sufficiently externally to the fluid dynamics. If we deliberately add species to an ecology until it achieves a certain resilience to drought, on the other hand, it would not be considered to *self*-organise that dynamical form transition. The system could also not be considered to simply have been organised either, because the outcome may not be increased organisation — see 'organisation' below. (Though, speaking colloquially, we could say that we 'organised' it, we would only sometimes be technically correct.)

microscopic-originated perturbations and similar external disturbances.) The iron becomes a 2-level system, (1) that below the level of the lattice, the individual ions and electrons, obeying their dynamical interaction laws, and (2) that at the lattice level with its fermi conduction band where electrons stream through the lattice, the lattice collectively vibrates, and so on. This is a dynamically well-defined and grounded notion of 'level', all other uses are for gravitation (e.g. level table) and measurement (e.g. flood level) or are metaphorical (e.g. abstraction level) or confused. The constraint sense of 'level' is the only use made of the term in this essay and [Hooker-b,c, this volume].

Emergence. When the outcome of dynamical interaction among system components is surprising or unexpected or too complex to readily understand, scientists are apt to talk about emergent patterns. However this is a vague, shifting and subjective approach. (In colloquial usage 'emerge' often means no more than 'becomes perceptible'.) Limiting emergence to the appearance of a phenomenon that could not have been predicted from knowing only the pair-wise dynamical interactions of components sharpens it somewhat. But this still ties the definition of what is a physical property to a non-dynamical criterion (prediction). Indeed, since prediction is often limited, many behaviours would count as emergent just on these grounds. But there is instead the alternative option of pursuing a dynamical criterion. A clear, wide criterion would be to identify emergence with bifurcation generally, a clear narrower one would be to identify it with just self-organisation. In each case a new dynamical form does come into being. Other criteria are possible (see section 2 above) but these are the clearest, simplest dynamically grounded ones. And, if the epistemic and semantic criteria are removed from the section 2 list, then plausibly the remaining criteria all work off one or other of these two dynamical criteria. [Hooker-c, this volume] argues for the wider usage on the grounds that emergence is concerned with new dynamical character as outcome, not the process producing that result.

Constraints — enabling and coordinated. The term 'constraint' implies limitation, most generally in the present context it refers to limited access to dynamical states. Equivalently, it means reducing degrees of freedom by limiting dynamical trajectories to sub-sets of state space. This is the common *dis*abling sense of the term. But constraints can at the same time also be *en*abling, they can provide access to new states unavailable to the unconstrained system. Equivalently, by coordinatedly decreasing degrees of freedom they provide access to dynamical trajectories inaccessible to the unconstrained system. Thus a skeleton is a disabling constraint, for example limiting the movements of limbs (cf. an octopus), but by providing a jointed frame of rigid components for muscular attachments it also acts to enable a huge range of articulated motions and leverages, transforming an organism's accessible niche, initiating armor and predator/prey races, and so on. Each of the eight great transitions in evolutionary history [Maynard-Smith and Szathmary, 1995], e.g. the emergence of multi-cellular organisms, marks a new coordination of constraints. By permitting reliable cooperation instead of competition and reliable inheritance of the fruits of cooperation, the new coordinations created new complexity and opened up vast new possibilities (cf. [Sterelny, 2007]). Coordinated constraints can work their way around physical laws. For instance, while no single pump can lift water higher than 10 metres, trees lift it many times this by physically linking together (coordinating) many cellular pumps.

It is possible to obtain complex dynamics in simple systems (such as logistic reproduction). However, plausibly the only way in which the complex properties to follow can be obtained is through complex coordination of constraints of the kind neural, muscular and skeletal coordinations exemplify. These have their origins in the complex coordination of biochemical products and gradients that allow intra-cellular chemistry to support cellular maintenance. We are here far from the holonomic constraints and, as cellular regeneration shows, the no-work on constraints condition of standard analytical mechanics and deep into the domain of multiple state-dependent, interacting non-holonomic constraints.

Intrinsically global coherence. In analytic (Lagrangian) dynamics the globalness or otherwise of constraints is not directly considered. A holonomic constraint provides an inherently global geometrical constraint on motion in the sense of being specified everywhere, but not in the sense of demanding internal global coordination of variables. Some holonomic constraints may force component motions to be globally correlated, others will not. The same applies to non-holonomic constraints. Moreover, these can be partial rather than global, with a dynamic network of constraints structuring system dynamics, as in the cell. But if a system is to perform a global function, e.g. metabolic regeneration or amoebic chemotaxing up sugar gradients, then this will force a global organisation of its components to achieve it. Thus underlying a global functionality must be a global constraint on dynamics that ensures realisation of the function. This must be so even when this constraint is realised through a network of state-dependent, interacting non-holonomic constraints, e.g. a network of work-constraint cycles as in the cell [Kauffman, 2000]. Multiple global functionalities, characteristic of living systems, require multiple underlying global constraints, and these will normally be significantly but subtly interrelated in order to allow multiplexing (many component roles combining to realise a single function) and multitasking (the one component playing roles in realising many functions). Multiplexing and multitasking are attractive because they reduce the number of required components while increasing system functionality and adaptability, and possibly evolvability, albeit at the expense of increasing system complexity, and possibly also increasing system instability and/or rigidity.

Order and organisation. The constraints underlying global functionality require global organisation as distinct from global order. A high degree of orderedness means internal uniformity (e.g. a crystal) while functional organisation requires inter-articulation of distinct components (e.g. in a motor vehicle engine).

The root notion of order is that derived from algorithmic complexity theory: the orderedness of a pattern is the inverse of the length of its shortest, most compressed, complete description; hence gases, being internally random, are disordered and regular crystals, being internally uniform, are highly ordered, but neither displays any functional organisation. Sometimes complexity is taken to be measured by the inverse of algorithmic orderedness, but this leaves gases the most complex systems; in short, it ignores organisation, the key to living complexity. Machines and living things are organised because their parts are relatively unique and each part plays distinctive and essential roles in the whole. The colloquial use of 'organisation' is broad and vague, though its core examples are functionally organised (engines, firms, rescue teams, ...). In this essay and [Hooker-b,c, this volume], the use of 'organisation' is restricted to functional organisation as characterised above.

In another use of 'order' entirely, talk of high order features refers to features characterised by high order relations, that is, relations among relations among ... Then organisation is a particular kind of ordering in this sense, involving relatively high order relations that characterise many nestings of correlations within correlations. (Think of correlations within and between the motor, electrical management and drive chain modules of a car.). That is, an organised system displays a non-redundant global ordering relation of relatively high order, including global (sub)relations characterising global functions. For this reason organised systems must be less highly ordered than are crystals (but are obviously more highly ordered than a gas). A system's organisational depth is measured by the degree of nesting of sub-ordering relations within its global ordering relation (cf. cells within organs within bodies within communities). Living systems are deeply organised. However, organisational depth also does not fully capture complexity.[43,44]

---

[43]Because, e.g., it does not capture the distinctiveness of nested relations and the top-down relatively global constraints that modulate them. Also, in dropping the algorithmic conception, it loses 'horizontal' relational complexity. Gell-Mann [1994] discusses effective complexity and logical depth (see [Bennett, 1985; 1992]) and Type 2 theories (see [Marr, 1982]) as other possibilities for measuring organised complexity but neither is satisfactory for various reasons he notices — fundamentally for the above reasons. For general discussion of these issues see [Collier and Hooker, 1999, sections III and VI].

[44]Shalizi [2001; *et al.*, 2004] also claims to provide a measure of organisation, in this case a statistical measure. What is actually measured is called system complexity, designated 'C', and is, roughly, the minimum information required to specify those past system behavioural features (variables plus inputs) that make a difference to future system behaviour. (Shalizi calls these the 'causal states' of the system; inputs are not mentioned but without these, reflecting increasing environmental responsiveness, the effect of increasingly nested correlations might simply be more constrained behaviour and so reduced complexity.) Shalizi has a more general conception of organisation in mind — something like number of partial interdependencies or correlations — than the functional organisation characteristic of metabolism which is concerned with sets of partial correlations satisfying global constraints (so that they sustain functionalities). This latter is what this essay understands as organisation proper. For this the central issue is whether the C measure differentiates between organisation (proper) and other kinds of interdependencies among variables, e.g. those in turbulent flows. I don't see that it is equipped to do so, since it simply counts interdependencies of all kinds. An aeroplane and a turbulent river could be judged alike in terms of numbers of required variables for behavioural prediction, but would internally be organisationally very different. Unsurprisingly (because C picks up only interdependency),

Modularity. A system contains a module if (and only if), to a sufficiently good approximation (e.g. to capture essential system functionality), its dynamics can be expressed as an interactive product, the dynamical product of its intra-modular dynamics and its inter-modular dynamics.[45] Three kinds of modularity can be distinguished, spatial or 'horizontal', level or 'vertical', and process modularity, labelled respectively S, L, and P modularity. Smodularity obtains when there is a principled division of a system into contemporaneous spatial modules such that the system dynamics is expressible as the product of the individual module dynamics and their interactions. This is how we currently design and model buildings and machines of all kinds (from homes to hotels, typewriters to television sets) and how we usually attempt to model both biological populations (the modules being the phenotypes) and often their individual members (the modules being internal organs, or cells). Lmodularity, in contradistinction, obtains when a system's dynamics may be decomposed into the interactive product of its dynamics at different system constraint levels. This is often how business management functionally analyses a firm. (Often that organisation will express a management hierarchy and be graphically represented vertically, often realising the functional roles vertically in a building — hence the alternative 'vertical' label.) It is also often an important part of machine design (cf. motor vehicle electrical regulation and drive chain modules) and of the scale analysis of organisms (cells, organs, organism). Pmodularity obtains when a system's dynamics may be decomposed into the interactive product of its process dynamics, and is characteristic of the analysis of organisms and complex machines into mechanisms, such as cellular respiration and pulp mill regulation.[46]

---

this parallels dynamical phase space blindness to organisation, where a pendulum and a feedback oscillator are both represented by the same wave equation. C may provide a measure of interdependency, but not of organisation (proper) — and thus not of complexity either in the sense later articulated. Of course, if the behavioural data were to include sufficient data about the internal state, as it happens the pendulum and the feedback oscillator could be behaviourally distinguished even though the inherent organisational blindness persists. Exploiting that idea, if it could be assumed that sufficient internal observation would result in every distinct organisational form having its own distinct C measure, then that would suffice to at least identify the organised systems, if not order them distinctively. (Cf. the fate of Behaviourism in relation to Cognitivism, section 4.2.4 below.) Cyclical temporal interdependencies of the kinds that characterise many metabolic processes (e.g. Krebs cycling) can certainly be detected with appropriate internal observational data (using temporal autocorrelation in the data stream), and also their temporal nestings. Indeed these interdependencies will already form part of the C total if the scope of observation includes them. But disaggregating them within total C may be very difficult.

[45]Then all system components, at whatever level of analysis, are modules, with basic components being those taken to have no internal dynamics and fixed inter-modular dynamical interactions. This is a physicist's definition. For an inequivalent and weaker but related computational definition in terms of network connectivity, see [Green and Leishman, this volume, 'Encapsulation and modularity']. The conceptions are inequivalent because interaction strength is not equivalent to network connectivity or even to information transmission densities. The dynamical conception is stronger because it can distinguish L from S and P modularity, which is difficult to do in purely network connectivity terms. As an information transmission conception the network conception is most closely related to P modularity.

[46]Note that each of S and L modularities define a relative, interactive conception of 'part'. (This is clearer for the network interaction conception, see note 45.) But parts may interact strongly,

As motor vehicle design illustrates, all three modularities may be combined, at least to significant extent, in current simple engineering designs. Here S and L modularity will create the constraints to enable corresponding functions in simple, reliable ways (though in the process disabling many others). The earlier viable system movement that sprang from cybernetics and general systems theory (see Beer and others, section 2) relied on such designs. But modular processes may also cut across levels and spread throughout a system and to that extent exclude S and L modularity. Modularity of any kind reduces system complexity, by decreasing dynamical degrees of freedom, while increasing functional and possibly developmental reliability and ease of repair. Like any coordinated constraint it will in general both disable and enable and thus have a complex relationship to higher order system properties like multiplexing and multitasking, adaptability and evolvability. Nor is there any simple relationship to reduction; simplification and decomposition aid reduction but if the decomposing modularity is achieved through self-organisation (e.g. in a Bénard cell system) then it also thwarts reduction.[47]

Hierarchy. Hierarchy proper is asymmetry of level (vertical) control in a sufficiently Lmodular system. While common in machine design and as underlying principle in organism and institutional design, pure hierarchy is in fact the exception, and is rarely more than partial in living systems. A higher level constrains lower level dynamics (as a crystal lattice constrains the behaviour of its atomic constituents), will often regulate it through feedback (e.g. coherence of crystal vi-

so neither S nor L (nor P) modularity is equivalent to an ontological part/whole relation, though on occasion they may have this effect. Likewise, none of them corresponds to a general/special (genus/species) hierarchy, though on occasion they may have this effect. (These latter distinctions may be less clear in the network conception.) Hmodularity in [Collier and Hooker, 1999] = Smodularity here.

[47]Modularity is one form of system composition, that is, of the relationship(s) between parts and whole. Another — aggregate composition — satisfies the condition that the whole is the simple sum of its parts. Various attempts have been made to decompose the non-aggregative systems, e.g. into nearly and minimally decomposable compositions [Simon, 1969; 1997] or component and integrative compositions [Bechtel and Richardson, 1993]. These ideas have to be applied carefully. For Simon's idea of near-decomposition, e.g., a Bénard cell system can be approximately described as a combination of separate cells, at least *while the producing conditions apply and the system is kept away from thresholds*. But move near or over a threshold and the continuing underlying global complexity of the system makes itself apparent. Again, one reading of the component/integrative distinction, e.g. by [Boogerd, *et al.*, 2005] who discusses it for cell biology, is in terms of the degree of alteration of a component's interactive capacities on entering a compositional relationship. But this applies only to sufficiently complex components like proteins (and organisms entering a culture — see section 4.1.2 below); the wider condition is simply top down constraint, corresponding to the alternative reading of the distinction, and providing a significantly different categorising of composition conditions. In sum, while there is a general intuitive sense of increasing mutual constraint from aggregative to nearly-decomposable/component to minimally-decomposable/integrative compositions, there are many different forms of mutual constraint, e.g. global constraints versus modularity, including all the collective properties discussed in this section. Whence there is no obvious simple ordering of their manifold possible compositions. The upshot is that these distinctions don't seem amenable to sharp dynamical characterisation and will not be pursued.

brations) and sometimes it will also control the lower levels in important respects (top down control, e.g. brain control of muscle). But it will also typically be true that lower level dynamics will constrain higher levels (as electron orbital dynamics constrains crystal angles), may regulate them through feedback (e.g. in catalysis of chemical reactions) and might control certain aspects of the higher level (bottom up control, e.g. indirect control of volume Hebbian learning through local NO release). And in many cases there will be no control asymmetry involved, simply mutual constraint through interaction, e.g. of oscillatory behaviour in a system of small oscillating springs connected to a common rigid, but moveable, bar. This latter interaction-only condition will be the common case among Smodular components at the same level, e.g. among cells of the same organ. Whenever there is a control asymmetry we shall speak of hierarchy relationships, with the direction of the hierarchy being the direction of control asymmetry.[48]

Path-dependence and historicity. Path-dependence occurs whenever there is positive amplification, for then initially nearby dynamical trajectories subsequently diverge as a function of small differences in their initial conditions, so the path taken depends on precisely where the first step began. A notable sub-class of path-dependencies are those where, once begun, development along a certain path itself becomes increasingly entrenched. This applies where, e.g., an initial fluctuation is amplified and entrenched, especially where that entrenchment involves a bifurcation that reinforces the irreversibility of the development. Examples include a particular impurity site of first freezing or of rolling boiling, a first genetic mutation that yields a distinctive kind of adaptive advantage or a first oil discovery or shop in a new suburb that transforms a local economy. These cases exhibit a clear sense of historical possibilities exploited, and correlatively of others foregone, and their resulting paths are often said to 'fix' their initial historical conditions.[49] By contrast, for stable systems in an attractor basin there is no overall path-dependence since the same outcome occurs (capture by the attractor) for all beginning points (initial conditions).

Constraint duality and super-system formation. Coordinated constraints that enable while disabling, e.g. the disabling movement constraints imposed by a skeleton and its enabling of locomotion and leverage, exhibit general constraint duality. The notion has a specific application to forming systems into

[48] Cf. [Eldredge, 1985; Nicolis, 1986; Pattee, 1973; Salthe, 1985; 1989]. Dyke's [1988] usage is confined to the rare special case where constraint is one-way only. Commonly among living and human engineered systems a hierarchy is specified by assembling cohesive, Lmodular combinations of Smodular components, e.g. building organs from cells and bodies from organs, cf. [Ravasz, *et al.*, 2002]. But it is possible to have at least dynamically distributed feedback regulation that doesn't require S or L modularity, e.g. the distributed rate dependency phase separation of Belusov-Zhabotinsky chemical reaction structures.

[49] However, it would stretch the notion of *physical* constraint to vacuity to call these initial conditions path constraints, because there is no cohesive force involved that grounds the constraint.

super-systems through mutual interaction. System constraints may contribute to enabling super-system capacities, for example the role of mitochondria in eukaryote energy production. Conversely, super-system constraints may free up system constraints, for example wherever multi-cellular capacities permit member cells to specialise. But it has a wider application in considering social community formation. We can gain a crude measure of the importance of socialisation to a species by considering the ratio of *usable* individual parametric plasticity (i.e. adaptiveness) between isolate and communal states. For simpler creatures of lesser neural capacities and more rigid social organisation, such as the insects, the relation is typically negative, individual capacities are sacrificed to communal cohesion and function. Oppositely, humans increase their coherently usable individual capacities enormously through collective culture, even while contributing to communal capacities. Unless we humans have a sophisticated, high quality cultural environment in which to develop there will be vast reaches of our somatic, especially neural, organisational space that we cannot use because it is not accessible to us. Thus for humans there is a positive relationship between individual and communal capacities. Coupled constructively, each enables the other, and together they largely (not wholly!) dominate those coupled competitively. (We can speculatively picture this effect increasing through the mammalian line as brain size, intense socialisation and intentional action increase together.) This makes all the difference to the power of culture, to its significance in adaptive evolution, and to the intricacy and globalness of its organised dynamics (cf. Section 4.1.2 below).

Coordinated spatial and temporal differentiation with functional organisation. Multiplexed, multitasked functions cannot all be realised simultaneously at every location, the resulting interference would render reliable performance impossible. It is necessary then to distribute the realising dynamical activities spatially and temporally so that each local area over each process cycle is restricted to a coherent set of concurrent activities. Moreover, these distributions have to be subtly organised so that each function is realised at convenient locations and times for receiving its inputs and also useful locations and times to contribute its outputs. Similarly, cellular metabolism requires a group of closed self-reproducing processes to recreate the constraints for each process from the products of other processes — Kauffman's [2000] work-constraint cycles — and this too requires subtle spatial and temporal differentiation to achieve reliability and effectiveness.

Multi-level and multi-order functional organisation. Metabolism, for example, refers to the organised network of biochemical interactions that convert input matter and negentropy (food and water) into usable forms and direct their flows to various parts of the body as required, for example for cellular respiration. The individual biochemical reactions are largely known. However it remains a challenge to characterise multi-level processes like respiration, comprising processes from intra-cellular Krebs Cycles to somatic cardio-vascular provision of oxygen and removal of carbon dioxide. These processes must be made coherent across the entire body for respiration to work. The overall process of respiration is *multi-level*: involving sub-cellular to organism coordination, *multi-dimensional/plexed*: involving organised interactions among many body parameters, *multi-modal/tasked*: involved in many different bodily modes of operation (motor, cognition, stress, etc.), e.g. the cardio-vascular system simultaneously transports resources (oxygen etc.), wastes (carbon dioxide etc.), regulatory hormones, and so on, and *multi-phasic* (asynchronous and non-stationary): respiratory processes occur on many different timescales, with local parameters constantly changing functions of temporary activity while more global respiratory parameters are functions of the longer term developmental and subsequent functional history of the organism. In this conception, global coherence is a result of internal regulation at various functional orders and scales.

Autonomy. Autonomy is a particular global functional coherence. It is the internally organised capacity to acquire ordered free energy from the environment and direct it to replenish dissipated cellular structures, repair or avoid damage, and to actively regulate the directing organisation so as to sustain the very processes that accomplish these tasks. There are two broad cyclic processes involved, internal metabolic interaction and external environmental interaction, and these need to be coordinated: the environmental interaction cycle needs to deliver energy and material components to the organism in a usable form and at the times and locations the metabolism requires to complete its regeneration cycles. At the same time the metabolic cycle needs to include a capacity to regulate the metabolic organisation so that both the external interactive capacity and its own regenerative capacity are sustained. (E.g. a cell can alter an internal chemical balance or gradient by ionic intake or expulsion and respond to environmental deficiency by tumbling.) This organisational regulation needs to be coordinated with the basic regenerative and interaction cycles; it may be fused with the regenerative process. The presence of these two thus synchronised cyclic processes resulting in system regeneration is the broadest functional sense of what is meant by a system's being autonomous.

Though the detail, especially the dynamical boundaries, vary in graded ways across living organisms, this autonomy requirement picks out all and only living individuals — from cells, to multicellular organisms to various multi-organism communities, including many suitably organised business firms, cities and nations. But because only autonomous systems have their functions serve their own physical regeneration, in turn supporting their functioning (they are 'recursively self-maintenant' [Bickhard, 1993]), they represent a distinctively new category of complex system organisation. (Though all living systems are ipso facto non-linear, irreversible, open, self-organising, globally constrained, etc. systems, non-living systems may also manifest one or more of these properties, cf. [Fox Keller, 2007], but not autonomy.) For the same reason, in all autonomous systems the locus of living process regulation lies more wholly within them than in their environment — hence it provides a root sense of autonomy that supports the traditional sense.[50] Birds organise twigs to make nests, but twigs themselves have no tendency to organise nests or birds. Entities with a distinctive wholeness, individuality and perspective in the world, whose activities are willful, anticipative, deliberate, adaptive and normatively self-evaluated, are properly treated as genuine agents; autonomous systems are inherently all of those things (section 4.1.1 below).

Adaptation, adaptiveness and learning. An organism is adapted when it possesses an autonomy-satisfying set of traits in its life-environment. Conversely, an organism's ecological niche is comprised of the range of life-environments for which its traits provide satisfaction of autonomy. An organism's adaptiveness is its capacity to alter its specific traits in mutually coordinated ways so as to adapt to, that is, satisfy autonomy in, different life-environments. Humans can run as well as stand still and this enlarges the range of prey they can catch, predators and natural disasters they can evade, and social commerce they can sustain. Shifting from standing still to running involves coordinated changes in physiological processes, sensori-motor feedback/forward foci, etc. The set of coordinated trait variability ranges consistent with autonomy-satisfaction comprises an organisms' adaptive envelope. An organism's adaptability, adaptive potential, or adaptiveness is some measure of this set. Learning, understood most generally, is the application of adaptability to develop adaptations. Ecologies and evolving populations learn only in this broadest sense, through changes to the internal compositions of their populations. Understood more narrowly, learning is the broad process manifest through internal sensory, memory and motor regulation, that is, through neurally modulated behaviour. Organisms learn in the narrower sense, which generally offers more powerful, but less broadly applicable, problem-solving capacities. Broad and narrow learning processes can be complexly combined and it can be instructive to inquire how various community groups, from colonies and flocks to human business firms, cities and nations, learn.[51]

To these properties of complex systems is added two further properties that might equally be said to characterise our study of them, though here they are considered as properties of the systems themselves.

[50] On autonomy and self-maintenance see further section 4.1.1 below and [Bickhard, 1993; 2000b; Christensen and Hooker, 2000a; 2002; Moreno and Ruiz-Mirazo, 1999; Moreno and Lasa, 2003; Ruiz-Mirazo and Moreno, 2004; Moreno, 2007; Moreno *et. al.*, this volume], and references. This includes reference to the root preceding tradition [Varela, *et al.*, 1974; Varela 1979], now a main alternative, autopoiesis [Maturana and Varela 1980; Maturana 1981]. Self-governance lies at the core of our commonsense conception of autonomy. However, we are most familiar with the idea of autonomy as applied to persons and political governance. But these are sophisticated notions applied to sophisticated systems whose trappings may distract from fundamentals. We need to return to basic principles operating in all living systems to construct a naturalist notion that will 'grade up' across the evolutionary sequence to our sophisticated concept.

[51] For a more detailed articulation of a basic functional organisation underlying cognition, including anticipation, normative evaluation, self-direction and self-directed anticipative learning, see section 4.1.3 below and e.g. [Christensen and Hooker, 2000a; 2000b; Farrell and Hooker, 2007a; 2007b; 2009; Hooker 2009b; Hooker and Skewes, 2009].

Model centredness and model specificity/model plurality. Complex systems of the foregoing kinds are typically characterised by dynamical equations that lack analytical solutions. Thus it is necessary to explore their dynamics through computational simulation. This places computational modelling at the centre of their scientific investigation in a strong manner. Model centredness refers to this feature. It highlights the unique contribution of computers to cognition (all its other uses being pragmatic, if often valuable). Since all computer modelling is finite, all quantitative models have inherent uncertainties in them, e.g. 'rounding off' errors. Further, mathematical science contains many non-computable mathematical functions, that is, functions where information crucial to identifying it is lost with any degree of finite approximation.[52] This provides a practically ameliorable, but ultimately ineliminable, basis for uncertainty about system state whenever the system dynamics involves positive amplification. Moreover, different implementations or simulations of a given theoretical model can exist, e.g. employing different machine architectures and programming languages, and these need not be equivalent in expressive power, accessibility or performance (see, e.g., [Axelrod, 1997]). These variabilities represent a further source of uncertainty in modelling. They can be removed by demonstrating that the relevant outcomes are independent of implementation, or by restricting data to what is robust across implementations (or by requiring that there is available a special defence of a particular simulation), but this is seldom done, not least because the list of simulations to be checked is indeterminate. Uncertainties eventually translate into errors, if they are ignored or arbitarily resolved (e.g. by rounding off). Thus though computational numerical approximation represents a huge expansion of our capacity to know complex dynamics, it also represents a selective, but important, diminution in our knowledge capacity. On the other hand, the centrality of computational modelling provides the platform for making new complexes of modelling methods, e.g. use of evolutionary game theory and genetic algorithms. The full extent of these supported methods needs exploration.

Since it is frequently impossible to model the full complexity of systems, it is necessary to choose partial parametric models aimed at capturing the basis of some class of phenomena. Model specificity refers to the capacity to select parameter values so as to specialize the model to the characterization of some unique individual and/or situation. Conversely, model plurality refers to the capacity to capture the characterization of a plurality of individuals/ situations within its parameter ranges. These features are respectively the basis for deducing feature ranges in

[52]See e.g. [PourEl and Richards, 1989; Shipman, 1992] Many superposition or 'wave' phenomena (classical and quantum) are of this kind where wavelet information at indefinitely small scales is important to identifying the whole function. Here a mathematical function is a many:one map from a domain to a range, hence unique on the range. One distinctive merit of the proposal to model natural (e.g. biological) functions as input/output maps is that this relates them directly to mathematical functions and hence, via modelling, to dynamical maps and so to biochemical processes. For some instructive exploration of computation in science see Humphries 2004, the *Proceedings, Workshop on Physics and Computation*, IEEE, 1992 generally and `http://en.wikipedia.org/wiki/Digital_physics`, `http://wapedia.mobi/en/Church-Turing_thesis?t=9`).

individuals from more broadly characterised systems, e.g. populations, and for formulating valid generalisations across sets of individuals, e.g. across populations. And this in turn is the basis for condition-dependent laws that are the norm in all such complex systems domains.

Condition-dependent laws. Compared to the universal, invariant laws of physics, the local idiosyncratic behavioural patterns exhibited by many complex systems don't seem to qualify as laws. Of course biology and cognate sciences dealing with complex systems do use universal laws in constructing their models, e.g. if the elemental laws of chemistry did not operate the same everywhere, bio-chemistry and hence biology would be much harder than it already is. Even so, the complication arises from the fact that the nomic invariance largely occurs at the ion-ion interaction level but how n-body, k-component ion systems operate is often a sensitive function of the initial and constraint conditions, especially organisational conditions, obtaining. The bio-chemistry of carbon-hydrogen chains provides eloquent illustrations, for instance where folding history can alter subsequent interaction dynamics (e.g. via which binding sites are active). That is why no simple set of laws can be deduced in advance. However, the phenomenon is universal within science. Consider that, e.g., electric circuit dynamics is the outcome of many universally lawful component interactions, but it is the material circuit conditions that determine the outcome circuit law, e.g. whether it is oscillation, exponential decay or other. These circuit-design-dependent dynamics are properly considered law-like despite arising from specific material conditions (see any engineering text on oscillators or circuit design).

Moreover the point extends to laws conditioned on self-organisation, a distinctively complex systems circumstance. Pertinently, self-organisation precisely occurs because of the sensitivity of dynamical form to dynamical initial and constraint conditions (see above). But since self-organisation involves a new dynamical form, it is reasonable to say that it obeys new dynamical laws characteristic of that form. For instance, consider this condition: a cooling mold of liquid iron in contact with a heat reservoir of lower temperature. It leads to new emergent laws — rigid body (not fluid) dynamics, crystalline (not fluid) conduction of electricity, heat and sound. The universality requirement drops out.[53] Then we are free to see biology and cognate domains as replete with real laws, just as is physics. It is at most that they will be condition-dependent on highly intricate, perhaps idiosyncratic, and typically organisational, conditions.[54] (They will be 'special', as some philosophers say.) Also, often their form will be hard to predict but, so far at least, this distinguishes them from the situation in physics by at most a matter

[53]In any case, the requirement of universality, that is, to be specified independently of any initial and constraint conditions, is a conceit of fundamental physics, and perhaps ultimately not true there either considering that even fundamental laws evidently changed form as the big bang cosmos cooled.

[54]See [Polyani, 1968] which offers an early diagnosis of the condition-dependence of laws for biological organisms, and argues that their peculiar conditions (essentially autonomy, see above) makes reduction impossible.

of degree, not kind.[55]

This discussion provides a rich set of concepts and principles for characterising complex systems. For instance, Mayr [2004] provides a list of features distinctively characterizing biological systems and which he claims sets biology apart from other natural sciences: metabolism, regeneration, growth, replication, evolution, regulation, teleology (both developmental and behavioural). However, comparing his list with that of complex systems features above it is pretty clear that, at least in principle, complex systems provide resources for modelling, and hence explaining, each of them and molecular, systems and synthetic biology are between them well on the way to doing so. Again, the Santa Fe Institute was an early pioneer, dedicated to the explicit study of complex systems, its inaugural volume [Pines, 1985] providing a first cross-disciplinary preview of their diverse applicability. Its orientation then was summed up by Gell-man: the study of n-body dynamics of simple rules that generate complex behaviour, exhibited through such paradigm technical tools as cellular automata and Boolean nets. Today we can see that, while these systems comprise a distinctive group of typically counter-intuitive cases, and constitute rich worlds of phenomena[56], we now include within the scope of complexity all the dynamical behaviours associated with all of the complex systems features just reviewed. From this perspective the present volume is a characteristic successor to that early volume. The richness of its modelling and methodological and epistemological considerations reflect the hugely expanded conception of complex systems and vastly increased breadth and depth of their impact across all the sciences.

Despite the richness there are certainly further notions that might have been included. Some are constituted in more fundamental notions and are left to exposition elsewhere. Synergy [Haken, 1983; 1993], e.g., though of interest, results from some combination of positive feedback and coordinated constraints (in some cases with constraints and negative feedback to stabilise it). However, two fundamental notions concerning the basic metaphysics of complex systems are perhaps surprisingly absent: identity and complexity. They are absent because, like emergence and self-organisation, they are labels for a 'swamp' of diverse, and sometimes confused, ideas; e.g. Edmonds lists nearly 400 complexity references with great diversity among them, and that list is only partial.[57] Moreover, unlike emergence and self-organisation, there seems relatively little of a science-based, dynamical

[55] See further section 4.2.2 below. Note that, although they may seem to be closely related or even the same thing, condition-dependent laws are really the opposite of ceteris paribus laws [Earman, *et al.*, 2002]. The latter state that some reference condition occurs in standard circumstances, the exceptions picked up, otherwise unidentified, by the ceteris paribus collector. Sometimes this can be practically useful, but unless the exceptions are due to irresolvable noise, it represents a halt to scientific progress. Whereas the point of condition-dependence is to decompose the exceptions into differing conditions that change the law in systematic ways, neither random nor ignorable but at the heart of understanding the dynamics of the systems involved.

[56] Even sufficient for Wolfram to claim a 'new science' for cellular automata simulations, see [Wolfram, 2002].

[57] See `http://bruce.edmonds.name/combib/`, cf. [Bishop, this volume, section 2.5] and references.

nature that can sensibly be said about structuring these concepts.

Some remarks on identity are found in [Hooker-b, this volume, section 6.1B], while [Cumming and Collier, 2009] and [Bishop, this volume, section 2.7] show the difficulties involved in trying to develop more detailed systems identity criteria. As for the notion of complexity, the collection of concepts and principles gathered above should suffice to indicate the diversity and complexity of what needs to be covered — this is why its discussion was postponed until now. The same collection also suffices to show the insufficiency of any of the commoner simple ideas of complexity — number of degrees of freedom (component numbers), algorithmic incompressibility, levels, ... — to capture the notion by themselves. To approach its characterisation, first omit all epistemic notions as ultimately derivative considerations, e.g. those appealing to intelligibility and surprise, and all 'external' notions like controllability. Then at the least complexity has, I suggest, five quasi-independent dimensions to it: cardinality (component numbers), non-linearity (of interaction dynamics), disorderedness (algorithmic incompressibility), nested organisation (organisational depth) and global organisation. While I agree that these 'dimensions' are not fully mutually independent, it is also true that some are partially opposed and I do not know how to systematically characterise their inter- and counter- dependencies so as to produce a clearer, reduced dimensional analysis. I do not know of any one measure that can capture all these dimensions and I do not think anything less rich is adequate to the forms complexity can take.

In particular, trying to simplify the problem by leaving out globally organised constraints, when they are central to functional characterisation, much less by ignoring organisation entirely, is to aim at capturing a physics-centred class of cases at the expense of missing much of the special importance and challenge of complexity. Conversely, leaving out the forms of non-linearity involved may still allow structure and patterns to be discussed, e.g. bifurcation thresholds and attractors, but in omitting their quantitative character and distribution it misses understanding the roles and complications of noise, resilience and the like. Though aspects of this complex concept have large technical literatures associated with them, I suspect that at the present time the foregoing is the best that can be managed concerning characterising complexity in its largest meaning.

There remains then to tackle the question of what a science using such systems models and studying corresponding systems might be like in its foundations, philosophy and practice. Just this is of course the point of this volume. In [Hooker-b, this volume] some of the more general themes will be picked out. Meanwhile, there follow some particular topics that should ideally have found a place among the essays but did not do so.

4 SPECIAL TOPICS

4.1 Three interrelated complex systems ideas that challenge orthodoxies

Each of the complex systems ideas to follow is in its infancy. The first, autonomy, is sufficiently well developed to understand its significance, but its domain, biology and bio-sociality/robotics, has yet to really feel its impact. The second, cultural dynamics, has barely been identified and has yet to receive sustained exploration. The third, science dynamics, has a range of recognised models offering various expressions of it, but arguably these are all preliminary models on the way to a more mature expression. Nonetheless, these ideas represent the application of complex systems to three of the most basic features of human life, our constitution, culture and cognition. As such, they are worth exploring in themselves, to help correct the relative neglect of these domains within complex systems development thus far, and to expand the range of our understanding of what the introduction of complex systems means for scientific understanding, the theme of this book.

4.1.1 Autonomy

The first of these ideas is autonomy, the coordination and regulation of the internal metabolic interaction cycle and the external environmental interaction cycle so that the latter delivers in timely manner to the former the resources it requires to regenerate the organism (section 3). This idea, which is uniquely tied to complex organised systems, has manifold significances for science, of which four will be briefly reprised here.

The *first* was noted in section 3: it permits the demarcation of living systems. It is by no means the only approach to demarcating life and readers are encouraged to consider others. (See [Bedau, 2007] for a review of approaches.[58]) However it is deeply related to the unique organisation of living systems (cf. [Ruiz-Mirazo, *et al.*, 2004]), while not tied too closely to their chemical detail. And it uniquely picks out the living systems from within the wider domain of complex, organised, non-linear, dissipative (entropy increasing) and irreversible, chemical and biological systems. Whence it provides an unbiased, operational criterion of life hitherto missing and especially needed in exo-biology, where physical criteria (typically those above) are too wide and psychological criteria (intelligent, etc.) are too narrow. Moreover, though the detail, especially the dynamical boundaries, vary in graded ways across different living systems, autonomy has the advantage of not being confined to cellular organisms but extends to suitably organised social groups, where it picks out living individuals, from cells to multicellular organisms to various (by no means all) multi-organism communities, including many business firms, cities

[58] An informal, less structured but still useful starting review is `http://home.planet.nl/~gkorthof/korthof66.htm`.

and nations.[59] Many otherwise plausible-looking systems are not autonomous: all viruses, candles and hurricanes and the like, most or all ecologies, most social institutions (community sports, cafes, markets, ...), many business firms, and so on. These all lack sufficient organisational capacity to regulate themselves.[60] In consequence, Smith's hidden hand of the market has at best a weak existence (there is but a relatively scattered set of feedback loops, though some of them may be of great importance, and more are being added), as does Lovelock's Gaia for similar reasons. But at least we have a principled basis for analysing the vulnerabilities of these systems and hence for shaping our management policies toward them. Similarly, business firms too can benefit from a systematic understanding of their organisational deficiencies as viable enterprises, as can many other social institutions (though not all need to be so persistently focussed on survival as autonomy forces).

Although all this is so, there remains an important issue concerning precisely how to demarcate autonomous systems. The cell, but not the candle, regenerates itself (metabolism), including its own metabolic organisation to do this, and self-regulates that organisation to aid its continued viability (note 60). If complete closure in regeneration and regulation held (autopoiesis), this would provide a crisp demarcation of living systems. But the essence of the evolution of multi-cellular creatures has been the enlarging of their interactive capacities while (modestly) opening up both the closure of their metabolism, e.g. by amino acid imports in humans, and the flexibility of the regulatory closure of regeneration, e.g. by

[59]The close and older cousin of autonomy, autopoiesis, provides a similar demarcation. It will also share the other significances attributed to autonomy below. However because of its closure emphasis, as opposed to the interactive openness of autonomy (see below), the two will part company, somewhat over multi-cellular organisms, but especially over social organisations where autopoiesis finds much less application. Also, the formal closure of the planet could make it appear lifelike under autopoiesis, whereas its self-regulatory capacity seems too weak. I consider this another reason to favour autonomy over autopoiesis.

[60]A simple but useful comparison is that between a candle flame and a cell [Maynard-Smith, 1986]. A candle flame pre-heats its wax, permitting it to rise into the wick, and creates convection air currents delivering fresh oxygen, thereby supporting the burning process (flame). So it is, as Bickhard [1993] says, partially self-maintaining and (passively) recursively self-maintaining in these respects (that is, the same processes that maintain the flame thereby maintain the capacity to maintain the flame). In doing this a candle flame creates and sustains a thermodynamic asymmetry between itself and its environment, based in a (minimal) organisational asymmetry. This is essential for its continued existence as a far-from-equilibrium [ffe] thermodynamic system. The stationary cell does all this, but it also has two further organisational properties, (i) reproductive closure — it succeeds in regenerating all aspects of itself (metabolism), including its own metabolic organisation to do this, and (ii) it self- regulates that organisation to aid its continued viability, that is, it responds to distortions in its functioning with corrective measures, e.g. altering a chemical balance or gradient by ionic intake or expulsion, altering environment through tumbling. We might say that it is actively self-maintaining of its self-maintaining organisation (as opposed to the candle's passivity in this respect). It is these properties, not mastery of the environment per se, that mark off the autonomous systems. They create a distinctive organisational asymmetry between the cell and its environment that results in the locus of cellular regulation lying within the cell and not in its environment. But it is its interactive relations with the environment, essential to its ffe nature, that are the focus of its evolutionary development.

distorting metabolism to support interaction (extreme: burning internal organs to support marathon running). In the right circumstances, openness of these kinds contributes to fitness. (And note that we take in the amino acids despite having a regulatory membrane — skin; but a multi-cellular skin is more open that a cell membrane.) Indeed, interaction capacity in the form of medical science can now intervene throughout metabolism, export manufacture of body parts (from knees to hearts) to industry, and so on.[61] But this removes the crispness to the autonomy boundary. Every export of manufacture to the environment represents a loss of regulatory organisation to the system, at least of regeneration, but of organisation as well in cases like heart pacemakers. How much openness of cyclic processes can be tolerated before autonomy fails, before so much manufacture, and/or so much regulation of metabolic organisation, is exported to the environment that the locus of organisational regulation shifts from the organism to its environment?[62]

These are typical of boundary problems for open complex adaptive systems. They make autonomy analysis of social entities like firms, cities and nations particularly complex because the boundaries themselves have been rendered still more diffuse in favour of increased regulation of flow organisation. Future robotics, now that embodied design has begun at last to take hold (cf. below, [Barandiaran and Moreno, 2008]), will also provide challenging cases. Meanwhile, consider the following, partially related formulas (I) 'their thermodynamic asymmetry with the environment is explained by regulatory loci that are more within the system than in the system environment' and (II) 'actively regulates its regenerative organisation'. Used to determine the scope of autonomy, either of these formulas, a fortiori their conjunction, will suffice to clearly exclude the candle, hurricane, viruses and the like but clearly include organisms, and leave open the future border cases in robotics that are sure to come. I would prefer to have available more penetrat-

[61]At this point a difference between autopoiesis and autonomy highlights the subtleties at issue. (Autopoiesis is Maturana and Varela's term [1980; cf. Varela, *et al.*, 1974], the pioneers of this analysis, although they have also used 'autonomy' [Varela, 1979]. So in what follows please read 'autonomy' as the dual loop coherence described in section 3.) Maturana and Varela emphasise closure as they key distinguishing feature of autopoiesis; to put it crudely we might say that autopoiesis = complete reproductive closure + complete self-regulatory closure of reproductive closure. And they are impressed by the existence of the cell membrane as the regulatory device for maintaining the locus of regulation within the cell. This conception has the advantage that it is relatively sharply defined. [Rosen, 1991], e.g., even thought that he had shown that no such system had computable models, but [Mossio *et al.*, 2009; Stewart and Mossio, 2007] claims to have a counter-example. Pattee [1995; cf. 1996] thought that closure entailed semantics because it entailed self-'reading', and was the only way to create semantic content, hence concluded that the genome was the origin of meaning content in biology. But the burden of the preceding text is that this orientation is neither accurate nor captures the focus of evolutionary development. See also note 69 and text.

[62]The complexity involved in answering is increased by the fact that we cannot do this analysis organism by organism. First, medical capacity spreads the environmental regulatory capacity among all of us, it makes no sense to ask how much environmental regulatory capacity a particular human distinctively relies on. (This would already be true as soon as humans collectively impacted the spread of plants with the missing 9 amino acids — whether they knew it or not.) Second, it would already deny autonomy to each of our cells, since each relies on specialised cells in other organs to manufacture and deliver many of their required inputs.

ing and crisp characterisations of the relevant distinctions. I can't even be sure I've correctly characterised the problem or done justice to the alternative. I do think that these issues go to the heart of understanding biological (and social) organisation.

A *second* significance is that it introduces organisation as fundamental to life. Finite systems sustaining dynamical equilibria far-from-(static)-equilibrium must do so by irreversibly taking in ordered or low entropy energy and material components from their environment and exporting to it material components carrying dissipated, less ordered or higher entropy energy. These open systems must be organised: by the Morowitz theorem [Morowitz, 1978], they must have at least one, and typically will have many, closed-loop processes running within them. This result applies whether or not the systems are living, as earth's weather and its organisms illustrate. However, comparing living systems to reversible equilibrium, so also inanimate, systems highlights the distinctive character of living interactive organisation:

Comparative System Order			
Property	***System Kind***		
	GAS	CRYSTAL	CELL
Internal bonds	None	Rigid, passive	Adaptive, active
Directive ordering*	Very weak, simple	Very strong, simple	Moderate, very complex
Constraints	None	Local	Global
Organisation	None	None	Very high

* Directive ordering is spatio-temporally selective energy flow

Moreover, much or all of living organisation is to support organism functionality and it is thus organisation under these global functionality constraints.

Such global organisational requirements can in fact be met only by complex congeries of mechanisms subtly coordinated in both space and time and possibly emergent.[63] Whence autonomy and its realising organisation pose important challenges to scientific understanding. Science has as yet only weak tools for studying organisation and equally weak tools for studying global constraints, especially spatio-temporally extended global constraints like autonomy. Rosen argued that living systems were not mechanical, that they could not be reduced to congeries of mechanisms of the sorts found in machines ([Rosen, 1985a; 1985b; 1991], cf. note

[63]As self-regenerating systems, internally their cyclic processes must contribute to re-creating each other, that is, each process must partially regenerate the material constraints for themselves and/or others to work, requiring a highly organised web of cyclic process-constraint interdependencies — what [Kauffman, 2000] calls work-constraint cycles. Multi-cellular organisms perform the same overall tasks as do uni-cellular ones, only with an expanded range of self-regulatory capacities, for both internal interaction (e.g. the cardio-vascular resource delivery and waste removal system) and external interaction (e.g. neurally regulated sensory and neuro-muscular motor systems, etc.), to match their expanded regenerative requirements. They are models of self-regulation, including of active self-maintenance of their self-maintenance capacities.

61). The gist of his argument is that their functions, being required to include reproducing the system, are both cause and effect of the system (cf. [Coffman, this volume; Moreno, *et al.*, this volume]) and that this precluded their being machines. That is, the holistic, organisational features like autonomy that are central to being alive cannot be captured by analysis into machine mechanisms, which have a function but one provided externally and which does not contribute to the generation of the mechanism components. Rosen argued that these limitations, largely unrecognised and unexamined, represented a powerful limitation on the development of biological science, and ultimately on all other sciences and that meeting them would ultimately require transforming all science, from physics on.[64]

There is, as just seen, some point to Rosen's line of objection. Metabolic regeneration is central, does exhibit autonomous organisation and currently cannot be adequately modelled dynamically. However, it is also not magic, it has a dynamical basis in biochemistry. The problem is at least that the science of spatio-temporally organised biochemical systems is still developing. Since new tools to understand at least complex spatio-temporal organisation of interaction are currently being developed (e.g. using cellular automata models), the current lacunae might best be temporarily recognised as so many methodological challenges rather than apriori demonstrations of the separation of biology from natural science (cf. [Bechtel, 2007]). Even so, constructing workable biochemical models of cells, even very elementary cells, is difficult (note 67), and there is no inherent capacity in any of these tools to represent either organisation per se or globalness of constraints, nor therefore the roles of contingent coordinating enabling constraints, cf. [Hooker-c, this volume; Juarrero 2002; Van Orden *et al.*, this volume]. Accommodating these features might indeed force fundamental changes on the formulation of dynamics, as Rosen supposed, or it may simply be that, as the capacity to model spatio-temporal organisation grows, more and powerful such tools will sufficiently alleviate these problems.[65]

[64]Their final version [Rosen, 1991] is couched in an arcane modelling language, and the ground has shifted somewhat from the earlier focus on systems that have internal predictive/anticipative models of themselves and their environment - see the instructive essay [Pattee, 2007] and for Pattee's version see [Pattee, 1993].

[65]See further [Hooker-c, this volume], note 18 and text. It should also be noted that this issue is quite distinct from that of whether and how autonomy might have evolved. Indeed, Rosen considered that life could persist without evolution but not vice versa. It seems right that autonomous systems may persist without evolving, just more precariously and so less probably, and that models of genetic algorithms do not specifically require autonomous systems. In practice, improbable miracles aside, it is hard to see how the development of complex life forms could have proceeded except through some such mechanism as evolution. Finally, how this bears on the generation of autonomy in systems is a further issue. It may well be that autonomy is realisable in many different ways, may contain components that can interrelate in many different ways and may partially or wholly self-organise under many circumstances (e.g. hydrophillic/hydrophobic lipid membranes evidently self-organise blastula), all of which, and given only that, would make an evolutionary explanation more probable (though not necessarily easy). Alternatively, it may turn out that autonomy imposes very precise requirements on at least some biochemical relationships, in the sense that none of the preceding apply to them, making an evolutionary explanation,

Dually, the challenge posed to practical construction and regulation/ control in biology and robotics is equally deep because, if the account of autonomy (and of autonomy-based cognition) is even roughly correct, it provides a set of organisational requirements for this task that will prove far from simple to meet. For instance, despite using the label 'autonomous agent', there are at present no truly autonomous robots in this organisational sense. Robotics uses a very limited formal notion of autonomy (something like invariant dynamical form) and limited performance criteria (typically confined to a single task) and an equally limited satisfaction method. There has recently emerged an embodied functionality movement within robotics (see e.g. [Nolfi, this volume; Pfeiffer and Bongard, 2007]) where cognitive organisation is strongly shaped by the dynamics of body and environment, in ways that you would expect from an autonomy, interactive perspective. This represents a vast improvement over the computer-in-a-machine approach that had previously dominated. However it is as yet very far from even incorporating evaluative signals representing the body coherence of robots, let alone the complexity required for self-regeneration and the capacity for fluid management of multi-dimensional environmental and internal interaction processes in relation to that.[66] There is an associated need to bring work on self-assembling, self-repairing robots into relation with attempts to develop artificial autonomous systems, where modelling even very elementary cells that are dynamically stable and thermodynamically coherent is proving difficult.[67]

A *third* significance of autonomy is that it provides a naturalistic grounding for agency. Entities with a distinctive wholeness, individuality and perspective in the world, whose activities are willful, anticipative, deliberate, adaptive and normatively self-evaluated, are properly treated as genuine agents. Autonomous systems are inherently all of those things [Hooker and Skewes, 2009].

Self-regulation. Autonomous systems are strongly self-regulated in both their internal and external interaction, making themselves the distinctive primary locus of their regulation. And because the self-regulation is in service of maintaining an internally coherent whole, they have a distinct, individual reference point for their activity that provides them a distinctive perspective on the world.

Normative self-evaluation. Autonomous self-regeneration constitutes the funda-

given only that, less probable (cf. [Behe, 2000]). But (*pace* Behe) to know how unconditionally (or anyway less conditionally) probable or improbable an evolutionary history might be we first have to know a great deal about the natural production and organisation of relevant bio-molecules in the circumstances obtaining at the time, knowledge that is still in its infancy.

[66]Cf. [Christensen and Hooker, 2002; 2004] and notes 50, 51 references. While studies such as that by [Nolfi, this volume] have made progress on the fluid management of environmental interaction, these are still primitive devices when it comes to management of activity in relation to norm-derived goals. The problem in artificial life is still further from solution, since formal reproduction is not regenerative and is not the core of metabolism and thus metabolism-based action norms. See also [Moreno and Ruiz-Mirazo, 1999; Moreno, *et al.*, this volume].

[67]For self-assembling/repairing robots see e.g. [Groß and Dorigo, 2007; 2008] and `http://www.swarmanoid.org/index.php` and for protocell investigation see e.g. [Gánti, 2003; Ruiz-Mirazo and Moreno, 2004; Szathmary, 2005; Barandiaran and Ruiz-Mirazo, 2008], `http://www.ees.lanl.gov/protocells`, and references.

mental basis for normative evaluation because it is the sine qua non and reference point for all else. Autonomy is the condition against which the outcomes of system processes are measured for success or failure. In single cells the measurement is simply continued existence or not. Multicellular systems have developed many internal, partial and indirect surrogate indicators for autonomy satisfaction and its impending violation, often based around closure conditions for their important sub-processes, e.g. hunger (impending violation) and food satiation (satisfaction). It is these specific surrogate signals (cf. also thirst/fluid satiation, pain/pain-freeness) we think of as the basic, primitive norms guiding behaviour, but they are literally grounded in turn in the obtaining of autonomy, from which they derive their normative character.

Wilfulness. A will is the capacity to do work (that is, transform energy) in relation to the self whose will it is. The constitution of the autonomy constraint, which focuses directive organisation on generating behaviour to achieve self-regeneration, constitutes just such a distinctive capacity.

Action. The wilful performance of anticipative interactive activity against a normative evaluation criterion provides a root sense of action.

Adaptedness, Adaptiveness . An organism is adapted when it possesses an autonomy-satisfying set of traits in its life-environment. Conversely, an organism's ecological niche is comprised of the range of life-environments for which its traits provide satisfaction of autonomy. An organism's adaptiveness is its capacity to alter its specific traits in mutually coordinated ways so as to adapt to, that is, satisfy autonomy in, a wider range of life-environments than its current one.

The major challenge facing this approach to individuality is to link it to a detailed account of the globally coherent, spatio-temporally organised realising mechanisms, e.g. neural organisation, in an account of the evolution of organisms from elementary proto-cells to contemporary forms of life (cf. [Gánti, 2003; Bechtel, 2007; Bechtel and Abrahamsen, this volume]). This non-trivial task now constitutes a large research focus. Part of it is to re-think the nature of intentionality and intelligence on the basis of the evolution of neural tissue in both the visceral and central nervous systems (see [Moreno, this volume] and references). Here again the functional organisation of autonomy provides a scaffold for developing more detailed theories [Christensen and Hooker, 2002].

A *fourth* significance of autonomy is that its conception of agency just delineated fruitfully frames the evolution of intelligence and intentionality, thus also providing a framework for (organically) intelligent robotics. The underlying theme of the expansion of multi-cellular life forms is the massive expansion of interactive potential and competencies that will support more complex regulatory capacities and vice versa.[68] The increasing self-regulation of interactive capacity grounds their

[68]Multicellular organisms differ in at least three important respects from single cells: they have (i) increased substitution of environmental construction for internal construction (e.g. carnivores intake complex molecules, humans rely on environmental production of many essential amino acids), (ii) increased self-regulation of their food acquisition and damage avoidance (e.g. rapid or prolonged migration to track food resources, hiding or hole construction to escape predators) and (iii) increased capacity to self-regulate the modification of metabolism to suit both tempo-

rich adaptabilities that make them so successful, and so intentional and intelligent in the way they are successful.[69] Becoming more specific, there are three major aspects determining a system's anticipative capacities: the width of its interactive time window, the degree of articulation of the autonomy-related norms which it can use, and the high-order interactive relationships that it can effectively regulate. Between them, these features characterise the dimensions of intelligent/intentional capacity [Christensen and Hooker, 2002; 2000b], and their roughly joint evolution traces the emergence of mind.

In this development two key thresholds are marked by the emergence of (i) self-directedness — the capacity to evaluate and modify interaction with the environment so as to improve basic autonomy-referenced evaluation — and subsequently by (ii) self-directed anticipative learning — the capacity to extend and modify anticipation, evaluation and modification strategy in the light of past interaction so as to better satisfy autonomy-referenced evaluation.[70] Moreover, autonomous systems can also be provided with action-centred informational and semantic characterisations, to complete the sense of agency. Organism information is modelled as reduction in downstream process regulation uncertainty. ('Shall I do A or B? Given the result of my last interaction, B is the thing to do.') Organism semantics is that of the anticipated norm-referenced, autonomy-satisfaction provided by an action. Intentionality is then conceived as a high-order regulatory capacity for

rary activity (e.g. heart rate and blood re-direction for running) and permanent change (e.g. callousing, neuro-muscular compensation for injury). Underlying these is a fourth, more basic, way in which they differ, (iv) they have acquired the capacity to communally regulate the birth, specialisation and death of their members (cells). While in neurally more complex species they can show a myriad of forms, every viable community, including human communities, must acquire some form of these capacities. (Over the last century as human societies have become more developed, they seem to be experimenting with internal regulation of birth (aka the demographic transition), and specialisation (aka education) while decreasing regulation of death (no death penalty, euthanasia, + some regulation of war). Similarly, all the conceptions to follow grade back to the actions of single cells, though the stronger the self-directed, anticipative organisation involved the richer the semantic and informational structures sustained.

[69]The predominant effect of evolution is to expand the capacity for interaction with the environment, including both anticipating environmental courses of action and acting to modify the environment to shape its selection pressures. Here an issue arises concerning the nature and role of closures in biological organisation. The autonomy relation involves the closure of metabolic processes — since the outcome must be reproduction (notes 60, 61 references) — and the closure of the interaction process, since it requires feedback to have effect and to support learning. For Rosen [1985a] this last is the root of anticipation in interaction. There is widespread agreement about these closures, so long as they (the anticipatory loop especially) are construed to be compatible with the expansion of adaptive capacity. Beyond this, however, Pattee [1995; 1996] looks to specific self-referential closure within the cell as the basis of for-the-organism information and [Etxeberria and Moreno, 2001; Ruiz-Mirazo and Moreno, 2006] then extend closure further to that of neural function underlying intelligence, cf. [Pask, 1981]. Again, this might prove correct, but it is unclear that any defensible special notion of information is thus defined and it is noticeable for shifting the focus of attention to the obtaining of closure rather than to the expansion of interactive openeness to the environment which is the functional focus of intelligence and intentionality (cf. [Hooker, 2009b]). Closure is for practical ends, not for its own sake. The issue needs further close consideration.

[70]For some more detail see section 4.1.3 below, where it is also applied to the scientific process.

fluid, meaningful goal-directed management of interaction and intelligence as goal-directed management of problem-solving capacity. Intelligence and intentionality co-evolve making use of a common self-regulatory apparatus. This avoids the common but implausible split between the two, respectively into problem solving and referential capacities. Agent actions are conceived as whole extended processes adaptively aimed at a goal (cf. [Juarrero, 2002; Van Orden *et al.*, this volume]), structured and regulated through anticipative potentiation [Hooker and Skewes, 2009].[71]

In sum, autonomy promises to provide the broad organisational framework from within which a fully naturalised conception of organisms can be developed in terms of the naturalistic inter-twined emergences and mechanistic reductions that reveal their biochemical organisational depth. Of course, from a scientific point of view, the devil lies in providing the details. And, as illustrated, the challenges in doing so run deeper than simply coping with complications.

4.1.2 Cultural dynamics

The second of the ideas is that of cultural dynamics. Culture has largely been omitted from biological consideration, it essentially only appears as a source of specific 'memes' that impact genetic evolution, its only changes being in those memes.[72] Culture is a population-level phenomenon and a general memetic model of culture is popular because it appears to copy the established gene-centred biological theory of population dynamics, is relatively simple and marries simply to genetic dynamics when modelling evolution-culture interactions. But it is also a dynamically simplistic, partial conception of how culture is structured and works. The discussion of *super-system formation*, section 3 made clear that we are deeply embedded in culture, much more deeply embedded than memetic infection can express. The beauty of using complex dynamical models is that doing so opens up a wide variety of other interactive conceptions for consideration — e.g. rather than its being like catching a cold, perhaps our relation to culture is more like

[71] In this limited sense [Hooker and Skewes, 2009] can be understood as employing the notion of autonomy and the ensuing notions of intentionality and intelligence built on it in [Christensen and Hooker, 2002; 2004] to complement and complete the complex systems approach to action begun in [Juarrero, 2002] and extended in [Van Orden *et al.*, this volume]. In other respects there are many differences. Here the resulting interaction-centred semantics is very different from, and more powerful than, standard direct referential semantics, for it captures directly the unlimited implicit possibility content in our action-differentiated grasp on reality. Bickhard argues that in this way it resolves the frame problem, and is anyway ultimately the only coherent naturalist semantics, see e.g. [Bickhard and Terveen, 1995].

[72] For memetics, see e.g. [Aunger, 2000; Blackmore, 1999], cf. `http://en.wikipedia.org/wiki/Memetics`. Extant evolution-culture studies tend to emphasise one side of a dual-mode model (see e.g. [Lalande, *et al.*, 2000]), either focusing on genetic determination of culture — e.g. some sociobiology and evolutionary psychology — or conversely on the impact of cultural practices, e.g. mate selection and tool use, on gene frequencies through their modification of selection pressures. While understanding whatever constraints genetic inheritance places on accessible cultural expression, and conversely, are important, if difficult and controversial, studies, each provides only very partial insight into cultural dynamics itself.

being glued to a shared rubber sheet so that the movements of each refract to the whole and vice versa.[73]

Knowing that genes associated with flight capacity have increased (all that population genetics/memetics supplies) is not the same as providing an explanatory theory *of* flight, which in addition requires at least an aerodynamic account of what flight requires, a physiological account of how it is achieved and exploited, and an ecological account of when it is advantageous and why it can form stable niches. Without all these we cannot explain how and why flight developed (e.g. as opposed to gliding), what limitations it imposes (e.g. on energy budgets, hence food sources), what its range of expression is (cf. albatross versus humming birds), what are its characteristic failure modes (e.g. inappropriate pectoral blood supply for wind regime), etc. And without these understandings we cannot understand how its embodied expression might change. That is, we cannot formulate an adequate dynamics for it.

Further, explanation deeply involves integrated holistic processes that resist modelling as simple bundles of separate units, genetic or otherwise.[74] Like respiration (see section 3, *Multi-level* and *multi-order functional organisation*), culture is actually constituted by a widely diffused but socially integrative, multi-dimensional, multi-modal, multi-plexed, multi-producted, and multi-phasic complex of highly interactive and plastic, highly organised processes and concomitant states. Consider clothing as a typical human cultural feature: (i) Clothing serves many functions simultaneously (multi-plexed): body temperature control; injury protection; personal comfort; social role/status indication; aesthetic expression; individuality/deviance creation; ... In consequence it is also involved in modulating many different interactions simultaneously, e.g. interaction with the physical surroundings and in various aspects of social interaction. (ii) The realisation of

[73]To illustrate the richness available, here are some simple proto-cultural dynamical distinctions. (SHMO = simple harmonic oscillator; DCC = dynamically collective constraint.) Model 1: a set of independent SHMOs. System state = aggregate of individual states. No DCCs. All collective phenomena are patterns determined only by initial and constraint conditions. Social example: the distribution of objects in refuse. Model 2: model 1 + small, local pair-wise interactions between SHMOs. System state = perturbation of model 1 state by addition of local pair-wise corrections. Weak local DCCs responsible for collective wave-like perturbation propagation. For increased interaction strength &/or less local interaction, stronger &/or more global DCCs emerge generating further collective phenomena, e.g. entrainment, chaotic behaviour. Social example: pair-wise reflex interaction behaviour. Model 3: model 2 + interactions modified by SHMO integrative memory. System state = joint product of SHMO states and interaction states. Memory is some function of past interactions and constrains current interaction form and strength. Emergence of global DCCs constraining SHMO behaviour in relation to collective properties. Social example: pre-recording socially referenced behaviours. Model 4: model 3 + integrative memory referenced to a shared global field. System state = joint product of SHMO states, interaction states, and field state. Field interacts locally with all SHMOs (realised, e.g., by a rubber sheet to which they are attached or an electromagnetic field which their movements collectively generate). Emergence of strong global DCCs constraining SHMO behaviour in relation to collective properties based on inherent field dynamics. Social example: socially recorded referenced behaviours.

[74]See e.g. [Ahouse, 1998; Ahouse and Berwick, 1998; Christensen and Hooker, 1998; 1999; Depew and Weber, 1999; Griffiths, 1992; Jablonka and Lamb, 2005; Miklos, 1993; Raff, 1996].

clothing functionality is multi-order and multi-level, requiring people with appropriate internal attitudes, preferences and behaviours; influencing performative aspects of every social activity (e.g. performing authoritatively in a business suit); and involving all of the processes that make up the fabrication, fashion, fabric materials production (including primary production) and recycling, industries. Many different interactive processes thus combine in complex ways to constitute an act or tradition of clothing (multi-dimensional), from production, to performing with and evaluating, to recycling. (iii) There is a large variety of clothing products and product attributes (multi-producted), from swimsuits to business suits to space suits. (iv) Clothing is involved in many different biological and social modes (multi-modal): differentiating work and leisure, dangerous work from safe work (and many forms of work, e.g. priest from pilot), and so on. It is also involved in many industrial production and manufacturing modes, from agriculture to petrochemicals, many different distributional modes, from commercial catwalk fashion to charity, etc. (v) This complex of interactive processes persists on many different timescales (asynchronous), from multi-generations for the overall structure of production + wearing/ performing/ evaluating + recycling, to the sub-generational ephemera of fashion attitudes and products. As technology, work role and lifestyle requirements have changed various aspects of this complex have radically changed organisation (non-stationary).

And of course we are not simply absorbers of culture but equally act to change it, and on many different scales from home behaviours to the conduct of global institutions. A cultural feature is the joint complex product of many groups acting for many different reasons, while also partially shaping them all. Thus we enjoy a delicate shaped/shaping dynamical relationship to culture, re-making it while it remakes us. Social constructability is necessary for possessing a culturally suitable plasticity, else the global constraining character of culture would lock its members out of shaping it and lock it in, emasculating genuine cultural participation. Constructability of itself does not however ensure suitably shapable plasticity; termite society is constructed by termites but is not culturally plastic. Culture requires a particular relationship between the responsive, constructive capacities of individuals and the globally binding capacity of the emergent society; too little binding and the society falls apart into a mere aggregate, too much binding and culture is squeezed out by merely rigid habits. But our shaping by culture is significantly plastic still; e.g. 'stone age' people can be taught to fly jet planes and even to adopt the jet-set social world. And our cultures are equally plastic and locally responsive to us and our environment (cf. clothing above). In such cultures powerful adaptive individual-group dynamics characterise all orders of organisation.

Culture plays an intimate role in our lives, not simply as mental contents but as embedded in our personality structure, general assumptions and orientations, and our methods and supporting skills and throughout our economy. Understanding cultural change involves characterising such features as clothing dynamically, placing the interactive complex of these cultural features in their organismic, communal and ecological settings, and capturing the delicate shaped/ shaping interaction dynamics that makes human cultures so powerfully creative and adaptive. That parts of clothing, such as hats, can be artificially extracted and their changes recorded no more shows the legitimacy of disassembling cultural features into bundles of objects and ideas than the study of hearts does in the case of the evolution/development of respiration or flight. We should not, e.g., rush to evolutionary heritability conclusions about culture just from patterns of sequential repetition with modification; ripples on a shelving beach show these, as will all similar spatially constrained succession, like urban expansion, plus all processes of self-organisational re-assembly, such as rush-hour queues, all processes of path-dependent biased copying, such as housing design, etc.[75] Instead of the relatively external relationship of infection between cultural memes and persons we may contemplate culture as shaping constraints, mutual entrainment couplings, self-organising path-dependent features, and the like. The bad news arising from this is that our modelling of cultural dynamics is as embryonic as is our biological modelling generally in these respects. The good news is that culture reveals a fascinatingly complex dynamical reality for study.

Though capturing cultural dynamics is thus daunting, its biological roots provide clues for at least beginning. The shaped/shaping dynamics underlying culture is as old as life itself. Every living entity is internally organised so as to match boundary behaviour to those interaction modes with its environment that will deliver needed resources while avoiding injury. Boundary-mediated, inside is functionally shaped to outside — but in order to preserve inside autonomy invariant, that is, independent of outside. Conversely, outside is altered by, and often shaped by, inside-directed actions, creatures internally regulating this inside/outside tuning in ways their inanimate environment cannot. Beyond any general naturalist leanings, the deep-seatedness of this complex dynamic suggests that we start there when trying to understand cultural change. In this setting culture is modelled as one class of complex shared integrative features of biological communities and it is the dynamics of these that must be captured.[76] But human agents are cognitively and strategically powerful enough to possess self-directed anticipative learning (see below). Beyond the basic task of differentiating and capturing the relevant biological dynamics lies the challenge of understanding the implications of such capacities

[75]Beach slope and water viscosity are factors in the explanation of ripple succession, but very different factors from those generating genuine lineages, let alone from those reproducing generations as combinatorial assemblies. Whence genuine lineage heritability requires Wimsatt's generative entrenchment conditions [Wimsatt, 1999] to *at least* be embedded in his assumed condition of an ongoing reproductive process, and arguably also requires generated autonomy so as to provide a principled sense of within-lineage individuality and of regulatory entrenchment which is genuinely (re)generative. And this is just the beginning to exploring biologically relevant constraints. How much constraint, e.g., can other self-organisation processes apply and a selection process still be claimed operating? Some theorists (the range extends from Wimsatt to Pattee) may still hope to show that there are informational units playing gene-like roles inside every social process, but this remains open and ambitious.

[76]Hooker [2002], on which this discussion is based, offers 5 basic relationships among autonomous systems (creatures) whose dynamics, once captured, might form the basis for a first agent based modelling of cultural interaction.

for human cultural dynamics. Those dynamics are at work everywhere, but they are pre-eminently displayed in science, that supreme cognitive cultural creation.

4.1.3 Science dynamics

The third of the ideas is that of the dynamics of science. Science is part of human culture and you would therefore expect to model its dynamics in similar ways. It is clear, even to casual inspection, that science is thoroughly dynamic, transforming itself from bottom (data) to top (meta-method, metaphysics), moving through major upheavals — some called revolutions, like the classical to relativistic and quantum shifts — and speeding up the process from centuries initially (e.g. the classical to relativistic shift) to decades now (e.g. the genetics to systems/ synthetic biology shift). Moreover, this is not just a matter of grand theory change, there are other, equally or more important changes also involved. Consider, for instance, the evolving role of the senses in science. On the one side there has been an increasingly refined critique of natural sensory perception for its limitations (e.g. limited discriminations), biases (e.g. tracking delays) and imperfections (e.g. illusions). Technology and theory were essential here, e.g. the camera and optics for detecting perspectival bias, infra-red and x-ray technologies and backing theory for the use of the non-visible electromagnetic spectrum. On the other side there is the development of extensions to, and substitutions for, the senses, e.g. telescopes, microscopes, micrometers, x-ray photography. These allow us to confine use of our senses to those narrow circumstances where they work best (e.g. identifying and counting appropriately shaped and coloured human-sized objects). This dynamic has been going on for centuries, interacting with theoretical developments, but as often intiating them and overall uninterrupted by theoretical and other upheavals. Another similar dynamic has been the elaboration of method.[77] The impact of change of method is nicely illustrated by the impact of complex systems models on archaeology [Hooker-b, this volume, section 5.1.3]. These two processes bid fair to be the great 'bedrock' drivers of scientific change, with data and theoretical shifts playing higher order steering, and less frequently fundamental, roles.

Here a biological orientation provides a larger setting. Cognition evolved because adaptive control of behaviour was reinforced. Central to that control is anticipative intervention to alter the body-environmental relationship, not only by moving the body about but also, from early on, to disturb the environment in order to profit by its response and, then, to learn from its response and, subsequently, to deliberately modify the environment so as to reduce risk and enhance

[77]For review and references see e.g. `http://en.wikipedia.org/wiki/History_of_scientific_method`, [Blake, *et al.*, 1960; Oldroyd, 1989]. Many contemporary Companions, Encyclopaedias, etc. of philosophy of science do not consider the topic. In these references method is primarily considered to be concerned with theory-experiment relations, but it ought to be expanded to include experimental design and theory of errors, all critical factors in the conduct of science and hence of scientific learning about learning (e.g. [Mayo, 1996; Farrell and Hooker, 2009] and references). These latter two, plus instrumentation (theory and practice), form the 'engine room' of science.

reward (e.g. nests and burrows, food caches and mating bowers). From this flows the central role of technology in the scientific process. Our earliest technologies were provided by our bodies. Subsequently they evolved as exo-prostheses. Technologies are essentially amplifiers. With respect to science they (i) extend its information base through instrumentation, e.g. electron microscopes, (ii) extend accessible methodologies, e.g. through numerical approximation techniques and automated data processing, (iii) generate new concepts, e.g. computer models of cognition, (iv) extend epistemic processes, e.g. through supporting global communications and (v) provide the resource base for scientific activity, from economic surplus generally to rare earth metals and other specific resources. Conversely, the development of technology is directly impacted by science through (i) new theoretical conceptions, (ii) new theory-driven designs and (iii) performance evaluation and re-design learning (fluid mechanics gained more from the development of reliable aeroplanes than vice versa). This co-evolution of method, theory and technology is of the essence of science and a vivid demonstration of its open-ended dynamicism.

So we need to think of science in these respects as a dynamic system, transforming its own instrumental 'body' as it evolves/develops, in delicate and increasingly intimate interactions with its transformation of its experimental and theoretical practices and its epistemological evaluative processes. And of course in strong positive feedback interaction with its economic and social environment through the supply of applied science and the feedback of funding and supporting institutions. Through technology science transforms its social and natural environment (e.g. respectively motor vehicles, world agricultural gene trade). This is the great Change Engine that is science-technology-economy. Not to take a dynamic process view of science will be to miss all this. And more. Science also transforms the policy processes that contribute to the dynamics of its environment (e.g. development of economic modelling for policy determination). And through the development of various science studies and science policy studies, which together form the so-called science of science, it is also transforming its own institutional design and social relations (e.g. respectively, the trend to larger specialised research groups, governmental science advice processes) including its own social evaluation. This sophisticated and thorough transformative capacity is a crucial part of understanding science as an increasingly autonomous, dynamic cognitive system.

But this perspective is wholly foreign to the traditional epistemological conception, the logic-centred conception, of science. The traditional models of science that consolidated and dominated in the period 1920-70 were all versions of an abstract, formally (logically) structured, a-temporal machine. In its empiricist version it takes empirical data and mathematical truths as input and, using formal inductive logic, outputs a best theory plus a specification of the best next experiment on that theory (or, where appropriate, a set of equal best theories and next experiments). In its falsification (Popperian) version it takes hypotheses, empirical data and mathematical truths as input and, using deductive logic, outputs a pass/fail verdict together with a next severest test for a passing hypothesis and a next hypothesis re-setting for a fail. (But not the next hypothesis to test, being beyond deductive logic it is generated randomly with respect to the preceding.) In each case the rational decision making of individual scientists is modelled by a universal abstract rational agent following its formal procedure [Forster, this volume]. While this 'method machine' is assumed to be a computational process, this is in the timeless sense of logical sequence, not in the dynamical sense. Thus this dominant tradition was inimical to taking a dynamical approach to understanding science.

True, a very weak sense of timing could be had by tagging the sequence of theories generated with successive times, but this was always an abstract normative sequence, never a real historical one and was soon shown inadequate to the richness and messiness of actual history. The epistemic power of science lies in the way it builds, not just new descriptive theories, but new concepts and new methods, including new inference patterns, to suit what it discovers, *and* becomes better at doing all this as it goes along — no eternal formal inference machine can ever truly illuminate that. But the point here is that its formality prevents even recognising such transformative dynamics, which ultimately have to be carried by human creativity. It is diagnostic that any attempt to introduce psycho-social processes to the method machine was dismissed as confusing the merely empirical detail of implementation with the overriding normative task of determining epistemic warrant (or worse, claiming it relevant only when it represented interference with normative process). When Kuhn famously concluded that a scientific revolution could not be understood as such a formal machine process, it was not instead proposed to be some other kind of social learning process, but was instead demoted or dismissed as 'irrational', outside the study of scientific epistemology.[78]

The first important move away from this conception from the process point of view was the growing insistence by historians and some philosophers of science that they deal with individual scientists and laboratories in their social milieu, understanding their arguments and decision making as responses to their respective situations, e.g. [Galison, 1987; Shapin and Schaffer, 1985] However, these changes were conducted during the time in which the traditional model was under attack from all sides, various anthropological, historical, social, cultural, feminist and post-modern approaches jostling for attention. All of these derived from humanist sources, innocent of any dynamical sense even when using vaguely systems terms, and were predominantly understood in the humanist context. Thus the focus of discussion was on the social and symbolic nature and role of cognition, not on the implied dynamical nature of the processes involved.[79]

---

[78]See [Heunighan-Heune, 1993; Kuhn, 1962; Lakatos and Musgrave, 1970] and e.g. the Stanford Encyclopedia of Philosophy at `http://plato.stanford.edu/entries/thomas-kuhn/`.

[79]Latour, for instance, introduces laboratory studies conceived as anthropological field work, considers each individual decision maker present, even including investigated bacteria. See [Latour, 1987; Latour and Woolgar,1979]. But these are treated as social agents and modelled as making decisions in the light of information available, their understanding of the situation and their interests (aka utilities, values). This isn't yet dynamical, because social 'games' were thought of at the time as intentional and strategic, capacities that were tied to the non-dynamical

But from the later, larger perspective of the dynamical modelling of processes throughout the sciences, these moves can be seen as the first important moves toward a more dynamical modelling of scientific process itself. For if one wants to model science as some kind of interactive process, then the first step is to focus on the interactors, the active components of the system. Rather than simply accepting the stable collective features of the whole system (when it is 'running smoothly') as the basic structure for epistemology, as the traditional approach does, it would instead be viewed as an emergent macroscopic feature generated by actor interactions. Epistemological understanding, if any is to be had, needs to be re-grounded in the underlying actors and to characterise the collective level as emerging from their activities. But this radically reverses the natural position of the traditional approach since normative epistemology, being transcendent, holds 'above' even science as a whole and has a one-way relation of obligation on, and thus critical evaluator of, individual efforts. Unhappily, instead of siezing this reversal to open up dynamical process modelling of scientific process, the actors in this meta-science process drove it toward portraying science as simply one social 'game' among many others (the 'strong programme' in sociology of knowledge). Here there is no interest in game dynamics (cf. [Harms, this volume]), only in assessing the games' status. Old-style formal normative epistemology was to be abandoned, or at least suspended.

Set aside that historians and anthropologists are not equipped, or inherently concerned, to treat epistemology. That they could not find the macro epistemology of the traditional model among their agent micro constituents is in retrospect to be expected. It is expected because re-understanding the macro phenomena is a characteristic issue for any dynamical micro account of a stable macro feature, typically involving emergence and functional reduction. That it should engender talk of abandoning epistemology is diagnostic, not only of professional disinterest, but of the lack of this modelling perspective. Nonetheless, the response was understandable in context. Not only was attention distracted from dynamical modelling by a myriad contentious meta-perspectives, there were few models to follow. Even today, after multi-agent modelling is well established in social science, the vastly intricate organisation of science places it as yet well beyond any detailed modelling capacity that might illuminate its distinctive cognitive and epistemic capacities. Compared to other domains, there are only simple, crude models to expose that cannot do justice to the complexities and nuances of real historical and laboratory situations, or even of the abstract network structure of science. Nonetheless, a range of simplified intermediate models have been proposed pro tem for approaching the modelling of scientific activity in complex dynamical terms. These draw their inspiration from complex systems models in related domains and include cybernetic, matter-energy web, evolutionary and emergent multi-agent models.

world of thought and social norms. Yet it nonetheless represented a radical shift away from the traditional approach — so much so that Latour called for a moratorium on further traditional epistemological study of science. Today we might instead read much of Latour as incipient multi-agent dynamical network modelling.

These will be briefly reviewed in that order.

Cybernetic modelling. Beer was a prominent exponent who modelled an organisation as a collection of necessary functional modules that must all specifically dynamically interrelate so as to collectively constitute a viable system.[80] Largely focussed on business organisation, this model was nonetheless intended to be universally applicable and so applies to science as a viable organisation (cf. [Leonard, 1999]). Its functional categories are too general and its modularity too divisive to realistically model real scientific institutions, where functional multi-tasking and multi-plexing is too widespread and interrelationships are idiosyncratically both local and international. Nonetheless, it gives the sense of a dynamic regulatory system and validates the search for dynamical features like modularity and networks, communication, feedback and regulation.

Hooker [1995] represents a sustained attempt to provide a biologically grounded dynamical regulatory model of science inspired by the earlier cybernetic work. The fixed macro architectures of Beer's approach are eschewed in favour of more fluid biologically oriented process architectures, but regulatory functioning remains at the centre. Hooker shows how Rescher [1977], who shifted the focus of change from theory to method (in partial defiance of the eternal logic model), provides a valuable key to a still more thoroughly regulatory conception of science. Here scientific development is modelled as a multi-layered regulatory process spanning meta-method (including mathematics), meta/proto-physics, method, theory, experiment and data layers, all mutually interactive, each storing sophisticated regulatory information within and between them. Piaget is revealed as a pioneering interactive, regulatory biological theorist (see also [Hooker, 1992]). (His later 'stages' model of development is then viewed as a partial, crude aberrant return to a non-dynamical, logic-based framework.) From this perspective Piaget can then be understood as developing a conception of reason as a regulatory process that generalises well to a dynamical regulatory setting. And Hodges and Hooker show how even Popper had begun construction of a conception of social 'soft' regulation that provides a natural generalisation of engineering regulation suited to science as a social process (and quite at odds with his better known evolutionary views — see below).[81]

Matter-energy web modelling. Subsequently ecology developed a related interest in matter-energy flow food-web models (e.g. [Odum, 1971]). [Gao, 2005; Gao and Herfel, this volume]), for example, describe later work by Pahl-Wostl that constructs dynamical energy-material-information flow models forming spatio-temporally organised trophic dynamic modules in ecological networks. Pahl-Wostl argues that individual organisms' activities are the driving forces responsible for the emergence of this ecological organisation. Gao shows how this applies to science as an analo-

[80]See e.g. [Beer, 1972; 1979; Espejo and Harndon, 1989] and `http://en.wikipedia.org/wiki/Viable_System_Model`.

[81]Hodges [1997] develops in part an analysis of philosophies of science across the late twentieth century as essentially providing increasingly elaborated feedback quality controls on scientific development. For the discussions of Rescher's method dynamics, Piaget's regulatory rationality and Popper on soft control see chapters 4, 5 and 3 respectively of [Hooker, 1995].

gous networked system and provides insight into the laboratory, disciplinary and communicational organisation of science.

Evolutionary modelling. A rather different approach is represented by evolutionary models of scientific knowledge, known as evolutionary epistemologies. These are explicitly focused on the cognitive/epistemic content of science, unlike the preceding models, but share with the preceding models a first, rudimentary sense of a dynamics of science.[82] The simplest and most common of these models postulates a minimal, selective or partial analogy between genetic and knowledge processes, namely a common process of variation, selection and retention [VSR]. The idea is essentially that of the memetics model discussed under 4.1.2 above: cognitive memes evolve in direct analogy to the process for genes, with V = proposed ideas (hypotheses, etc.), S = experiment, R = accepted knowledge, so that VSR is the basic process of trial-and-error learning. The point of employing only a selective analogy is to avoid, e.g., having to find cognitive analogs for either the genotype/ phenotype distinction or for the failure of transmission of information from cognitive phenotype to genotype. Many prominent evolutionary epistemologists occupy this cautious position[83] and it has interesting applications, e.g. to steam locomotion [Cragg, 1989]. Nonetheless, it is dynamically the crudest model kind with all of the defects noted under the discussion of culture above. In particular, these are disembodied genes/memes without embedding phenotypes, communities and ecologies to shape their dynamics. But, as just noted above, we must expect the embodied capacities of scientists and their communal institutional organisation to play important roles in science dynamics. In fact, expanding phenotypic capacities results in increasing internal regulation of VSR processes,[84] a crucial development unavailable to these models.

An obvious augmentation of the basic model is to introduce a genotype/phenotype structure to cognitive processes so as to extend the analogy. Toulmin [1972], e.g. takes the units of variation to be individual concepts, the genotype a relevant constellation of ideas and the phenotype roughly a particular theory and perhaps its applications. For Hahlweg [1983] the genotype corresponds to the entire realm of language and the phenotypes are actual scientists with their epistemic commitments and practices. There are many other versions, including from evolutionary economics (e.g. [Nelson and Winter, 1982]), adding to the foregoing array such analogical genotypes as mnemotypes, rules and routines, and such analogical phenotypes as artifacts and institutional actions. Since we now have parallel, but apparently unconnected, processes running on very different biological and cogni-

[82] For overview and analysis see [Hahlweg and Hooker, 1989a; cf. Hooker, 1995, 1.2].

[83] See e.g. [Campbell, 1974; 1977; 1986; 1997; Campbell and Paller, 1989; Rescher, 1977; Popper 1979; 1984; 1987]. The crude VSR process model suits Popper because he regards hypothesis creation as effectively random. Campbell has the merit of recognising multi-layered organisation in both processes and that these are efficient ways to store sophisticated regulatory information (cf. [Harms, this volume; Hogeweg, 2002b; Hogeweg and Takeuchi, 2003]). However the basic crudities still overwhelm the position [Christensen and Hooker, 1999].

[84] Cf. [Harms, this volume] on phenotypic capacities and on this internalisation of VSR process see [Christensen and Hooker, 1998; Bickhard, 2001; 2002; Hooker, 2009b].

tive substrates[85], and with none of the predecessor defects addressed, these models need specific evidential justification if they are to be adopted.

An importantly different extension of the basic analogy adds phenotype capacities to both biological and cognitive evolutionary processes. This recognises that selection acts on the phenotype, not on the genotype directly. Conversely, variations cannot become eligible for environmental selection unless they have first led to a viable embryogenesis producing a viable phenotype. Hahlweg later developed a theory of this kind (see [Hahlweg and Hooker, 1989a, part II; Hahlweg, 1991]). Notable about this conception is the distinction between adaptation and adaptability (adaptive capacity), together with the thesis that in certain circumstances adaptive capacity can accumulate under VSR while adapted capacities remain non-accumulative. In itself these are welcome sophistications of each process that begin to relieve the charge of dynamical crudity, as well as speak to the huge accumulation of adaptive capacity represented by the history of science. However, it also draws attention to the concomitant necessity of introducing the material communities and ecologies that phenotypes use and reconstruct to modify selection pressures, the one concrete the other abstract, and in doing so re-emphasises the original awkward dual substrates parallelism left untouched.

It makes sense to solve the parallelism problem by doing away with two processes in favour of one. That is to view knowledge as simply one aspect of biological life that emerges as phenotypic cognitively-based capacities increase, and view science as its concrete institutional communal/species form. Evolution is also taken to be a multi-layered development of regulatory systems, from intra-cellular to ecological (local, regional and planetary). From this point of view science represents an extension of regulatory complexity driven by information expansion, intensification and compression (into laws). Mathematics vastly extends our theoretical regulatory capacity while technology vastly extends its material reach, together enlarging our possibilities for further regulatory expansion. This unified evolutionary orientation is essentially that used in Hooker [1995], reviewed above, so long as it be remembered that expanding phenotypic capacities results in increasing internal regulation of VSR processes, so it is that regulation that takes centre stage.

Emergent multi-agent modelling. One way to develop the general regulatory-evolutionary approach above further is to exploit the idea that the epistemic process of science emerges from the material process, analogously to the emergence of multi-agent dynamics or market pricing. An embryonic version of the former alternative is offered by Herfel and Hooker [1996; 1999]. They explore the general (abstract) productive dynamics of consensus and dissensus in a multi-layered regulatory system (each of meta-method, method, theory, data used to generate and

[85] Popper [1979], e.g., makes the gulf literal, placing the biological entities in his worlds 1 and 2 (respectively, material and psychological) and the intellectual entities in his world 3 (abstract). He connects world 3 to worlds 1 and 2 by a 'Principle of Transference' which simply baldly asserts that what is true logically in world 3 is true causally in worlds 1 and 2. But this *labels* the problem rather than solves it (cf. [Hooker, 1981a; 1995, chapter 3]).

critique adjacent layers), and the prospect of modelling scientific revolutions as analogous to phase transitions within the system. Analogously to Hahlweg's accumulatory adaptiveness, but now concrete, what is progressive over the longer term is the superfoliation of the regulatory structure itself. Shi [2001] provides a more direct model of the self-organised emergence of institutional rules and cognitive regulation in science from a multiplicity of partially locally coordinated interactions among diverse strategic researching agents. This is done within a broadly economic approach to interaction. The emergence of price in the market is the model for self-organised emergence of social regulation. Here individual scientists are seen to play several entrepreneurial and consumer roles (e.g. investing in experiments, consuming each other's information). Between them these agents generate and stabilise institutional functioning. This functioning is summed up in its three kinds of rules: resource distributive (e.g. dollars and devices), cognitive constitutional (e.g. reliability and reproducibility) and evaluative aggregative (e.g. propriety and prestige). This lays the foundations for more detailed multi-agent modelling in future. As these rapidly developing modelling techniques mature (themselves the product of science), a key issue will be whether earlier features are also to be understood as emergent within a Shi-style approach, e.g., that revolutions will be incorporated as self-organised constitutional change and the earlier energy-materials-information flow work will be understood as a product of resource distribution.

Another, complementary, way to develop the position is to consider the methodological organisation of research activity, as founded in individual capacities. Recall that autonomy (see section 4.1.1 above) is a global organisation of sensory and motor capacities linked to satisfying metabolic requirements, and that these capacities have become more sophisticated over time, especially in the mammalian evolutionary lineage, currently culminating with humans. Two key organisational steps toward human cognition are the functional capacities of self-directedness and self-directed anticipative learning [SDAL]. These capacities are characterised by an increasingly sophisticated regulation of success and error feedback that effectively constitutes an internally regulated VSR process.[86] In particular, SDAL gives rise naturally to a methodology for solving open problems, the basic kind of relatively ill-defined problem faced in research, and in fact throughout life. (However such problems are notoriously inaccessible to logical method and to formal cognitive science more generally.) SDAL provides the basic model for the organisation of research activity. Science can then be understood as a complex of partially and multiply interrelated SDAL cyclings, driven by individuals operating in the institutional settings of laboratories, universities, societies, etc. This recent view has already garnered some instructive empirical support, but also poses special challenges to dynamical modelling (see below). It is now briefly presented.

Primitive creatures have stereotypical responses but more sophisticated creatures can shape their behaviour to suit their circumstances. The basic dimensions

[86]On SDAL see note 51, on internal VSR regulation see [Hooker, 2009b], cf. Bickhard's alternative conception of error feedback at note 84.

to this shaping capacity are the capacities to (i) dynamically anticipate the interaction process, (ii) evaluate interaction using normative signals and (iii) modify interaction in the light of (i) and (ii) to satisfy metabolic requirements. Organisms with this three-factor shaping capacity are *self-directed*.[87] Mosquitos hunting blood hosts are not self-directed, their behaviour is stereotypical. By contrast in cheetahs hunting prey we see the integration of evaluative and situational information with powerful anticipation to produce fluid goal-directed hunting interaction. Because cheetahs differentiate many kinds of variables (e.g. concealment, terrain, prey alertness, speed, agility, and aggressiveness) and anticipate their influence on hunting (e.g. on isolating a prey from its herd) they are able to act fluidly and appropriately in a complex and changing hunt. Cheetahs are powerfully self-directed hunters. Successful adaptive shaping is problem solving and these are the ingredients from which cognition is formed.[88]

Cheetahs are also powerful learners; a cheetah cub lacks most of the skills required for hunting and must acquire them through learning in practice. But powerful learning is only possible because the cheetah has the capacity to use evaluation of feedback to modify its own self-directedness, that is, modify its own anticipations, its evaluative signals and its existing behaviour modification procedures. Bumble bees are simple self-directed agents, since they can learn which flowers currently offer the best nectar rewards by evaluating the results of their flower searches, but they cannot learn to anticipate flowering patterns (e.g. by species, season and location), or modify or extend their operative norms, or modify their learning strategy. Cheetah cubs can do all of these things. They are *self-directed anticipative learners* [SDAL].[89] There is a virtuous feedback circle here, for as the cheetah gets better at differentiating the relevant factors in effective hunting it not only becomes better at hunting, it also becomes better able to recognise sources of error in its hunting technique and hence improve it. Cheetahs

[87]They are directed because they are thereby powerfully organised to direct interaction into autonomy satisfaction, and self-directed because these are internal processes and their locus of regulation lies primarily within the organism itself.

[88]There is thus no single 'mark of the mental', instead there is a group of capacities that become specialised in various ways through evolution over the long run and, as regulatory sophistication increases, also increasingly during individual development (using individually and communally regulated VSR processes). See [Christensen and Hooker, 2002].

[89]See [Christensen and Hooker, 2000b; 2002]. Introduced by Christensen [1999], the particular synergism of this converging cycle of improvement was a refinement of its preceding convergence models, see [Hooker, 1988; 1995, chapter 4]. Another more sophisticated example is that of a detective conducting a murder investigation. The detective uses clues from the murder scene to build a profile of the suspect and then uses this profile to further refine the direction and methods of the investigation. (A receipt may be enough to re-direct the search from lovers, using personal network construction methods, to business creditors, using financial analysis.) The profile tells the detective what the murderer is like and what types of clues to look for. This in turn sets new intermediate goals that focus the investigation. If the profile is at least partially accurate the modified investigation should uncover further evidence that in turn further refines the search process, ultimately (hopefully) culminating in capture of the murderer, and revealing how and why the crime happened. It is the interplay between the discovery of clues, the construction of a suspect profile and subsequent modification of the anticipative investigation strategy that provides the process its self-directing power.

learn how to learn better in the process of learning how to do better. SDAL uses interaction to acquire information about the nature of the task as well as of the specific performance to hand, and thereby improves performance.

Just this is the virtuous bootstrap that must be behind every capacity to solve open problems. When successful, SDAL results in a pushme-pullyou effect as learning is pushed forward by the construction of new anticipations and pulled forward by the environmental feedback generated, creating an unfolding self-directing learning sequence. Because of its characteristic self-improvement SDAL can begin with poor quality information, vague hypotheses, tentative methods and without specific success criteria, and conjointly refine these as the process proceeds. This makes SDAL powerful because it allows successful learning to arise in both rich and sparse cognitive conditions. These are the conditions met in scientific research. Uncertain about valid methods and relevant data, let alone true theories, researchers behave like cheetahs (and detectives — note 89). (I) They search for features that might provide useful clues and evaluate their significance against previously accepted criteria. (II) They then build a tentative theory-linked model of the situation based on the available evidence and accepted theoretical and practical options. (III) In that light they decide a tentative methodology for investigation. (V) They then rely on evaluating feedback from experiment to refine the model and investigative methods. If this continues successfully they will close in upon the real nature of the problem and its proper categories, theory and methods. If improvements stall, they return to (II), re-evaluate the now larger body of evidence, develop a new model and continue the cycle.

The SDAL model has been tested against a new analysis of the early history of ape language research and shown to provide deeper illumination of its processes than does conventional analysis, indeed to lead to an improved account of error management in science [Farrell and Hooker, 2007a; 2007b; 2009]. In this case the SDAL synergy between (i) learning improved interaction (experimentation) methods and (ii) learning how to learn about investigating intelligent, social animals, proved particularly powerful. It led to the transformation of the 'profile' of the nature and role of language in a social community in a way that opened up a radically new methodological space. This is the essence of SDAL learning.[90] Moreover, SDAL cycles can for this reason operate across so-called revolutionary periods of change as well as across the less transformative 'puzzle-solving' changes. This demonstrates a deeper SDAL-cycling organisation to science, reveals a rational sub-structure to scientific revolutions, and explains the way in which the learning capacity of science improves over time in tandem with the adequacy of its content.

These are early stages in developing an adequate dynamical model for science. An obvious next step is to attempt to combine the emergence model of dynamically stabilised macro organisation with the SDAL model of research process organisation and dynamics. Although dynamical in spirit, the realisation of any more

[90] It also focuses attention on methodological change as a, often the, principal driver of scientific change (as opposed to the primary emphasis on theory change in the standard approach). And the result is again a thorough individually and communally regulated scientific VSR process.

quantitative dynamics must likely await the development of multi-agent network models for these features. This will be no simple task. First, constructing a network dynamics will be challenging since these networks are ephemeral and highly idiosyncratic to content. For instance, at one point as research on ape language developed it became relevant to pay attention to the distinction between receptive and expressive competencies in the context of developmental linguistics and at another to expertise in non-linguistic behavioural signalling [Farrell and Hooker, 2007a; 2007b], extending the research network in these highly content-specialised ways for that time. In sum, scientists form relationships based on any and all of the specific methods they use, theoretical models they employ, problems they have, accessible useful data bases, and so on. Some of these will be long-term, but many will be ephemeral and opportunistic. They may involve specialists in other areas of their own discipline or domain or in 'distant' disciplines and domains (because, e.g., they share the same tools or related problems). And they can change in the same highly idiosyncratic ways. A single visit to another laboratory to learn a new technique can transform an entire research programme and its network. These networks will form no simple patterns, and if e.g. 'small world' phenomena indicate that a scientist talks a lot to just a few others, this will in itself be almost wholly uninformative about the importance and dynamics of the resulting science.

There is also the issue of representing the process of creativity within SDAL cycling. Any attempt to model it potentially reduces it to a mechanical process, something that has proven thoroughly elusive to date. It seems that it would do no better justice to it if it were confined to self-organisational processes. Finally, in the eternal formal method/AI machine conception, normativity derives from some realm radically disjoint from (and transcending) factual considerations. With that approach rejected, there arises the inherited issue of how to extract epistemic normativity from prima facie non-normative, because dynamical, processes. But meeting or failing autonomy requirements provides an inherent biological source of normativity, since this condition is a *sine qua non* for all else. This, elaborated through Piagetian developmental process and communal self-organisation of rules, can eventually support the elaboration of ideals like knowledge and truth and their operational surrogates like testability and confirmation that we in fact see in the historical record. Moreover, it renders them fallibly learnable and improvable, to allow modelling, and ultimately explanation, of that history. All together, this approach develops an appreciation of the open-ended complex knowledge generating system that is, in all its messy but wonderfully complex reality, science.[91]

[91]Bradbury says, science is itself a complex system because only a complex system can investigate a complex system [`http://www.tjurunga.com/biography/roger.html`]. This may well be so, but it is nontrivial to clearly show. It is tempting to appeal to some version of Ashby's Law of Requisite Variety (see e.g. `http://pespmc1.vub.ac.be/REQVAR.HTML`), but this has to be done carefully. For instance, since it is possible to control very complex systems with very simple (but rapid) curve-tracing procedures requiring few distinctions to deploy, if the goal is simple enough the variety that is requisite can be small. The amoeba controls its behaviour on the basis of just one distinction: membrane irritation or not, and this suffices for survival in very complex envi-

4.2 Six scientific domains where complex systems are challenging and enriching science

This book would have been more complete if, among the large number of further essays that could legitimately have been added, six in particular had been included covering complex systems in the domains of physics, chemistry, geology, engineering, neurophysiology/neuro-psychology and social science. These essays were diligently pursued but nonetheless did not eventuate. In what follows I provide very brief notes on these domains indicating at least some reasons why they ought to be included. I do not pretend to offer an integrated or penetrating overview — that would be to presume to an expertise I don't have. And, given that some of the literatures are already vast and some hidden away in 'corners', it would also be to presume to exploration time I don't have. These notes are simply designed to set out some issues that are distinctive to the domain and otherwise allow an interested reader to start burrowing into the literature for themselves.

4.2.1 Physics

Manifestations of complexity are surprisingly rare in traditional physics. Typically, it is possible to study the whole of mechanics — classical, relativistic and quantal — without meeting a single complex system. The primary reason for this is surely the clash between the presuppositions of the analytic core dynamics these approaches all share and the conditions under which complexity phenomena typically occur (see *Constraints*, section 3 and [Hooker-c, this volume]). Whence the primary topic for complex systems is to understand this issue and what might be required to resolve it. This is an extremely difficult matter on which few are able to offer systematic insight, even among those with relevant mathematical expertise. In the classical domain Newton's equations apply to all systems, so we can easily define functional solutions for them: the functions for a system S, no matter how complex, are whatever solve the S equations of motion. The trouble is, for almost all systems S, no one knows how to extract much information about S from these 'easy functions', especially for complex S. In short, we have no general the-

ronments. But if we take the goal of science to be maximal knowledge and maximal control, then we might appeal to an expanded form of the law along the lines: a scientific knower must have at least as much internal organisation as the system it wishes to know. Other applications include models whose behaviour can only be simulated not deduced, models so complex they can only be specified or constructed in the same time periods as they took to evolve, and models so complex they can only be investigated similarly to the investigation of reality. Under the rise of complex systems models, each of these cases is becoming increasingly frequent. Such phenomena are at least an antidote to the long diet of culturally inspired models of science. Jacob Broinowski, defending his approach to science, offers as a principle justification that "... language is a more telling, and a better model for science than is any mechanism." [Broinowski, 1965, p.176] With respect to the basis in language that Broinowski had in mind, it can now be understood how profoundly mistaken is this well intentioned pronouncement (though issues over mechanism persist). However, in keeping with the counter-intuitive character of the complex systems domain, note that in the light of [Lansing, this volume], should network dynamics develop into a mature pan-domain science including linguistics, it might yet still be true.

ory of classical non-linearity. The problem is only deepened when we contemplate the determined linearisation of the quantum operator dynamics confronting the equally determined non-linearisation of dynamical space-time in general relativity.

In practice, mathematical research is split among a number of disparate areas, including non-holonomic Lagrangian theory, effective Hamiltonians, non-linear and irreversible semi-groups, irreversible dissipative systems (initiated by Prigogine's group), irreversible and network thermodynamics, and so on (plus the diverse work on relativistic quantum field theory and quantum gravity). Nor is the bearing of these researches straightforward. One particular focus is on renormalisation (cf. [Batterman, 2002]) to deal with critical point phase transitions. However, the treatment is characteristically indirect, e.g. renormalisation does not directly model the non-linear dynamics involved but rather offers an elaborate strategy, based on self-similarity achieved in an idealised limit, to side-step doing so while still arriving at the requisite end states (cf. [Belletti, *et al.*, 2009]). Within complex systems dynamics a general mathematical characterisation of chaos (strange attractor dynamics) has been developed. This has its own interest because it is related to the self-similarity of renormalisation, to work on irreversible non-linear semigroups as characterising coarse-grained models of chaotic dynamics, and because it presents a rare inherent realisation of randomness (to appropriate sampling) even though the motions themselves are deterministic. Again the interest is primarily intellectual, finitude making literal chaos of dubious practical importance. Near-to-chaos criticality, however, has been given important roles — see, e.g., [Van Orden, *et al.*, this volume], and this 'locational' role may be the most important one, cf. note 39. In sum, it may be the case that there are some general laws to be discovered in this way, and we may be driven in the direction of infinite 'particle' or 'node' limits to extract them by the impenetrability of quantum field theory. But given the diversity of dynamics skated over, whether it points beyond mathematical artifices and artifacts, and precisely how and why, remains to be made clear.[92]

The depth and breadth of mathematics required to cover all this is enormous. An elegant review [Holmes, 2005] of the history of abstract geometrical characterisation of the 'flows' produced by systems of differential equations reveals in passing the thicket of directly related domains. In closing it reveals the large range of more distinct domains not addressed. Though these studies yield a general characterisation of the stabilities and instabilities of flows and in particular their kinds of bifurcations, and we appreciate the diversity of application, much remains to investigate, e.g. which bifurcations are predictable, whether any have continuous dynamical descriptions, and diversity of applications for most. The upshot is that

[92]For instance, Batterman [2002], in work deriving from renormalisation in condensed matter physics, holds that the structural universality in singular asymptotics represents the real basis of law-likeness (even generalised across networks of many kinds, cf. [Barabási, 2002; Poon and Andelman, 2006]). It is however unclear how general that work is, both because of the dynamical differences underpinning the asymptotics it spans (see [Hooker, 2004, section 3]) and because it still does not address global organisation and constraints (cf. [Hooker-c, this volume], note 18 and text).

there is relatively little communication among these various mathematical research domains. It may simply not be practically possible for any one person to be in a position to integrate these domains. And beyond this theoretical work there are more applied studies of individual complex systems dynamics of interest, now expanding very rapidly in virtually every science. Each of these studies involves its own skills and technical apparatus and is both dislocated from the foregoing theoretical work and every bit as internally diverse as it. Thus again it may not be practically possible for any one person to be in a position to integrate it. We are left then with a series of deep challenges to review and address.

4.2.2 Chemistry

An immediate class of cases are those many non-linear irreversible systems like the Belousov-Zhabotinsky system that display emergent spatio-temporal structure, such as oscillations, and auto-catalytic systems that self-organise a transformed chemical constitution.[93] Similarly, biochemical reaction systems with multiple possible energetic outcomes, e.g. multiple reactant products or multiple protein folding options, typically provide multiple path-dependent dynamics and self-organised end states sensitive to initial spatial ordering conditions (cf. Hogeweg [2000a; 2000b], healthy and BCC-damaged brains being one prominent case). Complex molecules already pose challenges to simple ideas of handedness (besides 'shoes' showing simple left/right mirror asymmetries (chiralities), there are non-handed asymmetries or 'potatoes'). They similarly challenge traditional notions of causality (since changes in molecular structure become a function of all participating chemical reactions).[94] As Ulanowicz [2003] remarks, becoming clear about all these matters is essential to understanding the chemical nature of life and its origins (cf. [Gánti, 2003; Bechtel, this volume]).

Beyond this, most chemical reactions are in practice irreversible thermodynamic processes, so currently defying analytical dynamical characterisation [Prigogine, 2003]. Moreover, chemical molecules are emergent dynamical structures, often with condition-dependent laws, e.g. as between various ionic forms of iron compound or forms of molecular symmetry. From this perspective, the whole of chemistry is a first 'special science' of complex systems beyond physics. Computational chemistry, where complex molecular reactions are digitally simulated, is now a cutting edge of chemistry and provides a wide variety of complex systems applications — see e.g. the review [Truhlar, 2008].

[93] See e.g. [Bechtel, this volume; Golbeter, 1996; Gray and Scott 1994] along with Kauffman's soups [Kauffman, 1993], Eigen's hypercycles [Eigen and Schuster, 1979; Eigen and Winkler, 1993] and Hakan's synergetics [Haken, 1983; 1993].

[94] See e.g. and respectively, [King, 2003] and [Ulanowicz, 2003] from a special *Annals of the New York Academy of Science* devoted to foundations and philosophy of chemistry.

4.2.3 Geoscience

Geoscience gained a foundational dynamics of form with the adoption of plate tectonic dynamics and this has subsequently been complemented by a range of process dynamics of the materials involved: magma formation, sedimentation, metamorphosis and like chemo-physical formation and transformation dynamics, together with earthquake, weathering and like physical dynamics (e.g. [Aharanov and Sparks, 1999; Keilis-Borok and Soloviev, 2003]). Complex phenomena abound, e.g. the breakup of super-plates such as Gondwana represent dynamical bifurcations in plate tectonic dynamics, while various rock formation processes such as cracking represent series of phase shifts. As modelling of various process dynamics has grown more sophisticated, complex system dynamics has increasingly come to the fore. Multi-scale modelling has received rather more attention in geoscience than it has in other sciences, for instance, because of the grand spread of interacting process scales typically involved. Similarly, the role of unique events has received rather more attention in geoscience than it has in other sciences, for instance, because of the centrality of the path-dependent dynamics they yield, e.g. events occasioning particular tectonic plate fracturing patterns that determine its subsequent cracking, weathering and sliding properties.

Oreskes provides an examination of complex geoscience modelling, identifying many of the difficulties involved in their use. The general difficulties will be reviewed in section 5.2 below. In the light of her analyses she argues that such models may be confirmed but never validated and urges that they never be used to make long term predictions because these are invariably wrong and thus undermine respect for science and its role in public decision making.[95] This represents a major challenge to geoscience methodology and deserves examination.

In addition, idiosyncratic individuality has always been a problem to geo*science* because of the scientific focus on discovering and applying simple general laws (compression, [Hooker-c, this volume]). Geo-logy, the universal laws of the earth, confronts the fact that the earth is a unique individual. That individual, like all living individuals, has a significantly idiosyncratic history that could have turned out rather differently had various component 'chance' events been different. This poses special methodological problems for geology: how does one investigate such an object, especially when the vast majority of its history is not directly accessible and feasible experiments cannot deal directly with the larger of its scales of operation? This issue has a close analogue in climate science (see [Snyder *et al.*, this volume]) and is analogous to investigating the developmental dynamics of societies and individuals in biology, where the same problems recur. How are these problems resolved and what is the impact on geo-science? Complex systems modelling provides the means to model idiosyncratic individuals through detailed model adjustment. Does this add distinctive methodological power (to the standard methods of multiple time sampling etc.) in geoscience? And, more widely, how has such modelling shifted the balance between evolutionary and complex sys-

[95]See [Oreskes, 2000a; 2000b; 2003; 2007; Oreskes and Fleming, 2000].

tems ordering principles in understanding our planetary history (cf. the equivalent debate in biology, see [Hooker-b, 2009, section 6.2.3] below).

4.2.4 Engineering

Engineering is so thoroughly suffused with non-linear dynamics and complex systems that to describe their characteristic presence is virtually to describe all engineering. From bridges and skyscrapers [civil and mechanical engineering] to electrical circuits [electrical and computer engineering] to pulp and paper mills and waste treatment plants [chemical and environmental engineering], even traffic flow dynamics [Schrekenberger and Wolf, 1998], all manifest complex non-linear dynamics. Even computers are in fact highly non-linear dynamical systems constrained so that their dynamical transitions model computation (sufficiently well). A simple electrical circuit with capacitors or inductances, e.g., is already a non-linear system and relatively simple versions of these form chaotic oscillators that show such features as sensitivity to initial conditions, strange attractors, rhythmic entrainment and phase change. Lingberg has studied the conditions that bring about chaotic dynamics in these circuits, with some instructive results.[96] Similar remarks could be repeated everywhere across engineering. Methods have had to follow suit. All feedback loops are dynamically non-linear, so control relationships are non-linear. But linear control theory, that is, the theory of the control of linear dynamical systems, formed into an elegant analytic theory 40 years ago (section 2 above). Non-linear control, including e.g. control of chaotic systems, however, is still developing its methods, with some general methods but often developing methods case by case, while controlled or guided self-organisation is also being developed.[97]

Recently engineering has noticeably expanded its models and methods towards the life sciences, even as the life sciences have moved toward it. First, particular methods were adopted and generalised from life sciences, e.g. neural network and genetic algorithm models used to model and control engineering systems, especially non-linear ones. Other examples among many are sensor network arrays that are intelligently organised and may mutually self-organise to perform complex functions and nano to micro insect-like molecular machines that perform designated functions, biomedical as well as industrial. Second, engineering methods were adopted by the life sciences, e.g. non-linear dynamical models of physiologies and ecologies and the spread of engineering design and dynamic control ideas and models into systems and synthetic biology. Systems and synthetic biology respectively aim to model cells and multi-cellular interaction as non-linear complex dynamical systems and to use such insights to design new artificial living systems.[98] These twin movements are allied to growing movements toward resiliently adaptable de-

[96] See e.g. [Cenys, 2003; Lindberg, 2004] and `http://server.oersted.dtu.dk/www/el/lindberg/el/public_html`.

[97] See e.g. [Prokopenko, 2008] and the GSO-2008 conference at `http://www.prokopenko.net/gso.html`.

[98] See e.g. [Fu, *et al.*, 2009] and note 18 references.

signs for building lighting, energy use and other features, and similarly for traffic flows, industrial processing and so on. The advance of control engineering into all these areas, especially adaptive control of physiologies (e.g. pacemakers), societies of robots, ecologies and socio-economic systems, including warfare (cf. [Ryan, this volume]) completes the process of entanglement between the two domains.

All of these developments are now in their infancy and are set to gather pace. Together they will eventually radically transform both engineering and the life sciences, ushering us into the long-anticipated Cyborg Era. Here there will arise many important scientific issues, e.g. how to develop improved participatory self-control processes for social groups (human and socially active robots) so that control of their own social systems may develop, including self-organise. (This issue has already arisen within organisational management.) There will also arise foundational/philosophy of science issues concerning the limits of engineering modelling, problems of predictability and anticipatory control, and the like. And beyond these issues, this merging of scientific disciplines and practices will raise (has already raised) deep existential, ethical and social issues that will deserve close study.

4.2.5 Neurophysiology/Neuropsychology/Psychophysiology

Science has been investigating neurons for 120 years, but useful dynamical models for networks of them are less than half that age. At that time we were entering the computer age and cognitive psychology was captured by abstract functionalism that was morphing from late behaviourism into computationalism. Behaviourism is a form of the input-output analysis common in engineering. Input-output analysis will always be applicable[99], but the only internal structure it can attribute to the modelled object is what can be introjected from external correlations among inputs and outputs (called intervening variables in Behaviourism). Behaviourism had the potential to expand into internal (neural) dynamical modelling and control, as engineering modelling has done. Unfortunately, this was blocked by its own positivist ideology ('there is no inside'), and then the shift to internal computer science models emphasised purely logically based architecture that was independent of implementation, cutting off neural modelling. (Computing devices could implement minds, so also could millions of people passing paper bit strings among themselves in the right ways.) Cognition was then normatively determined by its functional logical structure, everything else was simply implemented performance. Neurophysiology and beginning complex system dynamical modelling of neural

[99]Indeed, methodologically there is always a tension between achieving adequate internal modelling and employing a more agnostic and pragmatic approach that advocates use of simple standard models, like Kalman filters, whilst ever they work in practice, = for the input-output form of the original problem. Rather like (possibly exactly like in situ) behaviourism's intervening variables, these models often work as well or better than attempts at more realistic internal dynamical modelling, e.g. to discover patterns in sparse, high dimensional data and where there simply are no dynamical models available. See further the discussion of Breiman's similar argument at section 6.2.5 below that broadens the issue to machine learning generally.

systems progressed steadily during this time, but was held logically irrelevant to understanding cognition itself. (It was potentially useful in understanding practical performance competencies or malfunctions.) For its part, neurophysiologists were swamped with detailed data for which, with some exceptions (e.g. the discovery of visual edge detectors), it was impossible to see through the welter of detail to their cognitive significance.

Philosophy had no trouble following suit in all this since the rational mind as logic machine had long been dominant and was much easier than behaviourism to translate into computational terms. Its purest form was that of Putnam's early Turing machine functionalism,[100] extending from there into a general logical computational functionalism. This remains the dominant view and explains why cognitive philosophers, with rare exceptions (e.g. [Port and van Gelder, 1995]), have paid no attention to dynamical models or complex systems (cf. [Elias-Smith, 2009a]).

That era is now drawing to a close. A first step was the 1980's revival of Rosenblatt's 1950's perceptron in a generalised form that now contained one or more 'hidden' layers between the input and output layers, so-called 'neural nets'.[101] The essence of these non-linear models was their power to extract high-order patterns from complex data when it was difficult or impossible for humans to do so. The relevant point here is that the secret of their success was sub-categorial processing: they process information and store it in their internal ('hidden') layers in terms of non-logical (typically mathematical) functions of the input and output components and, in cognitive cases, in terms of non-logical functions of the input and output categories. In this way their hidden layers can represent information that cannot be expressed as any logical combination of inputs and outputs, and hence the output classification is not arrived at from the input by any process of purely logical analysis.[102] It shows that the underlying processes, invisible to everyday cognition, may be crucial to cognitive capacities and not simply affect the imple-

[100]See e.g. [Putnam, 1975], cf. [Coltheart, 2006a]. Turing machines are the structurally simplest universal computers, that is, in principle able to compute all computable functions. For the sake of intellectual propriety we should note immediately that (i) 'in principle' here is sweepingly abstract, it includes very many computations that could not be carried out in any material universe of our kind (cf. [Cherniak, 1986]) and (ii) mathematics, and so science, has many non-computable functions (see e.g. [PourEl and Richards, 1989; Shipman 1992] and `http://wapedia.mobi/en/Church-Turing_thesis?t=9`). These points appear simultaneously as demonstrations of the power of abstract thought and caveats concerning the reach of computation. To pursue them further, e.g. into undecidable and incomplete formal systems and Gödel-style computational models, is beyond this book's focus.

[101]The original perceptron was a two-layer input-output filter (classifier) adapted through supervised learning (that is, where errors produced on an initial input-output training set are fed back to modify the inter-layer relationships (here linear weights). This proved too simple to learn many classifications, which damped further research until the 'hidden' layer versions were shown to perform much better. See e.g. `http://en.wikipedia.org/wiki/Perceptron` references.

[102]For this reason Smolensky [1988, pp. 3-9] said they were operating in a sub-conceptual, sub-symbolic mode, but not yet the mode of dynamics proper. He labelled this a half-way level between symbolic and dynamic, but this runs together different notions [Hooker, 1995a, note 10]. See e.g. [Rummelhart and McLelland, 1986] and [Hooker, 1995a] for critical review and references.

mented performance of capacities already determined by logical structure. This is a first step away from conceptual computation as the basis of cognition and toward more generalised dynamical models.

A further step in this direction is inherent in the shift from node-focused neural net functions to a focus on dynamical states themselves. It is possible, e.g., to achieve substantial reduction in size for fault-tolerant networks by exploiting dynamical redundancy or degeneracy rather than the traditional node/connection redundancy (which increases network size).[103] To the extent that neuronal systems behaved in this way there would be a gain in economy (less nodes) and in resilience (fault-tolerance), both evolutionarily advantageous features. But understanding them would shift further away from any simple logical structure as foundational and decisively toward dynamically-based discriminations as the basis for agent-integrated functional capacities. This is the kind of step that has been recently taken by the embodied functionality movement within robotics — section 4.1.1 above. However, it remains an open challenge to really bring embodied robotics into a deep relationship to current brain-modelling inspired neuropsychology (cf. also section 4.2.6 below). On the other hand, that step away from symbolic computational models is reinforced by the emphasis of [Van Orden, *et al.*, this volume, secction 3.2] on the self-organised resolution of near-to-chaos criticality as a decision making/solving process and on the 'soft-assembly' (temporary self-organisation) of task-specific neural devices as intentional action preparation.

Although neuron-like only in their interconnectedness, 'neural nets' do form the first non-linear dynamical models to have significant cognitive application. Their advent also generated the first real philosophical revolt from computational functionalism. Churchland [1989; 1995] adopted these models to argue for a more conceptually liberated, more empirically connected philosophy of mind. It is still worth reading to understand the fault and flow lines of the philosophical and scientific shift still underway. However, today we know that neural nets form one class of learning machines, others include classification and regression trees and kernel-based learning machines, including support vector machines. Although applications, sometimes of combinations of these, are expanding rapidly, these methods have yet to be fully integrated with scientific practice and are themselves still developing.[104] These developments have created the rich domain of machine learning and statistical inference. This is a diverse and rapidly expanding domain of artificial learning processes, no longer explicitly tied to brain modelling and new

[103][Penfold, *et al.*, 1993], e.g., reports success in generating fault-tolerant networks for the XOR function (= exclusive 'or', aXORb = either a or b but not a and b) using a genetic algorithm and training selection. XOR has been considered a standard test for non-trivial functional capacity since it was shown that a (1-layer) perceptron could not learn the function. The fault-tolerant networks generated were half the size of the best node redundancy versions. This derived from that fact they were able to cycle through their state space and always settle into an XOR-able attractor after a time, despite the removal of a node and its connections. Cf. [Kohlmorgan, *et al.*, 1999; Zhou and Chen, 2003].

[104]See, e.g., [Hastie, *et al.*, 2008; Tan, *et al.*, 2006; Vapnik, 2001]. Neural nets may prove to have distinctive roles in control, and possibly in nervous system regulation, but this remains to be seen — see [Hooker-b, this volume, section 6.2.5].

for psychology. A key part of the Cyborg Era (4.2.4 above), how does this domain of new techniques relate to nervous system function and to its cognitive capacities and performance?

Today, non-linear dynamical models of the brain abound and there are ever-tightening interrelations between neurophysiology and cognitive psychology, their union encapsulated in 'neuropsychology' and like labels. This revolution has been largely brought about by new technologies for investigating brain function (cf. e.g. [Bressler and Kelso, 2001]), especially patterns of neuron firing and detailed anatomical tracing of neural pathways. To this is now added various forms of neurological imaging, e.g. functional magnetic resonance imaging. The earlier position is analogous to that of cell biology before the contemporary high throughput technologies that now provide the '-omics' data: in each case, though non-linear dynamical models of their respective components have been developed, it is very difficult to construct even roughly realistic complex systems models of their collective functioning.

However, even were that to be accomplished it would still be distant from the high-order functional descriptions that psychology employs. (Would it be analogous to high-order cell and multi-cellular physiology?). Neuronal firings may self-organise into systematic and complex patterns across the brain, but understanding how that might support cognitive or other functions is another matter. Coltheart [2006a; 2006b], for instance, has argued that neuro-imaging has contributed little constructive to cognitive psychology and much that is distracting or misleading (cf. [Elias-Smith, 2009b]). This is essentially because there are mostly no one-one neural-(cognitive)functional correlations to be discovered. The brain evidently employs ubiquitous multi-plexing and multi-tasking, with emergent (perhaps self-organised) constraints on these shifting on several timescales (cf. [Griffiths and Stotz, 2007] for genes). As with even simple machines of these kinds, this means that any high-order functional capacities to which psychology may appeal will be multiply neurally realised in condition-dependent ways. Conversely, the reverse engineering problem of identifying the psychology from the dynamics (much less from current neuroimaging) will prove essentially impenetrable until it is understood what are the dynamical principles on which the multi-plexing and multi-tasking are organised.

Are cognitively significant aspects of brain processes coordinated to larger scale, and/or longer process times, as emergent spiking dynamics might suggest (cf. [Elias-Smith, 2000; Port and van Gelder, 1995])? Is there a point in higher order neural function beyond which functional organisation is completely or mostly determined by the requirements of effective logical problem solving? (I suspect not, on the grounds that [Van Orden, *et al.*, this volume] provide, and because SDAL processes — 4.1.3 above — apply to the most advanced problems.) Such issues, allied to the current diversity of neuropsychological models of brain functional organisation, all claiming to capture the truth, points to the need to resolve the issue of proper methodology for neuropsychology and the place of neuro-imaging and like technologies within it. This is in turn a challenge shared with virtually every other domain of current complex systems investigation (see further [Hooker-b, this volume]).

This latter challenge is itself part of a larger task of insightfully interrelating work on neural networks, more physically oriented brain models and agent based social network models. In all of these classes of non-linear dynamical networks we discover certain phenomena, e.g. emergent collective dynamics and 'small worlds' hub-and-spoke structures. But how these are to be understood in each class may be different. Any network modelling faces the issue of determining to what extent the network dynamics is dominated by node capacities and to what extent it is dominated by connection structure (net topology) and connection strength and operational character. For boolean and adaptive neural nets, e.g., the node capacities are extremely limited (usually just summation or saturating summation) and the connection strengths are either constant (boolean) or simple functions of node values (adaptive neural) and no other operating constraints are applied. Again, many brain neuron net models concern spike timing only, throwing away all internal neural information. But neurons have much more complex non-linear axon dynamics than this (indeed, axon protein is said to possess vast computational power) and synaptic firing involves complex chemical transfers. It would, e.g., be instructive to use modelling tools to explore these interrelationships. (For instance, to generalise a tool like MARVEL that enables staged addition of physical features to abstract network models, see [van Zijderveld, 2007].) And while many social agent based models accord their human agents only simple capacities (e.g. a few independent, first order interaction rules), human agents possess autonomy-based independent evaluation and the capacity to form many highly context-dependent, long range and idiosyncratic relationships that can nonetheless have a powerful effect on the net dynamics. Agent rules simple enough to capture the dominant aspects of insect and bird swarming are unlikely to be able to illuminate any but the grossest aspects of the dynamics of scientific research or large scale project management networks. We have as yet scarcely crossed the threshold of understanding our new complex system modelling tools.

From the philosophical side there is an ongoing debate between those who take a reductionist approach to neuropsychological relations and those who adopt a mechanistic approach with more complex unification.[105] Both sides focus on much the same neurophysiological domains, long term potentiation/conditioning and memory, to provide evidence for their position. Especially if one takes on board the dual interrelations of emergence and reduction [Hooker-c, this volume], it is not clear that there need be any absolute opposition between these approaches (as opposed to the philosophical terms in which they are sometimes stated) — see [Hooker-b, this volume, section 6.2.3]. However Sullivan [2009], after examining in detail the experimental designs and protocols underlying these opposing claims, argues that there is such diversity among the protocols involved across current neuroscience that no conclusions either way can yet be drawn. This is not an unusual position for the investigation of any class of complex systems to find itself

[105]See, e.g., and respectively [Bickle, 2003; 2006; Churchland, 1986] and [Craver, 2002; 2007].

in today and reflects both the inherent methodological complexities involved and the early stage of methodological development. Sullivan concludes that the entire experimental domain requires a more careful and detailed analysis and she offers some new investigative directions for the philosophy of neuroscience. The future challenge is to develop a mature investigative methodology and to resolve how neuropsychology and cognate fields are to be philosophically understood.

These studies hook up in a natural manner with those starting from a more traditional psychological base but with theorising enlivened with complex systems ideas, such as [Smith and Sheya, this volume] and [Van Orden, *et al.*, this volume]. To these excellent introductions to a rich field I wish only to point to complementary work, such as that on emotional development, especially emotional self-regulation (see e.g. [Fogel, *et al.*, 2008; Lewis, 2005; Lewis and Todd, 2007; Lewis and Cook, 2007]). To this I add work applying ideas of developmental self-organisation, as well as less dramatic approaches using dynamical models, to identify specific processes [Heath, 2000; Ward, 2002] and dynamical modelling of individual patients as support for clinical treatment (e.g. [Heiby, *et al.*, 2003]), a new capacity afforded by parametric adjustment of the preceding dynamical models.

4.2.6 Social science

The recent explosive development of synthetic and systems biology is fuelled by the adaptation of complex systems models from chemical and electrical (especially control) engineering. This tactic works because for many purposes intra-cellular components can be treated as simple deterministic systems and the complexity located in their many organised interrelations. The tactic fails when it comes to modelling social systems because the components, persons, are themselves (extremely) complex systems. Economists faced this issue long before complex systems approaches became readily available. Their response was mainly to assume a universally representative ideal rational agent and to use its rationally deterministic behaviour as a surrogate for aggregate behaviour. Later they allowed for a more explicit averaging process, but with random behavioural variation and an emphasis on strong aggregation rules in the limit of large numbers (central limit theorems), to again extract representative aggregate outcomes, in analogy to statistical physics [Auyang, 1998]. The development of Traditional Game Theory relieved the constraints of simple determinism and collective uniformity by introducing a limited variety among players (namely variety in utility of outcomes) and a correlative capacity to discriminate among strategies. This still represents only limited relief, but even this much generates very rich modelling capacities (see e.g. [Harms, this volume]).

The relief is importantly limited because (I) persons have the capacities to construct complex internal models of both (i) their environment and their options within it and (ii) themselves and others in respect of their beliefs, motivations, values and limitations. In consequence, (iii) they take up complex and idiosyncratic strategic relationships (trading, research collaboration, villification, ...) with specific individuals both near and socially distant. Moreover, (II), they operate within a social millieu whose expectations vary from highly specific to vaguely wide and whose structure varies from highly labile to recalcitrantly entrenched (cf. 4.1.2 *Cultural dynamics* above). In consequence, people can interact (a) even when they have incompatible models of the game being played, (b) when they are simultaneously playing strategies within and about a game[106], and (c) when they are part, intentionally or not, of transforming the rules under which the primary game is being played (an important purpose of politics and like social regulation). In a social setting, e.g., often the most intelligent response to a Prisoner's Dilemma game (see [Harms, this volume]) is to seek to change the terms of the game. Modern economic institutions, for instance futures markets, can all be traced to particular historical innovations, they are not timeless features of our inherent rationality or social landscape. On this latter score, the relief offered by Traditional Game Theory is also limited because it represents games and strategies timelessly, their responses to others' moves simply a set of quasi-logical possibilities, whereas in real life timing may be everything.[107] Here the recent shift, more monumental than it may appear on the surface, from Traditional Game Theory to Dynamic Game Theory (let us call it), greatly expanded its capacity to represent strategies as sequentially distributed, with consequent increased representational richness and some surprising, and telling, outcomes (see e.g. [Harms, this volume]).

These developments are complemented by the introduction of ecological mod-

[106]See, e.g., Rapoport's wincingly poignant analysis of Otello and Desdemona as in fact playing different, incompatible games, even while each assumes that the other is also playing their game [Rapoport, 1966]. Here innocent Desdemona's perspective leads her to represent her situation in a game in which Othello is assumed to share her moral values and has as his dominant strategy believing her innocent, to which her best response is to affirm her innocence. (This is best, rather than falsely admitting guilt in the hope of a pardon, a response also of inferior expected utility so long as the probability of Othello's belief in her innocence is greater than .5.) But Othello's suspicion of her (whether based on the vice of paranoia or the virtue of respect for a fellow officer's advice) insures that his representation of their situation is expressed in a game where he and Desdemona do not share the same values and she has a dominant deception strategy to which his tragic best response is to believe her guilty. Here differences in judgement between Othello and Desdemona lead, not just to differences of game detail, such as in outcome utilities, nor even to differently available strategies within a game, but to divergent forms of the game they take themselves to be playing. Whether or not to believe Desdemona as an issue of strategy choice within a game is quite different from, though subtly related to, whether or not to believe her as part of conceptualising the decision problem itself. This situation also underlines the often crucial role of communication in social decision making (contrast Othello's attitude to communication with that of Desdemona, and that of Iargo). Though in Othello's case lack of communication had disastrous results, there are other situations where the inability to communicate improves the rationally accessible outcome, e.g. by removing the capacity to threaten or bluff and the like (e.g. [Rapoport, 1966, pp.126-7]). In sum, here we uncover a recurring theme in social process: rationality cannot consist simply in the formal playing out of specific games but is more fundamentally to do with the management of one's non-formal judgements of oneself and others that underlie formal game form itself.

[107]The 'meta-game', introduced by [Howard, 1971], that considers an extended game of several plays of an original game, can be analytically useful, but is still represented as timeless possibilities. (Note: a meta-game is not a supra-game about playing another game.)

elling (e.g. [Holland, 1995]) and the rise of social robotic agents (see [Nolfi, this volume]). Together they reinforce the contemporary shift toward the less structured and hence less confined agent-based network models now sweeping computational modelling in the social sciences (e.g. [Epstein, 2007; Ferber, 1999; Sawyer, 2005]). Cellular Automata models, employing local neighbour-neighbour interaction rules were among their earliest and best developed kinds (see [Wolfram, 2002]). Typically still focused around local interaction rules, these models follow the development of the interacting community over successive interactions (see also both [Lansing, this volume] and [Green and Leishman, this volume]). There is now a vast and rapidly expanding research literature, studying everything from group dynamics to institutional rule emergence, to macroeconomics. It is supported by international research institutions (societies, conferences, journals), and already with sufficient development to bring the social and robotics modellers together.[108] Here is the context where the self-organised emergence of social institutions is natural rather than unrepresentable (see section 5.3 below), where economics can be re-configured without the restriction to artificially rational, self-interested agents [Foster, this volume] and where the rules of science (distributive, constitutional and aggregative) can emerge from strategic actions by scientists instead of being eternally fixed ([Shi, 2001], see section 4.1.3 above).

Promising as this young field is, there is also a danger that this kind of social modelling is 'too easy'. Almost anything can be 'modelled' by throwing in some simple interaction rules and generating the usual slew of collective pattern features (cf. [McKelvey, 1999]). This runs the risk of undermining its insightfulness. Nonetheless, all the same methodological issues and criteria for sound procedure apply here as they do elsewhere. ([Goldspink, 2002], e.g., provides a nice review.) Increasing domain maturity will only winnow its products and improve its quality. However, as [Harms, this volume] argues, unless and until the self-organisation of social structure can be harnessed in explanatorily adequate ways there will remain a place for the multi-layered models he describes. Certainly, many are currently seeing much promise in a generalised 'network science' (cf. [Barabási and Bonabeau, 2003; Strogatz, 2001; Green, this volume]), an inheritor of the General Systems Theory mantle [Hofkirchner and Schafranek, this volume] that promised a trans-disciplinary systems framework for all sciences. This last has in recent times principally been prompted by the ubiquity of the 'small worlds' phenomenon, where most interconnections run through a few common nodes, greatly reducing node-node interaction distances [Barabási and Bonabeau, 2003]. When nodes are persons or groups of persons, this offers a richer, wider representation of asymmetrical interrelations than traditional tools like power hierarchies, and one with a more clearly underlying dynamical character. Because reducing interaction

---

[108] See generally e.g. [Antunes and Takadama, 2007; Epstein, 2007; Ferber, 1999] and e.g. `http://emil.istc.cnr.it/publications/`. On bringing together social and robotics social simulation see e.g. the First IJCAI Workshop on Social Simulation, `http://ss-ijcai2009.di.fc.ul.pt/`, during the 21st International Joint Conference on Artificial Intelligence, IJCAI-09, `http://ijcai-09.org/`.

distance can be represented as a form of economy or efficiency, the dynamics can be generalised to many other non-social applications.

Yet we should not hasten to suppose that any one simple pattern formation process captures all the complexities of social life. First, we should also recall that, as an exercise of power, those same common nodes can block interrelations as easily as facilitate them, e.g. block or bias communication (censorship, research domination, ...). To this we should add, second, that small world dynamics is by no means universally fixed. The internet, for instance, by vastly shortening communication interaction, allows a much richer plethora of partial small worlds to form and dissolve. By contrast, many other situations are characterised by long range and idiosyncratic relationships that can nonetheless have a powerful effect on the net dynamics, e.g. scientific research interactions (cf. Section 4.1.3, penultimate paragraph, and 4.2.5 above). Third, because networks are defined by none-node relations, especially if confined to local neighbourhood relations, they are as ill-equipped to represent organisation and global constraints as is phase space dynamics (section 5.1.1, [Hooker-b, this volume]). Thus it is but a tempting mistake to suppose that a city, business firm or research laboratory is simply a small world. They are small worlds, but that does not distinguish them from neighbourhood barbecues and internet twitters. In fact these entities posses significant autonomy and are no more merely small worlds than are biological organisms. Indeed, fourth, the appearance of small worlds phenomena might reflect no more than the limitations of agent finitude, here expressed in the number of connections that can be successfully maintained. Insofar as this is akin to the attention constraint to 7 ± 2 bits [Miller, 1956], it too is independent of subject matter and therefore offers no specific illumination (cf. [Fox Keller, 2005]).

Minimising interaction distance is attractive for finite agents and finitude is a powerful, many-sided constraint. As finite agents, persons are typically overwhelmed with the complexity of potential social interaction and sensibly seek ways to simplify their deliberations and decision making. This is accomplished primarily through the creation of shared expectation-setting social structure, simplified agent models (e.g. personality/ propensity profiles: irritable, generous, ...) and the use of simplified 'good enough' rational procedures ('rules of thumb', satisficing — see [Simon, 1947; 1969; Foster, this volume]), correcting these locally only when problems arise. Agent based modelling with finitude, especially satisficing, constraints is also thriving in social simulation (see, e.g, [Brock and Hommes, 1998; Hommes, 2001] on heterogeneous agents and bounded rationality). On the other hand, in focusing on this aspect, we should again not pre-emptively restrict agent social interaction capacities: the same finite agents that will use cheap and dirty heuristics to engage in Rapoport's fights and play Rapoport's games are also capable of the sensitive, flexible and rationally creative and constructive interaction called for by Rapoport's debates (see [Rapoport, 1970]). Whether and how all the human social richness will find its place in social simulation remains open, a huge and interesting challenge for future dynamical modelling — and in robotics as well (to add to their autonomy challenge, see 4.1.1 *Autonomy* and section 4.2.5 above).

ACKNOWLEDGMENTS

This essay and its closing companion essay have especially benefited from constructive comment on an earlier draft by Bill Bechtel, Mark Bickhard, Thomas Brinsmead, David Green, Jan Schmidt and Olaf Wolkenhauer. I am grateful for their substantial efforts, and to several other authors for particular helpful remarks. Above all I am grateful to my wife Jean whose practical generosity amid her own obligations made a long, time-consuming task so much easier and whose love enlightens everything.

BIBLIOGRAPHY

[Aharanov and Sparks, 1999] E. F. Aharonov and D. Sparks. Rigidity phase transition in granular packings. *Physical Review E*, 60, 68906896, 1999.

[Ahouse, 1998] J.C. Ahouse. The tragedy of apriori selectionism: Dennett and Gould on adaptationism. *Biology and Philosophy*, 133, 359-91, 1998.

[Ahouse and Berwick, 1998] J. C. Ahouse, R. C. Berwick. Darwin on the Mind. *Boston Review*, 23, 36-41, 1998.

[Andersen *et al.*, 2000] P. B. Andersen, P. V. Christiansen, C. Emmeche, N. O. Finnemann (eds.), *Downward Causation: Minds, Bodies, and Matter*, Aarhus University Press, Aarhus, 2000.

[Anderson, 1972] P.W. Anderson. More is different: broken symmetry and the nature of the hierarchical structure of science. *Science*, 177, 393–396, 1972. Reprinted in [Bedau and Humphries, 2008], pp. 221-230, 1008.

[Anderson *et al.*, 1988] P. W. Anderson, K. J. Arrow, D. Pines (eds.). *The economy as an evolving complex system.* Addison Wesley, Redwood City, Cal., 1998.

[Antunes and Takadama, 2007] L. Antunes and K. Takadama. (eds.). *MultiAgentBased Simulation VII*, Lecture Notes in Artificial Intelligence. Springer, Berlin, 2007.

[Arthur *et al.*, 1997] W. B. Arthur, S. N. Durlauf, D. A. Lane. *The economy as an evolving complex system II*, Addison-Wesley, Reading MA, 1997.

[Ashby, 1947] W. R. Ashby. Principles of the SelfOrganizing Dynamic System. *Journal of General Psychology*, 37, 125128, 1947. Also in H. Von Foerster and G. W. Zopf, Jr. (eds.), Principles of SelfOrganization, Pergamon, New York; republished in *Emergence: Complexity and Organization* (E:CO), 6(12), 102126, 2004.

[Ashby, 1952] W. R. Ashby. *Design for a Brain.* Chapman & Hall, London, 1952.

[Aunger, 2000] R. Aunger (ed.). *Darwinizing Culture: The State of Memetics as a Science.* Oxford University Press, Oxford, 2000.

[Auyang, 1998] S. Y. Auyang. *Foundations of complex systems theories in economics, biology and statistical physics.* Cambridge University Press, Cambridge, 1998.

[Axelrod, 1997] R. Axelrod. Advancing the art of simulation in the social sciences. *Complexity*, 3(2), 16- 22, 1997.

[Barabási, 2002] A.-L. Barabási. *Linked: the new science of networks.* Perseus Books, Cambridge, 2002.

[Barabási and Bonabeau, 2003] A.-L. Barabási and E. Bonabeau. Scale-free networks. *Scientific American*, 288(May), 50-59, 2003.

[Barandiaran and Moreno, 2008] X. Barandiaran and A. Moreno. Adaptivity: from metabolism to behavior. *Adaptive Behavior*, 16(5), 325-344, 2008.

[Barandiaran and Ruiz-Mirazo, 2008] X. Barandiaran and K. Ruiz-Mirazo, eds. *Modelling autonomy: simulating the essence of life and cognition. BioSystems*, 91(2), 2008.

[Batterman, 2002] R.W. Batterman. *The devil in the details: asymptotic reasoning in explanation, reduction and emergence.* MIT Press, Cambridge, Mass., 2002.

[Bechtel, 1989] W. Bechtel. An evolutionary perspective on the re-emergence of cell biology. In [Hahlweg and Hooker, 1989b], pp. 433-457, 1989.
[Bechtel, 2007] W. Bechtel. Biological mechanisms: organised to maintain autonomy. In [Boogerd, *et al.*, 2007], pp. 269-302, 2007.
[Bechtel and Richardson, 1993] W. Bechtel, R. Richardson. *Discovering complexity.* Princeton University Press, Princeton, 1993.
[Bedau, 1997] M. A. Bedau. Weak emergence. In J. Toberlin (ed.), Philosophical perspectives: mind, causation and world. Blackwell, Oxford, 375-399, 1997.
[Bedau, 2007] M. A. Bedau. What is life? In S. Sarkar, A. Plutynski (eds.), *A Companion to the Philosophy of Biology.* Blackwell, New York, 455471, 2007.
[Bedau and Humphries, 2008] M. Bedau and P. Humphries (eds.), *Emergence: contemporary readings in philosophy and science*, Bradford/MIT Press, Cambridge, Mass., 2008.
[Beer *et al.*, 1993] R. D. Beer, R. E. Ritzmann, T. McKenna (eds.). *Biological neural networks in invertebrate neuroethology and robotics.* Academic Press, New York, 1993.
[Beer, 1972] S. Beer. *The brain of the firm.* Allen Lane, The Penguin Press, London. 1972.
[Beer, 1979] S. Beer. *The heart of the enterprise.* John Wiley, New York, 1979.
[Behe, 2000] M. J. Behe. Self-organization and irreducibly complex systems: a reply to Shanks and Joplin. *Philosophy of Science*, 67, 155-162, 2000.
[Belletti *et al.*, 2009] F. Belletti, A. Cruz, L. A. Fernandez, A. GordilloGuerrero, M. Guidetti, A. Maiorano, F. Mantovani, E. Marinari, V. MartinMayor, J. Monforte, A. Muñoz Sudupe, D. Navarro, G. Parisi, S. PerezGaviro, J. J. RuizLorenzo, S. F. Schifano, D. Sciretti, A. Tarancon, R. Tripiccione, D. Yllanes. An InDepth View of the Microscopic Dynamics of Ising Spin Glasses at Fixed Temperature. *Journal of Statistical Physics*, 135(56), 1121-1158, 2009.
[Bennett, 1985] C. Bennett. Dissipation, information, computational complexity and the definition of organization. In [Pines, 1985], 1985.
[Bennett, 1992] C. Bennett. Logical depth and other algorithmically defined properties of finite objects. *Proceedings, Workshop on Physics and Computation*, IEEE, 75-77, 1992.
[Bennett, 1986] S. Bennett. *A history of control engineering, 1800-1930.* Peregrinus/ Institution of Electrical Engineers, London, 1986 (1^{st} edn, 1979).
[Bennett, 1993] S. Bennett. *A history of control engineering, 1930-1955.* Peregrinus/ Institution of Electrical Engineers, London, 1993.
[Bickhard, 1993] M. Bickhard. Representational content in humans and machines. *Journal of Experimental and Theoretical Artificial Intelligence*, 5, 285-333, 1993.
[Bickhard, 2000a] M.H. Bickhard. Emergence. In [Andersen, *et al.*, 2000], pp. 322-348, 2000.
[Bickhard, 2000b] M.H. Bickhard. Autonomy, function, and representation. In A. Etxebberia, A. Moreno, J. Umerez (eds.), *The contribution of artificial life and the sciences of complexity to the understanding of autonomous systems*, *Communication and Cognition*, Special Edn, 173-4, 111-131, 2000.
[Bickhard, 2001] M.H. Bickhard. Error dynamics: the dynamic emergence of error avoidance and error vicariants. *Journal of Experimental and Theoretical Artificial Intelligence*, 13, 199-209, 2001.
[Bickhard, 2002] M.H. Bickhard. Critical principles: on the negative side of rationality. *New Ideas in Psychology*, 20, 1-34, 2002.
[Bickhard, 2005] M.H. Bickhard. Interactivism: a manifesto. At `http://www.lehigh.edu/~mhb0/`.
[Bickhard, 2008] M.H. Bickhard. *The Whole Person: toward a naturalism of persons.* At `http://www.lehigh.edu/~mhb0/`.
[Bickhard and Terveen, 1995] M. Bickhard, L. Terveen. *Foundational issues in artificial intelligence and cognitive science: impasse and solution.* Elsevier, Amsterdam, 1995.
[Bickle, 2003] J. Bickle. *Philosophy and neuroscience: a ruthlessly reductive account.* Kluwer, Boston, 2003.
[Bickle, 2006] J. Bickle. Reducing mind to molecular pathways: explicating the reductionism implicit in current cellular and molecular neuroscience. *Synthese*, 151, 411-434, 2006.
[Blackmore, 1999] S. Blackmore. *The Meme Machine.* Oxford University Press, Oxford, 1999.
[Blake *et al.*, 1960] R. M. Blake, C. J. Ducasse, E. H. Madden. *Theories of scientific method.* University of Washington Press, Seattle, 1960.
[Bloch, 2003] A.M. Bloch. *Nonholonomic mechanics and control.* Springer, New York, 2003. See `http://www.cds.caltech.edu/mechanics_and_control/about_this_book/` for corrections and additions, incl. internet supplement.

[Blume and Durlauf, 2006] S. N. Durlauf. *The economy as an evolving complex system III*, Oxford University Press, Oxford, 2006.

[Boccaletti *et al.*, 2006] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D. U. Hwang. Complex networks: structure and dynamics. *Physics Reports*, 424, 175-308, 2006.

[Boogerd *et al.*, 2002] F. C. Boogerd, F. J., Bruggeman, C. Jonker, H. L. De Jong, A. Tamminga, J. Treur, H. V. Westerhoff, W. Wijngaards. Interlevel relations in computer science, biology, and psychology. *Philosophical Psychology*, 15(4), 463-472, 2002.

[Boogerd *et al.*, 2005] F. C. Boogerd, F. J., Bruggeman, R. C. Richardson, A. Stephan, H. V. Westerhoff. Emergence and its place in nature; a case study of biochemical networks. *Synthese* 145, 131-164, 2005.

[Boogerd *et al.*, 2007] F. Boogerd, F. Bruggeman, J-H. Hofmeyr, H. V. Westerhoff (eds.), *Systems Biology: Philosophical Foundations*, Amsterdam, Elsevier, 2007.

[Brading and Castellani, 2003] K. Brading and E. Castellani, eds. *Symmetries in physics: philosophical reflections*. Cambridge University Press, Cambridge, 2003.

[Braitenberg, 1984] V. Braitenberg. *Vehicles: experiments in synthetic psychology*. MIT Press, Cambridge MA, 1984.

[Bressler and Kelso, 2001] S. L. Bressler and S. Kelso. Cortical coordination dynamics and cognition. *Trends in Cognitive Sciences*, 5(1), 26-36, 2001.

[Brinsmead and Hooker, 2005] T. S. Brinsmead and C. A. Hooker. *Sustainabilities: a systematic framework and comparative analysis*, Research Report 53. Cooperative Research Centre for Coal in Sustainable Development, Brisbane, (pp.113), 2005.

[Brock and Hommes, 1998] W. A. Brock and C. H. Hommes. Heterogeneous beliefs and bifurcation routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22, 12351274, 1998.

[Bronowski, 1965] J. Bronowski. *The identity of man*. Heinemann, London, 1965.

[Brooks, 1991] R. A. Brooks. Intelligence without representation. *Artificial Intelligence* 47, 139-159, 1991.

[Brown and Harré, 1988] H. R. Brown and R. Harré (eds.). Philosophical foundations of quantum field theory. Clarendon Press, Oxford, 1988.

[Bruggeman *et al.*, 2002] F. J. Bruggeman, H. V. Westerhoff, and F. C. Boogerd. BioComplexity: A pluralist research strategy is necessary for a mechanistic explanation of the 'live' state. *Philosophical Psychology*, 15(4), 411 — 440, 2002.

[Campbell, 1974a] D. T. Campbell. Downward causation in hierarchically organised biological systems. In F. J. Ayala, T. Dobzhansky (eds.), *Studies in the philosophy of biology*. University of California Press, Berkeley, 179-186, 1974.

[Campbell, 1974b] D. T. Campbell. Evolutionary epistemology. In P. A. Schilpp (ed.), *The philosophy of Karl Popper*. Open Court, La Salle Ill., 1974.

[Campbell, 1977] D. T. Campbell. Comment on "The natural selection model of conceptual evolution". *Philosophy of Science*, 44, 502-507, 1977.

[Campbell, 1986] D. T. Campbell. Science's social system of validity-enhancing collective belief change and the problems of the social sciences. In D. W. Fiske, R. A. Shwader (eds.), *Pluralisms and subjectivities*. University of Chicago Press, Chicago, 1986.

[Campbell and Paller, 1989] D. T. Campbell, B.T. Paller. Extending Evolutionary Epistemology to 'Justifying' Scientific Beliefs. In [Hahlweg and Hooker, 1989b], pp. 231-257, 1989.

[Campbell, 1997] D. T. Campbell. From evolutionary epistemology via selection theory to a sociology of scientific validity. *Evolution and Cognition*, 3, 5-38, 1977.

[Cariani, 1992] P. Cariani. Some epistemological implications of devices which construct their own sensors and effectors. In F. Varela and P. Bourgine (eds.), *Towards a practice of autonomous systems*. MIT Press, Cambridge Mass., 1992.

[Cariani, 1993] P. Cariani. To evolve an ear. *Systems research*, 10(3), 19-33, 1993.

[Cash *et al.*, 2006] D. W. Cash, W. N. Adger, F. Berkes, P. Garden, L. Lebel, P. Olsson, L. Pritchard, O. Young. Scale and Cross-Scale Dynamics: Governance and Information in a Multilevel World. *Ecology and Society*, 11(2), 2006. `http://www.ecologyandsociety.org/`

[Cenys *et al.*, 2003] A. Cenys, A. Tamasevicius, A. Baziliauskas, R. Krivickas, E. Lindberg. Hyperchaos in coupled Colpitts oscillators. Chaos, Solitons & Fractals, 17(23), 349353, 2003.

[Cherniak, 1986] C. Cherniak. *Minimal rationality*. MIT Press, Cambridge, Mass., 1986.

[Christensen, 1999] W. D. Christensen. *An interactivist-constructivist approach to adaptive intelligence and agency*, PhD thesis, University of Newcastle, Australia, 1999.

[Christensen and Hooker, 1998] W. D. Christensen and C. A. Hooker. From cell to scientist: Toward an organisational theory of life and mind. In J. Bigelow (ed.), *Our Cultural Heritage*. Australian Academy of Humanities, University House, Canberra, 275-326, 1988.

[Christensen and Hooker, 1999] W. D. Christensen and C. A. Hooker. The organization of knowledge: beyond Campbell's evolutionary epistemology. *Philosophy of Science*, 66, Proceedings, PSA 1998, S237-S249, 1999.

[Christensen and Hooker, 2000a] W. D. Christensen and C. A. Hooker. Organised interactive construction: The nature of autonomy and the emergence of intelligence. In A. Etxebberia, A. Moreno, J. Umerez (eds.), *The Contribution of Artificial Life and the Sciences of Complexity to the Understanding of Autonomous Systems*, *Communication and Cognition*, special edn, 17(3-4), 133-158, 2000.

[Christensen and Hooker, 2000b] W. D. Christensen and C. A. Hooker. An Interactivist-Constructivist Approach to Intelligence: Self-directed Anticipative Learning. *Philosophical Psychology*, 13, 5-45, 2000.

[Christensen and Hooker, 2002] W. D. Christensen and C. A. Hooker. Self-directed agents. In J. MacIntosh (ed.), *Naturalism, Evolution and Intentionality*, Ottawa, *Canadian Journal of Philosophy*, Special Supplementary Volume 27, 19-52, 2002.

[Christensen and Hooker, 2004] W. D. Christensen and C. A. Hooker. Representation and the meaning of life. In H. Clapin, P. Staines, P. Slezak (eds.), *Representation in mind: new approaches to mental representation*. Sydney, Elsevier, 41-69, 2004.

[Churchland, 1989] P. M. Churchland. *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Bradford/MIT Press, Cambridge, Mass., 1989.

[Churchland, 1995] P. M. Churchland. *The Engine of Reason, the Seat of the Soul*, MIT Press, Cambridge Mass.,1995.

[Churchland, 1986] P. S. Churchland. *Neuro-philosophy: toward a unified understanding of the mind-brain*. MIT Press, Cambridge Mass., 1986.

[Clark, 1997] A. Clark. *Being There: Putting Brain, Body, and World Together Again*. Bradford/MIT Press, Boston, 1997.

[Collier and Hooker, 1999] J. D. Collier and C. A. Hooker. Complex Organised Dynamical Systems. *Open Systems and Information Dynamics*, 6, 241-302, 1999.

[Coltheart, 2006a] M. Coltheart. What has functional neuroimaging told us about the mind (so far)? *Cortex*, 42, 323-331, 2006.

[Coltheart, 2006b] M. Coltheart. Perhaps functional neuroimaging has not told us anything about the mind (so far). *Cortex*, 42, 422-427, 2006.

[Cragg, 1989] B. Cragg. Evolution of the steam engine. In [Hahlweg and Hooker, 1989b], pp. 313-356, 1989.

[Craver, 2002] C. Craver. Interlevel experiments and multilevel mechanisms in the neuroscience of memory. Philosophy of Science Supplement, 69, S83-S97.

[Craver, 2007] C. Craver. Explaining the brain: mechanisms and the mosaic unity of neuroscience. Oxford University Press, Oxford, 2007.

[Craver and Bechtel, 2007] C. F. Craver and W. Bechtel. Topdown causation without topdown causes. *Biology and Philosophy*, 22, 547563, 2007.

[Crutchfield, 1994] J. Crutchfield. The calculi of emergence: computation, dynamics and induction. *Physica D*, 75, 11-54, 1994.

[Cumming and Collier, 2009] G. Cumming and J. Collier. Change and identity in complex systems. In B. Brown, K. de Laplante, K. Peacock (eds.), *Philosophy of Ecology*. Elsevier, Amsterdam, 2009.

[Cunningham, 1972] M. Cunningham. *Intelligence: Its Origin and Development*. Academic Press, New York, 1972.

[Depew, Weber, 1999,] D. Depew, B. Weber. *Darwinism Evolving*. MIT, Cambridge Mass., 1999.

[Dieckmann *et al.*, 2000] U. Dieckmann, R. Law, J. Metz (eds.). *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*. Cambridge University Press, Cambridge, 2000.

[Dreyfus, 1996] H. Dreyfus. The current relevance of Merleau-Ponty's phenomenology of embodiment. H. Haber, G. Weiss (eds.), *Perspectives on embodiment*. Routledge, London, 1996.

[Dreyfus and Dreyfus, 1986] H. L. Dreyfus, S. E. Dreyfus. *Mind over machine*. Blackwell, Oxford, 1986.

[Dyke, 1988] C. Dyke. *The evolutionary dynamics of complex systems*. Oxford University Press, Oxford, 1988.

[Earman *et al.*,, 2002] J. Earman, C. Glymour, S. Mitchell. *Ceteris paribus laws.* Kluwer, Boston, 2002.

[Eigen and Schuster, 1979] M. Eigen and P. Schuster. *The Hypercycle: A principle of natural selforganization.* Springer, Berlin, 1979.

[Eigen and Winkler, 1993] M. Eigen and R. Winkler. *The Laws of the Game: How The Principles of Nature Govern Chance.* Princeton University Press, Princeton, 1993.

[Eldredge, 1985] N. Eldredge. *Unfinished Synthesis: Biological Hierarchies and Modern Evolutionary Thought.* Oxford University Press, New York, 1985.

[Elias-Smith, 2000] C. Eliasmith. Is the brain analog or digital? The solution and its consequences for cognitive science. *Cognitive Science Quarterly*, 1(2),147170, 2000.

[Elias-Smith, 2009a] Dynamics, control and cognition. In P. Robbins, M. Aydede (eds.), *Cambridge Handbook of Situated Cognition.* Cambridge University Press, Cambridge, 2009.

[Elias-Smith, 2009b] Computational neurosience. In P. Thagard (ed.), *Philosophy of Psychology and Cognitive Science.* Handbook of the Philosophy of Science, vol. 4, Elsevier, Amsterdam, 2009.

[Emmeche *et al.*, 2000] C. Emmeche, S. Koppe, F. Stjernfelt. Levels, emergence, and three versions of downward causation. In [Andersen, *et al.*, 2000], pp. 13-34, 2000.

[Epstein, 2007] J. M. Epstein. *Generative social science: studies in agentbased computational modeling*, Princeton University Press, Princeton, 2007.

[Espejo and Harnden, 1989] R. Espejo and R. Harnden (eds.). *The Viable System Model.* John Wiley, New York, 1989.

[Etxebberia, 1998] A. Exteberria. Embodiment of natural and artificial agents. Iin G. van der Vijver, S. N. Salthe and M. Delpos (eds.), *Evolutionary Systems.* Reidel, Dordrecht, 1998.

[Etxeberria and Moreno, 2001] A. Etxeberria and A. Moreno. From complexity to simplicity: nature and symbols. *BioSystems*, 60, Special issue on the Physics and Evolution of Symbols and Codes: Reflections on the Work of Howard Pattee, 149-157, 2001.

[Farrell and Hooker, 2007a] R Farrell and C. A. Hooker. Applying self-directed anticipative learning to science: agency and the interactive exploration of possibility space in ape language research. *Perspectives on Science*, 15(1), 86-123, 2007.

[Farrell and Hooker, 2007a] R Farrell and C. A. Hooker. Applying self-directed anticipative learning to science: learning how to learn across revolutions. *Perspectives on Science*, 15(2), 220-253, 2007.

[Farrell and Hooker, 2009] R. Farrell and C. A. Hooker. Applying self-directed anticipative learning to science: identifying error and severe testing through interactive exploration of possibility. *Foundations of Science*, 2009.

[Ferber, 1999] J. Ferber. *Multi-agent Systems: An Introduction to Distributed Artificial Intelligence,* Addison -Wesley, New York, 1999.

[Fogel *et al.*, 2008] A. Fogel, B. J. King, S. G. Shankar. *Human development in the twenty first century: visionary ideas from systems scientists.* Cambridge University Press: Cambridge, 2008.

[Fox Keller, 2005] E. Fox Keller. Revisiting 'scale-free' networks. *BioEssays*, 27, 1060–1068, 2005.

[Fox Keller, 2007] E. Fox Keller. The disappearance of function from 'self-organizing systems'. In [Boogerd, *et al.*, 2007], pp. 303-317, 2007.

[Fu and Panke, 2009] P. Fu and S. Panke. eds. *Systems Biology and Synthetic Biology*, Wiley, Hoboken N.J., 2009.

[Gánti, 2003] T. Gánti. *The principles of life.* Oxford University Press, New York, 2003.

[Gao, 2005] Y. Gao. *Ecologies of Scientific Practice: An Analysis of the Organization and Dynamics of Science.* PhD thesis, University of Newcastle, Newcastle, Australia, 2005.

[Galison1987,] P. Galison. *How experiments end.* Chicago: University of Chicago Press.

[Gell-Mann, 1994] M. Gell-Mann. *The Quark and the Jaguar: Adventures in the Simple and the Complex.* Henry Holt, New York, 1994.

[Glass, 1988] L. Glass, M. Mackey. *From Clocks to Chaos.* Princeton University Press, Princeton, 1988.

[Gleick, 1988] J. Gleick. *Chaos: making a new science.* Penguin, New York, 1988.

[Goldbeter, 1996] A. Goldbeter. *Biochemical Oscillations and Cellular Rhythms: the Molecular Bases of Periodic and Chaotic Behaviour.* Cambridge University Press: Cambridge, 1996.

[Goldspink, 2002] C. Goldspink. Methodological implications of complex systems approaches to sociality: simulation as a foundation for knowledge. *Journal of Artificial Societies and Social Simulation*, 5(1), 2002. Available at `http://jasss.soc.surrey.ac.uk/5/1/3.html`.

[Goldstein, 2002] J. Goldstein. The singular nature of emergent levels: suggestions for a theory of emergence. *Non-linear dynamics, psychology, and life sciences*, 6(4), 293-309, 2002.

[Goodwin, 1994] B. Goodwin. *How the leopard changed its spots*. Weidenfeld & Nicolsen, London, 1994.

[Goodwin and Saunders, 1989] B. Goodwin and P. Saunders. *Theoretical Biology: Epigenetic and Evolutionary Order from Complex Systems*. Edinburgh University Press, Edinburgh, 1989.

[Gray, 1992] R. Gray. Death of the gene: developmental systems strike back. In P. Griffiths (ed.), *Trees of Life: Essays in Philosophy of Biology*. Kluwer, Dordrecht, 1992.

[Gray and Scott, 1994] P. Gray and S. K. Scott. *Chemical Oscillations and Instabilities: Non-linear Chemical Kinetics*. Clarendon Press: Oxford England, 1994.

[Griffiths, 1992] P. E. Griffiths (ed.). *Trees of Life: Essays in Philosophy of Biology*. Kluwer, Dordrecht, 1992.

[Griffiths and Stotz, 2007] P. E. Griffiths and K. Stotz. Gene. In M. Ruse, D. Hull (eds.), *Cambridge Companion to Philosophy of Biology*. Cambridge University Press, Cambridge, 85-102, 2007.

[Groß and Dorigo, 2007] R. Groß and M. Dorigo. Fifty years of selfassembly experimentation. In W.M. Shen, H. Lipson, K. Stoy, M. Yim (eds.), Proceedings of the Workshop on Self-Reconfigurable Robots/Systems and Applications, USC Information Science Institute, 2007

[Groß and Dorigo, 2008] R. Groß and M. Dorigo. Selfassembly at the macroscopic scale. *Proceedings of IEEE*, 96(9), 14901508, 2008.

[Guckenheimer and Holmes, 1983] J. Guckenheimer and P. Holmes. *Non-linear oscillations, dynamical systems and bifurcations of vector fields*. Springer, New York, 1983.

[Gunderson *et al.*, 1997] L. Gunderson, C. Holling, L. Pritchard, G. Peterson. *Resilience in Ecosystems, Institutions and Societies*. Beijer Discussion Paper Series, No. 95. Beijer International Institute of Ecological Economics, Stockholm, 1997.

[Gunderson and Holling, 2002] L. Gunderson and C. Holling (eds.). *Panarchy: Understanding Transformations in Human and Natural Systems.* Washington, Island Press, 2002.

[Hahlweg, 1983] K. Hahlweg. *The evolution of science: a systems approach*. Ph.D. thesis, University of Western Ontario, Canada, 1983.

[Hahlweg, 1991] K. Hahlweg. On the notion of evolutionary progress. *Philosophy of science* 58, 436-51, 1991.

[Hahlweg and Hooker, 1989a] K. Hahlweg and C. A. Hooker. Evolutionary epistemology and philosophy of science. In [Hahlweg and Hooker, 1989b], 21-150.

[Hahlweg and Hooker, 1989b] K. Halhweg and C. A. Hooker (eds.). *Issues in Evolutionary Epistemology*. State University of New York Press, Albany N.Y., 1989.

[Haila and Levins, 1992] Y. Haila and R. Levins. *Humanity and Nature: Ecology, Science and Society*. Pluto Press, London, 1992.

[Haken, 1983] H. Haken. *Synergetics, an Introduction: Nonequilibrium Phase Transitions and SelfOrganization in Physics, Chemistry, and Biology*, 3rd rev. enl. edn. SpringerVerlag, New York, 1983.

[Haken, 1993] H. Haken. *Advanced Synergetics: Instability Hierarchies of SelfOrganizing Systems and Devices*. SpringerVerlag, New York, 1993.

[Hastie *et al.*, 2008] T. Hastie, R. Tibshirani, J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*, 2^{nd} edn. Springer, New York, 2008.

[Heath, 2000] R. Heath. *Nonlinear Dynamics: Techniques and Applications in Psychology*. Erlbaum, Mahwah, N.J., 2000.

[Heylighen, 1991] F. Heylighen Modelling emergence. *World futures (Journal of general evolution)*, 31 (Special Issue on Emergence, edited by G. Kampis), 89104, 1991.

[Herfel and Hooker, 1996] W. E. Herfel and C. A. Hooker. Cognitive dynamics and the development of science. In D. Ginev and R. S. Cohen (eds.), *Issues and images in the philosophy of science*. Kluwer, Boston, 127-172, 1996.

[Herfel and Hooker, 1999] W.E. Herfel and C. A. Hooker. From formal machine to social colony: toward a complex dynamical philosophy of science. In M. L. Dalla Chiara, R. Giuntini, F. Laudisa (eds.), *Language, Quantum, Music: Select Proceedings of the 10th International Congress of Logic, Methodology and Philosophy of Science*. Boston, Kluwer, 7-18, 1999.

[Heiby *et al.*, 2003] E. M. Heiby, I. Pagano, D. Blaine, K. Nelson, R. A. Heath. Modeling unipolar depression as a chaotic process. *Psychological Assessment*, 15, 426434, 2003.

[HoyningenHuene, 1993] P. HoyningenHuene. *Reconstructing Scientific Revolutions: Thomas S. Kuhn's Philosophy of Science*. University of Chicago Press, Chicago, 1993.

[Hodges, 1997] B. Hodges. *Retooling Popper: conventional decision and the rationality of improvement*. PhD thesis, University of Newcastle, Newcastle, Australia, 1997.

[Hogeweg, 2002a] P. Hogeweg. Computing an organism: on the interface between informatic and dynamic processes. *Biosystems*, 64, 97-109, 2002.

[Hogeweg, 2002b] P. Hogeweg. Multilevel processes in evolution and development: computational models and biological insights. In M. Lässig, A. Valleriani (eds.), *Biological Evolution and Statistical Physics*. Springer lecture notes in physics 585, Springer Verlag, Berlin, 217-239, 2002.

[Hogeweg and Takeuchi, 2003] P. Hogeweg and N. Takeuchi. Multilevel selection in models of prebiotic evolution: compartments and spatial self-organization. *Origins of Life and Evolution of the Biosphere*, 33, 375-403, 2003.

[Holland, 1995] J. Holland. *Hidden Order: How Adaptation Builds Complexity*. Addison Wesley, Redwood City, Cal.; Perseus Books, Cambridge MA, 1995.

[Holland 1998,] J. H. Holland. *Emergence: from chaos to order*. Addison Wesley, Redwood City, Cal.; Basic Books, New York, 1998.

[Holling, 1992] C. Holling. Cross-scale morphology, geometry and dynamics of ecosystems. *Ecological monographs,* 62(4), 447-502, 1992.

[Holmes, 2005] P. Holmes. Ninety plus thirty years of nonlinear dynamics: less is more and more is different. *International Journal of Bifurcation and Chaos*, 15(9), 2703-2716.

[Hommes, 2001] C. H. Hommes. Financial markets as nonlinear adaptive evolutionary systems. Quantitative Finance, 1, 14967, 2001.

[Hooker, 1978] C. A. Hooker. An Evolutionary Naturalist Realist Doctrine of Perception and Secondary Qualities. In C. W. Savage (ed.), M*innesota Studies in the Philosophy of Science* Vol. IX. University of Minnesota Press, Minneapolis, 405-440, 1978.

[Hooker, 1981a] C. A. Hooker. Formalist rationality: the limitations of Popper's theory of reason. *Metaphilosophy* 12, 247-266, 1981.

[Hooker, 1981b] C. A. Hooker. Towards a general theory of reduction. *Dialogue*, XX, Part I, Historical framework, 38-59, Part II, Identity and reduction, 201-36, Part III, Cross-categorial reduction, 496-529, 1981.

[Hooker, 1987] C. A. Hooker. *A RealisticTheory of Science*. State University of New York Press, Albany NY, 1987.

[Hooker, 1989] C. A. Hooker. From logical formalism to control system. in A. Fine, J. Leplin (eds.), *PSA 1988*, East Lansing, Philosophy of Science Association, 211-221, 1989.

[Hooker, 1994a] C. A. Hooker. Idealisation, naturalism, and rationality: some lessons from minimal rationality. *Synthese*, 99, 181-231, 1994.

[Hooker, 1994b] C. A. Hooker. From phenomena to metaphysics. In D. Prawitz, D. Westerstahl (eds.), *Logic and Philosophy of Science in Uppsala*. Dordrecht, Kluwer, 159-184, 1994.

[Hooker, 1994c] C. A. Hooker. Regulatory constructivism: on the relation between evolutionary epistemology and Piaget's genetic epistemology. *Biology and Philosophy*, 9, 197-244, 1994.

[Hooker, 1995a] C. A. Hooker. Toward a naturalised cognitive science. In R. Kitchener, W. O'Donohue (eds.), *Psychology and Philosophy*. Allyn and Bacon, London, 184-206, 1995.

[Hooker, 1995b] C. A. Hooker. *Reason, regulation and realism*. State University of New York Press, Albany, 1995.

[Hooker, 2000] C. A. Hooker. Unity of science. In W. H. Newton-Smith (ed.), 2000, *A Companion to the Philosophy of Science*. Blackwell, Oxford, 540-549, 2000.

[Hooker, 2002] C. A. Hooker. An integrating scaffold: toward an autonomy-theoretic modelling of cultural change. In M. Wheeler, J. Ziman (eds.), *The Evolution of Cultural Entities*. British Academy of Science & Oxford University Press, Oxford, 2002.

[Hooker, 2004] C. A. Hooker. Asymptotics, reduction and emergence. *British Journal for the Philosophy of Science*, 55, 435-479, 2004.

[Hooker, 2009a] C. A. Hooker. On fundamental implications of systems and synthetic biology. In [Fu and Panke, 2009], 2009.

[Hooker, 2009b] C. A. Hooker. Interaction and bio-cognitive order. *Synthese*, 166(3), 513-546, 2009.

[Hooker-b, this volume] Introduction to philosophy of complex systems. Part B: Scientific paradigm + philosophy of science for complex systems: a first presentation c. 2009. In this volume.

[Hooker-c, this volume] Conceptualising reduction, emergence and self-organisation in complex dynamical systems. In this volume.

[Howard, 1971] N. Howard. *Paradoxes of rationality: theory of metagames and political behaviour*. MIT Press, Cambridge Mass., 1971.

[Humphries, 1994] P. Humphries. Numerical experimentation. In P. Humphries (ed.), *Patrick Suppes, scientific philosopher*. vol. 2. Kluwer, Dordrecht, 103-118, 1994.

[Humphries, 2004] P. Humphries. *Extending ourselves: computational science, empiricism and scientific method*. Oxford University Press, Oxford, 2004.

[Icke, 1995] Icke. *The force of symmetry*. Cambridge University Press, Cambridge, 1995.

[Jablonka and Lamb, 1995] E. Jablonka, M. Lamb. *Epigenetic Inheritance and Evolution: The Lamarckian Dimension*. Oxford University Press, Oxford, 1995.

[Jablonka and Lamb, 2005] E. Jablonka and M. Lamb. *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. MIT Press, Cambridge, Mass., 2005.

[Juarrro, 2002] A. Juarrero. *Dynamics in action: intentional behavior as a complex system*. MIT Press, Cambridge, Mass., 2002.

[Kaplan and Glass, 1995] D. Kaplan and L. Glass. *Understanding Nonlinear Dynamics*. Springer-Verlag, New York, 1995.

[Kauffman, 1995] S. Kauffman. *At Home in the Universe*. Oxford University Press, Oxford, 1995.

[Kauffman, 2000] S. Kauffman. *Investigations*. Oxford University Press, New York, 2000.

[Keilis-Borok and Soloviev, 2003] V. I. Keilis-Borok and A. A. Soloviev , eds. Nonlinear dynamics of the lithosphere and earthquake prediction. New York, Springer, 2003.

[King, 2003] R. B. King. Chirality and Handedness: The Ruch "ShoePotato" Dichotomy in the RightLeft Classification Problem. Annals of the New York Academy of Science, 988, 158170, 2003.

[Kohlmorgan *et al.*, 1999] U. Kohlmorgen, H. Schmeck and K. Haase. Experiences with fine grained parallel genetic algorithms. Annals of Operations Research, 90, 203-219, 1999.

[Kuhn, 1962] T. S. Kuhn. *The Structure of Scientific Revolutions*, 1st. ed.n. University of Chicago Press, Chicago, 1962. 2^{nd} edn, with postscript, 1970.

[Kubic, 2003] A. Kubic. Toward a formalization of emergence. *Artificial Life*, 9(1), 41-65, 2003.

[Lakatos and Musgrave, 1970] I. Lakatos, A. Musgrave (eds.). *Criticism and the Growth of Knowledge*. Cambridge University Press, London, 1970.

[Lalonde *et al.*, 2000] K. N. Laland, J. Odling-Smee, M. W. Feldman. Niche Construction, Biological Evolution and Cultural Change. *Behavioral and Brain Sciences*, 23 (1), 131-146, 2000.

[Landau, 1937] L. D. Landau. Broken symmetry and phase transitions. *Transactions Physics Zeitschrift Sowjetunion*, 11, 26, 542, 1937.

[Latour, 1987] B. Latour. *Science in action*. Harvard University Press, Cambridge Mass., 1987.

[Latour and Woolgar, 1979] B. Latour and S. Woolgar. *Laboratory life: the social construction of scientific facts*. Sage, Beverley Hills, 1979.

[Layzer, 1990] D. Layzer. *Cosmogenesis: the growth of order in the universe*. Oxford University Press, New York, 1990.

[Leonard, 1999] A. Leonard. A Viable System Model: Consideration of Knowledge Management. *Journal of Knowledge Management Practice*, August 1999. Accessed at `http://www.tlainc.com/articl12.htm`, June 2009.

[Lewis, 2005] M. D. Lewis. Bridging emotion theory and neurobiology through dynamic systems modeling. *Behavioural and Brain Sciences*, 28, 168-245, 2005.

[Lewis and Todd, 2007] M. D. Lewis and R. M. Todd. The selfregulating brain: cortical-subcortical feedback and the development of intelligent action. *Cognitive Development*, 22, 406430, 2007.

[Lewis and Cook, 2007] M. D. Lewis, M. L. Cook. Changing habits of emotion regulation at transition points in infancy: A dynamic systems analysis. *Journal of Developmental Processes*, 3, 6789, 2007.

[Lindberg, 2004] E. Lindberg. On the mechanisms behind chaos. *Proceedings of the 12th Workshop on Nonlinear Dynamics of Electronic Systems*, NDES 2004. Evora, Portugal, 227230, 2004.

[Lorenz, 1997] H.-W. Lorenz. *Nonlinear Dynamical Economics and Chaotic Motion*. (1^{st} edn 1989, 2^{nd} rev. & enl. edn, 1993.) Springer, Berlin, 1997.

[Mainzer, 2005] K. Mainzer. *Symmetry and Complexity: the Spirit and Beauty of Non-linear Science*, World Scientific, Singapore, 2005.

[Marr, 1982] D. Marr. *Vision*. MIT Press, Cambridge, Mass.,1982.

[Maturana, 1981] H. R. Maturana. Autopoiesis. In M. Zeleny (ed.), *Autopoiesis: A theory of the living organization*. Westview Press, Boulder, 1981.

[Maturana and Varela, 1980] H. Maturana and F. Varela. *Autopoiesis and cognition*. Reidel, Dordrecht, 1980.

[May, 1976] R. May, ed. *Theoretical Ecology: Principles and Applications*. Saunders, Philadelphia, 1976.

[Mayo, 1996] D. Mayo. *Error and the Growth of Experimental Knowledge*. University of Chicago Press, Chicago, 1996.

[Maynard-Smith, 1986] J. Maynard-Smith. *The problems of biology*. Oxford University Press, New York, 1986.

[Maynard-Smith and Szathmary, 1995] J. Maynard-Smith and E. Szathmary. *The major transitions in evolution*. Freeman, New York, 1995.

[Mayr, 2004] E. Mayr *What Makes Biology Unique?* Cambridge University Press, Cambridge, 2004.

[McCullogh and Pitts, 1943] W. S. McCullogh and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5,115133, 1943. But cf. `http://www.dlsi.ua.es/~mlf/nnafmc/pbook/node10.html`.

[McKelvey, 1999] B. McKelvey. Complexity theory in organization science: seizing the promise or becoming a fad?. *Emergence*, 1(1), 5-32, 1999.

[McLaughlin, 1992] B. McLaughlin. The rise and fall of British emergentism. In A. Beckerman, H. Flor, J. Kim (eds.), *Emergence or reduction?: essays on the prospect of nonreductive physicalism*. de Gruyter, Berlin, 1992. Reprinted in [Bedau and Humphries, 2008, pp. 19-60].

[Merleau-Ponty, 1962] M. Merleau-Ponty. *Phenomenology of perception* (transl. C. Smith). Routledge & Kegan-Paul, London, 1962.

[Miklos, 1993] G. Miklos. Molecules and cognition: the latterday lessons of levels, language and *lac*. *Journal of Neurobiology*, 246, 842-890, 1993.

[Miller, 1956] G. A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 8197, 1956.

[Moreno, 2007] A. Moreno. A systemic approach to the origin of biological organization. In [Boogerd, *et al.*, 2007, pp. 243-268].

[Moreno and Ruiz-Mirazo, 1999,] A. Moreno and K. Ruiz-Mirazo. Metabolism and the problem of its universalization, *Biosystems*, 49, 45-61, 1999.

[Moreno and Lasa, 2003] A. Moreno and A. Lasa. From basic adaptivity to early mind: the origin and evolution of cognitive capacities, *Evolution and Cognition*, 9(1), 12-30, 2003.

[Morowitz, 1978] H. J. Morowitz. *Foundations of bioenergetics*. Academic Press, New York, 1978.

[Morowitz, 2002] H. J. Morowitz. *The Emergence of Everything*. Oxford University Press, Oxford, 2002.

[Morowitz and Singer, 1995] H. J. Morowitz, J. L. Singer (eds.). *The Mind, The Brain, and Complex Adaptive Systems*. Addison-Wesley, Reading Mass., 1995.

[Mossio *et al.*, 2009] M. Mossio, G. Longo, J. Stewart. A computable expression of closure to efficient causation, *Journal of Theoretical Biology*, in press, 2009. (`http://dx.doi.org/10.1016/j.jtbi.2008.12.012`)

[Nagel, 1961] E. Nagel. *The Structure of Science*. Routledge & Kegan Paul, London, 1961.

[Nelson and Winter, 1982] R. R. Nelson and S. G. Winter. *An Evolutionary Theory of Economic Change*. Harvard University Press, Harvard, 1982.

[Nicolis, 1986] J. S. Nicolis. *Dynamics of hierarchical systems: an evolutionary approach*. Springer, Berlin, 1986.

[O'Connor and Wong, 2002] T. O'Connor and H. Y. Wong. Emergence. Stanford Encyclopedia of Philosophy, `http://plato.stanford.edu/entries/properties-emergent/`, 2002.

[Odum, 1971] H. Odum. *Environment, Power and Society*. Wiley, New York, 1971.

[Oldroyd, 1989] D. Oldroyd. *The arch of knowledge*. University of NSW Press, Kensington N.S.W., 1989.

[Oreskes, 2000a] N. Oreskes. Why believe a computer? Models, measures, and meaning in the natural world. In J. S. Schneiderman (ed.) *The Earth Around Us: Maintaining a Livable Planet*. San Francisco, W.H. Freeman and Co., 7082, 2000a.

[Oreskes, 2000b] N. Oreskes. Why predict? Historical perspectives on prediction in the earth sciences. In D. Sarewitz, R. Pielke, Jr., R. Byerly, Jr. (eds.) *Prediction: Decisionmaking and the Future of Nature*. Washington, D.C., Island Press, 2340, 2000b.

[Oreskes, 2003] N. Oreskes. The role of quantitative models in science. In C. D. Canham, J. J. Cole, W. K. Lauenroth (eds.) *Models in Ecosystem Science*. Princeton, Princeton University Press, 1331, 2003.

[Oreskes, 2007] N. Oreskes. From scaling to simulation: Changing meanings and ambitions of models in the Earth sciences. In A. N. H. Creager, E. Lunbeck, M. N. Wise (eds.) *Science without Laws: Model Systems, Cases, and Exemplary Narratives*. Duke University Press, 2007.

[Oreskes and Fleming, 2000] N. Oreskes, J. R. Fleming, (eds.). Perspectives on Geophysics. Special Issue of Studies in the History and Philosophy of Modern Physics, 31B, September 2000.

[Oyama *et al.*, 2001] S. Oyama, P. Griffiths, R. Gray (eds.), *Cycles of Contingency: Developmental Systems and Evolution,* MIT Press, Cambridge Mass., 2001.

[Pahl-Wostl, 1995] C. Pahl-Wostl. *The Dynamic Nature of Ecosystem: Chaos and Order Entwined*. Wiley, New York, 1995.

[Pask, 1960] G. Pask. The natural history of networks. In M. C. Yovits, S. Cameron (eds.), *Self-organizing systems*. Pergamon, New York, 1960.

[Pask, 1981] G. Pask. Organisational closure of potentially conscious systems. In M. Zeleny (ed.), *Autopoiesis: A theory of living organisation*. North Holland, New York, 1981.

[Pattee, 1973] H. H. Pattee (ed.). *Hierarchy theory: the challenge of complex systems*. New York: Geo. Braziller, 1973.

[Pattee, 1993] H. H. Pattee. The limitations of formal models of measurement, control, and cognition. *Applied Math. and Computation*, 56, 111-130, 1993.

[Pattee, 1995] H. H. Pattee. Evolving self-reference: matter, symbols, and semantic closure. *Communication and Cognition — Artificial Intelligence*, 12 (1-2), 9-28, 1995.

[Pattee, 1996] H. H. Pattee. The problem of observables in models of biological organizations. In E. L. Khalil, K. E. Boulding (eds.), *Evolution, Order, and Complexity*. Routledge, London, 1996.

[Pattee, 2000] H. H. Pattee. Causation, control and the evolution of complexity. In [Andersen *et al.*, 2000, pp. 322-348].

[Pattee, 2007] H. H. Pattee. Laws, constraints and the modelling relation — history and interpretations. Chemistry & Biodiversity, 4, 2272-2278, 2007.

[Penfold *et al.*, 1993] H. B. Penfold, P. S. Champness, U. Kohlmorgen. Breeding neural nets using a genetic algorithm on a massively parallel computer. Research Report, Electrical Engineering, Universities of Karlsruhe and Newcastle, 1993.

[Pfeiffer and Bongard, 2007] R. Pfeifer and J. C. Bongard. *How the body shapes the way we think: a new view of intelligence*. MIT Press, Cambridge Mass., 2007.

[Piaget, 1971a] J. Piaget. *Insights and illusions of philosophy* (trans. W. Mays). Meridian, New York, 1971.

[Piaget, 1971b] J. Piaget. *Biology and knowledge: an essay on the relations between organic regulations and cognitive processes* (trans. B. Walsh). University of Chicago Press, Chicago, 1971.

[Piaget and Garcia, 1989] J. Piaget and R. Garcia. *Psychogenesis and the history of science* (trans. H. Feider). Columbia University Press, New York, 1989.

[Pines, 1985] D. Pines, ed. *Emerging Syntheses in Science, Proceedings of the founding workshops of the Santa Fe Institute*, Addison Wesley, Redwood California, 1985.

[Polanyi, 1968] M. Polanyi. Life's irreducible structure. *Science*, 160, 1308-1312, 1968.

[Poon and Andelman, 2006] W. C. K. Poon and D. Andelman. *Soft condensed matter physics in molecular and cell biology*. Taylor & Francis, London, 2006.

[Popper, 1972] K. Popper. *Conjectures and Refutations*. Routledge and Kegan Paul, London, 1972.

[Popper, 1979] K. Popper. *Objective knowledge: an evolutionary approach* (rev. ed.). Oxford University Press, Oxford, 1979.

[Popper, 1984] K. Popper. Evolutionary epistemology. In J. W. Pollard (ed.), *Evolutionary theory: paths into the future.* Wiley, Chichester, 1984.

[Popper, 1987] K. Popper. Natural selection and the emergence of mind. In G. Radnitzky, W. W. Bartley III (eds.), *Evolutionary epistemology, theory of rationality and sociology of knowledge.* Open Court, La Salle Ill., 1987.

[Port and van Gelder, 1995] R. Port, T. J. van Gelder. *Mind as Motion: Dynamics, Behavior, and Cognition.* MIT Press, Cambridge Mass., 1995.

[PourEl and Richards, 1989] M. PourEl and J. I. Richards. *Computability and Analysis in Physics*,Perspectives in Mathematical Logic. SpringerVerlag, Berlin, 1989.

[Prigogine, 1967] I. Prigogine. Thermodynamics of Irreversible Processes, 3^{rd} edn. Wiley/Interscience, New York, 1967.

[Prigogine, 1980] I. Prigogine. *From Being to Becoming*, W. H. Freeman, San Francisco, 1980.

[Prigogine, 2003] I. Prigogine. Chemical Kinetics and Dynamics. Annals of the New York Academy of Science, 988, 128132, 2003.

[Prigogine and Stengers, 1984] I. Prigogine and I. Stengers. *Order out of Chaos*, Shambhala, Boulder, Colorado, 1984.

[Prokopenko, 2008] M. Prokopenko, ed. *Advances in Applied Selforganizing Systems.* Advanced Information and Knowledge Processing Series, Springer, Berlin, 2008.

[Putnam, 1975] *Mind, Language and Reality.* Philosophical Papers, vol. 2. Cambridge University Press, Cambridge, 1975.

[Raff and Kaufman, 1983] R. Raff, T. Kaufman. *Embryos, Genes, and Evolution.* Macmillan, New York, 1983.

[Raff, 1996] R. A. Raff. *The shape of life: genes, development and the evolution of animal form.* University of Chicago Press, Chicago, 1996.

[Rapoport, 1966] A. Rapoport. *Two person game theory*, University of Michigan Press, Ann Arbor, 1966.

[Rapoport, 1970] A. Rapoport. *Fights, games and debates.* University of Michigan Press, Ann Arbor, 1970.

[Rashevsky, 1964] N. Rashevsky. *Some Medical Aspects of Mathematical Biology.* Thomas, Springfield, IL., 1964.

[Ravasz *et al.*, 2002] E. Ravasz, A.L. Somera, D.A. Mongru, Z.N. Oltavai, and A.L. Barabási. Hierarchical Organization of Modularity in Metabolic Networks. *Science*, 297, 15511555, 2002.

[Rescher, 1977] N. Rescher. *Methodological pragmatism.* Blackwell, London, 1977.

[Roos, 1927] C. F. Roos. Dynamical Economics. Proceedings of the National Academy of Science, 13, 145150, 1927.

[Rosen, 1985a] R. Rosen. Anticipatory *Systems.* Pergamon, New York, 1985.

[Rosen, 1985b] R. Rosen (ed.). *Theoretical biology and complexity.* Academic Press, Orlando, 1985.

[Rosen, 1991] R. Rosen. *Life Itself.* Columbia University Press, New York, 1991.

[Rosenblatt, 1958] F. Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386408, 1958.

[Ruiz-Mirazo *et al.*, 2004] K. Ruiz-Mirazo, J. Peretó, A. Moreno. A universal definition of life: autonomy and openended evolution. *Origins of Life and Evolution of the Biosphere*, 34(3), 323346, 2004.

[Ruiz-Mirazo and Moreno, 2004] K. Ruiz-Mirazo and A. Moreno. Basic autonomy as a fundamental step in the synthesis of life. *Artificial Life*, 10(3), 235-259, 2004.

[Ruiz-Mirazo and Moreno, 2006] K. Ruiz-Mirazo and A. Moreno. On the origins of information and its relevance for biological complexity. Biological Theory, 1(3), 227229, 2006.

[Rumelhart and McClelland, 1986] D. E. Rumelhart and J. L. McClelland (eds.). *Parallel Distributed Processing*, Vol. 1. MIT Press, Cambridge, Mass., 1986.

[Ryan, 2007] A. Ryan. Emergence is coupled to scope, not level. *Complexity*, 13(2), 67-77, 2007. Pre-print available at `http://arXiv.org/abs/nlin/0609011`.

[Salthe, 1985] S. N. Salthe, *Evolving hierarchical systems: their structure and representation.* Columbia University Press, New York, 1985.

[Salthe, 1989] S. N. Salthe. Self-organisation of/in hierarchically structured systems. *Systems Research* 6, 199-208, 1989.

[Sapp, 2003] J. Sapp. *Genesis: the evolution of biology.* Oxford University Press, Oxford, 2003.

[Sawyer, 2005] R. K. Sawyer. *Social emergence: societies as complex systems*, Cambridge University Press, Cambridge, 2005.

[Scheffer *et al.*, 2002] M. Scheffer, F. Westley, W. Brock, M. Holmgren. Dynamic interaction of societies and ecosystems — linking theories from ecology, economy and sociology. In L. Gunderson, C. Holling (eds.), *Panarchy: Understanding Transformations in Human and Natural Systems.* Island Press, Washington, 2002.

[Schreckenberg and Wolf, 1998] M. Schreckenberg and D. E. Wolf. *Traffic and Granular Flow.* Springer: Singapore, 1998.

[Scott, 2004] A. Scott. *Encyclopedia of nonlinear science.* Routledge, London, 2004.

[Shalizi, 2001] C. R. Shalizi. Causal architecture, complexity and self-organisation in time series and cellular automata. Phd thesis, Physics Department, University of Wisconsin, Madison, 2001. Available at `http://www.cscs.umich.edu/~crshalizi/thesis`.

[Shalizi, 2004] C. R. Shalizi, K. L. Shalizi, R. Haslinger. Quantifying selforganization with optimal predictors. *Physical Review Letters*, 93(11), 118701, 2004. Available at `http://arXiv.org/abs/nlin/0409024`.

[Shalizi, 2006] C. R. Shalizi. Methods and techniques of complex systems science: an overview. In T. Deisboeck, J. Kresh (eds.), *Complex systems science in bio-medicine.* Springer. New York, 2006. Also available at `http://arXiv.org/abs/nlin/0307015`

[Shapin and Schaffer, 1985] S. Shapin and S. Schaffer. *Leviathan and the AirPump: Hobbes, Boyle, and the Experimental Life.* Princeton University Press, Princeton, 1985.

[Shi, 2001] Y. Shi. *Economics of scientific knowledge: A rational choice neo-insitutionalist theory of science.* Edward Elgar, London, 2001. PhD thesis, 2000, Philosophy Department, University of Newcastle, Australia.

[Shipman, 1992] J. Shipman. Aspects of computability in physics. *Proceedings, Workshop on Physics and Computation*, IEEE, 299-314, 1992.

[Simon, 1947] H. Simon. *Administrative behavior*: A Study of DecisionMaking Processes in Administrative Organizations. Free Press, New York, 1947.

[Simon, 1969] H. Simon. *The sciences of the artificial.* MIT Press, Cambridge, Mass., 1969.

[Simon, 1997] H. A. Simon. *Models of bounded rationality, vol.3.* MIT Press, Cambridge, Mass., 1997.

[Skewes and Hooker, 2009] J. Skewes and C. A. Hooker. Bio-agency and the problem of action. *Biology and Philosophy*, 24(3), 283-300, 2009.

[Smithers, 1995] T. Smithers. Are autonomous agents information processing systems? In L. Steels, R. A. Brooks (eds.), *The Artificial Life Route to 'Artificial Intelligence': Building Situated Embodied Agents.* Erlbaum, Hillsdale NJ, 1995.

[Smolensky, 1988] P. Smolensky. On the proper treatment of connectionism. *Behavioral and Brain Science*, 11, 1-74, 1988.

[Solé *et al.*, 2002] R. Solé, I. Salazar, J. Garcia-Fernandez. Common pattern formation, modularity and phase transitions in a gene network model of morphogenesis. *Physica A*, 305, 640-647, 2002.

[Solé and Bascompte, 2006] R. Solé and J. Bascompte. *Self-organization in Complex Ecosystems.* Princeton University Press, Princeton, 2006.

[Sperry, 1969] R. W. Sperry. A modified concept of consciousness. *Psychology Review*, 76, 532-536, 1969.

[Sperry, 1986] R. W. Sperry. Macro versus microdetermination. *Philosophy of Science*, 53, 265-275, 1986.

[Sperry, 1987] R. W. Sperry. Consciousness and causality. In R. L. Gregory (ed.), *The Oxford Companion to the Mind.*S Oxford University Press, Oxford, 164166, 1987.

[Sterelny, 2007] K. Sterelny Social intelligence, human intelligence and niche construction. *Proceedings of the Royal Society*, London (Series B), 362(1480), 719-730, 2007.

[Stewart and Golubitsky, 1992] I. Stewart and M. Golubitsky. *Fearful Symmetry. Is God a Geometer?* Blackwell, Oxford, 1992.

[Stewart and Mossio, 2007] J. Stewart and M. Mossio. Is "Life" Computable? On the Simulation of Closure under Efficient Causation, Proceedings of the 4th International Conference on Enactive Interfaces, Grenoble, France, 271-276, 2007.

[Strogatz, 2001] S. Strogatz. Exploring complex networks. Nature, 410, 268-277, 2001.

[Sullivan, 2009] J. A. Sullivan. The multiplicity of experimental protocols: a challenge to reductionist and nonreductionist models of neuroscience. Synthese, 167, 511-539, 2009.

[Szathmáry, 2005] E. Szathmáry. Life: in search of the simplest cell. *Nature*, 433(3 Feb.), 469470, 2005.

[Tan *et al.*, 2006] P-N. Tan, M. Steinbach, V. Kumar. *Introduction to data mining.* Addison-Wesley, New York, 2006.
[Toulmin, 1972] S. Toulmin. *Human understanding.* Princeton University Press, Princeton NJ, 1972.
[Truhlar, 2008] D. G. Truhlar. Editorial: Molecular Modeling of Complex Chemical Systems. Journal of the American Chemical Society,130(50), 16824-16827, 2008.
[Ulanowicz, 1997] R. E. Ulanowicz. *Ecology: the Ascendant Perspective.* Columbia University Press, New York, 1997.
[Ulanowicz, 2003] R. E. Ulanowicz. On the Ordinality of Causes in Complex Autocatalytic Systems. *Annals of the New York Academy of Science*, 988, 154157, 2003.
[Van Fraassen, 1989] B. Van Fraassen. *Laws and Symmetry.* Oxford University Press, Oxford, 1989.
[van Zijderveld, 2007] E. J. A. van Zijderveld. MARVEL: Principles of a method for semi-qualitative system behaviour and policy analysis. Proceedings, 25th International System Dynamics Conference, Boston, 2007. Available at `http://systemdynamics.org/conferences/2007/proceed/papers/VAN{%}20Z213.pdf`
[Vapnik, 1995] V. Vapnik. *The nature of statistical learning theory.* Springer, New York, 1995.
[Varela, 1979] F. Varela. *Principles of biological autonomy.* Elsevier, New York, 1979.
[Varela *et al.*, 1974] F. Varela, H. R. Maturana, R. B. Uribe. Autopoiesis: the organization of living systems, its characterization and a model. *Biosystems*, 5, 187-196, 1974.
[Vygotsky, 1986] L. S. Vygotsky. *Thought and language.* MIT Press, Cambridge Mass., 1986.
[Walter, 1950] W. G. Walter. An imitation of life. *Scientific American*, 182(5), 42-45, 1950.
[Walter, 1953] W. G. Walter. *The Living Brain.* Duckworth, London, 1953.
[Ward, 2002] L. M. Ward. *Dynamical Cognitive Science.* MIT Press, Cambridge MA, 2002
[Westerhoff and Palsson, 2004] H. Westerhoff and B. Palsson. The evolution of molecular biology into systems biology. *Nature Biotechnology*, 22, 1249-1252, 2004.
[Wiener, 948] N. Wiener. *Cybernetics: Or the Control and Communication in the Animal and the Machine.* Paris, France: Librairie Hermann & Cie, and MIT Press, Cambridge, MA, 1948.
[Wimsatt, 1999] W. C. Wimsatt. Genes, memes and cultural hereditary. *Biology and Philosophy*, 14, 279-310, 1999.
[Wimsatt, 2007] W. Wimsatt. *Re-engineering philosophy for limited beings.* Harvard University Press, Cambridge, Mass., 2007.
[Winfree, 1987] A. Winfree. *When Time Breaks Down.* Princeton University Press, Princeton, 1987.
[Woese, 2004] C. Woese. A new biology for a new century. *Microbiology and Molecular Biology Reviews*, 68(2), 173-186, 2004.
[Wolfram, 2002] S. Wolfram. *A new kind of science.* Wolfram Media, Champaign Ill., 2002. See also `http://www.wolframscience.com/thebook.html`
[Zhou and Chen, 2003] Z-H. Zhou and S-F. Chen. Evolving fault-tolerant neural networks. *Neural Computing & Applications*, 11(34), 156160, 2003.

SYSTEMS AND PROCESS METAPHYSICS

Mark H. Bickhard

Systems engage in various kinds of processes, so it might seem that a process metaphysics would be a natural approach to complex systems. Indeed it is, but systems are usually construed in terms of the interactions of parts that are not themselves modeled as systems or processes. The parts collectively form the stable base among which the system processes occur. Adopting a genuine process metaphysics, however, forces more fundamental changes in approaches to complex systems:

It overturns several standard conceptual and explanatory defaults, and
It enables posing and exploring new questions and explanatory frameworks.

A BRIEF HISTORY

Process conceptions of the world have had a presence in Western thought at least since Heraclitus, but have been dominated and overshadowed by substance and atomistic metaphysical frameworks at least since Empedocles and Democritus. In fact, Parmenides argued against Heraclitus that, far from everything being flux, change is not possible at all: in order for A to change into B, A would have to disappear into nothingness and B emerge out of nothingness. Nothingness does not exist, so change does not exist. The nothingness that Parmenides was denying perhaps has a contemporary parallel in the notion of the nothingness outside of the universe (vacuum is not nothingness in this sense). It is not clear that such a notion makes any sense, and this was roughly the Parmenidean point. Furthermore, for the ancient Greeks, to think about something or to refer to something was akin to pointing to it, and it is not possible to point at nothing. For a modern parallel to this, consider the difficulties that Russell or Fodor (and many others) have had accounting for representing nothing or something that is false [Hylton, 1990].

In any case, Parmenides' argument was taken very seriously, and both the substance and the atomistic metaphysical frameworks were proposed as responses. Empedocles' substances of earth, air, fire, and water were unchanging in themselves, thus satisfying the Parmenidean constraint, and Democritus' atoms were similarly unchanging wholes [Graham, 2006; Guthrie, 1965; Wright, 1997]. In both cases, apparent changes were accounted for in terms of changes in the mixtures and structural configurations of the underlying basic realities.

Plato and Aristotle also took these issues very seriously. Aristotle, in particular, developed a very sophisticated framework in which what he called earth, air, fire,

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

and water could transform into each other [Gill, 1989], but there remained an underlying unchanging, thus Parmenidean satisfying, substrate of prime matter.[1]

These are the traditions that have dominated for over two millennia, and in many respects, still do. There is, however, a historical move away from substance models toward process models: almost every science has had an initial phase in which its basic phenomena were conceptualized in terms of some kind of substance — in which the central issues were to determine what kind of substance — but has moved beyond that to a recognition of those phenomena as processes. This shift is manifest in, for example, understanding fire in terms of phlogiston to understanding fire in terms of combustion, heat in terms of random kinetic motion rather than the substance caloric, life in terms of certain kinds of far from thermodynamic equilibrium processes rather than in terms of vital fluid, and so on. Sciences of the mind, arguably, have not yet made this transition [Bickhard, 2004] — I will have more to say about this below.

As mentioned, however, a thorough shift to a process metaphysical framework involves some deep and ramified conceptual changes in explanatory defaults and frameworks for questions. In this chapter, I will be outlining several of these.

1 THREE CONCEPTUAL SHIFTS

I begin with three basic consequences of substance frameworks, and their reversal in process approaches.

1.1 From stasis to change

The default for substances and Democritean "atoms" is stability. Change requires explanation, and there are no self-movers. This is reversed in a process view, with change always occurring, and it is the stabilities of organizations or patterns of process, if such should occur, that require explanation.

There are two basic categories of process stability. The first is what might be called energy well stabilities. These are process organizations that will remain stable so long as no above threshold energy impinges on them. Contemporary atoms would be a canonical example: they are constituted as organizations of process that can remain stable for cosmological time periods.

The second category of process stability is that of process organizations that are far from thermodynamic equilibrium. Unlike energy well stabilities, these require ongoing maintenance of their far from equilibrium conditions. Otherwise, they go to equilibrium and cease to exist.

Also in contrast to energy well stabilities, far from equilibrium stabilities cannot be isolated for significant periods of time. If an energy well stability is isolated, it goes to internal thermodynamic equilibrium and retains its stability. If a far from

[1]There are contemporary interpretations of Aristotle that do not attribute Prime Matter to him, but the assumption of an underlying unchanging substrate for change is maintained (e.g., [Gill, 1989, 2005]).

equilibrium process organization is isolated, it goes to equilibrium and ceases to exist.

Far from equilibrium processes can exhibit self-organization, in which pattern emerges as an intrinsic result of underlying processes. Such resultant self-organized patterns can be of fundamental importance in, for example, the self-organization of tissue differentiation in embryos.

Some systems self-organize in ways that result in making contributions to their own far from equilibrium stability — they contribute to their own maintenance, hence are (partly) *self maintaining*. A canonical example would be a candle flame: the flame organizes (and is constituted by) a flow of air that brings in fresh oxygen and gets rid of waste. The organization is driven, in standard conditions, by the maintenance of above combustion threshold temperature, which melts wax that can then percolate up the wick, and vaporizes wax in the wick which can then burn in the incoming air.

Still more complex systems can vary their activities in accordance with changes in their environment, or in their relationships with their environment, so that the condition of being self-maintaining is itself maintained: they are in this sense *recursively self-maintenant*. A canonical example would be a bacterium that can swim if it is headed up a sugar gradient and tumble if it is oriented down the gradient. The two activities make self maintenant contributions, but in differing circumstances.

Recursive self maintenance requires some sort of infrastructure for detecting relationships with the environment and for switching among available activities. Such infrastructure may itself be of energy well stability form, but in most living examples, it too is necessarily open system far from equilibrium process. It functions as infrastructure in virtue of cycling on slower time scales than processes such as detection and switching.

In complex living systems, the maintenance of infrastructure (metabolism) and the interactions with the environment that support overall self maintenance constitute *autonomy*. Autonomy, in this sense, is a graded phenomenon, with self maintenance and recursive self maintenance constituting the broadest and potentially simplest kinds. Autonomy is the property of being able to exploit the environment in the service of self maintenance — in the service of the stability of the far from equilibrium process organization [Bickhard, 2004; Christensen and Bickhard, 2002; Christensen and Hooker, 2000].

1.2 From unalterability to emergence

Empedoclean earth, air, fire, and water not only cannot change into one another, it is not possible to generate an emergent fifth substance. Reasonably so, since emergence was one of the phenomena that such a metaphysical framework was introduced in order to preclude.

Hume

A manifestation of this point that has been of fundamental historical importance is Hume's argument against the possibility of deriving norms from facts [Hume, 1978]. Hume argues that, if the premises of a valid reasoning contain only factual terms, then the conclusion can only involve factual terms. Therefore, no normative conclusions can validly be derived from factual premises.

The argument, as standardly rendered, is that any terms in the conclusion that are not already in the premises must have been introduced via definition, and the definitions must be in terms either of what is available in the premises or what has been previously defined in the course of the reasoning. In either case, any new terms in the conclusion can always be back-translated through their definitions, substituting the defining phrases or clauses for the defined terms, until all such new terms are eliminated in favor of those in the premises.

Since the premises, by assumption, contain only factual terms, via such a procedure any valid conclusion can also be stated in those same factual terms. Therefore, only factual conclusions can be validly derived from factual premises.

But Hume's argument depends on the assumption that definitions always permit back-translation, and this is false. In particular, implicit definition does not permit back-translation. Implicit definition has been on the scene for roughly a century, most forcefully introduced by Hilbert around the end of the 19^{th} century [Hilbert, 1971; Otero, 1970]. There has been a tendency to minimize or ignore it for various reasons, one of which is Beth's theorem which holds implicit definition and explicit definition to be of equal power [Doyle, 1985]. But Beth's theorem provides only an extensional equivalence of the two kinds of definition, and even that only in certain combinations of kinds of logic and classes of models [Kolaitis, 1990]. In general, implicit definition is often more powerful than explicit definition. For current purposes, it cannot be ignored [Bickhard, 2009a; in preparation; Hale and Wright, 2000; Quine, 1966; Shapiro, 1997; 2005] and it does not permit back-translation, so Hume's argument is unsound.

Note that Hume's argument, in its general form, precludes anything other than re-arrangements of terms already in the premises. In its general form, it precludes any kind of emergence, not just normativity. It is a manifestation of the restriction to combinatorics of a substance or atomistic metaphysical framework. But, because the argument is in fact invalid, this barrier to the metaphysical possibility of emergence is removed.

Jaegwon Kim

In modern form, this metaphysical block of emergence is perhaps most strongly argued for by Jaegwon Kim's argument that any higher level organization may well manifest its own particular properties of causal regularities, but that these will ultimately always and necessarily be causally epiphenomenal relative to the basic particles that are engaged in those processes. Particles are the only legitimate potential loci of causality; organization is the framework, the stage setting, in which the particles do their causal dance. Any manifestation of higher level organization is merely the working out of the particle interactions within that organization [Kim, 1991; 1998].

But, from the perspective of a process metaphysics, everything is process, and process is inherently organized. Furthermore, process has whatever causal powers that it does in part in virtue of its organization. In a process view, organization cannot be delegitimated as a potential locus of causality without eliminating causality from the universe.

But, if organization is a legitimate potential locus of non-epiphenomenal causality, then it is at least metaphysically possible that higher level organization, including that of brains, can exhibit interesting, important, and non-epiphenomenal emergences, such as, perhaps, various kinds of mental phenomena [Beckermann, Flohr, Kim, 1992; Bickhard, 2000; 2004; 2009a; in preparation; Teller, 1992].

In any case, the fundamental block against emergence that is inherent in substance frameworks is eliminated.

1.3 From a split metaphysics to potential integration

Positing a metaphysical realm of substances or atoms induces a fundamental split in the overall metaphysics of the world. In particular, the realm of substances or atoms is a realm that might be held to involve fact, cause, and other physicalistic properties and phenomena, but it excludes such phenomena as normativity, intentionality, and modality into a second metaphysical realm. It induces a split metaphysics.

Given such a split, there are only three basic possibilities — though, of course, unbounded potential variations on the three. One could posit some version of the two realms as fundamental, and attempt to account for the world in terms of them. Aristotle's substance and form, Descartes' two kinds of substances, Kant's two realms, and the realm of fact and science distinct from that of modality and normativity of analytic philosophy are examples. Or, one could attempt to account for everything in terms of the "mental" side of the split, yielding idealisms, such as for Hegel, Green, and Bradley. Or, finally, one could attempt to account for everything in terms of the physical realm, such as Hobbes, Hume (on many interpretations), Quine, and much of contemporary science and philosophy.

It might be tempting to try to account for the whole range of phenomena in terms of some kind of emergence of normative and mental phenomena out of non-normative phenomena, but emergence is excluded by the metaphysical frameworks that induce the split in the first place.

Adopting a process metaphysics, however, reverses the exclusion of emergence, and opens the possibility that normativity, intentionality, and other phenomena might be modeled as natural emergents in the world. This integrative program is, in fact, being pursued in contemporary work [Bickhard, 2004; 2009a; 2009b; in preparation].

2 SOME FURTHER CONSEQUENCES

The default of stasis, preclusion of emergence, and split into substance and normative realms are three of the most important consequences of the adoption of a substance or particle metaphysics — consequences that are undone by a shift to a process metaphysics. But they are not the only ones.

2.1 Barriers to further questioning

It makes no internal sense to ask why Empedoclean earth, air, fire, and water have the properties that they do, nor why they have the relationships among themselves, nor where they came from, and so on. They constitute a ground of metaphysics with which much can be done, but about which there is little that can be meaningfully questioned — at least from within that framework itself.

That has certainly not prevented such questioning, but the questions are necessarily of the metaphysical framework itself, not questions within that framework. This kind of barrier to further questioning is a further consequence that is reversed by the shift to a process framework. In general, it does make sense to ask of a process why it has the properties that it does or the relationships to other processes or where it came from. The possibility that the process in question is emergent from others by itself legitimates such questions as questions within the process metaphysical framework. Answers may or may not be discoverable, but there is no metaphysical barrier to asking the questions and seeking for answers.

2.2 Determinism and predictability

A process view lends itself naturally to consideration of chaotic dynamics, and, thus, to consideration of the differentiation between determinism and predictability that chaotic phenomena introduce: chaotic phenomena are fully deterministic, but cannot in principle be predicted into a far future, given any finite resolution of system state.

It should be noted, however, that this is a claim about the determinism and prediction of specific dynamic trajectories in full detail. Chaotic dynamics may well be predictable in more general senses, such as if the space of possible dynamic trajectories is organized into a few attractor basins, with, perhaps, chaotic attractors in those basins. Even in such a case, predictions about what the attractor possibilities are might be accurate.

2.3 Process physics

As mentioned, most sciences have made a historical shift from substance frameworks to process views of their subject matter. Sciences of mentality are delayed in this respect, possibly because the mental is the primary realm that encounters the split of normativity and intentionality from the rest of the world.

But the shift to process has also occurred in fundamental physics, so a shift to a metaphysics of process is consistent with and lends itself to consideration of this property of contemporary physics. In particular, according to quantum field theory, there are no particles. Everything is organizations of quantum field processes, and particle-like phenomena are results of the quantization of quantum field processes [Cao, 1999; Clifton, 1988; Halvorson and Clifton; 2002; Huggett, 2000; Kuhlmann, Lyre, Wayne, 2002; Weinberg, 1977]. This is akin to the quantization of the number of waves in a guitar string, and similarly gives no basis for assuming particles — there are no guitar sound particles, and no quantum field particles either. Everything is stable (or not so stable) organizations of processes.

We know, of course, that contemporary physics is incomplete and has to be wrong in crucial respects. But a return to a substance or particle framework is precluded by the empirical confirmation of multiple non-localities, and dynamic space-time and vacuum effects. Such phenomena are not consistent with the local independence and fixedness of particles and substances.

2.4 Thermodynamics

Thermodynamics does not require a process metaphysics, but a process framework puts thermodynamic considerations at the center of metaphysical issues. The distinction between energy well forms of stability and far from equilibrium forms of stability, for example, is an explicitly thermodynamic distinction. Elsewhere, it is argued that these distinctions underlie the basic emergences of normative biological function, intentionality, and other normative phenomena.

2.5 Historicity

'indexbiosphere Taking a process view permits new kinds of explorations of, for example, self-organization. If processes of self-organization are temporally extended, they may manifest powerful dependencies of later organization on earlier organization. That is, they may manifest powerful historicities.

The self-organization of the biosphere, maintained in a far from equilibrium condition for billions of years by the sun, with natural selection constituting the local processes by which that self-organization is driven, is a primary realm, though certainly not the only realm, of such historicity. Macro-evolution is constrained and enabled by multifarious such historistic dependencies, including various kinds of individuation and encapsulization, reproductive modularizations, ecosystem interdependencies, and so on [Bickhard, in preparation; Bickhard and Campbell, 2003].

3 CHALLENGES

A process metaphysics puts many traditional assumptions into question. It is not always clear that the questions any longer make any sense, and often certain that neither they nor possible answers can be understood in traditional ways. A central class of these have to do with issues of boundaries and individuation.

3.1 Boundaries and individuation

Boundaries

In conjunction with thermodynamic considerations, a process metaphysics overturns standard assumptions about the individuation of entities in terms of boundaries. For open, far from equilibrium systems in particular, it is not clear what form such questions or their answers should take.

For example, what are the boundaries of a whirlpool or a hurricane? Or a candle flame? The candle flame is an interesting first focus: here we find various kinds of phase changes, such as between the region that engages in combustion and the region that feeds and cleans up after that combustion. It might seem that this is similar to the phase change boundary that individuates a rock, but note that a rock also has a co-extensive boundary at which it can be isolated, and also a co-extensive boundary at which it can be pushed. A candle flame has no clear boundary at which it can be isolated, though a distant boundary might permit it to continue for a time, and it has no boundary at which it can be pushed at all.

In many cases, it makes good sense to ask why there are any boundaries at all, and how the ones that do appear can be explained. In general, boundaries are put into question within a process framework in the same sense in which stability is put into question.

In the biological realm, boundary and individuation questions take the forms of: Why are there cells? Organisms? Species? Ecosystems? What creates and what maintains such boundaries? What different kinds of boundaries are there? Do they always, necessarily, exist? Such questions tend strongly to evoke classic canonical examples in which the answers might appear to be clear, but single organisms of fully sexually reproducing and sexually isolated species are far from the whole story.

How many individuals are there, for example, in a field of crabgrass, in which some rooted clumps of grass are still connected to others via runners, and some have had their generating runners decay? How many individuals in a clone of birch trees? How many species of bacteria that engage in horizontal gene transfer? Or species that maintain stable hybrid zones? And so on? What counts as a second instance of an autocatalytic process, rather than just a further spread of the autocatalytic cycling into broader regions? Or how many fires are there when what appear to be two brush fires merge, or one fire splits? The point, simply, is that boundaries and individuations are not inherent to process, so any boundaries that do exist, or are at least posited, must be explained in terms of their natures, origins, and forms of maintenance.

Individuation

If not in terms of an underlying substance with clear and unique entity-boundaries, how can processes be individuated?[2] What remains of the notion of an entity?

Subject to the basic conservations, processes occur in regions of space-time.[3] Trajectories along which conservations are maintained constitutes a major class of (partially) individuated processes. Quantum and field phenomena render this only a partial identification in terms of point trajectories because the relevant conservation honoring processes may not be locatable at particular points. Instead, they may be spread out through regions which themselves have at best broad and non-distinct boundaries.

This class of process individuations illustrates some general principles of individuation: 1) individuation is in terms of types of patterns or organizations of processes that satisfy some (set of) criteria, and 2) instances of the relevant types are commonly differentiated in terms of location, though, as was discussed above, that criterion isn't always clear: it's essentially a boundary-dependent notion.

One generalization of conserved quantity individuations are patterns of process organization that tend to recur for reasons other than basic conservations. Unsupported rocks tend to roll down hillsides, and sufficient quantities of U-235 in close proximity tend to explode. Such processes, of course, honor conservations, but are more restricted. In terms of dynamic landscapes, conserved quantity processes have unbounded "attractor basins" — all processes honor them — while more restricted kinds of process will have varying regions of attraction of prior processes that will enter such patterns: processes that remove support to rocks on hillsides will lie in the roll-down-hill attractor region, and so on. Note that the sense in which such kinds of processes have attractor basins is not the same as for the usual usage of the term: in particular, once on the way in the "attracting" process, the ongoing processes may remain in that attracting type only so long as is involved in its reaching a natural completion — reaching the bottom of the hill, or dissipation of the explosion. The "attractor" in such cases may attract, but the dynamics do not necessarily remain in or on the attracting pattern for indefinite lengths of time.

In spite of their possible short temporal extent, process types with attractor basins nevertheless exemplify a special kind of stability or persistence: a recurrence persistence. *Instances* of these types may not have much temporal persistence, but the *types* of process patterns may persist in multiple recurrences of instances of the type. Recurrences of such types will be probable insofar as their attractor basins are large relative to the relevant space of possible process organizations. In the physical realm, electrons of high energy level will tend to emit photons and thereby fall to lower energy levels; in the biological realm, individual bacteria may or may not last long, but the types tend to persist via recurrence — in this case, via a historistic process of reproduction.

[2] For an approach toward a general process ontology, see [Seibt, 2004; 2009].

[3] I set aside here issues concerning, for example, the unboundedness of a quantum field, various quantum non-localities, and the possibility that space and time are in some sense emergents.

It may also be the case, of course, that instances of a process type may exhibit some form of persistence in themselves. In such cases, the attracting process trajectories exhibit some sort of closure, such that, once entered on such a trajectory, a process will tend to remain on it (or *at it*, in the case of a point attractor). A very important subclass of kinds of instance-persistences are those composed of intertwined trajectories of simpler instances, perhaps conservation instances, for which the topological knot formed by that intertwining is relatively persistent. This category is subdivided with respect to whether or not the "simpler" instances remain associated with the instance of the dynamic topological knot, or instead if the knot persists, but the constituent simpler instances change.

A first special subclass of these, then, are those topological knot patterns of process for which the process instances that make them up are stable not just as instances in themselves, but also in their *relationships to the knot.* That is, for which the constituting instances are stable *parts* of the overall process patterns. A major kind of this case are those for which the stability of the knot pattern is maintained by being in an energy well, as discussed earlier (e.g., quark and electron field processes in an atom; atoms in a molecule, etc.). These constitute the canonical examples of *cohesive* process patterns [Collier, 1988].[4]

On the other hand, there will also be relatively persistent topological knots of process that are stable qua knot, but for which the relationships with the process instances that constitute the knot are not stable. Far from thermodynamic equilibrium self-organized process patterns are canonical examples of this type. The constituent instances of atoms and molecules that make up such a process — e.g., a candle flame — are continually changing, and are *necessarily* continually changing.

A further interesting class of topological knot persistences are those in which the form of the dynamic topological knot or twist (pattern) will itself change over time, but it remains a descendent *as an instance* of some original persistent knot of processes. In general, this will occur when there is a spatio-temporal continuity from earlier instance to later instance. Here we find continuities through growth, development, evolution, and so on.

Spatio-temporal continuity, however, is not conceptually required so long as there is a dynamically real sense of dependent derivation. In such cases, the distinction between type of process pattern and instance of process pattern becomes unclear, or at least crossed: *whale* as a type has individual whales as instances, but *the whale* as an instance of a species (arguably) does not. In general, to the extent that the dependence derivational relation becomes increasingly informational, the distinction between instance and type becomes at best complex, and in some ways blurred [Griesemer, 2005]. What are the type-instance rela-

[4]Another subtype would be stabilities that result from mechanical connections that may be weaker than the energy well stability, and perhaps only exist in certain directions or forms, such as something hanging on a hook: the stability of the relationship is only with respect to the direction of potential falling, and is weak in an upward direction. In general, energy well stabilities do not necessarily have isotropic strengths: e.g., layers of graphite are much stronger within a layer than between layers.

tionships involved in a range of artifacts that are in part manufactured through identical or similar processes, and in part copied from one manufacturing firm to another, or partially copied within a firm from earlier patterns/instances, such as generations of automobile types? Primarily informational derivation tends to abstract dynamic patterns from constituted instances, and thereby collapse the pattern-instance distinction. It is because of this collapse that biological species are better understood as spatially distributed single instances of complex dynamic and changing processes — individuals — than as types of which single organisms are instances [Ghiselin, 1974; 1987; 1997; Hull, 1976].

Individuation and boundaries: Summary

Issues of individuation, thus, become varied and complex within a process metaphysics. Simple assumptions of bounded substances constituting entities simply do not apply, and certainly do not suffice, for the complexities actually encountered in the world. Note, in this regard, that very little about boundaries has been relevant to the discussion of individuation. We have individuation in terms of topological patterns that exhibit dynamically relevant properties, with persistence of various kinds being among the important such properties. Boundaries, should they exist, will be created within and by those topological patterns, and, when they do exist, can be of multifarious kinds and forms.[5]

A process metaphysics raises basic metaphysical issues about unity, individuation, and boundaries. They are (multiple kinds of) temporal phenomena of (some) processes — not inherent characteristics of what it is to exist.

3.2 Supervenience

A final topic that I will consider is that of supervenience. The intuition of supervenience is that higher level phenomena cannot differ unless their supporting lower level phenomena also differ. There may be something correct in this intuition, but a process metaphysics puts at least standard ways of construing supervenience into question too.

Most commonly, a supervenience base — that upon which some higher level phenomena are supposed to be supervenient — is defined in terms of the particles and their properties, and perhaps the relations among them, that are the mereological constituents of the supervenient system [Kim, 1991; 1998]. Within a particle framework, and so long as the canonical examples considered are energy well stabilities, this might appear to make sense.

But at least three considerations overturn such an approach. First, local versions of supervenience cannot handle relational phenomena — e.g., the longest pencil in the box may lose the status of being longest pencil even though nothing about the pencil itself changes. Just put a longer pencil into the box. Being the longest pencil in the box is not often of crucial importance, but other relational phenomena

[5] Including, for example, fractal boundaries.

are. Being in a far from equilibrium relation to the environment, for example, is a relational kind of property that similarly cannot be construed as being locally supervenient. And it is a property of fundamental importance to much of our worlds — including, not insignificantly, ourselves.

A second consideration is that far from equilibrium process organizations, such as a candle flame, require ongoing exchanges with that environment in order to maintain their far from equilibrium conditions. There is no fixed set of particles, even within a nominally particle view, that mereologically constitutes the flame.

A third, related consideration is the point made above about boundaries. Issues of boundary are not clear with respect to processes, and not all processes have clear boundaries of several differentiable sorts — and, if they do have two or more of them, they are not necessarily co-extensive. But, if boundaries are not clear, then what could constitute a supervenience base is also not clear.

Supervenience is an example of a contemporary notion that has been rendered in particle terms, and that cannot be simply translated into a process metaphysical framework [Bickhard, 2000; 2004]. More generally, a process framework puts many classical metaphysical assumptions into question.

4 CONCLUSION

A hurricane or a candle flame illustrates a number of changes forced by a shift to a process metaphysical framework. They are roughly constituted as a twist or knot in the topology of the flow of ongoing process. Note that the point here is not just that the hurricane or flame is *dependent on* such process organization, but that they are *constituted in* such process organization.

As such, they have no inherent boundaries, individuations, supervenience bases, and so on. They are not entities in any classical sense. Questions about such properties — their existence, emergence, nature, maintenance, etc. — cannot be taken for granted, as is (or appears to be) the case within a particle framework. Instead, questions about such phenomena become legitimate and important *scientific* questions, questions that are not well motivated by a substance or particle metaphysics. Such questions take on an especially central importance in realms of science that address inherently far from equilibrium phenomena, such as biology and studies of the brain and mind. These are the realms in which the limitations and failures of classical substance and particle presuppositions are most damaging.

Conversely, a process metaphysics maintains the historical trend in science toward process. It is consistent with contemporary foundational physics, and integrates thermodynamics in a central and natural way. It makes emergence a genuine metaphysical possibility, and, in particular, it renders normative emergence a class of phenomena that are scientifically addressable. It requires changes, such as the shift to a default of change rather than stasis, and it raises multiple questions about properties that have historically often been presupposed, such as individuations and boundaries. But it is arguably the only framework that offers a viable orientation for the scientific future.

ACKNOWLEDGEMENTS

Thanks to Cliff Hooker and Itay Shani for helpful comments on an earlier draft.

BIBLIOGRAPHY

[Beckermann *et al.*, 1992] A. Beckermann, H. Flohr, and J. Kim. *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*. Berlin, Walter de Gruyter, 1992.

[Bickhard, 2000] M. H. Bickhard. Emergence. In P. B. Andersen, C. Emmeche, N. O. Finnemann, P. V. Christiansen, eds., *Downward Causation*, pp. 322-348. Aarhus, Denmark: University of Aarhus Press, 2000.

[Bickhard, 2003] M. H. Bickhard. Variations in Variation and Selection: The Ubiquity of the Variation-and-Selective Retention Ratchet in Emergent Organizational Complexity, Part II: Quantum Field Theory. *Foundations of Science, 8*(3), 283-293, 2003.

[Bickhard, 2004] M. H. Bickhard. Process and Emergence: Normative Function and Representation. *Axiomathes — An International Journal in Ontology and Cognitive Systems, 14*, 135-169, 2004. Reprinted from: M. H. Bickhard. Process and Emergence: Normative Function and Representation. In J. Seibt, ed., *Process Theories: Crossdisciplinary Studies in Dynamic Categories.* (121-155). Dordrecht: Kluwer Academic, 2003.

[Bickhard, 2009a] M. H. Bickhard. The interactivist model. *Synthese* 166(3), 547-591, 2009a.

[Bickhard, 2009b] M. H. Bickhard. Interactivism. In J. Symons, P. Calvo, eds., *The Routledge Companion to Philosophy of Psychology*, pp. 346-359. London: Routledge, 2009.

[Bickhard, forthcoming] M. H. Bickhard. Interactive Knowing: The Metaphysics of Intentionality. In R. Poli, J. Seibt, and J. Symons, eds., *Theory and Applications of Ontology.* Dordrecht: Kluwer, forthcoming.

[Bickhard, in preparation] M. H. Bickhard. *The Whole Person: Toward a Naturalism of Persons — Contributions to an Ontological Psychology.* In preparation.

[Bickhard and Campbell, 2003] M. H. Bickhard and D. T. Campbell. Variations in Variation and Selection: The Ubiquity of the Variation-and-Selective Retention Ratchet in Emergent Organizational Complexity. *Foundations of Science, 8*(3), 215-282, 2003.

[Campbell, 1974] D. T. Campbell. Evolutionary Epistemology. In P. A. Schilpp, ed., *The Philosophy of Karl Popper*, pp. 413-463. LaSalle, IL: Open Court, 1974.

[Campbell, 1992] R. J. Campbell. *Truth and Historicity.* Oxford, Oxford University Press, 1992.

[Cao, 1999] T. Y. Cao. *Conceptual Foundations of Quantum Field Theory.* Cambridge: U. of Cambridge Press, 1999.

[Christensen and Bickhard, 2002] W. D. Christensen, M. H. Bickhard. The Process Dynamics of Normative Function. *Monist, 85*(1), 3-28, 2002.

[Christensen and Hooker, 2000] W. D. Christensen, C. A. Hooker. An Interactivist-Constructivist Approach to Intelligence: Self-Directed Anticipative Learning. *Philosophical Psychology, 13*(1), 5-45, 2000.

[Clifton, 1996] R. Clifton. *Perspectives on Quantum Reality.* Kluwer Academic, 1996.

[Collier, 1988] J. Collier. Supervenience and Reduction in Biological Hierarchies. In M. Matthen and B. Linsky, eds., *Philosophy and Biology: Supplementary Volume 14 of the Canadian Journal of Philosophy*, pp. 209-234. Calgary: University of Calgary Press, 1988.

[Doyle, 1985] J. Doyle. Circumscription and Implicit Definability. *Journal of Automated Reasoning, 1*, 391-405, 1985.

[Ghiselin, 1974] M. T. Ghiselin. A Radical Solution to the Species Problem. *Systematic Zoology, 23*, 536-544, 1974.

[Ghiselin, 1987] M. T. Ghiselin. Species Concepts, Individuality, and Objectivity. *Biology and Philosophy, 2*, 127-144, 1987.

[Ghiselin, 1997] M. T. Ghiselin. *Metaphysics and the Origin of Species.* SUNY, 1997.

[Gill, 1989] M-L. Gill. *Aristotle on Substance.* Princeton, Princeton University Press, 1989.

[Gill, 2005] M. L. Gill. Aristotle's Metaphysics Reconsidered. *Journal of the History of Philosophy, 43*(3), 223-251, 2005.

[Graham, 2006] D. W. Graham. *Explaining the Cosmos.* Princeton, NJ: Princeton University Press, 2006.

[Griesemer, 2005] J. R. Griesemer. The Informational Gene and the Substantial Body: On the Generalization of Evolutionary Theory by Abstraction. In M. R. Jones and N. Cartwright, eds., *Idealization XII: Correcting the Model, Idealization and Abstraction in the Sciences*, pp. 59–115. Poznan Studies in the Philosophy of the Sciences and Humanities: Vol. 86, Amsterdam: Rodopi, 2005.

[Guthrie, 1965] W. K. C. Guthrie. *A History of Greek Philosophy vol. II: The Presocratic Tradition from Parmenides to Democritus*. Cambridge: Cambridge University Press, 1965.

[Hale and Wright, 2000] B. Hale and C. Wright. Implicit Definition and the A Priori. In P. Boghossian and C. Peacocke, eds., *New Essays on the A Priori*, pp. 286-319. Oxford: Oxford University Press, 2000.

[Halvorson and Clifton, 2002] H. Halvorson, R. Clifton. No Place for Particles in Relativistic Quantum Theories? *Philosophy of Science, 69*(1), 1-28, 2002.

[Hilbert, 1971] D. Hilbert. *The Foundations of Geometry*. La Salle: Open Court, 1971.

[Hooker, 1995] C. A. Hooker. *Reason, Regulation and Realism: Toward a Naturalistic, Regulatory Systems Theory of Reason*. Albany, N.Y.: State University of New York Press, 1995.

[Huggett, 2000] N. Huggett. Philosophical Foundations of Quantum Field Theory. *The British Journal for the Philosophy of Science, 51*(supplement), 617-637, 2000.

[Hull, 1976] D. L. Hull. Are Species Really Individuals? *Systematic Zoology, 25*, 174-191, 1976.

[Hume, 1978] D. Hume. *A Treatise of Human Nature*. Index by L. A. Selby-Bigge; Notes by P. H. Nidditch. Oxford, Oxford University press, 1978.

[Hylton, 1990] P. Hylton. *Russell, Idealism, and the Emergence of Analytic Philosophy*. Oxford: Oxford University Press, 1990.

[Kim, 1991] J. Kim. Epiphenomenal and Supervenient Causation. In D. M. Rosenthal, ed., *The Nature of Mind*, pp. 257-265. Oxford University Press, 1991.

[Kim, 1998] J. Kim. 1998). *Mind in a Physical World*. MIT, 1998.

[Kolaitis, 1990] Ph. G. Kolaitis. Implicit Definability on Finite Structures and Unambiguous Computations. In *Proc. 5th IEEE LICS*, 168-180, 1990.

[Kuhlmann *et al.*, 2002] M. Kuhlmann, H. Lyre, and A. Wayne. *Ontological Aspects of Quantum Field Theory*. River Edge, NJ: World Scientific, 2002.

[Otero, 1970] M. H. Otero. Gergonne on Implicit Definition. *Philosophy and Phenomenological Research, 30*(4), 596-599, 1970.

[Quine, 1966] W. V. O. Quine. Implicit Definition Sustained. In W. V. O. Quine, ed., *The Ways of Paradox*, pp. 195-198. New York: Random House, 1966.

[Seibt, 2004] J. Seibt. *General Processes: A Study in Ontological Theory Construction*. Habilitations-Thesis, University of Kanstanz, Germany, 2004.

[Seibt, 2009] J. Seibt. Forms of Emergent Interaction in General Process Theory. *Synthese, 166*(3), 479-512, 2009.

[Shapiro, 1997] . Shapiro. *Philosophy of Mathematics: Structure and Ontology*. Oxford, 1997.

[Shapiro, 2005] S. Shapiro. *The Oxford Handbook of Philosophy of Mathematics and Logic*. Oxford: Oxford University Press, 2005.

[Teller, 1992] P. Teller. A Contemporary Look at Emergence. In A. Beckermann, H. Flohr, and J. Kim, eds., *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*, pp. 139-153. Berlin: Walter de Gruyter, 1992.

[Weinberg, 1977] S. Weinberg. The Search for Unity, Notes for a History of Quantum Field Theory. *Daedalus, 106*(4), 17-35, 1977.

[Wright, 1997] M. R. Wright. Empedocles. In C. C. W. Taylor, ed., *From the Beginning to Plato*, pp. 175-207. Routledge 1997.

METAPHYSICAL AND EPISTEMOLOGICAL ISSUES IN COMPLEX SYSTEMS

Robert C. Bishop

1 INTRODUCTION

There are multiple senses of complexity. For instance, something can be considered complex because it is complicated or intricate (think of an engine or a watch with many moving parts and interrelated systems). However, systems that have a complicated set of interacting parts may actually exhibit relatively simple behavior (this is the case for engines). In contrast, the notion of complexity of interest in "complexity studies" centers on systems whose behavior is nonlinear and typically exhibits self-organization and collective effects. Such behavior appears to be anything but simple (see sec. 2.5), yet it is the case that systems with relatively few interacting components can exhibit this kind of intricate behavior. For instance, phase changes in Ising models and other systems with large numbers of components are examples of systems exhibiting complex self-organizing behavior. But even three-particle billiard-ball-like systems can exhibit the requisite complex behavior. On the other hand, many n-body systems, where n is large, do not exhibit complexity (e.g., Brownian motion of molecules) because the interactions among the constituents are not of the right type. An important feature of the systems of interest in complexity studies is that the interactions of the system components be nonlinear. Still, characterizing the resulting behavior, and the complex systems which exhibit it, is one of the major challenges facing scientists studying complexity and calls for the development of new concepts and techniques (e.g., [Badii and Politi, 1997]).

A number of metaphysical and epistemological issues are raised by the investigation and behavior of complex systems. Before treating some of these issues, a characterization of nonlinear dynamics and complexity is given. Along with this background, some folklore about chaos and complexity will be discussed. Although some claim that chaos is ubiquitous and many take the signal feature of chaos to be exponential growth in uncertainty (parameterized by Lyapunov exponents, see sec. 2.4), these examples of folklore turn out to be misleading. They give rise to rather surprising further folklore that chaos and complexity spell the end of predictability and determinism. But when we see that Lyapunov exponents, at least in their global form, and measures for exponential divergence of trajectories only apply to infinitesimal quantities in the infinite time limit, this further folklore also

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

turns out to be misleading. Instead, the loss of linear superposition in nonlinear systems is one of the crucial features of complex systems. This latter feature is related to the fact that complex behavior is not limited to large multi-component systems, but can arise in fairly simple systems as well.

The impact of nonlinearity on predictability and determinism will be discussed including, briefly, the potential impact of quantum mechanics. Some have argued that chaos and complexity lead to radical revisions in our conception of determinism, namely that determinism is a layered concept (e.g., [Kellert, 1993]), but such arguments turn on misunderstandings of determinism and predictability and their subtle relations in the context of nonlinear dynamics. When the previously mentioned folklore is cleared away, the relationship among determinism, predictability and nonlinearity can be seen more clearly, but still contains some subtle features. Moreover, the lack of linear superposition in complex systems also has implications for confirmation, causation, reduction and emergence, and natural laws in nonlinear dynamics all of which raise important questions for the application of complex nonlinear models to actual-world problems.

2 NONLINEAR DYNAMICS: FOLKLORE AND SUBTLETIES

I will begin with a distinction that is immediately relevant to physical descriptions of states and properties known as the ontic/epistemic distinction tracing back at least to Erhard Scheibe [1964] and subsequently elaborated by others [Primas, 1990; 1994; Atmanspacher, 1994; 2002; d'Espagnat, 1994; Bishop, 2002). Roughly, ontic states and properties are features of physical systems as they are "when nobody is looking," whereas epistemic states and properties refer to features of physical systems as accessed empirically. An important special case of ontic states and properties are those that are deterministic and describable in terms of points in an appropriate state space (see secs. 2.3 and 3.2 below); whereas an important special case of epistemic states and properties are those that are describable in terms of probability distributions (or density operators) on some appropriate state space. The ontic/epistemic distinction helps eliminate of confusions which arise in the discussions of nonlinear dynamics and complexity as we will see.

2.1 Dynamical systems

Complexity and chaos are primarily understood as mathematical behaviors of *dynamical systems*. Dynamical systems are deterministic mathematical models, where time can be either a continuous or a discrete variable (a simple example would be the equation describing a pendulum swinging in a grandfather clock). Such models may be studied as purely mathematical objects or may be used to describe a target system (some kind of physical, ecological or financial system, say). Both qualitative and quantitative properties of such models are of interest to scientists.

The equations of a dynamical system are often referred to as dynamical or evolution equations describing the change in time of variables taken to adequately describe the target system. A complete specification of the initial state of such equations is referred to as the *initial conditions* for the model, while a characterization of the boundaries for the model domain are known as the *boundary conditions*. A simple example of a dynamical system would be the equations modeling a particular chemical reaction, where a set of equations relates the temperature, pressure, amounts of the various compounds and their reaction rates. The boundary condition might be that the container walls are maintained at a fixed temperature. The initial conditions would be the starting concentrations of the chemical compounds. The dynamical system would then be taken to describe the behavior of the chemical mixture over time.

2.2 Nonlinear dynamics and linear superposition

The dynamical systems of interest in complexity studies are *nonlinear*. A dynamical system is characterized as linear or nonlinear depending on the nature of the dynamical equations describing the target system. Consider a differential equation system $d\boldsymbol{x}/dt = \boldsymbol{F}\boldsymbol{x}$, where the set of variables $\boldsymbol{x} = x_1, x_2, ..., x_n$ might represent positions, momenta, chemical concentration or other key features of the target system. Suppose that $\boldsymbol{x}_1(t)$ and $\boldsymbol{x}_2(t)$ are solutions of our equation system. If the system of equations is linear, it is easy to show that $\boldsymbol{x}_3(t) = a\boldsymbol{x}_1(t) + b\boldsymbol{x}_2(t)$ is also a solution, where a and b are constants. This is known as the *principle of linear superposition*.

If the principle of linear superposition holds, then, roughly, a system behaves such that any multiplicative change in a variable, by a factor α say, implies a multiplicative or proportional change of its output by α. For example, if you start with your television at low volume and turn the volume control up one unit, the volume increases one unit. If you now turn the control up two units, the volume increases two units. These are examples of linear responses. In a nonlinear system, linear superposition fails and a system *need not* change proportionally to the change in a variable. If you turn your volume control up two units and the volume increases tenfold, this would be an example of a nonlinear response.

2.3 State space and the faithful model assumption

Dynamical systems involve a *state space*, an abstract mathematical space of points where each point represents a possible state of the system. An instantaneous state is taken to be characterized by the instantaneous values of the variables considered crucial for a complete description of the state. When the state of the system is fully characterized by position and momentum variables (often symbolized as q and p, respectively), the resulting space is often called a *phase space*. A model can be studied in state space by following its trajectory, which is a history of the model's behavior in terms of its state transitions from the initial state to some

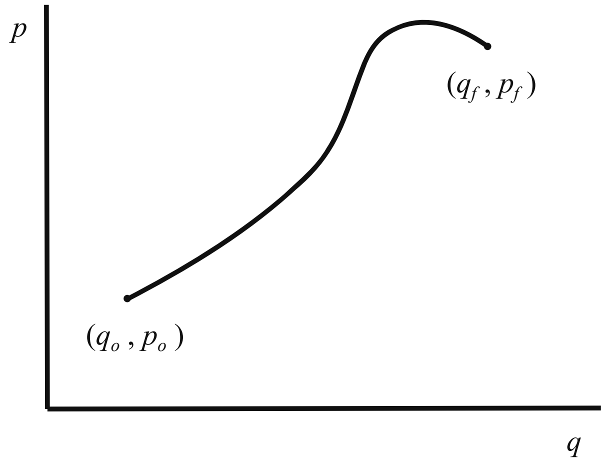


Figure 1.

chosen final state (Figure. 1). The evolution equations govern the path — the history of state transitions — of the system in state space.

There are some little noticed yet crucial assumptions being made in this account of dynamical systems and state spaces. Namely, that the actual state of a target system is accurately characterized by the values of the crucial state space variables and that a physical state corresponds via these values to a point in state space. These assumptions allow us to develop mathematical models for the evolution of these points in state space and to consider such models as representing the target systems of interest (perhaps through an isomorphism or some more complicated relation). In other words, we assume that our mathematical models are faithful representations of target systems and that the state spaces employed faithfully represent the space of actual possibilities of target systems. This package of assumptions is known as the *faithful model assumption* (e.g., [Bishop, 2005a; 2006]). In its idealized limit — *the perfect model scenario* [Judd and Smith, 2001]— it can license the (perhaps sloppy) slide between model talk and system talk (i.e., whatever is true of the model is also true of the target system and vice versa).

2.4 Sensitive dependence and Lyapunov exponents

One striking feature of chaos and complexity is their *sensitive dependence on initial conditions*: the property of a dynamical system to show possibly extremely different behavior with only the slightest of changes in initial conditions. A very popular measure of this sensitive dependence involves the explosive growth of the smallest uncertainties in the initial conditions of a nonlinear system. This explosive growth is often defined as an exponential parameterized by the largest *global*

Lyapunov exponent. These exponents arise naturally out of linear stability analysis of the trajectories of nonlinear evolution equations in a suitable state space. The infinite time limit plays an important role in this analysis, indicating that global Lyapunov exponents represent time-averaged quantities so that transient behavior has decayed. The existence of this limit is guaranteed by Oseledec's [1969] multiplicative ergodic theorem, which holds under mild conditions.

Imagine a small ball of points in state space around the initial conditions $\boldsymbol{x}(0)$. For any number $\delta > 0$ and for every slightly different initial condition $\boldsymbol{y}(0)$ in this small ball, exponential growth means the initial uncertainty, $|\boldsymbol{x}(0) - \boldsymbol{y}(0)| < \delta$, will evolve such that $|\boldsymbol{x}(t) - \boldsymbol{y}(t)| \approx |\boldsymbol{x}(0) - \boldsymbol{y}(0)|e^{\lambda t}$.[1] Here, λ is interpreted as the largest global Lyapunov exponent and is taken to represent the average rate of divergence of neighboring trajectories issuing forth from points very nearby $\boldsymbol{x}(0)$. If $\lambda > 0$, then the growth in uncertainty is exponential (if $\lambda < 0$, there is exponential convergence of neighboring trajectories). In general, such growth cannot go on forever. If the system is bounded in space and in momentum, there will be limits as to how far nearby trajectories can diverge from one another.

In most all philosophy literature and much physics literature, sensitive dependence as parameterized by global Lyapunov exponents is taken to be a distinguishing mark of chaotic dynamics. That is to say, exponential growth in the separation of neighboring trajectories characterized by λ is taken to be a property of a particular kind of dynamics that can only exist in nonlinear systems and models. However, there are problems with this folklore for defining sensitive dependence (and, hence, characterizing chaos and complexity using Lyapunov exponents).

One problem is that the definition of global Lyapunov exponents involves the infinite time limit. Strictly speaking, λ only characterizes growth in uncertainties as t increases without bounds, not for any finite time. At best, this would imply that sensitive dependence characterized by a global Lyapunov exponent can only hold for the large time limit. And this would further imply that chaotic phenomenon can only arise in this limit, contrary to what we take to be our best evidence. Furthermore, neither our models nor physical systems persist for infinite time, but an infinitely long time is required to verify the presumed exponential divergence of trajectories issuing from infinitesimally close points in state space.

The standard physicist's assumption that an infinite-time limit can be used to *effectively* represent some large but finite elapsed time will not do in the context of nonlinear dynamics either. When the finite-time Lyapunov exponents are calculated, they do not usually lead to on-average exponential growth as characterized by the global Lyapunov exponents (e.g., [Smith, *et al.*, 1999]). This is because the propagator — an operator evolving the uncertainty in some ball of points in state space forward in time — varies from point to point in state space for any finite times. The propagator is a function of the position $\boldsymbol{x}$ in state space and only approaches a constant in the infinite time limit. So local finite-time Lyapunov exponents vary from point to point in state space (whereas global Lyapunov ex-

[1] Technically, this kind of measure is taken to be valid in an appropriate state space for "almost all" points in the region around $\boldsymbol{x}(0)$ except a set of measure zero.

ponents do not). Therefore, trajectories diverge and converge from each other at various rates as they evolve in time, so that the uncertainty usually does not vary uniformly in the chaotic region of state space [Smith, *et al.*, 1999; Smith, 2000]. In contrast, global Lyapunov exponents are on-average global measures of uncertainty growth and imply that uncertainties grow uniformly (for $\lambda > 0$). Such uniform growth rarely occurs outside a few simple mathematical models (e.g., the baker's map).

For instance, the Lorenz, Moore-Spiegel, Rössler, Henon and Ikeda attractors all possess regions dominated by decreasing uncertainties in time, where uncertainties associated with different trajectories issuing forth from some small ball in state space shrink for the amount of time trajectories remain within such regions (e.g., [Smith, *et al.*, 1999, 2870-9; Ziehmann, *et al.*, 2000, 273-83]). What this means is that on-average exponential growth in uncertainties is not guaranteed for chaotic dynamics. Linear stability analysis indicates when nonlinearities can be expected to dominate the dynamics. And local finite-time Lyapunov exponents can indicate regions on an attractor where these nonlinearities will cause *all* uncertainties to decrease — cause trajectories to converge rather than diverge — so long as trajectories remain in those regions.

To summarize this first problem, the folklore that trajectories issuing forth from neighboring points in some ball in state space are guaranteed to diverge on-average exponentially in a chaotic region of state space is false in any sense other than for infinitesimal uncertainties in the infinite time limit.

The second problem with our folklore is that there simply is no implication that finite uncertainties will exhibit an on-average growth rate characterized by *any* Lyapunov exponents, local or global. As pointed out above, the linearized dynamics used to derive global Lyapunov exponents presupposes infinitesimal uncertainties. But when uncertainties are finite, linearized dynamics involving infinitesimals does not appropriately characterize the growth of finite uncertainties aside from telling us when nonlinearities should be expected to be important (this latter information is extremely useful however). Infinitesimal uncertainties can never become finite in finite time except through super-exponential growth. And even if infinitesimal uncertainties became finite after a finite time, that would presuppose the dynamics is unconfined; however, the interesting features of nonlinear dynamics usually take place in subregions of state space (e.g., on particular energy surfaces or in regions where attractors exist). Presupposing an unconfined dynamics, then, would be inconsistent with the features we are typically trying to capture.

One can ask whether the on-average exponential growth rate characterized by global Lyapunov exponents can ever be attributed legitimately to diverging trajectories if their separation is no longer infinitesimal. Examining simple models like the baker's map might seem to indicate yes. However, answering this question requires some care for more complicated systems like the Lorenz or Moore-Spiegel attractors. It can turn out to be the case that the rate of divergence in the finite separation between two nearby trajectories in a chaotic region changes character numerous times over the course of their winding around in state space, sometimes faster, sometimes slower than that calculated from global Lyapunov exponents, sometimes contracting, sometimes diverging [Smith, *et al.*, 1999; [Ziehmann, *et al.*, 2000]). In the long run, will some of these trajectories *effectively* diverge *as if* there was on-average exponential growth in uncertainties as characterized by global Lyapunov exponents? It is conjectured that the set of initial points in the state space exhibiting this behavior is a set of measure zero. This means that although there are an infinite number of points exhibiting this behavior, this set represents zero percent of the number of points composing the attractor. The details of the kinds of divergence (convergence) uncertainties undergo depend crucially on the detailed structure of the dynamics (i.e., it is determined point-by-point by local growth and convergence of finite uncertainties and not by any Lyapunov exponents).

In practice, however, all finite uncertainties saturate at the diameter of the attractor. The uncertainty reaches some maximum amount of spreading after a finite time and is not well quantified by measures derived from global Lyapunov exponents (e.g., [Lorenz, 1965]). So the folklore — that on-average exponential divergence of trajectories characterizes chaotic dynamics and complexity — is misleading for nonlinear systems. Therefore, drawing an inference from the presence of positive global Lyapunov exponents to the existence of on-average exponentially diverging trajectories for a dynamical system is shaky at best.

2.5 Complexity measures

Although several formal definitions of complexity have been proposed for characterizing random, chaotic and other forms of complex behavior, there is no consensus on which is the best definition, nor do these different definitions agree in picking out the same categories of behavior [Grassberger, 1989; Wackerbauer, *et al.*, 1994; Badii and Politi, 1997]. There is some evidence to suggest that different measures are useful for characterizing interesting behaviors of different systems for different purposes [Wackerbauer, *et al.*, 1994]. Perhaps this is not too surprising as it can be argued that complexity is just the kind of feature requiring a complex suite of tools and measures [Badii and Politi, 1997]. However, most of these complexity measures provide no intuitive access to the issues of emergence and causation at work in complex systems (some dynamical measures applicable in particular circumstances are exceptions). This is because most measures of complexity are formalized in terms of probabilities with no explicit reference to physical system variables (again, dynamical measures are an exception). Physical variables are implicitly involved in probabilistic measures because such variables are required to define the state space over which probability measures are defined.

Often it is more informative to characterize complex systems phenomenologically. Some of the most important features in these characterizations are:

- *Many-body systems.* Some systems exhibit complex behavior with as few as three constituents, while others require large numbers of constituents.

- *Broken symmetry.* Various kinds of symmetries, such as homogeneous arrangements in space, may exist before some parameter reaches a critical value, but not beyond.

- *Hierarchy.* There are levels or nested structures that may be distinguished, often requiring different descriptions at the different levels (e.g., large-scale motions in fluids vs. small-scale fluctuations).

- *Irreversibility.* Distinguishable hierarchies usually are indicators of or result from irreversible processes (e.g., diffusion, effusion).

- *Relations.* System constituents are coupled to each other via some kinds of relations, so are not mere aggregates like sand grain piles.

- *Situatedness.* The dynamics of the constituents usually depend upon the structures in which they are embedded as well as the environment and history of the system as a whole.

- *Integrity.* Systems display an organic unity of function which is absent if one of the constituents or internal structures is absent or if relations among the structures and constituents is broken.

- *Integration.* Various forms of structural/functional relations, such as feedback loops couple the components contributing crucially to maintaining system integrity.

- *Intricate behavior.* System behavior lies somewhere between simple order and total disorder such that it is difficult to describe and does not merely exhibit randomly produced structures.

- *Stability.* The organization and relational unity of the system is preserved under small perturbations and adaptive under moderate changes in its environment.

- *Observer relativity.* The complexity of systems depends on how we observe and describe them. Measures of and judgements about complexity are not independent of the observer and her choice of measurement apparatus [Grassberger, 1989; Crutchfield, 1994].

Such features of complex systems make the development of context-free measures of complexity unlikely (e.g., aside from observer relativity, the sense of order invoked in defining behavior as "intricate" depends on context). This can be illustrated by focusing on the nature of hierarchies in complex systems.

2.6 Hierarchies and sensitive dependence

The concept of *hierarchy* in the context of complex systems is of particular note. In some systems the hierarchy of physical forces and dynamical time scales (e.g., elementary particles, molecules, crystals) provide ontologically distinguishable levels of structure. In some cases the lower-level constituents may provide both necessary and sufficient conditions for the existence and behavior of the higher-level structures. In complex systems, however, levels of structure are often only epistemically distinguishable in terms of dynamical time scales. Furthermore, these levels are coupled to each other in such a way that at least some of the higher-level structures are not fully determined by, and even influence and constrain, the behavior of constituents in lower-level structures. That is, the lower-level constituents provide necessary but *not* sufficient conditions for the existence and behavior of some of the higher-level structures (cf. [Bishop, 2005b; 2008a; Bishop and Atmanspacher, 2006]). Moreover, the lower-level constituents may not even provide necessary and sufficient conditions for their own behavior if the higher-level structures and dynamics can constrain or otherwise influence the behavior of lower-level constituents (e.g., [Bishop, 2008a]). This latter kind of hierarchy is called a *control hierarchy* [Pattee, 1973, 75-9; Primas, 1983, 314-23]. Control hierarchies are distinguished from merely hierarchical structure like sand grain piles through the kinds of control they exert on lower-level structures and dynamics.

In complex systems, control hierarchies affect lower-level constituents primarily through constraints. The most important examples of constraints actively change the rate of reactions or other processes of constituents relative to the unconstrained situation (e.g., switches and catalysts). These constraints control lower-level constituents without removing all the latter's configurational degrees of freedom (in contrast to simple crystals, for instance). These top-down constraints may be external, due to the environment interacting with the system. Or such constraints may arise internally within the system due to the collective effects of its constituents or some other higher-level structural feature. Typically fundamental forces like gravity and electromagnetism are not explicitly identified with these latter internally generated constraints.

The notions of hierarchy and sensitive dependence allow us to formulate a more qualitative distinction between linear and nonlinear systems (though this characterization can also be made empirically precise — see [Busse, 1978], and [Cross and Hohenberg, 1993] for examples). Linear systems can be straightforwardly decomposed into and composed by subsystems (a consequence of the principle of linear superposition). For a concrete example of the principle of linear superposition, consider linear (harmonic) vibrations of a string which can be analyzed as a superposition of normal modes. These normal modes can be treated as uncoupled individual oscillators. The composition of the string's vibration out of these component vibrations is then analogous to aggregating these parts into a whole ("the whole is the sum of its parts"). The linear behavior of such systems in these cases is sometimes called *resultant* (in contrast with emergent).[2]

In nonlinear systems, by contrast, this straightforward idea of composition fails (a consequence of the failure of the principle of linear superposition). When the behaviors of the constituents of a system are highly coherent and correlated, the system cannot be treated even approximately as a collection of uncoupled individual parts ("the whole is *different* than the sum of its parts"). Rather, some particular global or nonlocal description[3] is required taking into account that individual constituents cannot be fully characterized without reference to larger-scale structures of the system. Rayleigh-Bénard convection, for instance, exhibits what is called *generalized rigidity* — the individual constituents are so correlated with all other constituents that no constituent of the system can be changed except by applying some change to the system as a whole. Such holistic behaviors are often referred to as *emergent* (in contrast with resultant).

The tight coupling between constituents in nonlinear systems is related to the nonseparability of the Hamiltonian. The latter is a function which corresponds to the total energy of the system and is related to the system's time evolution. Roughly, a Hamiltonian is separable if there exists a transformation carrying the Hamiltonian describing a system of N coupled constituents into N equations each describing the behavior of an individual system constituent. Otherwise, the Hamiltonian is nonseparable and the interactions within the system cannot be decomposed into interactions among only the individual components of the system.

In summary, linear systems can be decomposed into their constituent parts and the behavior of each component can be changed independently of the other components (which will then respond to the change introduced). Nonlinear systems often exhibit collective behavior where an individual system component cannot be isolated and its behavior changed independently of the rest of the system. Modifications of behaviors in a nonlinear system may have to take place at some higher hierarchical level or even at the level of the system as a whole.

2.7 Identity and individuation and a classical measurement problem

The interplay among hierarchical levels in nonlinear systems exhibiting complexity blur distinctions like part-whole, system-environment, constituent-level and so forth (e.g., cases where hierarchies are only distinguishable by differing time scales rather than by ontologically distinct features). The mathematical modeling of physical systems requires us to make distinctions between variables and parameters as well as between systems and their environments. However, when linear superposition is lost, systems can be exquisitely sensitive to the smallest of influences. A small change in the parameter of a model can result in significantly different behavior in its time evolution, making the difference between whether

[2]See [McLaughlin, 1982] for a discussion of the origin and history of the terms 'resultant' and 'emergent.'

[3]A nonlocal description in nonlinear dynamics denotes a description that necessarily must refer to wider system and environmental features in addition to local interactions of individual constituents with one another.

the system exhibits chaotic behavior or not, for instance. Parameters like the heat on a system's surface due to its environment may vary over time leading to wide variations in the time evolution of the system variables as well as temporal change in parameters. In such cases, the distinction between model variables and parameters tends to break down. Similarly, when a nonlinear system exhibits sensitive dependence, even the slightest change in the environment of a system can have a significant effect on the system's behavior. In such cases the distinction between system and environment breaks down. For instance, even the behavior of an electron at the 'edge' of the galaxy would affect a system of billiard balls undergoing continuous collisions [Crutchfield, 1994, p. 239], so the system/environment distinction becomes more a matter of pragmatically cutting up the 'system' and 'environment' in ways that are useful for our analysis.

All these subtleties raise questions about identity and individuation for complex systems. For instance, can a complex system somehow be identified as a distinct individual from its environment? Can various hierarchies of a complex system be individuated from each other? Asking these questions presupposes both that a distinct entity can be identified as well as individuated from other entities. Consider the so-called butterfly effect. Earth's weather is a complex system, but its potential sensitivity to the slightest changes of conditions leave its boundaries ill-defined if the flapping of a butterfly's wings in Argentina can cause a tornado in Texas three weeks later. Is the butterfly's flapping an internal or external source of wind and pressure disturbance? Turning towards space, is the magnetosphere surrounding the earth, which exhibits 'space weather' and shields the earth from lethal solar radiation, a distinct weather system or a qualitatively different extension of the Earth's weather system?

Traditional questions about identity and individuation revolve around numerical identity and the criteria for individuation and identity through time. It certainly seems plausible to consider butterflies, the Earth's weather and the Earth's magnetosphere (with its space weather) as numerically distinct systems (or as numerically distinct subsystems of a larger system). After all, according to Leibniz's principle of the identity of indiscernibles, these different "systems" do not share all their properties. On the other hand, systems are generally composed of subsystems that differ in properties, so given the lack of absolute boundaries between them, perhaps the best way to conceive of the butterfly-weather-magnetosphere system is as one very large complex system. As suggested by the phenomenological properties of complex systems (sec. 2.5), it is often the case that distinctions between parts and wholes, hierarchies and the like are pragmatic rather than absolute. There are further problems with identifying the boundary between the butterfly-weather-magnetosphere system and its solar system/galactic environment (e.g., electromagnetic fields and gravity extend over enormous distances in space).

Classical views of identity and individuation based on Leibniz's principle might be of some use in the pragmatic project of identifying complex systems and their components. However, these would only yield identification and individuation based on the kinds of questions scientific and other forms of inquiry raise and not a kind of objective ontology of distinct things (hence, many of our judgements about identity and individuation in nonlinear dynamics are epistemic rather than ontic). Whether the kinds of features described in secs. 2.5 and 2.6 imply there are no rigid designators, and hence complex systems represent a case of contingent identity and individuation [Kripke, 1980], is an open question.

Such features also raise questions about our epistemic access to complex systems. Obviously, some kind of cuts between observer and observed, and between system and environment have to be made. Along with this difficulty, there are clear epistemic difficulties confronting the measurement and study of complexity.

One epistemic difficulty is the mismatch between the accuracy or level of fine-grained access to the dynamics of a complex system and its underlying states and properties (i.e., the ontic/epistemic distinction). If a particular measurement apparatus only samples some even relatively fine-grained partition of the dynamical states of a complex system, the result will effectively be a mapping of (perhaps infinitely) many system states into a much smaller finite number of measurement apparatus states (e.g., [Crutchfield, 1994]). Such a mapping produces an apparent complexity — epistemic dynamical states in the measurement apparatus' projected space — that may not faithfully represent the complexity (or simplicity) of the system's actual dynamics — ontic states.

Another epistemic difficulty is that any measurement apparatus used to ascertain system states necessarily will introduce a small disturbance into complex systems that, in turn, will be amplified by sensitive dependence. No matter how fine-grained the measurement instrument, no matter how tiny the disturbance, this perturbation will produce an unpredictable influence on the future behavior of the system under study, resulting in limitations on our knowledge of a complex system's future. Along with the disturbance introduced to the complex system being measured, there is also a small uncertainty in the measurement apparatus itself. So the apparatus must also measure both itself and its disturbance perfectly for a full accounting of the exact state of the complex system being studied. This, in turn, leads to an infinite regress of measurements measuring measurements requiring the storage of the information of the entire universe's state within a subsystem of it, namely the measurement instrument. Because a system exhibiting sensitive dependence is involved and any measurement uncertainty will be amplified, an infinite amount of information stored in the measurement apparatus is required, which is physically impossible.

3 METAPHYSICAL AND EPISTEMOLOGICAL IMPLICATIONS

Complex systems, then, have rich implications for metaphysics and epistemology. Some of these implications for determinism, prediction, confirmation, causation, reductionism and emergence, and laws of nature will be surveyed here. I will begin with some epistemological implications that lead naturally into metaphysical issues.

3.1 Predictability and confirmation

So long as the uncertainty in ascertaining the initial state of a nonlinear system remains infinitesimal, there are no serious limitations on our ability to predict future states of such systems due to rapid growth in uncertainties.[4] In this sense, it is not the presence of a positive global Lyapunov exponent that signals predictability problems for nonlinear systems per se; rather it is the loss of linear superposition that leads to possible rapid growth in *finite* uncertainties in the measurement of initial states.[5] When the disturbance of the initial state due to the act of measurement is included, rapid growth in the total uncertainty in the initial state places impossibly severe constraints on the predictability of individual trajectories of systems or their components over various time scales.

The case of forecasting individual trajectories of complex systems receives the most attention in discussing the implications of chaos and complexity for predictability. For example, even under the perfect model scenario (sec 2.3), no matter how many observations of a system we make, we will always be faced with the problem that there will be a set of trajectories in the model state space that are indistinguishable from the actual trajectory of the target system [Judd and Smith, 2001]. Indeed, even for infinite past observations, we cannot eliminate the uncertainty in the epistemic states given some unknown ontic state of the target system. One important reason for this difficulty is traced back to the faithful model assumption (sec. 2.3). Suppose the nonlinear model state space (of a weather forecasting model, say) is a faithful representation of the possibilities lying in the physical space of the target system (Western European weather, say). No matter how fine-grained we make our model state space, it will still be the case that there are many different states of the target system (ontic states) that are mappable into the same state of the model state space (epistemic states). This means that there will always be many more target system states than there are model states.[6]

The constraints nonlinear systems place on the prediction of individual trajectories do not spell doom for predictability of systems exhibiting complex behavior, however. There are other statistical or probabilistic forms of prediction that can

---

[4] At least this is the case for chaos. In the more general context of nonlinear dynamics, such as the three-dimensional Navier-Stokes equations, it remains a grand prize challenge question as to whether there are solutions where infinitesimal uncertainties blowup on finite time scales. If answered in the affirmative, the loss of linear superposition would pose more potent challenges to prediction than the much ballyhooed chaos.

[5] More precisely, the loss of linear superposition is a necessary condition for rapid growth in uncertainties. Since nonlinear systems do not always exhibit rapid uncertainty growth, the detailed character of the actual parameter values and nonlinear dynamics has to be investigated for the conditions governing uncertainty dynamics.

[6] At least this is the case for any computational models since the equations have to be discretized. In those cases where we can develop a fully analytical model, in principle we could get an exact match between the number of possible model states and the number of target system states. Such analytical models are rare in complexity studies (many of the analytical models are toy models, like the baker's map, which, while illustrative of techniques, are misleading when it comes to metaphysical and ontological conclusions due to their simplicity).

be effectively applied to such systems (though these also have their limits; see [Smith, 2003]). For instance, one can apply various techniques for forming ensembles of initial states surrounding the assumed actual state of the target system and evolve these ensembles forward in time to forecast the behavior of the target system with specified measures for the uncertainties (e.g., [Judd and Smith, 2001] and references therein).

Moreover, faithfulness problems for our nonlinear mathematical models are problematic for standard approaches to confirming such models. We typically rely on the faithfulness of our mathematical models for our confirmation or verification of their efficacy in capturing behavior of target systems, but when the models are nonlinear and the target systems complex, faithfulness turns out to be inadequate for these standard confirmation practices.

Given a target system to be modeled (e.g., the weather over Western Europe), and invoking the faithful model assumption, there are two basic approaches to model confirmation discussed in the philosophical literature on modeling, focusing on individual trajectories and following a strategy known as piecemeal improvement.[7] These piecemeal strategies are also found in the work of scientists modeling actual-world systems and represent competing approaches vying for government funding (for an early discussion, see [Thompson, 1957]).

The first basic approach is to focus on successive refinements to the accuracy of the initial data fed into the model while keeping the model fixed (e.g., [Laymon, 1989, p. 359]). The intuition lying behind this approach is that if a model is faithful in reproducing the behavior of the target system to a high degree, refining the precision of the initial data fed to the model will lead to its behavior monotonically converging to the target system's behavior. This is to say that as the uncertainty in the initial data is reduced, a faithful model's behavior is expected to converge to the target system's behavior. Invoking the faithful model assumption, if one were to plot the trajectory of the target system in an appropriate state space, the model trajectory in the same state space would monotonically become more like the system trajectory as the data is refined. The second basic approach is to focus on successive refinements of the model while keeping the initial data fed into the model fixed (e.g., [Wimsatt, 1987]). The intuition here is that if a model is faithful in reproducing the behavior of the target system to a high degree, refining the model will produce an even better fit with the target system's behavior given good initial data. This is to say that if a model is faithful, successive model improvements will lead to its behavior monotonically converging to the target system's behavior. Again, invoking the faithful model assumption, if one were to plot the trajectory of the target system in an appropriate state space, the model trajectory in the same state space would monotonically become more like the system trajectory as the model is made more realistic.

What both of these basic approaches have in common is that piecemeal monotonic convergence of model behavior to target system behavior is a means of con-

[7] I will ignore bootstrapping approaches as they suffer similar problems, but only complicate the discussion (e.g., [Koperski, 1998]).

firming the model. By either improving the quality of the initial data or improving the quality of the model, the model in question reproduces the target system's behavior monotonically better and yields predictions of the future states of the target system that show monotonically less deviation with respect to the behavior of the target system. In this sense, monotonic convergence to the behavior of the target system is a key mark for whether the model is a good one. If monotonic convergence to the target system behavior is not found by pursuing either of these basic approaches, then the model is considered to be disconfirmed.

For linear models it is easy to see the intuitive appeal of such piecemeal strategies. After all, for linear systems of equations a small change in the magnitude of a variable is guaranteed to yield a proportional change in the output of the model. So by making piecemeal refinements to the initial data or to the linear model only proportional changes in model output are expected. If the linear model is faithful, then making small improvements "in the right direction" in either the quality of the initial data or the model itself can be tracked by improved model performance. The qualifier "in the right direction," drawing upon the faithful model assumption, means that the data quality really is increased or that the model really is more realistic (i.e., captures more features of the target system in an increasingly accurate way), and is signified by the model's monotonically improved performance with respect to the target system.[8]

However, both of these basic approaches to confirming models encounter serious difficulties when applied to nonlinear models exhibiting sensitive dependence [Koperski, 1998; Bishop, 2008b]. In the first instance, successive small refinements in the initial data fed into nonlinear models is not guaranteed to lead to any convergence between model behavior and target system behavior. Due to the loss of linear superposition, any small refinements in initial data can lead to non-proportional changes in model behavior rendering this piecemeal convergence strategy ineffective as a means for confirming the model. Even a refinement of the quality of the data "in the right direction" is not guaranteed to lead to the nonlinear model monotonically improving in capturing the target system's behavior. The small refinement in data quality may very well lead to the model behavior diverging away from the system's behavior.

In the second instance, keeping the data fixed but making successive refinements in nonlinear models is also not guaranteed to lead to any convergence between model behavior and target system behavior. Due to the loss of linear superposition, any small changes in the model, say by adding additional higher-order terms into the equations, can lead to non-proportional changes in model behavior for the same initial data, again rendering the convergence strategy ineffective as a means for confirming the model. Even if a small refinement to the model is made "in the right direction," there is no guarantee that the nonlinear model will monotonically improve in capturing the target system's behavior. The small refinement in the model may very well lead to the model behavior diverging away from the system's

[8]If one waits for long enough times piecemeal confirmation strategies will also fail for linear systems if there are imperfections in the data or models.

behavior.

So whereas for linear models piecemeal strategies might be expected to lead to better confirmed models (presuming the target system exhibits only stable linear behavior), no such expectation is justified for nonlinear models exhibiting sensitive dependence deployed in the characterization of nonlinear target systems. Even for a faithful nonlinear model, the smallest changes in either the initial data or the model itself may result in non-proportional changes in model output, an output that is not guaranteed to "move in the right direction" even if the small changes are made "in the right direction."[9]

Sticking with the individual trajectory approach, one might consider alternatives to these piecemeal confirmation strategies. One possibility is to turn to a Bayesian framework for confirmation, but similar problems arise here for nonlinear models exhibiting sensitive dependence. Given that there are no perfect models in the model class to which we would apply a Bayesian scheme and given the fact that imperfect models will fail to reproduce or predict target system behavior over time scales that may be long or short compared to our interests, there again is no guarantee that some kind of systematic improvement can be achieved for our nonlinear models.[10] Another approach is to seek trajectories issuing forth from the set of initial conditions in the model space — presumably the actual state of the target system has been mapped into this set — that areis consistent with all observations of the target system over the time period of interest. Given a faithful model, choose an initial condition consistent with the observational uncertainty that then yields a model trajectory passing within the observational uncertainty of the desired future observations. Models can then be judged as better or worse depending on the length of their shadowing times. Finding such trajectories that consistently shadow target system observations for longer and longer times under changes in either the initial data or the models themselves may be quite difficult, however. Furthermore, it is possible to construct models that can shadow any set of observations without those models having any physical correspondence to the target system (e.g., [Smith, 1997, p. 224-225]).

It seems that probabilistic models utilizing ensembles of trajectories or ensembles of probability distributions would allow for a clearer sense of confirmation. Yet, similar problems can crop up here as well. Ensemble forecasting models can give unique, but incorrect indications of the target system's future behavior or such models can give no unique indications of expected future behavior. And even when an ensemble model gives a relatively unique indication that tracks with the outcomes of a target system over a shorter time scale, its indications may diverge significantly from that time point forward.[11] Again, we face difficulties with formulating a systematic confirmation scheme.

[9]Of course, this lack of guarantee of monotonic improvement also raises questions about what "in the right direction" means, but I will ignore these difficulties here.

[10]Here I leave aside the problem that having no perfect model in our model class renders most Bayesian schemes ill-defined.

[11]See [Smith, 1997, pp. 236-237]) for an illuminating example of this in the context of weather forecasting.

While the above difficulties with determining when a nonlinear model is good causes problems for philosophers' desires to produce a systematic theory of confirmation for all models, this situation does not impede physicists and others from finding ways to improve their models and make determinations about how good their imperfect models are. However, it is also the case that these model builders do not follow some universal scheme for improving or confirming their models, but use a variety of techniques (e.g., [Smith, 1997]).

Last, but not least, there are ramifications here for the use of nonlinear models in the development and assessment of public policy. Policy formation and assessment often utilizes model forecasts and if the models and systems lying at the core of our policy deliberations are nonlinear, then policy assessment will be affected by the same lack of guarantee as model confirmation due to the loss of linear superposition. Suppose government officials are using a nonlinear model in the formulation of economic policies designed to keep GDP ever increasing while minimizing unemployment (among achieving other socio-economic goals). While it is true that there will be some uncertainty generated by running the model several times over slightly different data sets and parameter settings, assume policies taking these uncertainties into account to some degree can be fashioned. Once in place, the policies need assessment regarding their effectiveness and potential adverse effects, but such assessment will not be merely a function of looking at monthly or quarterly reports on GDP and employment data to see if targets are being met. The nonlinear economic model driving the policy decisions will need to be rerun to check if trends are indeed moving in the right direction and are on the right course with respect to the earlier forecasts. But, of course, the data fed into the model have now changed and there is no guarantee that the model will produce a forecast with this new data that fits well with the old forecasts used to craft the original policies. How, then, are policy makers to make reliable assessments of policies? The same problem that small changes in data or model in nonlinear contexts are not guaranteed to yield proportionate model outputs or monotonically improved model performance also plagues policy assessment using nonlinear models.

3.2 Determinism

Intuitively, one might think that if a system is deterministic, then it surely must be predictable, but the relationship between determinism and predictability is much too subtle to support this intuition [Bishop, 2003]. Predictability of systems has much to do with epistemic states while determinism has to do with ontic states. And while the characteristics of ontic states should have some implications for the character and behavior of epistemic states, it is difficult at best to draw any conclusions about the ontic states of a system based on our access to epistemic states. This is a fundamental reason why the often discussed unpredictability of chaotic and complex systems by itself does not undermine the determinism of the underlying ontic states of nonlinear systems in classical mechanics. So arguments like Karl Popper's [1950] to the effect that a breakdown in predictability implies a breakdown in determinism trade on confusing epistemic conclusions for ontic ones.

The trend since the Scientific Revolution has been to support the belief in metaphysical determinism by appealing to the determinism of theories and models from physics, though this strategy is not without its subtleties and surprises [Bishop, 2006]. A standard way of characterizing a mathematical model as deterministic is through the property *unique evolution*:

> Unique Evolution: A given state of a model is always followed (preceded) by the same history of state transitions.

The basic idea is that every time one returns the mathematical model to the same initial state (or any state in the history of state transitions), it will undergo the same history of transitions from state to state and likewise for the target system if the faithful model assumption holds. In other words the evolution will be unique given a specification of initial and boundary conditions.[12] For example the equations of motion for a frictionless pendulum will produce the same solution for the motion as long as the same initial velocity and initial position are chosen.

It is not uncommon to find arguments in the literature that purport to show that chaos and complexity tell against determinism. For example, in several books John Polkinghorne has pushed the claim that the kind of sensitive dependence exhibited by complex dynamical systems should lead us to view even the deterministically rigid world of classical mechanics as ontologically indeterministic. Here is an instance of this line of reasoning:

> The apparently deterministic proves to be intrinsically unpredictable. It is suggested that the natural interpretation of this exquisite sensitivity is to treat it, not merely as an epistemological barrier, but as an indication of the ontological openness of the world of complex dynamical systems [Polkinghorne, 1989, p. 43].

He attempts to make this line of thought plausible through demanding a close link between epistemology and ontology under a critical realist reading of the two.

If we remain at the level of dynamical systems — i.e., mathematics — then clearly there is a serious problem with this line of reasoning. Namely, the mathematical equations giving rise to the exquisite sensitivity and attendant predictability problems are deterministic in exactly the sense of unique evolution described above. So our ontic description in terms of these equations push in precisely the opposite direction that Polkinghorne pursues. Although it is true that apparent indeterminism can be generated if the state space one uses to analyze chaotic behavior is coarse-grained, this produces only an epistemic form of indeterminism. The underlying equations are still fully deterministic.

[12]Note that as formulated, unique evolution expresses state transitions in both directions (future and past). It can easily be recast to allow for unidirectional state transitions (future only or past only) if desired.

Instead, to raise questions about the determinism of real-world systems, one has to pursue the nature of these complex models and their implications as well as examine their presumed connection with target systems via the faithful model assumption. As pointed out in sec. 2 above, the mathematical modeling of actual-world systems requires us to make distinctions between variables and parameters as well as between systems and their boundaries. As we saw, these distinctions become problematic in the context of complex systems, where linear superposition is lost and such systems can be exquisitely sensitive to the smallest of influences. Such features raise questions about our epistemic access to systems and models in the investigation of complex systems, but they also raise questions about making sense of the supposed determinism of target systems. As an example, consider applying a deterministic mathematical model to forecasting the weather over Western Europe, where the identity and individuation of that system is questionable (sec. 2.7). What all do we have to include in this model to be able to make some reasonable pronouncement about whether Western Europe's weather is deterministic or not? Do we need only include a particular fluid mass over this particular continent, or over the earth's surface or that plus the stratosphere and magnetosphere, or .... And do we have to include every butterfly flapping its wings to get an identifiable target system?

There is a further problem in our application of deterministic models to actual-world complex systems and our belief that those systems are deterministic. Although the faithful model assumption appears fairly unproblematic in some simple contexts, if the system in question is nonlinear the faithful model assumption raises serious difficulties for inferring the determinism of the target system from the deterministic character of the model. For example, there is the problem that there will always be many more target system states than there are model states as described above (sec. 3.1).

More fundamentally, there is the problem of the mapping between the model and the target system itself. Even for a faithful model, we do not have a guarantee that the mapping between the model and the target system is one-to-one as we customarily assume. The mapping may actually be a many-to-one relation (e.g., several different nonlinear faithful models of the same target system as is the case with competing weather forecasting and climate prediction models) or a many-to-many relationship. For many classical mechanics problems — namely, where linear models or force functions are used in Newton's second law — the mapping between model and target system appears to be straightforwardly one-to-one with plausible empirical support. By contrast, in nonlinear contexts where one might be constructing a model from a data set generated by observing a system, there are potentially many nonlinear models that can be constructed, and each model may be as empirically adequate to the system behavior as any other. For the inference from the deterministic character of our mathematical model to the deterministic character of the target system to hold appears to require either a one-to-one relationship between a deterministic model and target system or that the entire model class in a many-to-one relation be deterministic.[13]

A different approach attempting to call the ontological determinism of the macroscopic world into question via complexity is the research on far-from-equilibrium systems by Ilya Prigogine and his Brussels-Austin Group [Antoniou and Prigogine, 1993; Petrosky and Prigogine, 1996; 1997; Prigogine, 1997].

Conventional physics describes physical systems using particle trajectories as a fundamental explanatory element of its models. If a system of particles is distributed uniformly in position in a region of space, the system is said to be in thermodynamic equilibrium (e.g. cream uniformly distributed throughout a cup of coffee). In contrast, a system is far-from-equilibrium (nonequilibrium) if the particles are arranged so that highly ordered structures appear (e.g. a cube of ice floating in tea). This means that the behavior of a model is derivable from the trajectories of the particles composing the model. The equations governing the motion of these particles are reversible with respect to time (they can be run backwards and forwards like a film). When there are too many particles involved to make these types of calculations feasible (as in gases or liquids), coarse-grained averaging procedures are used to develop a statistical picture of how the system behaves rather than focusing on the behavior of individual particles.

In contrast the Brussels-Austin approach views these systems in terms of models whose fundamental explanatory elements are distributions; that is to say, the arrangements of the particles are the fundamental explanatory elements and not the individual particles and trajectories.[14] The equations governing the behavior of these distributions are generally *irreversible* with respect to time. Moreover, focusing exclusively on distribution functions opens the possibility that macroscopic nonequilibrium models are irreducibly indeterministic, an indeterminism that has nothing to do with epistemic access to the system. If true, this would mean that probabilities are as much an ontologically fundamental element of the macroscopic world as they are of the microscopic.

One important insight of the Brussels-Austin Group shift away from trajectories to distributions as fundamental elements is that explanation also shifts from a local context (set of particle trajectories) to a global context (distribution of the entire set of particles). A system acting as a whole may produce collective effects that are not reducible to a summation of the trajectories and subelements composing the system [Petrosky and Prigogine, 1997; Bishop, 2004]. However, there are serious open questions about this approach, for instance what could be the physical source of such indeterminism[15] and what is the appropriate interpretation of the probabilistic distributions? Thus, the Brussels-Austin approach remains quite

[13]That this last requirement is nontrivial is exemplified in that different modeling teams will often submit proposals for the same project, where some propose deterministic models and others propose nondeterministic models.

[14]This does not imply, as some have erroneously thought (e.g., [Bricmont, 1995, 165-6]) that the Brussels-Austin Group argued there was no such thing as individual particle trajectories in such complex systems.

[15]One possibility is that this kind of indeterminism is ontologically emergent from the underlying dynamics (see sec 3.4 below).

speculative.

Another possible implication of complexity for determinism lies in the sensitivity of such systems to the smallest of disturbances. Some have argued that the sensitive dependence found in macroscopic chaotic systems opens such systems to the influence of quantum effects (e.g., [Hobbs, 1991; Kellert, 1993]). The line of thinking in these sensitive dependence arguments is that nonlinear chaotic systems whose initial states can be localized to a small patch of state space, because of quantum fluctuations, will have future states that can only be localized within a much larger patch of state space. For example, two isomorphic nonlinear systems of classical mechanics exhibiting sensitive dependence, whose initial states differ only in the quantum fluctuations within their initial conditions, will have future states that will differ significantly at later times. Since quantum mechanics sets a lower bound on the size of the patch of initial conditions, unique evolution must fail for such nonlinear chaotic systems.

This provocative line of argument is beset with a number of subtleties and difficulties, however. For example, there are difficult issues regarding the appropriate version of quantum mechanics (e.g., von Neumann, Bohmian or decoherence theories), the nature of quantum measurement theory (collapse vs. non-collapse theories), and the selection of the initial state characterizing the system that must be resolved before one can say clearly whether or not unique evolution is violated [Bishop, 2008c]. Just because quantum effects might influence macroscopic systems exhibiting sensitive dependence does not guarantee that determinism fails for such systems. Whether quantum interactions with such systems contribute *indeterministically* to the outcomes of these systems depends on the currently undecidable question of indeterminism in quantum mechanics, a resolution of the measurement problem, and a decision as to where to place the boundary between system and measurement.

Moreover, the possible constraints of nonlinear classical mechanics systems on the amplification of quantum effects must be considered on a case-by-case basis. For instance, damping due to friction can place constraints on how quickly amplification of quantum effects can take place before they are completely washed out [Bishop, 2008c]. And one has to investigate the local finite-time dynamics for each system because these may not yield any on-average growth in uncertainties (sec 2.4).

3.3 Causation

The analysis of determinism in complex systems is complicated by the fact that there are additional forms of causation arising in such systems that must be taken into account. Indeed, understanding whether a process is deterministic or not often depends upon understanding the underlying causal mechanism(s). There is no consensus account of what causes are; rather, there is a set of accounts — e.g. counterfactual, logical, probabilistic, process, regularity, structural (e.g., see [Sosa and Tooley, 1993]) — that each have strengths and weaknesses (and, perhaps, like definitions of complexity, have different applicability for different purposes). These accounts of causation require rethinking in the face of the richness of nonlinear dynamics. As indicated in section 2 above, chaos, complexity and self-organization are behaviors where complex wholes play important roles in constraining their parts. Such inter-level causation more generally has received little philosophical attention relative to bottom-up efficient modes of causation.

Immanuel Kant was one of the first to recognize the peculiarities of what we now call self-organization in living systems. He classifies such phenomena as *intrinsic physical ends* [Kant, 1980, p. 18] because they are in some sense both cause and effect of themselves. For instance, according to Kant, a tree "in the genus, now as effect, now as cause, continually generated from itself and likewise generating itself, . . . preserves itself generically" [Kant, 1980, p. 18]. An entity is an intrinsic physical end if "its parts, both as to their existence and form, are only possible by their relation to the whole" [Kant, 1980, p. 20]. Self-organizing systems, particularly organisms, are produced by, and in turn, produce the whole. Each part — such as they are distinguishable — exists in virtue of the activity of the other parts and the whole. Furthermore, each part exists for the sake of the other parts as well as the whole. "Only under those conditions and upon those terms can such a product be organized and self-organized *being*, and as such be called a physical end" [Kant, 1980, p. 22].

In Kant's view self-organization "has nothing analogous to any causality known to us" [Kant, 1980, p. 23] because the dominant concrete conception of causation available to him was that of external forces acting on systems generally through contact as exemplified in the model systems of Newtonian mechanics. Given his recognition that self-organizing systems required some kind of time-irreversible processes and that Newtonian dynamics was fully time-reversible, he relegated our lack of understanding how self-organization comes about to a limitation of reason [Kant, 1980, pp. 22-4]. Counterfactual, logical and regularity analyses of causation fare no better at penetrating this lack of understanding. While process and structural accounts each appear to have some pieces of the puzzle for understanding self-organization, process theories lack an adequate account of the structural constraints of wholes on parts, while structural theories lack an adequate account of processes.

Causation in complex systems has been given very little sustained analysis in the philosophy literature relative to causation in general ([Juarrero, 1999] is a notable exception). Probably this lack of attention is largely due to a widely shared assumption that causal analysis in complex systems is no different in kind than in typical metaphysics literature (save it is obviously more complex than the usual examples on which philosophers tend to focus). However, complexity raises difficult questions for thinking about causation when nonlinear inter-level relationships, rapid amplification of the smallest perturbations and so forth are present. For example, how are we to identify the causes at work in systems exhibiting sensitive dependence? What to do if quantum effects possibly can play causal roles in such systems (sec 3.2) or electrons dancing about in a distant galaxy possibly can play a causal role in such systems here on Earth (sec 2.7)? How far down do we have to go to identify all the causes at work in a complex macroscopic system (e.g., to butterfly wing flaps, to the atomic level or beyond)? Or how far do we have to extend a complex system to capture all of its causes (e.g., weather near the earth's surface or must we include the stratosphere, troposphere, magnetosphere, solar system, etc.)? There is a real problem here of specifying just what the causes in complex systems are aside from the trivial answer: everything! Finding principled ways to draw the boundary around the "crucial" or "dominant" causes at work in complex systems is difficult, to say the least because one of the lessons that nonlinear dynamics teaches us is that "small" causes are not insignificant (e.g., [Bishop, 2008a; 2008c]).

Hierarchies also raise questions about causation in complex systems (secs. 2.6 and 2.7). Typical metaphysical analyses consider all causation as "bottom up," where the system components are the only causal actors and systems as a whole have causal power in virtue of the causal powers of their constituents (e.g., [Kim, 2007]). But when control hierarchies act to limit or govern the causal influence of system components, is this not "top down?" Instances where lower levels provide necessary but insufficient conditions for the total behavior of wholes are rather routine in complexity systems. Moreover, when higher levels in a hierarchy and wholes act to constrain or direct the causal powers of constituents, even lower-level constituents in and of themselves turn out to not have necessary and sufficient conditions for governing all of their behavior (sec. 3.4). To repeat some key ideas of secs. 2.6 and 2.7 in slightly different language, in complex systems the formation of control hierarchies often comes about when a new form of dynamics arises that exhibits downward constraint on system constituents and is self-sustaining (e.g., [Hooker, 2004, pp. 449-477; Bishop, 2008c]). This kind of dynamical top-down constraint has a character more resembling Aristotle's notion of formal cause than efficient cause and has been largely unanalyzed by analytic philosophers (who tend to focus on logical and formal relationships among efficient causes in bottom-up constructions than on dynamics and dynamical relations).

3.4 Reduction and emergence

The issues of identity and individuation as well as causation in complex systems lead naturally to a discussion of reduction and emergence in complex systems. In rough outline, reductionist lore maintains that properties and behavior of systems as a whole are completely determined by the states and properties of their parts (ontic claim) or are explainable in terms of the states and properties of their parts (epistemic claim). Defenders of emergence deny one or both of these claims. The property of linear superposition plays an interesting role in the concepts of resultant and emergent forces in such systems. However, the loss of superposition and the possibilities for holism and constraining causation lead to the need to consider an alternative to the received views.

For instance, the lack of necessary and sufficient conditions for the behavior of lower-level constituents in complex systems directly challenges reductive atomism (e.g., control hierarchies). One of the core principles of atomistic physicalism as identified by Robert van Gulick is that "The only law-like regularities needed for the determination of macro features by micro features are those that govern the interactions of those micro features in all contexts, systemic or otherwise" [van Gulick, 2001, p. 18]. However, in complex systems control hierarchies and other inter-level causal relations are crucial to determining the behavior of system constituents. The behavior of constituents in such systems is *conditioned* by contexts in which the constituents are situated.

Recall the characterization of nonlinear systems and the linear superposition principle (sec. 2.2) and the relationship of the failure of that principle to hierarchical and large-scale structure behavior (sec. 2.6). When linear superposition holds, a system can be decomposed into its constituent parts and the behavior of each component is independent of the other components. This is the typical way in which philosophical proponents of reductionists tend to conceive of all systems. In contrast, when linear superposition breaks down, as it does for complex systems, such systems often exhibit behaviors reflecting the fact that individual system components are not independent of each other. Moreover, the behavior of individual system components are not even independent of the wholes (and various hierarchies in between). Hierarchies and wholes act to enable or constrain various possibilities for component behavior relative to what would be possible for the components if the hierarchies and wholes were absent.[16]

The interplay between parts and wholes in complex systems leads to the self-organization observed in such systems. Their sensitive dependence on the smallest of changes at the component level is partly guided by the inter-level causal relations in such systems (e.g., determining the characteristic features of convecting cells in Rayleigh-Bénard convection due to initial perturbations and instabilities in the system). This kind of behavior may be fruitfully captured by the concept of *contextual emergence* (see [Bishop, 2005b; Bishop and Atmanspacher, 2006]):

> The properties and behaviors of a system at a particular level (including its laws) offer necessary but not sufficient conditions for the properties and behaviors at a higher level.

The reference to necessary conditions at the lower level means that properties and behaviors of components at the higher level of a system may imply the properties and behaviors of components at the lower level. However, the converse is not true as the lower-level properties and behaviors do not offer sufficient conditions for the properties and behaviors of higher-level components. Contingent conditions specifying the context for the transition from the lower to the higher level of properties and behaviors are required to provide such sufficient conditions. In complex

[16][Juarrero, 1999; Silberstein and McGeever, 1999; Bishop, 2004; Bishop and Atmanspacher, 2006; Ellis, 2006; Bishop, 2008a]. In the language of Lagrangian mechanics the failure of the laws and conditions at the lower level to serve as both necessary and sufficient conditions for higher-level behavior is due to the failure of the constraints to be holonomic (see [Symon, 1971], sec 9.4, for a discussion of holonomicity).

systems, such contingent contexts are not given by the lower-level properties and behaviors alone. Moreover, the conditions for specifying a candidate reduction are not well defined until an appropriate contingent context is properly specified [Hooker, 2004, pp. 467-468].

In this sense, complex systems seem to validate intuitions many emergentists have that some form of holism plays an important causal role in complex systems that is missed by reductionist analyses. However, care is needed with typical emergentist slogans such as "The whole cannot be predicted from its parts," or "The whole cannot be explained from its parts" when it comes to contextual emergence like that exemplified in complex systems. Relevant information about the lower-level properties and behaviors of constituents *plus* the specification of an appropriate contingent context allows for the (in principle) prediction or explanation of higher-level properties and behaviors in many cases (e.g. [Primas, 1998; Bishop, 2005b; Bishop and Atmanspacher, 2006]). So complexity holds surprises for both proponents of reductionism and emergence.

3.5 Laws

While most metaphysicians focus on the "upward" flow of efficient causation from system components to system behavior as a whole, the possibilities for inter-level causal relationships in the dynamics of complex systems like convecting fluids present plausible examples of a "downward" flow of causation constraining the behavior of system components. Such behaviors clearly raise questions about the nature of laws in complex systems.

A very popular conception of science is that its goal is to discover natural laws often given form as universal statements in scientific theories. Philosophically there have been two main traditions for analyzing the nature of these laws. One tradition is the necessitarian tradition, where laws represent genuine necessities governing the regularities and patterns we find in reality. The other tradition is the regularity tradition, where laws are descriptive of regularities and patterns but no genuine underlying necessities exist. Although these two traditions involve differing relationships to the various theories of causation on offer, a very important class of laws in scientific theories have been causal laws, which govern or specify the history of state transitions a system will make given some initial starting conditions. Such causal laws often play important roles in philosophical analyses of science as well as metaphysical accounts of reality, however features like inter-level causation challenge an exclusive focus on this role.

For example, exclusive focus on causal laws might lead one to worry that fundamental laws of physics — viewed as causal laws — are being violated if the lowest-level constituents of systems (e.g., elementary particles, molecules) plus the fundamental laws are not jointly necessary and sufficient to fully determine the behaviors of these constituents (and, thereby, determine the behaviors of all higher-level constituents). Much is going on in this philosophical worry. Part of what is presupposed in this worry is an understanding of natural laws as being universal in the sense of being context free (as in the characterization of atomistic physicalism given above in sec. 3.4). Another presupposition in this worry is that the laws governing actual-world systems are only or primarily of the causal variety. Although perhaps under analyzed, science also includes structuring laws that govern or structure the range of possibilities, but do not necessarily specify which of those possibilities are actualized.

Rather than giving in to the worry, perhaps the lesson of nonlinear dynamics and complexity is that we should reconceive the primary role laws play in scientific theories. After all, scientific theories are notorious for eschewing language of the efficient causal type with which philosophers are enamored. What fundamental laws primarily do is carry out structuring functions, where the relevant space of possibilities is determined for the behaviors of the constituents guided by them, but such laws do not fully determine which of these possibilities in the allowable space are actualized. That can only be fully determined by concrete contexts into which the laws are coming to expression (which may involve higher-level causal and/or structuring laws). For instance, Newton's law of gravity determines the space of possible motions for an apple falling from a tree, but the concrete context where I reach out my hand and catch the apple actualizes a particular possibility among all those allowed even though the particular context is not included in the law of gravity.

The point, here, is very similar to that for contextual emergence above (sec. 3.4): while fundamental laws establish necessary conditions for the possible behaviors of objects, contingent contexts must be added in to establish jointly necessary and sufficient conditions for actual behaviors. If the fundamental laws play primarily structuring roles in nature, then concrete contexts are as important as the laws.[17] Empirically, this is consonant with the intricate and delicate behavior of complex systems. We should, then, resist the tendency to either appeal to only fundamental causal laws in our explanations or to pit causal laws against structuring laws in competing explanations. Sound explanations of complex systems are likely to involve both appeals to causal mechanisms and ordering/constraining structure via the interrelationships among the lower-level and higher-level dynamics[18]

If fundamental laws primarily structure the spaces of possibility they establish but do not fully determine the outcomes within this space, worries about violations of fundamental laws fade away. Hierarchies and wholes in complex systems act to constrain or direct the possibilities made available by lower-level laws as opposed to somehow violating those laws. Since such laws are part of the necessary conditions for the behavior of higher-level constituents, such laws cannot be violated by the behavior of higher-level constituents. Indeed, it is possible that the structuring function played by control hierarchies and wholes in complex systems are the result of some as yet unknown nonlinear dynamical laws, which may be causal

[17]Similar lessons about laws can be gleaned from continuum and statistical mechanics even in cases where systems are not exhibiting complexity.

[18]See [Chemero and Silberstein, 2008, sec. 4; Bishop, 2008b, sec. 5.2] for some discussion as to how these two levels might be brought together fruitfully in the service of explanation.

or structural as well as emergent with respect to the contexts where nonlinear interactions are dominant. Even so, fundamental laws still would be necessary for structuring the possibility space for emergent nonlinear laws, though particular features of contexts might be required for particular nonlinear laws to arise.

But complexity raises more questions for laws than the relative roles of structural vs. causal laws. It also raises questions about how we are to conceive laws that require us to fix appropriate boundary conditions for their corresponding equations to be well posed.[19] For example, the fluid equations governing Rayleigh-Bernard convection require the imposition of a constant temperature along the bottom plate of a container holding fluid so that a temperature difference can be established. In an idealized mathematical model of this situation, there is a principled choice to make for the boundary. However, when these same fluid equations are applied to model atmospheric weather, all of the boundary problems mentioned above in sec. 2 arise, where there is no longer any obvious choice for where to place the cut between weather system and boundary. Operationally it is fine that we can make pragmatic choices that give us well-posed equations to solve on a computer, but the foundational questions of the status of the laws in question and the determination of their boundaries remains unanswered. Perhaps such laws as philosophers have typically conceived them have no instantiations for complex systems because these systems lie outside the laws' domain of applicability, given the lack of an ontologically distinguishable boundary. Yet, we still have stable patterns where the dynamics governs the outcomes even if our philosophical analyses of laws come up short in characterizing the dynamics.[20]

Moreover, we typically connect laws with physical systems via models, which means that the faithful model assumption (sec. 2.3) is being invoked. Faithful yet imperfect models leave open questions about the applicability of the laws to phenomena exhibiting complexity. Even if we take the faithful model assumption to its extreme limit — the perfect model scenario — we run into problems since there are too many states indistinguishable from the actual state of the system yielding empirically indistinguishable trajectories in the model state space [Judd and Smith, 2001]. Add into the mix that the mapping between our nonlinear dynamical models could be many-to-one or many-to-many, and our philosophical judgements become difficult about the roles scientific laws play in the corresponding actual-world systems and which of these laws are fundamental and which are emergent (if any).

Empirically, it is hard to be sanguine about the necessitarian tradition on laws which, in turn, puts pressure on realist views of natural laws. On the other hand, the regularity tradition does not rest too comfortably either. The difficulties nonlinear dynamical systems raise for the notion of faithful models lead to analogous

[19]For an insightful survey of the delicate nature of the relationship between laws and boundary conditions, see [Wilson, 1990].

[20]Similar questions can be raised about the status of scaling laws and laws involving universality and order parameters, which are ubiquitous in literature on complex systems. The subtleties of such systems warrant care in being too quick to conclude that these laws are "merely epistemic" because we "lack access" to the underlying causal mechanisms at work in such systems.

worries about the role regularities play in actual-world systems, and about whether the regularities in these systems are fundamental and/or emergent. The regularities of complex systems are there, to be sure, and engage intense scientific study, but our current philosophical understanding of laws appears to be inadequate.

4 DISCUSSION

This brief survey of complexity and its implications suggests that there are challenges to our philosophical and scientific lore about the world. Nonlinear dynamics and the loss of linear superposition shed different light than is typically found in philosophical literature on identity and individuation, predictability, confirmation, determinism, causation, reduction and emergence, and natural laws, some of the bread and butter topics of metaphysics, epistemology and philosophy of science. There is much remaining to be explored about how nonlinear dynamics and complexity can challenge and enrich our understanding of metaphysics and epistemology as well as science and its practices.

On the one hand, nonlinear modeling and insights from complexity studies are in widespread use in the sciences (and increasingly in public policy). There is a genuine sense that nonlinear models have led to tremendous advances (e.g., in chemical studies, drug development, weather prediction, plasma physics). On the other hand, extracting firm results from such modeling is nontrivial, given that these results are context-dependent, qualified by some very strong idealizations and difficult to confirm. This is not the kind of picture of science presented in our textbooks and a lot of the philosophical literature. Moreover, the upshot of complex systems modeling for our pictures of the world, whether deterministic, causal or reductionistic, is challenging and nuanced.

Many philosophers and scientists have reflected on science and its methods with a set of assumptions that underestimate the complexity of science itself. Our assumption that faithful models are relatively unproblematic and relatively straightforward methodologically seemed to serve fine in scientific inquiry before the advent and explosion of nonlinear modeling and complexity studies. However, the latter have pushed our philosophical and scientific understanding to their limits as described above.

One of the important challenges to our thinking about science and its methods that complex system make abundantly clear is clarifying our understanding of the strengths and weaknesses of nonlinear modeling. Roughly, on the strength side, there is greatly increased power to model actual-world phenomena that defies our analytical capabilities. In domains like fluid flow, weather forecasting and fusion reactor modeling, tuning our models to the best available empirical data sets has proved quite useful for refining our models. On the limitation side, there is the challenge of understanding how to extract useful, reliable information from imperfect and difficult-to-confirm models that do have some faithfulness to the target systems of interest. For instance, even when we have tuned our nonlinear models to the best data sets, there is still a great deal of inadequacy in those data sets (no data set is ever perfect) that is translated into our models; this is over and above the inadequacy inherent in our faithful models. At the end of the process there still remains a great deal of uncertainty in the results of our models. To what extent do modelers and end users who receive model outputs as deliverables understand the limitations of these models? To what extent do they understand the uncertainties inherent in the modeling results? What strategies can modelers use to extract trustworthy information in the midst of such model inadequacy and uncertainty? What changes do end users, such as public policy officials, need to make so that their reasoning reflects the limitations and uncertainty of the modeling science on which they are relying? These latter questions are some of the most difficult and pressing philosophical and methodological issues raised by complex systems and their models. Efforts by philosophers, scientists, public policy experts and others to answer such questions would be effort well spent.

Any insights gleaned in exploring these questions will no doubt further our thinking in metaphysics and epistemology about determinism, causation, reduction/emergence, natural laws and confirmation in the context of complex systems. One example that is relevant to the latter issues is raised by the possibilities for there to be a many-to-one relationship between different nonlinear models and a single target system: Are such mathematical models simulating the target system or merely mimicking its behavior? To be simulating a system suggests that there is some actual correspondence between the model and the target system it is designed to capture. In contrast, if a model is merely mimicking the behavior of a target system, there is no guarantee that the model has any genuine correspondence to the actual properties of the target system. The model merely imitates behavior. This question becomes particularly important for modern techniques of building nonlinear dynamical models from large time series data sets (e.g., [Smith, 1992]), for example the sunspot record or the daily closing value of a particular stock for some specific period of time. In such cases, after performing some tests on the data set, modelers set about their work constructing mathematical models that reproduce the time series as their output. When multiple such models, each conceptually and mathematically different, reproduce the target system behavior with roughly equal accuracy (and inadequacy!), are such models simulating target systems, or only mimicking them? Here, all the questions about understanding model limitations and uncertainties, strategies for extracting useful information and how to reason about public policy or other implications based on model output are raised. In addition, for the philosophers further questions about what these models are telling us about our world and our access to that world are also raised with their attendant implications for metaphysics and epistemology.

The metaphysical and epistemological implications of complex systems is very rich, indeed.

ACKNOWLEDGEMENTS

Insightful discussions with Cliff Hooker, Jeffery Koperski, Michael Silberstein and Leonard Smith are all gratefully acknowledged. Remaining weaknesses are my sole responsibility.

BIBLIOGRAPHY

[Antoniou and Prigogine, 1993] I. Antoniou and I. Prigogine. Intrinsic Irreversibility and Integrability of Dynamics. *Physica A*, 192, 443-464, 1993.

[Atmanspacher, 2002] H. Atmanspacher. Determinism Is Ontic, Determinability Is Epistemic. In H. Atmanspacher and R. C. Bishop, eds., *Between Chance and Choice: Interdisciplinary Perspectives on Determinism*. Imprint Academic, Thorverton, UK, 2002.

[Badii and Politi, 1997] R. Badii and A. Politi. *Complexity: Hierarchical Structures and Scaling in Physics*. Cambridge University Press, Cambridge, 1997.

[Bishop, 2002] R. C. Bishop. Deterministic and Indeterministic Descriptions. In H. Atmanspacher, R. Bishop (eds.), *Between Chance and Choice: Interdisciplinary Perspectives on Determinism*. Imprint Academic, Thorverton, UK, 2002.

[Bishop, 2003] R. C. Bishop. On Separating Prediction from Determinism," *Erkenntnis* 58, 169-188, 2003.

[Bishop, 2004] R. C. Bishop. Nonequilibrium Statistical Mechanics Brussels-Austin Style," *Studies in History and Philosophy of Modern Physics* 35,1-30, 2004.

[Bishop, 2005a] R. C. Bishop. Anvil or Onion? Determinism as a Layered Concept," *Erkenntnis* 63, 55-71, 2005.

[Bishop, 2005b] R. C. Bishop. Patching Physics and Chemistry Together," *Philosophy of Science* 72, 710-722, 2005.

[Bishop, 2006] R. C. Bishop. Determinism and Indeterminism. In D. M. Borchert (ed. in chief), *Encyclopedia of Philosophy, Second Edition*. Macmillian Reference, Farmington Hills, MI, 3, 29-35, 2006.

[Bishop, 2008a] R. C. Bishop. Downward Causation in Fluid Convection. *Synthese* 160, 229-248, 2008.

[Bishop, 2008b] R. C. Bishop. Chaos. *Stanford Encyclopedia of Philosophy*, `http://plato.stanford.edu/entries/chaos/`, 2008.

[Bishop, 2008c] R. C. Bishop. What Could Be Worse than the Butterfly Effect?", *Canadian Journal of Philosophy* 38, 519-548, 2008.

[Bishop and Atmanspacher, 2006] R. C. Bishop and H. Atmanspacher. Contextual Emergence in the Description of Properties. *Foundations of Physics* 36, 1753-1777, 2006.

[Bricmont, 1995] J. Bricmont. Science of Chaos or Chaos in Science? *Physicalia Magazine* 17, 159-208, 1995.

[Busse, 1978] F. Busse. Non-Linear Properties of Thermal Convection. *Reports on Progress in Physics* 41, 1929-1967, 1978.

[Chemero and Silberstein, 2008] A. Chemero and M. Silberstein. After the Philosophy of Mind: Replacing Scholasticism with Science. *Philosophy of Science* 75, 1-27, 2008.

[Cross and Hohenberg, 1993] M. Cross and P. Hohenberg. Pattern Formation Outside of Equilibrium. *Reviews of Modern Physics* 65, 851-1112, 1993.

[Crutchfield, 1994] J. Crutchfield. Observing Complexity and the Complexity of Observation. In H. Atmanspacher and G. Dalenoort, eds., *Inside Versus Outside*. Springer, Berlin, 235-272, 1994.

[d'Espagnat, 1994] B. d'Espagnat. *Veiled Reality: An Analysis of Present-Day Quantum Mechanical Concepts*. Addison-Wesley, Reading, MA, 1994.

[Ellis, 2006] G. F. R. Ellis. Physics and the Real World. *Foundations of Physics* 36, 227-262, 2006.

[Grassberger, 1989] P. Grassberger. Problems in Quantifying Self-Generated Complexity. *Helvetica Physica Acta* 62, 489-511, 1989.

[Hobbs, 1991] J. Hobbs. Chaos and Indeterminism. *Canadian Journal of Philosophy* 21,141-164, 1991.

[Hooker, 2004] C. A. Hooker. Asymptotics, Reduction and Emergence. *British Journal for the Philosophy of Science* 55, 435-479, 2004.

[Judd and Smith, 2001] K. Judd and L. Smith. Indistinguishable States I: Perfect Model Scenario." *Physica D* 151, 125-141, 2001.

[Juarrero, 1999] A. Juarrero. *Dynamics in Action: Intentional Behavior As a Complex System.* MIT Press, Cambridge, MA, 1999.

[Kant, 1980] I. Kant. *The Critique of Judgement*, trans. J. C. Meredith. Oxford University Press, Oxford, 1980.

[Kellert, 1993] S. Kellert. *In the Wake of Chaos.* Chicago University Press, Chicago, 1993.

[Kim, 2007] J. Kim. *Physicalism, or Something Near Enough.* Princeton University Press, Princeton, 2007.

[Kripke, 1980] S. Kripke. *Naming and Necessity.* Basil Blackwell, Oxford, 1980.

[Koperski, 1998] J. Koperski. Models, Confirmation and Chaos. *Philosophy of Science* 65, 624-648, 1998.

[Lorenz, 1965] E. N. Lorenz. A Study of the Predictability of a 28-Variable Atmospheric Model. *Tellus* 17, 321-33, 1965.

[McLaughlin, 1982] B. McLaughlin. British Emergentism. In A. Beckermann, H. Flohr, and J. Kim, eds., *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism.* Walter de Gruyter, Berlin, 49-93, 1982.

[Oseledec, 1969] V. I. Oseledec. A Multiplicative Ergodic Theorem: Lyapunov Characteristic Number for Dynamical Systems. *Transactions of the Moscow Mathematical Society* 19, 197-232, 1969.

[Pattee, 1973] H. Pattee. The Physical Basis and Origin of Hierarchical Control. In H. Pattee (ed.), *Hierarchy Theory: The Challenge of Complex Systems.* George Braziller, New York, 69-108, 1973.

[Petrosky and Prigogine, 1996] T. Petrosky and I. Prigogine. Poincaré Resonances and the Extension of Classical Dynamics. *Chaos, Solitons & Fractals* 7, 441-497, 1996.

[Petrosky and Prigogine, 1997] T. Petrosky and I. Prigogine. The Extension of Classical Dynamics for Unstable Hamiltonian Systems. *Computers & Mathematics with Applications* 34, 1-44, 1997.

[Polkinghorne, 1989] J. C. Polkinghorne. *Science and Creation: The Search for Understanding.* Shambhala, Boston, 1989.

[Popper, 1950] K. Popper. Indeterminism in Quantum Physics and in Classical Physics. *The British Journal for the Philosophy of Science* 1, 117-133, 1950.

[Prigogine, 1997] I. Prigogine. *The End of Certainty: Time, Chaos, and the New Laws of Nature.* The Free Press, New York, 1997.

[Primas, 1983] H. Primas. *Chemistry, Quantum Mechanics and Reductionism: Perspectives in Theoretical Chemistry.* Springer, Berlin, 1983.

[Primas, 1990] H. Primas. Mathematical and Philosophical Questions in the Theory of Open and Macroscopic Quantum Systems. In A. Miller, ed., *Sixty-Two Years of Uncertainty.* Plenum, New York, 233-57, 1990.

[Primas, 1994] H. Primas. Endo- and Exotheories of Matter. In H. Atmanspacher and G. Dalenroot, eds., *Inside Versus Outside.* Springer-Verlag, Berlin, 163-93, 1994.

[Primas, 1998] H. Primas. Emergence in Exact Natural Sciences. *Acta Polytechnica Scandinavica* 91, 8398, 1998.

[Silberstein and McGeever, 1999] M. Silberstein and J. McGeever. The Search for Ontological Emergence. *The Philosophical Quarterly* 49, 182-200, 1999.

[Scheibe, 1964] E. Scheibe. *The Logical Analysis of Quantum Mechanics.* Pergamon Press, Oxford, 1964.

[Smith, 1992] L. A. Smith. Identification and Prediction of Low Dimensional Dynamics. *Physica D* 58, 50-76, 1992.

[Smith, 1997] L. A. Smith. Past and Present Variability of the Solar-Terrestrial System: Measurement, Data Analysis and Theoretical Models. In G. C. Castagnoli, A. Provenzale (eds.), *Proceedings of the Enrico Fermi International School of Physics, Course CXXXIII.* IOS, Amsterdam, 177-246, 1997.

[Smith, 2000] L. A. Smith. Disentangling Uncertainty and Error: On the Predictability of Nonlinear Systems. In A. Mees, ed., *Nonlinear Dynamics and Statistics.* Birkhauser, Boston, 31-64, 2000.

[Smith, 2003] L. A. Smith. Predictability Past Predictability Present," in T. Palmer (ed.), *Seminar on Predictability of Weather and Climate, 9-13 September 2002*. European Centre for Medium-Range Weather Forecasts, Reading, UK, 219-242, 2003.

[Smith, *et al.*, 1999] L. A. Smith, C. Ziehmann, and K. Fraedrich. Uncertainty Dynamics and Predictability in Chaotic Systems. *Quarterly Journal of the Royal Meteorological Society* 125: 2855-86, 1999.

[Sosa and Tooley, 1993] E. Sosa and M. Tooley (eds.). *Causation*. Oxford University Press, Oxford, 1993.

[Symon, 1971] K. Symon. *Mechanics* (3^{rd} ed.). Addison-Wesley, Reading, MA, 1971.

[Thompson, 1957] P. D. Thompson. Uncertainty of Initial State as a Factor in the Predictability of Large Scale Atmospheric Flow Patterns. *Tellus* 9, 275 — 295, 1957.

[Van Gulick, 2001] R. Van Gulick. Reduction, Emergence and Other Recent Options on the Mind/Body Problem: A Philosophic Overview. *Journal of Consciousness Studies* 8,1-34, 2001.

[Wackerbauer *et al.*, 1994] R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurths and H. Scheingraber. A comparative classification of complexity measures. *Chaos, Solitons, and Fractals*, 4: 133-173, 1994.

[Wilson, 1990] M. Wilson. Law Along the Frontier: Differential Equations and Their Boundary Conditions. *PSA 1990, Vol. 2*, 565-575, 1990.

[Ziehmann, *et al.*, 2000] C. Ziehmann, L. A. Smith, and J. Kurths. Localized Lyapunov Exponents and the Prediction of Predictability. *Physics Letters A* 271, 237-51, 2000.

COMPUTING AND COMPLEXITY — NETWORKS, NATURE AND VIRTUAL WORLDS

David G. Green, Tania Leishman

INTRODUCTION

By any measure you care to use, the scale and complexity of modern computing and information systems are growing at explosive rates. In the early 1990s, shortly after the World Wide Web was founded, all the web sites in the world could be listed on a single page. Less than 20 years later, the Web was estimated to contain some 108,810,385 distinct web sites [Netcraft, 2007], providing between them many billions of pages of information.

In the 1970s, when computers were still huge machines housed in special vaults, data on almost any topic was a rare and expensive commodity. By the early twenty-first century it had become abundant and cheap. The growth of biological information is typical. Genbank [Bilofsky and Burks, 1988], the main US repository for bioinformation, was established in 1982 with entries for 606 DNA sequences. Twenty-five years later, in February 2007, the number of entries had grown to over 10 million sequences (containing over 100 billion nucleotides) and included the complete DNA sequences for more than a dozen species [NCBI, 2007].

The data processing capacity of computers has increased in a similarly spectacular fashion. In 1955, the IBM STRETCH computer performed around 5000 floating point operations per second. Fifty years later, the world's fastest supercomputers were more than fifty billion times as fast, achieving up to 280.6 trillion floating point operations per second [TOP500, 2008].

As computers became more powerful, they also diversified. Specialized computers perform all manner of applications, from digital watches to mobile phones. Instead of a single, large, central computer, it is now possible to have many computers, each carrying out local functions and local control. As a result, computers are now found everywhere in people's lives. The natural conclusion of the trend towards small, specialized computers is the idea of *pervasive computing*. That is, computing is everywhere. This idea extends the notion of complexity in computer design and function, linking it to the complexity of issues in everyday human activity.

An important benefit of computers has been to save human time and effort. Advances in technology have led to changes in the way computers are used. As

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

they became smaller and less expensive, computers were no longer limited to large organizations. By the 1980s, the advent of personal computers led to a rapid expansion of the user base. By the early 21^{st} Century, computers had become standard equipment for almost every home and business in the western world.

The transition from data poor to a data rich society marks a huge change in the way information is handled. It posed huge challenges for information technology, which had to learn how to deal with enormous volumes of information and to solve the complex problems that arise in interpreting vast quantities of data. Meeting this challenge has led to many new ideas and approaches that are motivated by complexity and exploit many of the insights gained from complex systems science.

Complexity has had a profound influence on the evolution of computing and information processing. Its influence arises both from the nature of computers themselves, and from the nature of the applications for which computers are used.

Complexity is the richness of structure and behaviour often seen in large systems. It arises in many aspects of computing, including the design of software, hardware, and communications networks. Complexity also leads to many kinds of problems that need to be dealt with, such as: cascading faults, emergent properties, robustness and reliability, and critical behaviour.

Modern computing is intimately bound up with complexity theory. The need to deal with complex problems has motivated many ideas and issues in computing. Conversely, computing, and computational ideas, have also played a prominent role in the development of complexity theory. In applications of information technology, complexity arises from the need to handle enormous volumes of information and to solve ever larger and more intricate problems (e.g. multi-objective optimisation). Combinatorial explosions in large datasets create problems of high complexity and pose a major challenge for computer science. Several key ideas from complexity theory have dominated the struggle to cope with the inherent complexity of computing and information. Perhaps the first of these ideas was *encapsulation.*

ENCAPSULATION AND MODULARITY

What do you find when you walk into a typical house? A house that was one single large space would be unusual. It would also make living difficult. So we carve up our homes according to the ways we live. A house consists of rooms; each specialized for a different activity: washing, sleeping, cooking, eating, and so on. And the possessions are sorted into the room most appropriate to them. By this division of activity we reduce the complexity of everyday living.

This simple approach — “divide and rule” — is the most common way of dealing with complexity. Traditionally, people deal with complexity by trying to avoid it. And the most common way of avoiding complexity is by encapsulation. You divide a large, complex problem into smaller and simpler ones. That is, you break down a large system into smaller, more manageable units by encapsulating related items together in distinct compartments (units, modules). The “divide and rule” method appears in many contexts. It is so common in everyday life that most people are not even aware of it: handbags have pockets, libraries have sections, and large companies have divisions, departments and sections.

The rationale for the above approach is that compartmentalizing a problem breaks a large system or problem into smaller parts. It reduces the complexity by reducing the number of connections within the system. Note that there is a fine distinction between encapsulation, which usually refers to enclosing and isolating some set of objects, and modularization, which goes a step further by turning an encapsulated set of objects into a single, reproducible unit.

Not surprisingly, the divide and rule approach has dominated computing. To understand how encapsulation reduces complexity, look at the issues involved in networking computers. Imagine you have (say) 100 computers to organize into a data communications network. If you tried to link each machine to every other machine directly, then 4,950 connections would be needed. Such an arrangement would achieve maximum speed, but at a huge cost in wiring, most of which would never be used.

On the other hand, if you arrange the computers in a loop, with each computer passing on a message to the next one in the ring, then only 100 connections are needed, but communications would be slower. On average it would take 25 hops (maximum 50 hops) from computer to computer for a message to get from its origin to the destination machine. If you arrange the computers in the form of a binary tree, then again 100 connections are needed, but the maximum distance between machines falls to 14 hops ($2 \log_2 128$). Finally, if you organize the computers into sub-domains of 10 machines each (a form of encapsulation), with one machine acting as a router for each domain, then the maximum distance falls to just 4 hops (Figure 1).

The efficiency of a modular communications network rises dramatically as the size of the network increases. For 1000 computers, under the conditions described above, the maximum distances are 500 for a ring, 20 for a binary tree, and 6 for a hierarchy of clusters. For a million computers, these numbers increase to 500,000, 40, and 12 respectively.

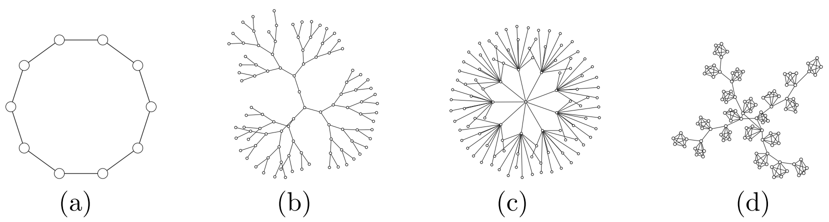


Figure 1. Four ways of arranging a communications network between computers (nodes): (a) a ring; (b) a tree (binary), with the root at the centre; (c) a hierarchy (tree) of domains, with one node in each acting as a router; (d) a hierarchy of modules, well-connected internally, but minimal connections between modules.

Networks and modules

The idea of a network is fundamental in both computing and complexity theory. A *network* is a collection of *nodes* (objects) and *edges* (links) that join pairs of nodes. The *connectivity* (or *edge density*) of a network is the number of edges as a proportion of the maximum possible number of edges. A network is *connected* if there is a *path* (series of edges) from each node to every other node.

The importance of networks in complexity theory stems from their universality. This is guaranteed by the following two theorems, which show that networks are implicit in the common representations of complex systems [Green, 1993; 2000]:

THEOREM 1. *The patterns of dependencies in matrix models, dynamical systems and cellular automata are all isomorphic to networks.*

THEOREM 2. *In any array of deterministic automata with a finite number of states, the state space forms a network.*

Together these theorems show that networks are implicit in both the structure and behaviour of almost any complex system. In later sections, we will discuss the implication, namely that properties of networks underlie many features of complex systems.

The ideas of networks and encapsulation are closely linked. A *module* is a set of nodes within a network and having the property that its internal connectivity is high but its connectivity with the rest of the network is low (Figure 1d). Modules form networks in two ways. In a *whole-part* hierarchy, sets of modules are linked together to form larger modules. For instance, the input, output and processing modules are parts that make up a whole computer. In a *gen-spec* ("general-special") relationship, different classes of modules (or objects) inherit properties from one another. A laptop is a special kind of personal computer, which is a special kind of digital computer, which is a special kind of information processor.

The importance of networks in computing arises from their intimate relationship to modular structures. A program is a network of processes (or objects). A supercomputer is usually a network of processors (e.g. clusters or grids). The Internet is a network of interlinked computer networks. Distributed data warehouses are organized networks of information objects.

The idea of encapsulation, in one form or another, turns up almost everywhere in computing. It took hold in programming as a way of coping with the complexity of computer source code. Programmers face a major problem. Since the earliest days of computers, writing programs has always been plagued by the difficulty of eliminating errors. It is almost impossible to write a large program that is error free. Even a program of modest size and complexity is bound to need debugging.

Problems arise both from mistakes by the programmer, and from fatal errors that emerge when different pieces of code interact in unexpected ways. The frequency of errors in a program increases exponentially with the number of lines of code. This problem, a symptom of complexity, at first made it almost impossible to write long programs. The solution, as the first programmers quickly learned, is to encapsulate pieces of code into manageable units (e.g. macros, functions, subroutines, and procedures). Small units of code are easy to perfect and can then be reused wherever a similar operation is needed.

Not only does the use of encapsulation reduce errors. It also speeds up the business of writing a program enormously. By creating names for particular pieces of code, a programmer merely needs to write in the name and have the computer substitute the actual code later. The practice of labelling sections of code in this way soon led to the idea of compilers and to high level programming languages.

The principle of encapsulation (and modularity) has remained a cornerstone of information system design ever since. For instance, object-oriented design treats systems as consisting of inter-related information objects, within which most attributes and methods are hidden from the rest of the system. This prevents interactions between objects from producing unpredictable errors and other unexpected effects.

In both computer science and information technology, modularity has been one of the first lines of defence against rampant complexity. However, as explained earlier, modularity is a way of reducing complexity. As programs, systems and problems became more and more complex; computer science had to find ways to deal with a host of problems posed by increasing complexity. Some of the ideas that arose in the course of this quest have deeply influenced the development of complexity theory. indexcombinatorial complexity

TIME, SPACE AND COMPLEXITY

Necessity is the mother of invention, and in computing it has given birth to a whole tribe. One of the greatest needs in computing and information technology is the vexed question of limited time and space. These two issues underlie many of the ideas and methods adopted in computer science. Note that we are not talking here about time and space in the usual sense, but computing time and memory space.

In computation, there is often a trade-off between processing time and memory. If one is limited, then you can reduce the need by using more of the other. For instance, if a computer needs to do the same calculation again and again, then it is a waste of time repeating the entire procedure every time. If you store the results in (say) a cache, then instead of doing an entire calculation, the computer simply looks up the answer.

In information science, the term "complexity" is usually taken to mean "computational complexity". The idea of carving up a large problem into smaller ones enters computing in the everlasting quest for more time and space. As we saw earlier, encapsulation in complex environments often helps achieve greater efficiency in both speed and memory. More generally, the ideas of time and space motivate two approaches to complexity that arose in the context of computation: *combinatorial complexity* (time) and *computational complexity* (space).

Combinatorial complexity

The first issue is how the computer time needed to solve a problem increases in relation to the size of the problem. Simple problems grow in simple fashion, but complex problems grow alarmingly. Some examples will make this clearer. Suppose you want to find the largest item from a list. If you double the size of the list, then you double the time needed. If you want to find the two most similar people in a list, then the time taken increases in proportion to the square of the number of people in the list. These problems grow in "polynomial time" (referred to as class $\boldsymbol{P}$). However, if you want to find the best arrangement of items (e.g. a schedule), then the time required grows in exponential fashion.

Perhaps the most famous of these complex ("hard") problems is the travelling salesman problem. A salesman must visit a number of towns and wants to know the shortest route that visits each town once. With just 30 towns, even using every computer in the world, and checking a million cases a second, the time required to check all possible solutions would be greater than the age of the universe.

Computational complexity

The second issue is how the computer memory space needed to solve a problem increases in relation to the information required for the problem. *Computational complexity* concerns the relationship between complexity and information. Suppose that you need to transmit two strings:

String #1: GQIRNCMSHWLBPMWUQHYE

String #2: ABABABABABABABABABAB

The first is a random string, whereas the second contains a repeated pattern. To transmit the first, you need to send the entire 20 symbols. However the second can be compressed to a formula:
such as

Signal #1: GQIRNCMSHWLBPMWUQHYE

Signal #2: 10(AB)

Considerations such as the above led Kolmogorov [1965] and Chaitin [1966] independently to propose taking the length of a signal as a measure of complexity. Here the signal is understood to mean a program plus data. So in the above example the program is the string "10()" and the data string is "AB". The complexity of a string is then the length of the shortest program that generates it. Ideas of this kind led Wallace [Wallace and Boulton, 1968] to introduce the use of "Minimum Message Length" (MML) as a computational, and often more precise, alternative to many kinds of statistical testing [Wallace, 2005; Dowe, 2008].

Parallel processing

What do supercomputers have in common with factories and supermarkets? They all have a need for faster processing. In their efforts to increase speed, and to handle a greater volume of information, supercomputers adopt the same strategies as factory production lines and supermarket checkouts.

One strategy to increase speed is to carry out several operations at once. There are several ways to do this. In a factory assembly line, items (e.g. cars) move in a line along which different steps are performed in turn. In *pipeline*, or *vector processing*, a computer uses this approach by applying a sequence of operations to a string of data items. However, a more common strategy for parallel processing is to emulate a supermarket checkout. Instead of trying to make the checkout faster, you add more checkouts. A notable early example of this was the Connection Machine [Hillis, 1982], in which up to 64,000 processors could carry out the same operation at once. However, for SIMD[1] problems, the processing units need not be housed in a single machine. This principle led to the idea of processing clusters, in which different machines are linked so that problems become distributed between them, and grids, in which large processing problems are passed to available processing units.

Parallel computing is straightforward if all the records to be processed are independent (e.g. calculating interest for different bank accounts), but the problem quickly becomes complex if the records need to interact with each other (e.g. simulating movement of air across the earth's surface). In this case the processors have to pass messages from one to another, which poses the problem of how to provide appropriately functioning communication channels between different processors. It also slows down the processing. The complexity of communication networks, mentioned earlier, therefore becomes crucial.

The idea of parallel computing has found several natural extensions, especially *distributed computing*. That is, processing is spread across computers at different locations. This idea has motivated many different ways of organizing complex information storage and data processing. In a grid, for instance, computing resources are shared across a network of computers; in much the same way as electricity is supplied across a power grid.

In distributed data warehouses, large volumes of information are spread across machines in a network. Usually each computer specializes on one type of data. Network agents provide a looser version of the warehouse idea. Networks of software agents collaborate with each other to tap into known sources of information and pool their information.

Networks of agents are closely associated with the idea of *emergent computation*. This means that desired results arise by interactions between the agents. Several developing areas of technology exploit this idea. In robotics, Rodney Brooks [1986; 1991] and others pioneered the idea of distributed intelligence, whereby simple local

[1]SIMD stands for "Single Instruction — Multiple Data." In other words, the same instruction is carried out on many different pieces of data simultaneously.

processes combine to produce intelligent behaviour. In a robot insect, for instance, simple controllers for each of the legs interact to coordinate walking. In sensor arrays, sets of sensors interact to provide more robust or reliable outputs than a single sensor alone. Examples of applications include military surveillance and swarms of microscopic sensors in nanotechnology.

VIRTUAL LABORATORIES

Science tends to be like a drunk who has lost his car keys late a night and searches for them under the street lamp, not because he lost them there, but because it is the only place where he can see. Likewise, science tends to study what it can study, rather than what it should. For most of the history of science, this was true of complexity: science had no means of dealing with it, so it was largely ignored. One of the greatest effects of modern computing on science had been to make it possible to attack complex questions that previously were intractable. By making simulation modelling practical, computers have helped to advance complexity theory in several ways.

Virtual experiments

There are many experiments that you would never do in real life. You would never, for instance, set off a nuclear bomb in a large city just to examine its effects. You would never release Ebola virus on a population to see whether your containment measures would work. You would never let a trainee pilot put an airliner into a flat spin to see whether he can get out of it. In each case, the consequences of failure would be terrible, not to say expensive. You would never cut down large tracts of rainforest to see whether it regenerates. In each case, the experiment itself could destroy the very thing you are trying to understand and protect.

Many experiments such as the above are far too costly, too destructive or too time-consuming to consider doing in real life. On the other hand, you can do many of them routinely inside a computer. It is safer, and less expensive, for instance, for trainee pilots to make mistakes in a simulator rather than risk crashing a real aircraft.

Simulation comes into its own when dealing with complex systems. The increasing power of computers makes it possible to investigate complex matters that could not be addressed previously. In particular, the memory available in modern computing clusters is now approaching the size needed to deal with the huge numbers of elements that comprise many organic systems. This potential has motivated proposals for grand challenges, such as modelling the human brain in detail. (See below).

Dynamics as computation

One of the first uses of computer simulation was to extend the ability of mathematics to deal with complex dynamics. Dynamic systems are traditionally expressed as systems of differential equations, but complex phenomena give rise to non-linear equations, which are generally not solvable. Computer simulation provides a way around the problem. Instead of solving the equations analytically, answers are calculated iteratively.

The use of computer simulation to extend mathematics raises an important issue. Mathematical models typically treat processes as continuous. In contrast computer simulations typically update system dynamics in discrete time intervals.

The discrete nature of computation has led science into new areas. One implication is that computed solutions are approximations, never precise. Studies into the implications of the errors involved led Yoshisuke Ueda [1961/2001] and Edward Lorenz [1963] to discover the sensitivity to initial conditions displayed by chaotic systems. Likewise the discrete, iterative nature of computation lent itself to development of fractal theory by Mandelbrot and colleagues [Mandelbrot, 1982], in which repetition of processes (self-similarity) on different scales plays a central part.

Computer simulation models can also capture qualitative aspects of the internal organization of a system that cannot be captured in a formula. A good example can be found in the approaches taken in models of organic growth. The "traditional" approach uses reaction-diffusion models (expressed as differential equations) that model the global patterns of growth rates in terms of biochemical gradients. Computational approaches led to the use of computational models, which attempted to capture the local organization of growth at the level of individual cells or modules (e.g. branches and leaves in plants). This approach, pioneered by Aristid Lindenmayer [1968] expressed the local organization of growth as sets of syntactic rules, known as L-systems. The two schools usually capture completely different aspects of the same system.

The two can be reconciled by adapting the process of computation. To avoid introducing artefacts, simulation models normally update every part of a modelled system synchronously and in parallel. In effect they treat simulation like a movie: display one frame while constructing the next. However, in many processes in nature, including growth, events are actually *asynchronous*: the local rates of change can vary from place to place [Cornforth, *et al.*, 2005]. So the two kinds of models can be united by using biochemical gradients to set the clock rate for L-system models.

An important lesson of computational models is that interactions at the local scale can have a huge effect on global behaviour [May and Martin, 1975]. In models of social networks, for instance, numerical models imply a steady rate of penetration of novel opinions. In contrast, agent models change in fits and starts: often they maintain a steady state for a time and then suddenly flip to a different state [Green, *et al.*, 2006]. In models of media influence, for instance, agent models initially change much more slowly than the equivalent mathematical ones, but then become much faster once a critical threshold had been reached [Bransden and Green, 2005].

Why is a starfish like an atomic bomb?

Analogies are always dangerous. It is easy to be misled by superficial similarities. It is easy to be fooled into assuming that because an analogy can be drawn between two things they must behave alike. Superficial analogies usually disappear when fleshed out with precise details. An important benefit of simulation is that is forces people to state details precisely, to be rigorous about their assumptions. It is easy to say, for instance, that an epidemic spreads from person to person. But to simulate an epidemic, you have to pin down exactly how, when and where one person infects another. The precise details are important. As we saw in the previous section, a slight change in a single detail can make a vast difference to the outcome.

On the other hand, the process of fleshing out details can itself be instructive. Simulations often reveal deep similarities between systems that are superficially very different. Why is a starfish outbreak like an atomic bomb? Because both are examples of percolation processes, as also are epidemics, fire spread and a host of other phenomena. This underlying property of percolation, which they all share in common, becomes obvious in the act of building a simulation model.

Simulation models put deep similarities into sharp relief. The reason for this is that to simulate a system you have to represent it in the memory of a computer. When systems are represented in this way, certain kinds of models turn up again and again. Boolean networks, for instance, have been used to implement simulation models of systems as varied as switching circuits, genetic regulatory networks and social opinion. These sorts of coincidences highlight the similarities, and differences, between what would otherwise appear to be completely different systems.

What is more, there are evident similarities between different kinds of models. For instance a cellular automaton with two states is also a Boolean network; it just has a regular pattern of connections of a particular kind. So too are spin glass models, which are used to model the formation of atomic lattices, such as glasses. Similarities and patterns of this sort have played an important role in the development of complexity theory.

The idea of simulations as analogies raises the question of validity. That is, does the simulation model accurately reflect events in the real world? A large simulation model, for instance, could have literally hundreds of parameters, each of which needs to be represented correctly and calibrated accurately. There is also the problem of model complexity. A large simulation with many non-linear relationships is likely to behave in a highly disordered fashion and it can be difficult to tease out which relationships produce a given kind of behaviour. Taken together, these problems sometimes lead to a phenomenon known as "garbage in, garbage out."

Table 1. Differences between mathematical and informatics modelling paradigms

Aspect/use	***Mathematical***	***Informatics***
Model representation	Formulae	Simulation
Model development	Curve fitting	Network analysis
Model presentation	Graphical plots	Virtual reality
Data resources	Rare & expensive	Abundant & cheap
Planning	Forecasting	Scenarios
Control	Optimisation	Adaptation
Interpretation	Algebraic	Sensitivity, resilience

Many modellers have noted that because the behaviour of complex systems is often chaotic and unpredictable, the approaches used to model them are of necessity different from those used to model simple systems. A model that cannot predict the future behaviour of a chaotic system may still reveal useful insights. A simulation of an epidemic, for instance, may not be able to predict exactly where the disease will spread, but it could help to identify what action would be needed to contain the spread of the epidemic under various scenarios.

The question of a model's validity therefore reduces to its *adequacy* for a particular purpose. Simulation models, and especially agent based models, may be adequate for one application but not for another. In short-term forecasting, for instance, an auto-regressive time series model might prove adequate and would be usually be much simpler and quicker to build than a complex simulation. In general, the uses of simulation models can be categorized as explanation, prediction and control. *Explanatory models* are used to investigate how a system works. *Predictive models* are used for planning in the light of future behaviour of a system. *Control models* are used to manage systems by constraining them to behave like a desired model. The advantage of these distinctions is that they help to narrow down which features a simulation needs to include, and which can be safely ignored.

The need to cope with systems for which long-term forecasting is impossible has given rise to a simulation based methodology that has parallels with many of the traditional procedures associated with mathematical modelling (Table 1). Instead of a mathematical formula, you have a simulation as the model. Because the systems are often complex and inherently unpredictable, studying scenarios replaces forecasting. Sensitivity studies (including both resilience and robustness analysis), in which the modeller examines the effect of systematically varying parameters, replace mathematical analysis of algebraic structure. And optimization is likely to involve adaptive methods rather than (say) solving formulae for points of inflexion.

The world inside a computer?

In 1972, a group of scientists calling themselves the Club of Rome published the results of their "*Limits to Growth*" simulation model of world population, environment and economics [Meadows, *et al.*, 1972]. The model's dire predictions of looming environmental disaster startled people. The backlash was severe. Hostile critics picked at every detail of the model's assumptions, trying to invalidate its message. The issue has been hotly debated ever since.

Despite its controversial nature, or perhaps because of it, the Limits to Growth model made one point emphatically obvious: simulation models are capable of making important contributions about complex problems of global importance. The model also highlighted the idea of building entire virtual worlds inside a computer. In subsequent decades, models of global climate have become both the fashion and the focus of international debate about environmental sustainability.

Is it possible to build a world inside a computer? General Circulation Models (GCM), for instance, simulate the entire earth's surface, as well as several layers of the atmosphere and oceans [Randall, 2000]. Virtual reality goes even further. As every player of video games knows, you enter a virtual world in which you can move around and which responds to what you do. In Douglas Adams's novel "Hitchhikers Guide to the Galaxy," a race of aliens builds a giant computer that turns out to be the Earth. Simulation turns this on its head by putting the world, or at least parts of it, inside a computer. In the movie *The Matrix*, people spend their lives in a virtual world that is complete in every detail, while their bodies lie entombed in life support capsules.

The increasing power of computers makes the idea of the world inside a computer more and more seductive. Table 2 shows the size of various real world systems. As the table shows, the size and speed of modern computers compare with that of natural systems, but it is the complexity of natural systems that make them difficult to model. The table also shows that the memory capacities of large supercomputer clusters compare favourably with large systems. At the time of writing, the world's population is approximately 6.5 billion people. Large computers have many times this capacity in bytes. The human genome contains only 3.3 billion base pairs, and only some 30,000 genes. So in principle, it would be possible to represent the entire genome or the entire human race inside a computer.

However the real problem is that the systems are much more complex than numbers suggest. To simulate anything useful about people, for instance, you need to know something about them, about the way they behave, and about the ways they interact with one another. You may be able to record the entire genome in memory, but the functioning genome has many complications, such as the way complex protein molecules fold and unfold when their sub-regions become active.

The idea of building the world inside a computer finds expression in several areas and has given rise to entirely new fields of research. Artificial Life (Alife), for instance, is a growing field of research in which simulation models are used to investigate biological questions. One of the advantages of virtual creatures in vir-

Table 2. Comparative sizes of various systems

System	*"Size" (number of elements)*
Human body	$\sim 1 \times 10^{14}$ cells
Human brain	$\sim 4 \times 10^{10}$ neurons
Human genome	$\sim 3 \times 10^{9}$ base pairs
Earth	$\sim 8 \times 10^{6}$ sq. km
Supercomputers	$\sim 1 \times 10^{14}$ bytes *

* Note: The memory capacity of supercomputers continues to increase rapidly (TOP500, 2008).

tual worlds is that we are not restricted to the living things we see around us. Nor are we restricted to the world we live in. Alife research is free to investigate "life as it could be" [Langton, 1989]. In so doing Alife throws new light on living things in the real world. By exploring the possibilities for life in alternative scenarios it can illuminate the principles and processes underlying real living systems.

NATURAL COMPUTATION

Every era tends to look at the world in terms of its preoccupations. During the Industrial Revolution, a preoccupation with machinery was associated with a mechanistic view of the world. The human body, for instance, was seen as a machine: the heart was a pump, the digestive system provided the fuel supply, and so on. Likewise the solar system was like a vast clock, with planets wheeling in their orbits counting out the years, decades and millennia.

The Information Revolution led to a whole new set of preoccupations. From TVs and telephones to digital music, people are surrounded by information every day of their lives. Laptops and personal computers are in almost every home. Communication, data processing and information management have become vital commercial issues.

It is not surprising, then, that the Information Revolution has led people to see analogies with information processing all around them. The human body is replete with systems for information processing. The brain is a computer. The senses provide data input. Speech is a communication medium. And DNA is like a computer program, encoding the information necessary to build new life.

One of the most fruitful and influential ideas to arise in computer science is *natural computation*. This is the view that processes in nature can be viewed as forms of computation. Many of the analogies that we can draw about nature and computation do have a valid basis. The sequence of bases on DNA, for instance, really does act like a set of coded instructions on a tape. And there really are uncanny similarities with computing in the way the DNA code is processed. After the DNA code is copied onto RNA, ribosomes attach themselves to the RNA. As they move along the RNA strand reading the bases, they output one amino acid for each reading frame of three bases. The resulting strings of amino acids form the proteins that are the substance of our bodies.

The idea of natural computation has grown into a new scientific paradigm and has proved to be a rich source of new insights about nature. Many processes in nature exhibit key characteristics of computation, especially discrete units or steps and repetition according to a fixed set of rules. Although the processes may be highly complex, their regularity makes them highly amenable to simulation (e.g. Lindemayer's L-system models described earlier).

Modellers have drawn many fruitful analogies between living things and computing systems. Examples include gene expression and switching networks, disease outbreaks and computer viruses, social networks and telecommunication networks, stock markets and software agents, and intercellular signalling and multiprocessing arrays.

In biology, this new paradigm has led to a host of insights about biological systems, especially through the increasing use of simulation and "virtual experiments", giving rise to new fields of research, such as synthetic biology's creation of designed organisms optimised for drug production and other industrial functions as well as Artificial Life, mentioned earlier. In bioinformatics, the idea of encapsulation is seen in such ideas as controller genes and motifs [Bairoch, 1991; Hansen, 2003]; and the idea of networks finds expression in the idea of genetic regulatory networks [Bower and Bolouri, 2001; Boutanev, *et al.*, 2002]. In sociology, societies can be viewed as networks of interacting agents.

At the same time as the natural computation paradigm has influenced biology and other sciences, it has also influenced computing. As the complexity of computers and computational problems has increased, traditional ideas and approaches have often proved inadequate to cope. Computer science has therefore sought inspiration from nature, which has evolved numerous ways of coping with extreme complexity. The result is that modern computing often resembles biology. Nature has evolved many ways of dealing with complex phenomena and of solving many kinds of complex problems. Investigating these "solutions" has proved to be a fruitful source of new insights about the nature of complexity, and about ways of managing complex systems. This activity led to a host of new ideas in computing, as evidenced by the names of new fields of research, such as *artificial life* (mentioned above), *cellular automata*, *evolutionary computing*, *neural networks*, and *swarm intelligence*.

Computing by chance

Ever since Archimedes shouted "Eureka" in his bath, serendipity, accidental discovery, has played an important part in science. The ability of computers to process huge volumes of information has raised the role of serendipity from a rare bonus to a standard method of investigation [Green, 2004]. The field of "data mining", for instance, centres on the idea of searching for "nuggets" of knowledge hidden within mountains of information.

Another, less obvious, way that serendipity has influenced information science is in *Evolutionary Computing* (EC). Rather than solving a problem deductively, the idea is to evolve a solution by trying out randomly varying features and retaining the best varieties. In essence this approach relies on a succession of lucky accidents that yield better and better varieties. One of the most active areas of research inspired by nature, Evolutionary Computing is a generic term for a large range of problem-solving methods that go by many names, mostly based on differences in their approach. One of the earliest approaches was to adapt finite state automata so as to improve their performance when solving some problem [Fogel *et al.*, 1966]. However, the most widely used and best known is the *genetic algorithm* (GA) [Holland, 1975]. GAs turn a problem into a set of "genes" that each represents some parameter, feature or assumption. You then breed populations of pseudo-organisms possessing those genes, choosing the "fittest" organisms in each generation from which to breed the next generation. The "fitness" represents an overall feature or function that you want to optimize. There are many variants, including methods to overcome technical difficulties (e.g. [Kirley, *et al.*, 1998]) and applications, such as genetic programming [Koza, 1992], to procedural problems.

As optimization methods, GAs combine aspects of both local and global search[2], but are not guaranteed to find an optimum and may be slow. There is a tendency among users to apply them in *ad hoc* fashion to any problem, without regard to its nature or the availability of faster, more suitable methods. On the other hand, they make it possible to find answers to problems, even highly complex problems, without making simplifying assumptions.

Along with different methods, natural computation also embodies different attitudes towards problem solving. In optimization, for instance, the traditional approach is to find the very best possible solution. However, natural systems usually occur in complex environments within which there are many competing issues. So living things typically need to solve multi-objective optimization problems. These usually involve trade-offs: a gain on one objective involves a loss on others. In other words, you cannot maximize everything, so instead you try to find a solution that is *adequate* for your needs. Computational methods such as genetic algorithms (GA) embody this idea. GAs cannot be guaranteed to find the absolute maximum, even for a simple function, but they are good at finding adequate solutions, even for very complex problems.

Thinking in crowds?

In the late 1980s, Rodney Brooks [1986; 1991] initiated an important change in direction in robotics by showing that intelligence, or at least intelligent behaviour, did not need to be the product of a central "brain." Instead it could emerge out

[2] A good analogy is trying to find the point of greatest elevation in a landscape. In *local search* (also called *exploitation*), you try to find the top of the hill you are on. In *global search* (also called *exploration*), you try to find the tallest hill.

of interactions amongst many simple and distributed processes. Brooks showed, for instance that when an insect walks, the action emerges out of feedback that coordinates simple motions by the individual legs. If a leg is pointing backwards, then it needs to lift and move forward. Simple responses such as this do not need to be referred back to a central "brain;" they can be controlled locally. This kind of decentralized control is another case of modular processing.

The idea that intelligent behaviour can emerge out of a complex network of interactions amongst many simple agents has since been extended and generalized from robots to many other systems. The field of swarm intelligence, for instance, deals with the ways in which useful global behaviour emerges out of simple local rules of interaction between large assemblages of agents. A flock of birds emerges when flying birds coordinate their direction and distance from their neighbours. An ant colony emerges because individual ants obey simple rules such as pick up any scrap you find and place with any other scrap you find. Computer science has adapted such examples to create useful algorithms, such as the *Ant Sort Algorithm*, that create order in dynamically changing systems. The idea of swarm intelligence is important for emerging new areas of computing. In nanotechnology, for instance, there is a need to understand what behaviour would emerge when multitudes of simple nanobots, each having limited "thinking power", are set to work on a problem.

Another approach to distributed intelligence is to try to mimic brain operation at the level of neurons. Rosenblatt [1958] introduced the idea of an *Artificial Neural Network (ANN).* This consists of layers of simple processors that pass information from layer to layer. In a simple feed-forward network, an input layer receives data from the outside and passes on processed information to successive internal "hidden" layers until some result emerges from an output layer.

An advantage of ANNs is that they can emulate a wide variety of functions without prior knowledge of their form. Instead of setting up a function of a particular form, ANNs are trained by presenting cases to them. This flexibility has led to many applications. On the other hand, the training process also contributes to some of the shortcomings of ANNs.

The greatest of these is the inability of neural networks to model systems with a very rich range of behaviour. This problem led researchers to investigate modular neural networks. These are ANNs in which some "nodes" in the network are themselves ANNs that perform some part of the behaviour.

Although inspired by biology, ANNs behave differently from living brains. Artificial neural networks typically converge to particular behaviour. In contrast, Freeman [1992] showed that living neural systems (e.g. cat's brains) are prone to respond chaotically. Freeman suggested that chaos provides a source of novelty which allows the brain to make fine discriminations between stimuli.

GRAND CHALLENGES AND THE FUTURE

Big questions get big answers. Major scientific endeavours, or "grand challenges", have long been important sources of inspiration and motivation in computing research. Challenges, such as global climate models and mapping the human genome were large-scale research programmes that sought to model huge systems of great complexity. They could not have succeeded without high performance computers.

Many organizations have proposed grand challenges as ways of focussing effort on complex problems of great significance, scale, timeliness and potential impact. In some cases they provide a way of capitalizing on advances in computing power. In others, solving the problem itself expands the limits of computer science and computing technology. One of the most notable sources of grand challenges is the UK Computing Research Committee, for which posing grand challenges is a major on-going activity. By 2004, their list included seven ambitious challenges [UKCRC, 2008]. For instance, the first challenge on the list, titled *In vivo-in silico,* is typically ambitious. It aims to produce high fidelity models that replicate the development and behaviour of simple plants and animals *in all their detail.* As we saw earlier (Table 2), challenges such as this not only require great computing power, but also a deep understanding of the complexities of the organisms involved.

It is likely that practical challenges such as these will continue to influence advances in computing and complexity science well into the future. As computers continue to evolve and spread, they become more like the systems they are built to serve. And as systems become more complex, so too computing theory becomes increasingly bound up with complex theory. At the present time this trend seems set to continue indefinitely.

The importance of challenge problems is that solving them almost invariably leads to new theories or fields of study. For instance, the 23 challenge problems posed by David Hilbert in 1900 profoundly influenced the course of mathematical research during the 20^{th} Century. Perhaps the most important discovery to arise from them, as far as computing is concerned, was Gödel's disproof of Hilbert's second problem, which was to prove that the axioms of arithmetic are logically consistent. Gödel's *Incompleteness Theorem* [Gödel, 1931] states that in any formal system there are always statements that can neither be proved nor disproved. This result leads almost directly to Turing's *Halting Problem.* If a computer can solve a problem, then it will halt after a finite number of steps. But can you know how long it will take, or even if it will ever halt? This problem in turn inspired research on computational complexity, which we discussed earlier.

New models of computation

The traditional abstract model of a computer is an automaton [Hopcroft and Ullman, 1969], a mathematical formalization of the original Turing Machine. An *automaton* is a device with inputs, outputs, internal states and a program that dictates how the object responds to combinations of states and inputs. Over several decades, other ideas have been introduced that effectively extend this model to deal with the complexities encountered in modern computing.

Many of these new ideas concern the distribution and organization of information and computation amongst many elements. First, the notion of objects has become formalized in computer science. *Objects* are information structures that possess "attributes" and "methods", some of which are internal to the object ("hidden"), while others are visible ("public") to outside. Ole-Johan Dahl and Kristen Nygaard [1966] designed the first programming language (SIMULA) that employed objects explicitly. *Object-oriented programming* has since become standard in software development, being embodied in widely used languages, such as C++ and Java.

Object-oriented methods constitute a useful contribution of computer science to the practical management of complexity. They not only apply to information but also to the analysis and modelling of many kinds of complex systems. The *Uniform Modelling Language*, for instance, provides formal methods for designing information systems around classes of objects [Larman, 2005]. Classes of object combine to form networks by sharing attributes or methods. They also combine to form larger classes of objects in "gen-spec" and "whole-part" hierarchies[3].

Besides providing practical models for analysing complex systems, the concept of networks of objects also provides models for emergent processing by networks of simple elements. Artificial neural networks, described earlier, provide one such model. Other models include cellular automata and agent-based models.

In all of the above models, the processing elements, or "agents", are tightly constrained. The connections between elements are normally pre-defined and fixed. Agent-based models often occupy intermediate positions with intra-agent relationships allowed some latitude to vary according to social rules of varying looseness while swarm intelligence, introduced earlier, provides an extreme: the individual processing elements or "agents" are only loosely connected and their interactions are intermittent and ever-changing.

Cellular Automata and other multi-agent systems

Cellular automata (CAs) are arrays of cells (1D, 2D or 3D) in which each cell is an identically programmed automaton. The idea was first proposed by John von Neumann, but became popular after John Conway introduced the Game of Life [Gardner, 1970]. The most important feature of a CA is that automata in the array interact with their neighbours. State changes in a cell depend not only on its current state, but also on the states of neighbouring cells. This feature has made CAs a popular vehicle for studying the emergence of properties in complex systems.

[3]We can use books to illustrate these ideas. The object classes "publication", "book", and "novel" form a general-special hierarchy. A book is a *whole* whose *parts* include title page, chapters, and index.

Besides CAs, the idea of information processing in networks appears in many different guises in computing. Often the same idea has arisen in different fields under different names and with different questions. Boolean networks, for instance, consist of automata that have a single state that takes only the two values ON or OFF (or TRUE/FALSE or 1/0). They correspond to switching circuits in engineering and have been used as models of genetic self-regulation [Kauffman, 1991] and the emergence of social consensus [Green and Bransden, 2006]. In essence they are also closely related to spin glasses, represented in the Ising [McCoy, Wu, 1973] and Potts models [Potts, 1952]. These are models of magnetic material in which the components (usually called sites) prefer to align in the same direction with some of their neighbours, and in the opposite direction with others. This frustrates the system from reaching minimal energy, leading to a rugged energy landscape over the state space.

TOWARDS A SCIENCE OF COMPLEXITY

One of the barriers hindering development of a coherent theory of complexity is the fragmented nature of research on the topic. Different fields of research approach complex phenomena independently. Each field has developed different models and different methods, with different terminologies and different insights. For example, computation considers the richness of an automaton's behaviour; the network theory of complexity considers the richness of connections in networks (e.g. combinatorial complexity), and nonlinear dynamics deals (amongst other things) with the richness of variety in individual relationships, e.g. those that generate chaotic systems.

A crucial question is whether different kinds of complexity, such as those mentioned above, are truly different? Or are they manifestations of deeper underlying principles? Some aspects of different approaches are relatively easy to reconcile. For instance, the variables associated with a dynamic system can be interpreted as forming the nodes of a network. The edges of the network are defined by interaction coefficients in the equations that define the behaviour of the system. Likewise, any network can be represented as a matrix, where each node has a corresponding row and column, and the edges are represented by non-zero entries in the appropriate cells (see Figure 2).

Universality of the network model

As mentioned earlier, mappings such as the above show that networks are inherent in the structure of all complex systems [Green, 1993; 2000]. An important result of the natural computation paradigm has been to stress the potential of the network model to reconcile different complexity paradigms. In doing so, it raises the prospect of a unified science of complexity.

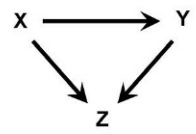
Instead of regarding "computation" as a set of recipes for building glorified calculating machines, we can regard it as an abstract model for processes of all

$$\frac{dx}{dt} = Ax$$
$$\frac{dy}{dt} = Bx + Cy$$
$$\frac{dz}{dt} = Dx + Ey + Gz$$

(a)

$$\begin{pmatrix} A & 0 & 0 \\ B & C & 0 \\ D & E & F \end{pmatrix}$$

(b)



(c)

Figure 2. (a) A system of equations in three variabls X, Y and Z. (b) The coefficient matrix derived from (a). (c) The network corresponding to (b).

kinds. We can see this in the way that the network model of complexity provides a new interpretation of computation, as well as many other processes. Besides representing organization and structure, networks can also represent processes [Green, 1993; 2000]. In a game of chess, for instance, the state of the game at any stage is given by the arrangement of pieces on the board. Each move made by the players shifts the board from one state to another. So if we take each arrangement of the board to be a node, then each move is an edge joining pairs of nodes. Represented in this way, the game becomes a pathway winding through a giant network of all possible moves and arrangements.[4] By extension, we can treat any process that proceeds in discrete steps as computation.

In similar fashion, the state space of a computer program can be viewed as a network in which states of the program (values of variables etc) are the nodes, and transitions between states provide the edges (Figure 3). This equivalence has many implications. First it suggests a different view of computation itself. For a dynamic system, computation is any connected path through the set of all possible states of the system. This shows that the dynamics of any system defines an automaton. Likewise, the behaviour of an automaton can be seen as a network in which its states are nodes in a network of possible states and the transitions between states define its edges.

Mappings between paradigms, such as those described above, make it possible to reinterpret well-known phenomena in new ways. For instance, all three approaches mentioned earlier identify phase transitions separated by critical regions. Chaos, for instance, can be seen as extreme richness in the behaviour of an automaton [Langton, 1990]. And richness in an automaton's behaviour can be seen as high density of connections within the state space. These two observations together imply that non-linear interactions, and chaos in particular, amount to well-connected networks of states. Considerations such as these have provided numerous insights. One implication is that period doubling, observed in the transition to chaos, is

[4]There are 20 possible opening moves in chess, and 20 possible responses. So after a single move by both players, there are 400 possible arrangements of pieces on the board. After two moves, the number of possible states balloons out to over 160,000 possible states and continues to escalate rapidly. The game of Go is even more extreme, with over 130,321 possible states after a single move by both players.

related to the connectivity avalanche that occurs in networks. Another is whether there is a link between computability and complexity.

Consequences of the network model

The universal nature of networks led researchers to explore the structure of state spaces for computational systems. Much attention focussed on the behaviour of cellular automata (CAs). For instance, Wuensche pointed out that many CAs have Garden of Eden states — states that cannot be reached from any other state. In other words these states are not computable [Wuensche and Lesser, 1992; Wuensche, 1994].

Another crucial line of studies linked the density of connections in a CA's state space (e.g. [Langton, 1990]) with different classes of CA behaviour (e.g. [Wolfram, 1984; 1986]). Given that the state space of an automaton is a network, a phase change in connectivity will occur as the number of connections between states increases. This phase change is associated with different kinds of behaviour. If the state space is poorly connected, then the system will quickly "freeze" into a simple attractor — i.e. a constant state or cycle. In contrast, an automaton with a richly connected state space may exhibit chaotic, gas-like behaviour.

Considerations such as these led Langton [1990] to focus attention on the critical region between simple and chaotic behaviour. In this critical region, which Langton called the *Edge of Chaos*, are found automata that exhibit the most interesting behaviour. Langton and others speculated that universal computation occurs only in systems whose state spaces sit within this critical region.

One immediate inference is that the critical region should play an important role in many kinds of processes. The idea provided theoretical support for the theory of *self-organized criticality* (SOC), which was developed independently (e.g. [Bak and Chen, 1988]). SOC postulates that complex systems evolve into a critical state and stay there. In disturbed systems, SOC is characterized by $1/f$ noise. That is, the frequency f of disturbances is inversely proportional to their magnitude.

Another mechanism, *Dual Phase Evolution* is also associated with phase changes in a state space [Green, *et al.*, 2006]. In dual phase evolution, systems evolve by flipping between the connected and disconnected phases. Outside influences (e.g. disturbances) trigger phase shifts and each phase is dominated by different processes (e.g. selection versus variation). Many optimization algorithms exploit phase changes in mediating between global and local search. Simulated annealing, for instance, mimics the cooling of metal. In annealing, heating metals allows the atoms considerable freedom of movement within an atomic lattice. As the metal cools, the atoms are increasingly constrained until their positions in the lattice become fixed. Likewise in simulated annealing, parameters of the solution are at first allowed to change freely, so the solution can jump wildly about in the "fitness landscape". However, the "cooling schedule" gradually constricts changes in parameter values, until the solution becomes trapped on a single fitness "peak."

New paradigms

Two essential requirements for a field of science are a coherent body of theory and a set of standard methodologies. There are good grounds for suggesting that the idea of computation in networks of many simple elements constitutes both the essence of modern complexity theory and an important new paradigm for understanding complex systems.

First, as we saw earlier, the idea of networks is fundamental in complexity theory. All systems have underlying networks in their structure and behaviour [Green, 1993; 2000]. Many properties of networks, such as connectivity phase changes and topologies, are now well-known and account for complex phenomena in a wide range of systems.

Secondly, techniques such as object modelling and simulation provide methods for representing and studying complex networks. Models of computation in networks (e.g. cellular automata) provide well-established representations with well-known properties. More generally, a body of methods has grown up around simulation modelling to make it a rigorous investigative technique. As we saw earlier, Table 1 summarizes some of the ways in which the emerging informatics paradigm differs from traditional mathematical modelling.

Emergence and reproducibility

A final example of the merging of computing and complexity theory is the issue of self-reference and its implications. A statement such as

"*This sentence is written in English*"

is *self-referential*, it refers to itself. In practical applications, formal languages are normally *first order*, that is they are incapable of self-reference. Mathematical logic has long grappled with the complex issues that self-reference leads to in second order languages.

The problem of self-reference extends into issues of biological complexity. During the 19^{th} Century, biologists struggled with the self-replication problem in trying to understand human reproduction. How do organisms produce copies of themselves? Attempts to answer the question led to the idea of the "homunculus." The suggestion was that eggs and sperm contained *homunculi*, tiny bodies that could grow into complete babies. However, the idea immediately led to a paradox: each homunculus would need to contain an infinite series of smaller homunculi, nested like Russian dolls, to reproduce future generations.

Self-replication is a central question in complexity theory. It is a crucial issue, for instance, in trying to understand self-organization and emergence. How can a part of a system contain a complete copy of the entire system? The paradox is resolved by observing that copying does not require a complete blueprint. Just as an ant colony emerges from the interaction of ants with their environment, so a complex system can emerge from information embodied in the environment that does the copying.

Von Neumann [1966] suggested that to be capable of self-replication, automata needed to exceed some threshold of complexity. However, Lionel Penrose [1959] demonstrated that under suitable conditions, even very simple structures can self-replicate. In his Tierra model, Tom Ray [1991] produced a virtual universe in which automata could not only replicate themselves, but also evolve into a multitude of forms. Some even evolved to become pseudo-viruses, replicating themselves by exploiting the reproductive code of other "creatures".

The above ideas find practical application in efforts to understand the human genome. When the *Human Genome Project* set out to map the entire human genome in the 1990s, biologists predicted that the number of genes could be as high as 2 million. Most people were therefore surprised to learn that the genome contains less than 25,000 genes [ORNL, 2008]. Their surprise arises from a simplistic view of gene function: the assumption that one gene codes for one attribute. Studies of Boolean networks had already shown that even relatively small sets of genetic "switches" can combine in myriads of ways and are capable of highly complex behaviour [Kauffman, 1991]. The years immediately following completion of the Human Genome Project therefore saw an explosion of research on genetic regulatory networks. Computational ideas such as modularity are finding expression in studies of genetic clustering (e.g. [Boutanaev, *et al.*, 2002]) and its role in evolution (e.g. [Hartwell, *et al.*, 1990; Hansen, 2003]).

This shift marks a dawning interest of biologists in the emergent properties of complex networks. It parallels similar shifts in other fields, such as engineering and economics, which likewise deal increasingly with interactions within large, complex networks. This upsurge of interest in many different fields has helped to stimulate the convergence, described above, of ideas in computing and complexity theory. It provides strong motivation for developing a science of complexity. Key topics, such as computation within complex networks of agents, are likely to become leading research areas during the Twenty-First Century.

BIBLIOGRAPHY

[Bairoch, 1991] A. Bairoch. PROSITE: a dictionary of sites and patterns in proteins. *Nucleic Acids Research*, 19, 2241-2245, 1991. http://kr.expasy.org/prosite/.

[Bak and Chen, 1988] P. Bak and K. Chen. Self-organized criticality. *Physical Review* A 38, 364-374, 1988.

[Bilofsky and Burks, 1988] H. S. Bilofsky and C. Burks. The GenBank genetic sequence data bank. *Nucl. Acids Res.*, 16, 1861-1863, 1988.

[Boutanaev, *et al.*, 2002] A. M. Boutanaev, A. I. Kalmykova, Y. Y. Shevelyov, and D. I. Nurminsky. Large clusters of co-expressed genes in the *Drosophila* genome. *Nature* 420, 666-669, 2002.

[Bower and Bolouri, 2001] J. M. Bower and H. Bolouri, eds. *Computational modeling of genetic and biochemical networks*, Computational Molecular Biology Series, MIT Press, Cambridge, Mass., 2001.

[Bransden and Green, 2005] T. G. Bransden and D. G. Green. Getting along with your neighbours - emergent cooperation in networks of adaptive agents. In A. Ohuchi, K. Suzuki, M. Gen, D. G. Green (eds.), *Workshop on Intelligent and Evolutionary Systems* (IES2005), Future University, Hakodate, Japan. http://www.waseda.jp/sem-genlab/~IES2005/.

[Brooks, 1986] R. A. Brooks. *Achieving artificial intelligence through building robots.* AI Memo 899, MIT, Cambridge, Mass., 1986.

[Brooks, 1991] R. A. Brooks. Intelligence without representation. *Artificial Intelligence* 47, 139-159, 1991.

[Chaitin, 1966] G. J. Chaitin. On the lengths of programs for computing binary sequences. *J. Assoc. Compuing Machinery*, 13, 547-569, 1966.

[Cornforth *et al.*, 2005] D. Cornforth, D. G. Green, D. Newth, and M.R . Kirley. Ordered asynchronous processes in multi-agent systems. *Physica D: Nonlinear Phenomena*, 204, 70-82, 2005.

[Dahl and Nygaard, 1966] O-J Dahl and K. Nygaard. SIMULA - An ALGOL-based simulation language. *Communications of ACMI*, 9(9), 671-678, 1966.

[Dowe, 2008] D. L. Dowe. Foreword re C. S. Wallace. *Computer Journal* 51(5), 523-560, 2008.

[Fogel *et al.*, 1966] L. J. Fogel, A. J. Owens, and M. J. Walsh. *Artificial intelligence through simulated evolution.* J. Wiley & Sons, London, 1966.

[Freeman, 1992] W. J. Freeman. Tutorial on neurobiology: from single neurons to brain chaos. *International Journal of Bifurcation and Chaos*, 2(3), 451-482, 1992.

[Gardner, 1970] M. Gardner. Mathematical Games: The fantastic combinations of John Conway's new solitaire game "Life". *Scientific American* **223**, 120–123, 1970.

[Gödel, 1931] K. Gödel. Über formal unentscheidbare sätze der principia mathematica und verwandter systeme, I. *Monatshefte für Mathematik und Physik*, 38: 173-98, 1931.

[Green, 1993] D. G. Green. Emergent behaviour in biological systems. In D. G. Green and T. J. Bossomaier, eds. *Complex systems - from biology to computation*, IOS Press, Amsterdam, 25-36, 1993.

[Green, 2000] D. G. Green. Self-organization in complex systems. In D. G. Green and T. J. Bossomaier, eds. *Complex systems*, Cambridge University Press, Cambridge, 7-41, 2000.

[Green, 2004] D. G. Green. *The serendipity machine - a voyage of discovery through the unexpected world of computers.* Allen and Unwin, Sydney, 2004. `http://www.csse.monash.edu.au/~dgreen/books/serendipity/`.

[Green and Bransden, 2006] D. G. Green and T. G. Bransden. Complexity theory. In *McGraw-Hill Encyclopedia of science and technology*, McGraw-Hill, New York. 507-511, 2006.

[Green *et al.*, 2006] D. G. Green, T. G. Leishman, and S. Sadedin. Dual phase evolution – a mechanism for self-organization in complex systems. *InterJournal*, New England Complex Systems Institute, Boston, 1-8, 2006.

[Hansen, 2003] T. F. Hansen. Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolution. *Biosystems*, 69, 83-94, 2003.

[Hartwell *et al.*, 1999] L. H. Hartwell, J. J. Hopfield, S. Leibler, and A. W. Murray. From molecular to modular cell biology. *Nature*, 402, C47-C52, 1999.

[Hillis, 1982] D. Hillis. New computer architectures and their relationship to physics or Why CS is no good. *International Journal of Theoretical Physics* 21(3/4), 255-262, 1982.

[Holland, 1975] J. Holland. *Adaptation in natural and artificial systems.* University of Michigan Press, Ann Arbor, 1975.

[Hopcroft and Ullman, 1969] J. E. Hopcroft and J. D. Ullman. *Formal languages and their relation to automata.* Addison-Wesley, Reading, MA., 1969.

[Kauffman, 1991] S. A. Kauffman. Antichaos and adaptation, *Scientific American* 265(2), 64-70, 1991.

[Kirley *et al.*, 1998] M. Kirley, X. Li, and D.G. Green. Investigation of a cellular genetic algorithm that mimics evolution in a landscape. In SEAL98, *Proceedings of the 2nd Conferenec on Simulated evolution and Learning.* X.Yao, R. McKay, C. Newton, J-H. Kim, T. Furusashi (eds.), University of New South Wales, Canberra, 93-100, 1998.

[Kolmogorov, 1965] A. N. Kolmogorov. Three approaches to the quantitative definition of information. *Proc. Inform. Transmission* 1, 4-7, 1965.

[Koza, 1992] J. R. Koza. *Genetic programming.* MIT Press, Cambridge, Mass., 1992.

[Langton, 1989] C. G. Langton, ed. *Artificial life.* Addison-Wesley, Reading, Mass., 1989.

[Langton, 1990] C. G. Langton. Computation at the edge of chaos: phase transitions and emergent computation. *Physica D* 42(1-3), 12-37, 1990.

[Larman, 2005] C. Larman. *Applying UML and patterns - introduction to OOA/D & iterative development*, 3^{rd} edn. Prentice Hall PTR, N.J., 2005.

[Lindenmayer, 1968] A. Lindenmayer. Mathematical models for cellular interaction in development I. Filaments with one-sided inputs. *Journal of Theoretical Biology* 18, 280-289, 1968.

[Lorenz, 1963] E. N. Lorenz. Deterministic non-periodic flow. *Journal of the Atmospheric Sciences* 20, 130–141, 1963.

[Mandelbrot, 1982] B. Mandelbrot. *The fractal geometry of nature*. W. H.Freeman, San Fancisco, 1982.

[May and Martin, 1975] R. M. May and B. Martin. Voting models incorporating interactions between voters. *Public Choice* 22(1), 37-53, 1975.

[McCoy and Wu, 1973] B. M. McCoy and T. T. Wu. *The two-dimensional ising model*. Harvard University Press, Cambridge, Mass., 1973.

[Meadows *et al.*, 1972] D. H. Meadows, D. L. Meadows, J. Randers, and W. W. Behrens. *The limits to growth*. New York, Universe Books, 1972.

[NCBI, 2007] National Centre for Bioinformatics. *Genbank*. `http://www.ncbi.nlm.nih.gov/`, 2007.

[Netcraft, 2007] `http://www.boutell.com/newfaq/misc/sizeofweb.html`, 2007.

[ORNL, 2008] How many genes are in the human genome? *Human Genome Project*. `http://www.ornl.gov/sci/techresources/Human_Genome/faq/genenumber.shtml`, 2008.

[Penrose, 1959] L. S. Penrose. Self-reproducing machines. *Scientific American* 200 (6), 105-114, 1959.

[Potts, 1952] R. B. Potts. Some generalized order-disorder transformations. *Proceedings of the Cambridge Philosophical Society* **48**, 106-109, 1952.

[Randall, 2000] D. A. Randall, ed. *General circulation model development – past, present and future*. Academic Press, San Diego, 2000.

[Ray, 1991] T. S. Ray. An approach to the synthesis of life. In C. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, eds. *Artificial Life II*. Santa Fe Institute Studies in the Science of Complexity, 11, 371-408, Addison-Wesley, Redwood City, California, 1991.

[Rosenblatt, 1958] F. Rosenblatt. The perceptron, a probabilistic mode for information, organization and storage in the brain. *Psychological Review* **65**, 368-408, 1958.

[TOP500, 2008] TOP500 Supercomputer sites. `http://www.top500.org/`, 2008.

[Ueda, 1961/2001] Y. Ueda. Recounted in R. H. Abraham and Y. Ueda, eds. *The Chaos Avant-Garde: Memoirs of the Early Days of Chaos Theory*. World Scientific Publishing Co., Singapore, 2001.

[UKCRC, 2008] UK Computing Research Committee. *Grand challenges in e-Science*. `http://www.ukcrc.org.uk/grand_challenges/index.cfm`, 2008.

[von Neumann, 1966] J. von Neumann. *The theory of self-reproducing automata*, A. Burks (ed.). University of Illinois Press, Urbana, Illinois, 1966.

[Wallace and Boulton, 1968] C. S. Wallace and D. M. Boulton. An information measure for classification. *Computer Journal* 11(2), 185-194, 1968.

[Wallace, 2005] C. S. Wallace. *Statistical and inductive inference by minimum message length*. In M. Jordan, J. Kleinberg, and B. Scholkopf, eds. Springer Science+Business Media, Inc., New York, 2005.

[Wolfram, 1984] S. Wolfram. Cellular automata as models of complexity. *Nature*, 331, 419, 1984.

[Wolfram, 1986] S. Wolfram. *Theory and applications of cellular automata*. World Scientific, Singapore, 1986.

[Wolfram, 2002] S. Wolfram. *A new kind of science*. Wolfram Media Inc., 2002. `http://www.wolframscience.com/`

[Wuensche and Lesser, 1992] A. Wuensche and M. J. Lesser. *The global dynamics of cellular automata; an atlas of basin of attraction fields of one-dimensional cellular automata*. Addison-Wesley, Reading MA., 1992.

[Wuensche, 1994] A. Wuensche. *Complexity in one-d cellular automata: gliders, basin of attraction and the z parameter*. Santa Fe Institute working paper 94-04-025, 1994.

EVOLUTIONARY GAMES AND THE MODELING OF COMPLEX SYSTEMS

William Harms

The human "spirit", which can be taken to include morality, reason, and the peculiar non-causal pointing relationship between thoughts and things-in-the-world, has been the one element most difficult to integrate into the scientific world-view. No doubt this is due in part to the ongoing turf war between science and religion, but it is also due to the peculiarity of the supposed "rules" of reason and ethics, and of the intentionality of thought and language as seen from scientific materialism. Historically, materialists have tended to emphasize this difficulty, rejecting objective morality and non-causal mental powers as vestigial superstitions, relics compatible with faith but not with science.

Over the last several decades, the evolutionary modeling of human behavior and language has revealed a very different picture than the received wisdom of informal intuitions and philosophical tradition. As we develop the tools necessary for a realistic model of the full complexities of human interdependency, it is beginning to appear that evolution makes us good and forges the bonds between thoughts and things, that random variation and natural selection is wiser than even the most idealized rationality, and that most things are the way they are not simply because history has left them that way, but because they inhabit equilibria of dynamical processes that cannot even be understood without mathematical and computational aids. The world is complex and non-linear, and an adequate understanding must involve tools that share those properties. Evolutionary Game Theory is probably the most common kind of such modeling, merging seamlessly with models from ecology and evolutionary theory proper, and finding participants from numerous disciplines. To some extent, the most persuasive evidence that "the naturalistic turn" was a good idea has come from this kind of modeling. Nonetheless, if we are to extend this success to models fully adequate for modeling cultural change, further development is needed.

What evolutionary game theory is

Game theory is the formal study of strategic interactions — situations in which the best choice or action depends on the choices or actions of others. In its original formulation, it focused on the choices of hypothetical "ideally rational agents" who can be counted on to make the best choices to maximize their payoffs, without limits on intelligence or the intrusion of emotional biases [von Neuman, Morgenstern, 1947]. Its primary method of analysis is the inspection of game payoff matrices

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

for equilibria, choice combinations from which no player can benefit by unilateral deviation (the so-called "Nash" equilibrium). It achieved considerable notoriety as a think-tank tool for analyzing strategies of deterrence during the cold war. Of more philosophical interest has been the gradual discovery that apparently simple coordination problems like the celebrated Prisoner's Dilemma can only be solved by Game Theory's ideally rational agents with some artificiality. This has led to considerable elaboration on the Nash equilibrium concept, as well as to more sophisticated characterizations of strategic situations.

Evolutionary Game Theory (EGT) takes Game Theory's schematic numerical representations of strategic interactions and interprets them as payoffs driving the propagation of strategies in biological or cultural evolution, rather than as anticipatory incentives for ideally rational agents. Virtually any technique of evolutionary analysis can and has been applied to this end. These techniques include the Replicator Dynamics — a derivative of R.A. Fisher's dynamics for population genetics — in both continuous and discrete forms and implemented both computationally and as systems of differential equations [Hofbauer, Sigmund, 1988]; as well as the famous round-robin computer tournaments run by Robert Axelrod [1984]; cellular automata; the analysis of evolutionarily stable strategies [Maynard Smith, 1982]; and models of gene/culture co-evolution [Boyd and Richerson, 1985]. Drawing theorists from numerous disciplines, its most common subject has been the evolution of cooperation, altruism and ethics.

It will be helpful to distinguish two fundamentally different kinds of evolutionary modeling: population (type-frequency) modeling and agent-based modeling. What they have in common is that both are intended as tools for understanding evolutionary change in a rigorous way. They differ greatly in the specifics.

Population models present an abstract representation of the relationship between a population and its environment, driven deterministically by mathematically expressed dynamical laws. The essential components of a population model are a list (or vector) of the relative frequencies of various types within a population, and another list of the "fitnesses" of the various types. Natural selection is simulated by multiplying the frequency of each type by its current fitness, and "normalizing" the frequencies so they all add up to one. If p_i is the frequency of type i and w_i is the fitness of type i, then the new frequency of each type (p_i') is:

$$p_i' = p_i \frac{w_i}{\sum\limits_j p_j w_j}$$

Typical elaboration consists of having the fitness (w) of each type be dependent on the frequencies (p) of other types, e.g. via one of game theory's payoff matrices, or of simulating mutation and immigration by adding some small value to a type's frequency. In the following frequency dependent fitness equation, the w_{ij} are just the payoffs of i playing against j from (e.g.) a game's payoff matrix.[1]

[1]It is usual to make this more manageable by noticing that the normalizing denominator in the last equation is just mean fitness.

$$p'_i = p_i \frac{\sum p_j w_{ij}}{\overline{w}}$$

In these population-vector models, the difference between selection and variation is simply that selection multiplies, and variation adds to, a type's frequency. The following is a simple uniform mutation equation. If n is the number of types, and m is a mutation rate between zero and one,

$$p'_i = p_i(1-m) + \frac{m}{n}.$$

One of the advantages of population models is that they can be studied both computationally and mathematically, and therefore important discoveries can be expressed as mathematical relationships. The downside is that assumptions sometimes get introduced into these models more for the purpose of tractability than realism.

Agent-based models, on the other hand, create analogous "toy systems" which consist of schematic individuals which interact with a schematic environment (and each other) through a series of computational events. The most familiar of these is the cellular automaton, which considers a fixed number of individuals arranged on a grid, each interacting only with its immediate neighbors. Such models are useful for studying cultural propagation of strategies and opinions, where the number and location of individuals remains fixed and simply alter their properties in response to the influence of neighbors. It is also possible to build agent-based models with mobile agents, variable population size, and so forth. Unlike population models, the specifics of agent-based models are unlimited by prior mathematical specifications of the system, so one can really build whatever features and complexity one wants into them. One common difficulty is knowing which of the plethora of features added for the sake of realism is responsible for an interesting result. Presumably an adequate modeling methodology will require both sorts of models.

The methodological importance of the shift from rational-choice Game Theory to evolutionary Game Theory should not be underestimated. At first glance the familiar payoff matrices are still there, so one might imagine that we are simply studying the same strategic situations with new tools, and this is true to some extent. However, an important shift has occurred in *what* is being modeled. Rational-choice games are modeling an idealized choice process, that is, a process in which the fundamental units are a rational agent's representations of a strategic situation, and the best choice is subject to the agent's epistemic situation. In evolutionary models, on the other hand, what is being modeled is not an agent's representation of anything, but rather the net effect of actual physical processes on the physical persistence and propagation of a strategy. In short, rational-choice games model thinking while evolutionary games model physical processes. This constitutes a major shift in the direction of "naturalism", the turning over of philosophical questions to materialist scientific methods.

Nonetheless, as a methodological shift, one might argue that it does not go far enough. For while it may be true that the game matrix is now interpreted as a summary of the balance of physical forces, nonetheless, it is just that — a summary. The payoffs that appear in the cells of the matrix can not be interpreted directly as physical quantities, but are rather averages of the effects that the presence of one strategy or type has on another. This is much like the net force vector on a Newtonian free-body resulting from a combination of other forces like gravity, momentum, and resistance. Moreover, while the matrices may holistically generate non-linear dynamics, the effect of individual payoff terms are computed linearly, limiting the complexity of the effect of one type on another. Whether this criticism is telling depends to some extent on one's goals for a modeling framework — do any differences in outcome made by neglected non-linear interactions matter to the purpose, whether the purpose is pragmatic or the cognitive purpose of obtaining a complete representation of the target system? We will return to this question later.

Types of non-linearities in evolutionary game theory

Understanding the complexities of the evolution of human culture and behavior requires formal models which share those complexities. Evolutionary Game Theory's models are generally non-linear and frequently behave counter-intuitively, for a variety of reasons, thus providing evidence that they are moving in the right direction. First, despite aspersions sometimes cast on population genetics by complexity enthusiasts, the Replicator Dynamics itself is non-linear due to the complexity of the payoff terms, and to the customary iteration of the equations in their discrete form.[2] This frequently results in systems equilibrating to periodic "limit cycles" rather than stable equilibria consisting of only one or a particular mix of strategies. These results are often referred to as "Red Queen Evolution", after Lewis Carroll's character who had to run faster and faster in order to stay still.

Chaotic attractors have also been demonstrated under Replicator Dynamics with at least four strategies, both analytically and computationally [Skyrms, 1992]. Chaos allows us to understand at least in principle how complexity can emerge deterministically from simple recursive processes, though there is good reason to suspect that much of the complexity we see around us is the result of unorganized "noise," rather than deterministic and mathematically precise recursions. Limit cycles, though less alluring than chaos, may be more significant in the long run, offering plausible models of a variety of periodic phenomena.

[2]Repeated multiplication of a frequency by the same fitness results in exponential rather than linear growth. Meanwhile, the sensitivity of the fitnesses to the frequencies of other types introduces additional complexity. For instance, notice that Fisher's Fundamental Theorem — that fitness increases monotonically — does not hold even for simple symmetrical games like the Prisoner's Dilemma. (The theorem requires the fully symmetrical payoff matrices characteristic of allele pairing.) Selection drives fitness down in the Prisoner's Dilemma.

Agent-based models introduce a second source of non-linearity. Whereas the Replicator Dynamics models population indexes of continuous quantities, agent-based models deal with relatively small populations of discrete individuals, whose birth, demise, or change of strategy introduces small jumps in the payoff values driving the process. Such populations always, in small measure, fail to evolve precisely according to the payoffs. While this is easily conceived of as noise or inaccuracy in the models, these small non-linearities are useful for modeling phenomena like genetic drift which are critical to models like Sewall-Wright's "shifting balance" evolution.[3] Rather than being mere artifacts of the modeling technique, non-linearities due to discreteness are related to real-world phenomena in important ways.

It may be that the most important kind of non-linear complexity in human evolution is something we are just beginning to explore: the interaction between biological and cultural evolution. Whether it be the way pain mechanisms influence clothing fashions, how sexual instincts affect cultural taboos, or how the reliability of the senses drives the evolution of scientific knowledge (each of which has dramatic reciprocal effects on human genetic fitnesses), one must understand both biological and cultural change in order to understand human beings, and ultimately we must understand them as interacting dynamical processes. Some noted preliminary models exist, but we are desperately in need of a general modeling framework for understanding this interaction. I will make some suggestions below about how this might be accomplished, and what difficulties we may encounter.

What EGT modeling has taught us so far

Perhaps the simplest and most dramatic result of EGT so far is the change in how we view the relationship between evolution and morality. Historically, it has been taken as self-evident that the kind of self-restraint and self-sacrifice that morality often requires of us cannot be a product of the material world (which is a matter of what is, rather than what ought to be) and especially not of evolution by natural selection, which is fundamentally driven by competition and self-serving behavior [Williams, 1994]. The study of the Prisoner's Dilemma over the last few decades has created a dramatically different view.

The Prisoner's Dilemma is the single best known and influential game in Game Theory, and constitutes the primary tool for studying the evolution of cooperation. For those not familiar with it, it is simplest to imagine some act which benefits another more than it costs the actor. Sharing excess food is a good example. Those who perform the act are called "cooperators" and those who do not are called "defectors". If the cost is 1 and the benefit is 4, and if one is equally likely to be on either end of the act, then we can easily calculate the expected payoff

[3]Selection drives large populations with small fixed mutation rates toward local optima. Wright's "shifting balance" model suggested that the stochastic fluctuations of fitness (good or bad "luck") in small groups (and subgroups) might allow populations under selection to escape from local peaks in the adaptive landscape. If so, this would increase the ability of the population to find global optima.

for four kinds of encounters. If, for instance, we let DC be the average payoff to a defector interacting with a cooperator, then DC = 2, CC = 1.5, DD = 0 and CD = -0.5. This satisfies the definition of the Prisoner's Dilemma: DC>CC>DD>CD with the added requirement that $2\times$CC > DC+CD. (The latter requirement insures that one can't do as well as cooperators by taking turns being exploited.) This cost-benefit method of generating a PD explains the centrality of the PD for understanding cooperation.

The Prisoner's Dilemma has been studied extensively, both as a rational choice game and as an evolutionary game. The results on the rational choice side have been disappointing. It turns out that even the most idealized rationality is fundamentally a device for the calculable satisfaction of personal preferences, and as such will never do the right thing simply because it is right. (The apparent exception is where the individual has a prior preference for doing the right thing, but then rationality merely serves it and thus cannot account for it.)[4]

In the simplest cases (the large panmictic populations of the basic Replicator Dynamics) natural selection reaches the same conclusion as rational choice: nice guys finish last. However, when additional elements of realism are added to the models, as when individuals are related or spatially located or tend to interact more often with some than with others for some other reason, it seems that efficient self-sacrifice can be adaptive in these "structured" populations. The curious thing is that, contrary to what philosophers have long supposed, reason does not give rise to morality in any straightforward way. It is easy, on the other hand, to show a variety of ways in which "hard-wired" instincts for self-sacrificing behavior can be adaptive. To this, critics have responded that inasmuch as "moral" instincts are present due to adaptive advantage, they can not be genuinely altruistic in the way that genuine morality requires. Proponents [Sober and Wilson, 1998] have responded that what we are normally worried about vis a vis morality is the proximal psychological mechanisms involved in social behavior, not whether the self-sacrifice of individuals somehow ends up being a good thing in general. Is the individual calculating advantage, or are they simply acting because that's the kind of person they are? Arguably, in the absence of a clearly satisfying account of reason-based morality, strong moral instincts suggest themselves as the real basis of what we are talking about when we talk about genuine, as opposed to calculated, moral behavior.

One result of all this is that it begins to appear to many that fundamental morality (as opposed to the specifics of moral codes) may be more plausibly based on instincts and evolutionarily structured emotions than on reason. This prompts a fundamental rethinking of who we are and which parts of ourselves are to be celebrated and which denigrated.[5] It seems possible that real human goodness

[4]There are those, of course, who argue that this simply shows the inadequacy of the default model of rationality favored by decision theorists and economists [McClennen, 1990; Green, Shapiro, 1994].

[5]David Hume, of course, should get credit for being the first to make a big point of this. See especially his *An Enquiry Concerning the Principles of Morals*.

comes from the side of us we share with other animals, and much-celebrated reason, while making us unique, is a bookkeeper, an optimizer, but not necessarily the source of what we have always recognized as the best in us.

The application of evolutionary games to problems of epistemology is still in its infancy, though there is reason for optimism. As our understanding of ourselves as products of evolution deepens, it is quite natural to think that many of the traditional problems of epistemology will have evolutionary solutions. That natural selection insures the general reliability of the senses and memory has become something of a truism. For instance, Quine [1969] suggested that natural selection insures that our subjective quality space resembles that of the world. Deeper questions regarding how concepts can be understood to "cut nature at its seams" may at least be approachable by considering the biological functions of thought and language. Presumably, even if concepts (e.g. "is edible") are biological constructs, the use of these concepts in controlling action picks out real divisions in nature, though they be relative to our needs and our behavioral repertoire. Such insights are not difficult to find. What is difficult is finding the tools and the model of inquiry necessary to transform what is merely a philosophical "stance" into a productive component of the general scientific study of cognition and language.

One of the more exciting recent developments in evolutionary Game Theory is the proliferation of signaling game models.[6] Following David Lewis' [1969] approach to modeling the formation of linguistic conventions, Skyrms [1996] and others [Huttegger, 2007; Barrett, 2007] may be laying the formal foundations for a biology-friendly correspondence semantics, consistent with certain current philosophical trends [Millikan, 2005]. Though current work tends to focus on the social-convention forming interpretation of these models, careful consideration shows implications for core problems of reference as well. Roughly, the idea is that the history of adaptive contributions of a signaling system induces a partition on the space of possible world-states, finding action-relevant divisions in nature via these historical interactions [Harms 2004]. Insofar as elements of brain architecture can be considered as co-evolved "senders and receivers", the models will apply to the meanings of thoughts as well as signals, and this without supposing any prior semantic properties in the systems [Millikan, 1984]. If this project works out, then it may be that evolution can explain the other big outstanding philosophical puzzle: the nature of the "directedness" of language and thought.

Multi-level selection models

One of the reasons for the remarkable popularity of evolutionary Game Theory is the fact that it can model biological and cultural transmission with equal facility, to the extent that theorists are often unconcerned with whether they are talking about Darwinian fitness payoffs driving differential reproduction or about cultural fitness payoffs driving differential imitation. Continuing with the example of food

[6] Notice that these models differ from the signaling literature in animal behavior in postulating cooperative rather than competitive payoffs. Cf. [Maynard Smith and Harper, 2003].

sharing, one can imagine two very different ways that it might propagate in a population. It might be the case that a food-sharing instinct triggered by familiarity propagates via biological inheritance due to kin-selection. On the other hand, it might be a learned strategy propagated via selective imitation, where individuals observe that those who share food do better overall (e.g. due to reciprocity.) One can use the same equations and the same payoff matrices to model either process.

It is instructive to compare population models to memetic approaches to cultural evolution. Richard Dawkins in his influential *The Selfish Gene* proposed that "progressive" Darwinian evolution will occur anytime there is a "replicator", an entity that produces mostly accurate copies of itself [Dawkins, 1976]. He used the term *meme* to designate a cultural replicator, which he conceived of as a kind of information virus that leaps from brain to brain according to how good it is at getting brains to make copies of it. The implication was that in order to apply Darwinian concepts to culture, these *memes* must exist. The sub-discipline generated by this suggestion is called "memetics", and its principle challenge remains the specification of what sorts of cultural entities fit the bill.[7]

One way to extend the domain of a theory is by analogy, and this is what memetics attempts to do. The other way to proceed is to examine the conceptual requirements of the theory one wishes to apply, which in this case is the more productive approach. The broad flexibility of evolutionary Game Theory stems from the fact that it need assume no replicators, but only that one has a "population" of things that can be counted and categorized, and that changes of the frequencies of those things are driven either by environmental interaction (selection) or additive causes like mutation and immigration (variation). In general, any cause of frequency shift can be represented in some version of the "replicator" dynamics, and as such those dynamics apply to any population regardless of whether its members are replicators in any sense at all. All of which suggests that rather than the Darwinian concepts of selection and variation only applying to a relatively small class of entities, they can be used to apply to just about anything. In a sense, this is bad news since "Darwinian evolution" doesn't explain the accumulation of adaptive complexity without a lot of physical details regarding why the fitnesses and patterns of variation of a population are the way they are. On the other hand, it seems to be very good news from the point of view of building a comprehensive formal framework for modeling gene-culture co-evolution. Why not just build two Replicator Dynamics models, and add in the kind of interdependencies that seem to hold between biological and cultural change?

Donald T. Campbell was the first to propose this kind of "selection hierarchy" with any detail [Campbell, 1987]. He believed that Darwinian evolutionary processes are ubiquitous, and suggested ten different levels in a "nested hierarchy of blind variation and selective retention processes" that underlies the accumulation of scientific knowledge.[8] He termed this project "evolutionary epistemology," and

[7]The now-defunct "Journal of Memetics" archives can be found here: `http://cfpm.org/jom-emit/`

[8]These ten levels seem to have come from a sort of survey of selection and variation processes

argued vigorously for the rigorous generality of the concepts of variation and retention. If we take something like Campbell's proposal as a long term goal, perhaps a modularized version of the replicator dynamics could be used to rigorously explore the behavior of this kind of selection hierarchy.[9]

My goal in this section is to explain why modularizing the Replicator Dynamics will not supply a fully adequate implementation of Campbell's selection hierarchies, despite the generality of its concepts of selection and variation. The limitation is not that there are items of interest that do not evolve according to the Darwinian algorithm, but that individuals become of critical importance in modeling the dependencies between biological and cultural evolution, and individuals do not appear in population models. Why they are so critical is a little difficult to explain.

Consider the relationship between thermal distress and human clothing styles. Suppose that the discomfort caused by extreme temperatures has a simple genetic basis, and it affects people's clothing choices. Clothing manufacturers present new sorts of garments to the public (variation), and those that proliferate tend not to create excessive thermal distress (selection). In this case, the genetically determined thermal distress functions as one of a number of selection mechanisms on the evolution of clothing. In return, thermally appropriate clothing has some incremental effect on the survival chances of those carrying the genes. (It helps here to imagine a cold environment and a relatively fashion-indifferent population.) In such a case, genetic distributions influence fitness differences in clothing, and clothing distributions influence genetic fitnesses. This kind of (typically non-linear) "fitness feedback loop" is exactly the kind of relationship we need to understand in order to see how and in what way cultural evolution serves biological ends. Let us resist the temptation to elaborate this model in the direction of realism, for the problem is already apparent, and will plague any more elaborate model as well.

The problem is this: while it is true that in the dual level model the genetic frequencies affect cultural fitnesses and cultural frequencies affect genetic fitnesses, this arrangement does not fully capture the two-way interrelationship we are interested in; in the real world genetic traits do not affect the culture uniformly, and cultural distributions do not affect genetic fitnesses uniformly. Rather, it is the particular individuals with the genes who make the safe choices and who in turn derive the genetic fitness benefit of the cultural distribution.

What this shows is that we cannot just pick the populations that interest us (genetic and cultural) and write some equations to model the dependencies between them, as proposed in [Harms, 2004]. On this simple approach we will systematically underestimate the selective force on genetics due to cultural evolution. Suppose, for example, that a new genetically determined sensor mechanism for some toxin results in a shift in dietary patterns in some culture. The primary ben-

that had been modeled or studied, rather than from a systematic theory of cognition and control.

[9] Note that the "levels" of selection here are of a very different kind than the levels of scale that come up in the "group selection" literature. There, the levels are those of scale, rather than those of a control hierarchy.

efit would be to the individuals carrying the sensor (in modifying their own diets) with a weaker secondary effect on others, say, because they were imitated or not served food with the toxin. The simple two-level vector model would show that the frequency of toxic food went down, and that benefits would be distributed equally to all genotypes carried by members of the culture. Thus the selective advantage to the toxin-avoiding gene is underestimated.

One way around this is to multiply the number of populations. Suppose there are two versions of the thermal distress allele. Then we distinguish two cultural populations — the distribution of clothing worn by those with version A and that worn by those with version B. This will allow us, in very simple cases like this, to localize the genetic fitness advantages of cultural trends. But, as our model becomes more complex, the number of required populations will increase to the extent that there is no longer any computational advantage to be had from the population approach over direct agent-based models. So the problem with modularizing the replicator dynamics is that far too many sub-populations are required to localize the downward flow of fitness advantages.

It is thus that the advantage of agent-based models becomes clear. It is now fairly common practice to implement even simple evolutionary games via agent based models, and such models commonly include both reproductive life cycles *and* interactions between individuals. Relatively little additional complexity is required to model simultaneous genetic and cultural evolution *without* losing track of the role of individuals as the loci of gene/culture interaction. Which is to say, the future study of gene-culture co-evolution must include agent based models. But are they a general modeling solution?

As noted above, agent-based models suffer from their own difficulties. It was once the case that they were too computationally intensive to be practical, but the power of modern object-oriented programming techniques and desktop computers has for the moment far outstripped the sophistication of the questions we are asking. The problem, or perhaps it is a challenge, is that agent-based models tend to be idiosyncratic and opaque to all but the author of the model. The number of arbitrary parameters required to set up even the most rudimentary life cycle is a bit surprising, and even setting aside the tendency of modelers to add their own pet elements of "realism," it may be that we will never see the kind of canonical formulations that have made evolutionary Game Theory so productive thus far. Arguably, without the duplication of results or at least the thorough criticism of cognizant peers, models of this sort may be doomed to an ongoing crisis of credibility.

Multi-model methodology

Whether or not the dilemma indicated above is grounds for despair depends largely on what one expects modeling gene-culture co-evolution to accomplish. On the one hand, it appears that canonical models need to be algebraic, and such models cannot accommodate the crucial role of individuals as the loci of fitness advantage. On the other, agent based-models are idiosyncratic and opaque, and thus cannot be canonical. If what one wants is a model of gene-culture co-evolution that is both tractable and complete, it is very hard to see what that could possibly look like. It may be, however, that such a desire is both unrealistic and outmoded.

It might be said that there are two basic attitudes toward modeling. One frequently detects, in certain areas of physics (cosmology, String theory) and in areas of complexity research like chaos theory and the metaphysics of information, the hope that we may someday (soon) discover the equation that generated the Universe. Or perhaps, it will be a single pattern, iterated over and over again at different scales, that explains everything simply and elegantly.

To workers in the biological and socials sciences who must deal with what is given rather than with the simplest and most fundamental things discoverable, this hope will seem foreign, even unintelligible. Models in these broad disciplines tend to be "local", in the sense of only covering a very small part of the overall fabric of the system being studied. Their complexity is a delicate balance between comprehensibility and breadth of practical applications. The general modeling strategy is one of modularity, where it may not be possible to connect up the modules in any systematic way.[10] What counts as understanding is not an overarching picture of the universally driving mathematics, but a facility with the available models — knowing when to apply them and how to deal with their idiosyncrasies. It is also the case that a good deal of what goes on in physics — one thinks of the simplified local models of classical mechanics — falls as much more into this second category than the first. The application of physics to engineering problems certainly does.

Campbell's "nested hierarchy of blind variation and selective retention processes" certainly had the flavor of the overarching theory of everything, and for a moment the flexibility of application of evolutionary models may have tempted us to think that his was a big picture which could be made both tractable and precise. The hope was a fleeting one, as we saw. This does not mean that these models are not part of the general solution for the modeling of complex social and biological systems.

If evolutionary models fail to provide lossless modular scalability, they have proven themselves to provide something that we need far more than that. Much as standard statistical tests have given workers-in-the-field ways of avoiding ordinary human measurement biases, evolutionary models have provided generations of researchers with a way of thinking about non-linear dynamics in biological and social systems, and with rigorous yet accessible models of equilibria and basins of attraction. Added to the track record of important discoveries about the nature of sociality and meaning, evolutionary game models will remain an essential part of the effort to model complex social and biological systems.

This does open up an important foundational challenge: if both agent-based and population/vector models are to be part of our mature methodology, what is the relationship between the two? Here it is helpful to remember the general methodological structure of the biological and social sciences prior to the easy availability of computer modeling. Omitting informal theorizing, one might say

[10]Rosenberg [1994] suggests that there may be reasons why biology *has* to be modular.

there are three major components of scientific inquiry: (1) formal (algebraic) models, (2) intrusive observations (experiments and physical models), and (3) passive (empirical) observation. It is easiest to imagine each component as a link in a methodological chain. A robust methodology requires a certain kind of interaction between each link. Formal models suggest experiments, whose results suggest reciprocal modifications in the formal models. Experimental results and empirical observation have a similar reciprocal relationship, each suggesting new things to try in the others.

In this simple chain-link picture, one can think of the role of experiment as mediating between extremes of theory and the world, helping us to bring them closer into conformity. (Recall that the original Aristotelian model of science did not include experimentation at all.) I would like to suggest that population vector and agent based models form additional links in this chain.[11]

If algebraic models (1) are the most abstract and satisfying links in the chain of our scientific theorizing, they suffer from either over-simplification (but are tractable) or from intractability when systems made more realistic get too complex. Population-vector models and the associated "Monte-Carlo methods" allow us to study the behavior of complex dynamical systems away from equilibrium (the basins of attraction), and in many cases find equilibria with greater ease than with analytic methods alone. As such, they form the new number (2) link in the methodological chain. Agent based models are a step closer to the jittery stochasticity of the real world, being a sort of computational experiment. Thus they form a new number (3), with physical experiment as (4) and empirical observation as (5).

If this makes sense, then the relationship between vector and agent models is not that one is truer and trumps the other, but that they form two new links in the chain of tools that allow us to bring our most abstract theories into conformity with the world as we observe it. As we move closer to mathematics, our understanding is clearer (and simpler). As we move toward the world, our models are more complex, our confusions abound. Each link should suggest modifications of the adjacent ones. It may be that, in the end, we will see that our understanding of the world exists distributed throughout the chain, rather than being the exclusive property of the most abstract link.[12]

Formalism in philosophy

There has always been a certain kind of philosopher attracted to the most rigorous and abstract tools. Historically, the tools have been those of mathematics and deductive logic. The twentieth century saw the introduction of set theory (shown to be the true foundations of math and logic by Russell and Whitehead [1925] and

[11] I am indebted to a talk some years ago by Danny Grunbaum (University of Washington Oceanography) for the basic idea.

[12] Doubtless, we would find that the real structure of healthy methodological interdependency is more complex than the simple chain imagined here.

later probability theory, which played a crucial role in the development of rational choice theory and technical approaches to ethics. What we have been considering is the emergence of a new kind of enquiry in philosophy, made possible by the dramatic improvement in computers and the associated dynamical models. It is probably too early to say that dynamical modeling is displacing logic and axiomatic systems as the tool of choice in technical philosophy, but it is certainly on the rise, and it may be that it will end up reinforcing different philosophical values than axiomatic systems.

What I have in mind is that axiomatic systems speak more to the Rationalists in us, encouraging us to turn inward, and to hope that somehow what seems *necessary* to us must be true of the world.[13] The potentially odd thing about this is the number of technical philosophers who are fans of today's science, and who may therefore be of a more Empiricist than Rationalist leaning. What is promising about the new tools of evolutionary modeling is that they are not at the most abstract end of the chain, but rather encourage us to engage in the messy business of fitting theory to the world, fostering our inner Empiricists, and perhaps, in the process, we will end up weaving the traditional philosophical concerns about meaning and normativity back into the fabric of science.

BIBLIOGRAPHY

[Axelrod, 1984] R. Axelrod. *The Evolution of Cooperation*. New York, Basic Books, 1984.

[Barrett, 2007] J. A. Barrett. Dynamic Partitioning and the Conventionality of Kinds. *Philosophy of Science*, 74, 527-46, 2007.

[Boyd and Richerson, 1985] R. Boyd and P. J. Richerson. *Culture and the Evolutionary Process.* Chicago, The University of Chicago Press, 1985.

[Campbell, 1987] D. T. Campbell. Evolutionary Epistemology. In Radnitzky and Bartley, eds. *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge.* La Salle, Illinois, Open Court, 47-89, 1987.

[Dawkins, 1976] R. Dawkins. The Selfish Gene. New York, Oxford University Press, 1976.

[Green and Shapiro, 1994] D. P. Green and I. Shapiro. *Pathologies of Rational Choice Theory : A Critique of Applications in Political Science*. New Haven, Yale University Press, 1994.

[Harms, 2004] W. F. Harms. *Information and Meaning in Evolutionary Processes*, Cambridge Studies in Philosophy and Biology. Cambridge, Cambridge University Press, 2004.

[Hofbauer and Sigmund, 1988] J. Hofbauer and K. Sigmund. *The Theory of Evolution and Dynamical Systems.* Cambridge, Cambridge University Press, 1988.

[Huttegger, 2007] S. Huttegger. Evolution and the Explanation of Meaning, *Philosophy of Science*, 74(1), 2007.

[Lewis, 1969] D. Lewis. *Convention.* Cambridge MA, Harvard University Press, 1969.

[Maynard Smith, 1982] J. Maynard Smith. *Evolution and the Theory of Games.* Cambridge, Cambridge University Press, 1982.

[Maynard Smith and Harper, 2003] J. Maynard Smith and D. Harper. *Animal Signals.* Oxford Series in Ecology and Evolution. Oxford, Oxford University Press, 2003.

[McClennen, 1990] E. F. McClennen. *Rationality and Dynamic Choice : Foundational Explorations.* Cambridge, Press Syndicate of the University of Cambridge, 1990.

[Millikan, 2005] R. G. Millikan. *Language : A Biological Model.* Oxford, Oxford University Press, 2005.

[13]This is, of course, the fundamental assumption behind *metaphysics*, and genuine metaphysics has been generally discredited since Hume and Kant.

[Millikan, 1984] R. Millikan. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge MA, MIT Press, 1984.

[Quine, 1969] W. V. O. Quine. Epistemology Naturalized. In *Ontological Relativity and Other Essays.* New York, Columbia University Press, 69-90, 1969.

[Rosenberg, 1994] A. Rosenberg. *Instrumental Biology, or, the Disunity of Science*, Science and Its Conceptual Foundations. Chicago, University of Chicago Press, 1994.

[Russell and Whitehead, 1925] B. Russell and A. N. Whitehead. *Principia Mathematica*. 2d ed. Cambridge, Cambridge University Press, 1925.

[Skyrms, 1992] B. Skyrms. Chaos and the Explanatory Significance of Equilibrium: Strange Attractors. In Evolutionary Game Dynamics. *PSA 1992 vol.2.* Philosophy of Science Association, 374-94, 1992.

[Skyrms, 1996] B. Skyrms. *Evolution of the Social Contract*. Cambridge, Cambridge University Press, 1996.

[Sober and Wilson, 1998] E. Sober and D. S. Wilson. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge MA, Harvard, 1998.

[von Neuman and Morgenstern, 1947] J., von Neuman and O. Morgenstern. *The Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1947.

[Williams, 1994] G. C. Williams. A Sociobiological Expansion of Evolution and Ethics. In Paradis and Williams, eds. *Evolution and Ethics* Princeton, NJ: Princeton University Press, 179-214, 1994.

GENERAL SYSTEM THEORY

Wolfgang Hofkirchner and Matthias Schafranek

1 INTRODUCTION

The origins of General System Theory (GST) date back to Ludwig von Bertalanffy's (1901–1972) interwar studies whose continuation after World War II found a friendly reception in the Anglo-American hemisphere. At the same time while there was also some mutual interaction with developments in operations research, engineering science, management science and cybernetics, partly independent and partly overlapping domains that shared a connection with United States think tanks like the *RAND Corporation* (cf. [Davidson, 1983; Hammond, 2003]).

The term "General System Theory" (GST) was coined by Bertalanffy himself. By that term his "Allgemeine Systemlehre" was translated into English [Bertalanffy, 1950], even though for a long time Bertalanffy resisted translating "Lehre" ("doctrine") as "Theory". Later on, during his 1954-5 stay at the *Center for Advanced Study in the Behavioral Sciences*, Stanford University, California (which had been formed with James Grier Miller as spearhead), he joined others in using the plural "Systems" when, together with Kenneth E. Boulding (1910–1993), Ralph W. Gerard (1900–1974) and Anatol Rapoport (1911–2007), they held a seminar on "General Systems". Subsequently they founded the *Society for the Advancement of General Systems Theory*, soon renamed the *Society for General Systems Research* and since 1988 again renamed the *International Society for the Systems Sciences*. They also edited the *General Systems Yearbooks*. Nonetheless, Bertalanffy continued to use "system" in singular and plural interchangeably, as well as the German translation of "GST" into "Allgemeine Systemtheorie" (cf. [Bertalanffy, 1972a; 1972b]).

The main reason why he initially hesitated to call his subject a theory was that he wanted to make it distinct from traditional disciplinary theories in science like radio astronomy or theory of emotions that have a rather restricted purpose, object of investigation and method by which they are constructed. GST, in contradistinction, was intended to be more than that: a theory of homologies/isomorphisms that are characteristic of the organisation of "wholes" — systems — wherever they might appear, whether in the domains of physico-chemical sciences, life sciences, social sciences or humanities. Because of that GST, as a theory of general principles of systems, was considered a kind of metatheory, cutting across, and capable of unifying, the diverse specialties, including offering a new *weltanschauung*, a new world view — hence a new paradigm for all sciences and humanities. (That is why

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

Pouvreau and Drack [2007, 282-283], argue that "general systemology" would be a more appropriate name for GST.) However, *weltanschauung* goes beyond mere descriptions of the world ('world pictures') to include values and it tries to make world pictures consistent with values concerning the relationship of humans to the world.

While this was an ambitious goal, when Bertalanffy announced his GST in reality it rather described a research programme and a framework than a fully-fledged body of theoretical knowledge, even though he and other colleagues had already made some, and would make further, substantial contributions to theorising systems (cf. [Bertalanffy, 1972a, 186]). Yet in 1968 Rapoport wrote: "General system theory is best described not as theory in the sense that this word is used in science but, rather, as a program or a direction in the contemporary philosophy of science" [Rapoport, 1968, 452]. Still today critics bemoan that even Bertalanffy's book *General System Theory* [1968] and Rapoport's book *General System Theory* [1986] by no means provide a concise overview, but demonstrate the rather unsystematic and unelaborated state of GST (cf. [Müller, 1996]). Others argue, however, that in spite of that, a substantial body of knowledge has been achieved that paved the way for, and often anticipated, findings more recently made known by the sciences of complexity (cf. [Hammond, 2003; Hofkirchner, 2005]). Kurt Richardson says in the introduction to the reprint of Boulding's article on General Systems Theory of 1956, "The modern complexity movement is in some ways quite different from the general systems movement ... [But] ... Complex systems thinkers share a lot of the aims and ambitions of the original general systems movement" [Richardson, 2004, 127].

What are these aims and ambitions of GST, on closer scrutiny?

According to Rapoport, the central aim of GST is to integrate the analytical and "organismic" as well as the descriptive and normative traditions of systems thinking. "Our aim will be to show that far from being incompatible, these views are complementary, revealing different aspects of a unified approach to system theory", [Rapoport, 1986, 7].

Regarding the difference between the analytical and the "organismic", Rapoport distinguished between two fundamentally different understandings of the term "system" [Rapoport, 1968]. On the one hand, there is the analytical comprehension that originates from the tradition of so-called exact natural as well as engineering sciences. A conception of a theory of systems in this tradition is tied to the use of mathematical models and methods that have originally been developed within the context of the investigation of the non-living world. On the other hand, we have the so-called "organismic" approach to the system idea, a term originally introduced by Bertalanffy. Rapoport relates it to the diverse phenomena of the living and social worlds as integrated wholes that are investigated by different life and social sciences respectively. Regarding the difference between the descriptive and the normative, Rapoport distinguished between two views of knowledge as answers to the question "how" and answers to the question "what for". In terms of the descriptive view, a system is perceived as interesting in its own right. Curiosity is guided here by the question of how a system actually works. By contrast, the normative approach perceives a system as an entity that exists for something; and this "for something" either refers to an instrumental attitude where a system is considered as specifically designed to serve some goal (and is judged by its performance), or to a non-instrumental view where a system is perceived as something whose continued existence is an end in itself.

By aiming to bring together the analytical and the organismic as well as the descriptive and the normative, GST was considered by its proponents a major step in the quest for a unity of science. "Its major goal is to generate a new type of unity of science: not a unity based on the reduction of the concepts, methods or even laws of all sciences to the ones of a single science regarded as more essential; but rather a formal unity based on the generality and ubiquity of the system concept and on the 'isomorphisms' it induces between sciences of which the logical and methodological autonomy is guaranteed", [Pouvreau, Drack, 2007, 283]. "The need for general systems theory is accentuated by the present sociological situation in science", wrote Boulding [1956/2004, 198], "The more science breaks into sub-groups, and the less communication is possible among the disciplines, ... the greater chance there is that the total growth of knowledge is being slowed down by the loss of relevant communications", [Boulding, 1956/2004,198-199]. Therefore GST would develop generalised ears to provide specialists with advantages due to the awareness of similarities among widely different empirical fields. But in order to do so it had to pursue another aim too — that of acknowledging that "the total of observable phenomena ... shows a structural uniformity" [Bertalanffy, 1950, 234] that is not only the result of construction but also an independent feature of the world.

Thus there are three aims of GST belonging to three different, albeit intricately interwoven, fields, each of which implies philosophical assumptions:

- the aim of bringing together the analytical method and a proper scientific method for investigating the organismic realm by postulating isomorphisms between different disciplines; this belongs to the field of epistemology (see section 2 below);

- the aim of relating real-world systems to each other by postulating isomorphisms between them; this belongs to the field of ontology (see section 3 below);

- and the aim of reconciling the world of facts with the world of values by considering GST a tool for intervening in systems in a humane, humanistic, that is, nonmechanistic and nonreductionistic, way; this belongs to the field of ethics (see section 4 below).

What then, from a contemporary point of view, may be considered the lasting achievements of GST in these fields?

2 EPISTEMOLOGICAL IMPLICATIONS

Bertalanffy called his epistemological position "Perspectivism". It was a kind of realism based upon evolutionary thinking. This is a position neither naively materialistic nor radically constructivist. It seems to be a common denominator of GST and deserves attention, in particular, if compared with later developments of the systems movement. Much can be learned from the way of thinking Bertalanffy and companions applied when addressing reduction and the rather idealistic and anthropomorphising opposite of reduction which might be called projection.

2.1 Realism, "Evolutionary Epistemology", Perspectivism

Bertalanffy made his realism distinct from positivism and empirism. As he pointed out, "one cannot base and develop any science on sole experience and induction" [Bertalanffy, 1927b, 304].[1] "The determination of the natural laws is the product of a kind of alchemy between observation, 'constructive activity of the reason' and an intuition 'similar to the artistic vision"'. (ibid. 304) "A theoretical model is a conceptual construction", "every scientific theory can be conceived as a conceptual model" [Bertalanffy, 1965, 300]. However, "the actual world ... does allow the application of our intellectual constructions" [Bertalanffy 1950, 222]. Bertalanffy holds "that the world (i.e. the total of observable phenomena) shows a structural uniformity, manifesting itself by isomorphic traces of order in its different levels or realms" [Bertalanffy 1968, 87]. Knowledge about these isomorphies is made possible in as far as the structure of the cognitive ability is isomorphic to the structure of reality. That is, isomorphism is a fundamental condition of the adequacy of thought to reality. It is not required that the categories of experience fully correspond to the real universe, even less that they represent it completely. It is sufficient that a certain degree of isomorphism exists between the experienced world and the real world, so that experience can guide the organism in such a way as to preserve its existence (cf. [Bertalanffy 1955a, 257]). Bertalanffy followed Konrad Lorenz in that the so-called "'a priori' forms of intuition and categories are organic functions, based upon corporeal and even machine-like structures in the sense organs and the nervous system, which have evolved as adaptation in the millions of years of evolution. Hence they are fitted to the 'real' world" [Bertalanffy 1955a, 256]. In this respect, GST seemingly anticipated fundamental assumptions of what was later subsumed under the labels of Evolutionary Epistemology and Evolutionary Psychology more generally.

Bertalanffy called his special epistemological view "perspectivism". By that term he meant "that no knowledge grasps the ultimate reality, and that it can only mirror some aspects of reality in more or less appropriate models" [Bertalanffy, 1965, 301]. Notwithstanding the fact that our experience and thinking appear

[1] English translation quoted after Pouvreau and Drack 2007, 304. In what follows all quotations for Bertalanffy's German language works referenced (1927-49, 1965) derive from this source and follow their English translation; all quotation page numbers used in the text, outside of reference brackets, refer to this source.

to be determined by biological as well as cultural factors, this human bondage is stripped by a process of progressive de-anthropomorphisation of our world picture making our perspectivist knowledge of reality, though only relatively, true. For example, physical constants such as Loschmidt's number and the like "represent certain aspects of reality, independent of biological, theoretical or cultural biases" [Bertalanffy, 1955a, 258, cf. 262]. Bertalanffy developed such a stance against the constructivist view of Jacob von Uexküll ([von Uexküll, 1920], cf. [Bertalanffy, 1955a, 255-6]).

2.2 Anti-reductionism, Anti-holism (Anti-mechanicism, Anti-vitalism)

Bertalanffy's work on a theoretical biology lies at the foundation of the modern scientific approach of systems thinking. The chief controversy marring theoretical biology at his time was the deep cleft between mechanicism and vitalism, where mechanicism was the materialistic approach that tries to reduce life phenomena to phenomena that can be explained by physics and vitalism was the idealistic conviction that that which transcends being explained by physics is something metaphysical. GST was born when Bertalanffy, in an attempt to overcome that deep cleft, formulated laws of organisation ruling biota as well as other ordered entities. By deliberating on the shortcomings of both positions, Bertalanffy developed a third view that tried to integrate the reasonable aspects of each of the two perspectives on life. He called it the "organismic" perspective. This view took over the notion of wholeness from the vitalist standpoint by fundamentally accepting the relative autonomy of the living world. Thus, it refused the neo-positivist notion of a mechanistic generation of form in organisms (as Friedrich Waismann most prominently put it forward) and the possibility of a complete reduction of life to physico-chemical processes. However, at the same time Bertalanffy's organismic stance adopted the mechanistic critique of the vitalistic idea of a supra-material, transcendent entelechy (as Hans Driesch most prominently advanced it). Actually, by searching for a tenable notion of wholeness Bertalanffy cleared this concept from its anthropomorphic implications and tried to put it on the firm ground of exact scientific thinking.

Bertalanffy laid the cornerstone for such an understanding within theoretical biology by advancing essential categories regarding the relation between open and closed systems, between causality and organised complexity, and the role of entropy. In doing so, he generalised the laws formulated to grasp biota as organised systems and found himself in the position to make them successfully apply to different domains such as medicine, psychology, psychotherapy, and so on. "It seems legitimate to ask for a theory, not of systems of a more or less special kind, but of universal principles applying to systems in general . . . , irrespective of whether they are of physical, biological or sociological nature" [Bertalanffy, 1955b, 31].

Bertalanffy not only disavowed reduction to physics and chemistry, which placed him in sharp contrast to attempts *en vogue* in the Vienna Circle at the time, he also explicitly repudiated biologism in relation to the explanation of social phenomena: "This does not imply 'biologism', i.e., reduction of social to biological concepts, but indicates system principles applying in both fields" [Bertalanffy, 1968, 125]. Besides his disapproval of, so to say, vertical reductionism regarding social science, he also argued against, so to say, horizontal reductionism as well. In discarding the summative concept of systems as mere aggregates, criticising the methodological individualism then abundent in social sciences as doomed to fail because of the practically insurmountable number of elements and interactions individuals might be involved in and because of its losing sight of the autonomy of systems due to the feedback the system exerts on the elements (cf. [Müller, 1996, 72-73])

On the other hand, he did not fall into the trap of holism. He was aware that there was a delicately balanced relationship between these two ways of approaching the world. Regarding living systems, he described the situation as follows: "What we can do is only: firstly, isolate the single process and determine it in a physico-chemical way... But this just gives us no knowledge about the biological problem, the reciprocal dependence of the single processes; or secondly, we can at one swoop define the whole event in the organism by an integral law ..., with which we nonetheless again have to forego the physico-chemical determination in the detail..." [Bertalanffy 1932, 307].

2.3 Mathematics and Non-formalism

This anti-reductionist as well as anti-holist stance has consequences for divining the relationship between, on the one hand, the analytical method and its use of mathematics and, on the other hand, the exclusive use of ordinary language in case "the general principles ... cannot be formulated in mathematical terms" [Bertalanffy, 1950, 221] as a scientific method as well. Bertalanffy shared this conviction with other representatives of GST, above all, with Rapoport.

Concerning the integration of the analytical and organismic approach, GST tried to achieve unity on an abstract level of valid analogy building based on structural similarities between diverse kinds of real world systems. Hence, the language of mathematics plays an important role in this integration process. As Rapoport put it in his last big synopsis of the GST programme [1986, 30]: "There is no escaping the conclusion that understanding (at least in its scientific sense) entails analysis." GST aims at extracting principal formal system properties that all kinds of real world systems (as investigated in the empirical sciences) have in common and tries to categorise abstract types of systems according to these formal properties.

In this context, new mathematical techniques were developed to cover all classes of real-world systems on a formal basis of representation. Regarding goal-seeking and self-controlling forms of behaviour, mathematical models of cybernetics based on the concept of various forms of information-controlled feedback became available. Additionally, the information concept opened up new directions to develop general notions of order and disorder, thus, to gain new formal measures of organisation. Generally, the information concept became the central unifying concept underlying the working of organismic systems. It reintroduced teleological ideas into the theory of physical processes, and established itself as a second unifying dimension alongside the energy concept.

Furthermore, the general scheme of a system of differential equations as formal representation of real-world systems became an ever more refined tool. In this regard, there are essentially two things to mention. First, by using such a representation of the concept of a whole the latter lost its metaphysical connotations. This becomes immediately clear when referring to Hall and Fagan's classical paper *Definition of system*. "The degree of 'wholeness' of the system is determined by the nature of the functions $f_1, \ldots, f_n$. If each of these functions depends strongly on each of the variables, the system shows a high degree of wholeness; a change in any variable then affects appreciable changes in the rest" [Fagan, Hall, 1956, 26]. Thus, the idea of interdependence became the central criterion for comprehending wholeness. "The more tightly interwoven is the network, the more organized is the system comprised by the relations" [Rapoport, 1970, 5]). Second, the scheme of a system of differential equations gave rise to further formal representation of important additional characteristics of organised systems. In this regard, Bertalanffy gave some examples that are derivable from such a mathematical expression of systems (cf. [Bertalanffy, 1968, chapter 3]). Among others, he referred to the phenomenon of segregation, where an initial system splits up into a group of independent units; a process he called "progressive mechanization". The subsystem that has so far acted as the leading hub looses its organising capacity. The opposite case is "progressive centralization" denoting development of new dependencies, leading to new functional units. In the course of the evolution of a system, both processes can combine such that a functional differentiation occurs; specialised and relatively autonomous subsystems coalesce on a more comprehensive level to a more effective structure. Related to these processes is also the essential concept of hierarchy. Together with further examples (growth, competition, finality) Bertalanffy showed how some general system ideas beyond the mechanistic level could be subject to formal consideration as well.

Finally, the concept of mathematical equilibrium that is tied to the identification of the extreme values of system equations has to be mentioned here. This procedure was equated with the definition of various general stability conditions related to the solution character of the system of equations. Thus, the methods of calculus were used in the GST programme for attaining general schemes for solving extreme value problems. Such schemes opened up new possibilities for the analysis of the stability of systems by asking for the robustness of equation systems in terms of varying initial values as well as of fluctuations of parametric boundary conditions. Consequently, the analysis of stability provided a general characterisation of systems behaviour as well.

However, as Rapoport pointed out in discussing stability in terms of equifinality related to the open systems aspect, "The question of the existence of a steady state independent of initial conditions is only one of many questions one can ask with reference to the behaviour of a system" [Rapoport, 1970, 8]. Other crucial questions were related to the number of stable states and the various types of steady states. The notion of equilibrium/stability as an exact concept was not related to the constituent entities of a system but referred to its overall organisation. As W. R. Ashby [1952, 57] put it, "The fact that the stability of a system is a property of the system as a whole is related to the fact that the presence of stability always implies some co-ordination of the actions between the parts." Additionally, as a mathematical idea with a clear appeal to wholeness it did not any longer refer to any of the teleological and metaphysialc connotations the concept of wholeness was traditionally confronted with.

The representation of systems as mathematical models in the form of differential equations, together with the possibility of a rigorous formal characterisation of a system's organisation as a holistic attribute, made this formal method a central instrument of choice within the GST. Representative for the overall programme and its proponents is the following quote of Bertalanffy: "... systems can be defined by certain families of differential equations and if, in the usual way of mathematical reasoning, more specified conditions are introduced, many important properties can be found of systems in general and more special cases" [Bertalanffy, 1968, 38].

This leads us, finally, to the central philosophy of science aim that the GST programme was pursuing — the possibility of an integration of all fields of science raised by an abstract mathematical representation of real-world systems. Due to its high degree of abstraction, a generalised mathematical model of the above mentioned form does not represent a specific system; rather, it represents the set of all systems with the same structural and dynamic relations. Thus, the method of mathematical analogy building seemed to be the most natural foundation of GST, as this implies analogy construction according to the most rigorous, i.e. formal, standards. This method was denoted the extraction and study of systems "isomorphism". Rapoport [1968, 455] gave the following definition of the term: "Two mathematical objects are isomorphic if there exists a one-to-one correspondence between the elements of one and those of the other and if the relations among the elements are preserved by the same correspondence." Thus, a system that is specified as a particular mathematical model is isomorphic to all systems that are specified by the models of the same type; and consequentially, a general classification of systems derives from a classification of mathematical models.

This methodological aspect of GST provided a shift from the specific nature of systems to their mathematical structure. It should lead to a new integration of knowledge on an abstract level of consideration; specialised concepts and terminology regarding the investigation of real world systems in the diverse fields of science should become interchangeable in the sense of a mutual translation on the structural level. "If a term enters as a homologous variable or parameter in two or more isomorphic models, then the term plays the same part in the respective theories" [Rapoport, 1968, 456]. GST regarded the reoccurrence of identical structural patterns in the theories of diverse disciplines as a sign of common organisation; they are not visible on the specialised level of scientific practice.

Nevertheless, GST was also fully aware of the fact that mathematical analysis does not reach into all fields of science, as formal-mathematical description has limits in terms of adequate representation of highly complex phenomena. Hence, GST considered the relationship of the analytical and organismic approach as a complementary one. Although achieving formal validity of analogies through reliance on mathematics is a central goal, this programme also acknowledges qualitative ways of analogy construction. Organismic phenomena often escape a reasonable mathematical representation, nonetheless analogous comparisons in a qualitative form can still bring about a significant heuristic potential for hypotheses generation. Thus, proponents of the modern framework of GST also referred to the heuristics of non-formal analogy building regarding the complexity of organismic systems [Gerard, 1958a; 1958b].

3 ONTOLOGICAL IMPLICATIONS

For Bertalanffy it was a fact that, ultimately, all sciences are concerned with systems [Bertalanffy, 1950, 223]. He distinguished real-world systems from systems in the mind. "What is to be defined and described as system is not a question with an obvious or trivial answer. It will be readily agreed that a galaxy, a dog, a cell and an atom are *real systems*; that is, entities perceived in or inferred from observation, and existing independently from an observer. On the other hand there are *conceptual systems* such as logic, mathematics ... which essentially are symbolic constructs; with *abstracted systems* (science) as a subclass of the latter, i.e. conceptual systems corresponding with reality" [Bertalanffy, 1968, XIX-XX]. It is worth noting that this is in contradistinction to some subjectivistic, constructivistic system theoretical concepts of today that exclude real-world systems from scientific endeavour.

What are the features of these real-world systems according to GST?

3.1 Self-organisation, Organised Complexity

Though the term "self-organisation" entered the scientific discourse only at the end of the 1950s, it might well be said that the concept itself was anticipated by Bertalanffy years before. Pouvreau and Drack [2007, 302] mention that Bertalanffy was strongly influenced by Julius Schaxel in Jena "who strives from 1919 on for the development of a theoretical biology worthy of this name and able to open a third way between 'mechanicism' and 'vitalism'." Publications he edited in *Abhandlungen zur theoretischen Biologie* recognised "self-organization as an inherent and materially immanent principle of life" [Pouvreau and Drack, 2007, 302]. Müller writes that Bertalanffy interpreted the phenomena in question as self-organisation processes [Müller, 1996, 87].

Rapoport's notion of "organized complexity" may also be regarded as term for depicting the essential feature of self-organising systems not to be found in mechanical systems. Rapoport came up with a simple general classification scheme of real world systems that also became the starting point for the further unfolding of the GST programme. It represents a triadic classification that builds upon principal kinds of phenomena that have been subject to scientific investigation according to the crucial criterion of organisation. The three groups are the following: organised simplicity, organised complexity, and chaotic complexity, which today would be re-phrased as something like simple, complex organised and complex random. Any phenomenon investigated by an empirical science can be subsumed under one of these categories.

Regarding organised simplicity, Rapoport and Horvath gave the following description. "The organization of a system is simple if the system is a serial of an additive complex of components, each of which is understood" [Rapoport, Horvath, 1959, 89]. As examples, they referred to the notion of a machine as a deterministic time-linear chain of events and to a propagating wave consisting of additively super-imposed sinusoidal components. On the other hand, concerning systems of chaotic complexity, "... the number of entities involved is so vast that the interactions can be described in terms of continuously distributed quantities or gradients, i.e., do not need to be specifically identified with regard to the individual entities" (ibid., 89). The problem with this description is that it is much too broad as it stands since it comprises all sufficiently dense or sufficiently random n-body systems. However, Rapoport and Horvath linked the explanation of this second class of systems to the probability methods applied in statistical mechanics. Thus, given this context we can adjust the quoted statement and narrow the group of chaotic complexity to systems of equilibrium gases and fluids in random motion only.

The crucial point here is that science in the tradition of the mathematical methodology developed in physics had only comprised systems in terms of organised simplicity as well as chaotic complexity. Thus, the exact science approach in its classical form did not cover the phenomena comprising the category of organised complexity. It is obvious that this class covered all the diverse phenomena of the living world that have been the early starting point of systems thinking in the organismic understanding of Bertalanffy. Concepts like a living organism, survival, reproduction, development, life cycle, etc. simply had no place within the early exact sciences. In fact, they had to be excluded from research to free educated thinking from the teleological Aristotelian philosophy and to prepare the ground for modern science (see [Rapoport, 1966, 3]). Thus, we can claim the category of organised complexity as the fully fledged ('genuine') field of the system idea.

GST faced the problem of defining systems in a way that did not exclude systems exhibiting organised complexity. In a first step, Rapoport came up with a working definition of "system" that relates to the analytical tradition of this concept: "I accept the definition of a system as (1) something consisting of a set (finite or infinite) of entities (2) among which a set of relations is specified, so that (3) deductions are possible from some relations to others or from the relations among entities to the behavior or the history of the system" [Rapoport, 1968, 453]. The analytical bias of this definition is due to the following characteristics. First, it carries out a total separation of substance and form; the criteria for something being a system are that it consists of definable sets of constituent parts as well as of their relations that together make up a certain structural form, regardless of the specific content of the phenomenon in question. Second, this comprehension, then, makes it possible to come up with the application of formal mathematical language for a more or less rigorous representation of a system's structure and dynamics via deductive means. Thus, in the exact sciences, a material system is defined by a series of numbers that change in time, i.e. by the variables ascribed to the constituent parts (e.g. mass, distance, concentration); and the technique of differential calculus is used to make precise statements about how these quantities and their rate of change are related. Together, these two things make up a mathematical model in the form of a system of differential equations.

An analytical approach understands phenomena as systems by analysing the whole into measurable quantities and their deterministic or probabilistic law-governed interactions and re-synthesising from these elements; in that sense it follows the core procedure of the exact mathematical sciences. However, in order to include systems of organised complexity too, a second step was necessary to get rid of shortcomings in the traditional approach due to the one-sided hypostatising of linear deterministic relations. This aspect refers to the disentanglement of the analytic approach from the idea of a mere additive and linear thinking. In this regard, Rapoport clearly pointed out that the classical *dictum* of the holists — "the whole is greater than the sum of its parts" — can be interpreted as a rejection of the additivity notion. But at the same time he argued that the additive way of thinking should not be equated with analytical thinking *per se*, but with "pitfalls of elementalism" that are often used for a mistaken interpretation of analytic thinking [Rapoport, 1986, 9-11]. Especially, these pitfalls comprise the linkage of causes and effects in a simple linear chain of consecutive events, a neglect of the context dependency of stated causal relationships, and a fixation on unidirectionality of causation, thus, a non-awareness of the fact that cause-effect relations are often invertible. The origins of these pitfalls probably lie in an indefensible generalisation of a common analytical techniques of classical science. In this context, the idea of analysis was to isolate and understand the parts of a phenomenon and, subsequently, to add up the results in a mathematically linear way. Rapoport and Horvath stressed that "... in presenting the analytical view, it is natural to fall into the language appropriate for describing a mechanism — 'the working of the parts'. The implied hope is that it is possible to 'build up' the understanding of a complexity by 'superimposing' the workings of the various parts" [Rapoport and Horvath, 1959, 87]. This hope of the classical picture can be fulfilled in the case of rather simple mechanical systems. Typically, the latter can be exemplified as wholes where the operation can be broken up into a temporal chain of events, all connected by determinate causal relations (like machines); or where the possibility to add up the fundamental quantities is given, as perturbing influences are so small that they may be widely neglected (as in the solar system case).

Thus, concerning the system category of organised complexity, Rapoport stressed that the emphasis lies "... on the circumstances that the addition of a new entity introduces not only the relation of this entity to all others but also modifies the relations among all the other entities" [Rapoport, 1966, 5]. Consequently, we have the notion of a system as an interwoven network of relations beyond linearity and additivity. Nonetheless, a mathematical approach to such a notion is possible and brought forward by Rapoport as well. "Mathematically an 'organized complexity' can be viewed as a set of objects or events whose description involves many variables, among which there are strong mutual interdependencies, so that the resulting system of equations cannot be solved 'piece-meal' ..." (ibid., 4).

A complex organised system — standing in the organismic tradition of system thinking — might be comprised by a qualitative approach to systems. From this "soft" perspective, Rapoport defines an organismic system: "According to a 'soft' definition, a system is a portion of the world that is perceived as a unit and that is able to maintain its 'identity' in spite of change going on in it" [Rapoport, 1970, 22]. We can immediately notice the difference to the analytic notion. Whereas in the latter an ascription of specific values to an analytical mathematical structure constituting a state is the central definition criterion, the organismic conception stresses the idea of identity. A system in this regard is essentially an identity-preserving phenomenon, something that keeps constancy amidst change; this idea underlies the organisimic notion of a whole and it reveals itself in an act of cognition.

From this standpoint, the living organism is the prototype of a system. The material composition of an organism is constantly changing through metabolism, yet the organism maintains its identity. In fact, this is true of organisations and institutions of the human social realm too.

While the comprehension of organismic phenomena in a teleological way explicitly brought into focus the whole, it had to be replaced since its anthropomorphic underpinning did not conform to scientific standards of thinking. The question, then, was how the essential idea of goal directedness could be represented in a formal way and, as such, be freed from teleological connotations. In this regard, "A goal in its general sense is simply some end state to which a system tends by virtue of its structural organization ..." [Rapoport, 1970, 8].

3.2 Evolution

Self-organisation has a diachronous and a synchronous aspect. The diachronous aspect refers to the evolution of systems, the synchronous aspect to systems' hierarchies.

Regarding the diachronous aspect, it goes without saying that Bertalanffy shared the basic assumption that reality is dynamical. As a biologist he supported the idea of evolution.

Concerning development as part of evolution, Bertalanffy describes an inherent trend toward the rise of complexity. He does so by framing his second organismic principle. It is the "principle of progressive organization". It comprises a segregation process as a disintegrative trend (specialisation or differentiation which is well-known in Niklas Luhmann's social systems theory) and centralisation as an integrative trend. Later, Bertalanffy called these processes, after Richard Woltereck, "anamorphosis" [Bertalanffy, 1949].

As Davidson [1983] points out, Bertalanffy said that both the scientific view and a religious, mystical view reveal the same idea when the first is referring to *homo sapiens* as by now the ultimate product of terrestrial evolution and the second is underlining that it is God who becomes aware of himself in the course of evolution. This assumption anticipates the idea of system theorist Bela H. Banathy [2000] and others that circumscribes the shift from the evolution of consciousness towards a conscious evolution.

3.3 Hierarchy, Emergence and Downward Causation

Regarding the basic structural assumption of Bertalanffy's early GST, it is clear that he laid the foundations for what was later called "hierarchical system theory". Hierarchy has two fields of application:

- one field is within a system where the term circumscribes the part-whole relationship (intrasystemic aspect) and
- another field comprises the relation between systems that represent different stages of the overall evolution of nature and society (intersystemic aspect).

3.3.1 Intrasystemic Hierarchy

Bertalanffy takes Nicholas of Cusa's idea "*ex omnibus partibus relucet totum*" ("each part reflects the whole") as point of departure. It is well justified to look upon this assumption as something that later on became known as "downward causation" (cf. [Campbell, 1974]) which is closely related to emergent effects. As early as in 1928 Bertalanffy wrote with regard to the organism: "The characteristic of the organism is first that it is more than the sum of its parts and second that the single processes are ordered for the maintenance of the whole" ([Bertalanffy 1928, 305]). Bertalanffy discovered that there is "maintenance of the organized system in a dynamical pseudo-equilibrium through the change of its components" [Bertalanffy, 1932, 309]. By that he was aware of what Haken (e.g., [1996]) called, initially, the "enslaving principle" by which he meant the ordering of parameters along different levels evident not only in living systems but also in material ones like Laser light — the higher the level the slower the development of the respective parameter. It is noteworthy that this has also been partly paralleled by Maturana and Varela's concept of autopoiesis by which they referred to living systems — the difference is that they abandon the hierarchical perspective, since they talk about networks of nodes only that produce each other and leave out the question of the whole (e.g., [Varela, *et al.*, 1974]). But here Bertalanffy gets quite clear: organismic conceptions in biology "assert the necessity of investigating not only parts but also the relations of organization resulting from a dynamic interaction and manifesting themselves by the difference in behavior of parts in isolation and in the whole organism" [Bertalanffy, 1950, 219-220].

Thus Bertalanffy distinguishes in this respect between two levels, that is, the level of parts and the level of the whole, and he distinguishes between the dynamic interaction of parts, and the relations of organisation. And it seems obvious that he locates the interaction on the parts' level and the relations on the whole's level. And he considers the following relationship between the interaction and the relations: the relations, on the one hand, result from the interaction and, on the other, are manifest in the behaviour of the parts in that the behaviour is different from the behaviour when in isolation. It follows that there are two processes in organisms/organisations/systems: one bottom-up in which interaction on the level of the parts result in relations on the level of the whole, and one top-down in which relations on the level of the whole manifest themselves on the level of the parts, viz., in their behaviour.

Hence Bertalanffy shared both the concept of emergence and the concept of downward causation in the following ways.

- As to emergence, he repudiated its metaphysical connotation. But nevertheless he emphasises the difference in quality between the two levels and supposes the irreducibility of the quality of the higher level to the quality of the lower level when he claims "with assurance that even if we would completely know the single substances in the organism, this problem of organization, the essential trait of life, would not yet be solved" [Bertalanffy, 1932, 306]. "But if we know *all* the components brought together and *all the relations existing between them*, then the higher levels are derivable from their components" [Bertalanffy, 1932, 308] which is obvious since the relations express the new quality and belong thus to the higher levels.

- As to the dominance exerted from the higher levels on the lower levels, he reinstated final causality in a certain respect only, for he equally discountenanced teleology. "What in the whole denotes a causal equilibrium process, appears for the part as a teleological event" [Bertalanffy, 1929, 306].

3.3.2 Intersystemic Hierarchy

Bertalanffy abstracted these ideas from applying to living systems exclusively and affirmed that "we are certainly able to establish scientific laws for the different levels or strata of reality" [Bertalanffy, 1950, 233-234]. So he could arrive at the conclusion: "Reality, in the modern conception, appears as a tremendous hierarchical order of organized entities, leading, in a superposition of many levels, from physical and chemical to biological and sociological systems." "When emphasizing general structural isomorphies of different levels, it asserts, at the same time, their autonomy and possession of specific laws" [Bertalanffy, 1950, 234]. "Speaking in the way of gross oversimplification", Bertalanffy conceived of three major levels:

"physical nature; organisms; and human behavior, individual and social." And here "the notion of emergence is essentially correct: each higher level presents new features that surpass those of the lower levels" [Bertalanffy, 1959, 67].

In that way his idea of the hierarchical ordering of processes within living systems extended to the conception of a hierarchical order of system classes, based on their level of organisation and complexity which is, in turn, related to the evolutionary stage systems inhabit. For example, Boulding suggested an accordant hierarchical scheme, comprising an eleven-level construction [Boulding, 1956, 14-16]. According to Boulding, one central function of such a scheme was to give some ideas of gaps in both theoretical and empirical knowledge. Regarding the question of the availability of adequate formal theoretical models, it is worth stating that according to him such representations merely extended up to the fourth level of his scheme — i.e. from the level of static structure through the clockwork level of simple dynamic systems through the level of the control mechanism to the level of self-maintaining structure. The realm of definitive living systems started not before the fifth level; hence, that meant that no genuine formal theoretical model for the higher levels existed. Nevertheless, as each level incorporates those below it, "... much valuable information and insights can be obtained by applying low-level systems to high-level subject matter" [Boulding, 1956, 17]. Later Koestler coined the terms "holarchy" and "holon". "Holarchy" denotes the hierarchy of "holons". A holon is a self-contained whole made up of its subordinate parts, while, in return, it is itself a part dependent of another whole [Koestler, 1967].

4 ETHICAL IMPLICATIONS

Concerning the integration of the descriptive and normative approach, GST was from the outset considered a tool for intervention in the world we live in. In the social sphere it is so far social systems that are blamed for undesired and unintended consequences. "Contemplating contemporary history in the making, it is difficult to ascribe its irrationality and bestiality solely to individuals. Rather we seem to be victims of 'historical forces' – whatever this may mean. Events seem to involve more than just individual decisions and actions and to be determined more by sociocultural 'systems', be these prejudices, ideologies, pressure groups, social trends, growth and decay of civilizations, or what not" [Bertalanffy, 1968, 6].

Since Bertalanffy grew up in the Viennese post-*fin-de-siècle* atmosphere of cultural criticism, it is unsurprising that he shared the descriptions of crises that were said to abound in all spheres of life and the norms that were implicit in these descriptions. For Bertalanffy crisis also got a grip on science. In 1927 he wrote, "the mechanistic epoch . . . , whose hope it was to create a happy future for mankind by means of mechanics and technology . . . may today come to its end" [Bertalanffy, 1927a, 285]. When announcing what we today, after Kuhn, would term a paradigm shift in scientific thinking, albeit one that comprises all science — natural sciences and social and human sciences — Bertalanffy expresses in his

1928 *Kritische Theorie der Formbildung* his value-laden stance and his conviction that there is an intrinsic link between science and society: "The technical age is about to become disgusted with itself — let us hope that it will be followed by an organismic one, which opens new doors to the future of mankind" [Bertalanffy, 1928, 288]. The mediator between science and society is *weltanschauung.* "The organismic conception, first grown on the ground of biology, is in the position to broaden to a general world view" [Bertalanffy, 1934].

At the end of his life, and after the Nazi atrocities of World War II, which he meanwhile had had to witness, Bertalanffy shared the very same idea: "The nineteenth and first half of the twentieth century conceived of the *world as chaos.* ... The mechanistic world view, taking the play of physical particles as ultimate reality, found its expression in a civilization which glorifies physical technology that has led eventually to the catastrophes of our time" [Bertalanffy, 1968, 198]. Bertalanffy then devoted his thoughts to the future of humanity. However, he grew more pessimistic as to the fate of civilisation and humankind. He admitted that Oswald Spengler in his writings had omitted that our civilisation has disposal of the technologies required for overcoming many plagues that have beleaguered mankind and that we are empowered today to act upon challenges globally. But he did not rule out the possibility of extinction. "We know precisely and scientifically what the effects of pollution, waste of natural resources, the population explosion, the armaments race, etc., are going to be" [Bertalanffy, 1968, 6]. However: "We seem to follow some tragic historical necessity," though there would be an evolutionary necessity to fight the dangers coming along iwith utilitarian common sense and the military-industrial complex.

This possible design of future humanity is the meaning of "unity-through-diversity" in the context of the social task of the GST. Bertalanffy identified the causes of environmental pollution, waste of natural resources, population explosion, arms race, and so on, not in psychic features of wicked people that are in power, but in systemic features of the civilisation, in the design of socio-cultural systems. System theoretical insights are to be applied to contribute to that aim. Bertalanffy's GST is a humanistic one. Thus all his descriptions of humans and social systems serve the function to help to formulate guidelines for acting in ways that support humane norms and values.

In that respect, Bertalanffy's co-workers in GST joined him. According to Bertalanffy it was the reductionistic robot view of humans in behaviourism that was closely connected to the militarism of the post-war era. GST inspired the development of peace research. Boulding and Rapoport established the Center for Peace Research and Conflict Resolution at the University of Michigan in 1956 [Hammond, 2003]. Both are well-known for their protests against the Vietnam War and for their criticism of the military-industrial complex of the United States. Rapoport is known for his development of non-zero sum models in game theory [Rapoport, 1992]. Boulding was among the first economists to include ecological considerations in economic theory. By that it is clear that GST did not abide by the *dictum* of value-free science.

5 CONCLUSION

"GST" turns out to be the name for systems science *in statu nascendi* from which many ramifications followed in the course of the history of systems science. The complex systems approach as the most recent development of the new paradigm seems to have more in common with the original ideas than other ramifications and more than today acknowledged. This holds for epistemological, ontological and ethical aspects of philosophical implications as well.

In its aiming for generalisations, GST is thus heading towards a state of science called in our days "trans-disciplinarity". The term "trans-disciplinarity" is used to define a concept that goes beyond the meaning of multi- and even inter-disciplinarity. While multi-disciplinarity would mean the unrelated coexistence of mono-disciplinary accounts and inter-disciplinarity the casual establishment of relations between mono-disciplines without having feedback loops that have a lasting impact on their repertoire of methods and concepts, trans-disciplinarity comes into play when each discipline is engaged in the collaborative undertaking of constructing a common base of methods and concepts, of which its own methods and concepts can be understood as kind of instantiations. Trans-disciplinarity does thereby not mean the abolition of disciplinary knowledge but grasping for a bigger picture. In fact, GST and systems science, aware of the aims set out by GST, are the trans-disciplinary science *per se*.

BIBLIOGRAPHY

[Ashby, 1952] W. R. Ashby. *Design for a brain: the origin of adaptive behavior*. Wiley, New York, 1952.

[Banathy, 2000] B. H. Banathy. *Guided Evolution of Society: A Systems View*. Kluwer, New York, 2000.

[Bertalanffy, 1927a] L. von Bertalanffy. Über die neue Lebensauffassung. In: Annalen der Philosophie und philosophischen Kritik 6, 250-264, 1927.

[Bertalanffy, 1927b] L. von Bertalanffy. Über die Bedeutung der Umwälzungen in der Physik für die Biologie. Studien über theoretische Biologie II. *Biologisches Zentralblatt*, 47, 653-662, 1927.

[Bertalanffy, 1928] L. von Bertalanffy. *Kritische Theorie der Formbildung*. Borntraeger, Berlin, 1928.

[Bertalanffy, 1929] L. von Bertalanffy. Vorschlag zweier sehr allgemeiner biologischer Gesetze. Studien über theoretische Biologie III. *Biologisches Zentralblatt*, 49, 83-111, 1929.

[Bertalanffy, 1932] L. von Bertalanffy. *Theoretische Biologie*, 1. Band. Gebrüder Bornträger, Berlin, 1932.

[Bertalanffy, 1934] L. von Bertalanffy. Wandlungen des biologischen Denkens. *Neue Jahrbücher für Wissenschaft und Jugendbildung*, 10, 339-366, 1934.

[Bertalanffy, 1937] L. von Bertalanffy. *Das Gefüge des Lebens*. Teubner, Leizig, 1937.

[Bertalanffy, 1949] L. von Bertalanffy. *Das biologische Weltbild*. Francke, Bern, 1949.

[Bertalanffy, 1950] L. von Bertalanffy. An Outline of General System Theory. *British Journal for the Philosophy of Science*, 1(2), 134-165, 1950.

[Bertalanffy, 1955a] L. von Bertalanffy. An essay on the relativity of categories. *Philosophy of Science*, 225, 243-263, 1955.

[Bertalanffy, 1955b] L. von Bertalanffy. General System Theory. *General Systems*, 1, 1-10, 1955.

[Bertalanffy, 1959] L. von Bertalanffy. Human Values in a Changing World. In A. H. Maslow (ed.), *New Knowledge in Human Values*, Harper & Brothers, New York, 65–74, 1959.

[Bertalanffy, 1965] L. von Bertalanffy. Zur Geschichte theoretischer Modelle in der Biologie. *Studium Generale*, 18, 290-298, 1965.

[Bertalanffy, 1968] L. von Bertalanffy. General System Theory. Foundations, Development, Applications. Braziller, New York, 1968.

[Bertalanffy, 1972a] L. von Bertalanffy. Response. In E. Laszlo (ed.), *The relevance of General Systems Theory — Papers presented to L. von Bertalanffy on his 70th birthday*, Braziller, New York, 181-191, 1972.

[Bertalanffy, 1972b] L. von Bertalanffy. The History and Status of General Systems Theory. In G. Klir (ed.), *Trends in General Systems Theory*, Wiley, New York, 21–41, 1972.

[Boulding, 1956] K. Boulding. General system theory: the skeleton of a science. *General Systems*, 1, 11-17, 1956.

[Boulding, 1956/2004] K. Boulding. General Systems Theory — The Skeleton of Science. *Management Science.*.2(3), 197-208, 1956. Reprinted in E:CO. 6(1-2), 127-139, 2004.

[Campbell, 1974] D. T. Campbell. 'Downward causation' in hierarchically organized biological systems. In F. J. Ayala, T. Dobzhansky (eds.), *Studies in the Philosophy of Biology.* MacMillan, London, 179-186, 1974.

[Davidson, 1983] M. Davidson. *Uncommon sense: The life and thought of Ludwig von Bertalanffy, Father of general systems theory.* Tarcher, Los Angeles, 1983.

[Gerard, 1958a] R. W. Gerard. Concepts and principles of biology. *Behavioral Sciences.* 3(2), 95-102, 1958.

[Gerard, 1958b] R. W. Gerard. Units and concepts of biology. *Behavioral Sciences.* 3(2), 197-206, 1958.

[Hall and Fagan, 1956] A. D. Hall and R. E. Fagan. Definition of system. *General Systems*, 1, 18-28, 1956.

[Haken, 1996] H. Haken. The slaving principle revisited. *Physica D*, 97, 95-103, 1996.

[Hammond, 2003] D. Hammond. *The Science of Synthesis. Exploring the Social Implications of General Systems Theory.* University Press of Colorado, Boulder, 2003.

[Hofkirchner, 2005] W. Hofkirchner. Ludwig von Bertalanffy. Forerunner of Evolutionary Systems Theory. In J. Gu and G. Chroust, (eds.), *The New Role of Systems Sciences For a Knowledge-based Society*, Proceedings of the First World Congress of the International Federation for Systems Research, Kobe, Japan. CD-ROM (ISBN 4-903092-02-X), 6, 2005.

[Klir, 1972] G. Klir, ed. *Trends in General Systems Theory.* Wiley Interscience, New York, 1972.

[Koestler, 1967] A. Koestler. *The ghost in the machine.* Hutchinson, London, 1967.

[Laszlo, 1972] I. Laszlo, ed. *The Relevance of General Systems Theory Papers Presented to Ludwig von Bertalanffy on His Seventieth Birthday.* George Braziller, New York, 1972.

[Müller, 1996] K. Müller. *Allgemeine Systemtheorie.* Westdeutscher Verlag, Opladen, 1996.

[Pouvreau and Drack, 2007] D. Pouvreau and M. Drack. On the history of Ludwig von Bertalanffy's "General Systemology", and on its relationship to cybernetics. *International Journal of General Systems.* 36(3), 281-337, 2007.

[Rapoport and Horvath, 1959] A. Rapoport and W. Horvath. Thoughts on organization theory and a review of two conferences. *General Systems*,, 4, 87-93, 1959.

[Rapoport, 1966] A. Rapoport. Mathematical aspects of general systems analysis. *General Systems*, 11, 3-11, 1966.

[Rapoport, 1968] A. Rapoport. General systems theory. *International Encyclopedia of the Social Sciences*, 15, 452-458, 1968.

[Rapoport, 1970] A. Rapoport. Modern systems theory: an outlook for coping with change. *General Systems*, 15, 15-25.

[Rapoport, 1986] A. Rapoport. *General system theory: essential concepts & applications.* Cybernetics and systems series. Abacus Press, Tunbridge Wells, Kent, 1986.

[Rapoport, 1992] A. Rapoport. *Peace — an idea whose time has come.* University of Michigan, Ann Arbor, 1992
Die systemische Sicht. In A. Rapoport, H. Lübbe, M. Thürkauf (eds.), *Weltbilder - Wissen und Glauben*, Hochschule St. Gallen für Wirtschafts-, Rechts- und Sozialwissenschaft, St. Gallen, 3-27, 1992.

[Richardson, 2004] K. Richardson. Introductory remarks. E:CO, 6(1-2), 127, 2004.

[Spengler, 1973] O. Spengler. *Der Untergang des Abendlandes.* Beck, München, 1973

[Varela, *et al.*, 1974] F. Varela, H. Maturana, R. Uribe. Autopoiesis: The organization of living systems, its characterization and a model. *BioSystems*, 5, 187-196, 1974.

[von Uexküll, 1920] J. von Uexküll. Umwelt und Innenwelt der Tiere. Springer, Berlin, 1920.

CONCEPTUALISING REDUCTION, EMERGENCE AND SELF-ORGANISATION IN COMPLEX DYNAMICAL SYSTEMS

Cliff Hooker

1 SETTING THE SCENE[1]

Fortunately for science, algorithmically compressed representation based on meso-accessible data goes some distance in this cosmos: diverse and complex data can often be shown to express the same, often relatively simple, law and diverse phenomena are often underlain by the same set of laws. Clouds, snow, ice and rain, for instance, all turn out to be re-arrangements of water molecules, just as Democritus had hoped, governed by their common molecular properties and the basic laws of physics. Science provides algorithmically compressed representation whenever it is able to organise information under laws: a single set of initial conditions added to a single law statement in equation form (e.g. Newton's 2^{nd} law) suffices to encapsulate an entire dynamics. A final compression occurs whenever the solution of the characterising dynamical equations has a closed-form analytic representation in terms of known mathematical functions, such as form the core mathematical dynamics of simple systems.[2]

Maximising the reach of explicit compression sums up the drive supporting the orthodox vision of the world as formal machine and scientific method as logical algorithm for the construction of its deepest compressed description. Physics is inherently universal: it is the study of those laws of nature (if any) that hold for all physical states. Over the 300+ years since Newton, this study has proven remarkably successful, until now we speak hopefully of including even the universe itself under 'physical states' and obtaining a Theory of Everything. On that basis scientists understandably erected an ideal: show the physics algorithmic compression

[1]Key terms used here will be explained in the sequel.

[2]McAllister [2003] points out that, since actual empirical data will in general contain noise as well as systematic information, it cannot in general be compressed. Thus the compressible data on which science runs has to be actual data with noise removed (along with any exogenous systematic biases). Of course, achieving such cleaned up data can prove difficult — it is part of what the theory of errors in scientific method is about (cf. [Farrell and Hooker 2009]); but this is an epistemic issue and it is set aside here. It turns out that scientific method is much more complex (and interesting!) than the neat logical algorithms to which the philosophers had hoped to reduce it. Instead it requires many artful decisions if it is to be executed powerfully (cf. [Farrell and Hooker, 2007a; 2007b; Hooker, 1995]. To reduce it to algorithms is at least to reduce all judgement, including creativity, to algorithms, a far-future prospect at best.

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

universal, that is, reduce all other sciences to applications of the laws of physics. One form of this vision is already clear in the Greek atomist tradition where complexity is constrained to spatio-temporal arrangements; another equally venerable form is found in the Greek plenum, later field, theory tradition where complexity also involves dynamical creation and annihilation.[3]

To that vision philosophy of science has contributed a criterion for its realization, crafted with the logical tools whose use it takes to be obligatory: to *re*duce X (laws, phenomena in some other science) to physics, show how to *de*duce X from the laws of physics plus any required initial and constraint ('boundary') conditions, that is, show how all the other sciences are applied physics.[4] Within this perspective reduction represents increased compression since hitherto independent laws are now represented as deductive consequences of their reducing laws. Conversely, emergence represents a constraint on compression. It is helpful to approach emergence and reduction from this perspective because it ultimately makes it easier to recognise the importance of the shift to dynamical criteria for emergence and reduction, permitting these phenomena to be characterised independently of such formal issues.

This orthodox vision has to date carried us so remarkably far in knowing our cosmos as to seem an incredibly unlikely miracle for such finite, fallible creatures as us. It is true that we have played our part in this, massively enlarging our observational niche through invention of instrumental prostheses and massively increasing our logical machinery through the invention of mathematics and computation. But we have still needed the enormous advantage of a penetrable world to work with.[5] Moreover, note immediately this limitation to the maximally explicit version of the vision: the dynamics of many complex systems have no analytic representations and so cannot be given explicit analytic compressed representations, their detail can at best be represented extensionally by computational simulation. And in fact every era since Newton has worried about what might be the principled limits of the compression programme. The last century was consumed by trying to understand the implications of quantum mechanics — itself a triumph of the compression programme — for limits on expanding our reach any further down in scale (cf. note 3).

The emergence of complex systems presents a new challenge to unlimited compression, this time in the form of dynamical emergence and condition-dependent laws. Dynamical emergence cannot yet be brought within our analytical dynamics, because it represents a change in the currently fundamental dynamical form. In consequence, the laws of the emergent dynamics, whose form may depend on the particular constraint conditions, must then be separately described and applied. (This is so even where the occurrence of emergence is predictable.) Whence these

[3]See [Hooker, 1973; 1974]. It is presently obscure whether QM represents a 3^{rd} alternative, see [Hooker, 1973; 1991].

[4]In recent times a major expression of this ideal was the Encyclopaedia of Unified Science, see [Neurath *et al.*, 1971].

[5]Just how much luck is needed is shown, e.g., by what is required to get astronomy, and hence Newtonian mechanics — our first sophisticated theory — started,.see [Hooker, 1994].

laws and their generative conditions cannot be further compressed. (Moreover, at least all self-organisation involving a critical point process is evidently computationally impenetrable and so cannot be more than partially simulated either.) In short, limits to the reach of compression are substantially determined by the complexity of dynamics in relation to our mathematical representation tools. These limits bite at the same location as the 'scandal' of complex systems (see below). But it does not follow that they must be expressed solely or even primarily in terms of compression.

Reduction is concerned first with ontology (that is, with what exists), in particular with the ontological relationship between phenomena described in apparently different ways. What, for example, is the relation between clouds, snow, ice and rain, and water molecules in various dynamical arrangements (that is, under various initial and constraint conditions)? And, more challengingly, what is the relation between physiologically described function and biochemically described dynamical states and processes? The obvious response to make in each case is that the two are one and the same; that, for example, aerobic cellular respiration is nothing but ATP synthesis through glycolysis, Krebs cycling and electron transport. This is reduction by identification. The physiological function of respiration is identically reduced to, is identical to and so nothing other than, the dynamical system process of ATP synthesis through glycolysis, Krebs cycling and electron transport.[6] And this ontological relationship hinges on the dynamics involved.

Because relationships are expressed in language, there is the issue of how identifying descriptions of the putative two phenomena relate under reduction. The answer must ultimately be that they refer to the same thing (co-reference) but establishing the conditions for that is non-trivial and can be controversial. And the answer is bound up with that to the epistemic issue of what warrants affirming reduction. The answers to these issues, I suggest, must ultimately appeal to dynamical criteria, not only to purely logical criteria as usually presumed by philosophers.[7] Roughly, reduction obtains when the same one dynamics occurs and its affirmation is warranted when there is sufficient support for affirming same dynamics. It is then clear that the prevalence of condition-dependent dynamical laws [Hooker-a, b, this volume] means that no attempt to spell out these conditions in abstract generality will suffice, the identifications will always need to invoke the specific local systems conditions that determine the dynamics. Nagel's general appeal to some kind of abstract deductive relationship ultimately remains, but only when it is clarified dynamically and refracted through the dynamical models;

[6]All this assumes that the biochemical systems models involved are empirically supported and predictively and explanatorily adequate, an assumption made throughout this discussion The issue of when and why that assumption is reasonable is again just the general issue of the nature of scientific method at large, see note 2. Also, this essay is concerned with empirical functions generally, not with the biological notion of proper function. The proper treatment of this latter notion is in fact given in terms of the notion of biological autonomy, see [Hooker-a, this volume, section 4.1.1].

[7]This was originally argued in [Hooker, 1981]; see further below.

it is too weak to stand by itself.[8] Moreover, only through appeal to dynamics can the relation between a new emergent existent and its constituting components be explained so as to coherently both give distinctive existence to the emergent entity and not compromise the fundamentalness of the components, especially for self-organised emergents (see below). And only in that way can the concomitant subtle entwinement of emergence with reduction to yield a coherent naturalism be achieved (see below). This is why complex systems are so important to reduction and emergence: uniquely in them we find the subtle dynamical relationships that confront our formal efforts and drive us to improved understanding.

2 REDUCTION IN COMPLEX SYSTEMS: THE BASICS

There is a large philosophical literature on reduction in science, some of it proclaiming it and much arguing against it. The latter is especially prevalent in the domains of biology and other sciences concerned with internally complex system components where functionality (e.g. respiration) and its more intentional teleological forms is important to system integrity. Yet, from a scientific point of view it would be anomalous to claim anything less than a reduction, for example to claim instead just a correlation between the occurrence of functional and biochemical systems properties. Doing that would leave unexplained duplicate realities, one functional and the other dynamical. Against the advice of Occam's razor, it would leave two realms mirroring each other but running in parallel, for no reason more substantive than the different descriptive languages used, the one of functions (purely physical, but ultimately also including strategies, purposes, intentions and communication) and the other dynamical. Though among the debaters, in what follows I try to briefly summarise the state of philosophical debate from a commonsense scientist-friendly point of view, in order to focus on the underlying substantive issues at stake, especially those that concern complex systems.

General objections. Perhaps surprisingly, one group of philosophical objections to reduction in general argues that correlation must be accepted because identification is impossible. These arguments largely turn on semantic (meaning) considerations. To-be-reduced states are held to be characterised by novel properties that

[8]Recently, Batterman tried to provide an alternative formal analysis of inter-theory reduction that was both more specifically tied to theoretical structure and yet still fully general by appealing to mathematical asymptotics, but this too ultimately fails because it parts company with the dynamics involved. See [Batterman, 2002] and for the critique see [Hooker, 2004]. Batterman appeals to the fact that a class of singular asymptotics shares the same formal structure, claiming that universality as the basis for nomic force. While not denying the mathematical facts, [Hooker, 2004] argues that this class covers nomicly disparate cases, from changes in molecular interactions through those in non-interacting light wave structures to transforms of kinematical possibility structures, and it is hard to see what nomic basis these disparate cases could share in common. Batterman's position requires understanding the universality as a formal artifact and it is not clear how to present a nomic basis for it in those terms or what its ramifications might be elsewhere in physics. This remains an unresolved issue.

are not properties of the reducing substrate, e.g. macroscopic solidity and rigidity, much less colour, have been considered not properties of molecules, individually or collectively. Then the two sets of properties are held to have different meanings. And then it seems logically impossible to deduce the former from descriptions of the latter, as identificatory reduction requires. Again, talk of functioning, like respiring, the argument goes, has a very different meaning from talk of biochemical states and processes, so the two can never be identified, even if they are correlated.[9] The proper response to these kinds of objection is to point out, first, that they rely on apriori claims about the fundamentalness of semantics whereas what is known scientifically about language suggests that current semantics are better treated as themselves shifting dynamical emergents, not a priori constraints. In this spirit, second, there is an attractive alternative semantic basis to hand that obviates these problems, namely a same-dynamical-role criterion of property identity. This in turn supports a same-dynamics criterion of thing (object or process) identity and thus the identification of functions with their corresponding dynamical processes.[10]

Another group of arguments turn on the fact that the parallel mirroring is often not precise. Often there will be particular phenomenological conditions (for example, 'respiration') that do not nicely reduce to exactly corresponding underlying conditions (for example, ATP synthesis) of exactly the same scope. This is so because there are anaerobic organisms and various energy storage molecules, but it is also true because of complex dynamics. For instance, even Kepler's laws of planetary motion do not reduce exactly to a theorem of Newtonian mechanics, because planet-planet interactions produce small deviations from Kepler's generalizations. (This is the case under present conditions, but they may produce far larger deviations under other conditions, especially over long time periods.) Such complications will commonly arise wherever a more complex dynamics underlies more macroscopic/phenomenological observations. There are also cases where large mismatches occur. These are so large in the relationship of phlogiston chemistry to oxygen chemistry, e.g., that scientists deny that phlogiston exists even if its postulation served to codify a number of chemical relationships that survive the replacement. And there are intermediate cases, for example the imperfections

[9]Conversely, Nagel once argued that if water is defined as H_2O then it is reducible, but not if it is not. See [Collier and Muller, 1998].

[10]This is argued in [Hooker, 1981, Part II]. What we know scientifically about language is that it is a recent evolutionary development, is characterised by rapid dynamical shifts in vocabulary, syntax and semantics as historical conditions change, and has a fundamentally action-centred intentional basis. In this light, a same-dynamical-role criterion of property identity is a far more plausible basis for semantics than so-called speaker intuitions; these have time and again been shown to simply project communal habit or vivid personal experience as cosmic truths. Scientists have had to learn the hard way that our concepts, even our seemingly most basic ones like hardness, simultaneity and consciousness, have to be continually reconstructed as we learn more because the world proves to be so deeply counter-intuitive. In consequence, scientists are wedded to constructing a unified dynamical conception of the world, not to naïve linguistic intuitions. Definitions are treated as works-in-progress and settling the 'right' ones is left until after a field matures.

of the thermodynamics-statistical mechanics relation.

In all these and other cases, the proper scientific response is that a warrant for reduction is ultimately concerned with establishing the capacity to systematically replace one kind of description with another kind that is equally or more precise, and equally or more predictively and explanatorily powerful when embedded in its scientific context. This satisfies the key cognitive aims of science. Reduction by identification forms one extreme of the reduction spectrum, where component ontology as well as relational structure is conserved under the replacement. The other extreme is occupied by cases like phlogiston where some significant relational structure, but not ontology, is conserved.[11] Labels are only useful to the extent they clarify, so in this case either of two labelling schemes is satisfactory: (a) label the entire continuum 'replacement' and retain 'reduction' for the identificatory extreme, or (b) retain 'reduction' for the entire replacement continuum, 'identificatory reduction' for its identificatory extreme and 'replacement reduction' for the opposite extreme. Either scheme locates the increasing discrepancy that characterises the relationship as one moves away from the identicatory extreme. This is what concerns scientists who frequently rely on the reduced theory because it is (typically) simpler and more immediately measurable, but who are concerned with understanding and managing the errors involved in doing so.

Local objections. These general issues aside, there are also various 'local' objections to reduction to consider. An important part of the philosophical objection to specifically biological reduction, for example, has really been to geneticism, to the idea that organisms could be reduced to just a collection of genes and gene-determined traits. Modern biology supports this objection, DNA is one biochemical component among many — if with a distinguishable role — and it is the dynamical system of all of them that is the reduction candidate for physiology. Again, various kinds of systems, e.g. those showing path dependencies, including systems whose parts play roles that depend on the specific system history, have been thought to raise difficulties for reduction because of their historical individuality. But it is easy to construct simple, clearly physical machines that also have those features, removing these objections.[12]

Objections to functional reduction. Setting aside such local objections as well, there remains only those objections that are specific to reduction of functions to systems dynamics. Here too there are various general objections of the semantic and mismatch kinds to deal with. One common semantic objection argues that if some property P (e.g. 'pumps blood') is defined as having the causal role or

[11]Beyond that, sheer discontinuous replacement would occur, but it is hard to think of a substantial case in science where the replaced theory once had significant support (as opposed to simply being one of a number of initial hypotheses about some matter that were eliminated through experiment). For the replacement view see Part I of [Hooker, 1981] and, more informally, [Churchland, 1979]. Churchland's elegant overall strategy, more subtle but powerful than it may appear, is itself explained in [Hooker, 2006].

[12]Any machine that assigns activity to parts as a function of their past total work hours and present system demand schedule will present this feature — see [Hooker, 1981, Part III], cf. e.g. [Miller *et al.*, 2000].

function of producing Bs (outputs) from As (inputs) then P cannot explain the producing of Bs from As. But why not? The fact is that an A-to-B transform occurs, hence the corresponding function is real and the explanation of A-to-B transformation appeals to the transform process, hence to the function. Indeed, it is better to avoid this (mis-)use of definition, replacing it by factual characterisation. In science, definitions are typically derivative and approximate, fallible and temporary, not basic (cf. Pierce), simply ready useful characterisations.[13] A further common objection is that our commonsense day-to-day function talk is rather imprecise for marrying up to dynamical systems specifications, e.g. compare 'is boiling' with all the specific ways fluids may convect. Another objection is that vague function descriptions can seem to cut across what turn out to be the dynamical process distinctions. For example, bird flight involves a lift function and a propulsion function, but the two functions cannot easily be separated into dynamically distinct components, as they can in contemporary aircraft. Objections of both these sorts can be resolved through a little careful analysis of language.[14]

There is also an inherent under-determination by any function, taken in isolation, of its correct embedding (that is, with which specific dynamical processes it is identical). While this has sometimes been taken as a fundamental objection to reduction, it ultimately reduces to a pragmatic issue of sufficient data. The problem is nicely illustrated in the case of the output of a network of electrical generators having a frequency variation less than that of any one generator. Some kind of feedback governing (i.e. regulatory) process is at work, but is it a real governor or simply the functional appearance of one at the network level? This latter is possible because connecting the electrical generators in parallel automatically creates a phase-stabilising mutual interaction among them without the need for a real governor.[15] This question is resolved by gathering other data about the network — this is the point of the unification criterion below.

Conditions for function-to-dynamics reduction. Setting these objections aside as well finally brings us directly to the substantive conditions for function to dynamical process reduction. And here a complex systems framework plays an important role.[16] For generality of application within complex systems, functions are considered under their most general aspect as maps carrying inputs to uniquely specified outputs, even if they are colloquially labelled as 'pumps blood' and the like. They

[13]See note 10 and text. The metaphilosophical slogan here is: the bearable lightness of semantics. Since it is derivative and approximate, not basic, do philosophy by dynamics, not semantics, and let the semantics follow later, reflecting what has been dynamically established.

[14]See [Hooker, 1981, Part III] and, briefly, [Hooker 2004, Part V case I and case II end]. It would be possible to separate out static structural reduction — compositional reduction of one structure, say a rigid body, to another, say an atomic lattice — from functional reduction (cf. [Causey, 1977]). However, the overall issues turn out to be similar and most of the interesting cases are suppressed because they involve process-dependent structures, e,g, in the cell.

[15]For this example see [Dewan, 1976] and further [Hooker, 1981, Part III, pp. 508-511].

[16]This was recognized early by Hooker, see [1981], especially Part III, where the approach to reduction that follows was first developed. The construction there was intentionally cast in a complex systems context at a time when little philosophical attention was paid to this field (though not none, e.g. [Wimsatt , 1974] is an honourable exception).

are to be reductively identified with the dynamical processes that dynamically carry the (relevant, dynamically characterised) inputs to the (relevant, dynamically characterised) outputs, grouped together as the mechanisms that deliver the functionality (see 3 below). For example, cellular respiration, crudely globally specified, is the function that takes food and water as inputs and outputs stored energy and expired breath. Corresponding to this in the molecular description is a dynamical process — that is, a map carried by (biochemical) dynamical laws, constraints and initial conditions — that takes oxygen and glucose as inputs and yields ATP (and rarely, other forms of chemical energy storage) and carbon dioxide as outputs. Then the obvious requirement for identificational reduction is that the respiration functional map be embeddable into the corresponding biochemical process map without distortion (homomorphically embeddable). A further coherence condition is equally obvious: the collection of all such embedded dynamical maps, together with any non-functional data concerning the system, e.g. concerning its structure, should provide a single coherently unified biochemical cellular model that preserves or increases predictive and explanatory power.[17]

As the respiration example suggests, the dramatic example is that of cellular biology. On this score, Hooker [1981, Part III, pp. 515-517] considers the schematic reduction of Mendelian to molecular genetics from this point of view. Essentially, Mendelian genetics stands to molecular genetics as an input/output theory of some system (genes in, traits out) stands to a detailed internal dynamical theory of the system (DNA + cellular organisation in → biosynthetic pathways + multi-cellular developmental dynamics → spatially organised protein complexes out). This way of posing the relationship already shifts the focus from genes as isolatable objects to genes as functional units in complex system processes (cf. [Griffiths, Stotz, 2007]). Recalling the electrical governor example above, in the Mendelian case a sentence such as 'X has a gene for characteristic A which is dominant' would not be perspicuously analysable as 'X has a component gene y that causes X to be A and y is dominant', but instead as 'There is some causal process P within X such that P causes X to be A under conditions C and X has P because of C', where C specifies the operative cellular constraints, including genome structure. This is a crucial shift, it corresponds to the end of genes-as-phenotype-determining-objects and the emergence of what is today the exploding fields of systems and synthetic biology, where the focus is on the complex regulatory mechanisms constituting the biosynthetic pathways. It is these that will determine whether and how the

[17]See [Hooker, 1981, Part III] and, briefly, [Hooker, 2004, Part V]. The basic reduction requirement, that functional maps are mirrored by dynamical maps, is in fact just the application of Nagel's deductive reduction conception, rightly understood. Nagel [1961] shows how scientists arrive at reduction of a law L_2 or property P_2 of theory T_2 respectively to a law L_1 or property P_1 of theory T_1 by first showing how to choose conditions (real or idealised) under which it is possible to construct in T_1 a law L_1 or property P_1 that will mirror (be a relevantly isomorphic *dynamical* image of) the dynamical behaviour of L_2 or P_2. From that the reduction is shown possible through the identification of L_2 or P_2 with the mirroring L_1 or P_1. Indeed, rather than having priority, the requisite 'bridging' conditions can be deduced from the mirroring condition, and then asserted as identities on the basis that doing so will achieve a reduction, supported in that light by claims of spatio-temporal coincidence or appeal to Occam's razor.

cell will be recaptured as a complex dynamical system and the above coherence condition on successful reduction thus satisfied. As Hooker [1981, p.515] concluded two decades earlier: "The search for a reductive base for Mendelian genetics is now the search for the inner (in fact, molecular) mechanisms of genotypic-to-phenotypic production. ... genes are not things but complexes of mechanisms."

In that setting, and paraphrasing Hooker 1981, pp.515-516, what the reduction of Mendelian genetics to molecular genetics requires is that (1) the possibility, relative stability and conditions of change, of the cellular structures mediating the processes involved is explained by the basic biochemical laws, (2) as a result there is available a characterisation of the relevant initial and boundary conditions (other than structure) for any given system such that (3) the set of molecular mechanisms is unambiguously specified, (4) every true Mendelian sentence (in a suitably coevolved theory) has a corresponding condition realised within the complex of mechanisms and (5) for a specified Mendelian input/output relation and initial conditions a unique complex mechanism (biosynthetic pathway or complex of pathways) is selected such that (6) nomic Mendelian input/output relations (e.g. epistatic ones) are preserved and (7) the properties of the molecular model fit together in such a way that the ability to introduce Mendelian complex predicates such as 'is dominant' is explained, even though these predicates do not designate distinctive molecular properties at any level of structure. Though in 1981 they were just a gleam in the eye of molecular biologists, constructing such molecular models of complex mechanisms is what contemporary systems and synthetic biology are slowly making possible, aided by their high throughput experimental techniques. Despite its being in advance of the sequencing of genomes, constructing cellular models of its dynamical biosynthetic mechanisms is, as many biologists already knew in 1981 and Hooker reiterated, the really hard part of the reduction.

3 REDUCTION AND MECHANISM

This way of setting up the reduction criteria was designed to be appropriate for reduction within complex systems of the sort illustrated by respiration above, with reduction in simpler cases being simplified special cases. In fact, the reduction was explicitly designed to reduce functions to internal mechanisms as the specific kinds of processes that underlie functions (see [Hooker, 1981, Part III, p. 505]). Anticipating the later interest in mechanisms (note 19), mechanisms were understood there as law-like in their operation ("laws for the specific mechanisms"). And, as an explanatory requirement on their adequacy deriving from the further coherence condition on reduction (text to note 17), it was demanded that the operative mechanism laws should explain all of the empirical functional interrelations, e.g. dominance and epistatic relations among Mendelian genes (see [Hooker, 1981, Part III, p. 517]).

Following the example of respiration, specifying the mechanisms underlying functionality in complex systems can involve many different inputs and outputs appearing at different stages of a process and at different levels of a multi-level system, sophisticated coordinated constraint relationships and multiple phase-shifted feedback/forward relations. A more familiar illustration is the reduction of automotive engine functions to complex organised engineering relationships among inputs, outputs and internal components from the chemical fuel, mechanical drive and electrical regulatory sub-systems. Such complex dynamical processes are essential if functions based on global organisation of the kind engines and organisms display are to be properly captured since no simple sequences of causal production relations can capture the global interrelations that constitute them. Capturing global constraints requires instead a complex network of dynamical interrelations with the requisite closure pathway structure (see [Hooker-a, this volume, section 3]). In these cases the active components in these networks can themselves be altered by their roles, even destroyed and re-constituted by them (in cells, not current engines). Thus mechanisms involving such irreducible constraints are not specifiable in terms of some fixed set of components, but rather in terms of their constituting dynamical processes. It is in this context that investigation of reduction within detailed models for specific systems are essential and valuable.[18]

As this discussion suggests, the embedding criterion essentially captures recent conceptions of a function to mechanism reduction, reducing both the cell and multi-cellular organisms to complexes of mechanisms.[19] Bechtel [2007] contrasts traditional universal law centred conceptions of explanation, reduction and generalisation with those appropriate to mechanisms and this may seem to challenge the present approach. He characterizes mechanistic explanation not in terms of logical inference but in terms of showing how a phenomenon is produced. Reduction is construed, not as a matter of deriving one set of laws from another, but as showing how parts and operations within a mechanism enable the whole mechanism to respond to conditions in its environment in specific ways. Finally, he characterizes generalization not in terms of universally quantified linguistic statements but in terms of similarities between model systems and other instances which share many of the same parts, operations, and modes of organisation, albeit often with some changes. (In many cases, the relations between systems are understood as features

[18] See, e.g., [Boogerd *et al.*, 2005], cf. [Bruggeman, 2002; Boogerd *et al.*, 2007]. It is typical of physicists that they never deal with globally organised systems in their textbooks and hence tend to ignore global constraints and organisation, or assume that they can be reduced to separable momentary, local piece-wise interactions. If, for instance, the ambitions for condensed matter physics include direct modelling of the cell as simply a dynamical state like any other, as [Barham, 2004] supposes (cf. [Amaral and Ottino, 2004, sec. 3.2]), then they run directly into the issue of how to satisfy global organisational constraints, ones that are self-organised and adaptive at that. The challenge then is to understand how this project could possibly be carried out in a principled manner rather than ad hocly. Interestingly, Barham also doesn't see the problem. Moreover, he cites [Moreno and Ruiz-Mirazo, 1999] without noting that these researchers make such constraints central in the form of autonomy. Ironically for Barham, the autonomy global constraint is the foundation of natural value in that tradition — see [Hooker-a, this volume, section 4.1.1].

[19] On mechanisms see recently [Bechtel, 2007; Bechtel and Abrahamsen, 2005; Craver and Bechtel, 2006; Machamer *et al.*, 2000].

that have been conserved through processes of descent with modification.)

This is all appropriate, but once the condition-dependent character of laws is understood (see Hooker-a, b, this volume]), so that mechanisms are understood as possessing law-like operation, it will be appreciated that the above differences do not mark out a category opposed to nomic operation and explanation. Rather, they capture the entry of dynamical organization conditions into specification of dynamical process laws, in contrast with traditional laws that are taken to have minimal or no such conditions. Then the dynamical embedding approach provides a unifying basis for treating both kinds of laws as variants of one another, mechanisms having more complex, materially constrained dynamical processes that realise their functions (see section above): explanation appeals to the organisation of dynamical transform relations, reduction to the whole organised complex of dynamical processes and generalisation appeals to functionally relevant similarity of dynamical organisation.

In the general formulation above, Hooker [1981, Part III, pp. 503-505] identifies three kinds of descriptions of processes, the set of *mechanisms* (level L'_3) constructed from the basic physical process dynamics (level L_3), as the reducing mechanisms for the *functions* (Level L_1). L_3 is itself constructed piecewise from L_2, the basic *interaction/constraint dynamics* for the system. This is a very general approach to function and mechanism; there are as many functions and matching mechanisms as there are maps from any set of inputs to any set of outputs of the system. Let us then call these *basic* functions and mechanisms. While these provide a fully general basis for articulating function-to-mechanism reduction they are too undiscriminating to fully characterise productive scientific analysis of component mechanisms (cf. [Bechtel, Richardson, 1993; Bechtel, 2007]). A major challenge to mechanism theory posed by complex systems is to identify the explanatorily interesting mechanisms within the basic ones, that is, those mechanisms that realize the explanatorily important functions, whether of cells, cars, or cities.

Essentially this requires figuring out a principled basis for decomposing the total basic function of a system (the composition of all input/output maps for the system) into sub-functions. This is possible to the extent that the system is globally linear and locally modular, since then the total map can be written without loss of generality as a product of modular sub-maps. (The decomposition problem then recurs for each module.) But for non-modular systems, this is an important unsolved problem — and not just for philosophy, but for science as well.[20] Cellular biologists, control engineers and others concerned with complex systems would like to reduce the dimensionality of the system models and data they deal with and this again requires decomposing the total basic function in some principled way. But complex systems throw up organisational constraints that

[20]Rosen (1991) calls decomposable systems *synthetic*, and those that are not *analytic*. He argues that the former are mechanical, but there is a clear sense in which analytic systems are not mechanical. His argument is somewhat obscure and contains both gaps and unsupported speculations, but in spirit it has much in common with the approach taken here.

simple linear mechanisms cannot cope with, especially feedback/forward phase-lag loop relationships and global organisational constraints like the requirement for the functional integration of automotive functions so as to simultaneously satisfy performance standards (acceleration and braking, driving control, ride quality, ...) or the equivalent global functional integration (autonomy) that defines life.[21]

According to contemporary analyses (note 19), mechanisms are characterised by four key features, operationality — the job they do (e.g. pump blood), componentiality — their component parts whose causal articulation structures them, causal regularity — their reliable operation, and organisation — the interrelatedness in their articulation. But global biological organisation challenges this overly 'mechanical' conception: components are often not stable but variously created and dissolved by the processes themselves and the globally coherent organisation this requires for overall persistence in turn requires a conception of globally coherent mechanisms. Mechanisms are conceived as organized processes, but a serious incorporation of organisation within them remains an outstanding issue.[22]

4 SELF-ORGANISATION AND EMERGENCE

This issue also forms one facet of a serious outstanding problem posed by complex systems for a theory of reduction, viz. the occurrence of system levels. Following the dyamical approach to levels provided in [Hooker-a, this volume, section 3], levels proper are characterised by the presence of relatively macroscopic dynamical constraints. It follows that a level acts to 'top down' constrain component behaviour, in just the way that the iron lattice crystal does to create Fermi-band electrical current and lattice dissipative heat conduction. These constraints may be formed in many ways, e.g. both as a rigid crystalline structure emerges during the cooling of a liquid (iron bar) and as a liquid is heated to form Bénard convection cells. In the iron bar a macroscopic ionic lattice constrains the dynamics of microscopic component ions and electrons, constituting a distinctive dynamical level above them. In the Bénard convection case the macroscopic cell formation constrains the motions of all the molecular fluid constituents to same flow directions at all adjacent boundaries (cf. [Bishop, this volume, section 2.6]). Thus the formation of levels is directly linked to two other key complex systems features, self-organisation and emergence. Unravelling these issues will require introducing conditions under which reduction fails in a particular way. But it will then permit the account of reduction to be completed in an elegant, naturalistic manner, with the perhaps unexpected feature that irreducibility is a prerequisite condition for

[21]On this notion of autonomy see references [Hooker-a, this volume, note 50].

[22]For some discussion of the issues posed for mechanism see Bechtel 2007 and especially Bechtel herein. In addition, the identification of the functions required for internal functioning is complicated by the role played by environmental features in generating organised behavioural patterns: ant and termite colonies, e.g., show complex organised behavioural patterns but these are largely environmentally generated, their individuals obeying only simple, myopic local 'to-do' rules.

coherent complex systems reducibility, the two conditions being intimately intertwined.[23]

In all systems it is true that the interacting components together create a dynamics that would not otherwise be present. When the outcome is surprising or unexpected or too complex to readily understand, scientists are apt to talk about emergent patterns.[24] When the outcome is more complicated or subtle behaviour, and the dynamics is entirely internal to the system, it is said to be *self*-organised.

There are many reasons why leaving things like that is unsatisfactory, among them that (i) no significant feature is addressed, our subjective surprise or sense of complicatedness and the like keeps shifting with experience, and (ii) these sort of criterion are dynamically so weak as to trivialise these ideas. But when it comes to strengthening the requirement, there is currently huge diversity of opinion about how both self-organisation and emergence are to be understood. Two broad approaches to identifying something more penetrating can be distinguished, the one epistemic and the other dynamical.

We are following the dynamical approach, but first consider the epistemic alternative. The epistemic approach tightens up the subjectivity by adding a clause along the lines that emergence or self-organisation occurs when the resulting system dynamics could not have been predicted from the known interaction rules of the components. This approach is attractive because there are many complex behavioural patterns that arise from the simplest interaction rules, for example with social insects (hives of bees and termite mounds), city traffic and even simple population dynamics as reflected in the logistic equation. However, it still ties the definition of evidently physical properties to a cognitive test. And if prediction is restricted to logical deduction from dynamics then almost everything self-organises since the demand for analytic closed-form solutions fails for all bar the simplest sets of differential equations.[25] So we pass to the option of a dynamical criterion.

Two dynamical distinctions stand out, and fixing on them avoids a long detour through a tortuous literature. The distinguished differences are those of (i) bifur-

[23]This last was certainly unexpected in [Hooker, 1981], but not in [Collier, 1988; Collier and Muller 1998].

[24]An elegant way to systematise that talk has been provided by Ryan 2007, providing a powerful way to locate the phenomenon being referred to and identify the nature of claim being made about it, but it is largely agnostic about dynamical distinctions that might underlie it.

[25]As it stands, the text formulation is intolerably vague: Predicted by whom? Knowing what? Using what tools? And *prima facie* it makes an evidently ontological distinction (the existence of emergent behaviour) depend on a cognitive condition (human predictive capacity). If, in response, the criterion is instead formulated along the lines of 'cannot be derived from the set of interaction rules', or perhaps from just binary interaction rules, then these problems are lessened, but only to be replaced by the problem of what counts as an acceptable derivation. If derivation is restricted to logical deduction then almost everything self-organises since the demand for analytic closed-form solutions fails for almost all sets of differential equations. If derivation includes computational modelling of collective dynamics then almost all dynamics counts as derivable and nothing self-organises. Perhaps non-computable dynamics might be considered an exception, but since this occurs in quantum theory and other 'wave' dynamics as well as in critical-point bifurcations, it remains an insufficiently differentiated boundary. No satisfactory criterion of in-between scope is readily formulable.

cations, a structural instability leading to a shift in dynamical form, and (ii) the subset of bifurcations that lead to the establishment of a new system level. In the phase transition marking the formation of the iron bar, e.g., there is a bifurcation, but one in which a new level is formed. Specifically, a macroscopic pattern of inter-molecular relations is formed, the ionic crystal, which does thereafter have the power to constrain the movements of its molecular components through the formation of a new macro-scale force constituted in the ionic lattice bonds formed. Its formation alters not only individual component behaviour but also the specific dynamics under which they are now able to move: there are lattice vibrations and a Fermi conduction band in place of liquid molecular dynamics. That is, the phase change alters the force form of the dynamical equations that govern component behaviour. The new macro-scale force is able to retain the constraint relationship invariant under component fluctuations and exogenous perturbations, through lattice dissipation of these perturbing energies as sound and/or heat — that is, it is a top-down constraint.[26]

By contrast, the bifurcation in motion under gravity within a spherical bowl produced by raising a small mound in the bowl at some location introduces a shift in the spatial structure of the operative constraint force but no new macroscopic constraint. The same can be said of small alterations to avian flocking interactions and stigmergic rules for termite mound construction, both of which can nonetheless lead to shifts in collective behaviour. There is no dynamical constraint internal to either flocking birds or jamming motorists comparable to the iron crystal force that compels their members to wheel and turn just so, however complex their collective motion patterns. Finally, although from intersecting shallow waves on a gently undulating beach there emerges the most beautiful and intricate patterns, there is no comparable constraint of any sort formed by their interaction; shift the underlying sand structure and the dynamics can shift to entirely other patterns.

Between this last and the iron crystalline cases lie a variety of intermediate strengths of top-down constraint. For example, many social insect societies, like ants, are constrained by chemical reflexes and so exhibit some stronger constraint to their collective behaviours than bird flocking and jamming motorists, but it is still not comparable to the ferric crystal force. The crystalline cases are peculiar too in that their constituents are constrained in their total behaviour, whereas in other systems the top-down constraints can be expected to be more partial. For instance, the division of labour in bee hives seems much more collectively

[26]The iron bar is a new macro-scale level with respect to its molecular constituents with its own characteristic dynamical interaction form. All other talk of levels either concerns measurement (liquid level), gravitation (level surface), or is metaphorical (semantic, social, abstraction, theory ... levels) and can thus be paraphrased away — or is confused. (The use of 'level' in Hooker's L_i levels, 2 paragraphs above, pre-dates and violates this usage decision, referring only to descriptive 'levels'.) Note that the presence of a top-down constraint does not fully determine the specific dynamical form of the system; both the virtual and real electrical governor arrangements (n.9 and text) exhibit the same phase-stabilising top-down constraint. Distinguishing between them is the electrical engineering 'system identification' problem — now with equivalents across all sciences, e.g. it is the problem of identifying correct cellular biosynthetic pathway models from fast-throughput experimental data that is a current central issue in systems and synthetic biology.

constrained than does their foraging patterns, which are more like bird flocking, even given their dancing. Cells in multi-cellular organisms seem as or more tightly bound, at least in respect of reproduction, specialisation and death (the 3 functions any viable community must control to some degree), and possibly in respect to other aspects of inter-cellular signalling.

It is natural to choose bifurcation as the dynamical criterion of emergence, for then a new behavioural pattern develops, and one whose occurrence is dynamically grounded in a shift in dynamical form. Bifurcations can take many different dynamical forms, but emergence is not concerned with the specifics of the process, only with something genuinely new coming into being from essentially the same resources. Bifurcation satisfies these requirements while providing the widest dynamically well-grounded criterion. In particular, this criterion applies independently of the computational character of the formation process and end state. While failure of computational penetrability represents a common alternative approach to characterising emergence, its status is problematic at this time.[27] It may be regarded as a positive feature of the birfurcation criterion that it does not depend on resolving the issues.

Thus defined, emergence corresponds to a part of the supra-component pattern formation spectrum running from level formation to non-birfurcational pattern shift. Calling the entire spectrum 'emergence' would reduce the notion to simply behavioural change, eviscerating its content. More informative to call the entire spectrum 'pattern formation'. Confining 'emergence' to just the level-forming bifurcations would be to miss all those cases where the change is objectively grounded in a holistic change in dynamical form. This would leave these cases begging for a label that captures this sense of holistic transition. Thus neither of these is conducive to insight. Because the whole spectrum seems to have no other natural dynamical divisions, the intermediate usage — emergence = bifurcation — is ultimately most helpful.[28]

[27]The predominant assumption is focused on unpredictability and thus linking irreducibility to some form of formal impenetrability, in particular the absence of a computational simulation of the dynamics, e.g. [Collier, 2008]). Bifurcations involving a critical point satisfy computational impenetrability, it seems, but they are not necessary for either a bifurcation or constraint-formation criterion to be met. Within present analytical mechanics the shift in dynamical form across constraint formation bifurcation is sufficient of itself to disrupt predictability. However its essence lies in the ontological reality of the internal constraint formation, not in unpredictability. This is fortunate since to pursue the impenetrability approach requires uncovering the details of the reach of compression into the domain of complex systems, and to distinguish in-principle limits from current pragmatic limits. But in fact there is relatively little known about this very difficult area. Beyond affirming vaguely that our present analytic techniques do not carry us very far, and that specific fragments of progress are occurring, see e.g. [Israel, Goldenfeld, 2006] among many, there seems little to be done but await mathematical developments.

[28]The criteria suggested in [Collier, 1988; Collier and Hooker, 1999; Hooker, 2004] is that of the more stringent requirement of self-organisation — see below. While this has the attraction of simplicity, since emergence = self-organisation, this now seems to me not to do justice to the differing orientations of the two concepts. Conveniently, there are two well-grounded dynamical distinctions available, quite well suited to the purpose, and it seems good sense to make use of them.

There is no physical mystery about emergence because the dynamics itself gives rise to the emergent dynamical form. The only components involved, e.g. forming any macro structures, continue to be the original dynamical entities from whose interactions the constraint emerged. From the component level perspective, this is simply local interaction placing mutual constraints on component behaviour in a way that eliminates certain collective possibilities (e.g. flowing as a liquid) while creating others (e.g. rigid collective interaction, refracting sound waves). But the formation of a new, holistic dynamical form makes it clear that the new collective possibilities are as real as were the old ones, just different from them (cf. [Bishop, this volume, section 2.6] on non-separable Hamiltonians).

There is some inclination to require that emergence means sufficient separation of the emerged level from its substrate that it has an intuitively clear separate existence. In consequence, the continuing presence and efficacy of the dynamical components is taken to mean that emergence has failed, e.g. that to some sufficiently god-like mind the iron bar is really just its components after all. But such imaginative vanities are irrelevant as well as ill-founded. What matters is that the iron crystal constraint is dynamically real, as determined by its energetic consequences, even though generated endogenously. The centrality of the dynamical criterion of new dynamical form is underlined by the fact that were the components to be fields and not particles, there would be no unique components available, yet the same dynamical criterion would demarcate emergence.[29] Whether or not the dynamical interactions that give rise to the emergent constraint change the components in the process, it is the fact that it is a *dynamical* bond that makes it possible to assert both the distinctive dynamical (if you wish: causal) character of the emergent entity and that it is comprised of (dynamical) components with the power to give rise to and sustain it. In this way emergence is naturalised for science.[30]

It is equally natural to choose level-forming bifurcation as the criterion of self-organisation (in the respects and to the degree it applies), for just this characterises the coming into being of a new self in the form of a new dynamical existent. The iron top-down constraint formation constitutes the coming into being of a new, individuated capacity to do work, expressed both endogenously in dissipation of perturbations and exogenously in rigid-body action. It is the arrival of a new dynamical individual characterised by a new dynamical form. The character of the new individual is constituted by its capacity to do new work, expressed in its dynamical form.

---

[29] If there are unique, unchanging, spatio-temporally local, fundamental dynamical entities (for example, chemical ions as biochemical a-toms) then there is no emergence of new fundamental kinds, only existential dynamical emergents having these entities as ultimate components in various dynamical compounds. But top-down constraint formation of itself does not require this particular ontology. Fundamental non-linear fields would yield the same emergent result and there are no such local components, while mutant spatio-temporally local fundamental components would issue in fundamental kind emergence.

[30] This also ultimately invalidates Kim's treatment of diverse realisations of relatively macroscopic states and the conclusions he draws therefrom [Hooker, 2004, 5B; Wimsatt, 1994].

It will be observed that the same choice of terminology is at issue for 'self-organisation'. It might be allowed that self-organisation should be more broadly defined to capture simply the central idea that the resulting pattern is brought about through the interactions of the system components. The colloquial term 'organise', as in 'get organised', encourages this wide connotation. But again, since the entire specrum would then count as cases of self-organisation, this would reduce the notion to simply behavioural change, eviscerating its content. Alternatively, it might be argued that the term should be restricted to emergence in the sense of bifurcation, which remains weaker than that of top-down constraint formation. But this would include within self-organisation those systems whose bifurcation was produced by external manipulation, thus lacking an active 'self' in the process, and those systems that reduce their behavioural orderedness and/or complexity, thus lacking increased 'organisation' (even in the colloquial sense). While these outcomes are irrelevant to whether anything emerges, they still confuse any notion of self-organisation. Because the whole spectrum seems to have no other natural dynamical divisions, the narrow usage — self-organisation = level-forming bifurcation = level-forming emergence — is ultimately most helpful.[31]

This definition effectively replaces the idea of exhibiting more ordered and/or more complex behaviour as a characterising outcome of self-organisation by formation of a new level. For some this will be regarded as unsatisfactory. But at this time there does not seem to be any further dynamically well-grounded criterion that would do any better justice to this intuition. The obvious option to consider is adding some requirement concerning organisation proper (see [Hooker-a, this volume, section 3]). In this respect, it is worth noting that self-*organisation* need have little to do with organisation proper. This is as it should be. Organisation is a relational condition of systems where components play distinct roles but the roles are so interrelated as to produce a coherent global outcome. A simple illustration is found in the way the parts of a car engine are interrelated so as to deliver torque from fuel ignition; a profound example lies in intra-cellular organisation to produce a functioning organism. However, the end products of self-organisation need not involve the emergence of any organisation, as the case of crystallisation shows. Crystal formation is, rather, an instance of the formation of orderedness, rather than of organization. The unfortunately wide colloquial connotation of 'organise' conflates order and organization.[32]

With terminology settled we turn finally to the interrelationships between self-organised emergence and reduction in complex systems. A stable level is stabilised by damping or filtering component fluctuations, e.g. in the way that the iron crystal lattice dissipates perturbing fluctuations in lattice vibrations and the convecting cell dissipates them through molecular collisions. This insures that relatively

[31]The narrower usage is followed in [Collier, 1988; Collier and Hooker, 1999; Hooker, 2004].

[32]The conflation has a venerable history, e.g. while [Ashby, 1959] provides a constraint approach to what he calls organisation, what he defines is orderedness. The idea of global functional constraints, though implicitly available, had not come into focus and his formal approach excluded the dynamical bifurcations that underlie top-down constraint formation. For more on the order/organisation distinction see [Collier and Hooker, 1999], but cf. e.g. [Denbigh, 1975].

macroscopic properties have the stability we find them to have. This in turn making possible all the myriad processes that scaffold off them, e.g. the rigidity of iron components that make them fit to use in buildings and machines. And it is also on that general basis, and only on that basis, that we can track dynamical influences 'up' and 'down' through the component/supra-component levels, e.g. track the consequences of an ionisation event in a Geiger counter through to the output click [Hooker, 2004, Part 5B]. By providing higher level structure for well characterised lower level processes, macroscopic constraints underpin the reduction of the functions served to their dynamical process mechanisms. (And of course the constraints themselves and attendant structures reduce to dynamical compounds of the components whose interactions constitute them.). That is, it is the very stability of the emergent constraints that both make possible well-defined functions like the Geiger counter's detection and also the well-defined dynamical processes underlying them (e.g. ionisation and amplification in the Geiger counter) and in doing so provide the basis for functional reduction [Hooker, 2004, section 5B]. On the other hand, these macroscopic constraints herald the presence of a new dynamical existent created by an over-arching dynamical bond, so the new dynamical existent precisely cannot be reduced to components alone. It is an irreducible new being.

Thus, contrary to the standard view of reduction and emergence where they are opposed, this discussion shows that emergence underpins functional reduction, and reduction, both compositional and functional, in turn allows specification of the processes and conditions (initial and boundary) that underpin emergence. The two are thus intricately interwoven and mutually supportive.

5 CONDITIONS FOR IRREDUCIBLE EMERGENCE

Significantly, self-organisation is a process where dynamical form is no longer invariant across dynamical states but is rather a (mathematical) function of them. Despite advances in non-linear mathematics and computing, this remains an ill-understood process.[33] These circumstances set the scene for examining, first, what a minimalist logical account offered by philosophers might do to clarify emergence (section 5.1) and, second, what can be said about its dynamical conditions in relation to the foundations of classical dynamics (section 5.2).

5.1 Logical conditions for irreducible emergence: supervenience

Within contemporary philosophy the notion of supervenience is the dominant, and minimalist, approach to characterising emergence. A states are supervenient on states B if and only if every change in A requires a change in B (though not necessarily vice versa). For the intuitions driving this formulation, consider a standard macro-micro relationship: The iron bar is supervenient on its molecules, no

[33]For the spread of non-linear research see e.g. `http://arXiv.org/archive/nlin/`; Ma and Wang [2005] provides a mathematical introduction to bifurcations more generally.

properties of the macroscopic bar (A states) can change without the change being dynamically grounded in appropriate molecular changes (B states), though there will be many molecular changes that produce no macroscopic change. Supervenience specifies a many-one map (a mathematical function) from B to A (many different arrangements of iron molecules can produce the same bar). Thus A states correspond to sets of A-equivalent B states and A state transitions (say from A to A') correspond to sets of B transitions (all those allowed from any member of the $B(A)$ set to any member of the $B(A')$ set).

The supervenience relation is minimalist because nothing further is said about how the $B - A$ relationship works. There is no dynamical account of the $B - A$ relation specified, as there is for the iron bar, e.g., where we know which molecular changes are required to dynamically explain which macroscopic changes. Indeed, and quite unlike the iron bar, supervenience is strictly compatible with the A states being causally unconnected to the B states, instead just 'running in parallel'. (The A states might, e.g., be mental states in some spiritual 'ether' and the B states brain states.) Supervenience is also compatible with the $B-A$ mapping itself being a function of other variables, e.g. of time, though most commentators assume not. The desire to accommodate the parallelism option, together with philosophers favouring use of purely formal relations, may explain the minimalism. In any event, it means that we cannot hope to derive much enlightenment from applying the supervenience concept. The only interesting consequence of applying it is that fixing the B states (and the values of any other relevant variables) fixes the A phenomena. This expresses the degree of 'bottom-up' determinism built into supervenience.

It also makes the A states look parasitic on the B states, since the A level can't change without the B states changing. This can be taken to imply B causal leadership, and conclude that the A level has no causal powers other than those possessed by its B substrate. (The aggregate is an especially plausible example here.) An obvious next move is the application of Ockham's razor to remove A states as separate existents, through reduction to the B states. This would have supervenience support reduction and exclude emergence.[34] But this need not be so. As just noted, supervenience strictly says nothing about the actual physical relation of B states to A states, so it is quite compatible with supervenience that that relationship is, e.g., one of dynamical self-organisation. In that case there are dynamically based A causal powers that can constrain B dynamics and thus A states are irreducible to B states. There is no contradiction here, simply that supervenience is in fact compatible with any of the relationships, illustrated above, ranging from aggregation to self-organisation. This reveals the discriminatory poverty of the supervenience relation.

Dynamical analysis provides a much richer language in which to discuss the

[34]Kim takes B as the physical states and assumes determinism, to obtain this version of supervenience [Kim, 1978]: If all of the (determinate) physical facts are determined, then all (determinate) facts are determined. Kim supposes reduction to follow, but things are not so simple (see following text).

possibilities, distinguishing all the relationships across the range from aggregation to self-organisation. In particular, for self-organisation the dynamics itself shows how the (relatively) macro level constraint is determined by the states of its micro constituents and so is supervenient on them, yet can nonetheless also constitute a constraint on them. Here dynamics gives the constraint a subtle status that eludes conventional formal analysis, and supervenience. Understanding is only obtained when supervenience is replaced with (or perhaps enriched with) dynamical relations.

Thus dynamical determination, = there being only one dynamical possibility for the collective dynamical state/property, cannot be equated with logical determination, = the collective dynamical state/property is logically derivable from, can be expressed as a logical sum of, its constituent states/properties. The latter includes only the weaker, aggregation and simple pattern formation relations, while the former also includes the stronger bifurcational, especially self-organisation, relations where deduction fails because, as noted earlier, there is no single mathematical framework within which dynamical form shift is a well defined transformation.[35] Thus reduction fails in these cases but holds for pattern formation and aggregation cases.

5.2 Dynamical conditions for emergence

The ambitious aim here would be to fully characterise the situation of complex systems dynamics in relation to dynamical theory, but this is a task far too deep and complex, and unfinished and rapidly evolving, to contemplate. So the scope of the discussion is restricted to just classical dynamics (dynamics of classical mechanics) and the aim is restricted to characterizing the relation of some key general aspects of classical dynamics to some key general aspects of complex systems dynamics.

The classical Newtonian laws of motion constitute a very general dynamical framework for particle mechanics that in itself places no restrictions on the kinds of material systems involved, e.g. whether involving charged particles or dissipative forces like friction. It is always possible to attempt to mathematically analyse such systems from first principles in order to discover their behaviour, however success in this is not guaranteed and is in fact typically very hard or impossible beyond the simpler classes of dynamics, e.g. for systems undergoing phase changes (e.g. gas to liquid to solid).[36] However, over the past 250 years a generalised analytical framework has been constructed for classical dynamical analysis — the Lagrangian/Hamiltonian formalism — that directly or approximately covers a wide range of cases and serves as the core of analytical classical dynamics.[37] Moreover the underpinning principles have the further merit of providing the

[35]Indeed, these trajectories may be computationally strongly inaccessible, for example through all critical point phase transitions, and so certainly cannot be deduced from micro information.

[36]Cf. Humphries on the mathematical tractability of computational templates, section 3.5 of [Humphries, 2004] and the discussion in [Strogatz, 1994, p.11] and text.

[37]For a nice introduction see [Butterfield, 2004a; 2004b]. Classic texts here are [Goldstein 1950; Arnold, 1978]. Butterfield [2004a, note 6] offers a brief guide to others. [Bloch, 2003]

means to also treat continua and force fields, especially the electromagnetic field, and of furnishing principles for subsequently generalising mechanics to relativistic and quantum formalisms. Thus this analytical framework has not only come to constitute "classical mechanics", its core exemplifying our conception of purely mechanical systems, but it has come to constitute the foundational core of all general dynamics, the focus for attempts to understand relativistic and quantum theories and their union.

In this setting it may be thought that insofar as complex systems lie outside this core, as most of them do, it is simply because solving their dynamics is still too hard as yet and we are driven to piecemeal tricks, model by model. We *are* at present largely driven to piecemeal tricks, and may or may not be stuck there, but there is more at stake than progress in practical mathematical methods. Complex systems characteristically centre on dynamical conditions at variance with the assumptions underlying the classical core and thus face us with a fundamental dilemma, and a small scandal.

To understand what is involved, begin by noting that a classical mechanical system consists of material components (the system elements), interactive forces among these components which provide the intrinsic forces, and external constraints or boundary conditions placed on the system (e.g. for a gas or fluid, the walls of its container) which provide the extrinsic forces acting on it. In the absence of constraints there are only intrinsic forces and a Newtonian dynamics can be formulated with mutually independent dynamical variables expressing the intrinsic degrees of (dynamical) freedom of the dynamics. External constraints are typically represented as restricting the movements of a system (e.g. gas container walls) and are assumed to apply sufficient forces to achieve this. While the non-linearity of the intrinsic forces is the primary means through which dynamical complexity is produced, it is the character of the external constraint forces that is the key factor in the formation of an analytically tractable treatment of the dynamics. External constraints typically enter unknown forces into the dynamics, so that a determinate Newtonian dynamics cannot be specified, and they result in interdependencies among the intrinsic dynamical variables that have to be accommodated, so that an unambiguous representation of the dynamical possibilities cannot be formulated.

The standard dynamical formalisms for mechanics are those of the Lagrangian and Hamiltonian forms; here just the more general Lagrange formalism will be discussed. To construct a Lagrangian model for dynamics it is necessary to restrict consideration to those systems where the external constraint forces act orthogonally to all allowed system motions[38], so that the system does no work against external constraints (constraint force orthogonality). This defines a 'constraint (hyper) surface' in the system configuration space to which the constraint forces

offers an introduction to the treatment of D'Alembertian but non-holonomic constraints, see also [Flannery, 2005].

[38]These are the 'virtual' displacements of a system, as opposed to actual displacements over some small time interval occurring under the influence of the intrinsic forces as well.

are everywhere perpendicular (orthogonal). The dynamics on this surface is thus effectively separated from the external constraints, each is unaltered by the other throughout the system motions — this is expressed in D'Alembert's principle. The dynamics is then open to a purely intrinsic characterisation. It is, however, non-trivial to make good on this promise since the issues of interdependencies among variables and unknown external constraint forces remain unresolved.

If in addition the external constraints are holonomic — literally: express a whole or single law[39] — then the system dynamics may be re-formulated on their D'Alembertian constraint surface in terms of new generalised coordinate variables that are mutually independent. The dynamics now has the form of a free (unconstrained) system. The effect of the constraints, implicit in the geometry of the constraint surface, is now also implicit in the construction of the new variables for intrinsic motion on it. Lagrange equations of motion can then be formulated for the system. This resolves the variable interdependency problem introduced by constraints. We think of these systems as simply following a least-action path in a pre-specified purely geometric framework and hence as distinctively 'mechanical' in nature.[40] Further, the method of Lagrange multipliers permits solving the system dynamics on the constraint surface (that is, specifying the action geodesics) without knowing the external constraint forces. Rather, once the dynamics is known, the external constraint forces can be reconstructed as the forces they need to be to maintain the external constraints during the system motion. This resolves the problem of their being initially unknown.

More could easily be added. Theoretically, e.g., the Lagrangian form represents a simpler set of equations to be solved than is Newton's and the Hamiltonian formulation extends this trend. Practically, e.g., the familiar form of the Lagrangian as kinetic minus potential energy can be derived if the forces can be expressed as the gradient of a single function. Moreover, the Lagrange multiplier method extends to some classes of non-holonomic constraints as well (specifically semiholonomic and exact linear constraints [Flannery, 2005]) and there may be progress with others (e.g. cf. [Fernandez, Bloch, 2008; Krupková, 2009]). However, the

[39]Holonomic constraints may be written as some function of the space-time geometry in which the system moves. Specifically, they satisfy an equation of the form $f(r_1, r_2, \dots\ r_n, t) = 0$, where the r_i are system coordinates and t is time. This expresses the effect of the constraint forces while not specifying the forces themselves. (The forces are often known only after the main problem is solved.) While smooth (frictionless) sliding under gravity on a sloping plane is a case of holonomic constraint, a spherical bead rolling smoothly on the outside of a cylinder is not because the constraint alters its basic character when the bead falls off. Essentially, for the constraints to be holonomic means that they may be expressed purely geometrically, so that they are independent of the behaviour of the system. Independence fails in the case of the bead on the cylinder, there is a change of constraints at a space-time location determined by the bead's motion. (Note that the reverse relation does not hold, e.g. though independent of system behaviour, containment walls do not form holonomic constraints.)

[40]Especially if neither the external constraints nor the potential energy are time-dependent, the usual textbook case. But note that the intrinsic force potential is still produced by the system components themselves and any internal constraints will form in virtue of such forces; we need to be wary of claims that there is any sharp gulf separating mechanical systems from the more complex non-mechanical ones to which we shall shortly point.

general result above states the nub of the relevant matter here. It suffices only to add that D'Alembert's principle introduces the first of the variational formulations of dynamics whose extension stands at the core of generalizing dynamics to encompass the relativistic and quantum domains. These variational principles are considered to lie at the core of analytical dynamics.[41]

Nonetheless, for many complex systems the external constraints that apply depend on what the dynamical state is, so that constraint holonomicity fails, blocking the path to further re-construction of the dynamics. This would, it seems, be true for all globally constrained, functionally resilient (often called robust) systems that can adapt their process organisation to compensate for damage or other altered circumstances, as can living cells. Moreover, many systems where constraints are a function of state also do work on the constraints, physically altering them over time. Examples include (i) a river altering its own banks, an accumulative process where the current constraints (banks) are a function of the history of past flows (currents), (ii) intra-cellular biochemical reaction processes where molecular structures constraining some processes are the products of other processes and vice versa; and (iii) any self-organisation where the constraint formed becomes an external constraint for subsequent processes (Bénard cell and iron bar formation, etc.). In all these systems constraint orthogonality fails. With this failure the most basic precondition for achieving the core analytic construction fails. There is then no general analytical mathematical formalism available for dynamical behaviour. Moreover, most of these systems have proven recalcitrantly impenetrable to analysis and essentially each system has to be treated individually on its merits. There are stronger claims, e.g. that all chemical reactions are irreversible thermodynamic processes defying analytical dynamical characterisation [Prigogine, 2003], but those above suffice. It is among these that we find all systems exhibiting emergence and organised global constraints and many other of the characteristic features of complexity (see [Hooker-a, this volume, section 3]).

There is then a hiatus between those systems whose dynamical foundations we think we understand (Lagrangian systems) and those systems that manifest the features characteristic of complexity. D'Alembert's principle fails for the latter systems, undermining the applicability of the very variational apparatus that we take to underlie all fundamental dynamics. In this way, complex systems challenge the reach of our deepest analytical understanding of dynamics and thus present a fundamental dilemma about how to approach dynamics: retain the present approach and exclude complex systems or search for some new, more generous foundations for dynamics. There is also a whiff of scandal here as well, namely, the unfortunately increasingly common scandal of dynamics textbooks simply ignoring these deep problems. This is especially scandalous at a time when, as this

[41] In this process analogous conditions to those necessary for the representations (no virtual work, holonomicity) are assumed universally applicable, plus, for the Hamiltonian formulation, also the assumption of the mutual independence of coordinates and momenta [Goldstein, 1950, p.227]. This is a pivotal point for reflection on how Newtonian systems that violate these requirements, as emergent systems will do (below), could be represented in relativistic and quantum terms.

volume demonstrates, complex systems are having such a huge impact on science, including on the mathematical techniques for analysing dynamical systems. More important, however, is scandalous question-begging in favour of the existing approach by the commonest textbook response, which implies that there is only a pragmatic issue of mathematical resources involved.

The ground offered for the latter is that ultimately all systems, constraints as well as components, can be represented at the fundamental component level (however components themselves are represented). Thus all external constraints are then represented as forces deriving from yet further fundamental components. The gas container becomes a metallic lattice plus free electrons, and so on. These external components may then be added to those of the systems they constrain to form a dynamics that is constraint-free (constraint forces = 0) and hence Lagrangian methods suffice.[42] If we cannot solve these systems then it is simply because there are too many components involved, a pragmatic rather than a principled difficulty.

It is certainly the case that constraints can be shifted between external and internal status. Consider the iron bar again; its ion lattice formed as an internal constraint but, once formed it may be treated as an external constraint for lattice processes such as sound and heat propagation and Fermi band electrical conduction. However should the energy in these processes become sufficient to perturb the lattice ions sufficiently to do work on the lattice, then the lattice has to be again brought within the dynamics as an internal constraint. The argument is that ultimately this is true of all constraints.

About this argument, the following points are germane to estimating its persuasiveness. (1) This is at best an in-principle argument, a necessary condition for its coherence being a proof that the actions of the fundamental forces always permit definition of a suitable D'Alembertian surface. (I assume that the method of Lagrange multipliers then works.[43]) Given that many of these systems do real work

[42]See e.g. [Goldstein, 1950, p.14] for this view.

[43]Flannery [2005] emphasises the compatibility of the method of Lagrange multipliers with the D'Alembertian surface (i.e. its ability to identify a field of extremal flows on the D'Alembertian surface) as the key condition for the coherent formulation of a Lagrangian dynamics. Others (e.g. Butterfield) emphasise obtaining a D'Alembertian surface at all as the key condition, with the multiplier method treated only as a useful add-on. Lagrange multiplier methods are not often discussed in dynamical contexts. One immediate issue is whether these two conditions are in fact equivalent — does it follow that if a D'Alembertian surface is definable then the method of Lagrange multipliers works for it and, conversely, if the method of Lagrange multipliers is well defined then a corresponding D'Alembertian surface is defined? This has some interest in the light of the extensions of Lagrangian formulation to various classes of non-holonomic constraints, e.g. [Fernandez, Bloch, 2008; Krupková, 2009]. Another issue is whether there is a general characterisation of the classes of dynamical systems for which the D'Alembertian variational principle yields a minimum and for which it yields a maximum (the method itself only requires an extremal value), and a general explanation of these differences. Finally, there is the less defined issues of (a) how to insightfully characterise the respective contributions of the intrinsic dynamical non-linearities and the constraint structure to the generation of system dynamics that are analytically untreatable and (b) whether there is a general characterisation of the class of non-holonomicly constrained systems that are treatable with an extension of Lagrangian methods

on their constraints, it is not obvious how this proof would succeed. Until this proof is available, it remains a fact that none of the constrained non-D'Alembertian systems have a coherent Lagrangian mechanics specified. (2) Nor can any system enlargement detract from the dynamical reality of constraint formation. (It certainly cannot be argued that internal constraint formation is only a convenient illusion, since its reality is attested by the energetic difference it makes, expressed as a difference in the system work function.) In this light it is difficult to see how systems showing self-organised emergence could be reduced to presenting merely pragmatic barriers to knowledge of solutions. To take this latter view requires presuming that in the fundamental representation all top-down constraint formation becomes representable as a process within Lagrangian dynamics. Since precisely in such processes the system changes dynamical form, hence would change Lagrangian form, it is unclear how the Lagrangian apparatus could accommodate that requirement.[44] Thus the response begs the question against these arguments, without providing a demonstration that there is a real solution available.[45] The basic dilemma persists.

6 CONCLUSION

The logic machine vision is not dead. Condition-dependent laws still compress and dynamical equation sets still provide implicit compressed representations even when most of that information is not explicitly available without decompression. And, paradoxically, there is still the determined march of fundamental analytical dynamics expanding its compression reach toward a Theory of Everything - even while the more rapidly expanding domain of complex systems dynamics confronts its assumptions and its monolithicity. Nor does science fall apart into a dis-unified aggregate of particular cases since, with fundamental dynamics as a backbone,

(other than that they satisfy your requirement of the compatibility of the method of Lagrange multipliers with the D'Alembertian surface).

[44]It seems true that if a universal D'Alembertian, holonomic model is to apply then all the complex phenomena of complex systems, such as emergence, must be explicable in terms of dynamically valid coarse graining on the underlying fundamental dynamics. However, coarse graining in itself may not help much in developing a coherent mathematical account simply because its point, reflected in renormalization, is to avoid appeal to the dynamical details. Further, in those cases of singular asymptotics of bifurcation, it is hard to see how coarse graining can provide intelligible legitimacy to an underlying analytic model since that model would seem to require ultimate continuity of change, in contrast with the discontinuous change involved in singular asymptotics, not to mention also requiring motion solely within some single dynamical form.

[45]It also requires presuming that in the fundamental representation all global organisational constraints, such as those of autonomy, also become analytically representable. But again, the Lagrangian formalism has no apparatus for representing such constraints. To insist that they reduce to just pair-wise interactions among fundamental components is to again beg the question of their dynamical reality (their contribution to dynamical form). While it is not clear just how such constraints are to be represented, it seems clear that they must be more palpable than mere pair-wise correlations — think in this respect, of the imposed global coordination of Bénard convection cell circulation..

complex matching up of models across theoretical and empirical domains then articulates its model-structured skeleton. Here is included the delicately entwined dance of emergence and reduction providing constraints on compression that also permit its expansion. However, while the vision is not dead, it is currently substantially more complexly structured through model similarities and differences than that initially envisaged and we are left with deep questions about compression unresolved.

ACKNOWLEDGEMENTS

I am indebted to a discussion with Paul Humphries and to several with Gordon McLelland that clarified several key points.

BIBLIOGRAPHY

[Amaral and Ottino, 2004] L. A. N. Amaral and J. M. Ottino. Complex systems and networks: challenges and opportunities for chemical and biological engineers. *Chemical Engineering Science*, 59, 1653-1666, 2004.

[Arnold, 1978] V. I. Arnold. Mathematical methods of classical mechanics. Springer, Berlin, 1978.

[Ashby, 1952] W. R. Ashby. *Design for a Brain.* Chapman & Hall, London, 1952.

[Barham, 2004] J. Barham. The emergence of biological value. In W. A. Dembski, M. Ruse (eds.), *Debating Design*, Cambridge University Press, Cambridge, 210-226, 2004.

[Batterman, 2002] R.W. Batterman. *The devil in the details: asymptotic reasoning in explanation, reduction and emergence.* MIT Press, Cambridge, Mass., 2002.

[Bechtel, 2007] W. Bechtel. Biological mechanisms: organised to maintain autonomy. In

[Boogerd *et al.*, 2007,] , 269-302, 2007.

[Bechtel and Abrahamsen, 2005] W. Bechtel and A. Abrahamsen. Explanation: a mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 421-441, 2005.

[Bechtel and Richardson, 1993] W. Bechtel and R. Richardson. *Discovering complexity.* Princeton University Press, Princeton, 1993.

[Bloch, 2003,] A.M. Bloch. *Nonholonomic mechanics and control.* Springer, New York, 2003. See `http://www.cds.caltech.edu/mechanics_and_control/about_this_book/` for corrections and additions, incl. internet supplement.

[Boogerd *et al.*, 2005] F. C. Boogerd, F. J.Bruggeman, R. C.Richardson, A. Stephan, H. V. Westerhoff. Emergence and its place in nature: A case study of biochemical networks.

[Boogerd *et al.*, 2007] F. Boogerd, F. Bruggeman, J-H. Hofmeyr, H. V. Westerhoff (eds.) *Systems Biology: Philosophical Foundations*, Amsterdam, Elsevier, 2007.

[Bruggeman *et al.*, 2002] F. J. Bruggeman, H. V. Westerhoff, and F. C. Boogerd. BioComplexity: A pluralist research strategy is necessary for a mechanistic explanation of the 'live' state. *Philosophical Psychology*, 15(4), 411 - 440, 2002.

[Butterfield, 2004a] J. Butterfield. Between Laws and Models: Some Philosophical Morals of Lagrangian Mechanics, 2004a. Available at `http://philsci-archive.pitt.edu/archive/00001937/`. Accessed April 2009.

[Butterfield, 2004b] J. Butterfield. On Hamilton-Jacobi Theory as a Classical Root of Quantum Theory. In Elitzur, A.C., Dolev, S., and Kolenda, N. (editors), Quo Vadis Quantum Mechanics? Possible Developments in Quantum Theory in the 21st Century. New York, Springer, 2004b. Available at `http://philsci-archive.pitt.edu/archive/00001193/`. Accessed April 2009.

[Causey, 1977] R. L. Causey. *Unity of Science*, Reidel, Dordrecht, 1977.

[Chandrasekhar, 1961] S. Chandrasekhar. *Hydrodynamic and Hydromagnetic Stability.* Clarendon Press, Oxford, 1961.

[Churchland, 1979] P. Churchland. *Scientific realism and the plasticity of mind.* Cambridge University Press, London, 1979.

[Collier, 1988] J. D. Collier. Supervenience and reduction in biological hierarchies. In M. Matthen, B. Linsky (eds.) *Philosophy and Biology,* Canadian Journal of Philosophy, Supplementary Volume 14, Ottawa, 209-234, 1988.

[Collier, 2008] . J. Collier. A dynamical account of emergence. *Cybernetics and Human Knowing*, 15(3-4) 75-86, 2008.

[Collier and Hooker, 1999] J. D. Collier and C. A. Hooker. Complex Organised Dynamical Systems. *Open Systems and Information Dynamics*, 6, 241-302, 1999.

[Collier and Muller, 1998] J. Collier and S. Muller. The dynamical basis of emergence in natural hierarchies. In G. Farre, T. Oksala (eds.) *Emergence, Complexity, Hierarchy and Organization, Selected and Edited Papers from the ECHO III Conference, Acta Polytechnica Scandinavica, MA91.* Finish Academy of Technology, Helsinki, 1998.

[Denbigh, 1975] K. G. Denbigh. *An inventive universe.* Hutchinson, London, 1975

[Dewan, 1976] E. M. Dewan. Consciousness as an emergent causal agent in the context of control system theory. In G. Globus, G. Maxwell, I. Savodnik, (eds.) *Consciousness and the Brain*, Plenum, New York, 1976.

[Dupre, 1993] J. Dupré. *The disorder of things: metaphysical foundations of the disunity of science.* Harvard University Press, Cambridge Mass., 1993.

[Farrell and Hooker, 2007a] R. Farrell and C. A. Hooker. Applying self-directed anticipative learning to science: agency and the interactive exploration of possibility space in ape language research. *Perspectives on Science*, 15(1), 86-123, 2007.

[Farrell and Hooker, 2007b] R. Farrell and C. A. Hooker. Applying self-directed anticipative learning to science: learning how to learn across revolutions. *Perspectives on Science*, 15(2), 220-253, 2007.

[Farrell and Hooker, 2009] R. Farrell and C. A. Hooker. Applying self-directed anticipative learning to science: identifying error and severe testing through interactive exploration of possibility. *Foundations of Science*, 14(4), 249-271, 2009.

[Fernandez and Bloch, 2008] O. E. Fernandez and A. M. Bloch. Equivalence of the dynamics of nonholonomic and variational nonholonomic systems for certain initial data. *Journal of Physics A: Mathematical and Theoretical*, 41, 344005, 2008.

[Flannery, 2005] M. R. Flannery. The enigma of nonholonomic constraints. *American Journal of Physics*, 73(3), 265-272, 2005. Available at `www.physics.gatech.edu/people/faculty/flannery/publications/AJP73_March2005_265-272.pdf`

[Goldstein, 1950] H. Goldstein. *Classical mechanics.* Addison-Wesley, Reading, Mass., 1950.

[Griffiths and Stotz, 2007] P. E. Griffiths, K. Stotz. Gene. In M. Ruse, D. Hull (eds.) *Cambridge Companion to Philosophy of Biology*, Cambridge University Press, Cambridge, 85-102, 2007.

[Hooker, 1973] C. A. Hooker. Physics and metaphysics: a prolegomnon for the riddles of quantum theory. In C. A. Hooker (ed.) *Contemporary Research in the Foundations and Philosophy of Quantum Theory.* Reidel, Holland, 174-304, 1973.

[Hooker, 1974] C. A. Hooker. The Metaphysics of Science: Atoms versus Plena. *International Logic Review*, 9, 111-146, 1974.

[Hooker, 1981b] C. A. Hooker. Towards a general theory of reduction. *Dialogue*, XX, Part I, Historical framework, 38-59, Part II, Identity and reduction, 201-36, Part III, Cross-categorial reduction, 496-529, 1981.

[Hooker, 1991] C. A. Hooker. Physical Intelligibility, Projection, Objectivity and Completeness: The Divergent Ideals of Bohr and Einstein. *British Journal for the Philosophy of Science*, 42, 491-511, 1991.

[Hooker, 1994] C. A. Hooker. From phenomena to metaphysics. In D. Prawitz, D. Westerstahl (eds.) *Logic and Philosophy of Science in Uppsala*, Dordrecht, Kluwer, 159-184, 1994.

[Hooker, 1995] C. A. Hooker. *Reason, regulation and realism.* State University of New York Press, Albany, 1995.

[Hooker, 2004] C. A. Hooker. Asymptotics, reduction and emergence. *British Journal for the Philosophy of Science*, 55, 435-479, 2004.

[Hooker, 2006] C. A. Hooker. Reduction as Cognitive Strategy. In B. Keeley (ed.) *Paul Churchland*, Cambridge University Press, New York, 154-174, 2006.

[Hooker-a, this volume] Introduction to philosophy of complex systems. Part A: Towards framing philosophy of complex systems. In this volume.

[Hooker-b, this volume] Introduction to philosophy of complex systems. Part B: An initial scientific paradigm + philosophy of science for complex systems. In this volume.
[Humphries, 2004] P. Humphries. *Extending ourselves: computational science, empiricism and scientific method*. Oxford University Press, Oxford, 2004.
[Israeli, Goldenfeld, 2006,] N. Israeli, N. Goldenfeld. Coarse-graining of cellular automata, emergence, and the predictability of somplex systems. Physical Review E, 73**,** 26203-26220, 2006. Pre-print available at `http://arXiv.org/abs/nlin/0508033.`
[Kim, 1998] J. Kim. *Mind in a physical world: an essay on the mind-body problem and mental causation*, MIT Press, Cambridge Mass.,1998.
[Krupková, 2009] O. Krupková. The nonholonomic variational principle. *Journal of Physics A: Mathematical and Theoretical*, 42, 185201, 2009.
[McAllister, 2003] J.W. McAllister. Algorithmic randomness in empirical data. *Studies In History and Philosophy of Science Part A*, 34(3), 633-646, 2003.
[Marras, 2002] A. Marras. Kim on reduction. *Erkenntnis*, 57, 231-57, 2002.
[Miller *et al.*, 2000] J. F. Miller, D. Job, V. K. Vassilev. Principles in the evolutionary design of digital circuits - part I. *Journal of Genetic Programming and Evolvable Machines*, 1(1), 8–35, 2000.
[Moreno, 2007] A. Moreno. A systemic approach to the origin of biological organization. In [Boogerd *et al.*, 2007, 243-268], 2007.
[Moreno and Ruiz-Mirazo, 1999] A. Moreno and K. Ruiz-Mirazo. Metabolism and the problem of its universalization *Biosystems*, 49, 45-61, 1999.
[Nagel, 1961] E. Nagel. *The Structure of Science*. Harcourt, Brace and World, New York, 1961.
[Neurath *et al.*, 1971] O. Neurath, R. Carnap, C. Morris. Encyclopedia of unified science. University of Chicago Press, Chicago, 1971.
[Polanyi, 1968] M. Polanyi. Life's irreducible structure. *Science*, 160, 1308-1312, 1968.
[Prigogine, 2003] I. Prigogine. Chemical Kinetics and Dynamics. Annals of the New York Academy of Science, 988, 128-132, 2003.
[Rosen, 1991] R. Rosen. *Life Itself*. Columbia University Press, New York, 1991.
[Ryan, 2007] A. Ryan. Emergence is coupled to scope, not level. *Complexity*, 13(2), 67-77, 2007. Pre-print available at `http://arxiv.org/abs/nlin/0609011`.
[Strogatz, 2001] S. Strogatz. Exploring complex networks. Nature, 410**,** 268-277, 2001.
[Wimsatt, 1974] W. Wimsatt. Complexity and organization. In K. F. Schaffner, R. S. Cohen (eds.) *PSA-1972*, Proceedings of the Philosophy of Science Association, 67-86, Reidel, Dordrecht, 1974.

CHALLENGED BY INSTABILITY AND COMPLEXITY...

QUESTIONING CLASSIC STABILITY ASSUMPTIONS AND PRESUPPOSITIONS IN SCIENTIFIC METHODOLOGY

Jan C. Schmidt

1 INTRODUCTION: THE STABILITY ASSUMPTION IS UNSTABLE ...

Nonlinear Dynamics — including Complex Systems Theory, Chaos Theory, Synergetics, Dissipative Structures, Fractal Geometry, and Catastrophe Theory — is a young and fascinating field of scientific inquiry that spans many established disciplines (cf. [Mainzer, 1996]). However, it poses challenging problems for both scientific methodology and the philosophy of science. Methodological prerequisites as well as metaphysical assumptions are questioned, e.g., predictability, reproducibility, testability, explainability as well as lawlikeness (determinism/causation). The common denominator of all of these challenges is instability — that is the main thesis of this paper.

Since the advent of Nonlinear Dynamics and its advancement in the realm of physics in the 1960s — interlaced with methodological developments in computer technology and the computer's ability to numerically handle nonlinearity — further evidence for the existence and prevalence of unstable and complex phenomena in the physical world has emerged. Nonlinear systems, even those with just a few degrees of freedom, can exhibit static, dynamical and structural instabilities. Although instabilities call implicit metaphysical-methodological convictions and well-established underlying prerequisites of mathematical science into question, today they are not viewed in just a negative way. On the contrary, instabilities are highly valued — we find a *positivization of instabilities*: instabilities constitute the nomological nucleus of self-organization, pattern formation, growth processes, phase transitions and, also, the arrow of time (cf. [Schmidt, 2008a]). Without instability, there is no complexity and no change. The phenomena generated by underlying instabilities in nature, technology and society are manifest; we can observe these phenomena with our unaided senses. In fact, instability is the root of many homely phenomena in our day-to-day experience — for example, the onset of dripping from a tap or water freezing to ice in a refrigerator. Instability has to be regarded as an empirical fact of our life-world and beyond — not just as a contingent convention.

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

A reconsideration of the traditional methodological-metaphysical stability assumptions therefore seems to be indispensable. (a) In the past, stability was taken for granted as an implicit *a priori* condition to qualify a mathematical model as physically relevant or adequate. Stability seemed to be a key element underlying any kind of physical methodology: it was regarded as the sole possibility to guarantee the application of methods of approximation and, also, to deal with empirical and experimental uncertainties. (b) In addition to methodology, an underlying metaphysical conviction was pervasive throughout the history of physics, guiding the focus of interest and selecting the objects that were considered worth researching. Framing and conceptualizing nature as "nature" insofar as it is stable, time-invariant and symmetrical (metaphysics), was indeed a successful strategy to advance a specific physical knowledge (methodology). It is interesting to see that metaphysical convictions and methodological considerations are interlaced; there is no clear line between metaphysics and methodology, as will be shown in this paper.

Throughout history, stability metaphysics has always played a major role in science, beginning in ancient times with Plato's stability concept of the cosmos. In modern times, stability metaphysics can be found in the works of outstanding physicists such as Newton and Einstein. For instance, in his *Opticks* Newton did not trust his own nonlinear equations for three- and n-body systems which can potentially exhibit unstable solutions [Newton, 1730]. He required God's frequent supernatural intervention in order to stabilize the solar system. In the same vein, Einstein introduced ad hoc — without any empirical evidence or physical justification — the cosmological constant in the framework of General Relativity in order to guarantee a static and stable cosmos, "Einstein's cosmos" [Einstein, 1917]. Both examples, from Newton and Einstein, illustrate that metaphysical convictions — *what nature is!* — can be incredibly strong, even if they are in conflict with what is known about nature at the time.

Today, however, *ex post* and thanks to the advancement of Nonlinear Dynamics, we can identify a "dogma of stability" that has determined the selection (or construction) of both the objects and the models/theories in physics. "We shall question the conventional wisdom that stability is an essential property for models of physical systems. [...] The logic which supports the *stability dogma* is faulty." [Guckenheimer and Holmes, 1983, p. 259]: the stability assumption is itself unstable! Although the discovery history of instabilities traces back to physicists such as Newton, Laplace, Stokes, Maxwell, Poincaré and Duhem, physical objects were (and often still are) perceived and framed from the perspective of stability — even by the pioneers of instabilities. Throughout the history of exact sciences, instabilities were not acknowledged by the scientific community. This has been changing since the 1960s when physics began widening its methodological horizon — including getting rid of the restriction of methodology to stability requirements. The need to advance physical methodology emerged because instabilities have turned out to be so very fundamental in nature, technology, and even in social processes. In order to deal with instabilities, physicists have over the last 30 years successfully replaced the traditional quantitative, metrically oriented stability dogma by weaker, more qualitative topological characteristics. Many models (theories, laws) in Nonlinear Dynamics are unstable, "and we are confident that these [...] are realistic models [...] of corresponding physical systems" [Guckenheimer and Holmes, 1983, p. 259].

Nonlinear Dynamics shows that instability is not an *epi*phenomenon of minor relevance: instabilities are broadly present in our entire world. Discovering and acknowledging instabilities impels both a reconsideration of the metaphysical views that undergird the *stability dogma* and a revision of the methodological presuppositions. The outline of this paper is as follows: In section 2, I characterize instabilities and distinguish between three kinds of instability. In section 3, I focus on methodological problems and challenges caused by instabilities; the limitations of classical-modern sciences will be discussed. In section 4, I show how present-day physics manages, at least to some degree, to cope with instabilities.

Instabilities cannot be considered as exceptions within a stable world. Rather, it is the other way around: instabilities are the source of complexity, pattern formation and self-organization. This is why instabilities do not only appear in a negative light; a positive understanding emerges and shows challenging future prospects and perspectives for the rapidly progressing field of Nonlinear Dynamics — and beyond: for all mathematical sciences.

2 INSTABILITIES

2.1 Precursors of the Notion of "Instabilities" — and the Nucleus of Nonlinear Dynamics

For a long time, physicists disregarded instabilities. Nature was defined as "nature" by stability. Only those models that generate stable solutions were viewed as *good* mathematical models (laws, theories) of physical nature. The stability assumption dates back at least to the beginning of modern mathematical sciences in the 17^{th} century or, even further, to antique natural philosophy.

Since ancient times it was not doubted that instabilities allude to a deficit of knowledge. Nature, order, and stability were mainly used as synonyms. Plato's cosmos was structured by a Demiurge according to the idea of simple mathematical laws, intrinsic harmony, time-invariant order and universal stability. Even if harmony, simplicity and stability of nature seemed to be hidden at first glance, there was no doubt about their existence behind the apparent complexity, guaranteeing the timeless and infinite existence of nature as a whole. Plato's foundation of theoretical science was so influential that, from the beginning of modern science in the early 17^{th} century, mathematical invariants — from Platonic bodies to Newton's laws and Einstein's equations — were considered the core of nature. Modern science did not scrutinize whether nature is stable or not. Science operated on the implicit (normative) stability assumption: all mathematical laws, models and theories *have to* be stable. It is not surprising that, in consequence, nearly all laws show stability.

The disregard and ignorance of instabilities is remarkable because physicists have always been aware of the potential existence of unstable processes in nature as well as in models. The problem has been broadly identified ever since the development of hydrodynamics in the 1800s (e.g., Navier-Stokes equations) (cf. [Darrigol, 2006]). Poincaré's description of the solar system and Einstein's general relativity reveal several types of instabilities. And before then, instabilities were discussed in Newton's theory of the moon, and in Maxwell's "matter and motion". Maxwell in particular was a precursor in identifying instabilities. According to him, in physics "there is [... a stability] maxim which asserts 'That *like* causes produce *like* effects'. [But,] [t]his is only true when small variations in the initial circumstances produce only small variations in the final state of the system. [...] [T]here are other cases in which a small initial variation may produce a very great change in the final state of the system, as when the displacement of the 'points' causes a railway train to run into another instead of keeping its proper course" [Maxwell, 1991, pp. 13/14]. Maxwell was the first to identify the relevance of sensitive dependence on initial conditions — which is the nucleus of instability. However, the prevalence of instabilities was not obvious in the 19^{th} century. Instabilities were not broadly acknowledged at that time; they were regarded as exceptions within a stable world: instability just on the fringes of stability.

This changed during the 1960s — the decade of the beginning of the *structural scientific revolution* caused by the rapid advancement of microelectronics and the development of computer technology: Ed Lorenz discovered the "butterfly effect" [Lorenz, 1963; 1989], Hermann Haken developed his synergetics (cf. [Haken, 1977]), and Ilya Prigogine formulated the nonlinear thermodynamics of far-from-equilibrium dissipative structures [Prigogine and Glansdorff, 1971]. In the early 1970s, David Ruelle and Floris Takens presented an influential paper on solutions of the hydromechanical Navier-Stokes equations and coined the term "chaos" (metaphorically) for a certain class of solutions of partial differential equations [Ruelle and Takens, 1971]; in 1975, Li and Yorke published an epoch-making mathematical article on one-dimensional difference equations (maps) that highlights "chaos" explicitly in the title and provided a mathematically clear definition: "Period Three implies Chaos" [Li and Yorke, 1975; cf. May, 1976]. During the 1970s, Benoit Mandelbrot worked out his "fractal geometry" [Mandelbrot, 1991] and René Thom presented his "catastrophe theory" [Thom, 1975]. Many ideas were developed simultaneously — it was an inspiring and exciting time of change, inducing a *structural scientific* paradigm shift. The main overarching achievement was that the relation of stability and instability was reversed: no longer was instability viewed as a minor subset of stability. On the contrary, stability has become one tiny island in an ocean of instability; it is "order at the edge of chaos". One important lesson of these new theories for all mathematical sciences today is the fundamental role of instability. Instability — not stability — is the prevalent case [Holmes, 2005; Aubin and Dalmedico, 2002; Schmidt, 2008a; 2008b].

Instabilities are, in general, situations in which a system is on a razor's edge: criticalities, flip points, thresholds, watersheds, sharp wedges. They generate butterfly effects or sensitive dependencies, bifurcations, points of structural changes and phase transitions. The list of examples is extensive (cf. [Mainzer, 1996]): the emergence and onset of a chemical oscillation, the roll dynamics of a fluid in heat transfer, an enzyme kinetic reaction, a gear chattering, or turbulence of a flow. A fluid becomes viscous, ice crystallization emerges, a phase transition from the fluid to a gas phase takes place, a solid state becomes super-fluid, laser light issues, a water tap begins to drip, a bridge crashes down, an earthquake or a tsunami arises, a thermal conduction comes to rest and a convection sets in, e.g., Bénard instability. New patterns and structures appear; instabilities are *the necessary* condition for *novelty.* The various definitions and meanings of complexity — even if they do not refer to the genesis and evolution of a new pattern such as those in more geometric definitions of "complexity" via "dimensions" — refer directly or indirectly to instabilities [Atmanspacher *et al.*, 1992; Wackerbauer *et al.*, 1994].

Our world is essentially a world of dynamics, change and complexity. The mere empirical fact that self-organization and pattern formation is possible provides evidence that instabilities are not a mere convention but exist in the world *(minimal instability realism*). In a comprehensive paper the physicist J.S. Langer underlines the role of "instabilities for any kind of pattern formation" [Langer, 1980]. According to Werner Ebeling and Reiner Feistel, "self-organization is always induced by instability of the 'old' structure via small fluctuations" [Ebeling and Feistel, 1990, p. 46]. "This is why studying instability is of major importance" [ibid.]. Gregory Nicolis and Ilya Prigogine argue that "instabilities" are "necessary condition for self-organization" [Nicolis and Prigogine, 1977, pp. 3]. And Wolfgang Krohn and Günter Küppers emphasize that "instabilities are the driving force for systems evolution and development" [Krohn and Küppers, 1992, p. 3]. In addition to novelty, nonlinear scientists refer to temporality ("time's arrow"). Classical-modern science was "unable to tackle and to resolve the antagonism between reversibility and irreversibility. To accomplish this, it was necessary to introduce a new concept: the concept of instability" [Ebeling and Feistel, 1990, p. 197].

To summarize, instabilities are the common denominator of all theories in the realm of Nonlinear Dynamics. Complex phenomena, for instance, can be regarded as a derivative and consequence of instabilities. The priority of instability can also be shown for: deterministic chance, randomness, chaos, turbulence, catastrophes, fractals, self-organization/emergence, symmetry breaking, phase transitions, time's arrow/irreversibility, information loss or gain. The same argument, however, holds concerning the essential role of nonlinearity in these phenomena. But nonlinearity in itself does not challenge traditional physical methodology. Many systems are nonlinear but stable, for example simple two-dimensional planetary systems. They are, like linear systems, without chaos and turbulence, without self-organization and symmetry breaking. *Instability, and not nonlinearity by itself, makes the difference.* In the following, I will explicate "instabilities" in more detail. From an analytical perspective there are at least three cognate but different types of instability that can be distinguished: static, dynamical and structural instability.

2.2 Static Instability: Watersheds

The fact that small differences in the initial conditions can cause large effects in the future is not only a property of deterministic chaos: watershed or static instability does the same, and is more prevalent. Henri Poincaré referred to this type of instability at the end of the 19^{th} century [Poincaré, 1892]. Static unstable situations resemble a pen standing on its tip; after a while the pen will fall down to one side or another. A ball on a roof ridge will roll down to one side or the other. The same occurs with a mass of a physical oscillator at the unstable point of maximum potential energy. In a pinball machine the ball hits sharp wedges and jumps to the right or left. The quincunx ("Galton's apparatus") consists of rows of alternating pins that are the source of instabilities. A tiny ball that falls on a pin passes it either on the right or on the left — this happens to the ball several times on several rows; finally, we observe a random-looking (Gaussian) distribution of all tiny balls at the bottom. From a mathematical viewpoint, the main idea of instability traces back to Poincaré: "A very small cause which escapes our notice determines a considerable effect that we cannot fail to see, and then we say that the effect is due to chance" [Poincaré, 1914, pp. 56].[1] At points of static instability a sensitive dependence on initial conditions occurs. The alternative trajectories from two nearby initial points separate and will never again become neighbors. In the course of the evolution of the system, one raindrop falling near the "height of land" may be transported to the Mediterranean Sea, while its neighbor goes to the North Sea. Between the two raindrops there exists a watershed or a basin boundary. Many watersheds and unstable points govern playing dice. That is why instabilities also once served as an underlying idea for the development of classical probability theory. To Jacob Bernoulli, instabilities in coin tossing — and, based on this, the resulting binary sequence of 0 ("pitch") and 1 ("toss") — were paradigmatic for randomness. Instabilities generate nomological randomness or, to put it in other words, deterministic chance.

Cognate to, but different from static watershed instability is a type of instability that is based on the sensitive dependence of *all* points in the state space (*static continuous instability*). A prominent example is compound interest. Tiny differences in seed capital grow exponentially into enormous differences years later. The linear differential equation $dx/dt = ax$ yields a solution that can be determined in an analytical way; an integration is possible: $x(t) = x(0) \cdot e^{at}$ with a constant (Lyapunov) coefficient $a > 0$. Trajectories starting from a neighborhood reveal a static continuous divergence; in an open-ended state space they need never converge to each other again. However, a closed solution does not exist for this linear differential equation. If the initial conditions are known, then a calculation for all times is feasible and a prediction is possible.

[1] Compare Maxwell's [1991, p. 13] remark in 2.1 above of forty years before.

Thus, time evolution is of minor relevance for static watershed or static continuous instability. After a certain time, the initial points have left the neighborhood of the watershed: then the static watershed instability no longer has an influence on time evolution. The same holds for static continuous instability; here, instability is not an epistemological or methodological challenge.

2.3 Dynamical Instability: Chaos

According to Friedrich Nietzsche, the "underlying character of the world is chaos" [Nietzsche, 1930, p. 127]. Chaos is not associated with lawlessness and disorder but with dynamics and time evolution. Nietzsche's chaos differs from the messy *Tohuwabohu* of the Judeo-Christian tradition. This is why Nietzsche can be regarded as a precursor of our recent understanding of "chaos" based on dynamical instability. Referring to Nietzsche, Martin Heidegger advocates a positive understanding of chaos: Chaos means "dynamics, change, transition — whereas regularity seems to be hidden: we do not know the underlying law at first glance" [Heidegger, 1986, p. 87].[2] Clearly, Nietzsche and Heidegger did not define "chaos" in mathematical terms. It was up to mathematicians in the 1970s to provide a distinct definition (see 2.1 above).

Chaos is characterized by dynamical instability. The time evolution of a chaotic system is continuously on a knife's edge. Nearly all points in the state space exhibit the property of sensitivity: the trajectories separate exponentially by time evolution. Because chaotic attractors are bounded in the state space we find exponential separation merely as a mean value. For short time-scales, two trajectories may converge before they diverge again, and so on — an interplay of divergence and convergence among trajectories. However, dynamically unstable (chaotic) trajectories do not show an *exact* recurrence. Henri Poincaré, a precursor of modern Nonlinear Dynamics, studied the phenomenon of recurrence more than one hundred years ago. Today, we utilize Poincaré's fundamental ideas to construct a specific map in order to obtain insights into the time evolution of a dynamical system — the so-called "Poincaré map". The dynamically unstable trajectory of a chaotic attractor resembles a ball of wool. It covers the attractor densely. Nearly all points are reached approximately by the trajectory after a certain time, for instance in the case of a chaotic double pendulum.

Dynamical instability is the core of deterministic chaos. Chaos is associated with random-like behavior generated by deterministic models in a spatial-continuous state space in which every point is a watershed or divergence point: a random *pheno*type, in spite of a nomological *geno*type. Although Poincaré and later George Birkhoff [1927] were aware of the possibility of random-like behavior they did not have computer resources and numerical simulation tools to deal approximately with dynamically unstable systems. In the 1960s, Ed Lorenz, who discovered chaos in atmospheric dynamics through computer simulations, coined the

[2] Translation from German (JCS).

term “butterfly effect” to illustrate its real impact: for dynamically unstable systems a tiny event such as a butterfly’s wing flap in South America can cause a thunderstorm in the U.S. [Lorenz, 1963; 1989].

It is interesting to note that more than 15 competing notions and definitions of “chaos” exist (cf. [Brown and Chua, 1996; 1998]). One of the most relevant chaos definitions nowadays has been suggested by Robert Devaney [1987]. Devaney requires a system’s dynamics to exhibit three features in order to be called chaotic: (a) sensitive dependence on initial conditions (unpredictability); (b) topological transitivity (non-decomposability, wholeness); and (c) denseness of periodic points (elements of regularity). John Banks *et al.* [1992] show from a mathematical perspective for one-dimensional maps that sensitivity (a) follows from properties (b) and (c). So sensitivity does not seem to be a fundamental requirement in order to provide an adequate understanding of dynamical instability. From an empirical perspective, Devaney’s and Banks’ definitions — focusing on (b) and (c) — are regarded as interesting but hard to determine. Unlike mathematicians, physicists mostly advocate an understanding of chaos based on sensitivity coefficients. A classical sensitivity coefficient is Lyapunov’s coefficient that measures the average divergence of trajectories (“chaos coefficient”). For physics and its methodological need to deal with experimental uncertainties, a positive Lyapunov coefficient is one major characteristic of dynamical instability.

2.4 Structural Instability: Bifurcations, Thresholds, and Criticalities

There is another type of instability that has not received the attention it deserves — structural instability, which refers to changes in dynamical models, laws and equations. For example, it occurs in Rayleigh-Bénard convection where a fluid layer is confined between two thermally conducting plates and is heated from below, creating a vertical temperature gradient. For lower temperature gradients — and hence smaller density gradients — , an upward conductive heat transfer through horizontally uniform layers takes place. But if the temperature gradient is sufficiently large, a point of structural instability occurs. The hot fluid rises, leading to a circular convective flow in three-dimensional cells, which results in enhanced transport of heat between the two plates.

Structural instability does not refer to initial points and particular trajectories, but to the whole dynamical system, reflected in the mathematical structure of the model itself, for example in its Hamiltonian form and state space. In general, structural instability is expressed as follows: if one alters the structure of a model (equation, law) slightly, the overall dynamics changes qualitatively. The equivalence class of the model is *not* preserved; the disturbed and undisturbed models show different topological characteristics, for instance different types of orbits or, at least, different periodicities. Depending on the initial model class, altering a model can be performed by varying the model’s equations, the system parameters or exponents, by adding mathematical functions or by modifying the state space. This might describe, for instance, the onset of a chemical oscillation or of a fluid roll dynamics under heat transfer, the emergence of laser light or a water tap beginning to drip.

The first exact mathematical definition of structural instability traces back to Andronov and Pontryagin [1937]. They coined the term “robustness”, also meaning “coarse-grainedness” for structural stability, and indicaating the lack of structural instability. In addition, Birkhoff [1927] can also be regarded as a precursor of the concept of structural instability/structural stability. In his influential book *Dynamical Systems* he defines and distinguishes between different types of stability, such as trigonometric stability, stability of the first order, permanent stability, semi-permanent stability, unilateral stability, regions of stability, stability in the sense of Poisson (cf. [Birkhoff, 1927]). Some of Birkhoff’s definitions paved the way and gave rise to what would later be known as structural instability. When *dynamical systems theory* was further developed, Andronov’s and Pontryagin’s term “robustness” was replaced by “structural stability” with no content change in the mathematical definition. Today, structural stability, with its converse of structural instability, is a major part of bifurcation theory (cf. [Wiggins, 1988; Ruelle 1989]).

Unstable phenomena in nature are not just exceptions in an ocean of stability; rather, it is the other way around. According to Devaney, structural instability is most prevalent, even in classical physics. “Most dynamical systems that arise in classical mechanics are not structurally stable” [Devaney, 1987, p. 53]. Given its prevalence, structural instability deepens and extends our discussion. Since Plato, theories (models, laws, equations) have been considered as the nucleus of nature as well as the goal of science. In contrast to the mathematical body of theories, initial conditions — most important in cases of static and dynamical instability — were regarded as merely contingent factors or an outer appendix; structural instability does not refer to initial conditions but to theories.

3 METHODOLOGICAL CHALLENGES

3.1 What is being challenged?

It is common in *philosophy of science* to characterize exact science by its methodology. Although there is still an ongoing debate without any final consensus, the center of physical methodology can be considered as consisting of four interlaced elements that also highlight distinguishable objectives and different understandings of science: reproducibility/repeatability, predictability, testability, and describability/explainability. Peter Janich, for instance, identifies the “striking technical, prognostic and explanatory success” of physics (cf. [Janich 1997, p. 62]). Janich’s list might be complemented by elements such as the successful testing of physical theories in experiments (testability).[3]

[3]However, it is important to stress, as Janich does convincingly, the experimental/engineering aspects of physics: physics aims to produce experimentally reproducible phenomena, and it is based on experimentation (cf. [Hacking, 1983]).

Instabilities pose challenges to these well-established methodological elements of mathematical science and reveal several problems and limitations. In order to show the prevalence of instability-based issues for exact science in general, I will draw attention to the challenges relating to all four elements: reproducibility (3.2), predictability (3.3), testability (3.4), and explainability (3.5).

3.2 Problems with Reproducibility and Experimentation

The center of modern sciences, introduced by Galileo and Bacon, is the method-based experiment. By means of intentional interventions in nature we constitute, construct and control empirical facts of nature; these facts are mostly not given by themselves but have to be produced, made visible and measurable (cf. [Hacking, 1983]). According to Francis Bacon and Immanuel Kant, experiments serve as "interrogation techniques" to "force nature to reveal her secrets". Experiments — and not passive observation with unaided senses — guarantee empirical evidence, objectivity or, at least, intersubjectivity: Anybody performing a particular experiment in any place and at any time will obtain the same result — if she or he has acted in accordance with the norms of experimentation. Classical-modern science is based on the methodological assumption of person, location and time independency. It is presupposed that we can control the experimental setup including the relevant boundary conditions and, by this, sufficiently isolate the system from its environment to ensure this (*principle of isolation*): e.g., a bumble-bee on the planet Sirius as well as a falling leaf in another country has (*or should have*!) no influence on the experiment.

The distinguishing character of an experiment is therefore reproducibility. For a long time physicists have not doubted that reproduction is in principle feasible, even though it may not yet have been accomplished concerning a specific phenomenon under consideration. Jürgen Mittelstraß believes that "reproducibility [...] is the major scientific norm." "The requirement of reproducibility is an indispensable criterion to define 'science' as science: in fact, it is the necessary condition for the possibility of sciences and scientific knowledge!" [Mittelstraß, 1998, p. 107] Robert Batterman stresses that "any experiment in which a phenomenon is considered to be manifest must be repeatable" [Batterman, 2002, p. 57]. Gernot Böhme and Wolfgang van den Daele argue that the "methodological concept of science is the regular fact that includes the condition under which it can be reproduced by anybody at any time" [Böhme and van den Daele, 1977, p. 189]. Friedrich Hund states that exact science is the "realm of reproducible phenomena" [Hund, 1972, p. 274]. Wolfgang Pauli emphasizes that — in spite of astrophysics and cosmology — "exceptional and extraordinary events are far beyond the scope of physics; these events cannot be grasped by experiments" [Pauli, 1961, p. 94].

If we were to subscribe to these statements and restrict physics to reproducible phenomena, we would find ourselves in a dead end: physics and instability would be like fire and water. This is, however, not the case; certainly, it has to be admitted that the challenges induced by instabilities are severe. If infinitesimally small changes in the initial or boundary conditions grow rapidly with time, the experimental system with the physical object of interest cannot be isolated from its surrounding. The bumble-bee on Sirius might have an impact on the experimental result and therefore has to be taken into account — which obviously is impossible. Any unstable object performs its own singular trajectory. Instabilities convey unobservable small effects to the empirically accessible scales; they bridge different spatial domains, e.g., microcosm, mesocosm, and macrocosm, and by this, instabilities induce problems regarding experimentation. Repeatability is limited; reproducibility and the intentional (re-)production of events are difficult to achieve with unstable objects. Because of thermodynamic and quantum mechanical effects, initial and boundary conditions cannot be measured exactly and cannot be controlled in detail; empirical errors cannot be eliminated. Referring to the three-body problem and its instability, Poincaré emphasizes that "even if it were the case that the natural laws no longer had any secret for us,we could still only know the initial situation *approximately*" [Poincaré, 1914, p. 56]. Inherent inexactness causes severe problems for physics and, in fact, for all empirical sciences. Instability limits intentional action, production and re-production; the dynamics of unstable objects cannot be controlled by the experimenter. However, the lack of control is not just a pragmatic or epistemic boundary that could be overcome by improvement of methods and more advanced technology. It is part of nature; it is inherent in physical objects, and not just a challenge to knowledge: it is ontology rather than epistemology.

3.3 Boundaries of Predictability, Calculability and Mathematics

Predictability is considered an important qualifier of physics. Among those who cite predictability are instrumentalists and pragmatists, and sometimes others such as realists and transcendental epistemologists. Carl Friedrich von Weizsäcker argues that "determining and predicting the future" is the "major essence of physics" [Weizsäcker, 1974, p. 122]: Physics enables shortcuts to the future. In line with Weizsäcker, Michael Drieschner emphasizes: "We find 'prediction' as the key term for understanding and defining physics" [Drieschner, 2002, p. 90]. Einstein, Podolsky and Rosen regard predictability as the sole possibility for checking whether anything particular does indeed exist. "If, without in any way disturbing a system, we can predict [at least in principle] with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity" [Einstein *et al.*, 1935, p. 777]. According to Herbert Pietschmann, "the power and accuracy of prediction indicates whether to accept or to refute a theory under consideration" [Pietschmann, 1996, p. 166]. Aurell *et al.* argue: "The ability to predict has been the most important qualifier of what constitutes scientific knowledge" [Aurell *et al.*, 1997]. Wesley C. Salmon's main criterion for a good explanation is the success of predictions: Explaining means to "show that the event to be explained was to be expected" [Salmon, 1989, p. 119]. According to Stathis Psillos, scientific realists "regard

[...] predictively successful scientific theories as well-confirmed and approximately true of the world" [Psillos, 1999, p. xix]. In addition, Robert Shaw stresses that "physics owes its success to its ability to predict natural phenomena, thus allowing man a degree of control over his surroundings" [Shaw, 1981, p. 218]. It is well known that Auguste Comte once said: "Savior pour prévoir" (knowing in order to foresee) (cf. [Comte, 2006]).

Today, instabilities challenge any prediction-oriented understanding of exact science. Mario Bunge emphasizes the "immense mathematical difficulties with nonlinear laws and, in addition, with unstable solutions" [Bunge, 1987, p. 188].[4] A differential equation provides a general corpus for a law of nature but not a specific solution or a particular trajectory. The specific solution is not given with the differential equation itself; it has to be computed in order to enable predictions. Linear differential equations do not cause any trouble; most of them can be integrated analytically. In contrast, nonlinear differential equations often do not possess analytic solutions. Newton, Euler, Laplace, Poincaré and others were frustrated by this fact. Even if we were ever to succeed in finding a nonlinear theory of everything — the equation that governs the universe — , we likely would *not* be able to predict specific events in the far future. According to Maxwell, "it is manifest that the existence of unstable conditions renders impossible the prediction of future events, if our knowledge of the present state is only approximate, and not accurate" [Maxwell, 1873, p. 440].

In some cases, however, numerical algorithms can help to integrate nonlinear differential equations and in handling unstable solutions. In the case of *static instability,* if the watershed is not fractal or the initial states are not located within the neighborhood of the watershed, then numerical integration is feasible. Small perturbations do not make any difference in the final result. In other cases, two initial states differing by imperceptible amounts may evolve into two considerably different states. If, then, there is any observation error of the present state, an acceptable prediction of the state (within only small errors) in the distant future may well be impossible. Similar conclusions hold for *structural instability.* The most challenging problems occur in cases when *dynamical instability* is present. Dynamically unstable chaotic orbits cannot be approximated by numerical solutions; they are effectively non-computable. Robert Devaney stresses that "most dynamical systems that arise in classical mechanics are not stable. [...] These systems cannot be approximated" in the classic way of asymptotic convergence of mathematical functions (cf. [Devaney, 1987, p. 53]). No numerical solution is accurate enough to determine an unstable trajectory; instability requires an accuracy that is impossible in order to accomplish prediction tasks.

A *first* numerical boundary is due to the impossibility of digitizing real numbers, as is necessary to describe unstable trajectories. Digital computers only calculate and store results to finite precision, so there are rounding errors at every calculation step. In an unstable system, these errors will grow exponentially, and

[4] Bunge believes that instability makes it too difficult to assign determinism to unstable dynamics.

so the model's trajectory (when initialized from a particular state) will quickly differ significantly from the evolution of the (real exact) system. A *second* numerical boundary arises because the computer is not a formal logical system (Turing machine) but a physical machine. A physical machine is subject to physical, and especially thermodynamic, relativistic and quantum mechanical limitations. Of practical relevance are mainly thermodynamic limits: calculations, information processing and annihilation require thermodynamic energy and produce entropy. Depending on the thermodynamic costs one is prepared to pay, different numerical limits exist with regard to the quality of approximation. General limits are due to the maximum energy that is available in our cosmos for performing a calculation. The limitation of predictions is, therefore, built into the structure of the world — it is not an epistemological artifact of the way we are, nor is it just a methodological or pragmatic limitation that can be overcome; the limitation is not "human-dependent" in any subjective sense. When dealing with unstable systems, Mitchell Feigenbaum states: "You know the right equations but they're just not helpful" (quoted in: [Gleick, 1987, p. 174]). The differential equation of a deterministic system is effectively worthless since reliable predictions would require exact information.

Instabilities drive a wedge between (deterministic) equations (laws, theories, models) and predictability — with consequences for a deterministic worldview.[5] Despite Hume's skepticism, traditionally successful predictions have been considered as the main argument in favor of a deterministic worldview. Whether an effectively reduced predictability still might provide evidence and a convincing argument for determinism remains an open question.

3.4 Limits of Testability, Confirmability, and Objectivity

It seems to be a well-established position in philosophy of science to argue that science, insofar as it is science, has to meet the criterion of testability: both realists and empiricists — verificationists, confirmationists and falsificationists alike — consider empirical testability as methodologically essential to science, in particular in order to ensure objective truth or, at least, intersubjective evidence.

Ernst Mach advocates a strong empiricist view of science: "In domains where neither confirmation nor refutation does exist, we do not meet scientific standards: in fact, there is no science!" [Mach, 1988, p. 465] Pierre Duhem argues that the "main criterion for scientific evidence" is for a mathematical law to be in "accordance with experience" [Duhem, 1991, p. 22]. In line with this approach, Michael Redhead stresses that "testing of theories" means to make "predictions [based on mathematical equations of the theory] for comparison with experiment" [Redhead, 1980, p. 156]. Rudolf Carnap argues that "a statement to be meaningful" has to be based on empirical verification; "testability and meaning" turn out to be synonyms (cf. [Carnap, 1928]); "verifiability" — later Carnap preferred "confirmability" —

[5] In particular, between (deterministic) laws and (prediction-relevant single) trajectories.

is the major criterion to demarcate science from metaphysics. From another perspective, Karl R. Popper does not believe in confirmation at all but in refutation (cf. [Popper, 1934]). The growth of scientific knowledge is based on risky conjectures that (normatively: have to!) allow refutations. Most important are the so called *experimenta crucis* that can reveal the falsehood of a statement. Such decision supporting experiments are indispensable to qualify a theory as a scientific one — also showing the well-known asymmetry between the methodologically justified, because logically valid, refutation and the methodologically unjustified, because logically invalid, verification.

So, the requirement of testability as a criterion to qualify and distinguish science is broadly shared among critical rationalists, hypothetical realists and logical empiricists.[6] Any type of empirical test — for Carnap as well as for Popper — is based on a constant relation between the mathematical-theoretical side on the one hand and the empirical-experimental side on the other. If dynamical or structural instabilities are present, this relation is no longer given. The behavior of a single trajectory or a few orbits deduced from the mathematical model "cannot be compared with experiment, since any orbit is effectively non-correlated with any other and numerical round-off or experimental precision will make every orbit distinct" [Abarbanel *et al.*, 1993, p. 1334]. Similarly, Rueger and Sharp stress: "If we test a theory in this [quantitative] way we will not find a precise quantitative fit, and this is to be expected if the theory is true of the system" [Rueger and Sharp, 1996, p. 103]. Theory and experiment are separated into two disjunct worlds (cf. [Harrell and Glymour, 2002]).

Both the instability on the object's side and the instability on the model's side contribute to limiting testability: Because of instability, it is impossible to reproduce the object's behavior and it is hard to make predictions. One of these problems would be hard enough to cope with, but they emerge simultaneously. This is why Guckenheimer and Holmes emphasize: for any unstable model that refers to an unstable object "details of the dynamics, which do not persist in perturbations, may not correspond to testable [...] properties" [Guckenheimer and Holmes, 1983, p. 256]. A classical-modern test, based on quantification, is not possible — one enters the realm of uncertainty.

3.5 Challenges: Describability and Explainability

According to Descartes the objective of modern science is to "trace back vague, clouded and dark propositions step-by-step to simple and evident propositions [laws]" [Descartes, 1979, pp. 16/379]. Later, in 1948, Hempel and Oppenheim proposed the deductive-nomological model to describe (and to *evaluate*!) scientific explanations (covering-law model). A phenomenon is regarded as being explained by "subsuming it under general laws, i.e., by showing that it occurred in accordance with these laws, in virtue of the realization of certain specified conditions"

[6]It is still an ongoing debate whether testability is to be understood as a normative requirement or merely as a description of what in fact happens in science.

[Hempel, 1965, p. 246]. "Explanation in the special sciences involves subsumption under general laws" [Woodward, 2000, p. 197]. To put it in normative terms: experimental events have to be described by the shortcut of mathematical structure; data needs to be compressed by algorithms or laws (including Ockham's razor). In this rationalist tradition, the aim of science is to achieve a minimal, simple, non-redundant description of the world; the challenge for exact sciences is to find laws, models, algorithms as the major syntactic elements of any theory. The elimination of redundancy and the downsizing of description length is regarded as a necessary condition for explanatory success. According to Hertz, "all physicists agree upon the main task: Tracing the phenomena of nature back to simple laws of mechanics" [Hertz, 1963, p. xxv]. Sometimes this position is called explanatory or micro-reductionism or, more strongly, inter-theoretical reductionism. The history of the philosophy of science has mainly focused on these challenges, in particular on ontological, epistemological and methodological types of reductionism. The unification of three of the four fundamental theories in physics shows the success of explanatory reductionism.

Instabilities pose challenges to this kind of reductive explanation. Early considerations, however, date back to David Hume in the 18^{th} century. Hume anticipated problems of instability in his *Inquiry Concerning Human Understanding* — although he did not use the term "instability": "It is only on the discovery of extraordinary phenomena, such as earthquakes [...], that they [= men] find themselves at a loss to assign a proper cause, and to explain the manner in which the effect is produced by it. It is usual for men, in such difficulties, to have recourse to some invisible intelligent principle as the immediate cause of that event which surprises them" [Hume 1990, p. 69]. Today, physicists regard earthquakes as a paradigm of unstable natural processes. Similarly, Hume considered earthquakes to be unstable, irregular real phenomena of nature, also supporting his arguments against anti-naturalist positions. In addition to Hume, one hundred years later Maxwell pointed out that "in so far as the weather may be due to an unlimited assemblage of local instabilities it may not be amenable to a finite scheme of law at all" [Maxwell, 1991, p. 13]. In other words, "only in so far as stability subsists [...] principles of natural laws can be formulated" (ibid.). Only a few scientists in Maxwell's time were aware of this challenge. Nowadays, the challenge provoked by instabilities to any reductive and compression-oriented understanding of description and explanation is broadly acknowledged and deeply discussed within the context of Nonlinear Sciences. "Nonlinear dynamical systems theory [...] studies properties of physical behavior that are inaccessible to micro reductive analytical techniques", as Kellert emphasizes [Kellert, 1994, p. 115]. The "Devil is in the detail", Batterman [2002] stresses. Instabilities limit "approximate reasoning" and "reductive explainability" which is, in fact, the nucleus of the classic deductive-nomological account of explanation (cf. [Hooker, 2004]).

Unstable processes cannot be reductively condensed to a simple law. The effective incompressibility of data sequences generated by unstable processes is known from information and chaos theory (cf. [Chaitin, 1971; 1987; 2001]). The key notion here is that of informational *incompressibility*, which is linked to *essential unpredictability*. According to von Neumann's idea on complexity, a complex process is defined as one for which the simplest model is the process itself. The only way to determine the future of the system is to run it: there are no shortcuts and no compact laws. Insofar as instability underlies complexity, the simplest model of the unstable process is the process itself. Dynamical instability could also be regarded as the heart of deterministic random processes, e.g., various types of non-white noise. Although these random processes might be generated by a deterministic law, it is impossible with classical statistical tools to find and, further, to reconstruct the deterministic structure from the data sequences and to obtain a simple law or algorithm.[7]

Referring to dynamical and structural instabilities James Crutchfield *et al.* stress that "the hope that physics could be complete with an increasingly detailed understanding of fundamental physical forces and constituents is unfounded. The interaction of components on one scale can lead to complex global behavior on a larger scale that in general cannot be deduced from knowledge of the individual components" [Crutchfield *et al.*, 1986, p. 56]. Instabilities limit the elimination of redundancies and the possibility to compress data sequences. Unstable processes cannot be reduced to laws governing the microscopic level of atoms or molecules. Nietzsche might have gone too far ahead when emphasizing, "There are no laws of nature!" [Nietzsche, 1930, p. 127], but he seems to be right in stressing that there are limits to finding laws. An unstable world is only partly accessible and knowable. Our non-knowledge regarding the behavior of complex systems has thus nothing to do with a temporary insufficiency of our knowledge; it has everything to do with unstable characteristics of complex systems.

3.6 Questioning: Positions in and Concepts of Philosophy of Science

Insofar as traditional concepts in philosophy of science refer at least to one of the above-discussed four points and draw from these their main arguments, instabilities may also challenge traditional concepts: if we acknowledge the prevalence of instabilities we are not only faced with four methodological issues *within* mathematical sciences themselves but, in addition, traditional concepts in philosophy of science are called into question. Questioning does not mean rejecting these concepts in general; however, a need for further clarification emerges (cf. [Mitchell, 2002; Harrell and Glymour, 2002; Schmidt, 2008a]).

Reproducibility: Experiments constitute a point of reference for many concepts in philosophy of science, such as *experimentalism*, *methodological constructivism*, and, to some extent, *pragmatism*. One precursor of modern experimentalism was Francis Bacon who advocated an intervention- and action-oriented notion of science: forced by experiments, nature should unveil her nomological secrets. Later,

[7] Of course, nonlinear techniques such as nonlinear data analysis, phase space reconstruction, surrogate data analysis, and other tools provide some options to find deterministic structures and separate them from white noise.

Kant supported Bacon and argued in favor of an experimentation-oriented constructivist's point of view: nature is experimentally constructed and empirically conceptualized by human reason in order to impose laws. To the extent that intervention- and action-oriented concepts of science rely on strong, quantitatively oriented views of reproducibility ("strong causation"), these concepts are scruinized by instabilities. Arguably, instabilities do not only challenge experimentation but also human action in an unstable environment in general. However, action and planning theorists have not yet perceived this challenge or tackled the problems involved.

Predictability: Many *instrumentalists* refer to the predictive power of science. The success of scientific predictions is the main argument to support their position. In addition to instrumentalism there are other concepts which draw their arguments from predictive success. For instance, in the tradition of Kant, *transcendental epistemologists* regard prediction as a key element of science which is seen as a consequence of the necessary condition of the possibility of knowledge ("linear determinism") — and, therefore, of the world's structure: the possibility of knowledge is transcendentally guaranteed by a deterministic world. According to this concept, prediction means learning from the past for the future. Prediction and time's arrow are twins; time is a necessary condition for the possibility of scientific experience. For *realists* and *empiricists* predictions play an important and necessary but secondary role in their understanding of science. Predictive success is considered as an indispensable criterion of justification and the basis for deciding the truth of propositions or theories. However, realists and empiricists regard predictability only as a means, whereas for *instrumentalists* the objective and summit of science is prediction itself. Certainly, instabilities challenge all concepts referring to predictive success.

Testability (Refutability and Confirmability): Most *realists* and, from a different angle, many *empiricists* draw their arguments from the empirical testability of theoretically derived propositions. Realists' and empiricists' concepts are based on stronger claims than the above-mentioned. To advocate their concepts both predictability and reproducibility are necessary. Realists and empiricists claim that the theoretical side (predictability) and the experimental side (reproducibility) must be joined together: knowledge is qualified as *scientific* knowledge insofar as theories refer to empirical propositions and can be tested. In this sense theories approximate the empirical laws or propositions in order to provide empirical evidence for the theories. Contrary to this, instabilities call into question the possibility that such a constant relation between theory and experiment always exists. According to Rueger and Sharp, "theories of nonlinear dynamics appear to generate a peculiar [...] problem for the realist [...]: in dealing with chaotic systems improved input for the theory will in general not lead to better predictions about the features relevant for confirmation" [Rueger and Sharp, 1996, p. 94].

Describability and Explainability: Conventionalists and scientific *rationalists* refer to the descriptive and explanatory power of science. They draw their main arguments from condensed descriptions and the economy of thought. The Hempel-

Oppenheim scheme is the preferred type of what characterizes a satisfactory explanation. In fact, instabilities challenge all concepts that refer to condensed descriptions. An unstable time evolution can hardly be represented by a simple law.

It would be a further task to elaborate and to distinguish the variety of issues raised by each of the four points — reproducibility, predictability, testability and describability/explainability — and to show how severely the traditional concepts of philosophy of science are affected by instabilities. Of course, this does not mean rejecting these concepts in general; but the need for further development and clarification will be obvious. We will not go into further details here, but continue by discussing the extent to which Nonlinear Dynamics has managed to deal with instabilities.

4 HOW TO DEAL WITH INSTABILITIES?

4.1 Rejection and Restriction: The Stability Dogma

The above-discussed methodological issues underpin that there was at least some rationale behind the *implicit* restriction of exact science to stability. Throughout the history of science, physicists seem to have had an inkling of the methodological problems provoked by instabilities — that is why they have mainly focused on stable objects. In addition to methodology, the restriction of physics to stability was also supported by the underlying stability metaphysics that was mostly present during the historical development of physics — from Plato to Newton and Laplace to Einstein. Stability metaphysics was indeed very beneficial: from a historical perspective it is evident that the stability assumption was crucial for the outstanding progress and explanatory success of physics. Unstable objects were simply excluded from physics. By this, physics was prevented from confusion and physicists from frustration.

Even though physicists seem to have had some idea of the methodological issues provoked by instabilities — a most striking example is the debate surrounding the hydrodynamics and Navier-Stokes equations in the 19^{th} century (cf. [Darrigol, 2006; Aubin and Dalmedico, 2002; Holmes, 2005]) — historically it took a long time from the first discovery to a broader acknowledgment of instabilities. In the early 20^{th} century, the challenge turned out to be pressing: how to deal with instabilities? One classic possibility was to renew what had proven to be a successful assumption throughout the history of physics. So the climax of the methodological challenges posed by instabilities, as well as the very last attempt to reject them and to restrict physics to stability, can be found in the works of Pierre Duhem and Alexandr Andronov. In *Aim and Structure of Physical Theory* (1908; cf. [Duhem, 1991]) Duhem refers to Poincaré and Hadamard. Although (and because!) Duhem was aware of instabilities, he believed physics had to rely on stability. He restricted physical methodology — in particular the deductive-nomological structure of explanation — explicitly to stability. According to Duhem's methodology, deductions are theory-based predictions in order to reveal empirically accessible consequences of a physical theory (cf. [Duhem 1991, pp. 55]). In order to be "useful for physics", deductions have to take experimental uncertainties into account. Duhem's bundle concept, a key element in his hypothetico-deductive approach, comes into play here. Specific types of non-diverging bundles reflect the experimental uncertainty of facts and guarantee the application of well-established mathematical tools, such as quantitative error estimation, error theory, statistical analysis, hypothesis testing theory and approximation theory. A diverging bundle corresponds to an unstable trajectory and for this trajectory "mathematical deduction is and always will remain useless to the physicist; [... it] will not permit us to predict what [will ...] happen" [Duhem, 1991, p. 139].[8] According to Duhem, when faced with the methodological issues of making deductions that will satisfy the objectives of science, the physicist is normatively "bound to impose [...] rigorous conditions" on bundles: the requirement for non-divergence or, which is the equivalent, for approximability. "To be useful to the physicist, it must still be proved that the second proposition remains approximately exact when the first is only approximately true" [Duhem, 1991, p. 143]. That is, deductions are restricted to stable trajectories in order to enable predictions and to guarantee the soundness of conclusions — this is a strong normative requirement, equivalent to non-divergence. Stability is a necessary condition for approximation: deductions — if they are considered "physically relevant" — *must* preserve the likeness of the neighborhoods within the bundles; the second bundle (proposition) must be structurally similar to the first in order to guarantee coping with experimental uncertainties. *Physical relevance* is equated with stability. The historian Jules Vuillemin summarizes that "Duhem limits his reflections to extrapolations concerning the stability of systems" [Vuillemin, 1991, p. xxviii]. Unstable phenomena are not regarded as being located in a branch of science that is called "physics".[9]

In the 1930s, the Russian physicist Alexandr Andronov was in accord with Duhem in restricting physics to stable objects. O*n the one hand*, Andronov was a pioneer in the early awareness of the possibility of instability. *On the other hand,* he believed, like Duhem, that physics was threatened by instabilities and, as a result of such considerations, he argued in particular for (structural) stability: if stability cannot be taken for granted, then it has to be imposed. As a consequence he formulated a stability requirement which Guckenheimer and Holmes later called "stability dogma" [Guckenheimer and Holmes, 1983, pp. 256]. Accord-

[8] "Deduction is sovereign mistress, and everything has to be ordered by the rules she imposes" [ibid., 266]. A "deduction is of no use to the physicist so long as it is limited to asserting that a given *rigorously true* proposition has for its consequence the *rigorous accuracy* of some other proposition" [Duhem, 1991, p. 143].

[9] Further, Prigogine stresses that Duhem was pioneering in his reflection on instabilities, but–according to Prigogine–Duhem went too far in his general critique concerning the ultimate uselessness of "unstable deductions". Duhem assessed instabilities in just a negative sense, as a threat to classical-modern physical methodology. "At this point we are decisively of a different opinion" [Prigogine and Stengers, 1990, p. 316].

ing to Andronov, stability has to be normatively imposed in physical methodology. He asked: "Which properties must a dynamical system (model) possess in order to cope with a physical problem?" [Andronov *et al.*, 1965/1969, p. 403]: the "system must be structurally stable". The "stability dogma" was imposed as a strong methodological norm. This dogmatic stability requirement served also as a selection criterion for "physically relevant" objects.

Duhem and Andronov are threshold figures indicating a time of transition. Whenever a dogmatization occurs, this can also be regarded as a signal of a crisis: what has until now been an unquestionable implicit belief turns out to be questionable and becomes a matter of dispute. What was formerly implicit is made explicit. In order to counteract the problems raised and to reject the methodological crisis, the very first attempt is always to *re*-introduce dogmatically what seems to have been lost.

4.2 Complexity Characteristics and the Qualitative in Nonlinear Dynamics

The Duhem-Andronov position is no longer convincing. In fact, we can observe a *Gestalt*-switch in the methodology of physics and in the view of nature. Some of the most pressing problems posed by instabilities have been resolved thanks to the advancement of Nonlinear Dynamics over the past 30 years.[10] The very first step was to get rid of the stability dogma. After disregarding and damning instability over hundreds of years, the stability dogma was replaced by an approach employing more qualitative features relating to those physical properties of a theory and of an experiment which are relevant for the specific situation in question — for example, specific characteristics that are often called *complexity characteristics*. "The definition of physical relevance will clearly depend upon the specific problem. This is quite different from the original statement [...] of the stability dogma" [Guckenheimer and Holmes, 1983, p. 259]. It is widely acknowledged throughout the methodological debate on instability that weaker requirements are *sufficient* for physical methodology and, also, that they are *necessary* to gain access to unstable phenomena.

The weaker requirements are based on qualitative rather than on quantitative properties. For instance, Benoit Mandelbrot calls his Fractal Geometry a "qualitative theory of different kinds of similarities" [Mandelbrot, 1991] and René Thom stresses that his Catastrophe Theory is a "qualitative science of Morphogenesis" [Thom, 1975]. Morris Hirsch argues here that "dynamics is used more as a source of *qualitative* insight than making *quantitative* predictions" [Hirsch, 1984, p. 11]. All of these approaches have a common background: the "qualitative theory of differential equations" which can be traced back to George Birkhoff [1927]; some fundamental ideas originate from the work of Henri Poincaré [1892] and in the mathematical theory of topology and differential topology. According to the ad-

[10] However, a proper methodology of the sort that works for stable objects is, and will always remain, impossible.

vocates of instabilities, the qualitative plays a major role in the post-stability methodology; it does not confound a mathematical approach at all.

Instabilities provoke a strong type of *underdetermination*. Physicists have a choice and, therefore, they *have to* choose; they have to decide which of the qualitative complexity characteristics of a model and of the corresponding experiment are to be seen as physically relevant, whereas the selection is more or less pragmatic and depends upon the specific physical problem and on the goal of the modeling issue, i.e., on the objective, interest, situation, context, resource, and the like. The choice of a particular characteristic is the starting point for any empirical test of any unstable model in the field of Nonlinear Dynamics. Nancy Cartwright does not speak of characteristics, but similarly of "capacities" (cf. [Cartwright, 1994]). Cartwright elaborates her capacity concept by referring to unstable dynamics, such as the motion of a 1000 Dollar-bill from a tower down to earth (cf. [Cartwright, 1983]). Today, an increasing amount of research is carried out on formulating, exploring and justifying physically relevant qualitative capacities or characteristics [Wackerbauer *et al.*, 1994]. Not every one of these is suited for *all* physical problems in general. In certain contexts, *some* complexity characteristics are useless or meaningless; but there will be *other* characteristics enabling a classification, a detailed description, a partial explanation or a piecewise prediction of the dynamics of a specific physical model. The variety of these contextual qualitative characteristics is striking — some of them deserve to be mentioned:[11] entropies, information theory parameters, Lyapunov exponents, scaling exponents, lengths of transients, fractal dimensions[12], renormalization theory parameters, topological characteristics, structure of stable and unstable periodic trajectories, existence of chaos or hyperchaos, parameters of basin boundaries, types of bifurcations, parameters of chaos control theory, power and Fourier spectra, phenomenological analysis of patterns, etc.[13]

It is worth noting that the qualitative is not at all related to subjective impressions or to *qualia*-like perceptions. On the contrary, the qualitative still remains mathematical, but with a different meaning. Physicists do not require models to be quantitatively stable, i.e., to show a quantitative robustness of the dynamics or structure, but to possess some complexity characteristics that are invariant under topological transformations, i.e., be topologically equivalent or conjugate. These qualitative characteristics refer to the appearance (the shape or "Gestalt") of the phenomenon and the geometry of the pattern *after* the process of time evolution — and not solely to the bare equation. They are not grounded on *single* trajectories

[11] For further details see textbooks on nonlinear dynamics, modeling complexity and nonlinear data analysis (cf. [Parker and Chua, 1989; Ott, 1993]).

[12] Such as capacity dimension, correlation dimension, information dimension and global embedding dimension, etc.

[13] Some of these properties are well known from solid state physics, fluid and plasma physics, meteorology, or statistical thermodynamics. What is new is (a) that these properties are necessary to empirically test (complex dynamical) models with just a few degrees of freedom and (b) that the model-theoretical discussion is deepened by the acceptance that nature can be structurally unstable.

and *single* measurable events but on *all possible* dynamics of the *whole* dynamical system that can be depicted in the phase space, for instance of a chaotic attractor. The global dynamics cannot be described by making use of classical quantitative distance measures. Instead, the dynamics is characterized by means of one of the most advanced mathematical theories — differential topology and dynamical systems theory.

Even though instabilities do exist for a particular dynamical system, some of the above listed complexity characteristics will persist under perturbations or small errors — e.g., the instability itself — while others will not and vary dramatically. Persistence is a necessary requirement for any empirical test and for physical relevance. A sufficient requirement is what physicists call: *prevalence of a certain characteristic in the class of models.* Physicists do not require models to be stable but to possess some *prevalent* complexity characteristics. These characteristics constitute the basis for any empirical test of a physical model; they also serve as a foundation of the context of justification (cf. [Hooker, 2004]).

4.3 Nonlinear Data Analysis and the Shadowing Lemma — a New Kind of Calculability

When physicists test a model, they need to compare the theoretically deduced model output with the data measured in the real physical system. From a methodological perspective, *nonlinear data analysis* plays a central role in matching a model's data and the real experimental system's data (cf. [Takens 1985; Sauer *et al.*, 1991; Abarbanel *et al.*, 1993]).[14] This new tool to analyze data sequences from unstable dynamics has been developed in Nonlinear Dynamics since the 1980s. Nonlinear data analysis provides an excellent tool for the calculation of the above listed complexity characteristics from experimental time series. These empirically gained characteristics are confronted with theoretically deduced characteristics generated by the model. If (a) the model characteristic is proven to be prevalent in the model class and if (b) the empirically measured and the model-generated data — the qualitative complexity characteristics — match each other within error tolerances, physicists will accept the model as an adequate description of the empirical phenomenon.

The shift in physical methodology from a quantitative to a qualitative orientation is also exemplified with regard to calculability and predictability (cf. [Hirsch 1984; Jackson, 1989]). As explicated above, instabilities restrict the horizon of single trajectory oriented predictions. In order to deal with this problem, predictability is addressed explicitly and has become a major research topic in Nonlinear Dynamics (cf. [Aurell *et al.*, 1997; Badii, 1991]). During the last 30 years, nonlinear scientists have developed qualitative calculation tools. The most important of these is the "shadowing lemma", which addresses issues of calculability when a quantitative approach is impossible [Bowen, 1970; 1978; Coven *et al.*,

[14]Of course, other methods have been developed, for example comparing the structure of bifurcations, of stable and unstable periodic trajectories.

1988; Jackson, 1989; Peitgen *et al.*, 1992]. Besides "nonlinear data analysis", the "shadowing lemma" is the second mathematical tool available to deal with general problems in the classical quantitative approach. The issue the shadowing lemma addresses is: given a differential equation which describes an unstable dynamics, e.g., a chaotic attractor, there is no way of calculating and predicting the *true* trajectory. However, the shadowing lemma provides an encouraging way to deal with this situation. In many cases, shadow trajectories that can effectively be calculated do exist; they remain in the neighborhood of the non-predictable, unstable true orbit — although they cannot be derived as any simple approximation to that orbit. So, if we are not interested in an exact (quantitative) prediction, we can describe and calculate the global dynamics more or less in a qualitative way. The shadowing trajectories provide qualitatively similar (topological) properties that characterize the dynamical systems. The assumption required for the existence of such shadow trajectories is that the system is hyperbolic. How serious this restriction to hyperbolicity might be is still a matter of ongoing research.

4.4 Modeling Complex Systems and the Modeling Turn in Physics

Instabilities entail a model orientation and give rise to different kinds of explanations (cf. [Parker and Chua, 1989]). Traditionally, reductive explanations and condensed descriptions have been, and still are, associated with universal laws and fundamental theories. As explicated in the preceding section, instabilities put limitations on reductive explanations. This is why physicists, when accounting for unstable objects, hesitate to portray their work as discovering new theories or new laws of nature. They prefer to talk of "models" rather than "theories" or "laws" — regarding their approach as *modeling* systems, *model* construction and *model*-based explanations. The shift in terminology is not just rhetoric; rather, it indicates a *modeling turn* in physics.[15] In standard philosophy of science, models play, if any, a role of minor importance, e.g., as mere derivatives of universal theories in order to get in touch with empirical data (exceptions, overview see: [Hartmann and Frigg, 2006]). Contrary to the standard position, René Thom proposes an "Outline of a General Theory of Models" [Thom, 1975, p. 325]. Philosophers of science, such as Hans Poser, go further and identify "new general types of mathematical sciences" based on "model building and simulations" [Poser, 2001, p. 284]. And Stephen Kellert argues that Nonlinear Dynamics' "method of understanding" is nothing but "constructing, elaborating, and applying simple dynamical models" and "interpreting" them in different domains of application (cf. [Kellert, 1994, p. 85]). This type of explanation differs strongly from the deductive-nomological account of explanation.

Here we identify a *contextualism* which is based on complexity characteristics: in the realm of unstable objects, models can only be tested and justified for a certain context of application, also regarding modeling objectives and goals, but not for

[15] Cf. [Hartmann, 2009] and [Hartmann and Frigg, 2006] for an impressive introduction to the recent discussions about the term "model".

the "whole world". Therefore, it is common to discuss issues such as whether a certain "mathematical representation or model of the system [... is] a good one" or not, from a *pragmatist* perspective [Batterman, 2002, p. 57]. The criteria for a "good representation" of a model are based on the requirement that the model must provide some prevalent characteristics — and not just generic ones. It is interesting to note that an epistemological and methodological discussion on the structure of mathematical models is taking place *within* physics today, as part of the progress of physics itself. Well-established ideas of philosophy of science are not an add-on but can be considered as an intrinsic element of physics. Instability and complexity convey philosophical thoughts such as the reflection on the criteria for scientific evidence, truth or objectivity into the very heart of physics.

In the realm of instabilities, model-oriented explanations remain contextual in the sense that they are limited to a specific domain of application — for example, the impact oscillator model (see below). In addition, there is no clear and unique justification, but various ones. One of the interesting aspects in comparing traditional philosophical accounts of explanation with the understanding provided by Nonlinear Dynamics also results from a mismatch between the item to be explained in the former and the item to be understood in the latter. While, traditionally, philosophers commonly address scientific explanations of such things as "facts", or "events", Nonlinear Dynamics usually studies things such as behaviors, patterns, or bifurcations (cf. [Kellert, 1994, pp. 81f]). A shift in the focus of interest, and hence in what deserves an explanation, can be observed. According to Bergé *et al.* "the ultimate goal is to understand the origin and characteristics of all kinds of time-evolution encountered, including those which at first seem totally disorganized" [Bergé *et al.*, 1984, p. 102]. The type of explanation required and appreciated in Nonlinear Dynamics therefore differs from traditional physics: Nonlinear Dynamics favors a type of explanation "that is holistic, historical, and qualitative, eschewing deductive systems and causal mechanisms and laws" ([Kellert, 1994, 114], cf. [Mitchell, 2002; Schmidt, 2008a]).

4.5 "Context of Interest" and What is Worth Knowing . . .

Thus, Nonlinear Dynamics induces a shift in the focus of attention and in knowledge interests. The *context of interest* precedes both the context of discovery and the context of justification — a point which is mostly disregarded by standard positions in philosophy of science. Most important is that the context of interest shapes any research program, for example in the way problems are constituted and objects of scientific inquiry are framed. Decisions about relevance are, explicitly or implicitly, made: which objects and what phenomena deserve scientific research? What is of interest? What do we wish to know? What is a relevant and significant problem? Nonlinear Dynamics opens avenues to reflect explicitly on the context of interest and the purpose of knowledge. Not only reflections, but decision making processes are indispensable. According to René Thom, a "choice [. . .] of what is considered scientifically interesting [has to be made]. [. . . This] is certainly to a large extent arbitrary. [. . . Traditionally,] [p]hysics uses enormous machines to investigate situations that exist for less than 10^{-23} seconds [...]. But we can at least ask one question: many phenomena of common experience, in themselves trivial [...] – for example, the cracks in an old wall, the shape of a cloud, the path of a falling leaf, or the froth on a pint of beer – are very difficult to formalize, but is it not possible that a mathematical theory launched for such homely phenomena might, in the end, be more profitable for science?" [Thom, 1975, p. 9] Thom addresses these "phenomena of common experience" which have been "generally neglected" throughout the development of modern sciences since the 16^{th} century [ibid., p. 8]. The underlying reason for the neglect and disregard is clearly that "these phenomena are highly unstable, difficult to repeat, and hard to fit into a [universal] mathematical theory [...]" [ibid., p. 9]. The objective of Thom's catastrophe theory is to find a remedy. Thom tries to widen the focus and to readjust the context of interest in order to attract attention to phenomena of common experience — phenomena that are mostly unstable, dynamical, and complex. These phenomena are at home in the familiar mesocosm of our life-world — not just placed in the tiny microcosm or the huge macrocosm; we can observe them with our unaided senses. In order to approach these unstable phenomena in the mesocosm, Thom facilitates a mathematically-based morphology — a "general theory of morphogenesis" including the dynamics of forms and the emergence of pattern (cf. [ibid., pp. 101/124f]).

There are cognate *mesocosmic* approaches in Benoit Mandelbrot's "Fractal Geometry" (cf. [Mandelbrot, 1991], in Alfred Gierer's "physics of biological *Gestalt* shaping and structure formation" [Gierer, 1981] and in Hans Meinhardt's "pattern dynamics of sea shells" [Meinhardt, 1996]. For instance, Meinhardt shows that "the underlying mechanism that generates the beauty [of sea shells] is eminently dynamic. [...] A certain point on the shell represents a certain moment in its history" [Meinhardt, 1996, p. vii]. The dynamics of "sea shell pattern are not explicable on the basis of the elementary mechanisms in a straight forward manner" [ibid.]. Due to instabilities, no quantitative explanation is feasible, only a qualitative description which is based on computer simulations and graphical-visual representations. The historical gap between the development of these tools in the 1980s and 1990s and the recognition that there was a need for them by Poincaré and others 80 years before can probably be attributed to the absence of adequate computational power to analyze the unstable systems. The development of computer technology has opened up new opportunities for theorizing and experimentation that enable dealing with unstable systems, giving rise to more qualitative approaches. (a) *Theorizing:* The naked model equations do not say anything — the phenomena are mostly hidden and cannot be foreseen. The phenomena have to be generated by numerical simulations; in addition, they are, often, represented by visualization techniques. Simulations have become an indispensable part of any theory describing unstable systems. (b) *Experimentation:* Computers not only provide new ways of dealing with theoretical problems of unstable systems — most impressive is the emergence of a *new experimental culture*;

today, it is common to perform numerical or computer experiments [Ulam, 1952; Phillips, 1956; Lenhard, 2007]. So, in addition to the tremendous progress on a pragmatic level, numerical experiments and computer simulations indicate a shift in the principles of methodology. Hans Poser identifies "new criteria of scientific methodology" [Poser, 2001, p. 282]; according to Bernd-Olaf Küppers, numerical experimentation with unstable systems implies a change of "the inner constitution of mathematical sciences" [Küppers, 1992, p. 9].

To summarize: Although the limitations of stability-oriented physical methodology cannot be overthrown, Nonlinear Dynamics has opened up new ways to deal with instabilities. The classical methodological nucleus is being modified — in particular by referring to more qualitative instead of quantitative aspects. Most striking is that phenomena, patterns, and structures within the mesocosm attract attention.

4.6 Examples

The shift in scientific methodology — e.g., from the quantitative to the qualitative, from the stability dogma to prevalent characteristics — shall be illustrated by a few examples. For the sake of simplicity, the examples are taken from classical physics but nevertheless they characterize Nonlinear Dynamics in general. In addition, the shift in scientific methodology has an impact on modern and recent physics as well as on all other sciences making use of mathematics to design and to test models.

Duffing oscillator. In 1918, Georg Duffing introduced a nonlinear oscillator model with a cubic stiffness term in the standard differential equation of an idealized linear harmonic oscillator to describe the hardening spring effect observed in many engineering and physical problems (cf. [Duffing, 1918]). For a specific model of the model class, i.e., for specific parameters, Duffing's equation with external sinusoidal driving is structurally and dynamically unstable and hence the above discussed methodological difficulties occur. But Duffing and others were confident that the model is a realistic qualitative model of the real physical oscillator (cf. [Wiggins, 1990, pp. 29/153]). The essential qualitative properties are the existence and structure of the three equilibrium points of the spring (two of them are point attractors) and the dynamical pattern of the chaotic attractor. In addition, complexity characteristics such as the Lyapunov exponents, different kinds of fractal dimensions, and the structure of bifurcations provide a qualitative justification for Duffing's model (cf. [Moon, 1992, pp. 5.]).

Railway vehicle dynamics. Based on the partial knowledge of some physical laws, i.e., the conventional bogie model (Cooperrider, Vermeulen and Johnson), Hans True *et al.* have designed a more complex model of a moving railway car (cf. [Isaksen and True, 1997]). The essential prevalent characteristics to empirically test the model are: the oscillation frequency of the railway car in the model for low velocities, the critical velocity, supercritical bifurcations (Hopf bifurcation), chaos and the types of instability (reverse period doubling).

Impact oscillators. These kinds of oscillators consist of an oscillating mass that impacts with a fixed boundary when its amplitude is large enough. In between these impacts the dynamic is smooth and often linear but it derives a strong nonlinearity from the presence of the impact. Impact oscillators are very common in applied physics and engineering. S. Bishop worked on a specific impact oscillator and on model-theoretical problems (cf. [Bishop *et al.*, 1996]). According to Bishop *et al.*, sufficient empirical criteria to test the model of this impact oscillator are characteristics such as the type of bifurcation ("grazing bifurcation"), the existence of chaos, and the Fourier and Power spectra. Other characteristics cannot adequately be applied in this situation.

Nonlinear electronic circuit. An electronic circuit with a specific nonlinear diode is modeled by Leon Chua ("Chua-Circuit", cf. [Chua *et al.*, 1986]). Of course, physicists know much about physical details and partial laws — and indeed, this is an excellent basis to design a physical model. But how can physicists assure that the diode and the whole circuit is modeled correctly? This main model-theoretical question can only be answered in reference to the real physical system, i.e., by making use of measurements. K. Murali and M. Lakshmanan choose characteristics such as the existence of chaos (positive Lyapunov exponent), a reverse sequence of period adding, the structure of periodic windows, hysteresis effects, intermittency and the coexistence of attractors in order to test their model with experimental data (cf. [Murali and Lakshmanan, 1990]). Working with another type of the Chua-Circuit, G. Gunaratne *et al.* select specific pattern in phase space, such as the Arnold tongues, as the best characteristics to test the physical model (cf. [Gunaratne *et al.*, 1989]).indexreproducibility

5 CONCLUSION AND PROSPECTS

Since the advent and advancement of Nonlinear Dynamics in the late 1960s, we can observe a structural-paradigmatic shift in physical sciences. Instabilities are regarded as the sources of complexity, self-organization and pattern formation (cf. [Langer, 1980]). They form the unifying core of all Nonlinear Sciences that gives rise to this *Gestalt*-switch.

Instabilities also challenge the methodological nucleus of sciences: predictability, reproducibility, testability, and explainability. (a) When instabilities are present, the usefulness of a theory or a law for *predictions* is effectively limited; the principle of superposition does not hold, and — in most cases — the equations cannot be integrated and approximated by using analytic tools of mathematics. (b) Insofar as *reproducibility* is the core of an experiment, the experimental method of physics is challenged; there is no possibility to decompose unstable processes into parts and to isolate them from the environment. (c) Even worse is the attempt to bring together both theoretical and experimental events (trajectories, data, states). The quantitative *testability* of laws is limited. (d) In addition, unstable processes cannot be compressed to a simple law; *explainability* is limited.

In order to cope with these challenges, physicists have changed their perspective from a *quantitative*-metrical to a more *qualitative*-topological approach. The traditional quantitative-oriented *stability dogma* has been replaced by an approach employing more qualitative features and relating to those physical properties that are relevant for the specific situation in question — a pragmatic non-universalism or contextualism is emerging. These changes are advocates for strong positions such as a "non-reductive physicalism" [Schmidt, 2008a] and an "integrative pluralism" [Mitchell, 2002]. According to Nancy Cartwright, "we live in a world rich in different things, with different natures, behaving in different ways. The laws that describe this world are a patchwork, not a pyramid" [Cartwright, 1999, p. 1].

BIBLIOGRAPHY

[Abarbanel *et al.*, 1993] H. D. Abarbanel, R. Brown, J. J. Sidorowich, and L. S. Tsimring. The Analysis of Observed Chaotic Data in Physical Systems. *Rev. Mod. Phys.*, 65, 1331-1392, 1993.

[Abarbanel, 1996] H. D. Abarbanel. *Analysis of observed chaotic data.* Springer, New York/Berlin, 1996.

[Andronov and Pontryagin, 1937] A. Andronov and L. Pontryagin. Systèmes Grossiers. *Dokl. Akad. Nauk.* (Doklady) SSSR 14, 247-251, 1937.

[Andronow *et al.*, 1965/1969] A. Andronow, A. A. Witt, and S. E. Chaikin. *Theorie der Schwingungen,* Part I + II, Akademie Verlag, Berlin, 1965, 1969.

[Atmanspacher *et al.*, 1992] H. Atmanspacher, J. Kurths, H. Scheingraeber, R. Wackerbauer, and A. Witt. Complexity and meaning in nonlinear dynamical systems. *Open Syst. Inform. Dyn.* 1(2), 269-289, 1992.

[Aubin and Dalmedico, 2002] D. Aubin and A. D. Dalmedico. Writing the History of Dynamical Systems and Chaos: Longue Durée and Revolution, Disciplines and Cultures. *Historia Mathematica*, 29, 273-339, 2002.

[Aurell *et al.*, 1997] E. Aurell, G. Boffetta, A. Crisanti, G. Paladin, and A. Vulpiani. Predictability in the large: An extension of the concept of Lyapunov exponent. *J. Phys. A*, 30, 1-26, 1997.

[Badii, 1991] R. Badii. Quantitative characterization of complexity and predictability. *Phys. Lett. A*, 160, 372-377, 1991.

[Banks *et al.*, 1992] J. Banks, J. Brooks, G. Cairns, G. Davis, and P. Stacey. On Devaney's definition of chaos. *Amer. Math. Monthly*, 99, 332-334, 1992.

[Batterman, 2002] R. Batterman. *The Devil in the Detail: Asymptotic Reasoning in Explanation, Reduction and Emergence.* Oxford Uni Press, Oxford/New York, 2002.

[Bérge *et al.*, 1984] P. Bérge, Y. Pomeau, and C. Vidal. *Order within Chaos. Towards a Deterministic Approach to Turbulence.* Wiley & Sons, New York, 1984.

[Birkhoff, 1927] G. D. Birkhoff. *Dynamical Systems.* AMS Colloquium Publications, New York, 1927.

[Birkhoff and Lewis, 1935] G. D. Birkhoff and D. C. Lewis. Stability in Causal Systems. *Philosophy of Science*, 2(3), 304-333, 1935.

[Bishop *et al.*, 1996] S. R. Bishop, M. G. Thompson, and S. Foale. Prediction of period-1 impacts in a driven beam. *Proc. Roy. Soc. Lond. A*, 452, 2579-2592, 1996.

[Böhme and van den Daele, 1977] G. Böhme and W. van den Daele. Erfahrung als Programm – Über Strukturen vorparadigmatischer Wissenschaft. In: G. Böhme, W. van den Daele, W. Krohn (eds.). *Experimentelle Philosophie.* Suhrkamp, Frankfurt, 166-219, 1977.

[Bowen, 1970] R. Bowen. Markov partitions for Axiom A diffeomorphisms. *Amer. J. Math.*, 92, 725-747, 1970.

[Bowen, 1978] R. Bowen. On Axiom A Diffeomorphisms. *CBMS Regional Conference Series in Mathematics*, 35, Providence RI, 1978.

[Brown and Chua, 1996] R. Brown and L. O. Chua. Clarifying chaos: Examples and counterexamples. *Int. J. Bif. Chaos*, 6(2), 219-249, 1996.

[Brown and Chua, 1998] R. Brown and L. O. Chua. Clarifying Chaos II: Bernoulli chaos, zero Lyapunov exponents and strange attractors. *Int. J. Bif. Chaos*, 8(1), 1-32, 1998.

[Bunge, 1987] M. Bunge. *Kausalitaet, Geschichte und Probleme*. Mohr, Tübingen, 1987.

[Carnap, 1928] R. Carnap. *Der logische Aufbau der Welt*. Weltkreisverlag, Berlin, 1928.

[Cartwright, 1983] N. Cartwright. *How the laws of physics lie*. Oxford University Press, Oxford, 1983.

[Cartwright, 1994] N. Cartwright. *Nature's Capacities and their Measurement*. Oxford University Press, Oxford, 1994.

[Cartwright, 1999] N. Cartwright. *The dappled world: a study of the boundaries of science*. Cambridge University Press, Cambridge, 1999.

[Chaitin, 1971] G. J. Chaitin. Computational Complexity and Gödel's Incompleteness Theorem. *ACM SIGACT News*, 9, 11-12, 1971.

[Chaitin, 1987] G. J. Chaitin. *Algorithmic Information Theory*. Cambridge University Press, Cambridge, 1987.

[Chaitin, 2001] G. J. Chaitin. *Exploring Randomness*. Springer, London, 2001.

[Chua *et al.*, 1986] L. O. Chua, Y. Yao, and Q. Yang. Devil's staircaise route to chaos in a nonlinear circuit. *Int. J. Circuit Appl.*, 14, 315-329, 1986.

[Comte, 2006] A. Comte. Cours de Philosophie Positive. Bachelier, Paris, (1830-1842) 2006.

[Coven *et al.*, 1988] E. Coven, I. Kan, and J. A. Yorke. Pseudo-orbit shadowing in the family of tent maps. *Trans. Amer. Math. Soc.*, 308(1), 227-241, 1988.

[Crutchfield *et al.*, 1986] J. P. Crutchfield, J. D. Farmer, N. H. Packard, and R. S. Shaw. Chaos. *Scientific American*, 12, 46-57, 1986.

[Darrigol, 2006] O. Darrigol. *Worlds of Flow. A History of Hydrodynamics from Bernoullis to Prandtl*. Oxford University Press, Oxford, 2006.

[Descartes, 1979] R. Descartes. *Regeln zur Ausrichtung der Erkenntniskraft*. Felix Meixner, Hamburg, 1979.

[Devaney, 1987] R. L. Devaney. *An Introduction to Chaotic Dynamical Systems*. Addison Wesley, Redwood City Cal., 1987.

[Drieschner, 2002] M. Drieschner. *Moderne Naturphilosophie. Eine Einführung*. Mentis, Paderborn, 2002.

[Duffing, 1918] G. Duffing. *Erzwungene Schwingungen bei veraenderlicher Eigenfrequenz und ihre technische Bedeutung*. Vieweg Verlag, Braunschweig, 1918.

[Duhem, 1991] P. Duhem. *The Aim and Structure of Physical Theory*. Princeton University Press, Princeton, 1991 (1906).

[Ebeling and Feistel, 1990] W. Ebeling and R. Feistel. *Physik der Evolutionsprozesse*. Akademie Verlag, Berlin, 1990.

[Einstein, 1917] A. Einstein. Kosmologische Betrachtungen zur allgemeinen Relativitaetstheorie. *Sitzungsberichte der Koeniglich Preussischen Akademie der Wissenschaften, physikalisch-mathematische Klasse*, 142-152, 1917.

[Einstein *et al.*, 1935] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47, 777-780, 1935.

[Feigenbaum, 1978] M. J. Feigenbaum. Quantitative universality for a class of nonlinear transformations. *J. Stat. Phys.*, 19(1), 25-52, 1978.

[Gierer, 1981] A. Gierer. Physik der biologischen Gestaltbildung. *Naturwiss.* 68 (1981), 245-251, 1981.

[Gleick, 1987] J. Gleick. *Chaos: Making a New Science*. Viking Penguin, New York, 1987.

[Guckenheimer and Holmes, 1983] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, New York, 1983.

[Gunaratne *et al.*, 1989] G. H. Gunaratne, P. S. Linsay, and M. J. Vinson. Chaos beyond Onset: A Comparison of Theory and Experiment. *Phys. Rev. Lett.* 63(1), 1-4, 1989.

[Hacking, 1983] I. Hacking. *Representing and intervening. Introductory topics in the philosophy of natural sciences*. Cambridge University Press, Cambridge, 1989.

[Haken, 1977] H. Haken (ed.). *Synergetics. A Workshop*. Springer, Berlin, 1977.

[Haken, 1980] H. Haken. *Dynamics of Synergetic Systems*. Springer, Berlin, 1980.

[Harrell and Glymour, 2002] M. Harrell and C. Glymour. Confirmation and Chaos. *Philosophy of Science*, 69, 256-265, 2002.

[Hartmann and Frigg, 2006] S. Hartmann and R. Frigg. Models in Science. In: E. N. Zalta (ed.). *Stanford Encyclopedia of Philosophy*. Standford University Press, Standford, 2006.

[Hartmann, 2009] S. Hartmann. Modelle. In: H.-J. Sandkuehler. *Enzyklopaedie der Philosophie*, 2^{nd} ed. Meiner, Hamburg, 2009 (in press).

[Heidegger, 1986] M. Heidegger. *Nietzsches metaphysische Grundstellung im abendlaendischen Denken - die ewige Wiederkehr des Gleichen*; Gesamtausgabe. Vittorio Klostermann, Frankfurt, 1986.

[Hempel, 1965] C. G. Hempel. *Aspects of Scientific Explanation*. Free Press, New York, 1965.

[Hertz, 1963] H. Hertz. *Die Prinzipien der Mechanik. In neuem Zusammenhang dargestellt.* Wiss. Buchgesellschaft, Darmstadt, (1894) 1963.

[Hirsch, 1984] M. Hirsch. The Dynamical Systems Approach to Differential Equations. *Bull. Amer. Math. Soc.*, 11, 1-64, 1984.

[Hooker, 2004] C. A. Hooker. Asymptotics, Reduction and Emergence. *Brit. J. Phil. Sci.*, 55(3), 435-479, 2004.

[Holmes, 2005] P. Holmes. Ninety plus thirty years of nonlinear dynamics: More is different and less is more. *Int. J. Bif. Chaos*, 15 (9), 2703-2716, 2005.

[Hume, 1990] D. Hume. *Enquiries concerning human understanding.* Clarendon Press, Oxford, (1748) 1990.

[Hund, 1972] F. Hund. *Geschichte der physikalischen Begriffe*. BI, Mannheim, 1972.

[Isaksen and True, 1997] P. Isaksen and H. True. On the ultimate transition to chaos in the dynamics of Cooperrider's bogie. *Chaos, Solitions, Fractals*, 8(4), 559-581, 1997.

[Jackson, 1989] E. A. Jackson. *Perspectives of nonlinear dynamics.* Vol. 1, Cambridge University Press, Cambridge, 1989.

[Janich, 1997] P. Janich. *Kleine Philosophie der Naturwissenschaften.* Beck, Munich, 1997.

[Kellert, 1994] S. Kellert. *In the Wake of Chaos: Unpredictable Order in Dynamical Systems.* University of Chicago Press, Chicago, 1994.

[Krohn and Küppers, 1992] W. Krohn and G. Küppers (ed.). *Selbstorganisation. Aspekte einer wissenschaftlichen Revolution.* Vieweg, Braunschweig/Wiesbaden, 1992.

[Küppers, 1992] B.-O. Küppers. *Natur als Organismus. Schellings frühe Naturphilosophie und ihre Bedeutung für die moderne Biologie.* Suhrkamp, Frankfurt, 1992.

[Langer, 1980] J. S. Langer. Instabilities and Pattern Formation. *Rev. Mod. Phys.*, 52, 1-28, 1980.

[Lenhard, 2007] J. Lenhard. Computer Simulations: The Cooperation Between Experimenting and Modeling. *Philosophy of Science*, 74, 176-194, 2007.

[Li and Yorke, 1975] T.-Y. Li and J. A. Yorke. Period three implies chaos. *Amer. Math. Monthly*, 82(10), 985-992, 1975.

[Lorenz, 1963] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20, 130-141, 1963

[Lorenz, 1989] E. N. Lorenz. Computational chaos – A prelude to computational instability. *Physica D*, 35, 299-317, 1989.

[Mach, 1988] E. Mach. *Die Mechanik in ihrer Entwicklung.* Akademie Verlag, Leipzig/Darmstadt, 1988 (1883).

[Mainzer, 1996] K. Mainzer. *Thinking in complexity. The complex dynamics of matter, mind, and mankind.* Springer, Heidelberg/New York, 1996.

[Mainzer, 2005] K. Mainzer. *Symmetry and Complexity. The Spirit and Beauty of Nonlinear Science.* World Scientific Publisher, Singapore/New York, 2005.

[Mandelbrot, 1991] B. Mandelbrot. *Die fraktale Geometrie der Natur.* Birkhauser, Basel, 1991 (1977).

[Maxwell, 1873] J. C. Maxwell. Does the progress of Physical Science tend to give any advantage to the opinion of Necessity (or Determinism) over that the Contingency of Events and the Freedom of the Will? In L. Campbell, W. Garnett (ed.). *The Life of James Clerk Maxwell.* Johnson, New York, 434-444, 1989.

[Maxwell, 1991] J. C. Maxwell. *Matter and motion.* Dover Publications, New York, (1877) 1991.

[May, 1976] R. May. Simple Mathematical Models with very complicated Dynamics. *Nature*, 261, 459-467, 1976.

[Meinhardt, 1995] H. Meinhardt. *The algorithmic beauty of seashells.* Springer, Berlin, 1995.

[Mitchell, 2002] S. D. Mitchell. Integrative Pluralism. *Biology and Philosophy*, 17, 55-70, 2002.

[Mittelstraß, 1998] J. Mittelstraß. *Die Haeuser des Wissens.* Suhrkamp, Frankfurt, 1998.

[Moon, 1992] F. C. Moon. *Chaotic and fractal dynamics.* Wiley & Sons, New York, 1992.

[Moon, 1998] F. C. Moon (ed.). *Dynamics and Chaos in Manufacturing Processes.* Wiley & Sons, New York, 1998.

[Murali and Lakshmana, 1990] K. Murali and M. Lakshmana. Observation of many bifurcation sequences in a driven piecewise-linear circuit. *Phys. Lett. A*, 151(8), 412-419, 1990.

[Newton, 1730] I. Newton. *Opticks: or, a Treatise of the Reflections, Refractions, Inflections and Colours of Light.* Innys, New York, (1717) 1730.

[Nicolis and Prigogine, 1977] G. Nicolis and I. Prigogine. *Self-Organization in Nonequilibrium Systems. From Dissipative Structures to Order through Fluctuations.* Wiley & Sons, New York/London, 1977.

[Nietzsche, 1930] F. Nietzsche. Die *froehliche Wissenschaft.* Kroener, Leipzig, (1887) 1930.

[Parker and Chua, 1989] T. S. Parker, L. O. Chua. *Practical numerical algorithms for chaotic Systems.* Springer, New York, 1989.

[Pauli, 1961] W. Pauli. *Aufsaetze und Vortraege über Physik und Erkenntnistheorie.* Vieweg, Braunschweig, 1961.

[Peitgen *et al.*, 1992] H.-O. Peitgen, H. Jürgens, and D. Saupe. *Bausteine des Chaos – Fraktale.* Springer/Klett-Cotta, Berlin/Stuttgart, 1992.

[Pietschmann, 1996] H. Pietschmann. *Phaenomenologie der Naturwissenschaft.* Springer, Berlin, 1996.

[Phillips, 1956] N. Phillips. The General circulation of the Atmosphere: a Numerical Experiment. *Quat. J. R. Met. Soc.*, 82(352), 123-164, 1956.

[Poincaré, 1892] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste.* Gauthier-Villars et fils, Paris, 1892.

[Poincaré, 1914] H. Poincaré. *Wissenschaft und Methode.* Teubner, Leipzig, (1908) 1914.

[Popper, 1934] K. R. Popper. *Logik der Forschung.* Julius Springer, Wien (English translation: The Logic of Scientific Discovery. Routledge, London), 1934.

[Poser, 2001] H. Poser. *Wissenschaftstheorie.* Reclam, Stuttgart, 2001.

[Prigogine, 1992] I. Prigogine. *Vom Sein zum Werden.* Piper, Munich, (1979) 1992.

[Prigogine and Glansdorff, 1971] I. Prigogine and P. Glansdorff. *Thermodynamic Theory of Structure, Stability and Fluctuation.* Wiley & Sons, New York/London, 1971.

[Psillos, 1999] S. Psillos. *Scientific Realism. How Science tracks Truth.* Routledge, London/New York, 1999.

[Redhead, 1980] M. Redhead. Models in Physics. *Brit. J. Phil. Sci.*, 31, 145-163, 1980.

[Rueger and Sharp, 1996] A. Rueger and W. D. Sharp. Simple Theories of a Messy World: Truth and Explanatory Power in Nonlinear Dynamics. *Brit. J. Phil. Sci.*, 47, 93-112, 1996.

[Ruelle, 1989] D. Ruelle. *Elements of differentiable dynamics and bifurcation theory.* Academic Press, London, 1989.

[Ruelle and Takens, 1971] D. Ruelle, F. Takens. On the nature of turbulence. *Commun. Math. Phys.*, 20, 167-192, 1971.

[Salmon, 1989] W. Salmon. Four Decades of Scientific Explanation. In: P. Kitcher, W. Salmon (ed.). *Scientific Explanation.* University of Minnesota Press, Minnesota, 3-219, 1989.

[Sauer *et al.*, 1991] T. Sauer and J. A. Yorke, M. Casdagli. Embedology. *J. Stat. Phys.*, 77(3/4), 579-616, 1991.

[Schmidt, 2003] J. C. Schmidt. Zwischen Berechenbarkeit und Nichtberechenbarkeit. Die Thematisierung der Berechenbarkeit in der aktuellen Physik komplexer Systeme. *J. Gen. Phil. Sci.*, 34, 99-131, 2003.

[Schmidt, 2005] J. C. Schmidt. Zufall (mathematisch, physikalisch, biologisch). *Historisches Woerterbuch der Philosophie* 12, 1419-1424, 2005.

[Schmidt, 2008a] J. C. Schmidt. From Symmetry to Complexity: On Instabilities and the Unity in Diversity in Nonlinear Science. *Int. J. Bif. Chaos*, 18(4), 897-910, 2008.

[Schmidt, 2008b] J. C. Schmidt. *Instabilitaet in Natur und Wissenschaft. Eine Wissenschaftsphilosophie der nachmodernen Physik.* De Gruyter, Berlin, 2008.

[Shaw, 1981] R. Shaw. Modeling Chaotic Systems. In: H. Haken (ed.). *Chaos and Order in Nature.* Springer, New York, 218-231, 1981.

[Swinney and Gollub, 1981] H. L. Swinney and J. P. Gollub, (eds.) *Hydrodynamic Instabilities and the Transition to Turbulence.* Springer, Berlin, 1981.

[Takens, 1985] F. Takens. Distinguishing deterministic and random systems. In G. I. Barenblatt, G. Ioss, D. Joseph (eds.). *Nonlinear dynamics and turbulence*; Pitman, Boston, 314–333, 1985.

[Thom, 1975] R. Thom. *Structural Stability and Morphogenesis. An Outline of a General Theory of Models.* Benjamin, Reading/MA, 1975.

[Ulam, 1952] S. Ulam. Random Processes and Transformations. *Proc.* 1950 *Intl. Congress* on *Mathematics, American Mathematical Society* 2, Providence RI, 264-275, 1952.

[Vuillemin, 1991] J. Vuillemin. Introduction. In: P. Duhem. *The Aim and Structure of Physical Theory.* Princeton University Press, Princeton, xv-xxxiii, (1906) 1991.

[Wackerbauer *et al.*, 1994] R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurths, and H. Scheingraber. A Comparative Classification of Complexity Measures. *Chaos, Solitions & Fractals*, 4(1), 133-174.

[Weizsaecker, 1974] C. F. v. Weizsäcker. *Die Einheit der Natur.* dtv, Munich, 1974.

[Wiggins, 1988] S. Wiggins. *Global Bifurcations and Chaos. Analytical Methods.* Springer, New York, 1988.

[Wiggins, 1990] S. Wiggins. *Introduction to applied nonlinear dynamical systems and chaos.* Springer, New York, 1990.

[Woodward, 2000] J. Woodward. Explanation and invariance in the special sciences. *Brit. J. Phil. Sci.*, 51, 197-254, 2000.

Part II

Biology

COMPLEX BIOLOGICAL MECHANISMS: CYCLIC, OSCILLATORY, AND AUTONOMOUS

William Bechtel and Adele Abrahamsen

1 INTRODUCTION

The mechanistic perspective has dominated biological disciplines such as biochemistry, physiology, cell and molecular biology, and neuroscience, especially during the 20^{th} century. The primary strategy is reductionist: organisms are to be decomposed into component parts and operations at multiple levels. Researchers adopting this perspective have generated an enormous body of information about the mechanisms of life at scales ranging from the whole organism down to genetic and other molecular operations.

Repeatedly, though, critics have emerged to challenge the mechanistic project. In the 18^{th} and 19^{th} centuries vitalists complained that mechanistic approaches to biology could not explain some of the important features of organisms. Xavier Bichat [1805] famously declared that organisms "resist death" — that is, owing to a distinctive, inherent ability to maintain themselves (vital force), living systems during their lifespan manage to foil the physical forces that threaten their viability. Although no longer embracing the label "vitalist," 20^{th} century opponents alleged that mechanistic sciences failed to recognize that organisms are wholes, that is, organized systems with capacities very different from those of their constituent parts.[1]

In the past these opponents lacked research techniques and tools that could explain rather than just denote the phenomena that seemed to escape mechanistic explanation. The recent application of mathematical tools for analyzing network structures and complex dynamical interactions has opened such systemic properties to analysis, and this project of complex systems modeling has begun to take root in the small but important subdisciplines of systems biology and computational biology. Certain advocates of complex systems models align with the earlier critics of mechanism, presenting their proposals as supplanting the mechanistic

[1] Many of these critics appealed to *emergent phenomena* in arguing that wholes are not just the sum of their parts, but the notion of emergence has been difficult to explicate and to insulate from concerns of spooky metaphysics. For a particularly clear discussion of emergence, see [Boogerd *et al.*, 2005].

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

program. This is misguided. The tools of complex systems modeling provide a needed extension of, but do not supplant, substantive accounts of mechanisms. To adapt a turn of phrase from Kant, dynamical models without mechanistic grounding are empty, while mechanistic models without complex dynamics are blind.

The central thrust of the mechanistic approach is to account for a target phenomenon by identifying, to a first approximation, the component parts, operations, and organization of the responsible mechanism. Ironically, given their centrality in the life sciences, mechanisms and mechanistic explanations were not much discussed by 20^{th} century philosophers of science. Their legacy is the deductive-nomological framework and its focus on laws as the primary explanatory vehicle; for them, a scientific observation is explained by formally deriving it from laws and initial conditions. More recently a number of philosophers, focusing primarily on biology, have sought to characterize mechanistic explanations and to differentiate them from deductive-nomological explanations. Although the vocabulary sometimes differs, the key elements of a basic mechanistic explanation are (1) the identification of the working parts of the mechanism, (2) the determination of the operations they perform, and (3) an account of how the parts and operations are organized so that, under specific contextual conditions, the mechanism realizes the phenomenon of interest.[2]

An organism — even a mere virus or amoeba — comprises numerous biological mechanisms. The first step in investigating each mechanism is to identify some of its component parts and operations (a step often carried out in different disciplines at different times for parts vs. operations). These components are important because they both make possible and limit what can be accomplished in the larger system. The use of lipids as building blocks of membranes, of proteins as catalysts, and of phosphate bonds for storage of energy determine many of the fundamental characteristics of organisms and provide the resources for them to maintain themselves. That chromosomes are composed of DNA, with its double helix structure, "immediately suggests a possible copying mechanism" — the pithy concluding remark by Watson and Crick [1953]. Beyond this, the nature of the bonds between nucleotides creates the possibility of complex editing, so that different proteins can be synthesized at different times from a single DNA sequence. These are just a few examples of what can be gained by identifying and investigating specific components; each has characteristics that are important for understanding the processes that maintain life. The opposite strategy — attempting to theorize about organisms without consideration of their actual building blocks — can lead to empty models, exhibiting interesting properties but not actually characterizing the organisms of this world.

Identification of component parts and operations is thus seen to be a crucial first step. The focus of this paper, though, is the implications of complex systems modeling for mechanistic explanation in biology and our understanding of

[2][Bechtel and Richardson, 1993; Bechtel, 2006; Craver, 2007; Darden, 2006; Machamer *et al.*, 2000; Thagard, 2006]. For more on our own construal of mechanistic explanation and how it differs from nomological explanation, see [Bechtel and Abrahamsen, 2005].

it. These implications are substantial. The nonlinear and non-equilibrium nature of the interacting operations within organisms often is downplayed in initial proposals of how the parts and operations are organized so as to comprise a mechanism, but they are critical to the orchestration of operations that is required for the mechanism to perform its task. Moreover, the operations performed by the parts, and even the very identity of these parts, are affected by their interactions with other parts. Consequently, the characterization generated in other, typically simpler, contexts may have to be revised as researchers come to understand the dynamical interactions occurring within organisms [Boogerd *et al.*, 2005]. Openness to such recharacterization of parts and operations fortunately lies within the mechanistic framework — as does recharacterization of their organization, if that framework is appropriately extended. Consider that mechanistic research often begins with an extremely simple conception of organization. The components are thought to operate largely independently, with each feeding the product of its internal operations to another component that has limited if any impact on the earlier component. Simon [1980] spoke of such systems as *nearly decomposable*. Numerous systems that he cites do fit that description, but biological mechanisms properly conceived generally do not. Increased recognition of their complexity has prompted inquiry into previously neglected temporal dynamics and the implications for our understanding of how operations are orchestrated in real time.

In brief, we are claiming that mechanistic research has resources for self-correction sufficient to encompass complex dynamics — there is no need to choose between mechanistic and complexity-theoretic approaches. When researchers extend the basic mechanistic program to seriously address the orchestration of operations in real time, dynamical systems and complexity theory offer relevant new tools. To flesh this out, we examine the discovery processes that led from certain mechanistic accounts with relatively simple organization to later accounts that recognized the complex dynamics characteristic of biological systems. We begin by describing how biologists came to recognize the ubiquity of cyclic organization in biology, focusing primarily on biochemistry. We then address the dynamics of such systems. In some, values of variables fluctuate irregularly (perhaps randomly) when repeatedly measured over time. Others — of greatest interest here — produce oscillations approximating the periodicity of a harmonic oscillator, such as a pendulum. Systems producing regular changes of this kind (e.g., in the concentration of a metabolite across minutes, or in alertness across hours) are referred to as *biological oscillators*. Even when there are nontrivial variations in period and amplitude, powerful tools for analysis can be brought to bear by treating such systems as oscillators. It should be mentioned, finally, that a few biologists (e.g., [Skarda and Freeman, 1987] have proposed models incorporating *chaos* (dynamics that are highly irregular, but deterministic) to explain certain biological phenomena.

The full range of dynamics should hold interest and relevance to biologists, more so than steady-state accounts, and available tools for characterizing these dynamics include mathematical modeling with differential equations and (from dynamical systems theory) limit cycles, bifurcations, chaotic regimes, and more. We are gradually moving beyond the era in which biological oscillations were concealed by such practices as focusing on the mean concentration of the product of a biochemical reaction rather than retaining the pattern of values over time. While still in the minority, there is a growing community of researchers whose questions, procedures, data, and analytic techniques are directed to discovering and characterizing biological oscillations.

There is much to be gained from enhanced attention to cyclic organization and the resulting dynamics, especially oscillations. Equally important, though, is to ask what cyclic organization and oscillatory dynamics do for the organism. The short answer is that they provide invaluable resources for controlling and orchestrating biological operations. As to why such resources are so crucial, it has been suggested that organisms most fundamentally are systems far from equilibrium that must maintain themselves as such or die: *autonomous systems* in the lexicon of the theorists offering this characterization.[3] Autonomous systems are continuously active, constantly carrying out operations necessary to their self-maintenance. But different operations can be inconsistent and even inimitable to each other. For example (as detailed later), organisms use metabolic operations to harvest energy from foodstuffs taken in from the environment, and some of these are inconsistent with operations of protein synthesis. Some means of orchestration is therefore necessary. In human engineering this most often involves external controllers, but a more elegant solution is internal cycles that interact to produce coupled oscillations. There is evidence that the ubiquity of this design in organisms figures crucially in their ability to regulate and maintain themselves.

In confronting these three features of biological mechanisms — cyclic organization, oscillatory activity, and autonomy — researchers are moving towards what we call *dynamic mechanistic explanation*. This approach significantly extends and refocuses the philosophical account of mechanism. It retains the basic mechanistic commitment to identifying parts, operations, and simple organization, but gives equal attention to determining how the activity of mechanisms built from such parts and operations is orchestrated in real time. The result is a novel framework that integrates the mechanistic philosophy of science that arose in the 1990s with the previously independent movement to understand complex systems and their dynamics. In a final section we briefly discuss the challenges in integrating mechanistic and dynamical or complexity theoretic perspectives and address broader implications.

2 FROM SEQUENTIAL TO CYCLIC ORGANIZATION

Humans typically conceive of causal operations as involving one entity acting on another — a rock damages a car by hitting its windshield, or one molecule cat-

[3][Ruiz-Mirazo *et al.*, 2004; Bickhard, 2000; Christensen and Hooker, 2000; Collier and Hooker, 1999].

alyzes a reaction that changes another molecule (e.g., by oxidizing it or adding a phosphate group to it). Note that often there are changes to the entity taken to be the cause as well as to the one affected — the rock might split when it hits the car — but this tends to be minimized as we typically conceptualize change. Moreover, once multiple steps are involved, we tend to conceptualize them as occurring sequentially. Human manufacturing focuses on adding one component at a time to a partially constructed object (as in an assembly line) and usually presumes that the already installed components are not altered in the process.

This predilection for simple organization was clearly manifest in research on alcoholic fermentation, the biochemical process essential to brewers that transforms glucose into alcohol and carbon dioxide. The chemical composition of glucose ($C_6H_{12}O_6$ in modern symbolism) and alcohol (ethanol, C_2H_5OH) was known by the early 19^{th} century, when it was assumed that fermentation was an ordinary chemical reaction. The discovery in the 1830s of yeast and its role in fermentation raised the question of whether or not fermentation was a process carried out only in whole living cells. Pasteur vigorously advocated this position and also established that fermentation occurs only in anaerobic conditions. Compelling evidence that living cells were not required finally came in 1897, when Buchner produced fermentation in extracts made by grinding and filtering yeast cells. Since these chemical soups contained a great variety of molecules as well as subcellular organelles, Buchner's success gave rise to a new question: what component(s) of cells, retained in the cell-free extracts, might be responsible for fermentation?

Buchner offered an answer that illustrates a common initial move in explaining a phenomenon: attribute it to a single component when a more complex, multicomponent mechanism is actually responsible. Accordingly, Buchner suggested that a hypothetical enzyme he named *zymase*, acting on glucose, accounted for fermentation. (By then enzymes had been characterized as chemical catalysts within cells and the suffix *–ase* used to designate them.) Other investigators, however, posited that fermentation involved multiple reactions, each catalyzed by a different enzyme, and gathered evidence pointing to various possible intermediates. Over the next thirty years they pieced together reactions involving phosphorylations, dephosphorylations, and oxidations, as well as internal reorganizations and the splitting of a six-carbon molecule into two three-carbon ones. The same reactions (except for the final one in which pyruvate is converted to alcohol) were responsible for aerobic and anaerobic glycolysis.[4] Figure 1 illustrates how the biochemists who uncovered this glycolytic pathway conceptualized it as a sequence of reactions — the simplest possible temporal organization scheme. The involvement of ATP and NAD also received minimalist treatment, as side reactions appended to the linear backbone.

In the context of oxidative metabolism (which requires aerobic conditions), pyruvate is not converted to ethanol but rather is taken up by another system of reactions to be further catabolized to water and carbon dioxide. Researchers focused on this system pursued the same strategy as for glycolysis, seeking to

[4]See [Bechtel, 2006, chapter 3], for a review of these advances in biochemistry.

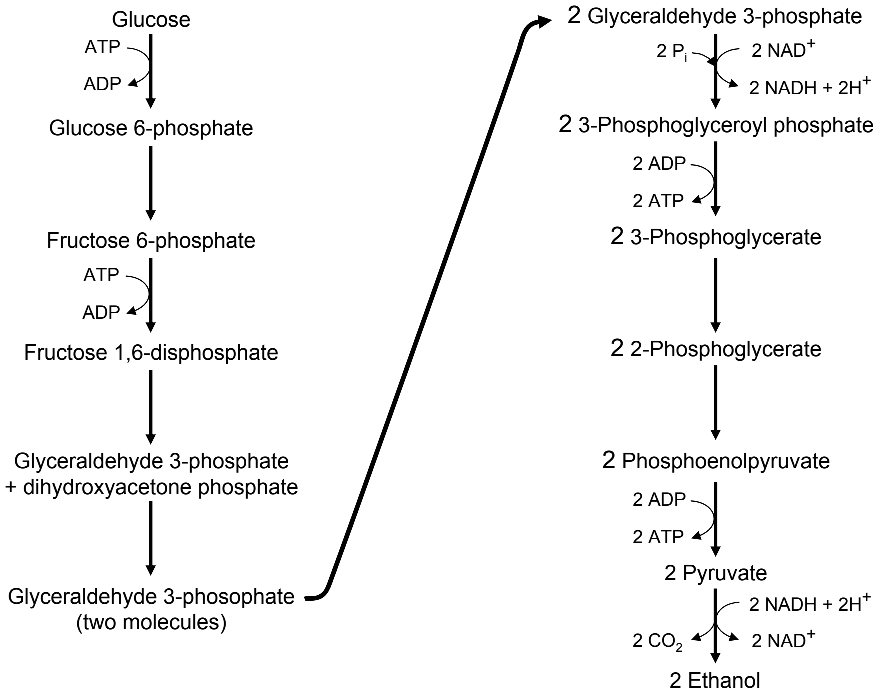


Figure 1. Glycolysis is represented as a sequence of chemical reactions

identify a sequence of molecular intermediates between an initial substrate and a final product. As before, each intermediate was assumed to be the product of one reaction and substrate of the next so as to fill in the sequence. Following upon Wieland's characterization of oxidative reactions as involving the removal and transfer of pairs of hydrogen atoms either to oxygen or to another hydrogen acceptor, Thunberg [1920] proposed a sequence of reactions, some involving oxidations, that led from succinic acid to acetic acid (with pyruvic acid as an intermediate rather than as an incoming product of glycolysis due to fragmentary knowledge of both pathways at this time):

Succinic acid → fumaric acid → malic acid → oxaloacetic acid → pyruvic acid → acetic acid

At this point Thunberg confronted a problem, since removal of two hydrogen atoms from acetic acid would not yield a known chemical compound. His solution was to propose that two molecules of acetic acid would combine; in the process each would surrender a hydrogen atom, yielding succinic acid. Necessity thus led Thunberg to close the sequence of reactions for which he had direct evidence into a cycle, but the implications were profound: a cyclic system of reactions helps resupply its own initial substrate. As it turned out, the first three reactions and the general claim of cyclic organization survived the test of time, but it was not until a landmark publication by Krebs and Johnson [1937] that a good, though still incomplete, account of this metabolic pathway was achieved. Figure 2 compares these two proposals. It can be seen that the initial substrate —- the one replenished at each turn of the cycle when an internal product reacts with an externally supplied product of glycolysis – in fact is citrate (citric acid), not succinic acid as in Thunberg's proposal.[5]

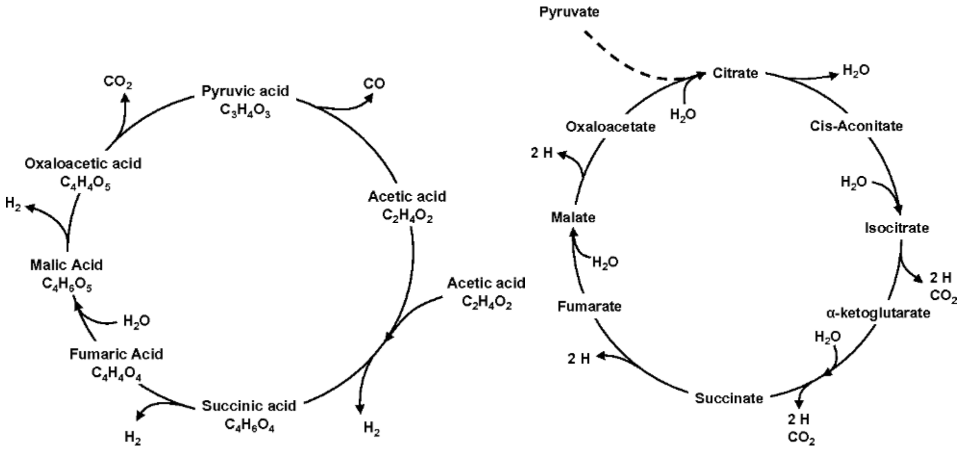


Figure 2. Two accounts of a key pathway of oxidative metabolism that recognized its cyclic organization. On the right is an early version of the Krebs cycle that was essentially correct, though incomplete. Its cyclic organization had been anticipated by Thunberg (1920), as shown on the left, but his conjecture that the crucial reaction produced succinic acid from acetic acid proved incorrect.

Hans Krebs had come to this project primed to find a cyclic solution, most directly by his own success in working out the ornithine cycle with Hanseleit in 1932. Though such cycles were born of chemical necessity, he took an interest in their functional significance and organization. Krebs [1946–8] proposed that they actually consisted of two levels of cycles. The outer, metabolic cycle repeatedly regenerates an initial substrate by means of a series of intermediate reactions, as shown in Figure 2 for the Krebs cycle and its initial substrate, citrate. Each of these reactions, though, depends upon an enzyme cycle that is simpler in that it

[5]The Krebs diagram lacks some important reactions, some discovered later and some detailing that the pairs of hydrogen atoms (2 H) were used to convert two molecules of NAD^+ to NADH or (in one reaction) FAD to $FADH_2$. It also masks debates regarding the precise role of citric acid that led to multiple names: citric acid cycle, tricarboxylic acid cycle, and simply Krebs cycle. The diagram does reflect a mid-century switch in reference from *succinic acid* to *succinate*, *citric acid* to *citrate*, etc. Both the Thunberg and Krebs diagrams must be understood as snapshots in what was a dynamic research area.

involves different forms of the same enzyme rather than a series of intermediates. He noted (p. 92) that metabolic cycles are "complex mechanisms which can be resolved into a chain of enzyme cycles" whereas enzyme cycles "cannot be further resolved into smaller cycles." Figure 3 shows how Krebs envisioned this as a cycle of cycles. Taking the enzyme cycle at the upper left as an example, the relevant substrate (malate) first binds to the enzyme (malic dehydrogenase), forming an "enzyme substrate complex" — the first step in the oxidation reaction achieved by this cycle. The enzyme takes two hydrogen atoms from malate and sends the product (oxaloacetate) to the next enzyme cycle (to the right), itself temporarily taking the form of dihydro malic dehydrogenase. The extra hydrogen atoms then combine with available oxygen to form water, leaving malic dehydrogenase free to begin the next turn of this cycle by again accepting a molecule of malate (sent from the preceding enzyme cycle as the product of a reaction with fumarate). The outer loop of metabolites (in which malate is an intermediate between fumarate and oxaloacetate, for example) is "on another plane of the chemical organisation of living matter" (p. 92) than the enzyme loops that create it. Krebs claimed that such complexly organized metabolic cycles are distinctive of life — in contrast to enzyme cycles, which are organized identically to inanimate catalytic cycles — and he was intrigued by how they enabled organisms to maintain themselves.

In the end, Krebs hinted at deeper reasons for cyclic organization than restoration of an initial state.[6] Nonetheless, a similar idea was pursued in much greater depth by the Hungarian chemist Tibor Gánti [1975], who sought to characterize the simplest chemical system that might exhibit the basic features of life. Like Maturana and Varela [1980], Gánti emphasized the need for such a system to maintain itself and identified cyclic organization as enabling a system, after it carries out a process, to be in the requisite state to perform the process again. This is true not just of biological systems but also of motors and other machines of human design. Gánti thought cycles were especially crucial for living organisms, though, because they must regularly recruit matter and energy from their environment and use it to build themselves (while expelling what they do not use as waste). Thus, he adopted an abstract characterization of the Krebs cycle as the core of his metabolic system and combined it with a limiting membrane (itself made by that system) that regulated the accumulation of metabolites. Together they constituted "a super-system" that could exhibit the fundamental biological properties of self-maintenance, growth, and reproduction.

Krebs anticipated a greater role for cycles as biochemists advanced their research: "Even if the specific meaning of cycles is still a puzzle, the fact that many processes have been found to be cycles suggests, as a working hypothesis, that other mechanisms as yet unknown might be cycles" (p. 98). He was right that the count of known cycles would increase, but might have found disappointing

[6]In particular, Figure 3 includes two kinds of reactions: irreversible (single-headed arrows) and reversible (bidirectional arrows); in consequence, the overall cycle of reactions is irreversible. Krebs conjectured that the inclusion of reversible reactions lent flexibility to the irreversible cycle in the face of changing requirements.

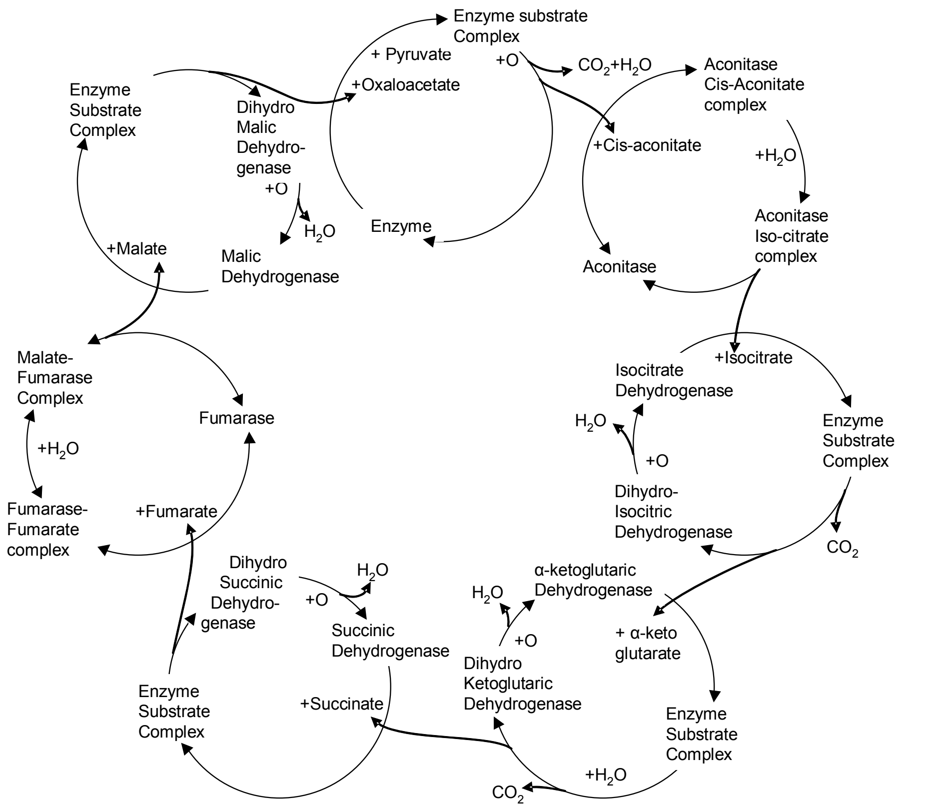


Figure 3. Krebs' [1946–48] characterization of the Krebs cycle as a cycle of cycles. (Note that citrate was omitted because its status as the initial substrate was temporarily in doubt.)

the limited pursuit of explanation. Attention to cyclic organization is discouraged even by notational conventions; serial sequences of reactions (as shown for glycolysis in Figure 1) are convenient, but also reflect and reinforce an essentially linear conceptual framework. Figure 4 conveys the limitations of a linear framework by comparing an abbreviated version of Figure 1 (left) to a rediagrammed version that reveals considerable cyclic organization (right). The simplest cycle is obtained by connecting the side-loop in which NAD^+ is reduced in the oxidation of glyceraldehyde-3-phosphate to the one in which NADH is oxidized in the reduction of pyruvate to alcohol. This illustrates how the hydrogen captured in the reduction reaction is thereby available downstream for consumption in the oxidation reaction (with NAD as carrier).

The ADP/ATP cycle is a bit more challenging to understand, in part because

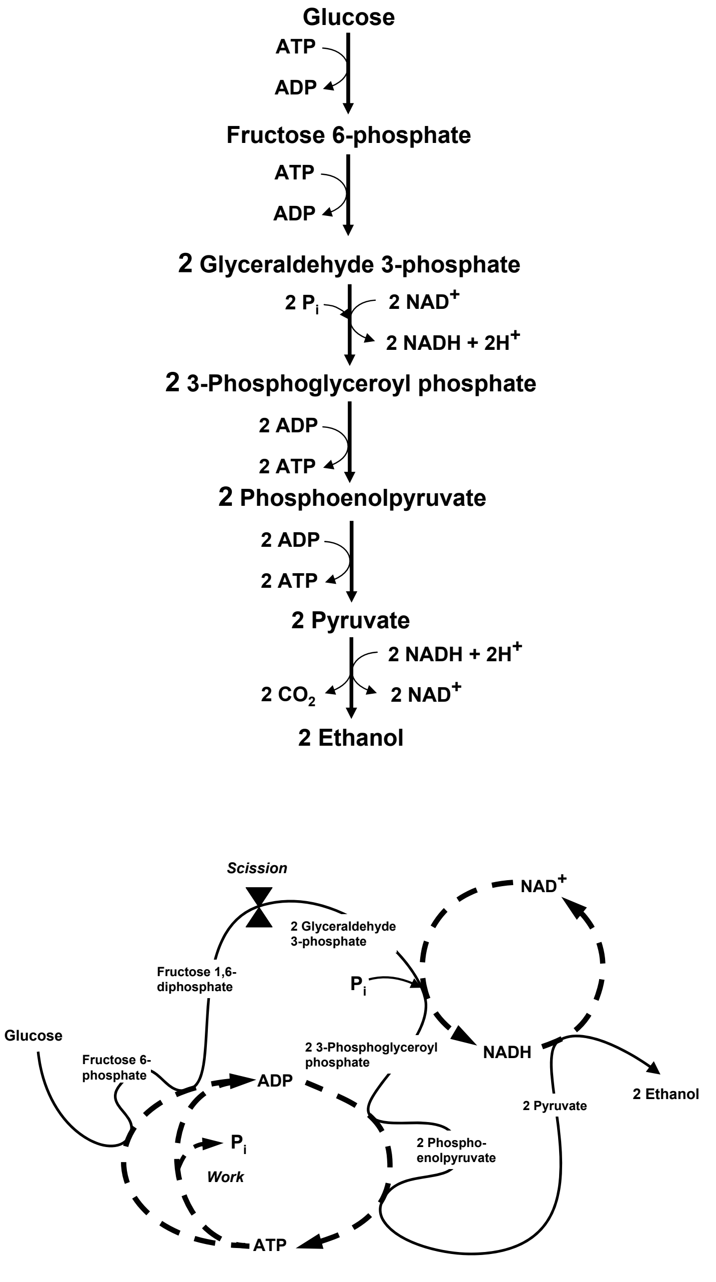


Figure 4. The linear schematization in Figure 1 is repeated here (showing only those reactions involving NAD or ATP) to contrast it with a re-representation of the glycolytic process in which the loops involving NAD and ATP are closed.

consumption of the energy stored in ATP's third phosphate bond (PO_4, also notated P_i) occurs earlier in glycolysis than the reactions that capture and store energy in such bonds. (No trick is involved; the ATP $\rightarrow$ ADP reactions take advantage of the supply of ATP in the milieu from earlier cycles of glycolysis or from other systems of reactions.) Moreover, the diagram is jigged to accommodate the fact that four different phosphorylation reactions are involved (two consumption and two storage). But the net result is that twice as much ATP is produced than consumed. The key to understanding this is the scission of what had been a single molecule into two molecules. The phosphorylation reactions that consume energy from ATP precede the scission and the dephosphorylation reactions that store energy in ATP follow it, thereby involving twice as many molecules. This makes two ATP molecules available to re-enter the glycolytic pathway (by phosphorylating first the glucose molecule and then the product of that reaction, fructose 6-phosphate) and leaves two additional ATP molecules available for other work (e.g., protein synthesis).

In brief, changing notation away from a sequential framework helps us appreciate the crucial role of cyclically organized processes. Figure 4 shows how the cycles involving NAD and ATP integrate the catabolic reactions of glycolysis into a coherent system, and hints at the dynamism of that system. Moreover, though not specifically diagrammed here, such cycles link glycolysis to other biochemical systems. ATP is used for protein synthesis and numerous other energy-consuming tasks, for example, and NADH gets shuttled into the mitochondrial matrix where it links to oxidative metabolism — especially to the electron transport chain, which uses hydrogen (electrons) carried by NADH from glycolysis and the Krebs cycle to power oxidative phosphorylation (an especially efficient conversion of ADP to ATP). This gives our biochemistry the character of Watts and Strogratz's [1998] *small worlds*: networks in which most links are between local components but a few more distant links serve to integrate the overall network. The overall metabolic system can be regarded as a small-world network. Each pathway has a series of local links, e.g., the sequential backbone of reactions in glycolysis, but the pathways are connected to each other by more distant links — especially those involving the NAD cycle. We will return to consider the role this might play in coordinating and regulating operations within the cell, after first considering oscillatory phenomena.

3 RECOGNIZING AND EXPLAINING OSCILLATORY PHENOMENA

The discovery of cyclic organization seems fairly straightforward: a sequence of reactions is found to close into a loop (the Krebs cycle) and/or to snake its way through reversible cycles that bring it into contact with other systems (a way of viewing the side reactions of glycolysis and the Krebs cycle). However, in nature such loops give rise to complex temporal dynamics. Investigators who move beyond identifying operations and sequences to consider how they unfold in real time find a wealth of phenomena to be explored. In particular, there has been an emerging awareness of the importance of oscillations in biological systems. Many oscillatory phenomena, such as the rhythmic flashing of fireflies and the beating of hearts, are obvious to the unaided senses. Others were not discovered until scientific instruments and research techniques were appropriately deployed by an attentive investigator. For example, neural oscillations have attracted substantial interest and novel proposals (see [Buzsáki, 2006], for discussion). It appears that oscillations are quite widespread in the biological world, from the intracellular chemical level all the way to the ecological level.

Despite this, biochemists and numerous other biological researchers have traditionally proceeded in a manner that blinds them to oscillations. Giving little thought to potential regularities across time that might be functional for the process of interest, but giving much thought to minimizing fluctuations regarded as noise in the data, they use preparations and techniques intended to create a steady-state system in close to equilibrium conditions. Moreover, focusing on summary measures such as mean and standard deviation conceals the dynamics of variation across time. Finding and explaining oscillations requires a major shift in thinking. We discuss three telling cases in which scientists have advanced evidence for oscillatory phenomena, identified the responsible mechanism, and investigated its characteristics as a biological oscillator. The first two cases involve *ultradian* oscillators (those with periods substantially shorter than 24 hours) and the third involves circadian oscillators (those with a period of approximately 24 hours). More specifically, these mechanisms produce the following phenomena:

1. ultradian oscillations in the glycolytic pathway discussed above;
2. ultradian oscillations separating glycolytic metabolism (during which DNA replication and protein synthesis occur) from oxidative metabolism;
3. circadian oscillations that coordinate the physiological processes and behavior of most organisms with the day-night oscillation in the environment.

In each case, once an oscillator was identified and characterized, the key question of its biological significance had to be addressed. The three cases are discussed in order from least to most satisfactory answers at our current level of knowledge: glycolytic oscillations with regular periodicity have not even been conclusively shown to occur under physiologically realistic conditions, whereas there is strong evidence that some circadian oscillators subserve important biological functions. (Perhaps coincidentally, cases 1, 2, and 3 also are ordered from shortest to longest period of oscillation.)

Glycolytic oscillations

Glycolysis provides a potent first example of the discovery and explanation of oscillations in what traditionally had been approached as a steady-state system. The initial discovery stemmed from Britton Chance's pioneering efforts to quantify biochemical processes. Working in Chance's laboratory, Amal Ghosh produced glycolysis by the usual method (adding the substrate, glucose, to suspensions of extracts from baker's yeast, which provide the necessary glycolytic enzymes). When he used Chance's spectrophotometric techniques to more closely examine the dynamics of the reaction sequence, he found that the concentration of NADH oscillated with a period of about 1 minute [Chance *et al.*, 1964]. The oscillations dampened rapidly, but Hess *et al.* [1966] developed a preparation in which oscillations of NADH continued for up to 22 hours. Within a few years, further tests revealed that the other reactants in glycolysis also showed periodic oscillations in their concentrations [Hess *et al.*, 1969]. Moreover, neighboring reactants in the glycolytic pathway oscillated together (i.e., in phase), whereas those on opposite sides of two major reactions were opposed (i.e., 180^o out of phase). The idealized graphical representation in Figure 5 shows that each reactant could be assigned to one of just four oscillatory patterns differing in relative phase, and that the phase offset ($\Delta\alpha$) for the top versus bottom inverse pairs varied with conditions — here, 70^o. By referring back to Figure 1, it can be seen where each subset of reactants resides in the glycolytic pathway — the first step in achieving a dynamic mechanistic explanation of the oscillatory phenomena.

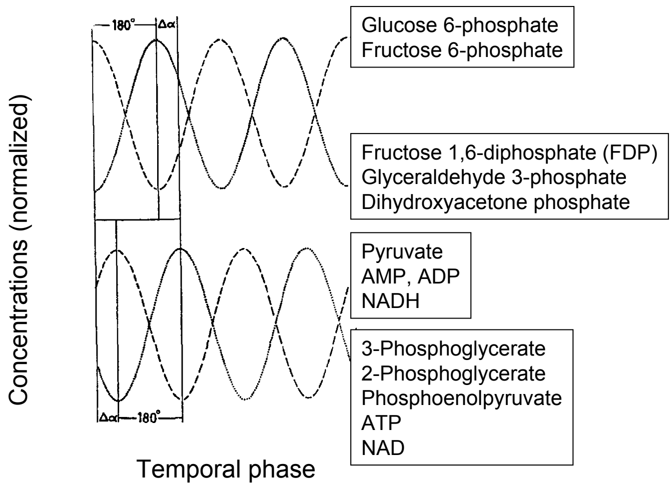


Figure 5. Idealized relative phases in the glycolytic oscillator. When each reactant's normalized concentration over time is plotted, they fall into four phase groups as shown. See text for description. Adapted from [Hess *et al.*, 1969, Figure 8].

The next step requires zooming in on the third reaction in Figure 1, in which the transfer of a phosphate group from ATP converts F6P to FDP and ATP to ADP. The fact that the first two sets of reactants in Figure 5 straddle this reaction with a phase offset of 180^o points to the enzyme involved, phosphofructokinase (PFK), as the critical factor in the oscillation [Hess *et al.*, 1969]. PFK is an allosteric enzyme — that is, an enzyme with binding sites not only for its substrate but also for other molecules that modulate its activity When these modulators are at high concentrations they more readily bind to PFK and hence have a greater effect, which is stimulatory for some modulators and inhibitory for others [Monod *et al.*, 1966]. It turns out that PFK is stimulated in this way by both products of the reaction it catalyzes (FDP and ADP) and also by AMP (made from ADP by removal of one of its two phosphate groups). Their binding to PFK serves a regulatory function, causing the entire glycolytic process to run faster — thus increasing the concentrations of NADH (immediately downstream) and ATP (further downstream). But high concentrations of ATP inhibit the reaction, contributing to its regulation by putting a long-range negative feedback loop into contention against the short-range positive feedback loops.

The times at which each loop has its maximum effect alternate, influenced in part by depletion and resupply of the substrate (F6P) and of the reactants in those loops. The interactions are quite complex, but some sense of the dynamics can be obtained by focusing on individual reactants. For example, concentrations of FDP (a product of the reaction) would tend to rise as the reaction runs faster (due in part to its own positive feedback loop) and then level off and fall as the reaction runs slower (due in part to inhibition from the ATP that had become plentiful when the reaction ran fast), then level off and rise again, and so forth. Concentrations of F6P (the substrate) would show the inverse pattern — its supply gets depleted as the reaction runs fast, which makes the reaction run slower, which allows the supply to build back up. As a final example, ATP's inhibitory effect results in less ATP being produced, which leads to less inhibition, which leads to more ATP, and so forth. In short, the feedback loops and other processes that deplete or resupply reactants dynamically interact. The net effect is a periodic oscillation in the rate of the PFK-catalyzed reaction and, due to its regulatory role, in the overall rate of glycolysis. This results in measurable oscillations in the concentrations of the various reactants in the glycolytic pathway — with phase offsets of 0^o, 70^o, or 180^o between different reactants (Figure 5) depending on where they fall in the pathway and the feedback loops (Figure 1).

We have qualitatively described the mechanism creating the oscillations, but equations specifying its operations quantitatively are necessary to account for period and amplitude of the oscillation and the conditions under which it will occur. Already in 1964 Joseph Higgins, working with Chance, published a computational model of glycolytic oscillation. It focused on the reaction catalyzed by PFK and just three of the factors known to influence it: availability of its substrate (F6P), positive feedback from the product (FDP), and removal of the product [Higgins, 1964, p. 994]. He succeeded in finding parameter value ranges in which concentrations of F6P and FDP oscillated with a phase offset close to 180^o. Shortly thereafter Sel'Kov [1968] argued that Higgins' model failed to generate a limit cycle[7] using physiologically realistic parameter values, and advanced an alterna-

[7] To show that the two reactants oscillate out of phase, it is sufficient to plot each across time

tive model that, most importantly, included the inhibitory ATP loop. With the primary source of tension now between this long-range negative feedback loop and the short-range positive feedback loops, the dynamics of this model much more closely resembled the known dynamics of the (in vitro) reaction. Thus, the PFK-catalyzed reaction alone could carry much of the explanatory burden for glycolytic oscillations, as shown also by Goldbeter and Leferver [1972] in a model comprising seven differential equations. Nonetheless, a different approach, constructing equations for all the reaction steps in glycolysis (involving 57 differential equations), was pursued by Garfinkel, Frenkel and Garfinkel. Notably, it included the other allosteric enzyme-catalyzed reaction in the glycolytic pathway, that in which ATP is resynthesized in the conversion of phosphoenolpyruvate to pyruvate. Not coincidentally, that is the reaction straddled by the third and fourth sets of reactants in Figure 5 which, like the first two sets, exhibit a phase offset of 180^o. But this reaction was not a major contributor to the overall rate of glycolysis. Including it (along with all the other reactions) yielded an account that, while more precise, was unwieldy.

None of these modelers was satisfied merely to find equations and parameter values that produced the targeted phenomena; instead, they coordinated mechanistic and dynamical systems approaches to explanation in drawing out the significance of their mathematical models. On the mechanistic side, the variables and parameters in their equations were anchored to specific parts and operations in a specific mechanistic account of glycolysis — not to global properties of the glycolytic pathway. On the dynamic side, in their pursuit of a deeper understanding of biological oscillations they used to good advantage tools for analysis of complex systems, such as limit cycles and bifurcations. We cannot develop that here, but a good early source is Gurel [1975]. In the 1980s and 1990s these applications were extended in a variety of ways. Hess [1997] provides a review of modeling and empirical work on cell-free glycolysis that reveals a full range of dynamic phenomena, from steady state to periodic to chaotic.

One extension involved the dynamics of coupled oscillators. It is a striking fact that when Huygens mounted several pendulums on the same non-rigid wall, they synchronized their oscillations. This required that an energetic product of at least one of these oscillators perturbs the oscillation of the others. Hence, when Chance *et al.* [1973] found that large populations of cells tended to synchronize their glycolytic oscillations, it raised the question of what cell product was responsible. More than 20 years later, Richard *et al.* [1996] determined that concentrations of acetaldehyde secreted into the extracellular milieu from individual cells oscillated at the same frequency as the glycolytic oscillations, and that adding acetaldehyde to a preparation could shift the phase. This pointed to it as the synchronizing

as in Figure 5. Equivalently, the concentrations of F6P and FDP can be plotted against each other (one point per timestep); this yields a repeating loop specific to the range of concentrations and their phase offset. To show that this loop is a limit cycle, it must also be the case that if the initial pair of values (F6P, FDP) lie off the loop, or if the system is perturbed, subsequent values follow a trajectory that brings them onto the loop.

agent between independent oscillating cells.

To return to the point with which we opened this section, glycolytic oscillations provide an excellent first example of how investigators came to appreciate that cyclic organization can give rise to unexpected temporal patterns. This involved two adjustments in how they construed glycolysis. First, since glycolysis is a continuous process, the default assumption had been that concentrations of the reactants would hold steady or change gradually with conditions, but in fact they oscillate with a period of approximately one minute. Second, glycolysis is not a purely linear sequence of reactions, but crucially involves cycles. This was illustrated very schematically in Figure 4, but it turned out that the ATP/ADP and $NADH/NAD^+$cycles as shown there were only part of the story. Computational modeling indicated that the key cycles were positive and negative feedback loops modulating the allosteric PFK-catalyzed reaction that converts F6P to FDP. To underscore this crucial point: it was not just that ATP obtained late in the pathway could supply a reaction early in the pathway (on the next turn of the cycle), but also that in so doing, ATP binds to PFK so as to inhibit it.

Britton Chance was particularly attracted to the glycolytic oscillator because he foresaw that it might be the basis for explaining the ability of organisms to endogenously keep time so as to produce behaviors at the appropriate time of day (circadian rhythms). Chance was to be doubly foiled in finding any confirmation for what he envisioned, however. First, as we discuss below, more recent research seeking a mechanistic explanation of circadian rhythms has pointed to gene expression, not glycolysis. Second, it proved difficult to get evidence that glycolytic oscillations occur under physiological conditions (even in whole yeast cells) and hence that they have functional significance in organisms.[8] But other oscillatory processes, with periods only slightly longer than those of glycolytic oscillators, appear to be important to physiological processes as they clearly do occur under physiological conditions and are demonstrably employed in regulating cellular processes. We turn next to these.

[8]In the 1970s there were a variety of proposals as to the functional significance of the glycolytic oscillator. It was thought, for example, that it might drive rhythmic contractions in slime molds or account for slow wave oscillations in insulin secreting β-cells (via a decrease in potassium conductance attributed to GAP dehydrogenase and phosphoglycerate kinase). Given the failure to find compelling evidence for any of these proposals, research on glycolytic oscillations declined after the 1970s. However, a new round of research was undertaken by Hans Westerhoff and his colleagues in the 1990s, spurred by the development of techniques that permitted measurement of metabolite concentrations in whole cells. Danø *et al.* [1999], for example, found a way to make measurements while continuously providing glucose and cyanide (to suppress oxidative metabolism) and removing waste, They determined that a stable attractor gave way to a limit cycle as the flow of substrate increased and that, if perturbed, the reactions showed a spiraling return to the unstable attractor — characteristics of a Hopf bifurcation. Richard *et al.* [1993] found that some of the metabolites generated after FDP — glyceraldehyde-3-phosphate (GAP), dihydroxyacetone phosphate, and phosphoenolpyruvate — either did not oscillate or did so with much smaller amplitudes. This suggested to them that the NADH oscillations were due, not to the reaction catalyzed by PFK, but rather to oscillations in the Gibbs energy of ATP hydrolysis (with the coupling achieved by GAP dehydrogenase and phosphoglycerate kinase).

Other ultradian oscillations

In addition to glycolytic oscillations with a periodicity of one minute, researchers were finding other ultradian oscillations in the biochemical processes within cells. (Rapp [1979] provides an atlas of oscillators discovered through the 1970s.) We will focus on findings of an oscillation in the overall metabolic cycle (i.e., alternations between glycolysis and oxidative metabolism) and the important claim that this oscillation is coupled both to the cell division cycle (based on measurements of the timing of DNA replication) and to gene expression (based on measurements of the timing of DNA transcription or protein synthesis). It has been easier to demonstrate such oscillations than to get consensus on their timing. The first reports regarding the metabolic cycle involved periods of just a few minutes, whereas those for protein synthesis were closer to one hour. But in brewer's yeast (*Saccharomyces cerevisiae*) grown under aerobic, glucose-restricted conditions, Satroutdinov *et al.* [1992] found a 40-minute metabolic cycle with oscillations in numerous reactants. Ethanol, for example, accumulated during the glycolytic (anaerobic) phase and was re-assimilated during the oxidative (aerobic) phase, whereas several other reactants showed the opposite oscillation (a phase offset of 180^o). As in the case of the much faster-running glycolytic oscillator, these oscillations were synchronized across cells via the action of diffusible substances such as acetaldehyde and H_2S [Sohn *et al.*, 2000]. This became a model system for David Lloyd and Douglas Murray, who found oscillations in concentrations of a host of reactants, including NAD and NADP, glutathione, ethanol, acetaldehyde, acetic acid, H_2S, and residual O_2(the oxygen that remains dissolved in the media after the organisms have drawn what they need). Lloyd and Murray [2005] referred to this suite of oscillations as the *ultradian metronome*:

> We propose that the 40-min oscillation percolates not only throughout the cellular network, including organelles, transcriptome, metabolome and proteome, but also throughout the entire population of organisms. This oscillatory state is not an exceptional curiosity found only in a peculiar system but, rather, a universal trait that is necessary for the maintenance of the robust metabolic auto-dynamic state characteristic of normally healthy cells. (p. 376)

Lloyd and Murray [2007] reported oscillations of this kind in cells from a variety of organisms (though with some variations from the 40-minute period) and found them to be coupled both to the cell division cycle and to gene expression. Specifically, the glycolytic phase of the metabolic cycle (the peak period for the reduced forms NADH and NADPH) coincides with DNA replication and transcription, whereas the oxidative metabolism phase (the peak period for the oxidized forms NAD^+ and $NADP^+$) coincides with the parts of those cycles in which DNA is intact. Lloyd and Murray [2007] also proposed a candidate mechanism for the coupling.

At its core is a reduction-oxidation (redox) cycle that constructs and breaks down disulfide bridges between two molecules of glutathione (or other thiols or small proteins). The resulting 40-minute period is robust through a wide range of temperature fluctuations, a phenomenon known as temperature compensation [Murray *et al.*, 2001].[9] Most important, this redox cycle mediates the couplings of interest. It links to metabolic pathways via NAD and NADP. That it links to DNA transcription (initiating gene expression) is evidenced by their examination of the transcripts of 5329 genes: 650 were maximally expressed during the oxidative phase and 4679 during the reductive phase [Murray *et al.*, 2007].[10] It links to DNA replication probabilistically: on any given oscillation only a subset of the cells initiate DNA replication (for cell division), but over about 8 hours all cells will replicate their DNA. We will return to the potential significance of this coupling in section 4.

One last finding bears mention. Tu and McKnight [2006] reported oscillations in the same systems, and the same couplings, but with periodicity of approximately 4 to 5 hours rather than 40 minutes — a disconcerting discrepancy that has not yet been resolved. One possibility is that each 4–5 hour oscillation contains within it (at a lower level) a number of 40-minute oscillations.

Circadian oscillations

We conclude this discussion of oscillations with perhaps the best known class of oscillatory phenomena in biology: circadian rhythms (*circa* = about + *dies* = day). A variety of physiological and behavioral manifestations have been reported from ancient times, apparently in nearly all life forms (e.g., body temperature in animals and the folding of leaves in plants; cave dwelling organisms are the most likely exception). Circadian rhythms did not become a focus of experimental investigation in biology until the work of Colin Pittendrigh [1960] and his contemporaries. The initial focus was to rigorously establish that the rhythms in animals were endogenously controlled by recording them after eliminating *Zeitgebers* (exogenous cues, such as daily light and temperature cycles). Experiments in caves showed that oscillations indeed were maintained under these conditions, albeit with periodicity varying somewhat from 24 hours. The first explanatory challenge was to find the endogenous mechanism(s) responsible for maintaining an approximately 24-hour rhythm. The second challenge was to find out how any such endogenous mechanism could be entrained by Zeitgebers so as to stay in synchrony with local

[9]Lloyd [2006] proposed that the central role of sulfur both in the synchronizing of rhythms between cells via H_2S and the building of disulfide bridges in the intracellular maintenance of the cycle could be a remnant of the origin of eukaryotic cells through a sulfur syntrophy between a-Proteobacterim and an archaeobacterial sulfide-producing host. These are progenitors of today's photosynthetic green sulfur bacteria that oxidized H_2S (either photosynthetically or using O_2) and basal Archeon, which reduced sulfur to H_2S. Such a proposal for the origin of modern mitochondria is advocated by Searcy [2003].

[10]The first suggestion of oscillations in gene expression stemmed from Soviet researcher Vsevolod Brodsky [1975; 2000]. Using UV-cytophotometry and microinterferometry to measure, among other things, RNA content, protein content, and amino acid incorporation into proteins, he identified oscillations ranging from 20 to 120 minutes, which he referred to as *circahoralian rhythms*.

day-night cycles, especially as they varied across seasons of the year, and how they could remain constant over a wide range of temperatures.

Both challenges sent researchers down to molecular biology to find answers. In the search for a molecular mechanism that could oscillate with an approximately 24-hour period, the first clue came from Konopka and Benzer [1971]. They identified a gene in *Drosophila* for which mutations resulted in shortened or lengthened rhythms or arrhythmic behavior, which they named *period* (*per*). The cloning of *per* in the 1980s led to the discovery that concentrations of its mRNA and protein oscillate in cells: specifically, *per*-mRNA peaks at the beginning of the night and the protein it codes for, PER, peaks about 6 hours later. Hardin *et al.* [1990] proposed a mechanism with a feedback loop to explain these phenomena, as shown schematically in Figure 6. First, transcription of the *per* gene generates *per* mRNA in the nucleus. These macromolecules are transported to the cytoplasm, where they are translated by ribosomes into molecules of the corresponding protein PER. After several hours PER molecules are transported back into the nucleus, where they suppress further transcription of *per*. This decreases the rate of synthesis of PER and hence also its transport into the nucleus. As the PER already present in the nucleus is broken down, *per* is released from inhibition and a new turn of the cycle begins. The elegant design of this mechanism has turned out to be applicable in a variety of other contexts. The general labels and pathway in Figure 6 (gene → mRNA and so forth) therefore qualify as a *mechanism schema*, in the terminology introduced by Machamer *et al.* [2000]. In this particular context, however, some important parts and operations in the mechanism were still unknown; for example, it was not understood how PER could suppress *per* transcription since PER molecules lack the necessary region for binding to DNA.

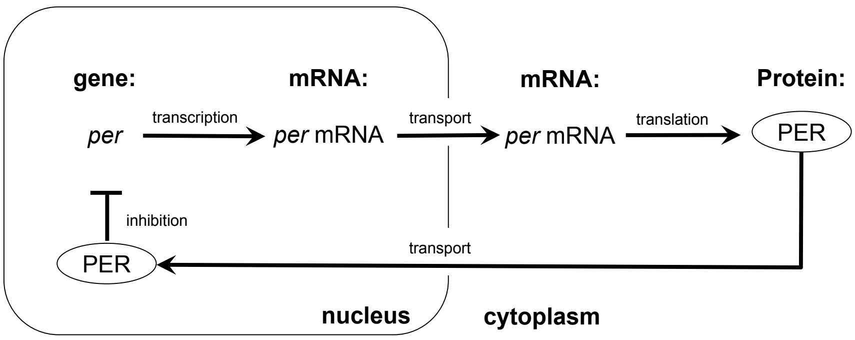
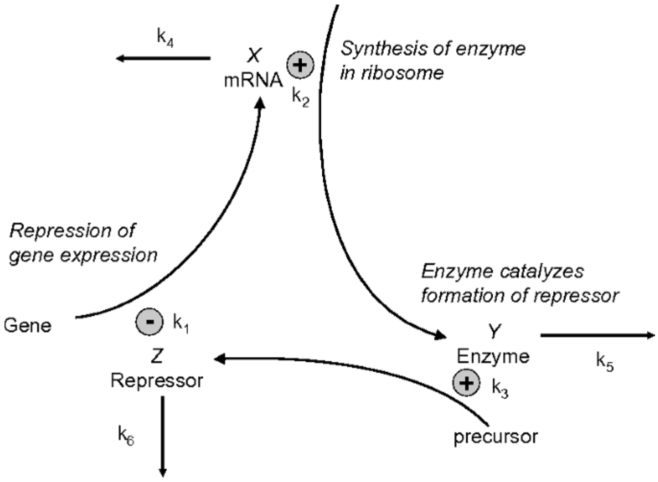


Figure 6. Hardin *et al.*'s [1990] mechanism for circadian oscillations in *Drosophila*. Expression of the gene *per* (transcription, transport and translation) produces the protein, PER, which is transported back into the nucleus. There PER inhibits further transcription of *per*. As this nuclear PER breaks down, *per* is released from inhibition and a new turn of the cycle begins.

Given the complexity of the interactions, mathematical modeling is needed to determine whether such a mechanism is actually capable of generating oscillations. Already in the 1960s, just as oscillatory phenomena were being discovered in living systems, Brian Goodwin [1965] offered an initial proposal. Inspired by the *operon* gene control mechanism proposed by Jacob and Monod [1961], he developed a system of equations that characterized a generalized version of that mechanism (Figure 7). Here two kinds of proteins collaborate to inhibit gene expression: (1) an enzyme, and (2) the product of a reaction catalyzed by that enzyme, which as a repressor molecule directly inhibits gene expression. The critical parameter for determining whether oscillations occur is n (also known as the Hill coefficient), which specifies the minimum number of interacting molecules needed to inhibit expression of the gene. Carrying out simulations on an analogue computer, Goodwin concluded that oscillations would arise with n equal to two or three. But subsequent simulations by Griffith [1968] determined that oscillations occurred only with $n > 9$, a condition that was deemed biologically unrealistic. However, if nonlinearities were introduced elsewhere (e.g., in the subtracted terms representing the removal of the various substrates from the system), it was possible to obtain oscillations with more realistic values of n. Accordingly, Goldbeter [1995b] developed his own initial model of the *Drosophila* circadian oscillator by modifying the Goodwin oscillator. By capturing the operations in the circadian mechanism shown in Figure 6 in a system of differential equations adapted from those in Figure 7, he achieved a 24-hour oscillation in concentrations of *per* mRNA and PER. Plotting these against each other over multiple cycles and conditions revealed a limit cycle (i.e., the two periodic oscillations with their particular phase offset acted as an attractor).

In the subsequent decade many additional components of the intracellular circadian oscillator in *Drosophila* were discovered and it was established that the oscillator in mammals utilizes homologues of many of the same components, albeit with some salient differences. The crucial cells for maintaining circadian rhythms in mammals had been localized in the 1970s to the suprachiasmatic nucleus (SCN), a midbrain structure of approximately 10,000 neurons, in each hemisphere, located just above the optic chiasm. Using neonatal rat cells cultured on a microelectrode array, Welsh *et al.* [1995] established that individual SCN neurons in culture sustained oscillations with a period of approximately 24 hours, though with a large standard deviation (1.2 hours). The considerable variability eventually prompted great interest, since circadian behavior in organisms is far more precise. After showing much less variability in measurements of period length for running wheel behavior in mice and for SCN slices compared with dispersed neurons, Herzog *et al.* [2004, p. 39] concluded: "Taken together, these results indicate that cell-cell interactions within the SCN synchronize SCN cells to each other and narrow the range of free-running periods expressed behaviorally." The same team subsequently advanced evidence that vasoactive intestinal protein (VIP) was the key synchronizing agent. SCN has two regions (core and shell), and a subset of cells in the core that release VIP are the only SCN cells that maintain sustained oscillations. It now



$$\frac{dX}{dt} = \frac{k_1}{Z^n + 1} - k_4 X$$
$$\frac{dY}{dt} = k_2 X - k_5 Y$$
$$\frac{dZ}{dt} = k_3 Y - k_6 Z$$

Figure 7. The Goodwin oscillator. The curved arrows specify that a gene, when not repressed, is translated into mRNA (X), which is transcribed into an enzyme (Y) in the ribosome, which then catalyzes the formation of the repressor protein (Z). The repressor then slows down the process that created it. The straight arrows indicate that the mRNA, enzyme and repressor molecules gradually break down. The rate of each operation is specified by a parameter ($k_1, k_2 \dots k_6$). The differential equations on the right give the rates at which concentrations of X, Y, and Z change over time (t).

appears that cells in the SCN shell are dependent on the VIP releasing cells for both continued oscillation and synchrony. Synchronizing of oscillators is known to be tricky and can often result in toroidal oscillations, deterministic chaos, or the coexistence of multiple attractors [Grebogi *et al.*, 1987]. A variety of simulations in the last few years have demonstrated that release of VIP is capable, at least in the models, of sustaining oscillations and producing synchronization. Moreover, using biologically plausible parameter values, Bernard *et al.* [2007] have replicated the empirical finding that shell oscillators tend to oscillate approximately 40 minutes ahead of core oscillators.

The problems of synchronization loom larger, however, when one considers responses to external inputs from Zeitgebers that are radically out of synchrony with the internal oscillators. The resulting disruptions are something human travelers experience when they cross multiple time zones. The effects of jetlag show up not just in sleep, but in a wide range of behavioral and physiological variables. Researchers learned that such variables are directly influenced by peripheral clocks in bodily organs and brain regions. Originally it appeared that these peripheral oscillators could not sustain oscillations when cut off from the SCN, so they were regarded as "slaves." However, there is now evidence that they can sustain their own oscillations but rely on complex couplings with the SCN for synchronization [Welsh *et al.*, 2004]. Recent simulations of relations between the SCN core and shell and between the shell and peripheral oscillators reveal complex responses when the system is perturbed by a six-hour change in day-night cycles (comparable to those experienced by travelers flying between North America and Europe). Empirical studies revealed that although cells in the SCN shell usually exhibit peaks in PER prior to those in the core, after a six hour light advance this order was reversed. Moreover, both advanced more than six hours (overshot the target adjustment) and it took several days to restore normal synchronization [Nakamura *et al.*, 2005]. In simulating the response of coupled oscillators representing both core and shell SCN oscillators and those in peripheral organs, Leise and Siegelmann [2006] found a very complex dynamical pattern including overshoots like those of the actual SCN oscillators. They also were successful in simulating the large number of cycles required before the peripheral oscillators returned to a normal relation to the SCN. (For further discussion of the roles of computational modeling and experimental research in achieving dynamic mechanistic explanations of circadian rhythms, see [Bechtel, in press-a].)

4 CYCLIC ORGANIZATION AND OSCILLATIONS AS FEATURES OF AUTONOMOUS SYSTEMS

One might treat the prevalence of cyclic organization and of oscillatory dynamics in living systems as simply accidents of the way biological systems happened to develop. But in fact both are of fundamental significance. One of the important features of living organisms is that they are systems far from thermodynamic equilibrium; to maintain themselves as such they must recruit matter and free energy from their environments and deploy them in the construction and repair of their own components. Insofar as such systems determine their own future existence by their success in constructing and maintaining themselves, they are referred to as *autonomous systems* (note 3). Ruiz-Mirazo and Moreno [2004] characterize basic autonomy in terms of

> the capacity of a system to *manage* the flow of matter and energy through it so that it can, at the same time, regulate, modify, and control: (i) internal self-constructive processes and (ii) processes of exchange with the environment. Thus, the system must be able to generate and regenerate all the constraints — including part of its boundary conditions — that define it as such, together with its own particular way of interacting with the environment. (p. 240)

An autonomous system is, of necessity, an active system — it must continually perform operations to maintain itself in a non-equilibrium relation with its environment. It contrasts with reactive systems that primarily respond to their environment. As Goodwin describes, the reactive perspective has been assumed in much biological research: "The traditional view of the cell as a biochemical system is that molecular populations move towards steady-state levels determined by the environment, and that when a steady state is reached the system maintains itself by a constant flow of intermediates. This view regards the cell as a passive system which changes state only in response to environmental stimuli" [Goodwin, 1965, p. 425]. Goodwin went on to show through simple models of feedback between reactions with nonlinear kinetics that spontaneous rhythmic activity was to be expected in cells and proposed: "This intrinsic rhythmic activity represents a type of biological energy which cells and organisms can use for organizing in time the staggering complexity of biochemical processes which make up living systems, thus achieving coherence and order in these activities. The interactions of nonlinear oscillators, illustrated in this paper, provide a dynamic basis for this self-organizing property of oscillating cellular control circuits" (p. 436).

To build and maintain themselves, living organisms require at their core a metabolic system that captures energy and builds the basic constituents of the organism itself. It also requires the management of a boundary so that substances needed by the system are admitted and waste products are expelled. These represent two of the three components of Gánti's [1975; 2003] proposal for a chemoton — a minimal system that manifests the basic features of living systems. Gánti's third component is a control system, which he implements through a mechanism for building polymers loosely inspired by DNA. Although such a component can play an important role in controlling a system [Griesemer & Szathmáry, 2008], considerable regulation can be achieved through cyclic organization and oscillatory processes without an external control system. It is an open question whether these are sufficient to realize all of the fundamental regulatory processes of life, or alternatively, whether entities comparable to genes are required to regulate even basic metabolism and movement of substances to and from the chemoton.[11]

The claim is that a particular biological cycle, regardless of whether it produces oscillations, provides a vehicle for regulating a system so that individual operations are performed at the time needed. One way to see this is to look at one of the feedback loops in glycolysis on its own, rather than in the usual competitive context known to produce oscillations. In particular, consider the negative loop in which

[11]One argument for the claim that something like genes are needed for effective control is that in all known organisms metabolic reactions and control over boundaries are achieved by complex proteins, and we have no account of how these structures could reliably develop in simple chemical systems via self-organization alone. Moreno (personal communication, September 2008) contends that in order to realize effective control, the controller must be at least partly dynamically decoupled from the system controlled. Before concluding that this is correct, though, we should explore further how much regulation can be achieved through interactions such as those that give rise to limit cycles that can be employed to segregate reactions in time (discussed below). At some point, living systems did begin to rely on structures such as genes as partially decoupled controllers. Clearly a significant consequence of relying on genes as control elements is that their stability enables them to be inherited and thereby provide the heritability needed in evolutionary processes including natural selection. It is important to note that genes, as well as the polymers generated in Gánti's chemoton, are static entities that do nothing on their own. Other components, including the apparatus for transcribing DNA into mRNA, editing the mRNA, and translating mRNA into proteins must also be inherited. Even if partly decoupled from what it controls, the actual control system is itself a dynamic system, not a static element.

high concentrations of ATP inhibit the PFK-catalyzed reaction. This ensures that glucose will not be metabolized unless energy is needed for other cell activities. Thus, even a single cycle enables at least some regulation of an important metabolic process.

When a mechanism's temporal dynamics do result in oscillations, these too can be used to regulate other mechanisms by coupling their activity. Goodwin [1963] proposed that oscillators provided a means of temporally segregating incompatible cellular events.[12] Recall that DNA replication and most transcription of genes occurs during the glycolytic phase, when oxygen consumption is low. Both DNA replication and transcription involve opening up the double helix structure and exposing the nucleic acids, which can be damaged by exposure to oxygen. Thus, by limiting these activities to periods when oxygen levels are low, DNA is protected [Lloyd and Murray, 2006; Tu and McKnight, 2006].

Circadian oscillations provide an even stronger illustration of the idea that oscillatory processes provide a means of segregating incompatible operations. A clear example is found in the cyanobacterium, *Synechococcus elongates*. The enzyme nitrogenase, critical for nitrogen fixation, is destroyed by oxygen, which the organism produces during photosynthesis. Its circadian oscillator ensures that nitrogen fixation and photosynthesis occur at different times, with photosynthesis proceeding during daylight hours, when the required sunlight is most likely to be available, and nitrogen fixation at night, when no oxygen is being produced.

Circadian oscillations also perform another control role: enabling physiological and behavioural processes to occur at optimal times of day. Examples include sleep during the night (for diurnal animals) sleep during the day (for nocturnal animals), and food foraging when prey are available. It might seem sufficient to rely on environmental cues for many of these activities, but appropriate performance often requires preparation before the environmental cue would be available.

5 CONCLUSION: IMPLICATIONS FOR MECHANISTIC SCIENCE

Mechanistic research has been extremely successful in identifying the parts, operations, and basic organization of a vast range of biological mechanisms. It has been less successful, though, in understanding the implications of various forms of organization — especially the temporal dynamics that orchestrate the functioning of biological mechanisms. Biochemists' basic strategy has been to put together linear sequences of reactions, moving to cyclic organization (e.g., the Krebs cycle) only as necessary. By focusing on near-equilibrium steady-state conditions and summary statistics (e.g., mean concentration of a metabolite), traditionally biologists have screened themselves off from oscillatory phenomena. We have suggested that the resulting mechanistic accounts are blind to crucial dynamics of

[12] Spatial differentiation of organelles is another way to obtain such segregation. Enzymes involved in breaking down cellular constituents, for example, are segregated in the lysosome so that they operate only on materials that have been transported for that purpose into that organelle. Temporal segregation can achieve the same purpose.

the systems they target. The new mechanistic philosophy of science has tended to parallel biology in this respect, emphasizing the discovery of component parts and operations and simple organizational schemes, and providing little systematic attention to orchestration.

Although explanation in biology remains primarily mechanistic, small clusters of investigators have confronted the complex dynamics that serve to orchestrate the functioning of biological mechanisms. We have described a few of these endeavors, focusing especially on explorations of the oscillatory dynamics that can result from some common types of cyclic organization (for other exemplars, see [Goldbeter, 1995a; Noble, 2006; Buzsáki, 2006; Ellner and Guckenheimer, 2006]). All of the cases we discussed were grounded in accounts of parts, operations, and organization of a particular biological mechanism but added concepts and tools of mathematical modeling and dynamical systems. Hence, they well exemplify the project of dynamic mechanistic explanation that we have endorsed.

Dynamic mechanistic explanation stands in contrast not only to purely mechanistic explanation but also to theoretical inquiries that emphasize complex dynamics in living systems conceived abstractly — at best neglecting but in some cases explicitly rejecting the mechanistic project. Artificial life research, for example, is conducted on a plane removed from research on actual biological mechanisms. While accounts oriented purely to complexity or dynamics can make unique and valuable contributions, they provide no understanding of how the dynamic relations are actually realized in living systems if they do not get anchored to component parts and operations of actual mechanisms. That is, they are empty. We contend that complexity and dynamical systems theory find their best use as tools in a more integrated endeavor.

Some theoretical biologists (e.g., [Kauffman, 2000]) have not only preferred to work on an abstract plane, but also aspired to achieve a unified, law-based theoretical framework. In the spirit of cybernetics and general systems theory, they direct themselves to the big picture that seems to be neglected in reductionistic inquiry. Again, this endeavor has produced some ingenious and valuable directions for further inquiry, but does not in itself achieve the integration we regard as crucial. The most promising contributions for near-term integration probably come not from comprehensive systems, but from particular proposed principles of organization: self-organization through positive feedback in non-equilibrium conditions, small-world organization, scale-free networks [Barabási and Bonabeau, 2003], and so forth.

A characteristic feature of modern biology is its particularity. Biochemical pathways, while showing common patterns across phyla, also reveal substantial differences that matter to the functioning of particular organisms. The same is true of circadian oscillators [Bechtel, in press-b]. The resulting extrapolation from studied models to other systems is very different from the generalization achieved by universal quantifiers in laws. Researchers do not know in advance which features change and which remain constant (or nearly so) when extrapolating, and must be prepared to modify specific parts and operations in mechanisms as they move to new instances. The same is likely to apply to the tools of complex systems analysis. That is, the general understanding of how small-worlds phenomena emerge from a few long-range connections in networks primarily constituted of short-range connections will need to be adapted given the particular long-range connections found in a given system. Complex systems analyses provide a rich toolkit for appreciating organization and orchestration of operations in biological mechanisms, and invoking these tools can illuminate how these mechanisms generate the rich phenomena biology seeks to explain. This will not, however, obviate the need to understand the particularity of any given mechanism.

Thus, we see an immediate future in which dynamic mechanistic researchers in biology will continue to offer piecemeal, context-specific accounts, even as they stretch them to incorporate dynamics. Systems biologists and philosophers of science can, and should, add insight and perspective while remaining grounded by examining a spectrum of these accounts. Such generalizations as we extract will be works in progress, with frequently modified contextual and other limitations, and will not readily be systematized — including those regarding cycles, oscillations and autonomy.

ACKNOWLEDGEMENTS

We thank Clifford Hooker, Xabier Barandiaran, Michael Grinfeld, Alvaro Moreno, and Kepa Ruiz-Mirazo for very helpful comments on an earlier version of this chapter.

BIBLIOGRAPHY

[Barabási and Bonabeau, 2003] A.-L. Barabási and E. Bonabeau. Scale-free networks. *Scientific American*(May), 50-59, 2003.

[Bechtel, 2006] W. Bechtel. *Discovering cell mechanisms: The creation of modern cell biology.* Cambridge: Cambridge University Press, 2006.

[Bechtel, in press-a] W. Bechtel. Computational modeling and mechanistic explanation: Understanding circadian rhythms. *Studies in History and Philosophy of Science Part A*, in press.

[Bechtel, in press-b] W. Bechtel. Generalization and discovery through conserved mechanisms: Cross species research on circadian oscillators. *Philosophy of Science*, in press.

[Bechtel and Abrahamsen, 2005] W. Bechtel and A. Abrahamsen. Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 421-441, 2005.

[Bechtel and Richardson, 1993] W. Bechtel and R. C. Richardson. *Discovering complexity: Decomposition and localization as strategies in scientific research.* Princeton, NJ: Princeton University Press, 1993.

[Bernard *et al.*, 2007] S. Bernard, D. Gonze, Čajavec, H. Herzel, and A. Kramer. Synchronization-induced rhythmicity of circadian oscillators in the suprachiasmatic nucleus. *PLoS Computational Biology*, 3(4), e68, 2007.

[Bichat, 1805] X. Bichat. *Recherches Physiologiques sur la Vie et la Mort*, 3rd edn. Paris: Machant, 1805.

[Bickhard, 2000] M. H. Bickhard. Autonomy, function and representation. *Communication and Cognition - Artificial Intelligence, 17*, 111-131, 2000.

[Boogerd *et al.*, 2005] F. C. Boogerd, F. J. Bruggeman, R. C. Richardson, A. Stephan, and H. V. Westerhoff. Emergence and its place in nature: A case study of biochemical networks. *Synthese, 145*, 131-164, 2005.

[Brodsky, 1975] V. Y. Brodsky. Protein synthesis rhythm. *Journal of Theoretical Biology, 55*(1), 167-200, 1975.

[Brodsky, 2000] V. Y. Brodsky. Circahoralian (ultradian) cellular rhythms: Initial studies and some results. *Russian Journal of Developmental Biology, 31*, 345-353, 2000.

[Buzsáki, 2006] G. Buzsáki. *Rhythms of the brain.* Oxford: Oxford University Press 2006.

[Chance *et al.*, 1964] B. Chance, R. W. Estabrook, & A. Ghosh. Damped sinusoidal oscillations of cytoplasmic reduced pyridine nucleotide in yeast cells. *Proceedings of the National Academy of Sciences, 51*(6), 1244-1251, 1964.

[Chance *et al.*, 1973] B. Chance, G. Williamson, J. Lee, L. Mela, D. DeVault, A. K. Ghosh, *et al.* Synchronisation phenomena in oscillation of yeast cells and isolated mitochondria. In B. Chance, E. K. Pye, A. K. Ghosh & B. Hess, eds., *Biological and biochemical oscillations* (pp. 285-300). London: Academic, 1973.

[Christensen and Hooker, 2000] W. D. Christensen and C. A. Hooker. Autonomy and the emergence of intelligence: Organised interactive construction. *Communication and Cognition - Artificial Intelligence, 17*, 133-157, 2000.

[Collier and Hooker, 1999] J. D. Collier and C. A. Hooker. Complexly organised dynamical systems. *Open Systems and Information Dynamics, 6*, 241-302, 1999.

[Craver, 2007] C. Craver. *Explaining the brain: What a science of the mind-brain could be.* New York: Oxford University Press, 2007.

[Danø *et al.*, 1999] S. Danø, P. G. Sørensen, and F. Hynne. Sustained oscillations in living cells. *Nature, 402*(6759), 320, 1999.

[Darden, 2006] L. Darden. *Reasoning in biological discoveries: Essays on mechanisms, interfield relations, and anomaly resolution.* Cambridge: Cambridge University Press, 2006.

[Ellner and Guckenheimer, 2006] S. P. Ellner and J. Guckenheimer. *Dynamic models in biology.* Princeton, NJ: Princeton University Press, 2006.

[Gánti, 1975] T. Gánti. Organization of chemical reactions into dividing and metabolizing units: The chemotons. *BioSystems, 7*, 15-21, 1975.

[Gánti, 2003] T. Gánti. *The principles of life.* New York: Oxford, 2003.

[Goldbeter, 1995a] A. Goldbeter. *Biochemical oscillations and cellular rhythms: The molecular bases of periodic and chaotic behaviour.* Cambridge: Cambridge University Press, 1995.

[Goldbeter, 1995b] A. Goldbeter. A model for circadian oscillations in the *Drosophila* Period protein (PER). *Proceedings of the Royal Society of London. B: Biological Sciences, 261*(1362), 319-324, 1995.

[Goldbeter and Lefever, 1972] A. Goldbeter and R. Lefever. Dissipative structures for an allosteric model application to glycolytic oscillations. *Biophysics Journal, 12*(10), 1302-1315, 1972.

[Goodwin, 1963] B. C. Goodwin. *Temporal organization in cells; a dynamic theory of cellular control processes.* New York: Academic Press, 1963.

[Goodwin, 1965] B. C. Goodwin. Oscillatory behavior in enzymatic control processes. *Advances in Enzyme Regulation, 3*, 425-428, 1965.

[Grebogi *et al.*, 1987] C. Grebogi, E. Ott, and J. A. Yorke. Chaos, strange attractors, and fractal basin boundaries in nonlinear dynamics. *Science, 238*(4827), 632-638, 1987.

[Griesemer and Szathmáry, 2008] J. R. Griesemer and E. Szathmáry. Gánti's chemoton model and life criteria. In S. Rasmussen, L. Chen, N. Packard, M. Bedau, D. Deamer, P. Stadler & D. Krakauer, eds., *Protocells: Bridging nonliving and living matter.* Cambridge, MA: MIT Press, 2008.

[Griffith, 1968] J. S. Griffith. Mathematics of cellular control processes I. Negative feedback to one gene. *Journal of Theoretical Biology, 20*(2), 202-208, 1968.

[Gurel, 1975] O. Gurel. Limit cycles and bifurcations in biochemical dynamics. *Biosystems, 7*(1), 83-91, 1975.

[Hardin *et al.*, 1990] P. E. Hardin, J. C. Hall, and M. Rosbash. Feedback of the *Drosophila period* gene product on circadian cycling of its messenger RNA levels. *Nature, 343*(6258), 536-540, 1990.

[Herzog *et al.*, 2004] E. D. Herzog, S. J. Aton, R. Numano, Y. Sakaki, and H. Tei. Temporal precision in the mammalian circadian system: A reliable clock from less reliable neurons. *Journal of Biological Rhythms, 19*(1), 35-46, 2004.

[Hess, 1997] B. Hess. Periodic patterns in biochemical reactions. *Quarterly Reviews of Biophysics, 30*(2), 121-176, 1997.

[Hess *et al.*, 1969] B. Hess, A. Boiteux, and J. Krüger. Cooperation of glycolytic enzymes. *Advances in Enzyme Regulation, 7*, 149-167, 1969.

[Hess *et al.*, 1966] B. Hess, K. Brand, and K. Pye. Continuous oscillations in a cell-free extract of *S. carlsbergensis*. *Biochemical and Biophysical Research Communications, 23*(1), 102-108, 1966.

[Higgins, 1964] J. Higgins. A chemical mechanism for oscillation of glycolytic intermediates in yeast cells. *Proceedings of the National Academy of Sciences, 51*(6), 989-994, 1964.

[Jacob and Monod, 1961] F. Jacob and J. Monod. Genetic regulatory systems in the synthesis of proteins. *Journal of Molecular Biology, 3*, 318-356, 1961.

[Kauffman, 2000] S. A. Kauffman. *Investigations*. Oxford: Oxford University Press, 2000.

[Konopka and Benzer, 1971] R. J. Konopka and S. Benzer. Clock mutants of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences (USA), 89*, 2112-2116, 1971.

[Krebs, 1946–8] H. A. Krebs. Cyclic Processes in Living Matter. *Enzymologia, 12*, 88-100, 1946–8.

[Krebs and Johnson, 1937] H. A. Krebs and W. A. Johnson. The role of citric acid in intermediate metabolism in animal tissues. *Enzymologia, 4*, 148-156.

[Leise and Siegelmann, 2006] T. Leise and H. Siegelmann. Dynamics of a multistage circadian system. *Journal of Biological Rhythms, 21*(4), 314-323, 2006.

[Lloyd, 2006] D. Lloyd. Hydrogen sulfide: clandestine microbial messenger? *Trends in Microbiology, 14*(10), 456-462, 2006.

[Lloyd and Murray, 2005] D. Lloyd and D. B. Murray. Ultradian metronome: timekeeper for orchestration of cellular coherence. *Trends in Biochemical Sciences, 30*(7), 373-377, 2005.

[Lloyd and Murray, 2006] D. Lloyd and D. B. Murray. The temporal architecture of eukaryotic growth. *FEBS Letters, 580*(12), 2830-2835, 2006.

[Lloyd and Murray, 2007] D. Lloyd and D. B. Murray. Redox rhythmicity: clocks at the core of temporal coherence. *BioEssays, 29*(5), 465-473, 2007.

[Machamer *et al.*, 2000] P. Machamer, L. Darden, and C. Craver. Thinking about mechanisms. *Philosophy of Science, 67*, 1-25, 2000.

[Maturana and Varela, 1980] H. R. Maturana and F. J. Varela. Autopoiesis: The organization of the living. In H. R. Maturana & F. J. Varela, eds., *Autopoiesis and cognition: The realization of the living* (pp. 59-138). Dordrecht: D. Reidel, 1980.

[Monod *et al.*, 1966] J. Monod, J. Wyman, and J.-P. Changeux. On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology, 12*, 88-118, 1966.

[Murray *et al.*, 2007] D. B. Murray, M. Beckmann, and H. Kitano. Regulation of yeast oscillatory dynamics. *Proceedings of the National Academy of Sciences, 104*(7), 2241-2246, 2007.

[Murray *et al.*, 2001] D. B. Murray, S. Roller, H. Kuriyama, and D. Lloyd. Clock control of ultradian respiratory oscillation found during yeast continuous culture. *Journal of Bacteriology, 183*(24), 7253-7259, 2001.

[Nakamura *et al.*, 2005] W. Nakamura, S. Yamazaki, N. N. Takasu, K. Mishima, and G. D. Block. Differential response of *Period 1* expression within the suprachiasmatic nucleus. *Journal of Neuroscience, 25*(23), 5481-5487, 2005.

[Noble, 2006] D. Noble. *The music of life: Biology beyond the genome*. New York: Oxford, 2006.

[Pittendrigh, 1960] C. S. Pittendrigh. Circadian rhythms and the circadian organization of living systems. *Cold Spring Harbor Symposia on Quantitative Biology, 25*, 159-184, 1960.

[Rapp, 1979] P. E. Rapp. An atlas of cellular oscillators. *Journal of Experimental Biology, 81*(1), 281-306, 1979.

[Richard *et al.*, 1996] P. Richard, B. Teusink, M. B. Hemker, K. van Dam, and H. V. Westerhoff. Sustained oscillations in free-energy state and hexose phosphates in yeast. *Yeast, 12*(8), 731-740, 1996.

[Richard *et al.*, 1993] P. Richard, B. Teusink, H. V. Westerhoff, and K. van Dam. Around the growth phase transition S. cerevisiae's make-up favours sustained oscillations of intracellular metabolites. *FEBS Letters, 318*(1), 80-82, 1993.

[Ruiz-Mirazo and Moreno, 2004] K. Ruiz-Mirazo and A. Moreno. Basic autonomy as a fundamental step in the synthesis of life. *Artificial Life, 10*, 235-259, 2004.

[Ruiz-Mirazo *et al.*, 2004] K. Ruiz-Mirazo, J. Peretó, and A. Moreno. A universal definition of life: Autonomy and open-ended evolution. *Origins of Life and Evolution of the Biosphere, 34*, 323-346, 2004.

[Satroutdinov *et al.*, 1992] A. D. Satroutdinov, H. Kuriyama, and H. Kobayashi. Oscillatory metabolism of *Saccharomyces cerevisiae* in continuous culture. *FEMS Microbiology Letters, 98*(1-3), 261-267, 1992.

[Searcy, 2003] D. G. Searcy. Metabolic integration during the evolutionary origin of mitochondria. *Cell Res, 13*(4), 229-238, 2003.

[Sel'Kov, 1968] E. E. Sel'Kov. Self-oscillations in glycolysis. 1. A simple kinetic model. *European Journal of Biochemistry, 4*(1), 79-86, 1968.

[Simon, 1980] H. A. Simon. *The sciences of the artificial. 2nd edn.* Cambridge, MA: MIT Press, 1980.

[Skarda and Freeman, 1987] C. A. Skarda and W. J. Freeman. How brains make chaos to make sense of the world. *Behavioral and Brain Sciences, 10*, 161-195, 1987.

[Sohn *et al.*, 2000] H.-Y. Sohn, D. B. Murray, and H. Kuriyama. Ultradian oscillation of *Saccharomyces cerevisiae* during aerobic continuous culture: hydrogen sulphide mediates population synchrony. *Yeast, 16*(13), 1185-1190, 2000.

[Thagard, 2006] P. Thagard. *Hot thought: Mechanisms and applications of emotional cognition.* Cambridge, MA: MIT Press, 2006.

[Thunberg, 1920] T. L. Thunberg. Zur Kenntnis des intermediären Stoffwechsels und der dabei wirksamen Enzyme. *Skandinavisches Archiv für Physiologie, 40*, 9-91, 1920.

[Tu and McKnight, 2006] B. P. Tu and S. L. McKnight. Metabolic cycles as an underlying basis of biological oscillations. *Nature Reviews Molecular Cell Biology, 7*(9), 696-701, 2006.

[Watson and Crick, 1953] J. D. Watson and F. H. C. Crick. Molecular structure of nucleic acids. *Nature, 171*, 737-738, 1953.

[Watts and Strogratz, 1998] D. Watts and S. Strogratz. Collective dynamics of small worlds. *Nature, 393*(440-442), 1998.

[Welsh *et al.*, 1995] D. K. Welsh, D. E. Logothetis, M. Meister, and S. M. Reppert. Individual neurons dissociated from rat suprachiasmatic nucleus express independently phased circadian firing rhythms. *Neuron, 14*(4), 697-706, 1995.

[Welsh *et al.*, 2004] D. K. Welsh, S.-H. Yoo, A. C. Liu, J. S. Takahashi, and S. A. Kay. Bioluminescence imaging of individual fibroblasts reveals persistent, independently phased circadian rhythms of clock gene expression. *Current Biology, 14*(24), 2289-2295, 2004.

ON CAUSALITY IN NONLINEAR COMPLEX SYSTEMS: THE DEVELOPMENTALIST PERSPECTIVE

James A. Coffman

SUMMARY

Science seeks to delineate causal relationships in an effort to explain empirical phenomena, with the ultimate goal being to understand, and whenever possible predict, events in the natural world. In the biological sciences, and especially biomedical science, causality is typically reduced to those molecular and cellular mechanisms that can be isolated in the laboratory and thence manipulated experimentally. However, increasing awareness of emergent phenomena produced by complexity and non-linearity has exposed the limitations of such reductionism. Events in nature are the outcome of processes carried out by complex systems of interactions produced by historical contingency within dissipative structures that are far from thermodynamic equilibrium. As such, they cannot be adequately explained in terms of lower level mechanisms that are elucidated under artificial laboratory conditions. Rather, a full causal explanation requires comprehensive examination of the flow networks and hierarchical relationships that define a system and the context within which it exists.

The fact that hierarchical context plays a critical role in determining the outcome of events reinvigorates Aristotelian conceptions of causality. One such perspective, which I refer to as *developmentalism*, views all non-random causality as a product of development at some level. Development ('self-organization') occurs via the selective agency of autocatalytic cycles inherent in certain configurations of processes, which competitively organizes a system as resources become limiting. In this view bottom-up causality (the concern of reductionism) holds sway mainly in immature systems, whereas top-down causality (organizational or informational constraint) dominates mature systems, the functioning of which is less dependent (and more constraining) on the activities of their lower-level parts. Extrapolating the developmentalist perspective to the limit, one might posit that the ultimate arbiters of causality, the 'laws of physics', are themselves no more than organizational constraints produced by (and contingent upon) the early development of the universe. The causal relationships that define chemistry and biology are more highly specified organizational constraints produced by later development.

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

Developmentalism helps resolve a number of long-standing dialectics concerned with causality, including reductionism/holism, orthogenesis/adaptation, and stasis/change.

In biological sciences, developmentalism engenders a discourse that overcomes barriers imposed by the still-dominant paradigms of molecular reductionism on the one hand and Darwinian evolution on the other. With regard to the former, it provides a better interpretive framework for the new science of 'systems-biology', which seeks to elucidate regulatory networks that control ontogeny, stem cell biology, and the etiology of disease. With regard to the latter, it provides an intelligible bridge between chemistry and biology, and hence an explanation for the natural origin of life. Finally, developmentalism, being an inherently ecological perspective, is well-suited as a paradigm for addressing problems of environmental management and sustainability.

1 INTRODUCTION

Scientific inquiry is motivated by a desire to understand the causal basis of empirical reality, which serves to enhance the predictability of nature. But although the word 'cause' is frequently used in scientific discourse, it is rarely if ever defined; the meaning of the term is held to be implicit and commonly understood.

In modern biomedical science, one of the pillars of this common understanding is the assumption of Reductionism: that causality can be reduced to molecular and cellular 'mechanisms', particularly those that can be isolated in the laboratory and thence manipulated experimentally. Macroscopic events (e.g., disease states such as cancer or infection) are generally thought to be caused by microscopic processes, the causes of which can be further reduced into even more microscopic processes. Whereas whatever predictable regularity manifested by these processes is attributed to mechanisms governed by inviolable, mathematically tractable laws of physics and chemistry, evolutionary change is explained in terms of Darwinian selection acting on random variation. Causal agency is thus assumed to be embodied by deterministic physico-chemical mechanisms that control macroscopic processes from the bottom-up, and which evolve over time as a result of low-level stochasticity being sculpted by the high-level filter of selection [Monod, 1972].

Western conceptualizations of causality beginning with Plato and Aristotle evolved through a succession of theological and philosophical discourses that eventually coalesced into the mechano-reductionistic determinism of modern science [Hulswit, 2004a; 2004b]. According to Hulswit [2004b], "the complex evolution of cause from the seventeenth century on is marked by the interplay between, at least, two radically different conceptions of cause: the Aristotelian-scholastic conception, according to which causes are the *active initiators of a change*, and the scientific conception, according to which causes are the *inactive nodes in a law-like implication chain*." In respect to this duality, we can still nevertheless classify the concerns of scientific inquiry according to the four causal categories established by Aristotle: material (substances), efficient (mechanisms), formal (circumstances), and final (needs). We might then say that formal causes compel efficient causes to serve as 'active initiators of a change' involving material causes that are 'inactive nodes in a law-like implication chain' manifesting *telos*, the logic of final causality.[1] Whereas reductionistic approaches seek to define material and efficient causes that answer the questions 'what is this and how does it work?', phenomenological approaches seek formal and final causes that answer the question 'why did this happen?'. Answers to the questions 'what is this and how does it work' are geared toward practicality: they provide a more immediate means for exerting some semblance of control over the composition and mechanical workings of nature. In contrast, answers to the question 'why did this happen' are (at least initially) more academic: they provide insight into the conditions and 'laws' of nature and the cosmos. Since the efficient/material operators of biology and chemistry are subject to the formal conditions and final logic of thermodynamics and cosmology, the reductionistic pre-occupation of biomedical science appears on the surface to be justified by conventional wisdom regarding the hierarchy of sciences: {Physics{Chemistry{Biology}}}.

There is nothing inherently wrong with a reductionistic approach to science; like phenomenology, it is an indispensable line of inquiry that has yielded valuable knowledge. The problem, as always, comes from the oversimplifications that arise from not recognizing the limitations inherent in a given approach, which usually stems from a failure to examine fundamental, untested assumptions about the nature of reality. When this happens, an otherwise reasonable discourse can develop into a dangerous caricature: a Frankenstein's monster. Some scientists may argue that studying philosophy and history is not relevant to their work, and therefore a waste of precious time. To that I would counter that insight provided by studies of philosophy and history provides the only defense against a discourse's development into caricature.

One assumption underlying the caricature of modern biomedical Reductionism (which although admittedly somewhat of a rhetorical straw man, is held near and dear by many) is that higher levels of empirical reality (e.g., human health) are determined by, and can therefore be *adequately* explained in terms of, lower level (molecular and cellular) mechanisms. Efficient/material causes are thereby privileged, leading to the denigration of phenomenology as 'descriptive' science that does not contribute to an understanding of causality. A related assumption is that causality can be reduced to linear chains of events, each event representing the predictable outcome of a mechanical process. Thus, biomedicine is heavily invested in studies of the biochemical 'pathways' of metabolism and intercellular signaling. To be sure, this endeavor has elucidated complex networks of interacting pathways, a vast and invaluable body of knowledge. But understanding has failed to keep pace with this knowledge, as causal interpretations of higher level processes are still often framed in terms of one pathway or another (or worse yet, the activity

[1] Note that material and formal causes are synchronic (i.e., a set of substances and their specific configuration at a given instant), whereas efficient and final causes are diachronic (i.e., a mechanism of action and its consequences over time) [Salthe, 2005a].

of one gene or another), and do not take into consideration causal ramifications entailed by the complexity of the system. Pathways are often described as being not only causally necessary, but *sufficient*, despite the fact that they are linked to (and thus contingent upon the context and activity of) myriad other pathways. A corollary of this assumption, which serves as the basis for the pharmaceutical-agricultural industry, is that blockade of a pathway can suffice as treatment for a pathological state that is dependent on it.

Reductionism developed as an extension of Naturalism, which rejects irrational or supernatural models of reality that are not empirically testable, such as those advanced by religion and mysticism. For reasons that go beyond the scope of this essay (see [Ulanowicz, 1997; 2009] for discussions), the scientific enterprise that grew out of the Enlightenment used Newton's *Principia* and its assumption of causal closure to equate Naturalism with the Mechanicism of Descartes, giving birth to what became the modern caricature of Reductionism; in the process, permissible causal explanations became largely relegated to the Aristotelian categories of material and efficient.[2] Thus, with the exception of workers at the fringes of cosmology and macro-evolutionary biology (fields that some consider not to be real 'science', as they do not permit laboratory experimentation), many modern Naturalists tend to shun phenomenological approaches that seek formal and final causes, sometimes even equating them with mysticism or religion (which, ironically, is something that has not been lost on the religion-motivated sophists working to place 'intelligent design' creationism on equal footing with evolutionary biology). However, Mechanicism is not the same thing as Naturalism. As I will argue, phenomenological aspects of causality provide important insights into the nature of empirical reality. Thus, one need not abandon principled Naturalism in order to adopt a full-spectrum and yet completely rational view of causality, a view that overcomes the blinders (and blunders) of Reductionism and its preoccupation with material and efficient causes. In fact, such a view is necessitated by the empirical behavior of non-linear complex systems.

It should be self-evident that fewer untestable assumptions make for better inquiry. Following others [Salthe, 1993; Ulanowicz, 1997], the thesis here is that the empirical facts of thermodynamics, information, complexity, nonlinearity, and historical contingency provide a striking insight into the nature of causality, one that allows us to discard many of the assumptions commonly held as prerequisites for scientific Naturalism. Ironically, this insight is implicit in our common use of language, and was at one time more widely appreciated. That is, predictable aspects of causality arise by *development*. This perspective, referred to here as 'Developmentalism' [Salthe, 2000], represents a synthesis that resolves a number of perennial dialectics on causality, including reductionism/holism, internal-

[2]Salthe (personal communication) notes that in Reductionistic science "formal causes were covertly present in models and equations, while final causes were covertly alive in variational principles such as least-action and the Second Law of Thermodynamics." It can also be argued that Natural Selection, which is commonly assumed to be the agency that establishes functional needs, acts as a final cause: that is, it is the source of Monod's 'Necessity' [Monod, 1972].

ism/externalism (orthogenesis/adaptation), and mechanicism/stochasticism (predetermination/randomness or stasis/change). In my opinion, widespread adoption of a developmentalist perspective would go a long way toward remedying the ecological and social maladies associated with the monstrous caricature that Reductionistic science has become.

2 THERMODYNAMICS AND INFORMATION

2.1 The Second Law as first principle

At the outset it seems quite reasonable to posit that everything that happens in nature is an expression energy flow. The laws of thermodynamics can therefore be taken as first principles in the Naturalist discourse on causality. The First Law says that the amount of energy in the universe is fixed; nothing comes from nothing, and nothing happens without balancing the equation. Therefore, energy flow occurs by transformation (as opposed to creation or destruction) associated with work and dissipation into heat. Work is governed by the Second Law, which says that it cannot be 100% efficient; hence, some of the energy used by work is transformed to configurations that cannot support that particular type of work. A thesis of this essay is that we can use the Second Law as first principle to deduce a good deal about the fundamental nature of causality. In turn, this allows us to discard many conventional assumptions, thereby increasing the depth, range, and accuracy of scientific inquiry.

In systems that are far from thermodynamic equilibrium (i.e., all natural systems), many things occur that seem at first glance to violate the Second Law, as epitomized by life on earth. While this has been one of the keystone arguments for anti-evolution sophistry, it is an illusion that disappears when one considers that life on earth is part of the Solar system, which provides a vast excess of useful energy. Moreover, there are countless local manifestations of the Second Law in living systems, as epitomized by an organism's continuous generation of metabolic waste. Nevertheless, recurrent arguments about life in reference to the Second Law suggest that there is much confusion about what the Law actually means. On the one hand, the meaning is simple: work is always inefficient, so anything that happens converts some of the energy that made it happen into waste (i.e., a less useful form of energy). On the other hand, it's not so simple; for now we must define 'waste', which is a contextually relative concept (one man's trash is, after all, another man's treasure). It can be argued that the latter fact is the basis for ecology, which in turn is the basis for life. For present purposes suffice it to say that, *contra* superficial appearances, life is essentially a complex manifestation of the Second Law applied to the flow of solar and geothermal energy over the surface of the earth. Or, said another way: life is caused *formally* by the intersection of earth with the energy gradients established by the sun and geothermal sources, and *finally* by the Second Law of Thermodynamics [Salthe, 2005b; Hoelzer *et al.*, 2006].

The fact that life fits comfortably with (and indeed, is explained by) the Second Law still begs the question of how this relates to the order and complex organization of living systems. In turning to this question, we introduce another key concept: *information.*

2.2 Information and its relationship to organization

Information is another intuitively simple concept that, having been defined in different ways, has been the subject of many confusing discussions. For the sake of the present argument we will limit ourselves to a simple definition that fits our intuition: *information is that which brings about a reduction of uncertainty or indeterminacy.* Information is a manifestation of asymmetry (i.e., reduction in degrees of freedom, mutuality between events, or expectation of one thing given another), and is the basis of predictability, the holy grail of the scientific enterprise. This definition is captured in the following quantitative expression:

(1) $I = H - H_r$

where H is the *a priori* uncertainty of an event (or indeterminacy of a system), and H_r is the residual uncertainty that remains following measurement of the event (or the residual indeterminacy following development of a system; see below). This allows us to further posit that organization is a source of information, since by definition it represents a state of relative determinacy. It follows that the act or process of organization creates information.

2.3 Entropy, information capacity, and complexity

As discussed above, the Second Law requires that some percentage of the energy used for any type of work be converted to waste energy that no longer has the capacity to support that particular type of work. The energy thus dissipated is quantified as *entropy.* Entropy produced as a result of work is a measure of the amount of randomization (or disorder) created by that work, which is the opposite of our definition of information. Whereas information measures a reduction of indeterminacy, entropy measures an increase in indeterminacy.

Returning to our definition of information (equation 1), we can see that the maximum amount of information that can be manifested by any event equals the maximum *a priori* uncertainty regarding that event. We can quantify this limit as the information capacity:

(2) $I_{max} = H$

In other words, the maximum amount of information that can be obtained by measurement is equal to the uncertainty associated with (indeterminacy of) the thing being measured. Therefore, by increasing indeterminacy, entropy production increases information capacity. This is reflected in the fact that Botzmann's equation for physical entropy ($S = \mathrm{k}\ln W$) is implicit in Shannon's equation for information capacity:

(3) $H = -k\Sigma p_i \log p_i$

where p_i is the probability of an event (which becomes smaller as H increases).

In seeking to understand the causal basis of empirical reality, science aims to increase predictability. It does so by interpreting[3] information obtained by way of measurement. This signifies an intimate relationship between information and causality; and indeed, it has been argued that causality is essentially a manifestation of information transfer [Collier, 1999]. If information is in turn taken to be organizational constraint [Kauffman *et al.*, 2007], then we might further surmise that *causality implies organization.* While this proposal might ruffle Reductionistic feathers, it is entirely consistent with recent formulations of ontogenetic causation in terms of organizationally constrained flow of genetic information (see e.g. [Davidson, 2006]).

This brings us to the subject of complexity, which like information, has been defined in many ways. Here we keep things as simple as possible by defining complexity as a measure of diversity. That is, the complexity of anything is simply the number of *different* parts (or events) that constitute that particular thing. Of course this can be a difficult concept to apply when the thing itself or its component parts are not easily distinguished, as is so often the case in living systems [Cumming and Collier, 2005]. However, there are some unambiguous and hence heuristically useful examples, with perhaps the best being the sequence complexity of nucleotides constituting the genetic material of an organism. The complexity of genomic nucleotide sequence is defined as the number of nucleotides within the total length of non-redundant (i.e., non-repetitive) sequence [Davidson, 1986]. This example is straightforward because it is only one dimensional.[4] But the same principle applies to higher dimensionalities.

Our definition of complexity fits well with our definitions of information and entropy, as it indicates that higher levels of complexity are invariably associated with greater *a priori* uncertainty, and hence higher information capacity. Thus, entropy and information capacity are directly related to complexity.

We now have a set of first principles that allow us to deduce the following basic premise: *in natural systems that are far from thermodynamic equilibrium, informational entropy production associated with work increases complexity and hence information capacity, the indeterminacy of which can be reduced through organization.* With this premise we can begin to investigate the nature of nonrandom causality in complex systems. To do so we must first develop some tools for dealing with the key concept of *organization.*

[3]The concept of interpretation (and its derivative *meaning*) is a key one that connects this discussion to the discourse on Semiotics, and to the fact that all knowledge is ultimately subjective. This is an important component of developmentalism that is not duly considered here. For discussions see Salthe [1993] and Hoffmeyer [1997].

[4]Note that because of the informed linkage between gene expression and phenotype, stage-specific differences in expressed sequence complexity can be used as a quantitative measure of developmental specification in ontogeny [Coffman, 2006].

3 ORGANIZATION: HIERARCHIES, FLOW NETWORKS AND DISSIPATIVE STRUCTURES

3.1 Scalar and specification hierarchies

Hierarchies, which are nested sets of inclusiveness, provide a contextual framework for modeling organization in terms of both scale (extensive spatiotemporal dimension) and specification (intensive integration).

Scalar hierarchies provide spatiotemporal context. We begin by noting that for any object of our attention this context occupies scales that are both vastly larger and vastly smaller than that object. Although at first glance this might seem to disallow comprehensive causal analysis, we further note that events that take place at scales distantly removed from that at which the object exists generally have less impact on the object than events at more proximal scales. For example, if the object of attention is an individual organism we can safely say that that organism is not likely to be affected by supernovas in distant galaxies or by quantum events within individual atoms inside one of its cells. On the other hand, the organism is quite often affected by events that occur at scales near that at which it exists, e.g. a hurricane or a cancerous growth. We can conclude that, practically speaking, adequate causal description of an organization in terms of scalar hierarchy should at least include the higher level context of its immediate environment and the lower level context of its internal milieu. For an organism, this would be, respectively, the environment within which it lives and its constituent cells (i.e.: [environment[organism[cells]]]). Thus, organization at any given 'focal' level involves both higher level 'boundary' conditions with slower dynamics and lower level 'initiating' conditions with faster dynamics [Salthe, 1993]. The scalar hierarchy describes the effective spatiotemporal domain of *dynamic processes* ("active initiators of a change"), which in terms of the Aristotelian causal categories involves material, efficient, and formal causes. For any given event in the material universe, boundary conditions establish formal causes (circumstances) that answer the question 'why (or how) did this happen?', whereas initiating conditions provide efficient and material causes (mechanisms and substances) that answer the question 'what was the underlying process?'. For example, if the organization in question is a human being with lung cancer, we might say that a formal cause of that is cancer cigarette smoke (a boundary condition established by the environmental context of the smoker), whereas an efficient cause is chronic inflammation of the lung and a material cause is oncogenically mutated DNA (initiating conditions provided by lung cell physiology and genetics, respectively).

Specification hierarchies provide integrative context. An example is the hierarchy of sciences {Physics{Chemistry{Biology}}} alluded to above. Inclusiveness within a specification hierarchy is a function of generality. Thus, all biological phenomena are also chemical phenomena, and all chemical phenomena are also physical phenomena, but not the other way around. The attributes that distinguish more highly specified sets within a specification hierarchy entail an increase in information. Biology is a specifically informed variant of chemistry, and chemistry is a specifically informed variant of physics. The information accrued at each ascending level of a specification hierarchy provides a set of rules that pertain specifically to that variant as well as all of its more highly specified sub-variants. Ascending the hierarchy of sciences we encounter first the rules of physics such as Gravity and the Second Law of Thermodynamics; then the rules of chemistry summarized in the Periodic Table of the Elements; and finally the varied rules of biology encoded genetically as regulatory networks that are only now being deciphered. The specification hierarchy describes levels (or stages) of *structural integration* ("inactive nodes on a law-like implication chain"), which models development [Salthe, 1993] and evokes material, formal and final causes. Thus anything in the material universe can be described in terms of generic material causes that answer the question 'what is this composed of?', as well as more specified formal causes (contextual relationships) and final causes (self-organizing 'attractors'[5]) that answer the question 'why does this particular thing exist?'. For example, in the example of the lung cancer, a final cause of the cancer might be taken to be the self-organizing attractor for the cancerous (proliferative, metastatic) phenotype and a formal cause is the gene regulatory network architecture that gives rise to that attractor [Huang and Ingber, 2006], whereas material causes are the DNA within which that particular regulatory network architecture is encoded as well as all of the other substances that constitute the cancer cell and its environment.

In sum, by using scalar and specification hierarchies to describe organization we avoid the pitfall of caricatured Reductionism while reviving Aristotelian causal categories. We reinforce the notion that adequate causal explanations must take into account hierarchical context, as illustrated by our example of lung cancer. While effective short-term treatments for particular instances of this pathology may be found by a pharmaceutical industry that focuses exclusively on its material or efficient causes, these successes may be outweighed in the long-term by the negative consequences associated with ignored formal causes (e.g., carcinogenic industrial pollution, complex developmental gene regulatory network architecture) and final causes (e.g., self-organizing attractors for secondary pathologies that may be induced by the treatment).

3.2 Flow networks and dissipative structures

Hierarchy provides a conceptual tool for describing both the extensive and intensive causal attributes of organization, but we need to relate organization to our first principles of thermodynamics and information. To do so we return to our primary postulate regarding causality: anything that happens is an expression of energy flow associated with work. Nonrandom causality emerges from organized patterns of energy flow. Science has provided us with abundant knowledge about such patterns, which often occur in the form of networks. In ecology, energy flows through trophic networks; in biochemistry, it flows through metabolic networks;

[5]See discussion below in section 4.2 on 'growth, development, and senescence'.

and in physical chemistry, it flows through sub-atomic and molecular networks. Generalizing from these examples, we can state that formally causal organization implies a *flow network*.

By definition, flow networks provide conduits for the flow of energy, which according to the Second Law can only occur in one direction (i.e., down a gradient passing from more to less useful). Simple thermodynamic systems that are near equilibrium (e.g., the classic mixing of two segregated gases) indicate that the entropy produced by equilibration is associated with the progressive disordering of the initial structure as energy flows down its gradient. As discussed above, this seems to be inconsistent with the fact that living systems maintain high levels of structured order. However, unlike the simple gas mixing experiment, life exists within an energy gradient that is far from thermodynamic equilibrium. Moreover, unlike the segregated gases, the structure of life is highly dynamic; it is a form of work maintained by a continuous flow of energy, and its constituent materials turn over more or less continuously. In the process, a good deal of entropy is produced and 'dissipated', as embodied for example by metabolic waste. Such *dissipative structures* [Prigogine, 1978] are not limited to biology; they are commonplace in complex systems that are far from thermodynamic equilibrium, with abiotic examples including tornadoes, tsunamis, and fires.

It is apparent that dissipative structures, like flow networks, are conduits for energy flow in circumstances that are far from thermodynamic equilibrium. But they are further distinguished by the fact that they have recognizable boundaries (which can nonetheless be difficult to define [Cumming and Collier, 2005]), setting them apart as distinct focal level organizations amenable to causal description in terms of scalar and specification hierarchies. Hence, the thesis developed thus far can be summarized and extended with following proposition: *causality is a manifestation of work constrained by information provided by organized flow networks that sometimes produce dissipative structures having distinct boundaries that enable hierarchical modeling.* Which raises the questions: how do organized flow networks come into existence, and from whence come the defining boundaries?

4 DEVELOPMENTAL ASCENDENCY AND EVOLUTIONARY CHANGE

4.1 Stochasticity, contingency, and mutuality

Whether events occur at random or are predetermined is a perennial metaphysical question that continues to vex scientists, philosophers, and theologians alike. It has been over a century since quantum physics established that stochasticity is an empirical reality and not a matter of insufficient knowledge. And although quantum events occur at very small scales, combinometric calculations based on the complexity of the universe indicate that unique (i.e. random) events are likely to be commonplace at all scales [Ulanowicz, 2009]. Thus, we can conclude that many events are not predetermined (at least, not in any way that is scientifically tractable).

And yet, much of reality is predictable. Indeed, Newton formulated his *Principia* with the premise that he was elucidating the lawful manner by which God determines the fate of the universe [Ulanowicz, 1997], and Einstein himself famously refused to believe that God leaves fate to chance. Even though Irreversible Thermodynamics, Quantum Physics and the Theory of Relativity undermine many Newtonian precepts, most scientists still believe that the predictable aspects of reality are mechanically determined; and to be sure, many things do indeed behave quite mechanically. But is such behavior finally attributable to universal Laws that are in essence Platonic ideals refractory to further explanation? Or is it a historical contingency that emerged in an intelligible way? We might begin to answer this question by considering that contingency is certainly an empirical reality: although many events occur at random, they do not occur in isolation and can therefore interact causally with other events, thereby constituting *processes* having historical trajectories.

Stochasticity and contingency are widely accepted as being self-evident, but by themselves they do not reveal the fundamental nature of causality and predictability. To get at the heart of the matter we have to consider a third empirical reality, that of *mutuality*: the fact that contingency often finds closure in configurations of transactional or mutually reinforcing (and hence dynamically nonlinear) processes. Physical forces such as gravity are cosmological manifestations of this reality. At higher levels of specification, chemical and biological systems manifest mutuality in the form of positive feedback or 'autocatalytic' cycles. Let us grant that such mutuality can initially emerge from stochastic events at any level in the material universe (there being no *a priori* reason to believe that it shouldn't), and examine the causal implications.

In an autocatalytic cycle (e.g. the reductive tricaboxylic acid or rTCA cycle, an example that is especially pertinent given that it establishes a bridge from chemistry to biology [Smith and Morowitz, 2004]), the dissipative activity of a process (e.g., the breakdown of citrate into acetate and oxaloacetate in the rTCA cycle) contributes to the activity of another process, which contributes to subsequent processes whose dissipative activity ultimately feeds back to contribute to the activity of the first process. Thus, each process within an autocatalytic cycle is energetically rewarded by the cycle. Note that this reward represents energy that each contributing process would otherwise have dissipated as entropy. Thus, a key attribute of an autocatalytic cycle is that it decreases the entropic 'wastefulness' of each contributing process.

If a process that contributes to an autocatalytic cycle becomes more productive (that is, manifests an increase in its energy throughput or activity), it is further rewarded via the cycle, reinforcing that particular process as well as the other processes within the cycle. Conversely, if a process within the cycle decreases its productivity, it receives less energy from the cycle and is thereby weakened. Although this could in principle diminish the activity of the cycle as a whole, this would not be the case if there are other 'redundant' processes within the cycle that can take up the slack by carrying out a similar activity (as might be expected in systems of high complexity); in so doing, the latter processes would be rewarded by the cycle. Thus, *sufficiently complex* autocatalytic cycles are not diminished by loss of productivity from any one of their constituent processes; whereas the underperforming process is always punished by a loss of energy. Because of this fact, *an autocatalytic cycle has an inexorable tendency to grow when energy resources are abundant, and to act as an agency for selection when resources become limiting* [Ulanowicz, 1986, 1997]. With these properties in mind, we can now return to the questions posed above concerning the emergence of spatiotemporally bounded organizations that are the basis of nonrandom causality.

4.2 Growth, development, and senescence

Suppose that, within a complex material (physical, chemical, or biological) system that is far from thermodynamic equilibrium, stochastically occurring events produce a configuration of processes that constitute an autocalalytic cycle. Initially, the incipient cycle would be controlled from the 'bottom-up' [Coffman, 2006]; that is, it would be an outcome of local (low-level) events and the processes contingent thereupon. In such an *immature* state, the cycle may be difficult to recognize within its environmental milieu. However, with a sufficient source of energy, once established the cycle would begin to grow; as a result, those processes that are part of the cycle will gradually come to dominate their environment.[6] At the same time, informational entropy production associated with growth will substantially increase the complexity (indeterminate information capacity) of the system, potentiating the emergence of new processes that increase the system's development capacity and degrees of freedom (Figure 1). Within this milieu, an increasingly complex array of processes will contribute more or less redundantly (and more or less effectively) to the cycle, endowing it with a high level of indeterminacy and hence lack of clear definition.

Now consider what will happen when growth of the cycle begins to reach the limit of its energy resources. At this point the processes that are most productive are favored over those redundant processes that are less productive; that is, competition for the limited resources intensifies, and the cycle becomes an *agency for selection*. As a result of competitive 'pruning' of processes, the most productive or 'fit' processes come to dominate the cycle, leading to a progressive reduction in redundancy (thereby reducing degrees of freedom, indeterminacy and information capacity), and concomitantly, and increase in definitive organization and information. This counteracts the continued increase in development capacity, causing the rate of this increase to decelerate (Figure 1). At this point, the cycle has developed to a state of *maturity* (a process also known as 'self-organization'), and would be controlling events from the 'top-down' through its *selective* agency [Coffman,

[6]Growth can be measured by considering that energy flow entails the transfer of material from one entity to another, or the transformation of one entity into another. Thus, if T_{ij} quantifies the transfer or transformation from entity i to entity j, then the total energy throughput of a system is: $T = \Sigma_{i,j} T_{ij}$. Growth manifests as an increase in T.

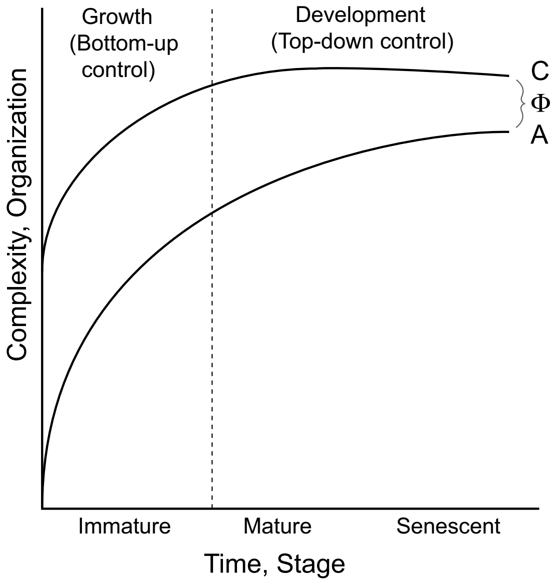


Figure 1. The canonical pattern of growth and development engendered by autocatalytic cycles. The increase in complexity and organization associated with the respective work functions 'development capacity' (C) and 'ascendency' (A) are shown. The residual indeterminate complexity, or 'overhead' (Φ) is the difference between the two curves. (Modified from Ulanowicz [1997] and Coffman [2006]; see section 4.2, footnote 7 for additional explanation.)

2006]; that is, it would represent a specified organization that establishes necessities of final causality (which in ecology are the determinants of fitness). This can be quantified as product of the cycle's growth and development, a work function that Robert Ulanowicz refers to as *Ascendency*.[7]

Although the forgoing is based on the premise that growth and development are a consequence of positive feedback cycles, it will be noted that antagonistic

[7]Ascendency can be measured by considering that development increases some joint probabilities at the expense of others, thereby increasing the extent to which a subset of conditional probabilities informs the system. This process can be quantified using the concept of information as defined above (equation 1), in which case I represents the developmental reduction in a system's indeterminacy, i.e., its *level of specification*. H quantifies the total complexity of the undetermined system prior to the onset of development, while H_r quantifies the system's residual undetermined complexity following some development [Coffman, 2006]. Ascendency then scales the information developed within a system to its growth, i.e. $A = T \times I$, giving a work function representing both the extensive scale and intensive organization of the system [Ulanowicz, 1986; 1997]. Scaling the total complexity to growth gives the system's development capacity, $C = T \times H$, while similarly scaling the residual indeterminacy gives the system's overhead, $\Phi = T \times H_r$, which sets the limit on its remaining developmental potential (including its ability to recover from perturbations; see Figure 1).

interactions (including *negative* feedback, i.e., inhibition of a process by its own activity) are a common feature of highly specified flow networks (e.g. genetic regulatory networks [Davidson, 2006]). In general, we can say that the consequence of such antagonism is to increase organization (information, definition) and sharpen boundaries. Since the selective agency of positive feedback favors the development of organization, it stands to reason that antagonistic interactions that serve to increase such organization will also play an important role in development.

The total set of positive (activating) and negative (inhibitory) interactions within a system can be described as a 'regulatory network' that constrains the development of information in a logical (and hence predictable) manner. An important property of such networks is that they often are the source of self-organizing 'attractors'. An attractor, which is a stable state toward which a developmental trajectory is inexorably drawn (e.g., the phenotype of an organism), is established by the regulatory network architecture; that is, by the set of logical rules (positive and negative interactions) that regulate the development of information within a self-organizing system [Huang and Ingber, 2006]. In essence, an attractor is a final cause accessed by the regulatory network, which is in turn a formal cause established by organization that developed via the selective agency of autocatalytic cycles.

At this point the astute reader will undoubtedly be thinking that the argument is beginning to sound suspiciously circular. On the one hand, autocatalytic cycles drive development of organization that takes the form of regulatory networks; on the other hand, regulatory networks are the source of attractors that control development. It is the 'chicken-or-the-egg' conundrum; how can both propositions be true? As in the case of the 'chicken-or-the egg' question, the answer can be found by accounting for hierarchy and historical contingency. Thus, in the specification hierarchy {amniote{bird{chicken}}}, the egg clearly comes first, and indeed was a pre-requisite for the evolutionary emergence of chickens. In terms of both scalar and specification hierarchies, systems develop *within* systems that evolve over time, an iterative, bootstrapping process wherein the organization that emerges developmentally within a given system is constrained by the information (regulatory networks) that developed previously in a larger, more general system. Over time this process gives rise to a specification hierarchy, as epitomized by that modeled in the hierarchy of sciences: {Physics{Chemistry{Biology}}}.

But we still haven't adequately described the initiation of developmental systems, beyond postulating that they can emerge as an outcome of stochastic events. Although this is undoubtedly true, given what we have just argued it is clear that many events occur as part of a developmentally constrained process, and are hence not random. Indeed, development would seem to lead inexorably away from randomness, ultimately producing a state of (homeo)stasis that precludes *change* and the initiation of new development. To understand why this is not so, we need to continue our investigation of the consequences of the developmental process.

Toward this end, we recall that autocatalytic cycles develop organization (and hence information) by selectively pruning away redundant processes that contribute less effectively to the cycle. The result is a progressive sequestration of resources into a smaller fraction of the total complexity; that is, a subset of processes emerges at the expense of others, which can ultimately lead to a loss of complexity. Nevertheless, the Second Law of Thermodynamics cannot be broken; although the efficiency of work is increased via development, it still must fall short of 100%, and entropy production continues. Recall that entropy production manifests as an increase in indeterminacy, i.e., an increase in disordered complexity. Thus, although developmental Ascendency manifests information as a reduction of indeterminacy, there is always an undercurrent of entropy production that counters this tendency, producing *overhead* that continues to increase the developmental capacity of the system [Ulanowicz, 1986, 1997]. And within this overhead, additional development can occur at lower levels of scale and/or higher levels of specification. So after a system matures, development continues; but not with a further reduction of complexity beyond that which accompanied the system's maturation, but now with an inexorable increase in the complexity of the system' organization (i.e., its *complicatedness*). This ultimate developmental trajectory into *senescence* continues to a limit imposed by the finitude of both useful energy resources and the potential for mutuality.

Because of their burden of organized complexity, senescent systems are paradoxically both stable and unstable. They are internally stable by virtue of their exquisitely baroque organization — the indeterminacy (overhead) associated with entropy production having been continuously reduced over time (converted to ascendency) by nested developmental systems. Eventually all of the energy flowing through the system is used just to maintain the system, at which point it becomes unstable to external perturbations and hence vulnerable to catastrophic collapse, as it lacks sufficient overhead to adapt. Therefore, although they can persist for relatively long periods, senescent systems inevitably collapse and are recycled [Salthe, 1993; Ulanowicz, 1997].

The collapse and recycling of senescent systems does not necessarily entail a complete loss of information however. In fact, as noted above senescent systems are a substrate that affords the maturation of processes occurring at lower scalar levels and/or higher levels of specification, and some of these may acquire a level of autonomy sufficient to allow them to persist and initiate the growth and development of new, perhaps even larger and more highly specified systems from the remains of the senescent system following its collapse [Gunderson and Holling, 2002]. It can be argued that this is exactly what happens in macroevolutionary transitions associated with ecological recovery following mass extinction, at ontogenetic transitions such as birth or metamorphosis, and during cellular mitosis [Coffman, 2006; Salthe, 1993]; indeed, it may even characterize the birth of our universe [Carroll, 2006]. In each of these cases, since the emergence of the new initiating systems was initially contingent upon the development of their 'parent' systems into senescence, it stands to reason that senescence represents a condition that is often ripe with the seeds of creative or progressive change.

4.3 Historical contingency, individuation, and creative evolution

Thus far, we can summarize our thesis on causality as follows: *in complex systems that are far from thermodynamic equilibrium, autocatalytic cycles arising by virtue of mutuality create information by selectively organizing a system, which in turn establishes attractors that constrain the flow of energy toward producing specific outcomes.* We have further proposed that this process ultimately leads to a senescent state that lays the groundwork for the emergence of new systems that develop even higher levels of specification. Here we extend this proposition by considering the development of complex systems within the broader context of a deeply stochastic universe.

The scientific discourse on causality tends to shy away from randomness because it is by definition refractory to causal explanation and predictability. And yet the complexity of the universe ensures that chance events are ubiquitous occurrences at all levels of scale. It seems quite possible that any certainty and predictability manifested by the universe (which is taken here to be an outcome of development) is balanced or even outweighed by its uncertainty and unpredictability. Given this, random chance makes an inescapable, irreducible contribution to causality.

It is clear that unique, unpredictable events can produce dramatic changes in the trajectory of a developing system. This leads to the proposal that evolution is the historical trajectory of change that occurs in and among developing systems as a result of capricious happenstance; in other words, evolution leads to the *individuation* of developing systems [Salthe, 1993]. Given that highly specified systems (e.g., embryos) are rigidly constrained by historically developed information (e.g., genes and their regulatory networks), it stands to reason that the evolution of novelty — the creation of *new* systems — is entirely dependent on the randomness of the universe. For organisms that segregate their germ lines early in ontogeny (which are by no means the majority; [Buss, 1987]) the implication is that the evolution of new species is initially dependent on random genetic mutations that occur within the germ line, an idea that is one of the foundations of neo-Darwinian evolutionary biology. For ecosystems, the implication is that evolution of new regimes (and their populating species) is dependent on random perturbations that cause the collapse of senescent systems, e.g. as occurs with mass extinctions [Sole *et al.*, 2002].

At any given level in a specification hierarchy, the evolutionary individuation of developing systems by virtue of historical contingency would be expected to produce unique variants of a generic process, thereby generating even higher levels of specification. This brings to mind the Platonic notion that things in the material world are 'flawed' manifestations of universal ideals. In Aristotelian terms, we might say that stochastic events constrained by the organizational context of a specified flow network provide circumstances (formal causes) that generate specific developmental attractors (final causes) which entrain the effects of physical-chemical mechanisms (material and efficient causes).

Viewed through the lens of this interpretive construct, complex abiotic flow networks and the dissipative structures they generate appear as predictably recurrent but unpredictably individuated instantiations of generic physico-chemical processes. Examples include galaxies, hurricanes, tornadoes, whirlpools, and other vortices produced by gravitational or thermal energy gradients. In contrast, biological systems present a somewhat different appearance by virtue of *semiosis* [Hoffmeyer, 1997]: the genetically-endowed ability to interpret, record, reproduce and transmit developmental information. In biology the evolutionary individuation of developing systems produces continuous branching lineages that retain information deeply rooted in history, all the while producing more highly specified sub-lineages. Thus, whereas final causality in abiotic systems can at most be generally attributed to the universal laws of physics and chemistry, in biological systems it can be specifically attributed to the highly developed, semiotically *interpretive* developmental-genetic regulatory networks that give rise to unique variants within each lineage.

5 DEVELOPMENTALISM: A METAPHYSICAL ALTERNATIVE TO REDUCTIONISM

5.1 A resolution of causal dialectics

I have argued that non-random causality is essentially information that accrues in complex systems via the self-organizing development engendered by mutuality. This perspective, which I refer to as *Developmentalism*, provides a naturalistic metaphysic that is an alternative to Reductionism. Importantly, it is a metaphysic that is consonant with emerging models of cosmology, which suggest that 'laws of physics' are no more (and no less) than organizational constraints created by the early development of our universe. If events had played out in a different way during the initial inflation of the universe (and there is no *a priori* reason to believe that they couldn't have, and perhaps actually have in alternate universes; that is, within the *multiverse* as a whole), the universe and its physical Laws would have developed quite differently [Carroll, 2006].

Developmentalism provides a synthesis that resolves many deeply rooted, perennial dialectics concerning causality. The thesis of reductionism that holds causality to be reducible to lower-level mechanisms is countered by the antithesis of holism (or organicism), which holds that causality can only be understood in terms of organizational constraints. The developmentalist synthesis argues that lower level causality predominates in immature systems, whereas higher level constraints are more important in mature systems. The thesis of internalism (orthogenesis) that holds internal developmental constraint (or drive [Arthur, 2001]) to be the predominant factor in evolution is countered by the antithesis of externalism (adaptation), which holds environmental factors to be of paramount importance [Gould, 2002] (the 'nature/nurture dichotomy). Developmentalism suggests that 'developmental constraint' is not a concept whose application is limited to the narrow ontogenetic focus of developmental biology, and that the adaptationist concerns of externalism can often (if not always) be related to higher level developmental (organizational) constraints viewed with a more expansive internalism [Salthe, 1993]. Finally, as discussed above, the thesis of mechanicism (Platonic stasis or certainty) that holds events to be determined by universal laws is countered by the antithesis of stochasticism (Heraclitusian flux or uncertainty), which holds that events occur at random. The synthesis of developmentalism maintains that in an uncertain universe manifesting both randomness and contingency, mutuality promotes the development of systems that create information bestowing a progressive increase in mechanical certainty.

5.2 Emerging developmentalism in the biological sciences

Although many biologists continue to work with the assumption that an adequate causal explanatory framework for biological systems will emerge as a 'sum of the parts' through reductionistic studies that are focused on single gene functions and linear pathways, an increasing number are beginning to appreciate that adequate causal explanation requires an understanding of irreducible, emergent properties of whole systems and networks and the attractors that they generate (e.g., [Bower and Bolouri, 2001; Davidson *et al.*, 2003; Davidson, 2006; Huang and Ingber, 2006; Chen *et al.*, 2007; Reid, 2007]). Thus, the caricature of Reductionism described in the Introduction is becoming more and more of a straw man due to the development of a new science of 'systems biology' [Auffray et al., 2003], which represents a paradigm shift growing out of the working realm of reductionistic science, but which is generally consistent with the developmentalist perspective advocated here.

Perhaps the most important contribution of developmentalism to biological science is that it undermines old assumptions that have imposed blinders on our view of reality. One such assumption is that of *uniformitarianism*, which asserts that processes occur today as they did in the past. This proposition was advanced by Charles Lyell as a means to erect a historical view of geology, and adopted by Charles Darwin as a way of doing the same for biology [Gould, 2002]. Like many Newtonian assumptions, the assumption of uniformitarianism was initially useful in that it provided a tool for opening a window that was previously closed (in this case, on earth's pre-history). But in the same way that irreversible thermodynamics, quantum mechanics, and relativity exposed the limits of the Newtonian assumptions, developmentalism exposes limits of uniformitarianism. This is particularly important for the discourses on evolution and evolutionary developmental biology ('evodevo'). Developmentalism views the genetic regulatory architectures that control the ontogeny of modern organisms as being highly specified systems that have developed to maturity through evolution [Coffman, 2006]. From this perspective it stands to reason that the causal dominance of genetics in ontogenetic programming and evolutionary dynamics would have been somewhat weaker (i.e., less mature) during the primordial history of life than it is now. Thus, it has been suggested that there was a 'pre-Mendelian, pre-Darwinian' period in the history of life, when animal ontongeny was less genetically constrained (i.e., less canalized) and more directly responsive to environmental factors [Newman, 2005] (note however that some of the core genetic regulatory circuitry, which is conserved throughout animals or clades of animals, would already have evolved [Davidson and Erwin, 2006]). Moreover, developmentalism suggests that morphological homoplasy and convergent evolution would be far more common than is generally assumed owing to the existence of (perhaps cryptic) developmental or structural attractors generated by ecological organization (an example would be the convergent evolution of morphologically similar feeding larva during the late pre-Cambrian [Dunn *et al.*, 2007]). This of course does nothing to alleviate the high level of uncertainty that characterizes the many heated discussions of homology.

Another significant contribution of developmentalism is that by acknowledging the primacy of mutuality, it overcomes the inability of Darwinism (with its primary focus on competition) to deal with life's origin, while at the same time providing a contextual explanation for the 'struggle for existence' underlying Darwinian selection [Coffman and Ulanowicz, 2007; Ulanowicz, 2009]. The ubiquity of mutuality in complex systems provides a natural bridge between chemistry and biology. Although modern biology is highly dependent on the function of nucleic acids as information carriers, developmentalism suggests that this function would have been selected by a pre-existing autocatalytic cycle. In other words, a nucleic acid produced (for example) as a byproduct of an autocatalytic reductive TCA cycle [Smith and Morowitz, 2004] would become an integral part of the cycle if it happened to both embody and reproduce some of the mutual information manifested by the cycle's ascendency, as for example could occur with catalytic RNAs [Copley *et al.*, 2007].

5.3 A world-view promoting ecology and sustainability

At present our planet is facing converging ecological and economic crises on a scale unprecedented in human history, and there is a high level of uncertainty regarding their resolution and the fate of civilization. Most, if not all of the problem is attributable to industrial economies that developed by virtue of energy flowing from fossil fuels. It seems reasonable to posit that the ascent of Industrialization was in no small way facilitated by Science's turn to mechanicistic Reductionism, which (in contrast to Developmentalism) is an inherently non-ecological construct. Although the central assumption underlying the modern 'global-economy' — that growth based on fossil fuel can continue *ad infinitum* — seems absurd from the developmentalist point-of-view, it is comfortably ignored by the narrow focus of Reductionism; indeed, it is quite remarkable how few learned people are questioning this assumption. And, whereas Developmentalism places a positive value on undeveloped indeterminacy ('overhead'), viewing it as a critical component of adaptability and hence sustainability [Ulanowicz *et al.*, 2009], Reductionism undervalues such indeterminacy, either viewing it as an obstacle to efficiency or 'progress', ignoring it, or denying that it exists. Predictably, industries founded on the latter approach are repeatedly surprised by unforeseen negative consequences of their labors (e.g., harmful side effects of drugs such as Vioxx or pesticides such as DDT, to mention but two of a long list of examples).

Nonetheless, there is a growing (albeit belated) level of appreciation for the interrelated problems of dwindling energy supplies caused by 'Hubbert's Peak', climate change, deteriorating ecosystems, and public health crises. From the perspective advocated here, all of these problems (as well as modern versions of the perennial problems associated with econo-demographics, social inequity, unrest and war) are formally caused by the *unregulated* development of a fossil-fuel-based global economy into maturity and beyond. On the one hand, the energy shortfall intensifies the competition associated with maturation of the global economy; on the other the informational entropy produced in the process (e.g. proliferation of new technologies, social disintegration, warfare, environmental pollution and greenhouse effects) continues to increase the indeterminate complexity of the system, creating opportunities within the industrial framework (e.g., for recycling operations and other means of increasing energy efficiency) that will allow for continued development of the global economy into senescence. Unfortunately, unlike the situation in natural systems that develop into long-term senescence within a stable environment built on energy flowing from an unlimited source (e.g., a rainforest), modern civilization (and all its specialized niches) will soon be faced with both an unstable environment and a rapid decline in energy flow. It is difficult to imagine any scenario that does not involve catastrophic social disintegration and loss of life.

Because there is still some residual (abeit decreasing) uncertainty associated with this pessimistic scenario, hope remains that civilization may yet access a more sustainable trajectory. What is needed is a maturation of our economic system, currently based on unrestricted growth fueled by consumerism (an approach that will rapidly lead to the terminal senescence described above), into a system that remains robust and sustainable by maintaining a reservoir of ecological and economic diversity that serves as developmental overhead [Ulanowicz *et al.*, 2009]. This will require a new socio-economic (and cultural) mentality, and probably cannot be accomplished without development of significant governmental (top-down) regulations that many find distasteful (and which are hence politically difficult). Nonetheless, there does appear to be an increasing awareness of the central role played by mutuality and autocatalytic cycles in causality. Thus, a developmentalist paradigm may eventually emerge as a lesson of history, and with it an understanding of the limits and consequences of growth as well as an ability to design more sustainable civilizations.

5.4 Toward a reconciliation of science and religion

Given the perennial skirmishes between science and religion over the past three centuries, as recently exemplified by the politically effective but ultimately futile attempts of certain religious groups to deny the fact of evolution, it is ironic that the mechanicistic Reductionism of modern science was in large part founded on Newton's *Principia*, a work that was inspired by religions faith [Ulanowicz, 1997; 2009]. Indeed, the fact that many atheists wield the deterministic perspective of Reductionism in defense of their belief in the non-existence of God, while theists tend to view religion as the only absolute in what would otherwise be an arbitrary universe, is quite telling: in an uncertain universe, absolute certainty holds a strong allure, whether it be dressed as science or religion.

Developmentalism offers a neutral, less polarizing ground that is conducive to both science and religion, but only for those who are willing to live with uncertainty. In accepting that random chance plays an irreducible and important role at all levels in the scalar hierarchy, developmentalism allows that many events in the material universe are not predictable, even for an entity with complete knowledge of current conditions and unlimited computational capacity. This leaves open the possibility of miracles as well as accidents. And, while developmentalism remains neutral concerning transcendental matters, it recognizes the ontological primacy of mutuality (and the evolutionarily creative development that it engenders), which has moral ramifications and might reasonably be viewed as a manifestation of 'God' for those who are so-inclined. Nevertheless, from a developmentalist perspective religion and science are both highly specified *humanistic* activities. While this fact does not invalidate either endeavor as a means for gleaning some aspects of truth, it strongly undermines literal interpretations and absolutist dogmas espoused by those claiming either scientific or religious authority.

ACKNOWLEDGEMENTS

I thank Bob Ulanowicz and Stan Salthe for helpful critiques of an initial draft of this essay, and Cliff Hooker for comments that resulted in its further improvement.

BIBLIOGRAPHY

[Arthur, 2001] W. Arthur. Developmental drive: an important determinant of the direction of phenotypic evolution. *Evol Dev* 3 (4), 271-8, 2001.

[Auffray *et al.*, 2003] C. Auffray, S. Imbeaud, M. Roux-Rouquie, and L. Hood. From functional genomics to systems biology: concepts and practices. *C R Biol* 326 (10-11), 879-92, 2003.

[Bower and Bolouri, 2001] J. M. Bower and H. Bolouri. Computational modeling of genetic and biochemical networks (. Ed.). MIT Press, Cambridge, MA, 2001.

[Buss, 1987] L. W. Buss. *The Evolution of Individuality.* Princeton University Press, Princeton, NJ, 1987.

[Carroll, 2006] S. M. Carroll. Is our Universe natural? *Nature* 440 (7088), 1132-6, 2006.

[Chen *et al.*, 2007] L. S. Chen, F. Emmert-Streib, and J. D. Storey. Harnessing naturally randomized transcription to infer regulatory relationships among genes. *Genome Biol* 8 (10), R219, 2007.

[Coffman, 2006] J. A. Coffman. Developmental ascendency: from bottom-up to top-down control. *Biol Theor* 1 (2), 165-178, 2006.

[Coffman and Ulanowicz, 2007] J. A. Coffman and R. E. Ulanowicz. Darwinian dynamics implies developmental ascendency: a matter of competing metaphysics. *Biol Theor* 2 (2), 179-180, 2007.

[Collier, 1999] J. Collier. Causation is the Transfer of Information. In H. Sankey, ed., *Causation and Laws of Nature*, pp. 279-331. Kluwer, Dordrecht, 1999.

[Copley *et al.*, 2007] . D. Copley, E. Smith, and H. Morowitz. J. The origin of the RNA world: Co-evolution of genes and metabolism. *Bioorg Chem* 35 (6), 430-43, 2007.

[Cumming and Collier, 2005] G. S. Cumming and J. Collier. Change and identity in complex systems. *Ecology and Society* 10 (1), 29 (online), 2005.

[Davidson, 1986] E. H. Davidson. *Gene Activity in Early Development*. Academic Press, San Diego, 1986.

[Davidson, 2006] E. H. Davidson. *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution*. Academic Press/Elsevier, San Diego, 2006.

[Davidson and Erwin, 2006] E. H. Davidson and D. H. Erwin. Gene regulatory networks and the evolution of animal body plans. *Science* 311 (5762), 796-800, 2006.

[Davidson *et al.*, 2003] . H. Davidson, D. R. McClay, and L. Hood. Regulatory gene networks and the properties of the developmental process. *Proc Natl Acad Sci U S A* 100 (4), 1475-80, 2003.

[Dunn *et al.*, 2007] . F. Dunn, V. N. Moy, L. M. Angerer, R. C. Angerer, R. L. Morris, and K. J. Peterson. Molecular paleoecology: using gene regulatory analysis to address the origins of complex life cycles in the late Precambrian. *Evol Dev* 9 (1), 10-24, 2007.

[Gould, 2002] S. J. Gould. *The Structure of Evolutionary Theory*. Belknap, Harvard University Press, Cambridge, MA, 2002.

[Gunderson and Holling, 2002] L. H. Gunderson and C. S. Holling. Panarchy: understanding transformations in human and natural systems. Island Press, Washington, D.C. USA, 2002.

[Hoelzer *et al.*, 2006] G. A. Hoelzer, E. Smith, and J. W. Pepper. On the logical relationship between natural selection and self-organization. *J Evol Biol* 19 (6), 1785-94, 2006.

[Hoffmeyer, 1997] . Hoffmeyer. Biosemiotics: towards a new synthesis in biology. *Eur J Semiotic Studies* 9 (2), 355-376, 1997.

[Huang and Ingber, 2006] S. Huang and D. E. Ingber. A non-genetic basis for cancer progression and metastasis: self-organizing attractors in cell regulatory networks. *Breast Dis* 26, 27-54, 2006.

[Hulswit, 2004a] . Hulswit. Causality and causation: the inadequacy of the received view. *SEED* 2, 3-23, 2004.

[Hulswit, 2004b] M. Hulswit. A short history of causation. *SEED* 3, 16-42, 2004.

[Kauffman *et al.*, 2007] S. Kauffman, R. K. Logan, R. Este, R. Goebel, D. Hobill, and I. Shmulevich. Propagating organization: an enquiry. *Biology and Philosophy* Online, 2007.

[Monod, 1972] J. Monod. *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology*. Vintage, 1972.

[Newman, 2005] S. A. Newman. The pre-Mendelian, pre-Darwinian world: Shifting relations between genetic and epigenetic mechanisms in early multicellular evolution. *J Biosci* 30, 75-85, 2005.

[Prigogine, 1978] I. Prigogine. Time, Structure, and Fluctuations. *Science* 201 (4358), 777-785, 1978.

[Reid, 2007] R. G. B. Reid. *Biological Emergences: Evolution by Natural Experiment*. MIT Press, Cambridge, MA, 2007.

[Salthe, 1993] S. N. Salthe. *Development and Evolution: Complexity and Change in Biology*. MIT Press, Cambridge, MA, 1993.

[Salthe, 2000] S. N. Salthe. Regaining the riches of a lost heritage: developmental systems theory and natural philosophy. `http://www.isss.org/2000meet/papers/20065.pdf`, 2000.

[Salthe, 2005a] S. N. Salthe. Meaning in Nature: placing biosemiotics in pansemiotics. *J. Biosemiotics* 1, 211-221, 2005.

[Salthe, 2005b] . N. Salthe. The natural philosophy of ecology: developmental systems ecology. *Ecological Complexity* 2, 1-19, 2005.

[Smith and Morowitz, 2004] E. Smith and H. J. Morowitz. Universality in intermediary metabolism. *Proc Natl Acad Sci U S A* 101, 13168-73, 2004.

[Sole *et al.*, 2002] R. V. Sole, J. M. Montoya, and D. H. Erwin. Recovery after mass extinction: evolutionary assembly in large-scale biosphere dynamics. *Philos Trans R Soc Lond B Biol Sci* 357, 697-707, 2002.

[Ulanowicz, 1986] R. E. Ulanowicz. *Growth and Development: Ecosystems Phenomenology.* Springer-Verlag, New York, 1986.

[Ulanowicz, 1997] R. E. Ulanowicz. *Ecology, the Ascendent Perspective.* Columbia University Press, New York, 1997.

[Ulanowicz, 2009] R. E. Ulanowicz. *A Third Window: Natural Foundations for Life.* Templeton Press, West Conshohocken, PA 2009.

[Ulanowicz *et al.*, 2009] R. E. Ulanowicz, S. J. Goerner, B. Lietaer, and R. Gomez. Quantifying sustainability: resilience, efficiency, and the return of information theory. *Ecological Complexity* 6 (1), 27-36, 2009.

THE IMPACT OF THE PARADIGM OF COMPLEXITY ON THE FOUNDATIONAL FRAMEWORKS OF BIOLOGY AND COGNITIVE SCIENCE

Alvaro Moreno, Kepa Ruiz-Mirazo and Xabier Barandiaran

1 INTRODUCTION: TWO TRADITIONAL WAYS OF DOING SCIENCE

According to the traditional nomological-deductive methodology of physics and chemistry [Hempel and Oppenheim, 1948], explaining a phenomenon means subsuming it under a law. Logic becomes then the glue of explanation and laws the primary explainers. Thus, the scientific study of a system would consist in the development of a logically sound model of it, once the relevant observables (state variables) are identified and the general laws governing their change (expressed as differential equations, state transition rules, maximization/minimization principles,...) are well determined, together with the initial or boundary conditions for each particular case. Often this also involves making a set of assumptions about the elementary components of the system (e.g., their structural and dynamic properties) and modes of local interaction. In this framework, predictability becomes the main goal and that is why research is carried out through the construction of accurate *mathematical models*. Thus, physics and chemistry have made most progress so far by focusing on systems that, either due to their intrinsic properties or to the conditions in which they are investigated, allow for very strong simplifying assumptions, under which, nevertheless, those highly idealized models of reality are deemed to be in good correspondence with reality itself.

Despite the enormous success that this methodology has had, the study of living and cognitive phenomena had to follow a very different road, because these phenomena are produced by systems whose underlying material structure and organization do not permit such crude approximations. Seen from the perspective of physics or chemistry, biological and cognitive systems are made of an enormous number of parts or elements interacting in non-linear and selective ways, which makes very difficult their tractability through mathematical models. In addition, many of those interacting elements are hierarchically organized, in a way that the “macroscopic” (observable) parts behave according to rules that cannot be, in practice, derived from simple principles at the level of their “microscopic” dynamics.

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

As a consequence, scientists have approached in a radically different way the study of these systems. Instead of looking for universally applicable laws or predictive models, biologists and cognitive scientists have searched to understand the behaviour of a living or cognitive organism by decomposing it into various parts, analysing them separately, and, then, investigating how these interrelate and affect one another within the whole system. Actually, both biologists and cognitive scientists have been able to develop rather successful research programs in their respective domains, consisting in the search for the specific *mechanisms*[1] that would explain the complex dynamic behaviour of those systems [Bechtel, 2006; 2007]. In this mechanistic context, *an explanation is viewed as the description of mappings between relevant functional operations and distinguishable structural components.* And that way of conceiving and applying the Cartesian analytic-decomposition method[2] has had, indeed, fruitful results: it has produced fundamental knowledge about the basic structure and — though in a lesser degree — also about the integration of many functional parts of living and cognitive systems (molecular and physiological in the first case, neuronal or abstract representational in the second one). However, this knowledge can only be considered as partial: it does not predict or fully explain the dynamic behaviour of such systems taken as complete entities, interacting in/with a variable and constraining environment; and neither does it allow building a general theory that could be applied to any system belonging to those domains of study.

Therefore, the development of science up to the 20^{th} century brought about two very different bodies of knowledge: in the realm where things are simple enough (though they are never as simple as they may seem), scientists were able to explain natural phenomena by means of fundamental interactions, laws of motion, basic dynamic principles and alike, generating mathematical models and theories that are predictive and can be checked through empirical techniques; whereas when the level of intricacy rises, not only in terms of the diversity of components of the system but also of their modes of organization (functional integration), the main contribution of science has been to identify and classify in a rigorous way the diverse types of material mechanisms underlying those phenomena. This, of course, included the elaboration of experimentally testable models, but with more of a justificatory or explanatory purpose than with real predictive power[3].

Both ways to do science have attained an unquestionable success. However, and for different reasons, there remained many other phenomena, not only in the field of bio-cognitive sciences but also in physics and chemistry (e.g., superconductivity, complex fluid convection patterns, reaction-diffusion propagating oscillations,

---

[1]A mechanism can be defined as a structure performing a function in virtue of its components parts, component operations, and their organization [Bechtel and Abrahamsen, 2005].

[2]Since Descartes, the workings of life and mind (except human conscious mind) have been likened to the working of machines and physiology has been seeking to interpret the organism as a complicated set of mechanisms.

[3]This lack of predictive power can also be attributed to theories that have tried to explain the appearance of those mechanisms — or their highly complex design — in historical terms, like Darwin's theory of evolution.

etc.), that were not amenable to any of these two traditional types of scientific approach. We are referring to what has been called "emergent" or "self-organizing" phenomena [Yates, 1987]. This type of non-linear, collective phenomenon, which was found to occur in a large variety of domains (physics, chemistry, biology, ecology, neuroscience, society, ...) shows the property that when a certain threshold in the number (and/or type) of interactions is reached, an emergent global pattern appears. Until recent times, these systems challenged scientific understanding because their behaviour could not be easily predicted, and the specific forms of those patterns were never fully explained through available mathematical models, based on fundamental laws of nature. Furthermore, the existence of a high number of variables interacting in non-linear ways precluded any possible reduction of that behaviour to a simple aggregation of the properties of the parts.

However new scientific approaches to this kind of phenomena have permitted a radical change. The advent of the sciences of complexity provided new tools and modelling techniques (such as cellular automata, genetic algorithms, Boolean networks, chaos and dynamical systems theory) to really tackle scientifically some of those problematic phenomena, and challenged, indeed, the simplified picture of science portrayed above. Interdisciplinary research combined with the increasing power of computer simulations made possible an unprecedented progress in our knowledge and comprehension of these systems. It is the concept of *network* and its mathematical and computational expressions that have provided the most fruitful metaphor and model for non-decomposable complex systems. We shall call 'network theory' the set of mathematical and computer simulation models and tools that have been developed to study network architectures and dynamics. Although there is no unified branch or corpus of mathematics that constitutes network theory, there exists however an increasingly indispensable 'tool-kit' of methods and disciplines that merge into what we might call network theory: this ranges from dynamical systems theory to network topology, from random boolean network models to coupled oscillators (see [Strogatz, 2001] for a survey). The study of networks with strongly and recurrently interacting components allowed scientists to deal with holistic systems, showing that, despite their variety, they share certain generic properties. For many, these advances seemed to reflect some common, universal architecture of complex systems (or, at least, the fact that they all belonged to a critical region 'at the edge of chaos' — [Langton, 1990; Lewin, 1992]).

So the blossoming of the sciences of complexity, among other important reasons, induced –and is still inducing– a profound change in biology and neuroscience toward less analytic and more synthetic-holistic views. For many, what these advances have shown is that the time is coming for converting biology and cognitive science into "true scientific disciplines", following the model of physical sciences. However, as we discuss below, it is still unclear whether the common features found through these network studies (e.g.: edge of chaos, self-organized criticality, power-law distributions, small worlds, etc...) actually respond to similar or completely different underlying, generative mechanisms. Indeed, when we examine more carefully biological and cognitive systems, a much wider variety of interacting components and non-linear interaction modes are found (as compared with any physical or chemical complex system). More precisely, biological and cognitive systems show an intricate and rich *functional* diversity within a *modular* and *hierarchical* dynamic organization. This organization is endowed with global/collective properties, like highly *robust self-maintenance,* and shows singular patterns of behaviour in their environments, like agency, multiscale adaptive flexibility, etc. Most of these features are missing from the complex dynamic models of the standard sciences of complexity, and suggests that biological and cognitive systems hide not only more complexity than physical systems, but rather different forms of it.

Acknowledging that new approaches and methodological tools are required to account for the specific types of organized complexity that generate biological or cognitive phenomena, we will discuss in what sense the 'paradigm of complexity' is deeply reframing our ways of conceiving science and philosophy of science itself. Our contribution is structured as follows: first, we explain more thoroughly how the challenge of complexity (i.e., the challenge of understanding scientifically holistic organizations) meant a very important step forward with regard to the traditional ways of doing biology or cognitive science. Then we explore how far did the 'complexity paradigm' actually go in terms of modelling biological and cognitive systems, highlighting the specificities of the latter and analysing the steps that need to be taken to capture them. And, finally, we recapitulate on the role that complexity sciences have played in the opening of what will probably constitute a new scientific era, in which up to now almost exclusively descriptive-explanatory disciplines, like biology and cognitive science, will turn into much more predictive and technologically fruitful parts of human intellectual endeavour.

2 THE CHALLENGE OF COMPLEXITY: UNDERSTANDING HOLISTIC ORGANIZATIONS

Seen from the modern perspective of complex systems, it is somewhat paradoxical that highly complex systems like biological and cognitive entities have been studied, for centuries, and rather successfully (at least in specific phenomenological domains), through atomistic and reductionistic methodologies, within the classical paradigm of mechanistic explanations. The mechanistic strategy assumes that the system under study is decomposable or 'nearly-decomposable' [Simon, 1962/1969]: i.e., that interactions between subsystem components do not change their structure and dynamics, so that the decomposition of the system in those sub-parts and their isolated study does not seriously threaten their stability and characteristic functioning when put back together. According to this working hypothesis (later we will explain why it can be adequate at certain levels of description), even when a system has a high number and a high diversity of parts and functions, the process of achieving a satisfactory mechanistic explanation of it does not require a change of methodological principles (although this might turn to be long and costly).[4]

Most researchers in both biology and cognitive science were probably aware that this methodology would not be able to provide models for predicting or reproducing the full, integrated behaviour of their objects of study, nor to provide a complete universal theory of them, beyond particular explanatory accounts. But it was assumed that this (if ever possible) had to come at a later stage. First of all one had to investigate whether a one-to-one mapping between structural and functional units in the system could be drawn and only later worry about more complicated relationships, or about how those units could be integrated. And here is where the 'organism-machine metaphor' came to play such an important conceptual role, as a second central working assumption, with great consequences. There were, to be sure, researchers ([Rachevsky, 1938; Bertalanffy, 1952; Elsasser, 1966] just to mention a few) that did not accept such naïve linear interpretations and reminded the scientific community of the necessity to adopt holistic, integral approaches in scientific strategies to deal with living or cognitive systems. But their actual impact until the end of the past century was very small, marginal at most.

Nevertheless, as knowledge on biological and cognitive systems became more detailed and fine-grained (i.e., as increasingly accurate biochemical and neurological understanding was gained at the level of basic constituents and elementary functions, thanks to the great advances in fields like molecular biology and neuroscience — see, e.g., [Morange, 2003; Bechtel, 2006; 2007]) researchers became progressively aware that components were acting in strongly holistic ways. For example, at the end of the last century the reductionist program based on molecular biology that was centred in the discovery of the structure of the genome faced a dead end, since the more details were known, the more evidence was accumulated that genetic components acted in a complex web of interactions. Thus, instead of focusing on the detection of genes, researchers started to look for the structure of regulatory networks at different levels (genomics, proteomics. . . .). In other words, as E. Fox Keller has pointed out [2005a], research in biology changed from a program inscribed in DNA analysis to a new distributed (namely, more holistic) program in which DNA, RNA and protein components operate alternatively as instructions and data.

This shift has not only been a consequence of the internal development of biol-

[4]What the strategy of decomposition and localization permits is to isolate parts and to study them separately, in terms of their functioning, so that an adequate understanding of the system can be attained. This method has led to very important progress in the understanding of the functioning of biological and cognitive systems (e.g., early understanding of metabolic pathways, reflex arcs and circulatory system). Atomistic reductionism can be understood as a class of mechanistic explanation where a one-to-one mapping between structure and function is established and where there is just a linear aggregation of localized functions, so that the properties of the whole can be reduced to or identified with the properties of the parts. What is meant by linear aggregation (or sum) is a sequence of operations (performed by well identifiable parts). Many human-made mechanisms operate in that way: a lever moves a gear that in turn makes another gear rotate which in turn displaces a latch and the door opens. But some natural mechanisms may also be described and modeled in a similar vein: a reflex arc, the snap trap of carnivore plants, etc.

ogy and cognitive science as research disciplines, but also of the already mentioned development of the sciences of complexity, which showed that the understanding of holistic systems could be closer than ever by means of strict scientific tools, especially through the application of new approaches to differential calculus (e.g., deterministic chaos, bifurcation theory, self-organized criticality) and network modelling tools that make use of computer algorithms and simulations [Waldrop, 1992; Lewin, 1992; Kauffman, 1993; Gell-Mann, 1994]. Of course, many researchers have kept working within the reductionist paradigm, which was — and will still continue to be — the central pillar of biological and cognitive sciences. But they have been, so to speak, 'filling in details'. And even if these details turned out to be very important (e.g., for medical or biotechnological applications), it does not seem that major breakthroughs in basic knowledge will come from there any more. What is still really missing, as announced above, is the combination, the *integration* of all that fine-grained accumulated knowledge, so that the *principles of organization* of those complex systems come to the surface. In other words, more global, holistic, approaches have to find their way through. Therefore, even if mechanistic explanations have not been abandoned, further progress demands that they be profoundly re-interpreted in systemic-organizational terms, as some authors have claimed in fields like cellular or neural physiology [Harold, 2001; Bechtel, 2006; 2007].

Nevertheless, the phenomenon of holistic organization is broader than biological, cognitive and social systems. Even relatively "simple" physical or chemical systems may show holistic properties. Within the classical paradigm of physical sciences the fact that there are special systems, whose properties cannot be statistically inferred, has been acknowledged since long ago [Weaver, 1948; Schrödinger, 1944]. Unlike gases and certain fluids and solids (traditional targets of physical sciences), where the global result of interacting parts can be averaged out, in these systems the non-linear and recurrent nature of the interactions between their components generates long-range correlations and global properties/patterns of dynamic behaviour, within and in relation with the environment, that hinder any straightforward connection between levels (micro-macro, individual-collective, ... — for a more detailed account see [Collier and Hooker, 1999]). These systems, which cannot be explained by means of the usual mathematical techniques and simplifying approximations in physics, are said to be *holistic* because they show global emergent properties that cannot be attributed to specific or well-distinguishable parts. And the phenomenon is *emergent* because the linear aggregation of those parts to understand the functioning of the system as a whole is not possible: 'the whole is more than the sum of the parts' as the main slogan of the sciences of complexity runs. According to Melanie Mitchell: "Roughly, the notion of emergence refers to the fact that the system's global behaviour is not only complex but arises from the collective actions of the simple components, and that the mapping from individual actions to collective behaviour is non-trivial" [Mitchell, 2006, p. 1195].

But, interestingly, the real cause of holism goes beyond the high number of components and the non-linearity of their local interactions. The system as an integrated whole determines in an important way how the parts behave. So, where does this "systemic causal power" come from? Instead of just an asymmetric, bottom-up, causal action from the constitutive components of the system to its global properties, a new type of process appears (in a reverse or opposite direction): the constraining action of supramolecular aggregates and long-range patterns on molecular dynamics, modifying or channelling low level interactions, so that a kind of feedback is created and a recurrent loop is eventually reached. That is why functional decomposition is not always possible or, more precisely, certain relevant properties of the system cannot really be reduced to an aggregation of its component parts. In those systems where we find holistic properties, the maintenance of the global stability is the result of a dynamical causal loop, such that, at least one macroscopic constraint (created by the underlying microscopic interactions) is causally necessary for the maintenance of such a loop. In other words, there is, in a minimal sense, a form of self-maintenance.[5] It is this causal circularity that makes it impossible for any specific participation in the global result to be localized in a particular component or set of components. Therefore, complex phenomena become extremely difficult — if not impossible — to trace through purely analytic mathematical models and decompositional mechanistic strategies.

So this was a challenge for traditional physical sciences, since emergent holism is not excluded from the bare physico-chemical world. The classical examples are the so-called 'dissipative structures' [Nicolis and Prigogine, 1977], like Bénard convection cells or the B-Z reactions. But there are other similar examples in nature, like tornados, or fire flames, in which the maintenance/stability of the global pattern of order requires a dynamical 'micro-macro' causal loop. In these cases, even if some structural decomposition of the system can be achieved, the mapping between structure and function is highly problematic because there is no one-to-one correlation between structure and function. In terms of the relationship between structure and function, there is a many-to-one mapping (many components perform a single collective function, or contribute to a single pattern) that cannot be accounted for by standard statistical approaches. From this situation, things can get even more intricate if: a) the emergent pattern may change (between different macroscopic configurations) depending on internal or environmental conditions[6]; b) intermediate-level patterns start to form; or, c) global self-organization

[5]As van Gulick [1993] has pointed out, most patterns in the world exist because they are self-maintaining; namely, these patterns are the result of circular processes: emerging organized structures exerting *selective action* on the lower level dynamics of their constitutive parts, which recursively contribute to the maintenance and preservation of those very organized systems.

[6]Feed-back loops occur not only among components of the system but also between the system and its environment. So much so that some characteristic behaviour of the system could be impossible to study if we consider it isolated from its environment. In those cases, causal attribution of observed behaviour to internal components is, obviously, not possible. It is always important to remember this other, not less relevant aspect of complexity, coming from the *interactive emergence* of some functional properties in a system. In fact, this is one of the major theoretical achievements that situated robotics has contributed to in the foundations of cognitive science, showing how complex adaptive behaviour could be the result of relatively simple –yet strongly interactive– control systems in robots [Brooks, 1991; Steels, 1991; Clark, 1997], whereas

is combined with local self-assembly or self-re-structuring processes, leading to new identifiable structural/functional components in the system. In these cases, the many-to-one mapping would become many-to-many, but we will address this a bit further on, when we deal with the specific complexity found in living and cognitive systems.

As already mentioned, the study of the emergent behaviour of holistic systems has been made possible thanks to the development of powerful computing models and machines. Although the dynamic processes leading to emergent behaviours are not analytically tractable, numerical methods, consisting in a fine-grained step-by-step update and recording of the state of all the interrelated variables of the system, allows the drawing of the state space of these systems. In this way, the evolution of the system is "synthetically" reproduced in the course of the simulation, rather than deduced for a given time value. Under these conditions, the prediction of the global property requires an enormous amount of parallel computation, where the complexity of the computational simulation is almost equivalent to that of the simulated system.[7] Since sensitivity to initial conditions might be critical, the process is repeated for a wide range of initial conditions and the behaviour of the system recorded. A full record of different simulations allows drawing the state space of the system (given certain boundary conditions and specific parameters) so that regular and reproducible patterns of behaviour can be observed. Under different network configuration parameters, the rules governing components or boundary conditions can be systematically studied through simulation, and a deeper understanding of the structure-function mapping is gained without a strict decomposition or a one-to-one localization.

Following this methodology, different tools and models have been developed to discover generic properties of complex systems, often without a specific target system to fit and correspond with [Barandiaran and Moreno, 2006a]. These models and tools are typically used to illustrate emergent properties, to discover universal patterns of connectivity in networks and their properties. Among the early models we should mention those that focused on the special dynamic behaviour of systems in a region in between 'order and chaos', for which, indeed, the standard statistical analysis did not work. The paradigmatic example is 'spin-glasses', which led to the development of 'Ising-type' of models (originally applied to ferromagnetic systems). And, essentially, the new approach was to represent the system as a discrete set of nodes, with very limited dynamic states (typically just two: 'on/off' or 'up/down') and variable connections among them, which affected in different ways their respective states. Through this basic strategy, one could deal with systems of

previous functionalist models required a high computational load to achieve the same behavioural performance.

[7]According to Kolmogorov and Chaitin, if the description of the structure cannot be shortened, it is said to be (maximally) complex (see, e.g.: [Li and Vitanyi, 1997]). However, there are also other criteria to measure complexity, like computational time or the difficulty of finding that compressed description. For instance, a fractal structure, according to these other criteria, would still be regarded as complex, whereas from the point of view of the mathematical algorithm that generates it, it would be very simple.

many components in a radically different way (as compared to previous statistical methods): namely, putting the emphasis on the different types/degrees of connectivity or interaction that may be established among components of a system and the effect that this has on its global properties [Weisbuch, 1986].

Actually, the classical works by Kauffman on complex gene regulation patterns (using Boolean networks) and autocatalytic sets (for a review, see [Kauffman, 1993]), or Cellular Automata [Wolfram, 1986] were based on this key modelling idea. The assumption that simple local rules among discrete, randomly connected elements could give rise to unexpected global behaviours was very fruitful, leading to many other simulation algorithms and models, which were applied to study phenomena at very different scales. Quite interestingly, all those phenomena were characterized by a critical transition point (most often related to a threshold of connectivity in the network — e.g., Langton's [1990] Lamda parameter) above which global properties emerged by some sort of 'percolation' effect. The success of these pioneer models triggered increasing interest in the properties of complex networks [Strogatz, 2001] and how these could capture essential features of systems as diverse as metabolic regulation pathways [Wagner and Fell, 2001; Ravasz, *et al.*, 2002], ecosystems [Montoya and Solé, 2002], brain networks [Sporns, *et al.*, 2000; Sporns, *et al.* 2005] or the world wide web [Barabasi and Albert, 1999], which were soon discovered to share some general topological properties (e.g., power-law distributions, scale-freeness, 'small worlds', etc.).

Thus, by the end of the century a whole new scientific research program was initiated and seemed to be capable of grasping, with tools based on the increasing power of computers, (at least certain types of) prototypic holistic systems. And not only grasping but also predicting/reproducing their dynamic behaviour, as well as quantifying to a good extent their degree of complexity. A possible interpretation of all this success is that, behind the apparent complexity and variety of those systems, there are some general principles (applicable to systems with a great number of interactions, or with particular patterns of connectivity) that reflect a common architecture emerging at various phenomenological levels.[8]

3 THE IMPACT OF NETWORK THINKING ON THE CONCEPTUAL CORE OF BIOLOGICAL AND COGNITIVE SCIENCES

Living — and, thereby, cognitive — systems necessarily originated and persisted on the basis of previous physico-chemical modes of complexity. In the same way as self-organizing reaction-diffusion patterns [Turing, 1952] appear in many biochemical processes of uni- and multi-cellular organisms, the type of holistic phe-

[8]A more critical view on these network modelling approaches (see, e.g.: [Fox Keller, 2005b]) tends not to consider their results so surprising or significant, after realizing that there are many ways, many mechanisms through which those general features can be generated in nature (and in the computer), and that the relevant information to understand complexity is contained precisely in those underlying mechanisms. We will see that this is quite true, when we examine, in the next section, the specific nature of biological and cognitive complex systems.

nomena we just described (let us call it 'primary holism') is ubiquitous in the biological and cognitive world: e.g., in the cell cycle, in the symmetry breaking phenomenon related to the appearance of morphogenetic fields in development, in the 'collective intelligence' of ant or bee colonies, in neural Central Pattern Generators, in the chaotic dynamics of the olfactory bulb neurons in rabbits... It basically reflects how an ensemble of interacting units produces a global property or pattern of behaviour that cannot be ascribed to any single of them, but to the whole ensemble. And this is both relevant for the biological and the cognitive spheres, revealing that one-to-one mappings between structural and functional parts are typically not possible. This new vision led to a number of revisions, acknowledging the power of collective, distributed organizations: the queen did not have to instruct the ants for them to operate in a coordinated way; cells did not have to contain highly explicit genetic instructions since they could form complex dynamic patterns and tissues just following local context-sensitive rules; neurons did not have to compute symbolic propositional states but could undergo distributed sub-symbolic processing; etc.

Before the advent of network theory and the new insights it brought about, the foundations of biological and cognitive sciences had to be based on those assumptions that could meet the demands of the mechanistic-decompositional methodology. These assumptions had enormous consequences on the way cognitive and biological systems were conceptualized, particularly at higher level theories (such as principles of biological evolution or philosophy of mind). The one-to-one mapping assumption, for instance, somehow forces a jump over many levels of complexity to establish a workable one-to-one representational relationships at different scales.

In biological systems this atomistic representationalism took the form of a one-to-one genotype-phenotype mapping whereby organisms were conceptualized as phenotypes encoded on genetic strings and evolution as a process of genetic variation and retention due to a selective force that propagated linearly from phenotypes to selfish genes [Dawkins, 1974]. Research on genetic regulatory networks [Lewis, 1992] has completely transformed the foundational framework of evolutionary biology together with the study of self-organization in developmental processes [Goodwin, *et al.*, 1983; Kauffman, 1993]: genes are now one among all the components and processes that produce organisms, within a network of internal and interactive dynamics that constitutes their life-cycle, and not a freely re-combinable encoding of biological forms. As a result, evolution does not operate upon an abstract functional phenotypic space but through a highly constrained (occasionally discontinuous) space of possible morphologies [Alberch, 1982], whose formation requires acknowledging the environmental, material, self-organized and often random processes that appear networked at different scales. Another assumption that was strongly revised was the sharp system-environment distinction, which was made necessary in order to map internal structures to environmental conditions. Thus, now it is not the environment that poses problems for genetically instructed phenotypes to solve, but a complex organism-environment interaction process that defines the adaptation and survival of an organism [Lewontin, 2000], including ecological and niche-construction processes [Laland, *et al.*, 1999].

The number of transformations that the sciences of complexity made in the foundations of biology was comparable to those made in cognitive science. Symbolic computational units and their rule based manipulation were postulated as theoretical primitives by functionalist approaches to the mind [Fodor, 1975; Newell, 1980; Block, 1980]. Cognition was primarily conceptualized as a set of inner symbol manipulation procedures on the side of the subject, mirroring external states of affairs (and a one-to-one mapping between internal abstract functions and environmental structures) in order to deliver planned actions to be executed in the world. The rise of artificial neural networks and parallel distributed processing in the 80's started to undermine the atomistic symbolic assumption leading to a subsymbolic processing paradigm where cognitive architectures became biologically grounded on the distributed nature of brain architectures [Rumelhart, *et al.*, 1986; Hopfield, 1982]. During the 90's the subject-object dichotomy came to be challenged by situated robotics, whose engineering principles exploited recurrent system-environment interactions to generate emergent functional behaviour [Brooks, 1991; Steels, 1991]. Around the same time, dynamical system approaches to cognition [Kelso, 1995; Port and van Gelder, 1995] started to conceptualize cognition as a dynamical process where cognitive behaviour and development emerged from dynamically coupled brain-body-environment systems [Beer, 1995; Thelen and Smith, 1994] challenging previous assumptions about the innateness of cognitive capacities [Elman, *et al.*, 1996] and their purely internalist and disembodied computational bases. Where previous mainstream neuroscience was mostly focused on finding neural correlates of cognitive representations and localizing functions, brain activity started to be understood as emerging from the collective dynamics of distributed neural ensembles at the edge of chaos [Skarda and Freeman, 1987; Tsuda, 2001] integrating its biological embodiment as a complex interplay between internal bioregulatory functions and sensorimotor activity [Damasio, 1994; Lewis, 2005]. Thus the clear cut subject-object dichotomy and the abstract atomistic computational framework came to be progressively substituted by an internally and interactively distributed network of dynamical processes capable of giving rise to context-sensitive, flexible and multiscale adaptive capacities, equally constrained and enabled by its biological embodiment.

Despite the early warnings about the consequences that reductionist approaches could have for conceptualizing biological and cognitive systems [Waddington, 1957; Bertalanffy, 1952; Merleau-Ponty, 1963], only when the network theory toolkit became available to explore the intricate and emergent structure-function mappings in metabolic, neural, ecological, developmental, behavioural and other types of phenomena, did the true complexity of life and cognition begin to be fully acknowledged and the old foundational framework progressively transformed, substituting its traditional simplifying assumptions by more encompassing and realistic ones.

4 SPECIFICITIES OF BIOLOGICAL AND COGNITIVE COMPLEXITY

Therefore, we can say that complexity sciences have already contributed in a significant way to a better understanding of living and cognitive systems, through the new insights and tools provided to deal with that primary or most elementary form of holism, as described in the previous section. However, the complexity of these systems goes far beyond there. It involves a much richer internal structure, with functional diversity integrated in second-order forms of holism. Biological and cognitive systems are certainly made of a large number of parts or elements acting in non-linear ways, but they also show other features that are absent in non-living complex systems: hierarchical organization, long-term sustainability, historicity, functional diversity, adaptivity and agency. Everywhere in biology and in cognitive science we deal with systems made of parts or elements with different functionalities acting in a selective and harmonized way, coordinating themselves at different time scales, interacting hierarchically in local networks, which form, in turn, global networks and, then, meta-networks... The organization of living systems consists in different nested and interconnected levels which, being somewhat self-organized in their local dynamics, depend globally one upon the others. This means that both the components and the sub-networks contribute to the existence, maintenance and propagation of the global organizations to which they belong. And, in turn, those global organizations contribute to the production, maintenance and propagation of (at least some of) their constitutive components.[9]

Thus, this complexity is quite different from what we can find in non-living systems, like the dynamical complexity of holistic non-linear networks in both physical and chemical domains. It is a type of complexity that goes beyond a mere increase in the "complicatedness" (i.e., an increase in the number and variety of components) of self-organizing systems: it involves qualitative changes in the form of organization of the system, by means of creating functionally differentiated structures and levels within it. So here we are dealing not with mere complexity but with *organized complexity* as Weaver [1948] called it. Nevertheless, by using this term we mean something more than Weaver's idea of systems consisting of many different elements interrelated into an organic whole; rather, we want to refer to the fact that these systems regulate and functionally manage their own complexity. As J. Mattick [2004] has recently pointed out, what really matters in biological evolution is not so much the generation of complexity, but its functional and selective control. Evolution shows that the appearance of increasingly complex systems goes together with the invention of different ways of utilising hierarchical regulation to manage internal functional variety.

Actually, what distinguishes biological and cognitive complexity from other forms of complexity or complicatedness lies in the role played by *mechanisms*

[9]Recurrent-functional relationships established between the parts and the whole of a system expand to interactions between the system and its environment. The latter are controlled by the system, so that they contribute to the system maintenance. Such a control is modulated by the internal and external (adaptability) conditions and constitutes the so-called *agent's behaviour*.

of regulatory control in the functioning of these systems [Christensen, 2007]. By the term "regulatory control" we mean an action performed by a certain part (component or substructure) of an organized system in a way that some processes occurring in it are functionally constrained: in other words, their dynamics is partially harnessed, or channelled in order to ensure or improve the maintenance of the system as a whole. For example, a metabolism is a self-maintaining organization that functionally modulates its ways of operating when perturbations occur (e.g., concentrations of internal metabolites are maintained despite changes in some of the inputs/outputs), whereas a whirl is a non-regulated self-maintaining organization (since a totally distributed holistic network, although it may be more or less robust, is not capable of self-regulation). Equally, neuronal and interactive self-organization may generate some interesting behavioural patterns; however, some means of action selection would also be required if the agent is to organize a complex behavioural repertoire according to a given set of priorities [Prescott, 2007]. Some early pioneers of the paradigm of situated and autonomous robotics (which was inspired, to a certain degree, in the sciences of complexity — particularly in the ideas of self-organization), after an initially celebrated success, soon became aware that self-organization alone faced a bottleneck of complexity growth and that modular and hierarchical control principles had to be reintroduced in order to scale up [Nolfi, 1997; Brooks, 1997].

This highly organized complexity cannot be the result of the "one-shot, order-for-free, kind of self-organization associated with the kinds of uniform non-linear dynamical systems that mathematicians usually study", as E. Fox Keller [2007, p. 316] has pointed out. Rather, it requires an organization structured in different (sub)systems interconnected at different levels and time scales.[10] In other words, it requires the creation, within the system, of *dynamically decoupled*[11] subsystems that can modify the parameters of other subsystem parts. Since these subsystems

[10]Organisms are internally organized in different networks and sub-networks of dynamical-metabolic processes. But, in addition, they are embedded in a more global evolutionary meta-network, articulated also on different levels and functioning at much wider spatial and temporal scales. Slow, phylogenetic processes are causally connected with faster, ontogenetic ones, generating a global long-term sequence of spiral-like cycles, that we call "evolution", which allows cumulative increases in complexity. The long-term sustainability of this complexity is based on the spatial and temporal interconnection and interdependence of the individual organisms, generating a wide meta-system (an eco-system or, even, a whole biosphere). An essential characteristic of these higher forms of complexity is, thus, their inclusion in a long-term evolutionary phenomenon. Life started as some minimal form of organization (population of prokaryotes) and developed in a continuous and causally interconnected way. As a consequence, *historicity* is also a fundamental component of biological and cognitive complexity.

[11]By the term "dynamical decoupling" [Moreno and Lasa, 2003; Ruiz-Mirazo *et al.*, 2007] we are referring to the organization of a system such that the basic, constitutive processes generate and sustain a relatively independent dynamic subsystem, which in turn acts selectively regulating those constitutive processes. This means that there are, at least, two organizational levels (although in highly complex biological systems there are many more): the constitutive system is the basic or "lower" level and the regulatory subsystem is the "higher" one. Significantly, the higher-level regulatory subsystem cannot exist without the regulated constitutive system: both levels causally depend upon each other, and therefore the whole system appears as an integrated unity.

work at different rates and with different operational rules, the system has an increased potential to explore new or alternative forms of global self-maintenance (that are not accessible to 'flat' systems without any hierarchy or modularity in their organization). In this way, the higher level subsystem creates a set of functional constraints on the lower-level dynamics. At the same time, the controlled level plays a fundamental role in the constitution and maintenance of the controller level (and therefore, of the whole system). For example, the genetic (sub)system in the cell acts — through the synthesis of specific enzymes — as a mechanism of control, harnessing the dynamics of the metabolic reactions, but in turn, metabolic processes contribute to the maintenance, reparation, replication and translation of genetic components. Another example is the nervous system, which controls metabolic processes (circulation, digestion, breathing, etc.) on the one hand but, on the other hand, is fabricated and maintained by the latter.

Therefore, biological and cognitive systems convey specific forms of complexity that, through holistic-emergent processes (which are continuously taking place), produce both dissipative patterns and new, more complex structures which, in turn, are bound to become selective functional constraints acting on the dynamic processes that underlie those holistic processes. The reason why those functional constraints can be described as mechanisms is that they act as distinguishable parts (or collections of parts) related to particular tasks (e.g., catalytic regulation) performed in the system. So both aspects are, thus, complementary: the holism of the global network of processes and the local control devices/actions that are required for the system to increase in complexity. Moreover, the newly created and functionally diverse constraints may give rise (once a certain degree of variety is reached) to new self-organizing holistic processes, which, in turn, may be functionally re-organized. In this way, an increase in organizational complexity can take the paradoxical form of an apparent "simplification" of the underlying complicatedness, giving rise to levels of organization in which a mechanistic decompositional strategy might be locally applicable. The idea, taken originally from Pattee [1973], would be that new hierarchical levels are created through a functional loss of details of the previous ones.

However, the key point here is that this complementarity between functional mechanisms and holism, is due to their causal circularity. Since a mechanism is an explanation of the functioning of a system in terms of a specific arrangement of parts, it always sends one back to another mechanism to explain that arrangement of parts, and so on indefinitely. Thus, causal circularity is the only possible solution to the infinite regress posed by mechanistic explanations. And here is where the main difference between what we mean by a mechanistic explanation in a man-made system and in a natural one lies. Actually, the organism-machine analogy brings biological and cognitive sciences closer to engineering than to physics or chemistry, as Polanyi [1968] sharply highlighted, arguing that both (organisms and machines) operate according to local rules (or 'boundary conditions') irreducible to general physico-chemical laws/principles.[12] But, whereas in man-made

[12]In fact, engineers care about the design of a device that operates as efficiently as possible in

organizations structure and function are causally asymmetric (a given structure generates a given function, though not conversely) in biological and cognitive systems *both structure and functions are engaged in a circular causal relation* (as other authors have already claimed — [Rosen, 1991]). Since what artificial machines do (i.e., their function) does not feed back to their own structure, machines must be externally designed and repaired.

In contrast to artificial machines, in biological (and cognitive) systems the organization is internally generated, so the structure is itself the cause and the result of their functions. In other words, biological and cognitive systems are organized so that for each required operation, there is a part (or set of interconnected parts) that is apt to perform it. For instance, the main emergent property of a living cell is its maintenance through the continuous renewal of its constitutive parts and through the management of the flows of matter and energy with its environment. Highly complicated components, like enzymes, accomplish specific functions, which, in turn, ultimately regenerate these enzymes from more simple building blocks. If we compare a living cell with a computer, we can see that, although in both cases there is a highly complex internal structure, the computer functioning does not contribute to the maintenance of its own structure, whereas the cellular functioning is the *cause* of its structure.

Quite interestingly, this property of causal closure in 'soft material automata' (as opposed to the rigid or fixed structure of relationships in traditional man-made machines) involves high rates of energy dissipation, so it requires the continuous production of work by the system [Ganti, 1987]. Thus, living systems, which are continuously and literally fabricating themselves, can only maintain their organization in far from equilibrium conditions by being material-thermodynamically open. Nevertheless, biological and cognitive systems are something more than self-maintaining organizations operating under specific matter-energy flow conditions. Rather, they recruit their own internal organization to actively create and maintain the internal and boundary conditions necessary for their own constitution: in other words, they are *autonomous* systems [Etxeberria, *et al.,* 2000; Ruiz-Mirazo and Moreno, 2004; Barandiaran and Moreno, 2006b; Barandiaran and Ruiz-Mirazo, 2008]. As a consequence, biological and cognitive systems are not only intrinsically functional, but *normatively* functional systems: they not only function in a certain complex way, but they *must* function in that way in order to ensure their continuing existence. The key notion here is that the structure and stability of living systems are not independent of their functions but, on the contrary, that the functional level feeds-back to the structural one. There is, thus, a circular co-dependence between the stability or self-maintenance of structures and their functions.

Now, what all this makes clear is that the development of models trying to capture (and potentially reproduce) the huge complexity of biological and cognitive systems cannot be achieved without a deeper understanding of the constraints or

---

order to achieve an external, human-defined, purpose; but they ignore the question of how such design could have appeared without another design (or without an ultimate human designer).

mechanisms that allow the functional reorganization of emergent collective patterns: in other words, how to convert these patterns into new functional substructures in the system; and how, within a global causal circularity, these new functional substructures exert top-down regulatory actions on other components of the system. Among other things, a deeper understanding of how a self-organizing system can re-use some sets of its initial components to achieve new functions, through a 'tinkering-like' process. But we will discuss this in the next section.

5 THE PRESENT OUTLOOK: FURTHER STEPS IN THE MODELLING OF BIOLOGICAL AND COGNITIVE SYSTEMS

So how could the specific features of biological and cognitive systems, in particular the emergence of new functional variables and their integration in complete operational units (i.e., the actual living/cognitive organisms), be modelled? The sciences of complexity should continue giving us new insights and tools to tackle this problem, through further advances in network theory and, particularly, the development of bottom-up simulation approaches. Up to now they have already provided, apart from a solid theoretical framework to explain first-order holistic systems (as remarked above), a new way to look into the organizational sub-structure of real living and cognitive systems. This mainly derives from the extensive use that is being made of network theory to try to extract relevant information from the huge data banks provided by research in cell and developmental biology, ecology and neuroscience.

Two main messages can be drawn from this survey, so far. The first is that modularity and cross-level processes play a very important role in the organization of these highly complex systems (see, e.g., [Hartwell, *et al.*, 1999; Solé, *et al.*, 2002a; 2002b]). In other words, that functional decomposition is justified but has to be carried out carefully, taking into account different levels of description and their possible interconnections. The second is that redundancy and degeneracy [Tononi, *et al.,* 1999; Edelman and Gally, 2001] are surely present at these different levels, so accurate decomposition will actually be very hard (if ever possible) to achieve. That is to say: in reality things are not as neat as they could be, since these systems appear to be the result of a long process of tinkering (not of a beautiful design made by an external, analytically inspired engineer). Indeed, nature builds up from what is already available: the metaphor of the tinker is quite appropriate [Jacob, 1977], and particularly well suited to describe how things have come together and evolved in the realm of complexity [Solé, *et al.*, 2002b].

This general picture explains why the analytic-mechanistic approach to living and cognitive systems has been rather successful so far: despite the complexity involved, the high degree of modularity allowed for decomposition and the search for one-to-one mappings between structural and functional components. But it also sets its limits: this type of Cartesian approach can only be, in the best case (i.e., if close enough to the 'near-decomposability' condition), a first approximation to the problem. Degeneracy and redundancy — without going into the technicalities of network theory — reflect the fact that many structural components in these systems may actually have the same function, or that a single component may have more than one function, so the one-to-one assumption, in general, does not hold. Real complexity out there, in the biological and cognitive worlds, is not only many-to-one but certainly 'many-to-many'. Or, rather, 'many-to-many-to-all', because the presence of hierarchical, inter-level relationships indicates that the problem of integrating all these modules in a second-order holistic organization cannot be circumvented. However, this is far from being a trivial problem.

Recent advances in network approaches are pointing at this, although they are not quite there yet. Researchers in the field are moving from the current stage, in which the focus has been on the topological/architectural features of networks — like the 'small worlds' found in metabolic networks [Wagner and Fell, 2001] or in food-webs [Montoya and Solé, 2002], as mentioned before — towards re-introducing true dynamics in their graph analysis and enhancing, in this way, the bottom-up explanatory potential of the resulting models. So it is clear that the key question is not just specifying the general rules for network growth or development (e.g., preferential attachment, duplication/rewiring, optimization, ...), but including in their models real dynamics (with fluctuations, noise, possible chaotic behaviour,...), as in the original spirit of complex systems research. For instance, the analysis of fluctuations in network dynamics can give relevant information to distinguish how much in it is due to self-organizing, internal processes, and how much is driven by external perturbations [Argollo de Menezes and Barabasi, 2004a; 2004b]. What is important to realise it that these new approaches should not only be applied to construct descriptive models of complex dynamics (like the epidemic spreading effects in scale-free networks — [Pastor-Satorras and Vespignani, 2001]) but also to derive new general principles that throw some light on the problem of the emergence of functionality (i.e., re-organization mechanisms, like regulatory controls), hierarchies and second-order types of holism.

We consider that in order to tackle this fundamental problem further importance has to be given to bottom-up simulations of systems in between self-organization and full-fledged living or cognitive systems. Besides, such theoretical simulation models should be complemented with real experiments in the lab: real experimental data must be incorporated into the simulations, and conversely. It is crucial that *in silico* and *in vitro* approaches inform and push each other in a much more fluent, synergetic way than up to date. This is the focus, for instance, of a recent special issue on the prospects for the synthesis of 'artificial cells' [Solé, *et al.*, 2007]. Another example of a hybrid modelling approach, coming from cognitive and neuro sciences, is given by the DARWIN robotic platform, developed by Edelman and collaborators [Krichmar and Edelman, 2005]. The control architecture of these robots, inspired by human neuroanatomical data, combines self-organized networks with modularized and hierarchically controlled processes, which all together — in continuous embodied interaction with the environment — are capable of reproducing cognitive phenomena such as categorization, associative learning, etc. Another example is Bray et al.'s model of *E. coli* chemotaxis integrating molecular Brownian motion and DNA molecular details to render a predictive model of bacterial behavior [Bray, *et al.*, 2007].

But the point is not only methodological. The more knowledge we gain on complex systems, the stronger need we have to resort to manipulation techniques; i.e., the more dependent theoretical investigations are on the operations upon the objects of study. Traditional mechanistic methodologies in biology and cognitive science have, of course, made an intensive use of manipulation to study natural systems. However, what is new is the increasing necessity of developing experimental programs aiming at synthetic reconstructions of living or cognitive mechanisms.[13] Thus, researchers are illuminated by an engineering ideal: search for the design of new genetic circuits and metabolic and neural networks. At the same time more and more scientists find that analytic de-composition needs to be complemented by synthetic re-composition (even up to the point of recomposing a different organism/neural system from deliberately modified versions of the parts). In this way, the synthetic approach is going to be increasingly used to deepen our understanding of life and mind, and to confront new problems that do not manifest through analysis.[14] Therefore, the question is whether understanding life and mind, building up models of biological and cognitive organization, and the actual fabrication of living/cognitive beings are not really convergent processes.

6 CONCLUSION: TOWARDS A NEW VIEW OF SCIENCE?

For centuries biological and cognitive systems have been studied by analytic decomposition, trying to determine how their parts are arranged so as to generate an observed behaviour, in a similar way as the parts in human-made machines are suitably arranged to perform an externally fixed goal. 40 years ago Polanyi [1968] pointed out that the local rules harnessing physical (or chemical) laws, which would explain both the organization of living beings and machines, were complex sets of boundary conditions irreducible to general physico-chemical laws/principles. But the fundamental question remains, in so far as natural systems are concerned: where do these boundary conditions — and therefore complex organization — come from? What are the main transitions in complexity that allow access to the minimal thresholds (for life and cognition)? How do these minimal forms develop and evolve without the presence and contribution of external intelligent designers?

[13]Thereby the growing interest in their origins, as well: by trying to understand the natural origins of the material mechanisms underlying biological and cognitive phenomena it will be possible to disentangle their fundamental nature, beyond descriptive accounts. This bottom-up, naturalization program is crucial not only for the future of complexity sciences, but also for the future of biology and cognitive sciences, if they are to remain proper scientific disciplines (not just technological enterprises).

[14]For example, synthetic biology is becoming the natural evolution of a set of procedures, like genetic engineering, required to advance in biological knowledge itself. The development of biology as a quantitative research field has simply made genetic engineering and synthetic biology unavoidable. As Rosen [1991] pointed out, the goal of understanding life is becoming more and more equivalent to that of fabricating it.

Darwin provided a general theoretical framework that could be used to solve this problem (or part of it). But evolutionary theory (the mechanism of natural selection) requires an initial form of organization endowed with a considerable degree of complexity [Ruiz-Mirazo, *et al.,* 2007]. Thus, although the problem of growth and diversification of complexity could require concepts and tools coming from evolutionary theories, the problem of the *origin* of complex organizations cannot really be solved in that way. Modern science, in turn, has provided us with sophisticated computer techniques to perform quantitative studies of a wide variety of networks, showing emergent properties out of densely interconnected elements. Nevertheless, the sciences of complexity have not yet developed a full-fledged research program to deal with hierarchical organizations, articulated on multiple functional parts making up robust self-maintaining systems. In other words, neither evolutionary theory nor self-organization theories, as they stand today, can account for the natural origin of mechanisms.

Finding a theory that may eventually bridge this gap will require much more interdisciplinary collaboration in research. What we are learning at present is that the study of highly complex systems pushes us to overcome a strictly specialized view, constrained within a single referential framework [Pask, 1960], and to adopt a true interdisciplinary approach. The elaboration of theories and models that lead to a deeper and more global understanding of biological and cognitive systems involves the integration of very different methods and experimental data, all of them required to study even "isolated" aspects of their functioning. Experience shows that when an explanation of a complex system is achieved in terms of their more simple, lower level components, what is gained is not really more knowledge on the low-level theories but rather a new interdisciplinary theory. As it is most obviously illustrated by the research carried out on prototypical case studies or model systems (*Mycoplasma, E. Coli, slime moulds, drosophila, C. elegans,...*), successful explanations will be achieved through the merging of models, techniques and data coming from different studies such as genomics, development, cell physiology, psychology, neurobiology, neural networks, etc.

Actually, this is the only way to combine the network-focused, holistic perspectives with the mechanistic ones. That is why progress in the study of highly complex systems will probably imply the convergence between the holistic and the mechanistic methodologies. The understanding of complex forms of holism will progressively allow (and, at the same time, will be progressively allowed by) the merging between mechanistic explanatory methodologies, based on reductionist decomposition, and the construction of simulation models. Through these simulation models it will be possible to make explicit (and interpret under new light) processes that not only give rise to those emergent mechanisms but also assemble them into coherent and adaptive wholes.

In sum, the success of the new sciences of complexity will reframe our ways of conceiving and tackling old, open problems of science, philosophy of science and philosophy in general. As Benner and Sismour express it [2005], referring to these recent changes in the field "synthesis drives discovery and understanding in ways that analysis cannot". Biology and cognitive science seem to be at a historical crossroads in which the fabrication and simulation of life-like and cognitive-like systems (or subsystems) is beginning to be feasible. The consequences of this are beyond our imagination, but we are probably witnessing the first steps of a new scientific era.

ACKNOWLEDGEMENTS

Funding for this work was provided by grants 9/UPV00003.230-15840/2004 from the University of the Basque Country, IT-250-07 from the Basque Government and FFI2008-06348-C02-01 and BFU2006-01951/BMC from the Ministry of Science and Education (MICINN). K Ruiz-Mirazo holds a *Ramon y Cajal* research fellowship and X. Barandiaran had the support of the doctoral fellowship BFI03371AE and currently holds a post-doctoral fellowship from *Ministerio de Educación y Ciencia, Programa Nacional de Movilidad de Recursos Humanos del Plan nacional de I-D+I 2008-2011.* We also acknowledge Cliff Hooker's editorial suggestions for improvement and William Bechtel's helpful comments.

BIBLIOGRAPHY

[Alberch, 1982] P. Alberch. Developmental constraints in evolutionary processes. In J. T. Bonner (ed.), *Evolution and Development*, Dahlen Conference, Springer Verlag, New York, 1982.

[Argollo de Menezes and Barabási, 2004a] M. Argollo de Menezes and A. L. Barabási. Fluctuations in Network Dynamics. *Physical Review Letters* 92 (2), 028701, 2004.

[Argollo de Menezes and Barabási, 2004b] M. Argollo de Menezes and A. L. Barabási. Separating Internal and External Dynamics of Complex Systems. *Physical Review Letters*, 93(6), 068701, 2004.

[Barabási and Albert, 1999] A. L. Barabási, R. Albert. Emergence of scaling in random networks. *Science*, 286, 509-512, 1999.

[Barandiaran and Moreno, 2006a] X. Barandiaran and A. Moreno. ALife models as epistemic artefacts. In Rocha, L.M., Yaeger, L.S., Bedau, M.A., Floreano, D., Goldstone, R.L. & Vespignani, A. (Eds.) *Proceedings of the 10^{th} International Conference on the Simulation and Synthesis of Living Systems (ALifeX).* MIT Press, Cambridge, MA,513 — 519, 2006.

[Barandiaran and Moreno, 2006b] X. Barandiaran and A. Moreno. On What Makes Certain Dynamical Systems Cognitive: A Minimally Cognitive Organization Program. *Adaptive Behavior*, **14**(2), 171-185, 2006.

[Barandiaran and Ruiz-Mirazo, 2008] X. Barandiaran and P. Ruiz-Mirazo, eds. Modelling autonomy. *BioSystems*, **91**(1), 2008.

[Bechtel, 2006] W. Bechtel. *Discovering Cell Mechanisms. The Creation of Modern Cell Biology.* Cambridge University Press, Cambridge, 2006.

[Bechtel, 2007] W. Bechtel. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience.* Erlbaum, NJ, 2007.

[Bechtel and Abrahamsen, 2005] W. Bechtel, A. Abrahamsen. Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 421–441, 2005.

[Beer, 1995] R. D. Beer. A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, 72, 73–215, 1995.

[Benner and Sismour, 2005] S. A. Benner and A. M. Sismour. Synthetic Biology. *Nat. Rev. Genet.* 6, 533–542, 2005.

[Bertalanffy, 1952] L. von Bertalanffy. *Problems of life: An evaluation of modern biological thought.* Watts & Co., London, 1952.

[Block, 1980a] N. Block. What is Functionalism? In N. Block (ed.), *Readings in the Philosophy of Psychology*, Cambridge MA, Harvard University Press, 1980.

[Bray *et al.*, 2007] D. Bray, L. Matthew, and K. Lipkow. The Chemotactic Behavior of Computer-Based Surrogate Bacteria. *Current Biology*, 17, 12–19, 2007.

[Brooks, 1991] R. A. Brooks. Intelligence without representation. *Artificial Intelligence Journal*, 47, 139–160, 1991.

[Brooks, 1997] R. A. Brooks. From Earwigs to Humans. *Robotics and Autonomous Systems*, 20(2–4), 291–304, 1997.

[Christensen, 2007] W. Christensen. The evolutionary origins of volition. In D. Spurrett, H. Kincaid, D. Ross, and L. Stephens (eds.), *Distributed Cognition and the Will: Individual Volition and Social Context.* MIT Press, Cambridge MA, 2007.

[Clark, 1997] A. Clark. *Being there. Putting Brain, Body, and World Together Again.* MIT Press, Cambridge, MA, 1997.

[Collier and Hooker, 1999] J. D. Collier and C. A. Hooker. Complexly organised dynamical systems. *Open Systems and Information Dynamics*, 6(3), 241-302, 1999.

[Damasio, 1994] A. R. Damasio. *Descartes' Error. Emotion, Reason and the Human Brain.* G. P. Putnam, New York, 1994.

[Dawkins, 1974] R. Dawkins. *The Selfish Gene.* Oxford University Press, Oxford, 1974.

[Edelman and Gally, 2001] G. M. Edelman and J. A. Gally. Degeneracy and complexity in biology systems. *Proc. Natl. Acad. Sci. USA*, 98, 13763–13768, 2001.

[Elman *et al.*, 1996] J. L. Elman, E. Bates, M. H., Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett. *Rethinking innateness: a connectionist perspective on development.* MIT Press, Cambridge MA, 1996.

[Elsasser, 1966] W. M. Elsasser. *Atom and organism.* Princeton Univ. Press, Princeton NJ, 1966.

[Etxeberria *et al.*, 2000] A. Etxeberria, A. Moreno, and J. Umerez, eds. The contribution of Artificial Life and the Sciences of Complexity to the Understanding of Autonomous Systems. *CC AI:Comunication and Cognition-Artificial Intelligence*, 17(3-4), 2000.

[Fodor, 1975] J. A. Fodor. *The language of thought.* Thomas Crowell, New York, 1975.

[Fox Keller, 2005a] E. Fox Keller. The century beyond the gene, *J. Biosci.*, 30, 3–10, 2005.

[Fox Keller, 2005b] E. Fox Keller. Revisiting 'scale-free' networks. *BioEssays*, 27, 1060–1068, 2005.

[Fox Keller, 2007] E. Fox Keller. The disappearance of function from 'self-organizing systems'. In: F. Boogerd, F. Bruggeman, J-H. Hofmeyr, and H. V. Westerhoff (eds.), *Systems Biology: Philosophical Foundations*, Elsevier, Ámsterdam, 303-317, 2007.

[Ganti, 1987] T. Ganti. *The principles of life.* Oxford University Press, Oxford, 2003. (Rev. with notes, from 1971 and OMIKK English 1987, edns.)

[Gell-Mann, 1994] M. Gell-Mann *The quark and the jaguar.* Freeman and co., New York, 1994.

[Goodwin *et al.*, 1983] B. Goodwin, N. Holder, and C. C. Wyllie. *Development and evolution.* Cambridge University Press, Cambridge, 1983.

[Harold, 2001] F. Harold. The way of the cell: molecules, organisms and the order of life. Oxford_University Press, Oxford, 2001.

[Hartwell *et al.*, 1999] L. H. Hartwell, J. J. Hopfield, S. Leibler, and A.W. Murray. From molecular to modular cell biology. *Nature*, 42 supp, c47–252, 1999.

[Hempel and Oppenheim, 1948] C. G. Hempel and P. Oppenheim. Studies in the logic of explanation. *Philosophy of Science,* 15, 137-175, 1948.

[Hopfield, 1982] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci.,* 79, 2554-2558, 1982.

[Jacob, 1977] F. Jacob. Evolution as tinkering. *Science*, 196, 1161–1166, 1977.

[Kauffman, 1993] S. Kauffman. *The origins of order: self-organization and selection in evolution.* Oxford Univ. Press, Oxford, 1993.

[Kelso, 1995] J. A. S. Kelso. *Dynamic patterns: The self-organization of brain and behavior.* MIT Press, Cambridge MA, 1995.

[Krichmar and Edelman, 2005] J. L. Krichmar, G. M. Edelman. Brain-based devices for the study of nervous systems and the development of intelligent machines. *Artificial Life*, 11, 63–77, 2005.

[Laland *et al.*, 1999] K. N. Laland, J. Odling-Smee, and M. W. Feldman. Evolutionary consequences of niche construction and their implications for ecology. *Proc. Natl Acad. Sci. USA*, 96(18), 10242–10247, 1999.

[Langton, 1990] C. G. Langton. Computation at the edge of chaos. *Physica D*, 42, 12-37, 1990.

[Lewin, 1992] R. Lewin. *Complexity: life at the edge of chaos.* MacMillan, New York, 1992.

[Lewis, 1992] E. B. Lewis. Clusters of master control genes regulate the development of higher organisms. *Journal of the American Medical Association*, 267, 1524–1531, 1992.

[Lewis, 2005] M. D. Lewis. Bridging emotion theory and neurobiology through dynamic systems modeling. *Behavioral and Brain Sciences*, 28, 169 — 194, 2005.

[Lewontin, 2000] R. Lewontin. *The triple helix.* Harvard University Press, Harvard, 2000.

[Li and Vitányi, 1997] M. Li and P. Vitányi, P. An introduction to Kolmogorv complexity and its applications, Springer, New York, 1997.

[Mattick, 2004] J. S. Mattick. The hidden genetic program of complex organisms. *Scientific American*, *(Oct.), 61–67, 2004.

[Merleau-Ponty, 1963] M. Merleau-Ponty. *The Structure of Behavior.* Pittsburgh PA, Dusquene University Press, Pittsburgh PA, 1963.

[Mitchell, 2006] M. Mitchell. Complex systems: Network thinking. *Artificial Intelligence*, 170(18), 1194 — 1212, 2006.

[Montoya and Solé, 2002] J. M. Montoya and R. V. Solé. Small world patterns in food webs. *J. Theor. Biol.*, 214, 405–412, 2002.

[Morange, 2003] M. Morange. *Histoire de la biologie moléculaire.* La Découverte, Paris, 2003.

[Moreno and Lasa, 2003] A. Moreno and A. Lasa. From Basic Adaptivity to Early Mind. *Evolution and Cognition*, 9(1), 12–30, 2003.

[Newell, 1980] A. Newell. Physical Symbol Systems. *Cognitive Science,* 4, 135–83, 1980.

[Nicolis and Prigogine, 1977] G. Nicolis and I. Prigogine. *Self-organization in Non-equilibrium Systems.* Wiley, New York, 1977.

[Nolfi, 1997] S. Nolfi. Using emergent modularity to develop control systems for mobile robots. *Adaptive Behavior*, 5(3-4), 343 — 364, 1997.

[Pask, 1960] G. Pask. The Natural History of Networks. In M. C. Yovits, S. Cameron (eds.), *Self-organizing Systems,* Pergamon Press, Oxford, 232-263, 1960.

[Pastor-Satorras and Vespignani, 2001] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14), 3200-3203, 2001.

[Pattee, 1973] H. Pattee. The Physical Basis and Origin of Hierarchical Control. In H. Pattee (ed.), *Hierarchy Theory. The Challenge of Complex Systems.* George Braziller, New York.

[Polanyi, 1968] M. Polanyi. Life's irreducible structure. *Science*, 160, 1308 — 1312, 1968.

[Port and van Gelder, 1995] R. F. Port and T. van Gelder, eds. *Mind as motion: Explorations in the dynamics of cognition.* The MIT Press, Cambridge MA, 1995.

[Prescott, 2007] T. J. Prescott. Forced moves or good tricks in design space? Landmarks in the evolution of neural mechanisms for action selection. *Adaptive Behavior,* 15(1), 9–31, 2007.

[Rachevsky, 1938] N. Rachevsky. *Mathematical biophysics: Physico-mathematical foundations of biology.* Dover Publications, New York, 1938.

[Ravesz *et al*, 2002] E. Ravesz, A. L. Somera, D. A. Mongru, Z. N, Oltvai, and A. L. Barabassi. Hierarchical Organization of Modularity in Metabolic Networks. *Science*, 297, 1551 — 1555, 2002.

[Rosen, 1991] R. Rosen *Life itself: A comprehensive inquiry into the nature, origin and fabrication of life.* Columbia Univ. Press, New York, 1991

[Ruiz-Mirazo and Moreno, 2004] K. Ruiz-Mirazo, and A. Moreno. Basic Autonomy as a fundamental step in the synthesis of life. *Artificial Life,* 10(3), 235–259, 2004.

[Ruiz-Mirazo *et al.*, 2007] K. Ruiz-Mirazo, J. Umerez, and A. Moreno. Enabling conditions for 'open-ended evolution. *Biology and Philosophy,* 23(1), 67–85, 2007.

[Rumelhart and McClelland, 1986] D. E. Rumelhart, J. L. McClelland, and the PDP Research Group. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vo.1: Foundations.* MIT Press, Cambridge Mass., 1986.

[Schrödinger, 1944] E. Schrödinger. *What is life? The physical aspect of the living cell.* Cambridge Univ. Press, Cambridge, 1944.

[Skarda and Freeman, 1987] C. A. Skarda and W. J. Freeman. How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences,* 10, 161–195, 1987.

[Simon, 1962] H. A. Simon The Architecture of Complexity. *Proceedings of the American Philosophical Society*, 106, 467–482, 1962. Reprinted in H. A. Simon, *The Sciences of the Artificial*, 193-229, 1969.

[Solé *et al.*, 2002a] R. V. Solé, I. Salazar, and J. Garcia-Fernandez. Common Pattern Formation, Modularity and Phase Transitions in a Gene Network Model of Morphogenesis. *Physica A*, 305, 640–647, 2002.

[Solé *et al.*, 2002b] R. V. Solé, R. Ferrer, J. M. Montoya, and S. Valverde. Selection, Tinkering and Emergence in Complex Networks. *Complexity,* 8, 20–33, 2002.

[Solé *et al.*, 2007] R. V. Solé, A. Munteanu, C. Rodriguez-Caso, and J. Macía. Synthetic protocell biology: from reproduction to computation. *Phil. Trans. Roy. Soc. Lond. B,* 362, 1727 — 1739, 2007.

[Sporns *et al.*, 2005] O. Sporns, G. Tononi, and R. Kotter. The human connectome: A structural description of the human brain. *PLoS Comput Biol,* 1(4), e42, 2005.

[Sporns *et al.*, 2000] O. Sporns, G. Tononi, and G. M. Edelman. Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cereb. Cortex,* 10, 127–141, 2000.

[Steels, 1991] L. Steels Towards a theory of emergent functionality. In J. Meyer, R. Wilson (eds.), *Simulation of Adaptive Behavior.* MIT Press, Cambridge MA, 451–461, 1991.

[Strogatz, 2001] S. H. Strogatz Exploring complex networks. *Nature,* 410, 268–277, 2001.

[Thelen and Smith, 1994] E. Thelen and L. B. Smith. *A dynamic systems approach to the development of cognition and action.* Bradford Books/ MIT Press, Cambridge MA, 1994.

[Tononi *et al.*, 1999] G. Tononi, O. Sporns, and G. M. Edelman. Measures of degeneracy and redundancy in biological networks. *Proc. Natl. Acad. Sci. USA,* 96, 3257-3262, 1999.

[Tsuda, 2001] I. Tsuda Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems. *Behavioral and Brain Sciences,* 24(5), 793–847, 2001.

[Turing, 1952] A. M. Turing The chemical basis of morphogenesis. *Philos. Trans. R. Soc. Lond. B.*, 237, 37–72, 1952.

[van Gulick, 1993] R. van Gulick. Who's in Charge Here? And Who's Doing All the Work? In J. Heil and A. R. Mele (eds.), *Mental Causation*, Clarendon Press, Oxford, 233-256, 1993.

[Waddington, 1957] C. H. Waddington. *The strategy of the genes.* Macmillan, London, 1957.

[Wagner and Fell, 2001] A. Wagner and D. Fell. The small world inside large metabolic networks. *Proc. R. Soc. Lond. B,* 268, 1803-1810, 2001.

[Waldrop, 1992] M. M. Waldrop. *Complexity: The emerging science at the edge of order and chaos.* Simon & Schuster, New York, 1992

[Weaver, 1948] W. Weaver Science and Complexity. *American Scientist,* 36(4), 536-544, 1948. (Reprinted in G. J. Klir. *Facets of System Science*, 449-456, 1991.)

[Weisbuch, 1986] G. Weisbuch. Networks of automata and biological organization. *Journal of Theoretical Biology*, 121, 255-267, 1986.

[Wolfram, 1986] S. Wolfram. *Theory and applications of cellular automata.* World Scientific, Singapore, 1986.

[Yates, 1987] F. E. Yates (ed.). *Self-organizing systems the emergence of order.* Plenum Press, New York, 1987.

COMPLEXITY IN ORGANISMAL EVOLUTION

Stuart A. Newman

1 INTRODUCTION

Biological complexity can be defined as patterned, finely-tuned relationships among an organism's parts (including its anatomical structures and molecules) and their associated functions, including the processes by which form and function are conveyed from one generation to the next. (The living world also contains complex, evolving ecological systems, but this article will focus only on individual organisms.) In Warren Weaver's terminology, rather than the *disorganized complexity* of gases or liquids, life exhibits *organized* complexity [Weaver, 1948]. Complexity is generally considered to increase over the course of organismal evolution [Bonner, 1988; McShea, 1996], but there is less agreement on what this means and how it happens [Salthe, 2008].

Organismal complexity can evolve on multiple organismal levels — cell lineage (i.e., the production of a certain number of cells with a certain distribution of fates), genetic, morphological (see Section 4), or biochemical, to name just a few — and there can be different trends on different levels [McShea, 2001]. For example, there is a tendency in the evolution of multicellular animals for cell lineages to remain relatively simple despite an overall increase in organismal complexity [Azevedo, *et al.*, 2005], and for morphological evolution to occur in fits and starts [Eldredge and Gould, 1997].

Painting in broad strokes, there are three distinct scenarios for the evolution of complexity, each of them embedded in scientific or philosophical controversy. In the first scenario, complexity in form, function and development, and the increases thereof, are manifestations of the accretion of incremental adaptations over evolutionary time. Specifically, in both the original and contemporary (i.e., "Modern Synthesis") versions of Charles Darwin's theory, organisms are thought to change their characteristics gradually, in response to natural selection. In this picture, such change has no inherent directionality and is, in principle, constrained only by the requirement that it produce viable organisms consistent with the biology of the antecedent forms. According to this model, the variation in phenotype that provides the raw material for natural selection is based exclusively on genetic differences among members of a population. The possible evolutionary steps that can be taken in any lineage, therefore, are only those phenotypic modifications

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

that occur as a consequence of genetic variation. Since within the neo-Darwinian framework there is no theory, or general characterization, of the range of possible phenotypic variations, the restriction in this model to gene-associated variation does not carry any implication of a priori limits on possible morphological, functional or developmental complexity.

In the second interpretation of the evolution of complexity, phenotypic variation is not uniquely tied to genetic variation, and genetic change (or other mechanisms ensuring the persistence of the new phenotype) might occur after, rather than coordinate with, production of novel organismal forms and functions. This view takes into account the phenomenon of phenotypic plasticity [Newman, 1994; West-Eberhard, 2003], by which environmental and other contextual factors, or nonlinear effects of developmental mechanisms, allow for the existence of more than one phenotype for organisms of a given genotype, or the same phenotype for different genotypes, i.e., *one-many* and *many-one* relationships among levels of organization ([Hull, 1974]; see also [Brigandt and Love, 2008] and references therein).

Since, under this "plasticity" scenario, causal agencies other than genes enter into determination of the phenotype, it might be thought that this view would envision a potential organismal complexity even greater than within the neo-Darwinian paradigm. As will be discussed below, however, key non-genetic determinants of development (physical mechanisms of morphogenesis, for example) have constraints and directionalities of their own that produce stereotypical outcomes, thus reducing the scope of phenotypic variation and of the genes' capacity to determne it.

This second interpretation, though entirely naturalistic, is not very Darwinian. Darwin himself famously stated that "[i]f it could be demonstrated that any complex organ existed, which could not possibly have been formed by numerous, successive, slight modifications, my theory would absolutely break down" [Darwin, 1859]. Although "could not possibly" is perhaps too a high bar for any biological counterexample to surmount, we now have several examples of complex anatomical structures, such the segmented vertebral column (see Section 5), that are demonstrably not formed by (and thus are unlikely to have evolved by means of) numerous, successive, slight modifications. Rather, they arise from complex dynamical mechanisms that are inherently malleable and capable of generating variant forms with minimal genetic change [Newman and Müller, 2000; Collier and Hooker, 1999].

A third interpretation of biological complexity, advocated by pre-Darwinian naturalists and newly asserted by the contemporary parascientific[1] movement known as Intelligent Design (ID), holds that some organismal features are "specified"

[1] I use this term in preference to "pseudoscientific" since, as will become apparent, I think that aspects of the critique mounted by some of ID's proponents against gradualist/adaptationist accounts of the complexity of biological systems, whatever their underlying motivation may be, have merit. To be absolutely clear, the value I see in teleological design notions resides entirely in their challenge to the simplistic rejection of goal-directed material processes in evolutionary theory, and not in any explanations ID has proffered, or seem likely to offer in the future.

[Dembski, 2002] and "irreducible" [Behe, 1996]. Advocates of ID, while accepting of selectionist scenarios in some domains, affirm that there are no plausible naturalistic scenarios that could have brought about certain subsystem interrelationships found in organisms. In this case, complexity would have required a conscious fabricator, either a supernatural ("disembodied" in ID) god or a technologically advanced, "embodied" extraterrestrial, to bring it about.

In the *Critique of Judgment* Immanuel Kant contended that however philosophically unjustified it would be to treat the analogy between living organisms and the products of purposive design as a "constitutive concept," i.e., a property of nature, it is an all-but-inevitable "regulative concept," or heuristic, in the study of biological systems [Kant, 1790; Guyer, 1998; 2004]. Thus, for logical completeness, rather than for any scientific solutions offered by ID, it seems desirable to permit the specter of design to intrude occasionally as a guard against facile mechanistic scenarios. Even an avowed materialist like Francis Crick, in the absence of plausible earth-bound geochemical scenarios, entertained (in the form of the hypothesis of exogenesis or "directed panspermia"), intelligent extraterrestrial origins for the genetic code [Crick, 1981].

More typical than Crick's provocative speculation is the reflexive resort to adaptationism [Dennett, 1995; Dawkins, 1996] to account for nearly every aspect of an organism's form and function, a tendency criticized by Stephen Jay Gould and Richard Lewontin [1979] as analogous to Kipling's "Just-So Stories." This jibe (although not advanced in support of any notion of externally imposed design), pointed to the all-too-frequent recourse to the Darwinian default in the face of difficult cases. As mentioned above, and discussed in some detail below, there are often better explanations for complex traits than the ones of gradualist supposition.

2 THE DYNAMICAL CONSTITUTION OF MULTICELLULAR ORGANISMS

Living cells have many thousands of different components that are connected to one another by architectural relationships and biochemical transformations. Cells are also thermodynamically open systems, meaning that both energy and matter are constantly passing back and forth through their semi-permeable membranes. These characteristics place whole cells in the category of "complex dynamical systems," entities that remain displaced from the dead state represented by chemical equilibrium and are thereby capable of exhibiting a wide array of counterintuitive, active, self-maintaining and self-reinforcing "emergent" behaviors [Prigogine and Stengers, 1984; Rosen, 1991; Kauffman, 1993; Yates, 1994; Collier and Hooker, 1999]. While there are a variety of mathematical and computational techniques to analyze such systems (see, e.g., [Strogatz, 1994; Flake, 1998; Kaneko, 2006]) at all scales of cellular organization, for reasons that will become clear this article will deal primarily with multicellular assemblies, and only briefly with subcellular multi-protein assemblies.

Cellular life on Earth emerged in both its prokaryotic (i.e., bacterial and archaean) and eukaryotic (i.e., protozoan, fungal, animal, plant) forms by the amalgamation of, and gene exchange between, disparate systems operating on multiple spatial and temporal scales [Woese, 1998; Margulis and Sagan, 2002; Cavalier-Smith, 2006; Lester, *et al.*, 2006; de Duve, *et al.*, 2007]. After more than 3 billion years of evolution [Knoll, 2003], during which cellular subsystems became structurally and functionally intertwined, it is an all but hopeless task to analyze whole cells as if they were unitary dynamical systems obeying a common set of relatively simple dynamical principles.

For multicellular morphogenesis, however, the situation is simpler. For all the cellular evolution that preceded the rise of the Metazoa (i.e., the multicellular animals) in the late Precambrian-early Cambrian period (~600 million years ago), there was clearly no selection for developmental mechanisms involved the shaping and patterning of cell aggregates or tissue primordia. Nonetheless, many of the genes of the "developmental-genetic toolkit" used by the embryos of all animal phyla to perform morphogenesis-specific activities are present in the choanozoans, non-colonial relatives of the Metazoa that are descended from the same single-celled ancestors [King, *et al.*, 2008; Shalchian-Tabrizi, *et al.*, 2008]. These include genes whose products mediate cell-cell adhesion, cell-contact-mediated regulation of cell state, gradient-mediated spatial regulation of cell fate, and formation of connective tissue, none of which function as such in the choanozoans.

What appears to have happened during the transition from choanozoan ancestors to metazoans is that pre-existing surface proteins began, after a few mutations or a change in the ionic environment [Kazmierczak, Kempe, 2004], to mediate cell-cell attachment. The consequent change in spatial scale created a context in which other pre-existing molecules were able to mobilize mesoscopic (i.e., "middle-scale") physical processes and effects that were irrelevant to objects of the size and composition of individual cells [Newman and Bhat, 2008; 2009].

Multicellular systems are "excitable" [Mikhailov, 1990], "soft" [de Gennes, 1982] materials, which exhibit or sustain surface tension, viscosity, elasticity, chemical gradients, and (via synchronization of biochemical oscillations) global coordination of physiological state. Such effects, which include "self-organizing" behaviors also seen in nonliving complex dynamical systems [Prigogine and Stengers, 1984; Kauffman, 1993], can generate organismal forms that are multilayered, hollow, elongated, segmented and appendaged. These are, in fact, the typical morphological motifs of metazoan body plans and organ forms [Newman and Müller, 2000; Newman, *et al.*, 2006].

To summarize, while individual cells may themselves be too multileveled and functionally integrated to exhibit the stereotypical behaviors and outcomes characteristic of complex dynamical materials, embryos (having a briefer evolutionary history since their inception as loosely bound aggregates of cells), continue to be formed to a recognizable degree by their originating organizational principles [Newman, Comper, 1990]. (See, for example, the discussion of vertebrate segmentation in Section 5). To the extent that evolution leads to developmental outcomes becoming "canalized" [Waddington, 1942] or "autonomized" [Müller and Newman, 1999; Hooker, 2009] by the elaboration and integration of regulatory mechanisms over multiple scales, embryos will tend to the condition of individual cells, i.e., functioning in a *sui generis* fashion rather than by more physically transparent "generic" means. The relevant point here, however, is that the primitive plasticity of developmental systems and the non-genetic sources of change associated with such plasticity, were central to the evolutionary history of multicellular organisms and continue to play determining albeit constrained roles in present-day embryos [Forgacs and Newman, 2005].

3 GENETIC *VS.* PHENOTYPIC COMPLEXITY

The Modern Synthesis entails that genetic systems will often grow more and more complex with evolution, since within a given evolutionary lineage new genetic changes are necessarily superimposed on older genetic systems. The implication of such genetic complexity, however, is unclear. The fact that all organisms, from single-celled bacteria to multicellular humans, have between a few thousand and a few tens of thousand genes, as well as countless interactions among the genes and their products, makes the rules underlying any correspondence between genotype and phenotype impossible to characterize, even using the most advanced computers currently or foreseeably available. Because genetic similarity is a nebulous concept when comparing different species [Marks, 2003], the notion that the tempo and mode of phenotypic evolution parallels the extent of genetic change is coming under increasing challenge [Schwartz and Maresca, 2006].

Thus, while genetic complexity increases (by some definitions) as a consequence of evolution, it does not always correspond to common measures of phenotypic complexity. For example, the metazoans, a taxonomic group containing forms as varied in complexity as earthworms, insects, starfish, clams and humans, differ over just a two-fold range in gene numbers. And while the morphological diversification of the metazoans over a half-billion years has been accompanied by changes in both the regulatory and protein-coding regions of key genes [Carroll, 2005; Davidson, 2006; Wray, 2007], as indicated above, most of the gene regulatory mechanisms and regulatory genes (the "toolkit") utilized in animal development preceded this diversification [Larroux *et al.*, 2008], and many preceded multicellularity itself [King *et al.*, 2008]. The rapidity of the diversification [Rokas *et al.*, 2005], moreover, argues against its occurrence having depended on the emergence of novel genetic mechanisms [Newman, 2006].

A fundamental premise of the neo-Darwinian model, nonetheless, is that evolutionary change should closely track genetic change. This notion was fortified by a concept first proposed in the 1940s that an organism's genome serves as a software-like "program" for its development (see, e.g., [Schrödinger, 1944]; discussed in [Kay, 2000]). Opponents of the very reality of large-scale evolution have exploited this mostly unquestioned representation of the standard model. William Dembski, for example, has attempted to use mathematical notions of algorithmic or program-size complexity (originally devised to characterize computational problems; [Chaitin, 2001]) to demonstrate the improbability of the emergence of organisms by a process of random search [Dembski, 2002]. Although the validity of Dembski's mathematical procedures have been questioned [Shallit and Elsberry, 2004], the fatal flaw of his analysis is actually the genetic program notion he shares with his neo-Darwinist opponents.

Bringing phenotypic and developmental plasticity (manifestations of the nature of organisms as complex dynamical systems) into the evolutionary picture eliminates the theoretical need for strict linkage between genetic and phenotypic complexity [West-Eberhard, 2003]. The existence of determinants of function and form beyond the gene, including *epigenetic* factors in both the narrow [Jablonka, Lamb, 1995] and broad [Newman and Müller, 2000; Salazar-Ciudad *et al.*, 2003; Jablonka, Lamb, 2005; Callebaut *et al.*, 2007] senses, introduces condition-dependent sources of variability and novelty that can become assimilated into the genetic repertoire "after the fact." The biological outcomes of such "phenotype precedes genotype" [Palmer, 2004] scenarios are complex amalgams of genetic mechanisms and modes of plasticity rather than "genetic programs."

4 VARIATION AND INNOVATION IN THE EVOLUTION OF MORPHOLOGICAL COMPLEXITY

With respect to shape and form in multicellular organisms, two types of evolutionary change can be distinguished: those that incrementally remold an existing structure and those that lead to structures not previously present, i.e., "novelties" [Müller and Newman, 2005]. Since nothing comes from nothing, there will sometimes be disagreements over the attribution of novelty to a biological structure [Moczek, 2008]. At the extremes, however, the cases are usually clear: the reshaping of birds' beaks occurs gradually, over time, with both modern and paleontological intermediates in evidence [Weiner, 1994]; the beak itself, however, appeared fairly suddenly in the fossil record.

Darwin's mechanism of evolution by selection for increased fitness provides a plausible account of the wide array of bird beak morphologies, and indeed the Galapagos finches, with beaks variously adapted to particular food resources, were a paradigmatic example for this investigator. The molecular mechanisms underlying the shaping and reshaping of the bird beak provide a straightforward causal basis for the incremental operation of natural selection. The epithelia (embryonic skin) covering the two growing masses of tissue of the developing bird face that form the beak secrete a protein known as BMP4. This protein acts as a "morphogen," (i.e., a molecule produced by certain cells that affects the developmental fate of cells at a distance from the source), which regulates the rate of growth of the underlying mesenchyme (embryonic connective tissue) that in turn forms the species-characteristic upper and lower jaws. The relative amounts and timing [Merrill *et al.*, 2008] of BMP4 production by the epithelium, and then by the mesenchyme, can explain differences in beak shapes not only among Darwin's finches

[Abzhanov *et al.*, 2004] but also between chickens and ducks [Wu *et al.*, 2004].

Because developmentally efficacious genes such as that specifying BMP4 can be regulated in a tissue-specific fashion [Shentu *et al.*, 2003], it is a reasonable assumption that finch populations will exhibit genetic variability in the embryonic expression of BMP4 in the developing beak tissues. When confronted with new seeds or insects, as a result, perhaps, of climatological changes, genetically different subpopulations of finches would exhibit different fitnesses, and the population means of various parameters of beak size and shape would thereby change by classic Darwinian natural selection.

For completeness, however, it should also be noted that BMP4, like some other developmental factors, is regulated in a tissue-specific fashion by vitamin A and other nutrients and dietary regimes [Baleato *et al.*, 2005; Villaneuve *et al.*, 2006; Abdel-Hakeem *et al.*, 2008; Battacharyya *et al.*, 2008]. If a group of adventurous finches explored new sources of food rich in these development-perturbing factors there would be no necessary good-fit between induced beak shape changes and the ability to exploit the new resources in the birds' offspring. But where there happened to be such a correspondence, this lineage would tend to remain in its chosen niche [Odling-Smee *et al.*, 2003], perhaps becoming genetically isolated from the rest of the founding population by drift, or by stabilizing selection.

This evolutionary scenario, termed the "Baldwin effect" [Crispo, 2007], can be reconciled with Darwinian mechanisms if selection closely tracks the environmentally induced phenotypic alteration. However, in cases where the new phenotype sets off in new ecological directions, as happens with "transgressive segregation" in plants, whereby hybrids with extreme or novel phenotypes establish themselves in new niches [Rieseberg *et al.*, 1999], the challenge to the standard explanatory model would be substantially greater.

While it is thus clear that evolutionary changes in the shaping of a structure, such as a finch's beak, can arise from either Darwinian or non-Darwinian mechanisms, it is less apparent how to decide whether such changes (which follow relatively continuous trajectories and are often reversible; [Grant and Grant, 2002]), represent real increases in biological complexity. Conversely, novel structures certainly add complexity to an organism's phenotype, but it is difficult to see how they could arise by purely Darwinian means. Part of the problem is logical: Darwinian evolution is based on differential fitness of organisms that are variable in given characters, but until the character exists it cannot be variable [Müller and Newman, 2005].

A further source of confusion concerning natural selection and phenotypic novelty arises from misinterpretation of a tenet of population genetics established by R. A. Fisher, one of the founders of the Modern Synthesis. Using an abstract "geometrical model" to represent deviation, due to gene mutation, of a phenotype from some optimal value, Fisher showed that "genes (i.e., alleles) of large effect" are unlikely to remain (become "fixed") within the population [Fisher, 1930]. This engendered the widely-held view that theoretical population genetics confirms Darwin's notion (encapsulated in the quotation above concerning "numerous, successive, slight modifications") that evolution overwhelming proceeds by gradual steps, and that macroevolution can only result form microevolution over long times.

In fact, Fisher's argument has nothing to do with phenotypic innovation, since it only pertains to "quantitative traits," i.e., those like the size and shape of the finch's beak discussed above which are transformed in a continuous fashion in response to genetic variation. The effects of genes associated with the choice between *discrete* traits (i.e., the direction of coiling of a snail's shell; the number of body segments) or with the presence of morphological novelties (feathers; limbs; segmentation itself) are not addressed by Fisher's geometric model [Clarke and Arthur, 2000; Orr, 2001, 2005], although once an innovation is in place it may become a quantitative trait.

Genes of large effect, moreover, are in fact harbored by natural populations, some influencing adaptively useful characters that appear repeatedly in independently evolving lineages, such as the body armor spines on the stickleback fish [Colosimo *et al.*, 2005], and others influencing features that are non-adaptive, such as the change in identity of body segments in the fruit fly *Drosophila* [Gibson and Hogness, 1996]. Significantly, however, it is not the alleles themselves that define the effect: in different biological systems or even in different sister species [Phinchongsakuldit *et al.*, 2004], the same genetic change can lead to markedly different quantitative or qualitative outcomes.

To return to phenotypic novelties: while they are not likely to arise directly by a Darwinian process of gradual adaptation for the reasons given above, they can potentially emerge suddenly in a population, even simulataneously in multiple individuals, if the organisms are stressed during embryogenesis, environmentally [Badyaev, 2005; Badyaev *et al.*, 2005; Goodman, 2008], mechanically [Müller and Streicher, 1989], or socially-endocrinologically [Trut *et al.*, 2009]. Since development is mediated by complex cellular-biochemical systems that typically exhibit nonlinear dynamics, small mutation-induced changes in the rate or strength of component interactions can lead to abrupt ("saltational") changes in morphological outcome by the crossing of a developmental threshold. In such cases the genetic manifestation will be an allele of large effect. Novelties could also arise as "side effects" [Müller, 1990] of readjustments (possibly due to selection for an entirely different trait) in the set of cell interactions that generate developmental forms and patterns [Salazar-Ciudad *et al.*, 2003], with no single allele uniquely associated with the modification. And, since the threshold-overcoming stress may originate in the ecological or social environment, a subpopulation of organisms could sustain the transformed phenotype with no genetic change at all, so long as the precipitating environmental effect persists.

When a novelty first appears it will not typically have an immediate function and may even compromise the fitness of its carriers in the ecological setting in which it originated. But occasionally the novel character will enable a new mode of life, initiating the exploration and construction of a new niche [Odling-Smee *et al.*, 2003], and selection for persistence (epigenetic or genetic assimilation; [West-

Eberhard, 2003; Jablonka, Lamb, 2006]) will likely follow. In this case the novelty will have acquired an adaptive function, becoming (in retrospect) a preadaptation, or an "exaptation" [Gould and Vrba, 1982].

To summarize some implications of the preceding discussion that will prove useful later: while organisms can indeed evolve by incremental adaptation to new conditions based on existing phenotypic variabilty, increase in complexity of a lineage is often associated with phenotypic innovation, and this can occur abruptly, thus deviating from classic Darwinian gradualism. Moreover, while the "genotype determines phenotype" scenario (neo-Darwinism) is a biologically reasonable model for generating the raw material of both gradual and saltational evolutionary change, so is the "phenotype precedes genotype" scenario. All of these ideas flow directly from relinquishing strict genetic determinism, recognizing the existence of phenotypic and developmental plasticity, and appreciating the nonlinear and self-organizing dynamics of developmental mechanisms.

5 GENERATION OF MORPHOLOGICAL COMPLEXITY DURING DEVELOPMENT

The discussion in the preceding sections raises the possibility that present-day embryos might continue to be formed, to some extent, by the self-organizing physical processes that shaped them at their evolutionary inception. The generation of the somites of vertebrate embryos, paired blocks of tissue along the central axis that eventually give rise to the vertebrae, ribs, and the muscles of the body wall, provides a confirmation of this expectation.

Within the presomitic mesoderm (the band of embryonic tissue that organizes into the somites), certain networks of interacting gene products exhibit dynamics in which concentrations of several of proteins fluctuate periodically with time and thus act as "biochemical clocks." The clock time is tracked by the levels of one or more of the key proteins, which reach their peak values every 1- 2 hours (depending on the species). At a certain stage of development the phases of the cellular clocks become synchronous at every axial level in the presomitic mesoderm, but slightly displaced from one another in a continuous fashion along the embryo's axis. A consequence of this is that the peak concentrations of the involved molecules rise and fall in a periodic fashion in a moving front from one end to the other of the developing body [Dequéant and Pourquié, 2008]. High levels of at least one of the oscillating molecules ("S") promote cell clustering, causing tissue to both sides of the central axis of the embryo to separate into arrays of somites.

Although peaks of S sweep continuously across the embryo due to the clock phase-offset phenomenon described above, the somites are nonetheless formed as discrete structures. The way this occurs is that the tip of the embryo's tail is the source of a morphogen (analogous to BMP4 involved in beak formation), known as FGF8. High concentrations of FGF8 inhibit the ability of S to induce somites, and initially only the region of the segmental plate just beneath the head is far enough from the source to become induced when the peak arrives. During the following cycle, however, the growth-elongation of the embryo renders a new region, just behind the earlier-formed somite, sufficiently distant from the source of FGF8 to escape from its inhibiting effect. When the peak of S arrives at this region a new pair of somites will form. This continues until elongation ceases and the characteristic species-number of somites ($\sim$30-50 in mammals, birds and fish) have formed. This mechanism of somitogenesis was predicted on observational and theoretical grounds two decades before molecular evidence for it was obtained [Cooke, Zeeman, 1976].

For a molecular-genetic network to act as an oscillator it must have an appropriate set of positive and negative feedbacks, and the relative magnitudes of the associated interaction parameters must fall within a confined portion of the space of physically possible values [Goldbeter, 1996]. Thus, even if all the genes and interactions are in place, small variations in the relevant rates, due to genetic mutations or external parameters like temperature, could make a non-oscillating system oscillate, or vice versa.

The outcome of the interaction between an oscillator and a gradient (or clock and wavefront), two ubiquitous features of developing tissues, makes comprehensible the fact that segmentation, as a morphological novelty, emerged multiple times, in a relatively sudden fashion, from unsegmented ancestors of modern segmented animals.[2] It also implies that the neo-Darwinian model of selection for adaptive advantage of incrementally different phenotypes is not required to account for the origin of such novelties.

The clock-and-wavefront mechanism not only makes plausible how segmentation could have arisen in particular evolutionary lineages, but suggests how segments are likely to have been added or lost during the course of evolution. A comparative study of somitogenesis in zebrafish, chicken, mouse and snake embryos showed that all of these species use the clock-and-wavefront mechanism to control somite number. In each of them, somitogenesis is brought to an end through a process in which the presomitic mesoderm, having first increased in size, gradually shrinks until all the somites characteristic of that species have formed. In the snake embryo, however, the ratio of the period of the segmentation clock to the growth rate of the presomitic mesoderm is much higher than in mammals or birds [Gomez *et al.*, 2008]. This leads to a greatly increased number (315 in corn snakes) of smaller-sized somites.

If different phenotypically discrete versions of a character (e.g., different numbers of somites) confer elevated fitness to an organism but intermediate versions of the character are less fit, the population is considered in the neo-Darwinian model to reside on a "rugged" adaptive landscape [Kauffman and Levin, 1987]. Understanding how the transition from one phenotype to another could be mediated by genetic change, when genes of large effect are disallowed, is problematic.

[2] It has not been established that invertebrates utilize the same dynamical mechanisms as vertebrates for segmentation, although there are experimental and theoretical suggestions that this might be the case [Salazar-Ciudad *et al.*, 2001; Damen *et al.*, 2005; Fujimoto *et al.*, 2008; Pueyo *et al.*, 2008].

The situation is helped by the recognition that the dynamical-developmental system which generates somites can produce widely disparate outcomes as a result of genetic changes that alter the segmentation clock period or the cell division rate of the presomitic mesoderm. But this entails moving beyond the gradualistic Darwinian framework.

Such systems, moreover, can exhibit environment-dependent plasticity: the initiating step in an evolutionary transition may not involve genetic change at all. It has long been known, for example, that the number of somites in mice [McLaren and Michie, 1958] and snakes [Osgood, 1978] can be experimentally altered by changes in the external conditions of gestation. These findings, puzzling from the gene-centric neo-Darwinian perspective, and therefore neglected in common evolutionary narratives, are understandable when the physical-dynamic nature of the somite-forming system is taken into account [Bhat and Newman, 2009].

Somitogenesis is just one example of the utility of a dynamical systems approach in understanding evolution and development of form. Morphogenesis of the vertebrate limbs [Newman and Müller, 2005; Newman and Bhat, 2007a], lungs [Torday and Rehan, 2007; Miura, 2008] and tooth crowns [Salazar-Ciudad *et al.*, 2002] have also been analyzed in these terms. Indeed, as noted in Section 2, all the constructional features of animal bodies and their organs, including multiple tissue layers, body cavities, tubular and branched structures, and appendages, which emerged around 600 million years ago, can potentially be understood by the mobilization of physical processes that were newly efficacious with the emergence of multicellular aggregates [Newman and Bhat, 2008; 2009].

6 BIOCHEMICAL PATHWAYS AND "MOLECULAR MACHINES": IRREDUCIBLE, REDUCIBLE, OR EMERGENT?

The previous discussion has shown that with regard to multicellular organisms, whether or not Darwinian mechanisms are at work, no supernatural forces need be invoked to explain the growth of complexity over the course of evolution. As noted in Section 2, the structural motifs of animals plausibly first emerged when the products of certain genes (the "developmental-genetic toolkit") that pre-existed multicellular life came, in the multicellular context, to mobilize mesoscale physical effects [Newman and Bhat, 2008; 2009].

Mesoscale physics does not apply to protein-protein interactions, however, and it is at this level that ID advocates have mounted an assault on the ability of naturalistic evolution to have produced intricate biochemical pathways (e.g., the regulated sequence of protein interactions leading to blood clotting) and "molecular machines" (e.g., the bacterial flagellum — the proteinaceous whip that allows cells that bear it to locomote).

While mesoscale systems can sustain self-organizing processes, proteins and other entities on the smaller micro- and nanoscales typically undergo "self-assembly." Self-organization, as discussed in Section 2, refers to processes in which spatial nonuniformity or other complex organizational motifs are maintained by a flux of mass or energy through a system [Prigogine and Stengers, 1984; Collier and Hooker, 1999]. Self-organization in chemical or biological systems is a *non-equilibirum*, generally *irreversible* phenomenon in which structures form by bulk rearrangement of matter dependent on dissipation of energy.

Self-assembly, in contrast, is what what occurs when oil and water separate or when a mineral crystal or snowflake forms. Such structures, which are seen in mesoscale or nanoscale systems, are stable without any input of energy or mass. Self-assembly is an *equilibrium*, often *reversible* phenomenon of components that interact with inherent specificity [Whitesides and Grzybowski, 2002]. Self-assembled structures, once they form, may of course mediate or participate in energy-consuming processes, and the "molecular machines" found inside and on the surfaces of cells have this character [Chiu *et al.*, 2006].

Many subcellular and extracellular systems result from the self-assembly of proteins and other macromolecules. The proteins represent parts of intricately organized structures and to all appearances would not participate in the stuctures they do — cytoskeletal and extracellular matrix fibers, ribosomes, nuclear pore complexes, multisubunit enzymes — if they did not have shapes precisely suited to their roles in the larger structures. If there are only one or a few different kinds of proteins involved (as in the actin subunits of cytoplamic microfilaments or the subunit chains of extraceullar collagen fibers), it is straightforward to envision how the subunits and the larger functional structures could have evolved gradually and in concert.

The bacterial flagellum is a different story. It is a self-assembled structure consisting of approximately 20 major protein components and 20-30 additional ones with scaffolding and other accessory roles [Berg, 2003; Macnab, 2003]. The flagellum functions with uncanny precision and efficiency. Imagining a coordinated evolutionary scenario for this macromolecular assemblage is particularly difficult since most of its proteins are essential to its functioning [Pallen and Matzke, 2006]. The structure, about 10 μm long in the bacterium *Escherichia coli*, rotates bidirectionally at about 100 Hz using proton motive force (and not ATP, as in eukaryotic cilia and flagella), as its energy source [Manson *et al.*, 1977]. The flagella of some other species that are powered by sodium ions rather than hydrogen ions can rotate at over 1500 Hz (reviewed by Matzke [2003]).

Perplexity concerning the evolution of the bacterial flagellum pre-dates the recent focus on it by ID advocates (who consider it "irreducibly complex"; [Behe, 1996]). Robert Macnab, a long-time flagellum researcher, stated in a 1978 review,

> ...one can only marvel at the intricacy, in a simple bacterium, of the total motor and sensory system which has been the subject of this review and remark that our concept of evolution by selective advantage must surely be an oversimplification. What advantage could derive, for example, from a 'preflagellum' (meaning a subset of its components), and yet what is the probability of 'simultaneous' development of the organelle at a level where it becomes advantageous? (Macnab, 1978; quoted in [Matzke, 2003])

There have been several imaginative attempts to grapple with this question [Cavalier-Smith, 1987; Rizzoti, 2000; Matzke, 2003; Pallen and Matzke, 2006; Liu and Ochman, 2007], none of which are entirely self-consistent or satisfactory (see [Matzke, 2003; Pallen and Matzke 2006] on earlier models, [Liu and Ochman, 2007] on [Pallen and Matzke, 2006], and [Doolittle and Zhaxybayeva, 2007] on [Liu and Ochman, 2007]). The disagreements hinge on major issues, such as the assessment of homology of key genes, the role of lateral gene transfer, the relationship of the flagellum to subsystems that may have had/acquired different functions as antecedents or simplified derivatives.

All of these molecular models seek gradualist Darwinian scenarios for flagellar evolution. Similarly to a well-known and carefully argued philosophical deconstruction of the ID perspective (titled, revealingly, "Irreducible complexity and Darwinian gradualism"; [Draper, 2002]), they do not challenge the standard gene-centric model of determination of the phenotype. Just as the notion of the genetic program is the conceptual embodiment of genetic determinism at the organismal level, the embodiments at the nanomolecular level (i.e., protein structure and function) are the tenets that the different proteins produced by different organisms result entirely from differences at the gene sequence level, i.e., that "information" flows exclusively from gene to protein (the "Central Dogma" [Crick, 1958]), and that a protein's shape is uniquely specified by its primary sequence ("Anfinsen's postulate" [Sela *et al.*, 1957]).

It is not my purpose here to provide a model for evolution of the flagellum, but I note that all earlier scenarios for this require changes in the functional roles of individual proteins and their macromolecular complexes over the course of evolution. But for an entirely naturalistic account neo-Darwinian gradualism will probably not be adequate. Evolutionary models of molecular machines would benefit, rather, from the acknowledgement that neither the Central Dogma nor Anfinsen's postulate are universally true.

In the first place, many proteins are intrinsically unstructured [Wright and Dyson, 1999] and can "moonlight," i.e., adopt different conformations and different functions depending on their binding to other proteins [Tompa *et al.*, 2005; Dyson and Wright, 2005]. If the definition of a distinct protein includes a characterization of its functional structure, at least for members of the intrinscially unstructured class, protein identity is not determined by its primary sequence.

But even for proteins that are not intrinsically unstructured, "silent polymorphisms," i.e., codon variations that do not change primary sequence, can influence folding pathways and final three-dimensional structure of a protein, and ultimately its function [Kimchi-Sarfaty *et al.*, 2007]. This phenomenon appears to be mediated by differences the rate at which the synonymous mRNAs are translated [Komar *et al.*, 1999]. Since this rate can be influenced by microenvironmental effects [Hoe and Goguen, 1993; Proud, 2007], and not only codon usage, it is an open question as to what proportion of a cell's proteins actually conform to Anfinsen's postulate [Newman and Bhat, 2007b].[3]

[3]Prion diseases, in which a cellular protein assumes an abnormal conformation, which then

As with the phenomena of multicellular development, then, the structure and function of proteins, and the multicomponent aggregrates into which they self-assemble, are more loosely determined by genes than is generally believed. The possibility of innovation disconnected from selection and adaptation, as well as the indication of plausible pathways of molecular exaptation, must enter into any causal model for the evolution of highly complex structures such as the flagellum.

7 CONCLUSIONS

Although genetic determinism characterized thinking about the evolution of organismal complexity throughout the 20^{th} century, it left many major questions concerning the origination and innovation of penotype, and particularly form, unsolved. At present there is a general (though not universal, cf. [Rosenberg, 2006]) acknowledgment among philosophers of biology that it has run its course [Wimsatt, 1979; 2007; Dupré, 1993; Mitchell, 2003; Brigandt and Love, 2008]. Neo-Darwinism, though an all-but-necessary concomitant of the assumptions that organisms are produced by gene-based developmental programs and that their protein components are uniquely consituted by their gene-encoded primary sequences, has been slower to fade from favor, but it too is under philosophical assault [Robert, 2004; Reid, 2007; Callebaut *et al.*, 2007].

Replacing the gene-centric view of evolution is an emerging "extended synthesis" [Müller, 2007; Pigliucci, 2007] in which the theory of genes is complemented by, and in some cases replaced by, a theory of forms. As discussed in this article, the solution of long-standing evolutionary puzzles concerning the rapid appearance of body-plans and organ forms with conservation of a molecular toolkit and the apparently abrupt recruitment to unprecedented functions of novel tissue motifs and protein complexes, can be be advanced by incorporating evidence for a causal disconnection between phenotype and genotype. This includes recognition that developmental systems and proteins (and their complexes) each have inherent dynamical properties that are only loosely related to the genes associated with them, and that each type of system is capable of undergoing changes in their form and function, often in a discontinuous fashion, through means other than genetic variation.

While the complex dynamical systems perspective has thus contributed to the expansion of evolutionary theory, the multiscale, multidetermined, multi-historied nature of organisms has up till now stood in the way of its accomplishing the 20^{th} century goal of a general theory of life [Schrödinger, 1945; Bertalanffy, 1969]. What it is most successful with are models for subsystems: intermediary metabolism, gene regulatory networks, multicellular development, neuronal networks, and so forth, each operating on its own level, with its characteristic principles, rates and

acts as an infectious agent by inducing normal copies of the protein to also misfold [Prusiner, 1998], are the best-known examples of violation of Anfinsen's postulate. Until recently they have been considered rare exceptions to the unique determination of a protein's phenotype by its gene.

modes of evolutionary change. The laws of emergence and interlevel causation continue to be elusive.

In the first phase of the scientific study of evolution, from Darwin through the end of the 20^{th} century, Kant's challenge to understand the causation of the whole with respect to its parts, without recourse to intelligent design [Kant, 1790; Guyer, 1998; 2004], was held to be put to rest by the theory of natural selection. Although this did not prove sufficient to account for complex multicellular forms, newer knowledge of morphodynamics has begun to supply the missing elements "beyond the gene" [Müller and Newman, 2003]. For the origin and evolutionary transformation of complex subcellular molecular machines, however, the required theory is still to be produced. But here, incisive criticism of existing models, even from ID quarters, provides a Kantian goad to extend the naturalistic program to mechanisms that will of necessity also incorporate the dynamical plasticity inherent in every material system.

ACKNOWLEDGMENTS

I am grateful to Ramray Bhat, Clifford Hooker and Eva Jablonka for their close readings of earlier versions of this article and suggestions for impovement.

BIBLIOGRAPHY

[Abdel-Hakeem *et al.*, 2008] A. K. Abdel-Hakeem, T. Q. Henry, T. R. Magee, M. Desai, M. G. Ross, R. Z. Mansano, J. S. Torday, and C. C. Nast. Mechanisms of impaired nephrogenesis with fetal growth restriction: altered renal transcription and growth factor expression. *Am J Obstet Gynecol*, **199,** 252 e1-7, 2008.

[Abzhanov *et al.*, 2004] A. Abzhanov, M. Protas, B. R. Grant, P. R. Grant, and C. J. Tabin. Bmp4 and morphological variation of beaks in Darwin's finches. *Science*, **305,** 1462-5, 2004.

[Azevedo *et al.*, 2005] R. B. Azevedo, R. Lohaus, V. Braun, M. Gumbel, M. Umamaheshwar, P. M. Agapow, W. Houthoofd, U. Platzer, G. Borgonie, H. P. Meinzer, and A. M. Leroi. The simplicity of metazoan cell lineages. *Nature*, **433,** 152-6, 2005.

[Badyaev, 2005] A. V. Badyaev. Stress-induced variation in evolution: from behavioural plasticity to genetic assimilation. *Proc Biol Sci*, **272,** 877-86, 2005.

[Badyaev *et al.*, 2005] A. V. Badyaev. K. R. Foresman, and R. L. Young. Evolution of morphological integration: developmental accommodation of stress-induced variation. *Am Nat*, **166,** 382-95, 2005.

[Baleato *et al.*, 2005] R. M. Baleato, R. J. Aitken, and S. D. Roman. Vitamin A regulation of BMP4 expression in the male germ line. *Dev Biol*, **286,** 78-90, 2005.

[Behe, 1996] M. J. Behe. *Darwin's black box: the biochemical challenge to evolution.* The Free Press, New York, 1996.

[Berg, 2003] H. C. Berg. The rotary motor of bacterial flagella. *Annu Rev Biochem* **72,** 19-54, 2003.

[Bertalanffy, 1969] L. v Bertalanffy. *General system theory: foundations, development, applications.* G. Braziller, New York, 1969.

[Bhat and Newman, 2009] R. Bhat, S. A. Newman. Snakes and ladders: the ups and downs of animal segmentation. *J. Biosci.*, 34**,** 163-6, 2009.

[Bhattacharyya *et al.*, 2007] S. Bhattacharyya, A. Borthakur, P. K. Dudeja, and J. K. Tobacman. Carrageenan reduces bone morphogenetic protein-4 (BMP4) and activates the Wnt/beta-catenin pathway in normal human colonocytes. *Dig. Dis. Sci.*, 52**,** 2766-74, 2007.

[Bonner, 1988] J. T. Bonner. *The evolution of complexity.* Princeton University Press, Princeton, 1988.

[Brigandt and Love, 2008] I. Brigandt and A. C. Love. Reductionism in biology. In E. N. Zalta (ed.), *The Stanford encyclopedia of philosophy* `http://plato.stanford.edu/entries/reduction-biology`, Summer 2008.

[Callebaut *et al.*, 2007] W. Callebaut, G. B. Müller, and S. A. Newman. The organismic systems approach. Streamlining the naturalistic agenda. In R. Sansom, R. N. Brandon (eds.), *Integrating evolution and development. From theory to practice*, MIT Press, Cambridge MA, 25-92, 2007.

[Carroll, 2005] S. B. Carroll. *Endless forms most beautiful: the new science of evo devo and the making of the animal kingdom.* W.W. Norton & Co., New York, 2005.

[Cavalier-Smith, 1987] T. Cavalier-Smith. The origin of cells: a symbiosis between genes, catalysts, and membranes. *Cold Spring Harb. Symp. Quant. Biol.*, 52, 805-24, 1987.

[Cavalier-Smith, 2006] T. Cavalier-Smith. Cell evolution and Earth history: stasis and revolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 361, 969-1006, 2006.

[Chaitin, 2001] G. J. Chaitin. *Exploring randomness,* Springer; London, New York, 2001.

[Chiu *et al.*, 2006] W. Chiu, M. L. Baker, and S. C. Almo. Structural biology of cellular machines. *Trends Cell. Biol.*, 16, 144-50, 2006.

[Clarke and Arthur, 2000] B. Clarke and W. Arthur. What constitutes a 'large' mutational change in phenotype? *Evol. Dev.*, 2, 238-40, 2000.

[Collier, Hooker, 1999,] J. Collier, C. A. Hooker. Complexly organised dynamical systems. *Open Systems and Information Dynamics*, 6, 111-136, 1999.

[Colisomo *et al.*, 2005] P. F. Colosimo, K. E. Hosemann, S. Balabhadra, G. Villarreal, Jr., M. Dickson, J. Grimwood, and J. Schmutz, *et al.* Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science*, 307, 1928-33, 2005.

[Conway Morris, 2006] S. Conway Morris. Darwin's dilemma: the realities of the Cambrian 'explosion'. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 361, 1069-83, 2006.

[Crick, 1958] F. H. Crick. On protein synthesis. *Symp. Soc. Exp. Biol.*, 12, 138-63, 1958.

[Crick, 1981] F. Crick. *Life itself: its origin and nature.* Simon and Schuster, New York, 1981.

[Crispo, 2007] E. Crispo. The Baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution*, 61, 2469-79, 2007.

[Damen *et al.*, 2005] W. G. Damen, R. Janssen, and N. M. Prpic. Pair rule gene orthologs in spider segmentation. *Evol. Dev.*, 7, 618-28, 2005.

[Davidson, 2006] E. H. Davidson. *The regulatory genome: gene regulatory networks in development and evolution.* Elsevier/Academic Press, Amsterdam; London, 2006.

[Darwin, 1959] C. Darwin. *On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life.* J. Murray, London, 1959.

[de Duve, 2007] C. de Duve. The origin of eukaryotes: a reappraisal. *Nat. Rev. Genet.*, 8, 395-403, 2007.

[de Gennes, 1992] P. G. de Gennes. Soft matter. *Science*, 256, 495-497, 1992.

[Dembski, 2002] W. A. Dembski. *No free lunch: why specified complexity cannot be purchased without intelligence.* Rowman & Littlefield, Lanham, MD, 2002.

[Dequéant and Pourquié, 2008,] M. L. Dequéant and O. Pourquié. Segmental patterning of the vertebrate embryonic axis. *Nat. Rev. Genetics*, 9, 370-82, 2008.

[Doolittle and Zhaxybayeva, 2007] W. F. Doolittle and O. Zhaxybayeva. Evolution: reducible complexity — the case for bacterial flagella. *Curr. Biol.*, 17, R510-2, 2007.

[Draper, 2002] P. Draper. Irreducible complexity and Darwinian gradualism: A reply to Michael J. Behe. *Faith and Philosophy*, 19, 3-21, 2002.

[Dupré, 1993] J. Dupré, *The disorder of things: metaphysical foundations of the disunity of science.* Harvard University Press, Cambridge, Mass, 1993.

[Dyson and Wright, 2005] H. J. Dyson and P. E. Wright. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.*, 6, 197-208, 2005.

[Eldredge and Gould, 1997] N. Eldredge and S. J. Gould. On punctuated equilibria. *Science*, 276, 338-41, 1997.

[Fisher, 1930] R. A. Fisher. *The genetical theory of natural selection.* Clarendon Press, Oxford, 1930.

[Flake, 1998] G. W. Flake. *The computational beauty of nature: computer explorations of fractals, chaos, complex systems, and adaptation.* MIT Press, Cambridge, Mass, 1998.

[Forgacs and Newman, 2005] G. Forgacs and S. A. Newman. *Biological physics of the developing embryo.* Cambridge Univ. Press, Cambridge, 2005.

[Fujimoto *et al.*, 2008] K. Fujimoto, S. Ishihara, and K. Kaneko. Network evolution of body plans. *PLoS ONE*, 3, e2772, 2008.

[Gibson and Hogness, 1996] G. Gibson and D. S. Hogness. Effect of polymorphism in the Drosophila regulatory gene *Ultrabithorax* on homeotic stability. *Science,* 271, 200-203, 1996.

[Goldbeter, 1996] A. Goldbeter. *Biochemical oscillations and cellular rhythms: the molecular bases of periodic and chaotic behaviour.* Cambridge University Press, Cambridge, 1996.

[Goodman, 2008] R. M. Goodman. Latent effects of egg incubation temperature on growth in the lizard *Anolis carolinensis. J. Exp. Zool. Part A Ecol. Genet. Physiol.*, 309, 525-33, 2008.

[Gomez *et al.*, 2008] C. Gomez, E. M. Özbudak, J. Wunderlich, D. Baumann, J. Lewis, and O Pourquié. Control of segment number in vertebrate embryos. *Nature*, 454, 335-9, 2008.

[Gould and Lewontin, 1979] S. J. Gould and R. C. Lewontin. The spandrels of San Marco and the panglossian paradigm. *Proc. Roy. Soc. London B*, 205, 581-598, 1979.

[Gould and Vrba, 1982] S. J. Gould and E. Vrba. Exaptation; a missing term in the science of form. *Paleobiology*, 8, 4-15, 1982.

[Guyer, 1998; 2004] P. Guyer. Kant, Immanuel. In E. Craig (ed.), *Routledge Encyclopedia of Philosophy.* London: Routledge. Retrieved August 04, 2008, from `http://www.rep.routledge.com/article/DB047SECT13`, 1998; 2004.

[Hoe and Goguen, 1993] N. P. Hoe and J. D. Goguen. Temperature sensing in Yersinia pestis: translation of the LcrF activator protein is thermally regulated. *J. Bacteriol.*, 175, 7901-9, 1993.

[Hooker, 2009] C. A. Hooker. Interaction and bio-cognitive order. *Synthese*, 166, 513-546, 2009.

[Kaneko, 2006] K. Kaneko. *Life: an introduction to complex systems biology.* Springer, Berlin; New York, 2006.

[Kant, 1790] I. Kant. *Critique of judgement* (trans., J.H. Bernard; 1951). Hafner, New York, 1790.

[Kauffman and Levin, 1987] S. Kauffman and S. Levin. Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol.*, 128, 11-45, 1987.

[Kauffman, 1993] Kauffman, S. A. *The origins of order.* Oxford University Press, New York, 1993.

[Kay, 2000] L. E. Kay. *Who wrote the book of life?: a history of the genetic code.* Stanford University Press, Stanford, Calif., 2000.

[Kazmierczak, 2004] J. Kazmierczak, S. Kempe. Calcium build-up in the Precambrian sea: A major promoter in the evolution of eukaryotic life. In J. Seckbach (ed.), *Origins*, Kluwer, Dordrecht, 329-345, 2004.

[Kimchi-Sarfaty *et al.*, 2007] C. Kimchi-Sarfaty, J. M. Oh, I. W. Kim, Z. E. Sauna, A. M. Calcagno, S. V. Ambudkar, and M. M. Gottesman. A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science*, 315, 525-8, 2007.

[King *et al.*, 2008] N. King, M. J. Westbrook, S. L. Young, , A. Kuo, M. Abedin, J. Chapman, and S. Fairclough, *et al.* The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature*, 451, 783-8, 2008.

[Knoll, 2003] A. H. Knoll. *Life on a young planet: the first three billion years of evolution on earth,* Princeton University Press: Princeton, N.J., 2003.

[Komar *et al.*, 1999] A. A. Komar, T. Lesnik, and C. Reiss. Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. *FEBS Lett.*, 462, 387-91, 1999.

[Larroux *et al.*, 2008] C. Larroux, G. N. Luke, P. Koopman, D. S. Rokhsar, S. M. Shimeld, and B. M. Degnan. Genesis and expansion of metazoan transcription factor gene classes. *Mol. Biol. Evol.*, 25, 980-96, 2008.

[Lester *et al.*, 2006] L. Lester, A. Meade, and M. Pagel. The slow road to the eukaryotic genome. *BioEssays*, 28, 57-64, 2006.

[Liu and Ochman, 2007] R. Liu and H. Ochman. Stepwise formation of the bacterial flagellar system. *Proc. Natl. Acad. Sci. U S A*, 104, 7116-21, 2007.

[Macnab, 1978] R. M. Macnab. Bacterial motility and chemotaxis: the molecular biology of a behavioral system. *CRC Crit. Rev. Biochem.*, 5, 291-341, 1978.

[Macnab, 2003] R. M. Macnab. How bacteria assemble flagella. *Ann. Rev. Microbiol*, 57, 77-100, 2003.

[Manson *et al.*, 1977] M. D. Manson, P. Tedesco, H. C. Berg, F. M. Harold, and C. Van der Drift. A protonmotive force drives bacterial flagella. *Proc. Natl. Acad. Sci. U S A*, 74**,** 3060-4, 1977.

[Margulis and Sagan, 2002] L. Margulis and D. Sagan. *Acquiring genomes: a theory of the origins of species.* Basic Books, New York, N.Y., 2002.

[Marks, 2002] J. Marks. *What it means to be 98% chimpanzee: apes, people, and their genes.* University of California Press, Berkeley, 2002.

[Matzke, 2003] N. J. Matzke. Evolution in (Brownian) space: a model for the origin of the bacterial flagellum. `http://www.talkdesign.org/faqs/flagellum.html`, 2003.

[McLaren and Michie, 1958] A. McLaren, D. Michie. An effect of the uterine environment upon skeletal morphology in the mouse. *Nature*, 181**,** 1147-1148, 1958.

[McShea, 1996] D. W. McShea. Metazoan complexity and evolution: Is there a trend? *Evolution*, 50**,** 477–492, 1996.

[McShea, 2001] D. W. McShea. Parts and integration: the consequences of hierarchy. In J. Jackson, S. Lidgard, K. McKinney (eds.) *Evolutionary patterns: growth, form, and tempo in the fossil record*, pp. 27–60. Univ. Chicago Press, Chicago IL, 2001.

[Meinhardt and Gierer, 2000] H. Meinhardt and A. Gierer. Pattern formation by local self-activation and lateral inhibition. *Bioessays*, 22**,** 753-60, 2000.

[Merrill *et al.*, 2008] A. E. Merrill, B. F. Eames, S. J. Weston, T. Heath, and R. A. Schneider. Mesenchyme-dependent BMP signaling directs the timing of mandibular osteogenesis. *Development*, 135**,** 1223-34, 2008.

[Mikhailov, 1990] A. S. Mikhailov. *Foundations of synergetics* I. Springer-Verlag, Berlin, 1990.

[Mitchell, 2003] S. D. Mitchell. *Biological complexity and integrative pluralism.* Cambridge University Press, Cambridge, UK ; New York, N.Y., 2003.

[Miura, 2008] T. Miura. Modeling lung branching morphogenesis. *Curr. Top. Dev. Biol.*, 81**,** 291-310, 2008.

[Moczek, 2008] A. P. Moczek. On the origins of novelty in development and evolution. *Bioessays*, 30**,** 432-47, 2008.

[Müller, 1990] G. B. Müller. Developmental mechanisms at the origin of morphological novelty: A side-effect hypothesis. In M. Nitecki (ed.), *Evolutionary innovations*, University of Chicago Press, Chicago, 99-130, 1990.

[Müller and Newman, 1999] G. B. Müller and S. A. Newman. Generation, integration, autonomy: three steps in the evolution of homology. *Novartis Found. Symp.*, 222**,** 65-73, 1999.

[Müller and Newman, 2003] G. B. Müller and S. A. Newman, eds. *Origination of organismal form: beyond the gene in developmental and evolutionary biology.* MIT Press, Cambridge Mass., 2003.

[Müller and Newman, 2005] G. B. Müller and S. A. Newman. The innovation triad: an EvoDevo agenda. *J. Exp. Zool. B (Mol. Dev. Evol.)*, 304**,** 487-503, 2005.

[Müller and Streicher, 1989] G. B. Müller and J. Streicher. Ontogeny of the syndesmosis tibiofibularis and the evolution of the bird hindlimb: a caenogenetic feature triggers phenotypic novelty. *Anat. Embryol.*, 179**,** 327-39, 1989.

[Newman, 1994] S. A. Newman. Generic physical mechanisms of tissue morphogenesis: A common basis for development and evolution. *Journal of Evolutionary Biology*, 7**,** 467-488, 1994.

[Newman, 2006] S. A. Newman. The developmental-genetic toolkit and the molecular homology-analogy paradox. *Biological Theory*, 1**,** 12-16, 2006.

[Newman and Bhat, 2007a] S. A. Newman and R. Bhat. Activator-inhibitor dynamics of vertebrate limb pattern formation. *Birth Defects Res C Embryo Today*, 81**,** 305-319, 2007a.

[Newman and Bhat, 2007b] S. A. Newman and R. Bhat. Genes and proteins: dogmas in decline. *J. Biosci.*, 32**,** 1041-3, 2007b.

[Newman and Bhat, 2008] S. A. Newman and R. Bhat. Dynamical patterning modules: physico-genetic determinants of morphological development and evolution. *Phys. Biol.,* 5, 15008. 2008.

[Newman and Bhat, 2009] S. A. Newman and R. Bhat. Dynamical patterning modules: a "pattern language" for development and evolution of multicellular form. *Int. J. Dev. Biol.,* 53, 693-705, 2009.

[Newman and Comper, 1990] S. A. Newman and W. D. Comper. 'Generic mechanisms of morphogenesis and pattern formation. *Development*, 110**,** 1-18, 1990.

[Newman and Müller, 2000] S. A. Newman and G. B. Müller (2000). Epigenetic mechanisms of character origination. *J. Exp. Zool. B (Mol. Dev. Evol.)*, 288**,** 304-17, 2000.

[Odling-Smee *et al.*, 2003] F. J. Odling-Smee, K. N. Laland, and M. W. Feldman *Niche construction: the neglected process in evolution*. Princeton University Press, Princeton N.J., 2003.

[Orr, 2001] H. A. Orr. The "sizes" of mutations fixed in phenotypic evolution: a response to Clarke and Arthur. *Evol. Dev.*, 3,121-3; discussion 124, 2001.

[Orr, 2005] H. A. Orr. Theories of adaptation: what they do and don't say. *Genetica*, 123**,** 3-13, 2005.

[Osgood, 1978] D. W. Osgood. Effects of temperature on the development of meristic characters in *Natrix fasciata*. *Copeia*, 1, 33-47, 1978.

[Pallen and Matzke, 2006] M. J. Pallen and N. J. Matzke. From *The Origin of Species* to the origin of bacterial flagella. *Nat. Rev. Microbiol.*, 4, 784-90, 2006.

[Palmer, 2004] A. R. Palmer. Symmetry breaking and the evolution of development. *Science*, 306, 828-33, 2004.

[Phinchongsakuldit *et al.*, 2004] J. Phinchongsakuldit, S. MacArthur, and J. F. Brookfield. Evolution of developmental genes: molecular microevolution of enhancer sequences at the Ubx locus in *Drosophila* and its impact on developmental phenotypes. *Mol. Biol. Evol.*, 21, 348-63, 2004.

[Pigliucci, 2007] M. Pigliucci. Do we need an extended evolutionary synthesis? *Evolution*, 61, 2743-9, 2007.

[Prigogine and Stengers, 1984] I. Prigogine and I. Stengers. *Order out of chaos : man's new dialogue with nature*. New Science Library: Distributed by Random House, Boulder Colo., 1984.

[Proud, 2007] C. G. Proud. Signalling to translation: how signal transduction pathways control the protein synthetic machinery. *Biochem. J.*, 403, 217-34, 2007.

[Prusiner, 1998] S. B. Prusiner. Prions. *Proc. Nat. Acad. Sci. USA*, 95**,** 13363-13383, 1998.

[Pueyo *et al.*, 2008] J. I. Pueyo, R. Lanfear, and J. P. Couso. Ancestral Notch-mediated segmentation revealed in the cockroach *Periplaneta Americana*. *Proc. Natl. Acad. Sci. USA*, 105, 16614–16619, 2008.

[Reid, 2007] R. G. B. Reid. *Biological emergences: evolution by natural experiment*. MIT Press, Cambridge, Mass., 2007.

[Rieseberg *et al.*, 1999] L. H. Rieseberg, M. A. Archer, and R. K. Wayne. Transgressive segregation, adaptation and speciation. *Heredity*, 83, 363-72, 1999.

[Rizzotti, 2000] M. Rizzotti. *Early evolution: from the appearance of the first cell to the first modern organisms*. Birkhäuser Verlag, Basel, 2000.

[Robert, 2004] J. S. Robert. *Embryology, epigenesis, and evolution: taking development seriously*. Cambridge University Press, New York, 2004.

[Rokas *et al.*, 2005] A. Rokas, D. Kruger, and S. B. Carroll. Animal evolution and the molecular signature of radiations compressed in time. *Science*, 310, 1933-8, 2005.

[Rosen, 1991] R. Rosen. *Life itself: a comprehensive inquiry into the nature, origin, and fabrication of life*. Columbia University Press, New York, 1991.

[Rosenberg, 2006] A. Rosenberg. *Darwinian reductionism, or, How to stop worrying and love molecular biology*. University of Chicago Press, Chicago, 2006.

[Salazar-Ciudad and Jernvall, 2002] I. Salazar-Ciudad and J. Jernvall. A gene network model accounting for development and evolution of mammalian teeth. *Proc. Nat. Acad. Sci. USA*, 99, 8116-20, 2002.

[Salazar-Ciudad *et al.*, 2003] I. Salazar-Ciudad, J. Jernvall, and S. A. Newman. Mechanisms of pattern formation in development and evolution. *Development*, 130, 2027-37, 2003.

[Salazar-Ciudad *et al.*, 2001] I. Salazar-Ciudad, R. Solé, and S. A. Newman. Phenotypic and dynamical transitions in model genetic networks. II. Application to the evolution of segmentation mechanisms. *Evol. & Develop.*, 3, 95-103, 2001.

[Salthe, 2008] S. N. Salthe. Natural selection in relation to complexity. *Artificial Life*, 14, 363–374. 2008.

[Schrödinger, 1945] E. Schrödinger. *What is life? The physical aspect of the living cell*. The Macmillan company, New York, 1945.

[Schwartz and Maresca, 2006] J. H. Schwartz, and B. Maresca. Do molecular clocks run at all? A critique of molecular systematics. *Biological Theory*, 1, 357-371, 2006.

[Sela *et al.*, 1957] M. Sela, F. H. White, Jr., and C. B. Anfinsen. Reductive cleavage of disulfide bridges in ribonuclease. *Science*, 125, 691-2, 1957.

[Shalchian-Tabrizi *et al.*, 2008] K. Shalchian-Tabrizi, M. A. Minge, M. Espelund, R. Orr, T. Ruden, K. S. Jakobsen, and T. Cavalier-Smith. Multigene phylogeny of choanozoa and the origin of animals. *PLoS ONE*, 3, e2098, 2008.

[Shallit and Elsberry, 2004] J. Shallit and W. Elsberry. Playing games with probability: Dembski's "Complex Specified Information." In M. Young and E. Taner (eds.), *Why intelligent design fails: A scientific critique of the new creationism*. Rutgers University Press, New Brunswick N. J., 2004.

[Shentu *et al.*, 2003] H. Shentu and H. J. Wen, G. M. Her, C. J. Huang, J. L. Wu, S. P. Hwang. Proximal upstream region of zebrafish bone morphogenetic protein 4 promoter directs heart expression of green fluorescent protein. *Genesis*, 37, 103-12, 2003.

[Strogatz, 1994] S. H. Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Perseus, Cambridge Mass., 1994.

[Tompa *et al.*, 2005] P. Tompa, C. Szasz, and L. Buday. Structural disorder throws new light on moonlighting. *Trends Biochem. Sci.*, 30, 484-9, 2005.

[Torday and Rehan, 2007] J. S. Torday, and V. K. Rehan. The evolutionary continuum from lung development to homeostasis and repair. *Am. J. Physiol. Lung Cell Mol. Physiol.*, 292, L608-11, 2007.

[Trut *et al.*, 2009] L. Trut, I. Oksina, and A. Kharlamova. Animal evolution during domestication: the domesticated fox as a model. *BioEssays*, 31, 349-360, 2009.

[Turing, 1952] A. Turing. The chemical basis of morphogenesis. *Phil. Trans. Roy. Soc. Lond. B*, 237, 37-72, 1952.

[Villeneuve *et al.*, 2006] L. A. Villeneuve, E. Gisbert, J. Moriceau, C. L. Cahu, and J. L. Zambonino Infante. Intake of high levels of vitamin A and polyunsaturated fatty acids during different developmental periods modifies the expression of morphogenesis genes in European sea bass (*Dicentrarchus labrax*). *Br. J. Nutr.*, 95, 677-87, 2006.

[Waddington, 1942] C. H. Waddington. Canalization of development and the inheritance of acquired characters. *Nature*, 150, 563-565, 1942.

[Weaver, 1948] W. Weaver. Science and complexity. *American Scientist*, 36, 536-544, 1948.

[Weiner, 1994] J. Weiner. *The beak of the finch: a story of evolution in our time*. Knopf, New York, 1994.

[West-Eberhard, 2003] M. J. West-Eberhard. *Developmental plasticity and evolution*. Oxford University Press, Oxford, 2003.

[Whitesides and Grzybowski, 2002] G. M. Whitesides, B. Grzybowski. Self-assembly at all scales. *Science*, 295, 2418-21, 2002.

[Wimsatt, 1979] W. C. Wimsatt. Reduction and reductionism. In P. D. Asquith, H. E. Kyburg, Jr. (eds.), *Current research in the philosophy of science*, Philosophy of Science Association, East Lansing Mich., 352—377, 1979.

[Wimsatt, 2007] W. C. Wimsatt. *Re-engineering philosophy for limited beings: piecewise approximations to reality*. Harvard University Press, Cambridge Mass., 2007.

[Woese, 1998] C. Woese. The universal ancestor. *Proc. Natl. Acad. Sci. USA*, 95, 6854-9, 1998.

[Wray, 2007] G. A. Wray. The evolutionary significance of cis-regulatory mutations. *Nat. Rev. Genet.*, 8, 206-16, 2007.

[Wright and Dyson, 1999] P. E. Wright and H. J. Dyson. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.,* 293**,** 321-31, 1999.

[Wu *et al.*, 2004] P. Wu, T. X. Jiang, S. Suksaweang, R. B. Widelitz, and C. M. Chuong. Molecular shaping of the beak. *Science,* **305**, 1465-1466, 2004.

[Yates, 1994] F. E. Yates. Order and complexity in dynamical systems: Homeodynamics as a generalized mechanics for biology. *Mathematical Comput. Modelling*, **19,** 49-74, 1994.

THE COMPLEXITY OF CELL-BIOLOGICAL SYSTEMS

Olaf Wolkenhauer and Allan Muir

PREFACE

The structure of the essay is as follows:

Introduction: The complexity of cell-biological systems has an "inherent" basis, related to the nature of cells (large number and variety of components, non-linear, spatio-temporal interactions, constant modification of the components) and arises also from the means we have for studying cells (technological and methodological limitations).

Section 1: Ontological and epistemological questions are intertwined: to study the nature of living systems, we require modeling and abstraction, which in turns requires assumptions and choices that will influence/constrain what we can know about cells. We study cells, organs and organisms at a particular, chosen *level*, are forced to select subsystems/parts of a larger whole, and pick a limited range of technologies to generate observations. The study of cell biological systems requires a pragmatic form of reductionism and the interpretation of data and models is subsequently linked to a *context*. The framework to develop (mathematical) models is systems theory.

Systems theory is the study of organization *per se*. While investigations into the *structural (material) organization* of molecules and cells have dominated molecular and cell biology to this day, with the emergence of systems biology there is a shift of focus towards an understanding of the *functional organization* of cells and cell populations, i.e., the processes ("laws" and "mechanisms") that determine the cell's or organ's behavior. The necessity to select a level and subsystem, leads inevitably to a conceptual close in the theory of dynamical systems: by classifying observables into dependent, independent and invariant ones (parameters) we draw a boundary between an interior and exterior.

The main challenge is then to assemble a coherent whole from an (partial) understanding of its parts. We argue that this is only possible through an iterative process of modeling, design of experiments, further modeling and so on, in which hypotheses about the whole guide interim models.

Section 2: Applying systems theory to molecular and cell biology, we seek an understanding of structural and functional organization of the subcellular and macroscale level. The cell's functional organization at subcellular level can be

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

grouped into three classes of processes: gene expression, metabolism, and cell signaling. This classification involves a range of technologies for each class, leading to an operational division.

Section 3: The preservation of genomic properties through evolution motivates the notion of "model organisms". Unfortunately, even the simplest model organism is very complex and we give examples of the practical considerations involved in understanding so called "pathways". Not only are pathways, cells, organs or organisms complex, the structures of knowing are similarly complex. The concept of a "gene" and the notion of a "pathway" are examples of tools for understanding that develop. Our discussion highlights the importance of discussing how we try to make sense of observations in molecular and cell biology.

Section 4: With hundreds or thousands of components that need to be considered as actors in a pathway/network/subsystem, regardless of how sophisticated the technologies are, our cognitive skills and mathematical tools seem very limited to two- dimensional visualizations and only a handful of system variables. We criticize the suggestion that methods of "artificial intelligence" could avoid thinking of the experimentalist — data do *not* speak for themselves. Mathematical modeling remains an art that (fortunately) cannot be automated.

Section 5: In systems theory objects and relations between objects have identical ontological status. We emphasized above the focus on the cell's behavior (functionality) as a consequence of spatio-temporal interactions of molecules. At any level of a organism, its subsystems are interacting objects whose relationships and properties are largely determined by their function in the whole. While we can study a liver cell in isolation to investigate its stimulus-response behavior, we will only understand the cell's function fully by considering the cell and its environment as an undivided whole. The whole-part relationship emerges as a major stumbling block in dealing with the complexity of cell biological systems.

Section 6: The cells key functions include growth, proliferation, differentiation and apoptosis. Mathematical modeling of any of these processes seeks simplifications to reduce their behavior to its essence, to extract a principle that serves as an explanation. We argue that mathematical modeling is the art of making appropriate assumptions, balancing necessary reductions/approximations due to experimental/methodological limitations with abstractions serving explanatory purposes. This is true for any level (atoms, molecules, cells, orgas, and organisms) and since at any level there is another level above or below, every model is rendered macroscopic or phenomenological. While physics-style mechanial models of interacting mass points are not meaningful in systems biology, the attribute "phenomenological" does not imply arbitrariness in the construction of these models and their explanatory power — to paraphrase G. E. Box: all models are wrong, some are useful.

Section 7: In previous sections we highlighted the fact that living systems can be investigated at different levels, but also processes at subcellular and macroscale can be understood in terms of organizational levels (e.g. gene expression, metabolic networks and signal transduction pathways). The concept of a *domain of auton-*

omy for different levels, suggests a systems-theoretic framework to identify levels of bounded autonomy as subsystems that can be studied in relative isolation, while preserving a chance to understand the larger whole from knowledge about domains of autonomy. A rudimentary body of theory exists and we believe further research into such theoretical concepts and their application in systems biology could lead to practical tools, taking us a small step further in an attempt to resolve the tight whole-part relationship discussed in previous sections.

Section 8: While previous sections focused on dealing with complexity, the necessary reduction to subsystems will introduce uncertainty. The isolated view of subsystems, the necessity of ignoring observables, the inability to keep external variables constant in an experimental set-up, motivate stochastic model formalisms to capture uncertainty in form of *stochasticity*. While this form of stochasticity emerges from epistemological considerations, evolution is an example of purposeful randomness (required to generate alternatives/variations). We briefly discuss the semantics of deterministic vs. stochastic models.

Section 9 concludes our discussion with a summary of key points and an outlook on the field of systems biology. The tight whole-part relationship and the fact that ontological aspects of molecular and cell-biological systems are intertwined with epistemological questions lets us conclude that philosophers of science could actively contribute to the developments of the life sciences. There is a long history of dynamical and mathematical systems theory during which concepts of self-organization, emergence, feedback or system identification have been developed. Studying the difference between physical and biological systems, between living and non-living systems and studying the means by which we have investigated such systems, could improve our chances of managing the complexity of cell-biological systems. In the words of Ludwig Wittgenstein: "The fact that we can describe the motions of the world using Newtonian mechanics tells us nothing about the world. The fact that we do, does tell us something about the world."

INTRODUCTION

Cells are basic building blocks of living systems. Whether one considers a single cell, a colony of bacterial cells or populations of mammalian cells that form tissue, organs and whole organisms, the attribute "complex" is appropriate for any of these systems. An initial, intuitive analysis identifies for the complexity of living systems the following sources:

- Cells are composed of a very large number and variety of components interacting in space and time.
- Cell-biological systems are difficult to observe.
- The dynamic functioning of cells is of a nonlinear nature.
- Living systems are subject to continuous change.

The biologist Ernst Mayr [2004] argued that it is owing to their complexity, that biological systems have the capacity to reproduce, replicate, grow, adapt and evolve: new biological properties can emerge from others. The process that exemplifies the dynamic nature of cell-biological systems is the cell cycle. The *cell cycle* is the series of events leading to the cell's replication. Initially the cell grows, accumulating nutrients and duplicating its components needed for "mitosis", the phase during which the cell is duplicating its DNA. Finally, the cell splits itself into two distinct "daughter" cells. The cell-division cycle is a vital process that underlies the development of an organism and the maintenance of tissue. The cell cycle is a process subject to tight control, which includes the detection and repair of genetic damage, and provision of various checks to prevent uncontrolled cell division. The molecular events that control the cell cycle are ordered and directional; that is, each process occurs in a sequential fashion and it is impossible to "reverse" the cycle [Morgan, 2006].

A striking feature of living systems (organisms) is that their parts interact, modify and create themselves so as to realize an autonomous self-fabricated, self-organized whole. In living systems nothing remains constant; everything is in a perpetual state of transformation; everything comes from other things and gives rise to other things. This is true for the macromolecules making up a cell, as well as for the cells that form organs, which in turn make up whole organisms. A feature of living systems is that "the whole is in the parts", that is, each cell contains a copy of the genetic information for the whole organism. Related is the notion of *emergence*, which asserts that the whole is also more than the (logical) sum of its parts: The system as a whole displays a behavior that could not be predicted from studying its subsystems. Taken together, this provides the basis for a tight whole-part relationship, which is observed at each level of the system — molecules, organelles, cells, organs, organisms. As we shall discuss in detail below, the complexity of cell-biological systems, forces us to study subsystems/parts, raising the question of what we can then learn about the system as a whole?

The term "self" in "self-organization" suggests a form of closure of the system to outside influences, a kind of autonomy. Physically a cell is an open system which relies on a constant exchange of energy, matter and information with its environment. The closure of a living system is therefore not with respect to material causes but with respect to *efficient causation*:[1] In a living system each part/process is at once cause and effect, a means and an end — the cell is a self-organizing biochemical system that fabricates itself [Wolkenhauer, *et al.*, 2007]. Living systems are thus self-referential, every part owes its existence/explanation to the organization of the remaining parts. Through cyclic self-organizing and self-maintaining processes, parts generate the whole as does the whole define the context in which its parts function. The principle of autonomous self-fabrication is a, if not *the*, fundamental property that distinguishes living from non-living

[1]Following Aristotles classification of causes, the material cause of a phenomenon is related to the matter of what something is made of, efficient causation refers to the source of change or process that produces something.

systems. "Autopoiesis"[2] is a terminological framework to discuss the nature of living systems.[3] The ideas are due to Humberto R. Maturana and Francisco J. Varela [1987]. The boundary of an autopoietic system (e.g. the membrane in the case of cells), that is the space in which its components exist, is actively produced by the network of processes that define the system. While acknowledging the importance of self-organization/self-fabrication in living systems, we shall not discuss this theme in greater detail. Instead we refer to an early discussion of these concepts by Immanuel Kant [1892]:[4]

> "In such a natural product as this every part is thought of owing its presence to the agency of all the remaining parts, and also existing for the sake of the others and of the whole, that is an instrument, or organ. But this is not enough — for it might be an instrument of art. ... On the contrary the part must be an organ producing the other parts — each, consequently, reciprocally producing the others. No instrument of art can answer to this description, but only the instrument of that nature from whose resources the materials of every instrument are drawn — even the materials for instruments of art. Only under these conditions and upon these terms can such a product be an organized and self-organizing being, and, as such, be called a natural end."

We define the *structural organization* of a system as the configuration that relates material components of the system. Structural changes of the cell do not necessarily imply changes to the *functional organization*, that is, the processes that determine the cell's dynamic behavior, it's functioning. For example, the three-dimensional structure of a molecule indicates binding partners but does not define its "function". The function of a molecule is understood as its role in processes: changes in the population/concentration of a protein within a region of the cell underpin a process related to a cell function, i.e., growth, proliferation, differentiation (the process by which cells transform into specialized cell types) or apoptosis ("programmed" cell death). The behavior of such dynamical systems, realized through networks of biochemical reactions, would subsequently be analyzed with respect to the system's stability, responsiveness, robustness and sensitivity with respect to change in parameters. The function of a protein in such a network may thus be a contribution that leads to the network functioning as a "switch", "amplifier", "oscillator" or "filter". The theory of dynamical systems provides the conceptual framework to investigate the system's behavior as a result of feedback mechanisms [Tyson, *et al.*, 2003; Novák, *et al.*, 2008].

Studying biological cells we require technologies to observe their behavior and methodologies to construct models. Due to the complexity of cells, existing tech-

[2] The term *autopoiesis* is constructed from *autos*=self and *poiesis*=generation or production.

[3] We note that the emergence of life itself and its chemical origins are important questions relevant in this context. We refer to Luisi [2006] for a recent discussion of this issue.

[4] The quote of Kant ([1892], Section *Critique of Telelogical Judgement*, §65) can also be found in Ernst Cassirer [1981, p. 336].

nologies to generate quantitative measurements are restricted to a particular context in which they can be interpreted. Similarly, we lack suitable methodologies to analyze large scale nonlinear dynamical systems. There is no comprehensive approach to study the functioning of living systems at all levels (molecules, organelles, cells, organs, organisms). Therefore, by considering the multitude of technologies by which we generate measurements and the range of methodologies we employ towards an explanation of observed phenomena, organisms appear to be characterized by an unlimited set of qualities. The notion of a *qualitative infinity of nature* is due to the physicist David Bohm [1957], who noted that at any level at which we study a material system, the fundamental qualities and properties defining the modes of being of the system are limited in number. On the other hand, for every level there always appears to be another lower (respectively higher) level of description; such that the richness of properties and qualities apparently never approaches exhaustion.

Having settled on a particular level at which one investigates a living system (say the "cell level", studying spatio-temporal changes in protein concentrations), the large number of components forces us to decompose a larger whole into smaller, tractable parts for which we construct models. The process, by which a model is established, relies on assumptions and approximations. These may be for mathematical convenience but there is also the fact that in modeling we put a high value on simplicity: the aim of modeling is a reduction of complexity, an abstraction to extract essential properties of a system in a compact form, so as to formulate *generic principles*,[5] laws or mechanisms.

Taken together, the decision for a particular level, the limited range of technologies, the decomposition into subsystems, and the necessary simplifications in modeling a system, we find that the question of how a system functions and the process by which we describe the system are intertwined: what we can know about the system depends on the availability and choice of technologies and methodologies with which we probe the system. Adopting the words of the physicist Werner Heisenberg: "What we can observe is not nature itself, but nature exposed to our method of questioning".[6]

1 SOME ELEMENTS OF SYSTEMS THEORY

As a consequence of the complexity of biological systems, full understanding of cells and their function(ing) cannot be assured. Instead, simplified hypotheses must be formulated and tested by experiments. This requires a conceptual frame-

[5]The word "law" suggests universality of the principles discovered. A difference between physics, where one seeks "natural laws" and biology is that in biology everything seems to make sense only in a context (defined by the chosen organism, cell type, experimental set up, level etc.). Any system is composed of subsystems and is located in a super-system, leaving modeling to be an endless and always provisional exercise.

[6]In Werner Heisenberg: *Physics and Philosophy: The Revolution in Modern Science* [1958] Lectures delivered at University of St. Andrews, Scotland, Winter 1955-56.

work appropriate for making precise and empirically testable predictions. Such a framework is provided by (dynamical) systems theory. A central theme of this text is thus the role of modeling, as a means to formulate and test hypotheses. A mathematical *model* is a representation, a simplified version of the part of the biological system studied, one in which exact calculations and deductions are possible. An obvious priority in modeling is assurance that the model's behavior (established through numerical simulation or formal analysis) corresponds closely to the empirical behavior of the biological system, that the formal system (read "mathematical model") in some way resembles the behavior of the natural system. In addition to replication/reproduction of certain observed qualities or behavior, simplicity and mathematical tractability can be important criteria in developing a model.

A natural system (e.g. a cell or organism) is our interpretation of observable facts in the light of a formal system that we ourselves invent/construct. Understanding a complex system requires abstraction, reducing one type of reality to another. Mathematical modeling facilitates understanding through abstraction. If we are to describe the mechanisms/principles/laws by which the components of a system interact (and thereby realize the (sub)system functions), then the purpose of the model is to distill something complex to a simpler, essential aspect. Modeling does therefore imply for most cases a reduction of complexity; a model is then understood as an excerpt or selection from the biological system under consideration.

Abstraction can reveal truths, but never the complete truth. The growing understanding it affords will push us to the boundary of its validity, where it will eventually mislead if it is not continually re-examined. No abstraction is fully forced upon us, however evidently it appears to suggest itself. It always involves choice, those choices resulting from our standpoint — the position from which, and the purposes for which, we view the system — the level at which we are equipped to examine the system and how extensive are our powers so to do. Diverse approaches, by researchers with different expertise and interests, result in different abstractions which, taken together and made coherent with each other to form a more comprehensive abstraction, can enrich our understanding of the totality.

Fundamental philosophical issues present themselves quite concretely within a scientific discipline: ontology describes basic assumptions about the specific object of study; epistemology concerns itself with the tools we have for obtaining knowledge of that object. Maybe ontology is merely self-deluding epistemology. However the basic materialist presumption of science is that epistemology follows ontology: there is 'what is so' and then there is the extent to which we *know* 'what is so'. The separation, or better the "inseparability", between what the things are "in themselves" and how we represent observed phenomena, will haunt us throughout the essay. The mingling of ontological and epistemological problems tells us how important it is to reflect upon the modeling process itself, the diversity of approaches by which we make sense of observations and limitations these may pose on what we can know about the natural systems under consideration. As shall hopefully become clear towards the end of this essay, scientists in the life sciences (biotechnology, biomedicine, genomics, molecular and cell biology, systems biology) should or could benefit from interactions with philosophers of science.

A basic ontological assumption of all life sciences is that biological entities are self-regulating through closed causal loops — through *feedback*, that is. The appropriate language for discussing such matters is that developed in *systems theory* [Klir, 1991]. Systems theory is the study of organization *per se*, a general system being understood as a set of interrelated objects, organization being the form of interdependence of objects. For some authors systems theory, and as a consequence systems biology, is essentially the study of organization through *mathematical* analysis.[7] In this essay we shall treat systems theory as a branch of mathematics. An exposition of a rather general mathematical setting was given by Mesarovic and Takahara [1970; Mesarovic, *et al.*, 1975]. A more specialized framework is the theory of dynamical systems [Katok, 1995; Wiggins, 2003]. For a discussion of self-organization in physico-chemical (nonlinear, nonequilibrium) systems through nonlinear systems theory we refer to the book by Gregorie Ilya Nicolis and Ilya Gregorie Prigogine [1989]. Despite its long history, the theory of nonlinear dynamical systems continues to provide various challenges for practical applications in systems biology. In contrast, the study of linear dynamical systems has found numerous applications in the physical and engineering sciences [Kalman, *et al.*, 1969; Padulo, *et al.*, 1974].

A basic assumption in systems theory is that the natural system under consideration can exist in a set of distinct *states*. The experimentalist is probing the behavior of the system through stimulating it with *inputs* to produce observable responses, or *outputs*. It is presumed that the stimuli employed have some similarity with the conditions which the system experiences in its usual context. Using these *observables* the experimentalist may or may not be able to determine which state the system is in; in principle there may be an infinite number of observables necessary for an exact representation.[8] The degree to which he can be certain about the state depends on his choice of observables selected for consideration in the model, for which measurements are possible. This involves the resolution of the measurement devices and the design of experiments, providing suitable stimuli from which the system might be identified. So an observable in a mathematical model is formulated as a mapping that relates states of the system with numbers; the collection of observables which have been chosen is then a vector function of the state.

The *identification* of a system consists not just of finding the state but in determining how that state changes over time, as a function of the history of its

[7]With the (re-)emergence of systems biology as an active and supported research field there is a natural tendency of researchers to adopt definitions of systems biology that suit their interest. We here support the view that systems biology is a new paradigm, complementary but also distinct to activities in the fields of genomics and bioinformatics [Wolkenhauer and Mesarovic, 2005; Wolkenhauer, 2007].

[8]See also the ,,infinite number of qualtities of living systems" discussed by David Bohm [1957].

previous states and the inputs to which it has been subjected. This is reflected in similar functional relations between the observables over time which may be classified into

- Observables whose values remain fixed for every state, referred to as *parameters* of the model — of course, these features which are "fixed" for the given experimental context may well vary if that context is altered.[9]
- Observables that are determined by other observables and inputs to the system.

The definition of inputs, outputs and parameters, draws a boundary, separating the system from its environment. We shall discuss this conceptual closure further below. The choices of observables and the decision about their type also define the context in which the model is valid.

From these general and abstract considerations to the application of systems theory in molecular and cell biology one has to accept further assumptions. Most important for practical applications is the assumption that neither the relationship between observables, nor the parameter values change with time. The next step for a formal analysis and simulation of a dynamical system is that the set of abstract states is replaced by a *state space* with suitable mathematical properties. The mathematical model of the system subsequently encodes relationships between state variables, for which difference or differential equations are most commonly chosen. More radically, the notion of a state in such mathematical modeling can be considered as a secondary concept, being just the functional relation between stimulus and response. This is a purely external account, a severe ontological denial which refuses independent status to states of the system which now crucially depend on the observer (modeler) and his choices of experimental methods. The state of the system (model) is then an encoding of the past behavior of the system, sufficiently informative to form (together with knowledge of the current input value) the basis for the prediction of the output of the system.

Given that numerous assumptions, approximations and simplifications that are necessary for the use of mathematical modeling and computer simulations in practical applications, the phrase "the unreasonable effectiveness of mathematics in the natural sciences"[10] has been coined. Ludwig von Bertalanffy [1969] whose work laid the foundations for theoretical biology and systems biology wrote: "Considering the inconceivable complexity of processes even in a simple cell, it is little short of a miracle that the simplest possible model — namely, a linear equation between

[9]Experimentalists refer to variables such as temperature, pH etc., which they can manipulate in the experiment as ‚parameters'. In modelling biochemical reaction networks, environmental variables like temperature and pH are frequently assumed constant leading to constant values of rate coefficients. In a general mathematical model these rate coefficients are refered to as ‚parameters'.

[10]See, for example, Eugene Wigner's "The Unreasonable Effectiveness of Mathematics in the Natural Sciences," in *Communications in Pure and Applied Mathematics*, Vol. 13, No. I (February 1960).

two variables - actually applies in quite a general number of cases." In the light of recent suggestions for computer modeling projects that should lead to virtual cells, virtual organs and even an entire virtual human,[11] one should not forget the primary role of mathematical modeling in the understanding of complex systems: a reduction of complexity, that is, a simplification through abstraction. Abstraction, while unpopular in its mathematical form amongst experimentalists, serves a practical purpose — the reduction of complex relationships to their essence.

The modeling of cell biological systems with differential equations and the treatment of cells or subcellular processes as physical systems has been criticized, notably by Robert Rosen [1985; 1991; 2000] — see also George Kampis [2003]. Rosen [2000] defines "A system is simple if all its models are simulable. A system that is not simple, and that accordingly must have a nonsimulable model, is complex." Part of Rosen's work is dedicated to show that living systems are complex in the sense that they are not Turing computable. Turing-computability encompasses the class of recursive functions and the formalism of state-based Newtonian physics is just such a recursive formalism. What this entails is that state-based Newtonian physics applies within the realm of Turing-computability, only adequate for modeling simple systems; and conversely, are inadequate for modeling complex systems [Rosen, 1991].

A system theoretic framework formalizes the idea that everything exists only in relation to something else. Causation understood as the principle of explanation of change is thus treated as a relation, not between things, but between changes of states of the system under consideration. Life is considered a relation among molecules/cells and not a property of any molecule/cell. David Bohm [1957] writes: "In nature nothing remains constant. Everything is in a perpetual state of transformation, motion and change. [..] everything comes from other things and gives rise to other things. [A]s we study processes taking place under a wide range of conditions, we discover that inside of all of the complexity of change and transformation there are relationships that remain effectively constant. [..] The necessary relationships between objects, events, conditions, or other things at a given time and those at later times are then termed causal laws."

A central idea of systems theory is that the study of any system requires consideration both of its interior, the subsystems of which it is composed, and of its exterior, the context in which it normally operates as a component of a larger system. The wholeness of a system, its self-maintaining capability against external changes within tolerable ranges, is achieved through interior causal adjustments of its component parts. Thus the relation between part and whole manifests itself dualistically, through seeing the object of study as a whole composed of parts, and seeing it as being itself part of some larger whole and so on. Logically, such a view

[11] See the news brief ,,Systems biologists hatch plan for virtual human", *Nature* Vol. 451, 879, 20 February 2008. The so called ,,Tokyo Declaration" states that ,,Recent advances in Systems Biology indicate that the time is now ripe to initiate a grand challenge project to create over the next thirty years a comprehensive, molecules-based, multi-scale, computational model of the human ('the virtual human'), capable of simulating and predicting, with a reasonable degree of accuracy, the consequences of most of the perturbations that are relevant to healthcare."

implies a tower of systems seemingly unending in both directions; we will return in Section 7 to some implications of this multi-leveledness for biological systems.[12]

In order not to grapple with the whole of this potentially infinite reality at once, some conceptual closure will be an inevitable feature of any model. Modeling, therefore, will always impoverish reality. It can proceed in two ways: either collapsing the exterior into a simple characterization, when attention is focused on the internal mechanisms operating between its component parts; or black-boxing the interior, when the intervention of the exterior is the principal concern. In any model both of these simplifications will be present, allowing us to focus on the dynamics of selected variables with the rest of the totality, the elements excluded from the dynamics, appearing as phenomenologically described contingencies. Enlarging the system to account for the sources of these contingencies incorporates them into a larger model, but again there will be unexplained elements arising from the inside or outside of this larger system. A consistent adoption of the systems approach will alert us to the actual openness of all systems, as opposed to the necessarily closed models we employ [Muir, 1982].

There is an awkward confusion of terminology in this area. We will be referring throughout by the word "reduction" to the necessary selection which occurs when forming any model. This is to be distinguished from the philosophical stance of "reductionism" which insists that higher level variables be expressible only in terms of lower level ones, without the importation of emergent factors — factors which cannot be grounded in properties of the components. This is usually contrasted with "holism" which is dismissed as resting upon such emergence; we would suggest this term be used merely to remind us that each part of a system is constrained by the context of the whole. We will also allow ourselves use of the term "emergent", but only to express the possibility that new behaviors may appear when the scope of a model is extended to take fuller account of the exterior.

The reductionism/holism controversy becomes a non-issue whenever adequate conditions permitting closure of a system (model) are formulated. Indeed, the interrelation of interior and exterior of a system is reflected in the need to consider, in any model, how reductionist and holistic descriptions are related. When confronted with the problem of giving a reductionist explanation of a system's behavior, one is inevitably guided by one's knowledge of what one is trying to explain; one cannot expect, merely by sufficiently understanding the parts, to be able to assemble them to a coherent whole without, as in doing a jig-saw puzzle, some guidance from the overall picture.

[12]Due to the space limitations for an essay we will only be able to sketch personal perspectives derived from the literature and our own work. This will unfortunately neglect a vast number of relevant publications. With regard to theories of living systems we point towards James Grier Miller's magnum opus "Living Systems" [1978] and to the theory of Memory Evolutive Systems, rooted in category theory, beautifully presented in Jean-Paul Ehresmann and Andree C. Vanbremeersch book "Memory Evolutive Systems: Hierarchy, Emergence, Cognition" [2007].

2 THE CELL AND ITS COMPLEXITY

Research programmes in biology focussing on cells generally make an implicit assumption that this will allow us to draw conclusions about higher levels of the organisation (tissue, organs etc.). Our outline of systems theory suggests the formulation of two central questions for cell biology:[13]

> *How do the components within cells interact, so as to bring about the cell's structure and realize its functioning? (The cell's interior aspect)*
>
> *How do cells interact, so as to develop and maintain higher levels of structural and functional organization? (The cell's exterior aspect)*

In the following we assume that the functioning of the cell can be roughly divided into three major classes of processes [Alberts, *et al.*, 2008]: *Metabolism* describes those processes that construct and maintain the cell; processes that realize cell growth and the duplication of the genome before cell division. Metabolism is usually related to the energy household of the cell, divided into *catabolism*, yielding energy, and *anabolism* to describe processes that use this energy to construct the components of the cell (including proteins, organelles etc.) [Fell, 1997; Cornish-Bowden, 2004]. *Cell signaling* subsumes processes of inter- and intra-cell communication and the coordination of cell function [Kholodenko, 2006]. While cell signaling is realized through the generation, modification, degradation and translocation of molecules, the primary focus of research in this field is the explanation of signal transduction, information transfer, cellular decision making and "higher-level" coordination of basic cellular processes. *Gene expression and regulation* is here defined as the process by which information, encoded in the DNA, is transcribed and translated into a *gene product* (e.g. a protein).

A system's complexity arises from three factors: the quantity, the variety and the interconnectivity of its constituent elements. An essay by Warren Weaver [Weaver, 1948] is widely considered a founding text for thinking about complexity. Weaver distinguished between *disorganized complexity* and *organized complexity*. Disorganized complexity results merely from the quantitative aspect — having a very large number of similar parts. The interactions of the parts are perceived as "largely random" suggesting methods from statistical mechanics and probability theory to understand properties of the system as a whole. Gas molecules floating around in a container are a classical example of such disorganized complexity, where one is not interested in, or able to, trace/describe the trajectory of each material component using Newton's law of motion. Models of such systems are defined in terms of distributions, and predictions are usually expressed in terms of their mean values or standard deviations. (We return to questions of randomness in Section 8). Problems of *organized complexity* on the other hand are related to systems with properties that arise from non-similarity of a variety of parts from which it is composed and in which, moreover, the organization of the interacting

[13] cf. Ricard Solé and Brian Goodwin [2000].

parts cannot be derived from a study of the parts in isolation. Moreover, this internal complexity can be matched by a corresponding external complexity of the supersystem of which it may be a part.

While live cell imaging, tracing organelles or molecules in the cell, gives the impression of disorganized complexity, it is obviously a case of organized complexity we are dealing with in systems biology. Studying the functioning of cells, the experimentalists face various practical hurdles, discussed in the following section.

3 EXPERIMENTAL METHODOLOGY

The analysis of a system includes the design of experiments to generate data, the construction of models and the possibility for predictions about the system from a model.

Complexity, of itself, throws up immense challenges to experimental methodology: selection of some overall function of a system as a suitable object for study and identification of just a few relevant features from the immense variety of components present. In biological systems experimental difficulties arise not just from the three aspects of complexity but also from problems of size — technical issues of visibility, monitoring and control.

For example, *Escherichia coli* is one of many species of bacteria living in gut flora of mammals and which measures only two micrometer in length and a cell volume of 10^{-15} Litre, containing an estimated number of 2,600,000 proteins, generated from about 4252 protein coding genes.[14] In biology "Seeing is understanding", but making cellular processes measurable (visible) is an obvious technological challenge.

A *model organism* is a species (e.g. yeast, bacterial systems, worms, fish or flies) that is extensively studied to understand particular biological phenomena, with the expectation that discoveries made in the model organism will provide insights into the workings of other organisms. In particular, model organisms are widely used to explore potential causes and treatments for human disease when human experimentation would be unfeasible or unethical. This strategy is made possible by the common descent of all living organisms and the conservation of metabolic and developmental pathways and genetic information over the course of evolution. Model organisms are often chosen on the basis that they are amenable to experimental manipulation. This usually will include characteristics such as short life-cycle, techniques for genetic manipulation and non-specialist living requirements. The complexity of human cells leads then to a situation in which *E. coli*, a bacterial cell living in the lower intestines, is used as a model to study intracellular processes occurring in mammalian cells.

The threefold division of cellular processes into metabolism, signaling and gene expression associates with each a range of specialized technologies for generat-

[14] See `http://redpoll.pharmacy.ualberta.ca/CCDB/` for the *E.coli CyberCell* Database (Information accessed November 2007).

ing experimental data. The nature of the data can differ considerably, making their integration a challenge. At the methodological level, where one is trying to model and simulate these processes a range of approaches are used, depending on what type of pathway or network one is looking at. The biochemical reactions of metabolism are organized into *metabolic pathways*, while reaction networks to do with signaling are organized into *signal transduction pathways*. Genes are sometimes regarded as nodes in a *gene regulatory network*, with inputs being proteins such as transcription factors, and outputs being the level of gene expression.

The study of metabolism, cell signaling and gene expression requires a range of technologies, often leading to an operational division of researchers into “Omics” disciplines, including “metabolomics”, “proteomics” and “transcriptomics”. While there is an obvious relation between metabolism, signaling and gene expression, the complexity of the cell, specifically the technological difficulties of measuring these processes has forced researchers to specialize with obvious consequences for the overall endeavor – we can’t see the wood for the trees. Understanding neurodegenerative diseases or cancer requires the integration of knowledge and models of metabolism, signaling and gene expression.

To model inter and intracellular processes one requires quantitative spatio-temporal data for a relatively large number of components. At present these are not available, forcing us to handle uncertainty and “reduce” complexity. For practical purposes to do with technological limitations, but also with the time and money required to conduct the experiments, a subset of components is chosen. This leads to the pragmatic notion of *pathways* or *networks* as a selected subsystem of biochemical reactions (relevant to some cell function). For example, the Mitogen-activated protein (MAP) kinase signaling pathway is a system that responds to extracellular stimuli (mitogens) and is linked to various cellular activities, such as gene expression, mitosis, differentiation, and cell survival/apoptosis. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database[15] includes about 40 proteins in a graphical representation of one variant of this particular pathway. For most experiments one will, at the present time, only be able to focus on less than 10 proteins — which ones?

One criterion for the identification/separation of subsystems is based on different time scales. For metabolic systems where processes are assumed to be at steady state on the time scale of changes in metabolite pools, one can treat sets of enzymes as modules with a defined input and output. This approach is however not applicable to cell signaling systems where the transient behavior of the fast system (transfer of the signal) has consequences for the behavior of the slow system (subsequent gene expression), which then feeds back to the faster system after a delay. Examples include systems where the signal may be encoded in oscillations (e.g. Ca signaling) or transient movements between cellular compartments (e.g. NF-κB). Epidermal growth factor stimulation of the MAPK signaling cascade causes transients on time scales of minutes and relaxes to a new quasi-steady state within an hour but subsequent consequences can take much longer to emerge;

[15]See `http://www.genome.jp/kegg/` (Information accessed November 2007).

entry of cells into the cell cycle and commitment to cell division requires several hours sustained signaling, whilst receptor internalization and recycling and gene expression alter the concentrations of the components in the signaling system also on time scales of hours. To this day, most pathways are studied in isolation, while there are always many pathways that are relevant to any one cell function. The recent concept of "cross talk", more than anything else, is a proof of failure of our initial attempts to define a subsystem.

The question of how to identify subsystems (modules, pathways etc.) as functional subunits which possess, to some degree, bounded autonomy, and how one could subsequently integrate the knowledge and models achieved into a larger whole, are the two most important challenges for systems-theoretic research. A key problem is that in a complex system the whole is more than the sum of its isolated parts. In other words the interaction of subsystems can lead to emergent behavior irreducible to the system's constituent parts considered separately. Emergent behaviors occur through *interconnectivity* – intricate causal relations across different scales and feedback.

The difficulties listed above force the experimentalist/biologist to collaborate with technologists and modelers. While nowadays there is no doubt that interdisciplinary collaborations are necessary for advances in the life sciences, it is fair to say that most scientists would prefer to rely only on their own skills. While there may be some true "hybrids" with an equally good training in biology and mathematical modeling, the complexity of cell-biological systems requires specialization (in wet-lab experimentation, the development of technologies, data analysis, mathematical modeling and simulation). The need to combine different expertise, possibly across departments, universities, countries and cultures is a complex undertaking. People might study a particular MAPK cascade because their supervisor did the same, or because of what piece of equipment they have access to, what cell lines are available, etc. — the choice is made for historical and social reasons rather than scientific ones.

Finding the forest among the trees is a major challenge for today's life sciences as no individual, no single research group or institute can provide all of the necessary expertise, technologies and experimental systems. Studying living systems requires an interdisciplinary approach. The complexity of cells makes it necessary that highly specialized experts from different disciplines communicate and this often across long physical distances, countries and cultures. So we can add to our earlier list a further group of problems for the experimentalist, which arise from the social context, *interdisciplinarity:* specialisation, education, distances between institutes.

If epistemology is to follow ontology in the field of systems biology, the structures of knowing should mimic the structures of existence. So the structures of knowing will similarly be complex systems — inter-disciplinary teams exchanging information on what is known within the specialized spheres of each participant. This raises an intriguing philosophical question of the sense in which a group of people can be said to "understand" something. Must there always be an individual mind whose role is to be the overall understander — perhaps a convenor or chairperson who coordinates the separate understandings into a coherent whole, without necessarily knowing fully the details of any part. And what, anyway, is *understanding*: we will encounter later a related question of whether a computer program can deliver understanding by due processing of complex data.

Communication between participants in a collective scientific enterprise will be mirrored in complex knowledge structures, primarily now on the internet, reflecting the separate disciplines and how they interconnect. So the interrelated objects of a system may be material or informational. This leads us to a distinction between *natural systems* and *formal systems* [Rosen, 1985]. A natural system is a selected portion of the external world which we investigate through experiments and model through physico-chemical dynamics. A formal system is based on symbols, syntax and transformational rules of symbol manipulation. The *modeling relation* [Rosen, 1991] describes the process by which we establish congruence between the two systems; allowing us to study the formal system as a *model* of the natural system. Genomics has a peculiar status as a *natural information system*, where the dynamic modeling at the molecular level gives rise to rules taking the form of informational transformations.

4 DATA HANDLING

Studying complex systems one is forced into practical forms of reductionism. Our brains have evolved to cope with a relatively small number of pieces of information, dealing with processes taking place in the time scales of everyday events and for which linear/proportional relations may apply. Many nonlinearities, very fast or slow dynamics, as well as delayed responses are beyond our intuitive common sense; they surprise us.

Biological cells offer all of these sources of difficulty. However, there are now technologies available to measure the expression levels of thousands of genes simultaneously. The resultant gene expression data is initially stored in the form of a data matrix; the number of rows n, usually denoting genes, takes values up to 30,000, the number of columns m would denote either a few time points at which measurements were taken or a couple of different conditions. How does one recognize pattern in thousands of gene expression profiles? Plotting the raw data simultaneously and unordered would fill any computer screen or sheet of paper with thousands of lines, which almost certainly would not allow a detection of pattern by eyesight. Any analysis of high-dimensional data therefore requires a dramatic reduction/projection into lower dimensions allowing visualization and pre-grouping, in order that its structure may be grasped, in some intuitive sense. This is a necessary prerequisite for forming hypotheses leading to models. While 3D visualizations of data are possible on computer screens, this approach is largely for illustrative purposes. In practice, an analysis of experimental data is almost certainly conducted through plots in the plane fitting a sheet of paper, a computer screen or white/black-board.

A cynic might say that studying high-dimensional multivariate data does not only imply a form of reductionism into the two dimensions of computer screens and standard-sized sheets of paper, but scientists are also forced to reduce their understanding of complex systems into a 20 minute presentation and into about eight pages of a publication. Regardless of how complex a system is, the communication of research results will, in practice, almost always be limited to a short oral or written presentation that fits the constraints set by journals, conferences but also not exceeding attention span of an interested audience. So called "holistic approaches", so desirable in the life sciences, are at present wishful thinking.

We are led to enquire what tools we possess to aid this reduction. The usual suspects are mathematical modeling, statistics and computer programming. The first of these forms a proper basis for choosing or developing techniques from the other two, which is why in systems biology mathematical modeling has become an essential component: mathematical modeling is the refinement of common sense into the realm of complex systems. Most particularly for the present discussion, it possesses the appropriate language for describing and analyzing high dimensions.

The statistics we use to reduce the raw data to a comprehendible form require an underpinning mathematical analysis of their purpose. For example, we need to take care when interpreting data via the familiar two-dimensional reduction afforded by forming covariances between pairs of variables, since part of the essence of complexity is the interpenetration of many variables acting together. So some essential information will be lost by the very act of proceeding in this way. We are probably obliged to employ such methods, whenever more appropriate techniques are lacking, but we should remain aware of the desirability of their justification in any new circumstances.

Once suitable methods for reduction have been settled on, we have the facilities offered by computer programming to deliver the resultant reduction. Once again, though, it might be dangerous to merely lift tools which have proved efficacious down from the shelf without at least attempting to understand their underpinning justification. The availability of high-throughput and whole-genome technologies, which generate gene expression profiles for thousands of genes, has tempted some researchers to speak of a "holistic" perspective on cellular function. Douglas Kell [2003] contrasts two strategies for understanding cell-biological systems with "omics-data": "The reductionist view would have it that if we can break the system into its component parts and understand them and their interactions in vitro, then we can reconstruct the system physically or intellectually. This might be seen as a 'bottom-up' approach. The holistic (or top-down) approach takes the opposite view, that the complexities and interactions in the intact system mean that we must study the system as a whole." During the heyday of genomics and bioinformatics it became common practice to collect data without preconceptions, doing "some" (any) experiment and then search the data for pattern in what is called "data-driven discovery". It then seemed that no theory, hypothesis or model was required for these "hypothesis-free fishing expeditions".

With all respect to the area of "artificial intelligence", the reason why high-throughput genomics and bioinformatics have been able to do without a more focused approach is that in the early days of genomics it has been rather easy to discover "something". To continue the metaphor of a fishing expedition, fishing in the North Sea or Atlantic during the 1960's did not require much of a hypothesis about how to find fish; putting out your net almost anywhere, would catch some fish. Nowadays, you need to plan your fishing expeditions more carefully, supported by a good understanding of the fish's behavior. It is naive to believe that "any" experiment would do. A living system, observed at steady state or equilibrium, is either dead or does not reveal the information necessary to understand its behavior. We can only reconstruct/model the mechanisms or interactions that generate the observed behavior of a dynamic system if we systematically perturb/stimulate it and then observe its response. Even the simplest stress response or knock-out experiment implies a hypothesis — if only about the fact that the induced changes alter the behavior.

Some members of the computer science and bioinformatics community, providing boats and tools for these hypothesis-free fishing expeditions, argue that rather than beginning with a hypothesis, "artificial intelligence" approaches would generate hypotheses, extract them from data by letting "data speak for themselves". These approaches would then be considered inductive, inferring a general law or principle from the observation of instances. The philosopher Karl Popper famously argued that induction is a myth. John F. Allen [2001a; 2001b] responded to these genomic and bioinformatics fantasies, and this led to a series of articles documenting this interesting debate [Gillies, 2001; Kelley, Scott, 2001; Smalheiser, 2002]. Allen argues that knowledge cannot arise *de novo* from computer-assisted analysis of biological data. What he comments on is the proposition that analysis of data can enlarge human understanding in the absence of any hypothesis or preconceived idea. "It makes little sense to insist on collecting genomic and structural data before you, or someone else, has posited an underlying mechanism. Without having an underlying mechanism — in essence an explanatory, tentative hypothesis — you have no basis on which to decide which data to collect. Data do not, and cannot, 'speak for themselves'." [Allen, 2001c].

Studying complex systems one has to be able to adapt or refine the goal or hypothesis as you go along. As Allen [2001b] writes: "Computers are necessary to analyze large data sets, but they are not sufficient. [..] Creativity consists of a willingness to consider the relevance of observations that have no apparent connection with the problem as it is viewed conventionally." The computer program can only act via the given algorithm and cannot venture beyond the ominous "background knowledge" which has been fed into the program.

Apart from practical considerations we face methodological challenges. For instance, uncertainty can be seen as a consequence of complexity ("complex systems are difficult to understand") but apart from epistemological aspects of dealing with uncertainty we shall later also consider ontological aspects of randomness in nature.

At present, the questions asked in systems biology are largely determined by the technologies used to generate data, by the (model) organism chosen, by the choice of a particular cell type and cell line, antibodies available etc. Take for example the question of how the Ras/Raf/MEK/ERK (MAPK) signaling pathway works. While most researchers would readily agree that this is a reasonable and important question, there are several problems with this approach. To begin with this pathway is not a module of bounded autonomy as discussed below. The question should not be how this particular chosen pathway works but how cells grow, differentiate, proliferate or die as this is in fact the process relevant to diseases (cancer research being the primary motivation for studies on MAPK pathways). One should therefore first identify a question or hypothesis about the functioning/behavior of the cell, then identify a suitable model organism, cell type, cell line, to decide upon a subset of components, a pathway and only then identify the technologies adequate to generate the data required.

Assuming there are functional modules and levels of bounded autonomy in cellular systems, how do we best identify their boundaries and constituent components in experiments? Given a selected subsystem, how do we then unravel feedback mechanisms giving rise to the observed dynamical behavior and how do we integrate knowledge and models of subsystems to understand the interconnectivity of organizational levels and explain emergent behavior? Clearly, the study of complex systems is not just about the nature of things (an ontological problem) but this research also raises questions about the way in which we generate knowledge.

5 THE CELL AS A SYSTEM

Let us return to our two central questions of systems biology:

1. *How do the components within cells interact, so as to bring about the cell's structure and realize its functioning?*
2. *How do cells interact, so as to develop and maintain higher levels of structural and functional organization?*

Studying living systems one is bound to confess that the more we learn about them the less we are prepared to generalize. The complexity of cellular systems, the difficulties in studying them in experiments, has led to high levels of specialization within disciplines, hindering the generalization of results. Living systems appear so complex, so diversified that no general statement can safely be made about them. While systems biology has emerged from the need to put the pieces of the puzzle together, it does not offer a holistic salvation in the sense that universal laws ("physics style") can be derived. Throughout the essay we have emphasized the irreducible wholeness of living systems but also the impossibility of a truly holistic approach in which we can study/observe a cell as a whole.

An important aspect of a systems-theoretic approach is that objects and relations between objects have identical ontological status: Life is a relation between molecules/cells and not a property of any molecule or cell. Paraphrasing Henri

Poincaré[16] we might say that a cell is built up of molecules, as a house is with stones but a soup of molecules is no more a cell than a heap of stones is a house. Organisms, their cells, genes, and proteins are complex collections of interacting objects whose relationships and properties are largely determined by their function in the whole. In living systems everything exists only in relation to something else. The cell or any subsystem of it, together with the associated environment, has to be understood as an undividable whole. This almost obvious fact is constantly ignored by the reductionism that is forced upon us by the complexity of cells. To use the language of physics, the cell is a many-body system in which non-local interactions between the constituent molecules exert influence on the locally analyzed components. In other words, the inter-relationships of the parts (sub-wholes) within a system depend crucially on the state of the whole, in a way that is not expressible in terms of the properties of the parts alone. The irreducible wholeness of living systems suggests principle limitations to what we can know about them.[17] We are forced into reduced representations, necessarily leaving things out. The uncertainty arising from reduced or approximate representations could be captured with stochastic models but this would not solve the problem of how to distinguish between an intrinsic feature of the natural system and methodological considerations.

Living systems are dynamic systems, they are constantly changing, almost every part of an organism being exchanged throughout its lifetime. In systems theory this is reflected in the interpretation of causality as the principle of explanation of change: causal entailment is not considered to be a relationship between things (genes, proteins, etc.) but a relationship between changes of states of things. Not only do cells dynamically respond to immediate external changes and stimuli, they are also subject to evolution, not only at a time scale that covers generations of the organism but also in the range of hours, days and weeks. We can distinguish between two dynamic principles of key importance in studying cell function: a system's ability to maintain its current state against external perturbations (e.g. homeostasis) leading to some form of *robustness* and the system's *responsiveness* to environmental cues, to adapt its state or even modify its biophysical make-up. The basis for all forms of regulation, control, adaptation and coordination is the notion of *feedback*. Feedback loops provide the system with information about its current state and possible divergence from a desirable state/trajectory. Based on the current values of system or state variables a change is induced to move the system into a "desirable state" or follow an "intended trajectory". For example, in development stem cells should grow, proliferate and differentiate but this implicitly assumes the existence of an objective. In complex dynamic systems the change of state is influenced not only by inputs to the system but also by an overall goal or

[16] "A collection of facts is no more a science than a heap of stones is a house" Henri Poincaré (Science and Hypothesis, 1908) or, also attributed to Poincaré: "The aim of science is not things in themselves but the relations between things; outside these relations there is no reality knowable."

[17] See George Kampis [2003] for a critique of state-space based (differential equation) models as descriptions of living systems.

objective (i.e. the distance between a current and desirable/reference state): Living systems are *anticipatory* [Rosen, 1985]. The existence of feedback mechanisms also highlights the importance of systematic perturbation studies as only then we will be able to unravel the structure of dynamic networks from stimulus-response data. A system that is self-organized or robust to external perturbations does not reveal its internal functional organization in simple observations but requires an experimental manipulation.

6 SYSTEMS BIOLOGY OF THE CELL

Molecular and cell biology to this day has been preoccupied with the identification and molecular characterization of cellular components, leaving little time to conceptualize biological information and to develop "theories". The recent interest in systems biology is associated with the hope that it will be possible to manage the complexity of cellular systems, leading to postulated generic *principles* that govern those cellular processes that underlie the development and (mal)functioning of cells, cell populations, tissues, organs and organisms.

The (re-)emergence[18] of systems biology over recent years signals a shift of focus from the study of the *structural organization* of cells towards an understanding of the *functional organization* of cells. By structural organization we refer to the physical structure and material basis of cells, including macromolecules (e.g. DNA,[19] enzymes,[20] and receptors[21]), organelles[22] as well as the outer cell wall (and inner membrane in eukaryotes). The functional organization of the cell refers to processes that determine the cell's activity (its dynamic behavior). The word "function" refers to a role defined by the context of a system or process. For example, the role of stem cells can be the regeneration of tissue. This provides the context for *cell differentiation* (a specialization of stem cells, turning them into a specialized cell type). Cell differentiation in turn is the context for various networks in which proteins interact in order to realize this function (or an aspect of it). The most important *cell functions* studied in systems biology include cell growth, cell proliferation, cell differentiation and cell death (apoptosis).

Not only are we forced to select a subset of proteins, respectively a subsystem, even if we could quantify larger number of components, the analytical tools for the

---

[18]The need for a research field of systems biology was first formulated by Mesarovic [1968].

[19]Deoxyribonucleic acid (DNA) is a macromolecule that encodes the genetic information used in the development and functioning of all known living organisms (virus being a special case). The entirety of hereditary information of an organism is also referred to as the *genome*.

[20]Enzymes are proteins that catalyze (accelerate) biochemical reactions. The vast majority of processes in a biological cell require enzymes to facilitate the modification or transformation of molecules. Inhibitors are molecules that decrease enzyme activity; activators are molecules that increase activity.

[21]In cell signalling a common mechanism for information transfer is the binding of signalling molecules (ligands) to receptor proteins on the outer cell membrane. The binding leads to a biochemical modification which transfers the information through a series of intracellular processes into the nucleus where the information can lead to changes in gene expression.

[22]An organelle is a specialized subunit within the cell.

analysis of such large, nonlinear models are missing. Proteins are modified (e.g. activated), each of these states adding to the number of variables in a mathematical model. A system with 10 components can subsequently lead to 20 or more system variables. The theory of nonlinear dynamic systems, the methodologies and tools available to identify models (their structure and parameter values) from experimental data, to investigate their behavior analytically or through numerical simulations remains to this day limited. We are once more forced to simplify for practical considerations (e.g. through linearization). The reduction of complexity through abstraction and modeling does however not only serve practical purposes. Studying complex systems we seek simplifications to reduce complex processes to an essential aspect of their functional organization, to extract a principle that serves as an explanation. Studying complex systems we are seeking general principles underlying the observations we make in experiments. Mathematical modeling is then the *art* of making "appropriate" assumptions, balancing necessary reductions due to methodological and experimental limitations with abstractions serving explanatory purposes.

The construction of dynamical models is informed by experimental data. In an ideal situation, experimental time course datasets can be used to identify the structure of a network (and hence of the equation that form the model) and parameter values can be directly estimated from time series. At present, there is a lack of technologies that allow us to quantify temporal changes of gene activity and changes in protein concentrations with sufficient accuracy/reproducibility, for a sufficient number of time points and for a larger number of molecules (and their activation states). There are on the other hand technologies that can detect thousands of proteins simultaneously (e.g. 2D gels) or indicate the activity of genes for whole genomes (e.g. microarray or gene chips). Such "Omics" data, coming from high-throughput and whole genome technologies, have been analyzed in the area of bioinformatics using methods from multivariate statistics, "machine learning" or "data mining". Their qualitative character has, so far, prevented the identification of models from dynamical systems theory.

The study of complex systems is difficult, pushing state-of-the-art technologies to their limits and demanding new methodologies to interpret data through modeling. We can distinguish between two complementary general aims for modeling: reproducing complexity for computer simulations and reducing complexity in models that encode general principles. In the first case we try to establish a detailed replica computer representation of a complex system. Typical examples are large-scale mechanical/physical models of engineering systems (say airplanes). Computer simulations would then allow the study of the system's behavior, predicting the behavior of the system under unobserved conditions. For as long as the system is mechanical, subject to Newtonian physics the parameter values for the computer model can be derived from "first-principles" (considering mechanical properties of the constituent components). The second principle aim of modeling is to simplify, reduce complexity to some general principle through abstraction. For cell biological systems we cannot develop microscopic models in which molecules are treated as mass-points, instead one models changes in molecular concentrations in a macroscopic sense. Since parameter values cannot be derived from "first (physical) principles", one could estimate them from time course data. As discussed above, state-of-the art technologies cannot — at present — deliver suitable datasets, nor is system identification simple for nonlinear spatio-temporal systems. Even if these macroscopic models may be phenomenological, this does not mean that the structure of the equations is arbitrary as in black-box modeling. The structure of the mathematical model encodes in this case a hypothesized principle. We are going to focus our discussion on the second type of models, which cannot be derived from first principles. Whatever the goal of modeling, it will be important to distinguish between the complexity of the natural system under consideration and the complexity of the effort by which we gather information and gain knowledge about the complex system.

7 THE MULTILEVELEDNESS OF CELL-BIOLOGICAL SYSTEMS

The structural (physical/material) organization of cells is the outcome of an elaborate self-organizing process, involving gene expression, regulation, signalling and metabolism. This structural organization of the cell then serves as an environment for the cell's functional organization, leading to growth, differentiation, proliferation and cell death (apoptosis). The cell is itself a component of a larger system with higher levels of structural and functional organization. For the human organism these can be summarized as follows.

At the spatial level of the entire organism the human body grows, reproduces and dies in time scales that can be years. The human body is made up of organs, which help to maintain, renew, repair or regenerate the organism and adapt it to its environment. Organs realize their functions over hours and weeks. Organs are made up of cells, which go through the cell cycle, grow, divide, specialise and die. These processes take place over minutes and hours, while intracellular biochemical reactions take place in seconds.

Multileveledness is a key organizing principle in complex systems where the responsibility for proper functioning of an overall system is shared by the subsystems that constitute the different levels. A fundamental property that is determined by interlevel relations is that the levels have the latitude to focus on their 'allocated' tasks, and which implies that each level must posses a *domain of autonomy* [Mesarovic, *et al.*, 1970]. If two levels each possess a domain of autonomy it means that each level has some range of behavior which is autonomous in the sense that the two levels do not affect each other through changes in these ranges; changes within the domain of autonomy of one level is perceived as "background" by the other level and vice versa. The influence of one level is treated as a bona fide signal by the other level whenever the receiving level is outside of its domain of autonomy.

Making a distinction between the interaction (signaling) and the interdependence of levels is useful in this respect. Although the levels, belonging to the same system, are interdependent in many ways, they are non-interacting (non-signaling) within their respective domains of autonomy. Identification of the domains of autonomy of pathways is therefore a major challenge for systems biology. Domains of normal behavior are delineated by tolerances. The system can become pathological either when the a function on a level strays outside of the domain of autonomy or when a tolerance on a level changes due to internal or external influences. Bounded autonomy provides *cross-level harmonization* and illustrates that nature's design is not optimization of the behavior over time but rather optimization of a system's organized complexity. A major challenge for systems biology is to develop methodologies and experimental designs that allow the identification of functional modules, the separation of subsystems and levels of bounded autonomy.[23]

The very existence of different levels for experiment and modeling suggests an ontological basis for the complexity reduction which is necessary for our understanding. Our epistemological efforts arise not merely from arbitrary selection of features of a system. The fact that we observe coherence of the parts into higher level structures, of which we can speak with a suitably-tailored language, suggests the ubiquity of automatic stabilization. The procedures we adopt arise from our attention naturally being drawn to objectively real coherent structures which exhibit bounded autonomy.

The kinds of stability encountered in the major part of current systems theory is entirely inappropriate to handle such phenomena. The most elaborated results of this theory have been developed, mainly in engineering contexts, in the framework of linear systems. In such systems the possible forms of stable states are very limited and will not allow the kind of complexity reduction we are seeking, a dramatic decrease in the number of dynamic variables which need to be considered.

Our attention should therefore be directed towards *non-linear* systems theory. Recall that the simplest kind of dynamical system — a finite-dimensional, deterministic, memoryless system — is described mathematically by a collection of first-order differential or difference equations. The collection of dependent variables, say n in number, are quantities characterizing the system's state which develops through time. Our principal concern is with complex systems in which n is large, but for which we can regard as significant only a few functional combinations of these. This may be achievable in the following way.

Suppose each orbit of interest to have a simple attractor — an attractor for an orbit being its asymptotic limit as time tends to infinity. If the time-scale for the dynamical events is short when compared with that employed by an observer of the system, the state will appear to that observer to jump rapidly on to the attractor and stay there. We need only assume that all the usual initial conditions have orbits going to the same attractor and the observer will see an apparent self-organization of the system to a coherent behavior. If the attractor, considered as a subspace of the state space of dynamic variables has low dimension, we can regard the motion within the attractor as a simple description of the system's trajectory in terms of just a few aggregate variables which play the role of coordinates within

[23]See Mesarovic *et al.* [2004] for a discussion of these ideas in systems biology.

that subspace. Even allowing for perturbations away from the attractor, these variables can approximately describe the actual position of the system's state when it moves within a sufficiently close neighborhood of the attractor. One might then hope that the perturbations around the stable values, remaining small in magnitude, may be handled by suitable statistical techniques.

We initially stated the intention to avoid ontological questions about the role of randomness in living systems, instead focusing on epistemological questions, i.e. uncertainty arising from reduced and approximate descriptions. As long as our models capture a level of the functional organization of cells at which randomness does not matter, we are fine. However, models should after all be a representation of what the things are "in themselves", which means that we cannot always ignore the role of randomness. For nonlinear dynamical systems randomness is most potent at bifurcation points and in systems sensitive to initial conditions. How can we then distinguish between intrinsic (possibly purposeful) randomness and a signal that is formed from a temporal average of a molecular concentration? What this problem suggests is the importance of stating the context in which a model is valid.

8 DEALING WITH UNCERTAINTY: RANDOMNESS, STOCHASTICITY

Observing a cell-biological system, irregularities and the absence of an obvious pattern/trend in data induce uncertainty in the analysis of the system. The first question is then whether this *randomness* is an inherent, possibly purposeful aspect of the system or whether it is a consequence of limitations in observing the system (the choice of subsystem looked at, components that are ignored or limitations to measurement technologies)? In either case, one may consider a *stochastic model* to describe the system in terms of probabilities [Ullah, *et al.*, 2007; 2008].

Note that our discussion will be limited to the level of cells, where we investigate the function(ing) of cells in terms of changes in the abundance of molecules within cells and consequences this may have for populations of interrelated cells [Raj, *et al.*, 2008]. The discussion of randomness in physics, specifically statistical mechanics, may thus be avoided in our present context. While thermal and perhaps quantum fluctuations may in fact influence events at the cellular level and above, instead of modeling them in detail we may, without losing essential cellular and higher order modeling power, represent their consequences by irreducible stochasticities. The cell is here considered an open, non-equilibrium system, with a constant flux of material and information in and out of the cell. At the level of single molecules, the irregular motion of atoms and molecular bonds within the system may well be relevant but will here be referred to as effects of the *microscopic level*. This includes thermal fluctuations and Brownian motion. Looking at changes in the concentration of molecules, following a clear trend that may well be described in terms of differential equations, such models may be referred to as *macroscopic*.

Where necessary, a stochastic model can be formulated comprising both the deterministic laws and the fluctuations about them. Such models are sometimes referred to as *mesoscopic models* [van Kampen, 1992]. Considering a system of interacting mass points, fluctuations in non-equilibrium systems do not arise from a probability distribution of the initial micro-state, but are continuously generated by the equations of motion of the molecules. While mesoscopic stochastic models are attractive theoretical concepts, in a practical context where such a (nonlinear) model and its parameter values would have to be extracted from experimental data, we face various problems (which are in part a reason for the wide use of ordinary differential equations).

We can illustrate the notions of microscopic, mesoscopic and macroscopic in the context of cell biology by considering gene expression, the process by which information of the genome is first transcribed into RNA before being translated into proteins. These two stages involve two levels, the transcription of a gene being microscopic compared to fluctuations in the concentration of the protein for which the gene encodes the information. While for the initiation of transcription, say through the binding of transcription factors, a stochastic model may be appropriate; changes in the concentrations of the proteins involved in the function of a single (e.g. cell cycle) may on the other hand be described macroscopically by ordinary differential equations. Taken together, the whole model is mesoscopic.

In many situations random fluctuations are sufficiently small to be ignored, allowing macroscopic equations to predict the behavior of a system with great accuracy. Cells however are "open systems", where the environment may force them into a stationary non-equilibrium state in which the system's dynamics bifurcate, the direction taken depending on the specific fluctuations that occur. Note that therefore the "randomness" of the fluctuations (which we can only describe in terms of probabilities) influences the behavior of the system of macroscopic equations most critically at specific bifurcation points, while other areas of the state space may be perfectly well approximated by macroscopic equations. Intrinsic noise from thermal fluctuations or transcriptional control could determine how the system at the macroscopic level goes through a bifurcation. Looking at a population of genetically identical cells in a homogenous environment, this leads to variability of cell states that may well be exploited by the biological system [Rao, *et al.*, 2002; Kærn, *et al.*, 2005; Shahrezaei, *et al.*, 2008]. The obvious context in which randomness has a function is generating diversity in evolution.

Looking at a single gene in a single cell, the initiation of transcription at its promoter site is driven by the association and dissociation of a very small number of molecules. This very low copy number of molecules has two consequences: the time of reaction events can only be described in terms of probabilities and changes in the number of molecules are discrete, with no obvious trend that could be approximated with a differential equation (see [Paulsson, 2005] for a review). The expression of a gene does however serve a function; say during the cell cycle, growth, differentiation or apoptosis of the cell. For example, in response to external stimuli, the cell may produce large quantities of a protein. This response, measured as an apparently smooth/monotonic change in concentration, appropriately described by differential equations. Small fluctuations around an obvious trend/mean are thus ignored. At this level we are aiming at a description of a pathway acting as a switch, filter, oscillator, amplifier, studying the network's behavior in terms of its robustness, responsiveness, sensitivity of the model to changes in parameters, transitions between steady states and bifurcations. A usual assumption in such rate equation models is that parameters (rate coefficients) are constants. Since these parameters are implicitly linked to environmental variables, such as temperature, pH level or water balance, fluctuations in these are considered negligible. The art of modeling is then to decide in the given context which modeling approach or combination thereof is most appropriate. Even if ordinary differential equations are chosen, noise can influence the onset of oscillations. An example, serving as a toy model for this, is the "Brusselator" [Blomberg, 2006]. Here one observes damped oscillations around the stationary point before the oscillation bifurcation occurs. Noise afflicts these damped oscillations, and this gives information about the bifurcation before it appears, in the region of a stable stationary point. Thus, noise provides information and details about the type of bifurcation that are not as clear in the basic differential equations.

As pointed out in previous sections, in experiments one can only study a limited number of components and generate data for them. The unavoidable conceptual closure in modeling and the neglect of system variables, will inevitably lead to uncertainty in the analysis of a complex system, providing an epistemological motivation for stochastic models.

David Bohm [1957] argued for the possibility that there might be an ever-recurring dialectic between causality and chance - or stochasticity and determinism in nature. If there could be an infinite number of levels of existence of matter then for each level which manifested itself stochastically there could be a level below to which that could be deterministically reduced: but, conversely, each deterministic level could reflect some average behaviour of a complex stochastic level below.

9 WHAT CAN WE KNOW ABOUT LIVING SYSTEMS?

As the complexity of a system increases, our ability to make precise and yet significant statements about its behaviour diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost exclusive characteristics.[24] Our understanding of complex systems arises from reducing one type of reality into another.[25] A complex system is by definition too complicated to be comprehended by just using everyday common sense. Studying complex systems through mathematical modelling is therefore to seek an understanding through abstraction. In other words, studying complex systems we put a high value on simplicity. The problem is that abstraction itself can be complicated to start with and thus abstraction is often not perceived as what it really is: simplification.

[24]The statement has been attributed to Lotfi Zadeh.
[25]The statement is attributed to Claude Levi-Strauss.

Dealing with complexity by reducing it in modelling suggests a loss of predictability. A model reduces a complex biological process to an essential aspect of its behavior, removing a general principle by which a cell functions from its experimental context of a particular culture, cell line or organism. All models are wrong, some are useful.[26] Abstraction lives from the fact that not everything that is there in a natural system needs to be modelled. To reduce or simplify a complex system into a tractable form requires however an understanding of the natural system in question. Simplicity appears thus to follow understanding; to understand a process we need to know it well. On the other hand, an understanding of a complex system requires simplification; we are dealing with an iterative, exploratory and creative process. Systems biology is indeed the *art* of making appropriate assumptions.

Systems biology signals a shift of focus away from molecular characterization towards an understanding of functional activity; away from studying the function of genes and proteins towards an understanding of cell function, supporting inferences about phenomena at the physiological level of cell populations, tissue, whole organs and whole organisms. A skeptic might argue that this is about the same as trying to predict the world economy from observations I make at my local superstore. While this endeavor seems impossible due to the complexity of cells, encouragement comes from the likes of Max Weber: "All historical experience confirms that men might not achieve the possible if they had not, time and time again, reached out for the impossible."; Mike Mesarovic: "It is less frustrating not to catch a big fish than it is not to catch a small fish - we might as well ask the big questions." and Richard Feynman: "We do not know where we are 'stupid' until we 'stick our neck out,' and so the whole idea is to put our neck out.".

If something seems impossible we improve our chances of success by trying it. In the meantime the interdisciplinary endeavor systems biology would benefit from the involvement of philosophers of science, discussing the process by which we model complex systems. First steps in this direction have been made [Fox-Keller, 2002], Boogerd et al.. 2007]. In an essay for the journal *Nature* Fox-Keller [2007] discusses the differences between physics and biology and asks whether biology does have physics-style laws that are universally applicable? When limits to the generality of findings are found in biology, this is usually not considered a problem and simply sets the context for the findings. Evelyn Fox-Keller asks whether exceptions to presumed laws are just a reminder of the complexity of biological systems or whether biologists should adopt a different attitude and systematically search for all-encompassing laws. She concludes: "Even though we cannot expect to find any laws governing the search for generalities in biology, some rough, pragmatic guidelines could be very useful indeed." System biologist may already benefit from reading the first pages of Popper [1959], where he quotes Novalis: "Hypotheses are nets: only he who casts will catch."

[26]The statement is attributed to George E.P. Box.

ACKNOWLEDGEMENTS

The authors would like to thank Cliff Hooker and Ludger Jansen for corrections and suggestions in the writing of this text. Discussions with Mike Mesarovic and Jannie Hofmeyr informed various parts of the essay. The result of our efforts is unfortunately little more than a documentation of our state of confusion about the consequences of cell biological complexity. We therefore take full responsibility for the remaining misinterpretations and misguided conclusions. O.Wolkenhauer acknowledges support from the ExCell project at the Center for Logic, Philosophy and History of Science, University of Rostock, funded by the regional government of Mecklenburg-Vorpommern, Ministry for Education, Science and Culture (Project Nr. 62327030). and support from the Stellenbosch Institute for Advanced Study (STIAS) in South Africa, providing funding and an exceptional environment to study the complexity of biological systems.

BIBLIOGRAPHY

[Alberts *et al.*, 2008] B. Alberts, A. Johnson, P. Walter, J. Lewis, M. Raff, and K. Roberts. *Molecular Biology of the Cell*, 5^{th} revised edn,. Taylor & Francis, London, 2008.

[Allen, 2001a] J. F. Allen. Bioinformatics and discovery: Induction beckons again. *BioEssays*, 23(1), 104–107, 2001.

[Allen, 2001b] J. F. Allen. In silico veritas. Data-mining and automated discovery: The truth is in there. *EMBO reports* 2(7), 542–544, 2001.

[Allen, 2001c] J. F. Allen. Hypothesis, induction and background knowledge. Data do not speak for themselves. Replies to Donald A. Gillies, Lawrence A. Kelley and Michael Scott. *BioEssays* 23(9), 862–862, 2001.

[Bertalanffy, 1969] L. von Bertalanffy. *General Systems Theory*. George Braziller, New York, 1969.

[Blomberg, 2006] C. Blomberg. Fluctuations for good and bad: The role of noise in living systems. *Physics of Life Reviews* 3, 133-161, 2006.

[Bohm, 1957] D. Bohm. *Causality and Chance in Modern Physics*, Routledge and Kegan Paul Ltd., London, 1957.

[Boogerd *et al.*, 2007] F. C. Boogerd, F. J. Bruggeman, J.-H. S. Hofmeyr, and H. V. Westerhoff, eds. *Systems Biology: Philosophical Foundations*. Elsevier, Amsterdam, 2007.

[Cassirer, 1981] E. Cassirer. *Kant's Life and Thought*. Yale University Press, New Haven, 1981.

[Cornish-Bowden, 2004] A. Cornish-Bowden. *Fundamentals of Enzyme Kinetics*, 3^{rd} revised edn. Portland Press, London, 2004.

[Ehresmann *et al.*, 2007] A. C. Ehresmann , and J.-P. Vanbremeersch. Memory Evolutive Systems: Hierarchy, Emergence, Cognition. *Studies in Multidisciplinarity* 4, Elsevier, Amsterdam, 2007.

[Fell, 1997] D. Fell. *Understanding the Control of Metabolism*. Portland Press, London, 1997.

[Fox Keller, 2002] E. Fox Keller. *Making Sense of Life: Explaining Biological Development with Models, Metaphors, and Machines*. Harvard University Press, Harvard, 2002.

[Fox Keller, 2007] E. Fox Keller. A clash of two cultures. *Nature* 445(8), 603, 2007.

[Gillies, 2001] D. A. Gillies. Popper and computer induction. *BioEssays*,23(9), 859–860, 2001.

[Kærn, *et al.*, 2005] M. Kærn, T. C. Elston, W. J. Blake, and J. J. Collins. Stochasticity in Gene Expression: From Theories to Phenotypes. *Nature Reviews Genetics* 6, 451-464, 2005.

[Kalman *et al.*, 1969] R. E. Kalman, P. L. Falb, and M. A. Arbib. Topics in Mathematical Systems Theory. McGraw-Hill Book Company, New York, 1969.

[van Kampen, 1992,] N. van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland/Elsevier, Amsterdam, 1992.

[Kampis, 2003] G. Kampis. *Self-Modifying Systems in Biology and Cognitive Science*. Pergamon Press, New York, 2003.

[Kant, 1892] I. Kant. *Critique of Judgement*. Translated with Introduction and Notes by J.H. Bernard, 2^{nd} revised edition, Macmillan, London, 1914.
[Katok and Hassselblatt, 1995] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, Cambridge, 1995.
[Kell, 2003] D. Kell. Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *BioEssays* 26(1), 990–105, 2003.
[Kelly and Scott, 2001] L. A. Kelley and M. Scott. On John Allen's critique of induction. *BioEssays* 23(9), 860–861, 2001.
[Kholodenko, 2006] B. N. Kholodenko. Cell-signalling dynamics in time and space. *Nature Reviews Molecular Cell Biology* 7, 165-176, 2006.
[Klir, 1991] G. J. Klir. *Facets of Systems Theory*. Plenum Press, New York, 1991.
[Luisi, 2006] P. L. Luisi. *The emergence of Life: From chemical origins to synthetic biology*. Cambridge University Press, Cambridge, 2006.
[Maturana and Varela, 1987] H. R. Maturana and F. J. Varela. *The Tree of Knowledge: The Biological Roots of Human Understanding*. Shambhala Publications, New York, 1987.
[Mayr, 2004] E. Mayr. *What Makes Biology Unique?* Cambridge University Press, Cambridge, 2004.
[Mesarovic, 1968,] M. D. Mesarovic. Systems Theory and Biology—View of a theoretician. *System Theory and Biology*, 59–87, 1968.
[Mesarovic *et al.*, 1970] M. D. Mesarovic, M. Macko, and Y. Takahara. *Theory of Hierarchical, Multilevel, Systems*. Academic Press, New York, 1970.
[Mesarovic and Tkahara, 1975] M. D. Mesarovic and Y. Takahara. *General Systems Theory: Mathematical Foundations*. Academic Press, New York, 1975.
[Mesarovic *et al.*, 2004] M. D. Mesarovic, S. N. Sreenath, and J. D. Keene. Search for organising principles: understanding in systems biology. *IEE Systems Biology* 1(1), 19-27, 2004.
[Morgan, 2006] D. Morgan. *The Cell Cycle: Principles of Control*. Oxford University Press, Oxford, 2006.
[Muir, 1982] A. Muir. Holism and reductionism are compatible. In S. Rose (ed.), A*gainst Biological Determinism*. Alison and Busby, London, 122–135, 1982.
[Nicolis and Prigogine, 1989] G. Nicolis and I. Prigogine. *Exploring Complexity*. W.H. Freeman and Company, New York, 1989.
[Novák and Tyson, 2008] B. Novák and J. J. Tyson. Design principles of biochemical oscillators. *Nature Reviews Molecular Biology* 9, 981-991, 2008.
[Padulo and Arbib, 1974] L. Padulo and M. A. Arbib. *Systems Theory*. W.B. Saunders, Philadelphia, 1974.
[Paulsson, 2005] J. Paulsson. Models of stochastic gene expression. *Physics of Life Reviews* 2, 157-175, 2005.
[Popper, 1959] K. Popper. *The Logic of Scientific Discovery*. Routledge, London (first published in 1959), 1992.
[Raj and van Oudenaarden, 2008] A. Raj and A. van Oudenaarden. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell* 135(17), 216-226, 2008.
[Rao *et al.*, 2002] C. V. Rao, D. M. Wolf, and A. P. Arkin. Control, exploitation and tolerance of intracellular noise. *Nature* 420, 231-237, 2002.
[Rosen, 1985] R. Rosen. *Anticipatory Systems*. Pergamon Press, New York, 1985.
[Rosen, 1991] R. Rosen. *Life Itself*. Columbia University Press, New York, 1991.
[Rosen, 2000] R. Rosen. *Essays on Life Itself*. Columbia University Press, New York, 2000.
[Shahrezaei *et al.*, 2008] V. Shahrezaei, J. F. Ollivier, and P. S. Swain. Colored extrinsic fluctuations and stochastic gene expression. *Molecular Systems Biology* 4 (196), doi:10.1038/msb.2008.312008
[Smalheiser, 2002] N. R. Smalheiser. Informatics and hypothesis-driven research. *EMBO reports* 3(8), 702, 2002.
[Solé and Goodwin, 2000] R. Solé and B. Goodwin. *Signs of Life — How Complexity pervades Biology*. Basic Books, New York, 2000.
[Tyson *et al.*, 2003] J. J. Tyson, K C. Chen, and B. Novák. Sniffers, buzzers, toggles, and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology* 15, 221-231, 2003.
[Ullah and Wolkenhauer, 2007] M. Ullah and O. Wolkenhauer. Family tree of Markov models in systems biology. *IET Systems Biology* 1(4), 247-254, 2007.

[Ullah and Wolkenhauer, 2008] M. Ullah, O. Wolkenhauer. Investigating the two-moment characterization of subcellular biochemical networks. *Journal of Theoretical Biology* doi:10.1016/j.jtbi.2009.05.022 (in press).

[Weaver, 1948] W. Weaver. Science and Complexity. *American Scientist* 36 (4) 536-44, 1948.

[Wiggins, 2003] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, 2^{nd} edn. Springer, New York, 2003.

[Wolkenhauer and Mesarovic, 2005] O. Wolkenhauer and M. Mesarovic. Feedback Dynamics and Cell Function: Why Systems Biology is called Systems Biology. *Molecular BioSystems* 1(1), 14-16, 2005.

[Wolkenhauer, 2007] O. Wolkenhauer. Defining Systems Biology: An Engineering Perspective. *IET Systems Biology* 1(4), 204-206, 2007.

[Wolkenhauer and Hofmeyr, 2007] O. Wolkenhauer and J.-H. S. Hofmeyr. An Abstract Cell Model that describes the Self-Organization of Cell Function in Living Systems. *Journal of Theoretical Biology* 246, 461-476, 2007.

Part III

Ecology

CONSTRUCTING POST-CLASSICAL ECOSYSTEMS ECOLOGY

Gao Yin and William Herfel

Ecosystems ecology, especially under the influence of the energetic ecosystems models of the Odum brothers, has been guided by mechanistic and equilibrium assumptions about ecosystems dynamics. The assumptions are seldom articulated, except when they are being critiqued. Nevertheless the assumptions are implicit in the models proposed in both theoretical and empirical work as well as in the definition of research problems. We term such a system of assumptions a 'classical dynamic perspective'. Everyone makes fundamental assumptions about the world.[1] A subset of these assumptions constitutes the dynamic perspective: the point from which one explains, models and predicts change in the natural world. In order to perform research, both theoretical and empirical, scientists must possess a fairly sophisticated dynamic perspective. In particular, it stipulates what to expect by default concerning change in the natural world, and what is therefore necessary to explain. In this way, what we term 'dynamic perspective' is similar to what Stephen Toulmin [1961] calls "ideals of natural order".[2]

We have already mentioned two such aspects of classical ecosystems ecology: the expectation by default of equilibrium, with its corresponding requirement of explanation for departures from equilibrium;[3] and the expectation that patterns of energy flow through an ecosystem will be mechanistic, with its corresponding requirement of explanation for novel behaviour.[4] We do not attempt a systematic articulation of the dynamic perspective of classical ecology in this article, although some aspects of the perspective will be discussed. Instead, we focus on articulating elements of a new dynamic perspective that is emerging from the application of self-organising and complex adaptive systems or CAS in ecology. We see the new dynamic perspective as the largest contribution that complex systems dynamics is currently making to ecology today. The article will proceed in three parts: in the first we will discuss some of the various complex dynamics traditions that form the intellectual backdrop for the new dynamic perspective in ecology; in the second we will articulate in detail the key elements of the dynamic perspective of

[1] Wittgenstein [1969] refers to such assumptions as "hinge propositions".

[2] Hal Brown [1977, p. 95] attractively cashes out such assumptions as "commitments".

[3] Solé and Bascompte [2006, p. 2] discuss equilibrium assumptions in terms of the Hardy-Weinberg Law.

[4] Ulanowicz [1997] identifies and critiques the mechanistic assumptions in the work within the Odumian framework.

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

post-classical ecosystem ecology; in the third we will discuss some important philosophic implications of applying complex systems dynamics to this new paradigm in ecology.

1 TRADITIONS IN COMPLEX SYSTEMS DYNAMICS

The scientific study of complex systems has rich and diverse origins. We do not hope to recount this history here, but we do want to call attention to the various traditions[5] in complex systems research and show how these traditions have come to influence contemporary research in ecology.

1.1 Far-from-equilibrium thermodynamics

Of particular relevance to ecology is the far-from-equilibrium thermodynamic tradition with its origins in the Brussels School lead by Ilya Prigogine. Strictly speaking, at least until 1931, classical thermodynamics only applied to either static or reversible systems. Calculations of entropy could only be done at the equilibrium endpoints of a process. Hence the term thermo*dynamics* is a bit of a misnomer for a discipline that perhaps more accurately ought to be termed thermo*statics.* A contemporary thermodynamics textbook puts this point vividly:

> Although we have seen that it is not *energy* but *availability* that is desirable, we will use the term energy degradation to imply a degrading of ability [of energy to do work]... We lose availability in a *real* process, and in a reversible process it merely remains the same. Loss of availability is subtle and difficult to detect, but predicting irreversibilities, degradation of energy, or increases in entropy in a general process is still more difficult. One of the important areas of research in thermodynamics is the development of more *general* and accurate equations or *theorems* that can allow these predictions of irreversibility—it is a crucial area since most of our present tools of thermodynamics can accurately be used only for reversible and static systems, which *indeed are not the real world* . [Rolle, 1999, pp. 260-1, *our emphasis*]

Although they are not noted in Rolle's textbook, there have been two important contributions to irreversible thermodynamics [Prigogine, 1980, p. 86]. The first is the theorem published in 1931 of Lars Onsager that allows the treatment of linear irreversible transformations as a series of small departures from equilibrium. The other is Prigogine's theorem of minimum entropy production. Prigogine [1980, p. 88] characterises his law this way:

[5]We make no attempt to be comprehensive; the field is huge, but we hope our account is sufficient to give the reader an idea of the topography of the field. We apologise in advance to anyone whose favourite development in the area we have left out.

> The theorem of minimum entropy production expresses a kind of "inertial" property of nonequilibrium systems. When given boundary conditions prevent the system from reaching thermodynamic equilibrium (i.e. zero entropy production) the system settles down in a state of "least dissipation."

It is important to note that both Onsager's and Prigogine's results only apply to irreversible systems near equilibrium (that is, in the linear regime). Prigogine [1980, p. 88] states,

> ...for many years great efforts were made to extend this theorem to systems farther from equilibrium. It came as a great surprise when it was shown that in systems far from equilibrium the thermodynamic behaviour could be quite different—in fact even directly opposite that predicted by the theorem of minimum entropy production.

Hence, far-from-equilibrium thermodynamics was born. The work of Prigogine and his Brussels School has involved constructing nonlinear *dynamic* models of such far-from-equilibrium systems. Drawing on work of researchers from a wide range of disciplines the dynamics of systems, including Bénard convection, the Belousov-Zhabotinsky reaction and cellular slime mould (*D. discoideum*) aggregation, have been explored.[6] This work has significant relevance to ecosystem ecology, and we will return to some of the fundamental issues raised by it below, but first we need to discuss other traditions in complex systems dynamics that have influenced ecology.

1.2 Chaos Theory

Chaos theory has had an enormous impact on all sciences and on popular culture as well. Thus it is not surprising that ecology has been influenced by its results. The study of deterministic chaos dates back at least to the "homoclinic tangles"[7] Poincare discovered in his solutions to the three-body problem. However it was not until the 1970's that scientists began to realise the significance of deterministic chaos with the advent of digital computers powerful enough to simulate the time evolution of low dimensional systems of nonlinear differential equations. An important early result of such simulations, the significance of which was only recognised a decade later, was a paper by meteorologist, Edward Lorenz [1963]. Lorenz's paper documents results of his simple (three ordinary differential equations) but nonlinear model of atmospheric flow. The results of Lorenz's computer simulation of the equations for critical parameter values turn out to be an articulation of the two main features of chaotic systems. The first is *aperiodicity*. In the critical regime, a trajectory generated by his model settles into an

[6]Nicolis [1989] provides a nice readable overview of this work.

[7]The term "homoclinic tangles" refers to complex trajectories that remain aperiodic for long time frames. Poincaré discovered these mathematical objects whilst attempting a solution to the three-body problem. Today these trajectories are known as "strange attractors".

attractor, that is, the trajectory remains confined to a region of state space, but it does not repeat. Such attractors have been termed *strange attractors.* In the same regime the trajectories generated also exhibit *initial conditions sensitivity*, that is, nearby trajectories diverge exponentially. Another way of putting this is that when repeatedly running the simulation with different, but arbitrarily close, starting points the resulting trajectories quickly become no more similar than two runs with starting points chosen at random from anywhere in the space. Initial conditions sensitivity has a profound implication for the prediction of chaotic systems: in order to accurately predict even the medium-term behaviour of such a system requires measuring the initial conditions with infinite accuracy. Thus, even though chaotic systems are governed by deterministic dynamic laws, predicting their medium-term trajectories is in principle impossible.[8] Both features only occur for some (critical) parameter values; however their existence is independent of initial conditions.[9] Since Lorenz's pioneering work, physicists have examined chaotic models for a wide-range of systems, most notably for turbulent fluid flow.

However, a paper published in 1976 has direct relevance to ecology. In that year, physicist turned mathematical ecologist, Robert May, published his study of the iterative behaviour of the logistic map in *Nature.* The logistic map[10] is a very simple discrete nonlinear mapping of a population (expressed as a variable, x, ranging from 0 to 1) onto itself. Despite its simplicity it exhibits a wide range of behaviour when parameter values, λ, are varied. For low values of λ, repeated iteration of all initial conditions >0 result in the map reaching a particular point value. As λ is increased eventually a cycle of period two emerges. Further increase in λ results in a *period doubling* cascade, whereby 4, 8, 16... cycles emerge. Eventually, for λ above a certain value aperiodic behaviour emerges. Mitchell Feigenbaum [1978] proved that such periodic cascades exhibit a universal structure for a wide range of mappings. The implication for mathematical modelling is profound. Even the simplest nonlinear mappings are capable of a range of behaviour, including behaviour as complex as chaos. Equally profound is the implication for population ecology: even a very simple model of population dynamics can yield chaos. In the early 1920's Lotka and Volterra (independently) showed that (even without

[8]Although the implications for prediction of chaotic systems have obvious empirical consequences, it would be a mistake to consider them merely empirical issues. Initial conditions sensitivity is a mathematical feature of the model. Scientists discovered this feature by comparing runs of computer simulations, not by comparing models to reality. (In fact, if the system is initial conditions sensitive this latter would be impossible over the long term.) Furthermore, successful prediction does not require only measuring initial conditions with infinite accuracy, but also *specifying* them with infinite precision. This is a mathematical impossibility.

[9]In a wide range of systems aperiodicity and initial conditions sensitivity occur simultaneously. However, there exist strange attractors that do not exhibit initial conditions sensitivity; as well some initial conditions sensitive systems do not exhibit strange attractors [Grebogi *et al.*, 1984]. This raises the issue of how exactly to define chaos. However, we are most interested in the large class of chaotic systems exhibiting both features.

[10]The logistic map maps current population onto population at the next time step with the following mapping: $\lambda x_n(1 - x_n)\Diamond \quad x_{n+1}$, where λ is a control parameter varying from 0 to 4.

external perturbation) a simple predator-prey model could oscillate under certain conditions. May showed that an even simpler system could behave chaotically. Is it not possible that real ecological dynamics would be observed to be at least equally complex?

1.3 Complex adaptive systems

The final tradition we need to discuss, before we move on to the philosophical implications this work has for ecology, has much more recent origins. Perhaps "the Santa Fe School" would be sufficient to identify the tradition we have in mind, although people not associated with the Santa Fe Institute have contributed to this work. This work includes cellular automata, Boolean networks, multi-agent systems, classifier systems etc. Each of these models share a common structure: they consist of individual elements (cells, programs, agents) arranged in some sort of network. Each element is capable of being in two or more states. What state an element is in at any given time depends upon the state, at the previous time step, of one or more other elements in the network that constitute any particular element's 'neighbourhood'. There is a set of transition rules that determine the next state, given the current state of elements in any given element's neighbourhood. In such models initial conditions are assigned to the elements, and the system is run iteratively over time. What is remarkable about such systems is that very simple rules can yield a wide range of behaviour, including very complex dynamic pattern formations.

A very simple example is John Conway's "Game of Life". The Game of Life consists of a two dimensional array of cells. Each cell has two possible states, occupied and unoccupied. A cell's neighbourhood consists of the eight cells adjoining it. Consistent with the game's name the transition rules are expressed in terms of survival, death, and birth, as follows:

1. *Survival*: Occupied cells with two or three occupied cells in its neighbourhood at time, $t = n$, remain in the occupied state at $t = n + 1$.

2. *Death*: Occupied cells with four or more occupied cells in its neighbourhood at time, $t = n$, become unoccupied at $t = n + 1$ (overpopulation); occupied cells with one or less occupied cells in its neighbourhood at time, $t = n$, become unoccupied at $t = n + 1$ (isolation).

3. *Birth*: Unoccupied cells with exactly three occupied cells in its neighbourhood at time $t = n$, become occupied at $t = n + 1$.

This game can be played by hand on a checkerboard, on grid paper, etc. Apparently Conway originally played this game on the square tiled floor of his kitchen using dinner plates. Conway's game was introduced by Martin Gardner [1970] in a "Mathematical Games" article *Scientific American;* the original article includes a procedure for playing the game insuring the state transition rules are implemented correctly. For those who would find such a task tedious, the Game of

Life has been implemented in computer programs making it easy to observe the long-term emergent dynamics. What emerge from Conway's simple rules under certain conditions are patterns of enormous complexity. There are steady state structure (still lives), oscillators of several periods, and patterns that propagate in very complicated patterns for many (in some case over 1000) iterations. Verbal description really cannot do justice to the possibilities this game contains. Those who are unfamiliar with the behaviour of such cellular automata are encouraged to find a computer implementation of the Game of Life and observe for themselves.[11]

One of the most important issues raised by this work involves the issue of *self-organisation*. There is no universally accepted definition of the term, but let us tentatively define self-organisation as the emergence of novel behaviour in systems of elements obeying simple rules. Many would add that the emergent behaviour in question is not derivable from even complete knowledge of the dynamics of the individual elements constituting the self-organising system. Self-organisation is considered an emergent property in the sense that it is a property of the system rather than a property of the individuals that constitute the system. We do not intend to discuss in detail the theory of emergence in this article,[12] however the implications of self-organisation form a large part of the contribution of complex systems to the philosophy of ecology. We will elaborate this contribution below.

We have still not settled on a descriptive term for the systems (of which the work of the Santa Fe School provides numerous examples) we are discussing here. Much of the work on the models we have been focusing on are characterised in terms of Complex *Adaptive* Systems. We have no objection to the term; however, there is no consensus about what it is that makes a complex system particularly *adaptive*. One might be tempted to link adaptation and selection. If some process of selection figures in the outcome of a process (a system or model) then we can call it an adaptive system. Simon Levin [2002, p. 4] seems to take this line when he states,

> In general, we will define complex adaptive systems by three properties:
>
> 1. diversity and individuality of components,
> 2. localized interactions among those components, and
> 3. an autonomous process that uses the outcomes of those interactions to *select* a subset of those components for replication or enhancement.

However, property (3) is not without its problems. In his seminal article Richard Lewontin [1978] argues that evolution by natural selection is possible in systems

[11]There are now in the public domain numerous versions, including downloadable executable programs and *java applets*. A web search will turn up several pages. Some of our favourites are available at `http://www.bitstorm.org/gameoflife`, `http://www.ibiblio.org/lifepatterns` and `http://en.wikipedia.org/wiki/Conway's_Game_of_Life`.

[12]See Collier and Hooker [1999] and Hooker [2004] and references therein for a thorough discussion of the issue.

obeying three postulates, variation, heritability and selection (VHS). These are sufficient to bring about evolution, i.e. change in the composition of the system. VHS processes hence explain evolutionary change, but not evolutionary *improvement.* However, he asserts that Darwinian evolution theory also involves a notion of *adaptation*, that through the process of change organisms become increasingly fitter. Thus Darwinian Theory requires a postulate of adaptation. He spells this out in terms of engineering analysis: for a trait to be adaptive one must be able to specify how the trait enhances survival. In other words, traits perform certain functions, and it is improved functionality that is adaptive. If Lewontin is correct, the presence of a selection mechanism alone is not sufficient to call a system adaptive.

On the other hand, others, perhaps most vocally Stuart Kauffman [1993; 1995], argue that a mechanism of selection is not *necessary* for a system to be adaptive. He argues that a substantial amount of the adaptation observed in the biological world is the result of self-organisation. Thus, for Kauffman, there are two mechanisms that result in a system becoming increasingly adapted, natural selection and self-organisation.

One way to resolve this issue would be to argue that self-organisation plays a role, but only as one of the processes involved at the level of variation. Self-organising processes may well provide sources of variation, but in the long term selection will act on these (self-organised) variants, and the outcome will eventually be determined by natural selection. This solution seems panadaptationist. As Gould and Lewontin [1979] discuss, selection does not act on individual traits, but on the ensemble of traits we call organisms. There are several possibilities.[13] For example, neutral (or even deleterious) traits may tag along when found in combination with adaptive ones, or nonadaptive traits that have become entrenched in the process of selection may turn out to be adaptive in the future when the environment changes.[14] Gould and Lewontin do not discuss such processes in terms of self-organisation. However, Meinhardt [1995], in his book, *The Algorithmic Beauty of Sea Shells* shows how the myriad of patterns exhibited by dozens of genera of mollusc can be modelled as the result of the self-organisation of the chemical reaction and diffusion processes that form the shell. It is unlikely that such patterns confer adaptive advantage [Meinhardt, 1995, p. 5]. A multitude of shell patterns coexist, and Meinhardt [1995, p. 10ff.] suggests that this may sometimes be the result of chance fluctuations occurring during the life-cycle of the mollusc. But suppose the patterns observed *did* have relevance to selection. Would we then count the adaptive case as a CAS and the other not? This also

[13] Here we note just two of the possibilities suggested by Gould and Lewontin [1979].

[14] We are not sure of their biological currency, but examples are not difficult to imagine. We could express these in a series of "just-*not*-so stories" (or perhaps "not-quite-so stories"). Homo sapiens seems to be a walking example: E.g. bipedalism may well have conferred advantages (freeing up the hands so the opposable thumb could do its work?); however, many of us are familiar with the disadvantage bipedalism has for the sciatic nerve. Presumably, the advantages for the opposable thumb outweighed the disadvantages for the sciatic nerve. A similar story could be told for the big brain. Good for thinking, not so good for the birth mother's pelvis...

raises another issue. Are CAS's confined to living systems? On the surface at least, it seems that the mechanism involved in the formation of patterns found on seashells, is different in degree rather than kind from the process that forms the patterns found in inorganic systems.

This debate certainly has not been resolved in the scientific community, and we do not intend to offer a solution here. However, the self-organisation hypothesis seems at least plausible. There are certainly dynamically complex models that are proposed as CAS models and that do not employ a selection mechanism. On the other hand, there are CAS models (e.g. genetic algorithms) where the selection mechanism is explicit. Let us adopt the term Complex Adaptive Systems to name all such work, whilst acknowledging that there is currently no consensus on what constitutes being adaptive in the context of complex systems. We must leave open such issues as whether selection is necessary to properly call a system adaptive, what the relevance of the origin of the feature in question is to its status as an adaptation, and whether it is proper to call such structures as the B-Z reaction, slime mould aggregation and sand dunes Complex Adaptive Systems.

We end this section by identifying one more feature of complex dynamic models in general. The mathematical techniques for modelling complex systems fall into two distinct categories. Standard dynamic models employ (normally ordinary) differential equations to construct models of systems dynamics. These equations enable us to follow the "flow" of the system variables through state space. Such systems are continuous in space and time. The network models of complex adaptive systems instead employ transition rules characterising the interaction between individual elements. These systems are discrete in space and time. The differences in the mathematical properties of these types of system can cause subtle and difficult issues. For instance, technically, finite discrete deterministic systems cannot be aperiodic. Eventually the system runs out of states and once a state reappears the entire sequence repeats. Nevertheless, these systems are often characterised as "chaotic". Here chaos is defined as an *extremely high-order* periodic attractor. If the period of oscillation is greater than half the total number of points available in the space defined by the dimensions of the system, that is if on average more than half of the points in the space are visited by any given trajectory, the system for practical purposes is considered periodic. On the other hand, generally, the nonlinear differential equations employed in complex dynamic models do not yield closed-form solutions. Hence modellers do not attempt to solve the equations; instead they simulate the dynamics on a digital computer. A range of tested approximation techniques are available for such simulation. However, in complex systems, especially when capable of chaos, we have to be very careful with approximation techniques. Approximation is problematic for systems that require absolute precision to predict. Different techniques can yield very different results.[15] It would take us too far afield to explore how these issues are addressed. However, we need to be aware of these issues.

---

[15]Stewart [1989] provides a summary of this issue.

2 THE DYNAMIC PERSPECTIVE OF POST-CLASSICAL ECOLOGY

A novel dynamic perspective is emerging in ecology as models from complex systems dynamics are employed in ecology. In particular, models embracing network dynamics, self-organisation and complex adaptation, which we will explore in this part of the article, provide new insight at a fundamental level into the workings of ecosystems. As this perspective is adopted a truly organic approach to ecosystems ecology comes into focus with the *self-organising dissipative structure* replacing the *heat engine* as its central metaphor.

2.1 Network-dynamics: self-organising flux networks in ecosystems

The energetic framework has been dominant in ecosystem ecology since its development and articulation by the Odum brothers and their followers during 1950–70s [Golley, 1993]. Within this framework, ecosystems were construed as energetic systems, with biological properties and interactions "black boxed". Instead of modelling individual organisms, such work focussed on trophic groups linked by a food chain or web which was expected to behave mechanically [Golley, 1993, p. 80]. The mechanistic framework characteristic of much of the science of its day offers the machine or heat engine as it central metaphor. However it has come under increasing attack recently because of its limited success in providing conceptual tools that are necessary when dealing with the "unruly complexity" [Taylor, 2005] of ecological systems [Peters, 1991; Sagoff, 1997; Ulanowicz, 1997].

Ecologist Robert Ulanowicz [1997, p. 3] criticises the mechanistic framework, arguing that it cannot support adequate modelling and analysis of the self-organising and self-regulating organisations emergent in ecosystems. He develops a phenomenological description of the structure, i.e. the formal pathway network of the energy and material flow of ecosystems, based on far-from-equilibrium thermodynamics and information theory. He argues that such a phenomenological model can offer a more adequate framework to model and analyse the self-organising order emerging from ecosystems:

> The flows make possible the higher level behaviours, which in turn help to order and coordinate the flows. So reflexive is this couple that the description of one member lies implicit in the description of the other element. It is in this reflexive sense that a key postulate in the development of the current thesis should be understood; to thermodynamically describe an ecosystem, it is sufficient to quantify the underlying networks of matter and energy flows. A more general form of the postulate would read: *the networks of flows of energy and material provide a sufficient description of far from equilibrium systems* . [Ulanowicz, 1986, p. 30, e*mphasis* in original]

In doing this, Ulanowicz makes a deliberate effort to distinguish two states: thermodynamic equilibrium and a dynamical steady state in order to establish the fact that ecosystems, even when they have an unchanging balance that would conventionally be regarded as an equilibrium state, are not at a thermodynamic equilibrium but instead in a dynamic steady state with constant energy and material throughput. According to Ulanowicz [1997, p. 26], "only if no dissipation is occurring within the system—that is, if no entropy is being created—can it be regarded as at equilibrium." In this sense, Ulanowicz [1997, p. 26] argues that "no living systems are among these [thermodynamic equilibrium systems]." Ecosystems, as organised ensembles of living systems, certainly are not at thermodynamical equilibrium. Instead, they are far from equilibrium systems with a dynamically stable flux network. Furthermore, these flux networks are emergent self-organising dissipative structures [Ulanowicz, 1986, p. 57].

To illustrate how such networks emerge in real ecosystems Ulanowicz [1997, p. 43] analyses a typical autocatalytic process. He reports that a variety of species of aquatic vascular plants belonging to the genus *Utricularia,* or the bladderwort family, are found in freshwater lakes, especially in subtropical, nutrient-poor lakes or wetlands. The plant does not possess feeding roots that absorb nutrients from their surroundings directly. They get nutrients by trapping small motile animals collectively known as zooplankton. They are able to do this because there is always a film of bacteria, diatoms, a blue-green algae that collectively is known as *periphyton* growing on the leaves of the bladderwort, and this *periphyton* film attracts zooplankton to feed on it. There is also evidence that some species of bladderwort secrete mucous polysaccharides to bind algae to the leaf surface attracting bacteria [Ulanowicz, 1997, p. 43). While zooplankton feed on the *periphyton* film on bladderwort leaves, occasionally they trigger the hair of the plant and get stuck in the utricle, their decomposition releasing nutrient that is then absorbed by the bladderwort.

These three types of organisms make up a stable energy and material flow pathway network sustained by their activities; that is, bladderwort leaves provide a substrate surface for *periphyton* species to grow, the *periphyton* film then attract zooplankton to feed on them creating a potential nutrient source for bladderwort. This dynamic coupling creates surplus polysaccharide resources supplying enough additional *periphyton* such that there are more additional zooplankton sustained than are ingested; thus the numbers of all three species increase over what they would have been without this pathway. Each species provides the necessary conditions for the other two species participating in the close-looped activities sequence. Ulanowicz [1997, p. 41-6] then generalizes this kind of activity sequence cycle into a self-organising process, i.e., the autocatalytic processes that generate and sustain the energy and material flux network.

He further argues that in the development and growth of an ecosystem, the flux network will increase its organisation in terms of energy and material (EM) throughput. He calls this process 'ascendency' in contrast to the activities of disorganised components, which he labels 'overhead'. In Ulanowicz's [1997, p. 73] model, the ascendency is expressed by a variable called average mutual information (AMI), which "measures the average amount of constraints exerted upon an arbitrary quantum of currency in passing from any one compartment to the next."[16] Ascendancy is then defined as the product of AMI and the total system throughput [Ulanowicz, 1997, pp. 73-4]. By introducing these two concepts, Ulanowicz is able to quantitatively measure ecosystem flux networks as well as changes in such networks, i.e. the organisational changes of ecosystems, which is a major achievement.

A flux network is a complicated dissipative system for two reasons: 1. the structure is an emergent property that cannot be reduced to the sum of its component species; 2. the structure requires constant energy and material throughput to maintain its existence. Another important point made by Ulanowicz [1986, p. 54] about the nature of flux networks is that they can be regarded as relatively independent of their constituents. Ulanowicz [1997, p. 56] argues, "We can even conceive of a totally self-consistent and coherent of body of phenomenological observation that explicitly mentions only agencies at the focal level." In other words, his conclusion is that when constructing a dynamic ecosystem model, it is sufficient to exclusively focus on the organisation of the focal level, i.e., the flux network itself without taking into account the biological features of its constituent organisms. Ulanowicz [1997, p. 56] illustrates his argument with an analogy to thermodynamics:

> There exists a school of thermodynamicists, for example, that insists upon the sufficiency of macroscopic narration (which is not intended to constitute full explanation). As a student in chemical engineering science, I was made to toe this party line. If, in response to any question on thermodynamics, a student should utter or write the words "atom" or "molecule," the answer was summarily judged incorrect.
>
> It becomes permissible, therefore to concentrate on description at the focal level, while keeping implicit most of the contributions from finer scales. We are free, for example, to consider the growth and development of ecosystems without explicitly mentioning genes or the DNA embedded in them.

This position shows the link between Ulanowicz's approach with the earlier approach where the biological features of ecosystem constituents are black-boxed into trophic compartments in terms of energy and material throughput. We shall return to this issue in later sections, but first we shall examine a dissipative structure that comprises a simple model for self-organisation in ecosystems.

2.2 Dissipative structures and complex adaptive systems

A paradigm dissipative structure, discussed repeatedly in the work of the Brussels School, is Bénard convection, an experiment involving a shallow vessel with

[16] "Currency" refers to the energy and material that are essential to the growth of organisms within an ecosystem. "Compartment" refers to the organisms populations made up a trophic level in the food web of an ecosystem.

a source of heat below it. Initially the fluid and vessel are in thermodynamic equilibrium with their environment. In the equilibrium state, macroscopically, the fluid is homogeneous and motionless whilst, microscopically, the molecules of fluid pseudorandomly bump into one another. The experiment proceeds by gradually increasing the temperature of the heat source, inducing an energy gradient across the system. This energy gradient drives the system away from equilibrium. At low gradients energy is dissipated by conduction. Microscopic motion increases at the bottom of the fluid, and this increased motion is transmitted upward through the fluid eventually heating the air above it. During conduction, on average, there is no net upward movement of individual molecules of fluid; the fluid's viscosity confines each molecule to its layer.

Increasing the energy gradient moves the system farther from equilibrium. Once a critical threshold of gradient is crossed conduction becomes unstable, viscous force is overcome and a new means of energy transport emerges: convection. The form convection takes depends upon the boundary conditions of the system. In the most celebrated version of Bénard's experiment the vessel is cylindrical and open at the top. In this case, the convective structure that emerges is a lattice of hexagonal convection cells. The molecules in the Bénard cells travel upward through the centre of each cell carrying with them the heat they received at the bottom of the vessel. Reaching the top of the cell, molecules cool moving horizontally, releasing energy to the air above the vessel. By the time they reach the outer edge of the cell the now cooler molecules sink, moving vertically. Once at the bottom they travel horizontally receiving energy from the vessel. By the time they reach the centre again they are hot enough to repeat the process again. The structure of Bénard convection is organised: it is a complex pattern capable of efficiently moving energy through the system.

One steeped in the mechanistic view of the world might be tempted to view the Bénard cell as a collection of tiny heat engines transferring energy through the system. But we hope to disabuse our readers of this mechanistic interpretation. There are two key features that distinguish the Bénard system from heat engines: self-organisation and path-dependence. Unlike a heat engine, the Bénard convection cells emerge spontaneously as the system responds to the throughput of energy. The configuration of the cells is determined by internal constraints on the system. Furthermore, although the general pattern of the cells is invariant, the details of the pattern are path-dependent. In other words, Bénard convection recurs at the same threshold in repeated runs of the experiment with the same setup; nevertheless, the precise location of the cells, and in particular where cells that deviate from hexagonal (which inevitably, aperiodically form) is contingent, the result of microscopic fluctuations at the onset of convection [Velarde and Normande, 1980; Nicolis, 1989]. These properties are also important for understanding ecosystem dynamics as well, and they will be discussed in detail in the next two sections, below. Ecosystems are certainly much more complex than Bénard convection. However, like Bénard convection, ecosystems are driven by energy throughput, as well as exhibiting self-organisation and path-dependence.

For these reasons the dissipative structure, as a central metaphor for ecology, is far superior to the heat engine.

2.3 Self-organisation and ecosystems

Ecologist Claudia Pahl-Wostl advocates the self-organising systems approach to ecology in her book *The Dynamic Nature of Ecosystems.* After discussing the self-organising dynamics of the stock market, she applies the concept to ecosystems:

> The situation in ecosystems may be perceived along similar lines. Organisms are selective with respect to the type and the meaning of environmental signals received, they have certain degrees of freedom on how to react. The infinite number of possible combinations of internal and external states renders situations unique and prevents the approach to a stable equilibrium point. The question arises, how systems are organised that combine a high potential for change with the maintenance of function. Such self-organising systems with a distributed and flexible "control" are in sharp contrast to most technical systems where a centralized control optimizes a desired function. [Pahl-Wostl, 1995, pp. 48-9].

She continues by discussing some contributions to the understanding of self-organisation from the field of "computational ecosystems" citing Huberman [1988]. Then she expresses scepticism about the state of the understanding (circa 1995) of self-organisation in ecology:

> As ecologists we have to admit to not being able to contribute much knowledge in this respect. Intuitively we may be aware of the decentralised nature of ecosystems, of the fact that some type of organization must be generated within the system. However, any such insights have not yet been incorporated into theoretical concepts nor in our approaches towards management. Notions such as top-down or bottom-up control resonate much more with centralized than with distributed control. I argue that our current perception of natural systems from organisms to ecosystems resembles much more the idea of machines than of genuine living systems. [Pahl-Wostl, 1995, p. 49]

Perhaps Pahl-Wostl's scepticism is overstated. Certainly the idea of decentralised control was not entirely absent from bioscience circa 1995. The mathematical model of the aggregation of *D. discoideum* developed by Keller and Segal [1970] and subsequent empirical work validating the model reported by Garfinkel [1987] was explicitly aimed at showing aggregation was not under centralised control. Furthermore, also explicitly central to Christopher Langton's work on "artificial life" was the notion of decentralised control:

> [Langton] would later codify the elements of this approach in defining the essential features of computer-based a-life models: ... there is no

> single program that directs all of the other programs; ... there are no rules in the system that dictate global behaviour; and any behaviour at higher levels than the individual programs is therefore emergent. [Levy, 1992, p. 106]

And the situation has certainly improved in ecology since 1995. There are now ecological theory texts (aside from Pahl-Wostl's) from the self-organisation perspective [Jørgensen and Müller, 2000; Jørgensen, 2002; Solé and Bascompte, 2006]. A 1998 issue of *Ecosystems* featured a special section on complex adaptive systems and ecology (see [Hartvigsen *et al.*, 1998]). The ideas of L. H. Gunderson, C. S. Holling [2002] and associates have been explicitly aimed at ecosystem management from a self-organising/complex adaptive systems perspective.

Nevertheless, a focus on self-organisation in ecology definitely has not entered mainstream research. And certainly the notion of self-organisation is one of the most important contributions the study of complex dynamics can contribute to ecology. In what follows we do not attempt a precise definition of self-organisation. Instead we focus on the conditions for self-organisation, and the implications for the study of ecology. We will formulate these with direct reference to the discrete models contributed by complex adaptive systems modellers because they are more straightforward, but we assume corresponding formulations would be possible for other types of models (e.g. those based on continuous differential equations).

We start by identifying the constitutive properties[17] of systems capable of self-organisation:

- The focus is on systems composed of *discrete individual elements*; each element is simple and capable of being in two or more discrete states.
- The elements are *interconnected*; collections of elements form an interacting network.
- The system's dynamics is the result of *local interactions*; generally these interactions are governed by simple deterministic rules.
- The system exhibits *distributed global control*, that is, the global dynamics of the system is not controlled by any single element, and any regulation in the system results from interactions between the components.

Perhaps the simplest example of such a system is Conway's Game of Life discussed earlier. What we saw in that example is that patterns of enormous complexity can emerge in such systems. Complex functions can be implemented in the game, including universal computation [Emmeche, 1994; Langton, 1992]. The important point is that these structures and functions are neither *imposed* nor *centrally controlled*. Hence, they are *self-organised*, emerging from the collective interaction of the elements of the system.

[17]There exist many such lists of conditions and properties [Hartvigsen *et al.*, 1998; Levin, 1998; Arthur *et al.*, 1997; Emmeche, 1994]. We were not satisfied with any of them. We have tried to provide a set of minimal conditions for self-organisation in complex adaptive systems.

On the surface at least it seems that ecosystems possess the constitutive properties we have identified for systems capable of self-organisation. And certainly ecologies appear to exhibit the spontaneous emergence of structure. This suggests that constructing models of self-organisation would be a fruitful avenue for the ecosystem ecology.

One issue that arises when proposing self-organising models for ecology is the implication for the role of Darwinian evolution by natural selection. As we discussed above, Levin makes an explicit "selection process" a condition for a complex adaptive systems model. If we follow this stipulation we have no problem; by definition, complex adaptive systems employ selection. However, this seems overly restrictive. Many "Santa Fe" type models (including Conway's Game of Life) employ no such selection mechanism.[18] Kauffman has devoted enormous effort to articulating the role of self-organisation in evolutionary systems. He argues that both processes are present in evolution, and that the processes are complementary.[19] Langton [1992] proposed that being poised at "the edge of chaos" makes systems more adaptive.[20] Kauffman has simulated coevolving systems self-organised converging on the "edge of chaos" critical state. This work points to the ubiquity of "self-organised criticality" in a wide range of complex systems. Per Bak and colleagues [Bak and Chen, 1991] coined the term to describe systems that are "tuned" to a state on the edge of instability. Their paradigm is a sand pile created by dropping sand one grain at a time; the pile gets higher and higher until it becomes unstable and an avalanche occurs. The occurrence of avalanches follows a $1/f$ power law [Bak and Chen, 1991]. The same power law occurs in the models of Kauffman and Langton.

Our purpose here is neither to articulate the details of this work nor to settle the issue of "panselectionism". However, we do make the following points: There is no denying the role of natural selection in evolution. The jury is still out on the respective roles natural selection and self-organisation play in evolution. We can see no reason but the adherence to Darwinian dogma to restrict complex adaptive systems models to those employing a selection mechanism. We would expect self-organisation to be an important phenomenon occurring more likely at ecosystem level where a wide range of factors[21] play a part in the dynamics.

Adopting the self-organising complex adaptive system as the exemplar of ecosystem dynamics reflects a dynamic perspective apt to push ecology forward. Viewed from the "post-classical" dynamic perspective ecologists will expect their (mathe-

---

[18] There is no doubt that there exist many examples of complex adaptive systems models explicitly employing selection mechanisms. Genetic algorithms [Holland, 1992] spring immediately to mind.

[19] See Depew and Weber [1995, ch. 16] for an excellent account.

[20] Crutchfield and colleagues showed a similar result for continuous dynamic systems. See Langton [1992] and references therein.

[21] Such factors include, but are not limited to, coevolution, body and habitat size, hunting or foraging range, as well as life span, reproductive cycle, and other temporal characteristics in organisms' life histories, such environmental heterogeneity as fractal properties in living surface or space, as well as spatial and temporal variation in light intensity, humidity, temperature, and nutrient supplies.

matical and empirical) models to possess the following characteristics:

Lack of steady state or dynamic equilibrium Ecosystems are far-from-equilibrium systems. Such systems are expected to be dynamic. Of course a steady state or dynamic equilibrium may emerge in such a system, but this would require explanation.

Novelty and unpredictability Ecosystems are complex and dynamic. Novel structures would be expected to emerge in such systems, and the details of their behaviour would be expected to be largely unpredictable except under special circumstances.

Nonoptimality Adaptive systems are by and large satisfiers not optimisers. Multiple adaptive peaks are typical in rugged fitness landscapes. Furthermore there seems to be a trade-off between efficiency and resilience in such systems [Ulanowicz, 1997; Gunderson and Holling, 2002]. As available energy utilisation becomes more efficient the ecosystem becomes "brittle", that is, it is more vulnerable to catastrophic disruption from external perturbation.

Cycling Fundamentally ecosystems exhibit cyclical behaviour resulting from the interaction of fast and slow variables of the system. Organisational and dynamic features of ecosystems need to be understood within its cyclical pattern. For example, Gunderson and Holling [2002, pp. 33-4] propose an adaptive cycle as the core metaphysical framework to describe the organisation and dynamics of ecological systems.

Emergent constraints An ecosystem is composed of a system of interacting elements. Over time constraints emerge from the interactions in the system that both restrict and enable future interactions. For example, [Lansing *et al.*, 1998, p. 378] identify processes of "system dependent selection" whereby internal factors can drive functional organisation.

Historicity Ecosystem dynamics are path-dependent. Small events become locked-in to constrain the future dynamics and emergence of structure in the ecosystem. This requires a natural historical approach to understand the dynamics.

Diversity The diversity of individual organisms' biological characteristics and environmental heterogeneity are the essential conditions for understanding the energy/material/information flow pathway networks in ecosystems.

Locality The characteristic dynamics and organisation of ecosystems are site-specific, which means that the system's dynamics and organisation vary from one location to another.

3 COMPLEX DYNAMICS AND THE PHILOSOPHY OF POST-CLASSICAL ECOSYSTEMS ECOLOGY

The characteristics highlighted in the previous section are the focus of understanding ecosystems from a self-organising dynamic perspective. The adoption of this dynamic perspective prompts a rethink of fundamental issues in the philosophy of ecology. Such issues include the role of contingency, the importance of spatiotemporal details and the utility of general laws in ecological science. It also suggests a new strategy for the application of complex dynamic models to pragmatic issues. We will briefly explore these issues in the rest of this article.

3.1 The necessity of contingency: path-dependence and historicity

The term path-dependence comes from economics. A system is path-dependent if its present state depends crucially on some event or events that occurred in its prior history. A class of path-dependent phenomena exhibit lock-in: up to a certain point in the evolution of the system a range of possible future states is available to the system, but after lock-in some aspect of the system remains in a steady state. Brian Arthur [1994] has proposed a simple model of such dynamics based on Polya processes. For simplicity of exposition we will consider the n = 2 Polya process; Arthur [1994, chs. 3 and 10] and colleagues have analysed Polya processes for n < 2 and using a wide range of feedback functions.

The Polya model consists of an infinite-capacity urn and an infinite supply of balls of two colours. We start with one ball of each colour. The process starts by choosing a ball at random. That ball chosen is replaced along with another of the same colour. We continue the process *ad infinitum.* The key question is what happens in the long run to the proportions of balls of each colour as the result of such a process. (Since the proportion of one colour (say red) equals one minus the proportion of the other colour (say blue), i.e. $P_{red} = 1 - P_{blue}$, we will only discus the proportion of one colour, blue.) Polya proved that every such process will result in the proportion converging on a particular value, x. He also proved that x is a random variable ranging from zero to one.

It helps to visualise the result of several runs of the model in simulation. In each case the proportion of blue balls fluctuates randomly for the first 100 or so iterations. As the number of balls starts to get large the proportion of blue balls settles down on a particular number, $0 < n < 1$, with only small fluctuations around it. Once the proportion converges it is locked-in; the system is constrained to small variations around that proportion. Each run throws up a different convergent proportion. Arthur takes this to mean that in such processes the result depends on it history; where the proportion converges is a function of the events that precede it. Economists have applied the insights from this model to a wide-range of case studies including the QWERTY keyboard, VHS vs. Beta videocassette systems, rail tack gauges, coal wagons, chemical pest control, nuclear reactors, information technology, geography and trade and institutional development [Puffert, 2003].

The implications of path dependence and lock-in for the economics of technology has been discussed at length [David, 1986; Herfel, 2007; Leibowitz and Margolis, 1995; Puffert, 2003]; the implications for political science have also been discussed [Page, 2006; Pierson, 2000]. The chief implication is that reference to the particular details of the economic *history* of the case in question is necessary to explain the adoption of a technology [David, 1986].

The Polya process model is chosen, not because of its direct relevance to ecology, but for its pedagogic value. It is the most straightforward model of path dependence in the literature. A wide range of complex systems exhibits path dependence. One example occurs in dissipative structures. E.g. in Bénard convection discussed above, macroscopic details of particular instantiations depend on microscopic fluctuations of the molecules comprising the system at the onset of convection [Nicolis, 1989, p. 319). Certainly, the notion of path dependence is relevant to biology in general and ecology in particular. As complex systems, ecosystems exhibit path-dependency and locking-in effects, i.e. some of the historical events, even though they may appear trivial when they occur, get locked in and their result amplified in the system over a period of time.

For example, Allison and Hobbs [2004] identify such a lock-in process operating in agricultural practices in the Western Australian ecosystem. They identify clearing of native vegetation as a major catalyst initiating a cycle of positive feedback between alterations of microclimate, the hydrological cycle, rising water tables and biodiversity reduction. Each element in the cycle reinforces the next forming a closed-loop process. Furthermore, such factors as agricultural intensification and rapid degradation of natural resources contributed to keeping the WA ecosystem locked in to this cycle. Once this cycle is established the ecosystem is irreversibly constrained to a limited set of trajectories. Species are lost and the ecosystem remains monocultural, vegetation maintained only through continued agriculture practice. Even if agriculture is abandoned the ecosystem remains degraded, with most native species never returning to the region. This sort of process recurs in various regions throughout the globe eventually resulting in forest, scrubland, and savannah replaced by desert. This process accompanies intensive agriculture with notable examples being the Sahara and the Western United States of America[22], as well as various regions in Australia. In general, such regions remain desert, rarely returning to their previously observed ecosystem structures.

Such path-dependency and locking-in effects manifest in the form of the historicity and locality of ecosystem EM flux network structures in terms of their specific spatiotemporal details, as occurs in Bénard cell formation, only in a more discernable manner. These features have been ignored by ecologists for both conceptual and practical reasons. As we noticed earlier, Lindeman and the Odum brothers and their followers black-boxed constituent organisms into trophic compartments of a food chain or web. Such strategies were also adopted by later ecologists, including Ulanowicz. While acknowledging the self-organising nature of the flux network, he nonetheless takes the same approach when conceptualiz-

[22]For a comprehensive analysis of this phenomenon in the USA see Reisner [1993].

ing the structure of the network [Ulanowicz, 1997, pp. 56-60]. For Ulanowicz, the autocatalytic process is the major driving force that imposes structure to an ecosystem [Ulanowicz, 1997, pp. 81-6].

Practical reasons for ignoring the constituent organisms' biological characteristics and their implications for the historicity and localities of ecosystem flux network are the difficulties of obtaining such detailed data, however important such data might be to determining the right dynamic model of the ecosystem in question. Dunne [2006, 71] commented on these problems recently:

> Such approaches have much greater data requirements than just trying to characterize network structure. In many cases it will be difficult to document the necessary link and species characteristics given time and monetary constraints. Also, many of the concerns about "binary" data apply to "quantitative" data, and may be even more complicated to resolve. For example, biomass flow along trophic links and levels of species abundances vary spato-temporally. A snapshot of a set of flows or abundances at a particular time and in a particular space ignores variability that may result in a very misleading picture of typical flow or abundance levels. Indeed, it may be difficult to characterize "typical" or average levels that are ecologically meaningful.

Conceptual and practical reasons notwithstanding, it is important that historicity and its consequences for locality are taken into account when modelling ecosystem dynamics because of the implications of such dynamic properties for understanding ecosystem behaviour. Ecosystem historicity has three implications in relation to how an ecosystem would respond to disturbances: i) The impact of any disturbance to an ecosystem is highly unpredictable. ii) What happened to an ecosystem in the past has a strong impact on how the system responds to a disturbance or stimulus. iii) The timing (e.g. with respect to developmental phase) and magnitude (e.g. with respect to threshold effects) of a disturbance is critical to how an ecosystem responds to the disturbance.

With respect to this last implication, a small change in a disturbance's magnitude at the threshold might drive the system from one type of dynamic state to a different one [Gunderson and Holling, 2002, pp. 27-30]. For example, Scheffer *et al.* [2002, pp. 196-7) show that, for an ecosystem at a critical time, i.e., when the system's stress level is near a critical threshold, a small increase in stress could have a catastrophic consequence. Such a threshold is strongly influenced by both an ecosystem's intrinsic resilience and the stage of a system's development.

3.2 Organodynamics: spatiotemporal details in ecology

There are expanding efforts to take *spatiotemporal details* of ecosystems properly into account and in this section we focus on more systematic work, here termed 'organodynamics' that is, the dynamics that emerge from strictly biological properties of relationships between the organisms inhabiting the ecosystem, especially their spatiotemporal aspects. These properties and relationships include, but are not limited to, coevolution, body and habitat size, hunting or foraging range, as well as life span, reproductive cycle, and other spatio-temporal characteristics in organisms' life histories, such environmental heterogeneity as fractal properties in living surface or space, as well as spatial and temporal variation in light intensity, humidity, temperature, and nutrient supplies. As we discussed earlier, ecologists often ignore the biological features of organisms when constructing their models. In a range of ecological research, such features are *averaged out*, and hence have no dynamic effect when modelling ecosystem dynamics and organisation [Ulanowicz, 1986; 1997].

As mentioned earlier, the shortcomings of such omission have been gradually acknowledged by some ecologists and efforts have been made to rectify the problem. Pahl-Wostl [1995], Keeling [2000], Solé and Bascompte [2006] and others have investigated the spatiotemporal patterns found in ecosystems. Computer simulations are used to investigate the spatiotemporal properties of ecosystem dynamics and organisation [Pahl-Wostl, 1995; Solé and Bascompte, 2006]. Ecologists equipped with far-from-equilibrium thermodynamics and complex adaptive systems theories have investigated how organisms' existential activities influence ecosystems' organisation and dynamical properties. Among them, Pahl-Wostl's work deserves a closer examination because it is the first to adopt a complex system approach to model ecosystem organisation and dynamics and for its profound philosophical implications for ecological investigation in general.

Pahl-Wostl's work, to some extent, is a development of Ulanowicz's mathematical model of EM flux pathway networks in ecosystems, adopting Ulanowicz's concept of ascendancy and its mathematical model to analyse them. However, the significance of her contribution to their research lies with her account of the detailed organodynamic processes that contribute to the emergence of these pathway networks and their organisational properties. In Ulanowicz's model, the EM flux network is imposed by an autocatalytic process that is relatively independent of its constituent organisms. Pahl-Wostl, however, argues that these flux networks self-organise. In order to make this point, she makes an explicit effort to distinguish the imposed organisation from self-organisation. According to Pahl-Wostl, self-organisation has the following properties: 1) it is endogenously generated; 2) it has distributed control; 3) it is a flexible network of interaction; 4) it has high potential to adapt and change; and 5) it is unpredictable. While the imposed organisation is different in the following aspects: 1) it is externally imposed; 2) it is centralized; 3) it has rigid networks of interaction; 4) it has little potential to adapt and change; and 5) it is predictable. She argues that the current understanding of ecological organisation is more of the imposed organisation rather than the self-organisation [Pahl-Wostl, 1995, pp. 50-1].

Pahl-Wostl focuses on the peculiar biological characteristics of the constituent organisms of an ecosystem and the impacts of such characteristics on the EM flow pathway network structure. She extends Ulanowicz's basic concept of ascendency and its mathematic model by: a) distinguishing self-organisation from imposed organisation and demonstrating the EM flow network organisation as self-organising rather than imposed organisation; b) incorporating organisms' specific biological characteristics and environmental heterogeneity into her model and explicating the detailed mechanisms of how such pathway network organisation emerges; c) focusing on the detailed biological/ecological processes that contribute to the emergence of the pathway network. While Ulanowicz claims that organisms' biological characteristics could be safely ignored when modelling ecosystem organisation and dynamics, Pahl-Wostl argues that the diversity in the biological properties of the constituent organisms and in the heterogeneity of their environment are *essential* for the emergence of the flow networks in ecosystems.

According to Pahl-Wostl, an organisms' biological characteristics which she terms as dynamic characteristics, including body weight, foraging range, living space or surface, dietary habit, mating habit, active time (hunting or nutrient absorbing) during the circadian or annual cycle, reproductive cycle, and life span, as well as environmental characteristics including spatial and temporal properties, temperature, water and nutrient supplies in the environment, have a strong impact on how the flow networks are organised. She demonstrates, with a series of computer simulations as well as real ecological examples, that the EM flow networks emerge spontaneously from the interaction between organisms and their immediate environment, which includes both physical surroundings and other organisms. Her work also shows that the diversity in organisms' dynamic characteristics and environmental heterogeneity are essential for such organisations to emerge [Pahl-Wostl, 1995, chs. 4-6].

The spatiotemporal organisation of flux networks, according to Pahl-Wostl's model, is constructed under two sets of constraints. Firstly, because ecological organisation emerges from the ongoing activities conducted by the constituent organisms that are simply "making a living,"[23] they are constrained by the dynamic characteristics of the constituent organism. Secondly, the organisms' existential activities are constrained by the dynamic properties of their immediate environment. These environmental properties include climate conditions, seasonal cycles, nutrient availability, light intensity, temperature range and rainfall range, the spatial distribution of food, nutrients, and water, etc., as well as other organisms that either share the physical environment or interact with them via symbiosis, competition or predation. Spatial and temporal organisation of EM pathway networks, in turn, is mediated through such properties hence bears the marks of these properties.

Based on Ulanowicz's ascendency concept and its mathematical formulation, Pahl-Wostl [1995, pp. 139-40] defines spatio-temporal organisation as "the internal network structure that organises itself at any moment in space and time to fill the envelope determined by these constraints in combination with the boundaries imposed by the physical environment". And the pathway network's spatial and temporal organisation is measured by the flow from one compartment to another

[23] "Making a living" is a term used by Kauffman [2000, ch. 7] to refer to the various activities an organism does in order to survive as well as reproduce themselves.

along the spatial s and temporal t dimensions:

$$I_{ts} = \sum_{j=0}^{n} \sum_{i=1}^{n+2} \sum_{t=1}^{r} \sum_{l=1}^{q} \frac{T_{jikl}}{T} \log \frac{T_{jikl} T_{jikl} T}{T_{ji..} T_{,jkl} T_{j.kl}}$$

where $\{T_{jikl}\}, j = 0, 1, 2, \ldots n; i = 1, 2, ..., n+2$, refers to the intercompartmental flow set over r time intervals within q spatial intervals. T refers to the total throughput of the system [Pahl-Wostl, 1995, pp. 59-60].

On the surface, this spatiotemporal organisation of pathway networks is only slightly different from Ulanowicz's formal flux networks. However, without the spatio-temporal heterogeneity in the network model, as will be seen later, one misses out on many key dynamic properties which are essential to explain the emergence of flux networks as well as their dynamic characteristics. We shall turn to these properties in the following section.

One such feature is the coexistence of seemingly competing species within an ecosystem shown by Pahl-Wostl via a series of computer simulations. She demonstrates that species that extract the same resources but with different biological characteristics can coexist, and they do not necessarily behave according to the accepted "mutually exclusive competition principle", i.e., all but one species that is the fittest among them would eventually become extinct. Instead, the species of the ecosystem organise themselves into rather complicated EMI pathway networks, which manifest in a spatiotemporally unique organisation. For example, one of the computer simulations shows that when 40 species extracting the same resources but with different temporal niches (being active at different times of the circadian or annual cycle) are put into one ecosystem, a six-species group emerges with a sophisticated temporally differentiated pathway network organisation as the dominant assembly and achieves the ecosystem's optimum of temporal ascendency (Asct).[24] Pahl-Wostl [1995, pp.98-100] claims that this "implies that the system organises itself into a configuration where the Asct is maximal."

Pahl-Wostl [1995, p. 102] also demonstrates, by computer simulation as well as real examples, that: "As the number of species is increased, specialization (= narrowing of the temporal niche) confers an advantage." And a species ensemble could lift the growth limits that were set by nutrient availability in its environment by simply speeding up resource cycling process. She suggests that specialization (or temporal differentiation in organisms' resources utilization) and the concomitant temporal organisation in the pathway network can be regarded as cooperative processes. "We can talk here about cooperation without invoking any notion of deliberate altruistic behaviour. Cooperation arises simply from the effects of positive feedback spirals in a dynamic network where an ensemble of species shares a limiting resource" [Pahl-Wostl, 1995, pp. 102]

These changes in the perceived sources of ecological organisation and dynamics not only bring new approaches to such traditional problems as ecosystem stability,

[24]Temporal ascendency is the measure of temporal organisation. For the mathematical definition, see Pahl-Wostl [1995, pp. 59-61].

complexity, competition and succession, but also raise a new set of problems that, whilst inaccessible within the conventional mechanistic framework, could be addressed from the new perspective. Such problems include individual species' niche construction and differentiation, the roles played by individual organisms' existential activities in constructing EMI flow pathway networks, ecosystem resilience, organisms' biological diversity and environmental heterogeneity and their roles in ecosystem organisation and dynamics.

3.3 Ecopluralism: Laws and Models

In the course of development of the study of complex systems dynamics there have been two distinct approaches. One approach is to tackle particular problems through the articulation of a model; another approach is to search for a general law that holds across a class of systems. Both approaches have a venerable position in the history of science. And of course the two approaches interact: laws are routinely used to provide models; modelling natural systems can yield general laws. A particular pattern has emerged in the course of the development of nonequilibrium thermodynamics: Prigogine proved the law of minimum entropy production, but he found that it only held near equilibrium. As effort was concentrated on what happens in cases far from equilibrium laws have not been forthcoming.[25] The later work of the Brussels School catalogues a range of dynamic models for far-from-equilibrium systems but has failed to propose new laws. One possibility is that the systems of interest in the far-from-equilibrium regime are those which generate novelty. If behaviour is *determined* by a general physical law it could hardly count as novel.

Another consideration is the trade-off between explanatory power and truth that Nancy Cartwright discusses in *How the Laws of Physics Lie*. In this book Cartwright [1983] compares general laws to phenomenological models in physics. General laws apply to a whole class of cases so their explanatory power is high, yet without numerous auxiliary assumptions the laws tell us little about particular cases. Since theories are measured against data from particular cases, the laws are light on truth content. Phenomenological laws are just the opposite: they are supported by experiment yet low on explanatory power. Leaving aside the question of truth, we would reformulate the trade-off like so: laws have high explanatory power and low information content; models have high information content but low explanatory power. Here information has its usual meaning in terms of resolution of ambiguity. Without the additional assumptions required to specify the model, a general law is compatible with a wide range of dynamic behaviour. Specifying the conditions of the model resolves ambiguity; with enough such constraints even a particular trajectory can be specified. A good example is the second law of thermodynamics. It tells us that entropy must increase in spontaneous transitions and transformations in isolated systems. However the law tells us nothing about

[25] Nevertheless, others, working outside the Brussels School, e.g. Swenson [1989], have proposed such laws.

the paths of these transitions and transformations.[26]

The trade-off between explanatory power and information content is particularly relevant when we are interested in the dynamics of complex systems. When the second law is applied to open systems, it implies that all entropy decreases within the system must be offset with even larger entropy increases in the system's environment. This law applies to everything. Nevertheless it tells us nothing of the details of the dynamic behaviour of such systems.

Ecology is not without such general laws. For instance, several versions of a "fourth law of thermodynamics" have been proposed. Most of these laws are refinements of Lotka's [1922a; b] principle which is an attempt to apply the concept of available energy to the theory of Darwinian evolution. Lotka [1922a, p. 147] states,

> It has been pointed out by Boltzmann that the fundamental object of contention in the life-struggle, in the evolution of the organic world, is available energy. In accord with this observation is the principle that, in the struggle for existence, the advantage must go to those organisms whose energy-capturing devices are most efficient in directing available energy into channels favourable to the preservation of the species.

A recent attempt to refine this law can be found in the work of theoretical ecologist Sven Jørgensen [2000]. Other possible laws for ecosystems include results from work on "self-organised criticality" [Bak and Chen, 1991; Kauffman, 1993; 1995; Langton, 1992][27]. Cycles of succession [Ulanowicz, 1997; Gunderson and Holling, 2002] and theories of ascendancy [Ulanowicz, 1997; Pahl-Wostl, 1995] may also be sources of general laws in ecology. It is beyond the scope of this article to attempt evaluation or even exposition of this work. However, we would ask if Cartwright's trade-off applies to physics (and we think the case is strong) would it not apply equally in ecology?

Although there have been no shortage of proposals for laws in the area of complex systems dynamics the significant contributions have been made through the construction of dynamic models, usually with the aid of computer simulation. As we stated above, most of the work of the Brussels School consists of dynamic models of particular systems. Chaos theory really is not a theory at all, but a family of models all with similar characteristics, i.e. initial conditions sensitivity and aperiodicity [Herfel, 1990, ch. 2]. And certainly complex adaptive systems research involves mainly computer models of particular systems. Even work by Bak and colleagues [Bak and Chen, 1991] on self-organised criticality is an attempt to generalise the power law characteristics of his sand pile model to a class of systems. We agree with science journalist John Horgan's [1995; 1996, ch. 8] scepticism towards "theories of everything". In an article on applying complex systems models to anthropological phenomena, Steve Lansing [2003, p. 18], quotes Horgan:

[26]Cartwright [1983, p. 65] sites work by C. A. Truesdell who critiques the Onsager approach along the same lines.

[27]Langton's results have been critiqued by Mitchell *et al.* [1994].

> ...as the philosopher Karl Popper pointed out, prediction is our best means of distinguishing science from pseudo-science. The history of 20th-century science should also give complexologists pause. Complexity is simply the latest in a long line of highly mathematical 'theories of almost everything' that have gripped the imaginations of scientists in this century.

However, and ironically, while the attempt to construct theories of everything does drive some of the work at the Santa Fe institute (Lansing's work is a noteworthy exception), and these sort of proclamations may capture the headlines in the popular press, complex dynamic models of particular phenomena make an important, and perhaps the largest, scientific contribution. Nevertheless, Horgan's remark does raise an important issue. A traditional role for scientific models is to forecast the future behaviour of the systems that they model. The predictive utility of complex systems dynamics in general, and ecological modelling in particular, has been called into question by a wide range of critics. It is to the question of prediction, in particular, and to the issue of the utility of complex dynamic models, in general, that we now turn.

Explanation and prediction of natural phenomena is no doubt an important job of scientific models. But complex systems dynamics introduces a need for subtlety when examining the issue of prediction. As Horgan [1995] elegantly puts it, "chaos turned out to refer to a restricted set of phenomena that evolve in predictably unpredictable ways." If a system is chaotic, initial conditions sensitivity insures that an accurate model will be severely limited in its ability to predict the local details of the system's dynamics. This issue has been explored in detail many times (e.g. [Herfel, 1990; Kellert, 1993]). Initial systems sensitivity explains precisely why we cannot hope for reliable local medium-term prediction of chaotic systems. Of course, complex dynamic models can contribute to the identification of under what conditions to expect chaos, what range and character of behaviour to expect in the chaotic regime, etc. And this is all valuable information.

Ecological models have been criticised for their lack of predictive utility.[28] As important as both prediction and understanding its limits are in science, mathematical models perform other useful functions.[29] Models allow scientists perform mathematical experiments. Models help scientists define and refine research questions. Models can generate new concepts. Models contribute exemplars to scientific paradigms.

The advent of efficient inexpensive computers has revolutionised mathematical modelling. The implementation of mathematical models in digital computers allows scientists to experiment with a wide-range of mathematical relationships, parameters and initial conditions. These models may not exactly correspond to the real world, but allow the exploration of the range of possible dynamics that

[28]Two visible examples are Lawton [2000] and Sagoff [2003]. See Simberloff [2004] and de Laplante and Odenbaugh [forthcoming] for critiques of these critiques.

[29]Hannon and Ruth [1997, pp. 22, 3] provide a list of modelling functions in their section entitled "Why Model?" Their list has influenced ours.

can be realised. It is perhaps this sort of experimentation that ecologist Richard Levins [1984, p. 20] had in mind when he stated:

> ...even the most flexible models have artificial assumptions. There is always room to doubt whether a result depends on the essentials of a model or on the details of the simplifying assumptions. This problem does not arise in the more familiar models, such as the geographic map, where we all know that contiguity on the map implies contiguity in reality, relative distances on the map correspond to relative distances in reality, but colour is arbitrary and a microscopic view of the map would only show the fibres of the paper on which it is printed. In the mathematical models of population biology, on the other hand, it is not obvious when we are using too high a magnification. Therefore we attempt to treat the same problem with several alternative models each with different simplifications but with a common biological assumption. Then, if these models, despite their different assumptions, lead to similar results, we have what we call a robust theorem that is relatively free of the details of the model. Hence our truth is the intersection of independent lies.

So according to Levins, there is an ironic twist on Cartwright's thesis: in population ecology at least, the models lie as well! It is only through experimentation with a range of models, each with different simplifying assumptions that ecologists can hope to realistically capture ecosystems' dynamics. Hence from the post-classical dynamic perspective ecology must be essentially *pluralistic*.

Models are also important because of the role they play in defining and refining research questions. What may be vague ideas when presented verbally can become precise when expressed in a mathematical relationship. Through the process of constructing and simulating a mathematical model precise hypotheses are articulated and refined. Hannon and Ruth [1997, p. 23] state, "...a good model is a good thought-organising device. Sometimes most of the value in modelling comes from a deeper understanding of the variables involved in a system that people are routinely struggling to control." In the process of modelling a system new hypotheses may arise. Hannon and Ruth [1997, p. 23] continue, "Good modelling stimulates further questions about the system behaviour, and in the long run, the applicability of any newly discovered principles." This leads to a third important function for models.

As Hannon and Ruth state, models can lead to new principles, but even more significantly perhaps they can lead to new concepts. We have discussed many of them throughout this paper: self-organisation, initial conditions sensitivity, self-organised criticality, edge of chaos. No matter how widely applicable the principles associated with these concepts turn out to be, the development of the concepts is a significant contribution to post-classical ecology.

Finally, following Kuhn [1962] we assert that models play a significant role in providing exemplars articulating a paradigm-like *dynamic perspective*, that is, to the use of a general kind of dynamics, in this case complex adaptive systems dynamics with all of its peculiar features just noted. A paradigm plays a role in structuring a research field and extends to cover a range of general assumptions (definitions of key terms, what sort of models are germane, what counts as an anomaly, etc.) In one sense a dynamic perspective is narrower: it applies only to what sorts of dynamics one expects. But in another sense it can be broader insofar as it can apply across a range of research fields. The dynamics generated by the exemplary models of self-organising complex adaptive systems, as well as the concepts generated by the models, play an important role in articulating the post-classical dynamic perspective.

3.4 Ecosystems pragmatics: Human-environment interaction

Some of the most pressing issues of our time involve the human impact on ecosystems. Problems range from global warming to reduction in biodiversity. In this section we would like to make some remarks about the role of complex dynamic ecosystems models in addressing the human impact on the environment. We will start by describing one very successful example, the work done by anthropologist Steve Lansing and ecologist Jim Kremer on Balinese agriculture.

Lansing started his career as an anthropologist of religion and went to Bali to research religious practices. After spending some time there he realised that Balinese religious practices played a significant role in their agriculture. The main crop in Bali is rice grown in terraces on the rugged slopes of volcanic mountains. The rice paddies are organised into *subaks*, a collection of paddies run by a collective of farmers. During the dry season the *subaks* are irrigated from water that flows from the crater lakes at the mountaintops. Since water is limited, careful coordination is required so that *subaks* downstream are not deprived of water whilst crops are in the paddies. There are well over 100 *subaks* in a given watershed. What Lansing discovered in his first trip to Bali is that the priests in a series of water temples are responsible for the coordination of water flow in the watershed.

Prior to Lansing's arrival in Bali the Indonesian government had started an aggressive campaign, based on Green Revolution technologies, to increase agriculture productivity in Bali. By the time Lansing arrived this program was failing. Crops were being decimated by pest infestations, and the water control situation was in chaos. Lansing hypothesised that a major reason for the failure was that agricultural "experts" from the West implementing the Green Revolution ignored the crucial role that the religious practices played in rice farming.

Upon returning to the University of Southern California, Lansing discussed the possibility of constructing a model of the ecology of the *subaks* in order to investigate the role of water flow regulation in rice production. This was the start of a twenty year collaboration aimed at understanding the interrelationships within Balinese agriculture. It took several years and several trips to Bali to collect data and construct the model. The results of the model are documented in Lansing and Kremer [1993]. Although the system is very complex Lansing and Kremer found production was limited by two main variables: pest damage and water availability. The key to optimum production was to coordinate cropping so that adjacent *subaks* were left fallow in such a way that pest populations did not escalate out of control by moving from a recently harvested *subak* to one where rice was just maturing. It is a nontrivial problem to coordinate cropping patterns with water availability. Traditionally solving this problem is the work of the water temple priests that was ignored by the agricultural experts. Lansing and Kremer's [1993] model was designed to test a range of scenarios in the form of cropping patterns. They report that the actual cropping patterns coordinated by the water temples comes very close to the optimal patterns predicted by the model [Lansing and Kremer, 1993].

Since the original work was done Lansing and Kremer turned their attention to articulating their model in complex adaptive systems terms. This work is described in detail in Lansing's [2006] *Perfect Order*. Lansing and Kremer also had some success using the results of their model to convince the Indonesian authorities and Western experts that the water temples played a crucial role in the agricultural system they were trying to "develop". Their model is a prime example of a useful ecological model from the self-organising system perspective: it has a well defined question; it makes no claims for any general principles; it is supported by empirical data, and its lessons have practical significance.

4 CONCLUSION

Given its topic and audience, this paper has addressed the issues raised by the application of complex dynamics to ecology at an abstract level. Nevertheless, at its best, complex dynamics is applied to concrete problems. One of the leading practical concerns of the day is the destruction inherent in the current state of human-environment interaction, primarily the question of how to create an economy that will sustain human beings without undermining the ecosystems that thus far, are necessary for the survival of all life. Although dealing with practical concerns may not be traditionally the strong suit of the philosopher, it is the application of the complex dynamic perspective to such practical issues that will yield the biggest impact for the future of ecology. Thus we conclude this essay with a short discussion of the practical implications of this work.

In the introduction to *How Nature Speaks*, Haila and Dyke discuss the necessity of avoiding the temptation to 'totalization' in ecological modelling. By 'totalisation' they mean the tendency to take perfectly adequate analogies and turn them into complete, theoretical, general explanations of some phenomenon. One area of fruitful analogising has been the application of ecological concepts to human activity, in particular economic dynamics. They state:

> In the following we make a preliminary analogy between organismal physiology and human subsistence ... [which] depends on the tuning together of a vast complex of parallel and nested reproductive cycles. Small-scale cycles are constituted by the daily acts of human individu-

> als, which, of course, are built on the physiological cycles of the human organism. Large-scale processes are constituted by the metabolism of the community with its surroundings, that is, the local economy [Haila and Dyke, 2006, p. 22].

They continue this discussion introducing the notion of an analogue model, showing how such models contribute to the understanding of such complex systems.

> [Analogue models] point at important dynamic features shared between systems that are different from each other. Thus a fruitful analogy also helps specify which differences matter. [Haila and Dyke, 2006, p. 23]

It is through an appreciation of complex dynamics that we can come to grips with the deep issues raised by ecology:

> ...our assessment of what we can do, the evaluation of the scope of our agency, has to be carried out in terms of our capacity to deal with complexity. A particularly intriguing set of questions arises when we start to think of our capacity to *create* complexity. For it may be entirely possible to create unmanageable bifurcations by ourselves—change our world from one where, by and large, we can manage our problems to one where we can no longer do so.... The more interesting cases are systems that we inadvertently put into place, processes we unthinkingly get going, that begin to take on a life of their own, that begins to gain control of at least some of their own boundary conditions. [Haila and Dyke, 2006, p. 26]

A good example of this sort of dynamic is found in the work of Lansing and Kremer discussed in detail above. Adopting 'green revolutionary' agriculture involves more than merely utilising hybrid seed. Such seemingly innocuous changes in practice as adopting new seed can be the first step in creating a bifurcation in culture resulting in irreversible changes in the scale of complexity at a biosocial level. It remains to be seen if and how Balinese agriculture will "gain control of at least some of its own boundary conditions" and survive these changes in dynamic.

The greatest achievements applying complex dynamics to ecology will come from the application of complex adaptive systems concepts to concrete ecological problems. It is only by adopting, not only the models of complex systems, but most importantly its *dynamic perspective* that will enable us to have the insight to deal with the new ecological challenges awaiting us in the Twenty-first Century.

BIBLIOGRAPHY

[Allison and Hobbs, 2004] H. E. Allison and R. J. Hobbs. Resilience, Adaptive Capacity, and the "Lock-in Trap" of the Western Australian Agricultural Region. *Ecology and Society* 9(1), 3, 2004. `http://www.ecologyandsociety.org/vol9/iss1/art3.`

[Arthur, 1994] W. B. Arthur. *Increasing Returns and Path Dependence in the Economy.* Michigan University Press, Ann Arbor, 1994.

[Arthur *et al.*, 1997] W. B. Arthur, S. N. Durlauf, and D. Lane. Introduction. In W. B. Arthur, S. N. Durlauf, D. Lane (eds.), *The Economy as an Evolving Complex System II.* Addison-Wesley, Reading, MA., 1–14, 1997.

[Bak and Chen, 1991] P. Bak and K. Chen. Self-Organized Criticality. *Scientific American* 264(1), 46-54, 1991.

[Brown, 1977] H. I. Brown. *Perception, Theory and Commitment: The New Philosophy of Science*. University of Chicago Press, Chicago, 1977.

[Cartwright, 1983] N. Cartwright. *How the Laws of Physics Lie*. Cambridge University Press, Cambridge, 1983.

[Collier and Hooker, 1999] J. D. Collier and C. A. Hooker. Complexly Organised Dynamical Systems. *Open Systems and Information Dynamics*, 6, 241-302, 1999.

[David, 1986] P. A. David. Understanding the Economics of QWERTY: The Necessity of History. In W.N. Parker, ed., *Economic History and the Modern Economist*, Oxford University Press, Oxford, 1986.

[de Laplante and Odenbaugh, forthcoming] K. de Laplante and J. Odenbaugh. What Isn't Wrong with Ecosystem Ecology. In R. A. Skipper Jr., C. Allen, R. A. Ankeny, C. F. Craver, L. Darden, G. Mikkelson, R. Richardson, eds., *Philosophy and the Life Sciences: A Reader*, MIT Press, Cambridge, Mass., Forthcoming. Available online at `http://www.lclark.edu/~jay/What{\%}20Isn{\%}27t{\%}20Wrong{\%}20with{\%}20Ecosystem{\%}20Ecology.pdf` (accessed 9 June 2009).

[Depew and Weber, 1995] D. J. Depew and B. H. Weber. *Darwinism Evolving: Systems Dynamics and the Genealogy of Natural Selection*, MIT Press, Cambridge, Mass., 1995.

[Dunne, 2003] J. A. Dunne. The network structure of food webs. In M. Pascual, J. A. Dunne, eds., *Ecological networks: linking structure to dynamics in food webs*, Oxford University Press, Oxford, 2003.

[Emmeche, 1994] C. Emmeche. *The Garden in the Machine: The Emerging Science of Artificial Life*. Princeton University Press, Princeton, 1994.

[Feigenbaum, 1978] M. J. Feigenbaum. Quantitative universality for a class of nonlinear transformations. *Journal of Statistical Physics* 19(1), 25-52, 1978.

[Gao, 2005] Y. Gao. *Ecologies of Scientific Practice: An Analysis of the Organization and Dynamics of Science*. PhD dissertation. The University of Newcastle, Callaghan, Australia, 2005.

[Gardner, 1970] M. Gardner. The fantastic combinations of John Conway's new solitaire game "life". *Scientific American* 223(4), 120-123, 1970.

[Garfinkel, 1987] A. Garfinkel. The slime mold Dictyostelium as a model of self-organization in social systems. In F. E. Yates, ed., *Self Organizing Systems: The Emergence of Order*, Plenum, New York, 181-212, 1987.

[Golley, 1993] F. Golley. *A History of the Ecosystem Concept in Ecology: More Than the Sum of the Parts*, Yale University Press, New Haven, 1993.

[Gould and Lewontin, 1979] S. J. Gould and R. C. Lewontin. The Spandrels of San Marco and the Panglossian Paradigm: A Critique Of The Adaptationist Programme. *Proceedings Of The Royal Society of London, Series B,* 205(1161), 581-598, 1979.

[Grebogi *et al.*, 1984] C. Grebogi, E. Ott, S. Pelikan, and J. A. Yorke. Strange attractors that are not chaotic. *Physica D: Nonlinear Phenomena*, 13(1-2), 261-268, 1984.

[Gunderson and Holling, 2002] L. H. Gunderson and C. S. Holling, eds. *Panarchy: Understanding Transformations in Human and Natural Systems*. Island Press, Washington, 2002.

[Hannon and Ruth, 1997] B. Hannon and M. Ruth. *Modeling Dynamic Biological Systems.* Springer-Verlag, New York, 1997.

[Hartvigsen *et al.*, 1998] G. Hartvigsen, A. Kinzig, and G. Peterson. Use and Analysis of Complex Adaptive Systems in Ecosystem Science: Overview of Special Section. *Ecosystems*, 1, 427–430, 1998.

[Herfel, 1990] W. E. Herfel. *Coming Attractions: Chaos and Complexity in Scientific Models.* PhD dissertation. Temple University, Philadelphia, 1990.

[Herfel, 2007] W. E. Herfel, Positive Feedback and Praxiology: Path Dependence in Action . In W. W. Gasparski, T. Airaksinen, eds., *Praxiology and the Philosophy of Technology*, in the series Praxiology: The International Annual of Practical Philosophy and Methodology, Transaction Publications, Piscataway, NJ, 2007.

[Hooker, 2004] C. A. Hooker. Asymptotics, Reduction and Emergence. *British Journal for The Philosophy of Science*, 55, 435-479, 2004.

[Holland, 1992] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press, Cambridge, Mass., 1992.

[Horgan, 1995] J. Horgan. From complexity to perplexity. *Scientific American* 272(6), 104-110, 1995.

[Horgan, 1996] J. Horgan. *The End of Science: Facing the Limits of Knowledge in the Twilight of the Scientific Age*. Little, Brown & Co., London, 1996.

[Huberman, 1988] B. A. Huberman (ed.). *The Ecology of Computation*. North-Holland, Amsterdam, 1988.

[Jørgensen, 2002] S. E. Jørgensen. *Integration of Ecosystem Theories: A Pattern*. 3^{rd} edn. Kluwer, Dordecht, 2002.

[Jørgensen and Müller, 2000] S. E. Jørgensen and F. Müller. *Handbook of Ecosystem Theories and Management*. Lewis Publishers, Boca Raton, 2000.

[Kauffman, 1993] S. A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, 1993.

[Kauffman, 1995] S. A. Kauffman. *At Home in the Universe: The Search for Laws of Self-Organization and Complexity*. Oxford University Press, New York, 1995.

[Kauffman, 2000] S. A. Kauffman. *Investigations*. Oxford University Press, New York, 2000.

[Keeling, 2000] M. J. Keeling. Evolutionary Dynamics in Spatial Host-parasite Systems. In U. Kieckmann, R. Law, J. A. J. Metz, eds., *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*. Cambridge University Press, Cambridge, 2000.

[Keller and Segel, 1970] E. Keller and L. Segel. Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26, 399-415, 1970.

[Kellert, 1993] S. H. Kellert. *In the Wake of Chaos: Unpredictable Order in Dynamical Systems*. University of Chicago Press, Chicago, 1993.

[Kuhn, 1962] T. S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, 1962.

[Langton, 1992] C. G. Langton. Life at the Edge of Chaos. In C. G. Langton, C. Taylor, J.D. Farmer, S. Rasmussen, eds., *Artificial Life II,* Santa Fe Institute Studies in the Sciences of Complexity, Proceedings Volume X, Addison-Wesley, Redwood City, CA, 1992.

[Lansing, 2003] J. S. Lansing. Complex Adaptive Systems. *Annual Review of Anthropology*, 32, 183–2, 2003.

[Lansing, 2006] J. S. Lansing. *Perfect Order: Recognizing Complexity in Bali*. Princeton University Press, Princeton, 2006.

[Lansing and Kremer, 1993] J. S. Lansing and J. N. Kremer. Emergent Properties of Balinese Water Temple Networks: Coadaptation on a Rugged Fitness Landscape. *American Anthropologist*, 95(1), 97-114, 1993.

[Lansing *et al.*, 1998] J. S.Lansing, J. N. Kremer, and B. B. Smuts. System-dependent Selection, Ecological Feedback and the Emergence of Functional Structure in Ecosystems. *Journal of Theoretical Biology*, 192, 377-391, 1998.

[Lawton, 2000.] J. H. Lawton. Community Ecology in a Changing World. In O. Kinne, ed., *Excellence in ecology*. Book 11. International Ecology Institute, Oldendorf/Luhe, 2000.

[Levin, 2002] S. A. Levin. Complex Adaptive Systems: Exploring the Known, the Unknown and the Unknowable. *Bulletin (New Series) of the American Mathematical Society*, 40(1), 3-19, 2002.

[Levin, 1998] S. A. Levin. Ecosystems and the Biosphere as Complex Adaptive Systems. *Ecosystems*, 1, 431–436, 1998.

[Levins, 1966] R. Levins. The Strategy of Model Building in Population Biology. *American Scientist*, 54. 421-431, 1966.

[Levy, 1992.] S. Levy. *Artificial Life: The Quest for a New Creation*. Pantheon, New York, 1992.

[Lewontin, 1978] R. C. Lewontin. Adaptation, *Scientific American.* 239(3), 212-228, 1978.

[Liebowitz and Margolis, 1995] S. J. Liebowitz and S. E. Margolis. Path Dependence, Lock-In, and History. *Journal of Law, Economics, and Organization*, 11, 204-26, 1995. `http://wwwpub.utdallas.edu/~liebowit/paths.html`

[Lindeman, 1942] R.L. Lindeman. The trophic-dynamic aspect of ecology. *Ecology*, 23, 399-418, 1942.

[Lorenz, 1963] E. N. Lorenz. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20, 130-41, 1963.

[Lotka, 1922a] A. J. Lotka. Contribution to the Energetics of Evolution. *Proceedings of the National Academy of Sciences*, 8(6), 147-51, 1922.

[Lotka, 1922b] A. J. Lotka. Natural Selection as a Physical Principle. *Proceedings of the National Academy of Sciences*. 8(6), 151-4, 1922.

[May, 1976] R. M. May. Simple Models with Very Complicated Dynamics. *Nature*, 261, 259-67, 1976.

[Meinhardt, 1995] H. Meinhardt. *The Algorithmic Beauty of Sea Shells*. Springer-Verlag, Berlin, 1995.

[Mitchell *et al.*, 1994] M. Mitchell, J. P. Crutchfield, and P. T. Hraber. Dynamics, Computation, and the "Edge of Chaos": A Re-Examination. In G. Cowan, D. Pines, D. Melzner (eds.). *Complexity: Metaphors, Models, and Reality*. Santa Fe Institute Studies in the Sciences of Complexity, Proceedings Volume 19. Addison-Wesley, Reading, MA., 1994.

[Nicolis, 1989] G. Nicolis. Physics of far-from-equilibrium systems and self-organisation. In P. Davies (ed.). *The New Physics*. Cambridge University Press, Cambridge. 316-347, 1989.

[Page, 2006] S. E. Page. Path Dependence. *Quarterly Journal of Political Science*, 1, 87–115, 2006.

[Pahl-Wostl, 1995] C. Pahl-Wostl. *The Dynamic Nature of Ecosystems*. Wiley, New York, 1995.

[Peters, 1991] R. H. Peters. *A Critique for Ecology*. Cambridge University Press, Cambridge, 1991.

[Pierson, 2000] P. Pierson. Path Dependence, Increasing Returns, and the Study of Politics. *American Political Science Review*, 94(2), 251-67, 2000.

[Prigogine, 1980] I. Prigogine. *From Being to Becoming: Time and Complexity in the Physical Sciences*. W. H. Freeman & Co., New York, 1980.

[Puffert, 2003] D. Puffert. Path Dependence. R. Whaples (ed.). EH.Net. Economic History Services. Encyclopedia.June 10, 2003. URL `http://eh.net/encyclopedia/article/puffert.path.dependence`. (Accessed 7 July 2009.)

[Reisner, 1993] M. Reisner. *Cadillac Desert: The American West and Its Disappearing Water*. Penguin, New York, 1993.

[Rolle, 1999] K. C. Rolle. *Thermodynamics and heat power*. Prentice Hall, Upper Saddler River, NJ, 1999.

[Sagoff, 2003] M. Sagoff. The Plaza and the Pendulum: Two Concepts of Ecological Science. *Biology and Philosophy*, 18, 529-52, 2003.

[Scheffer *et al.*, 2002] M. Scheffer, F. Westley, W. A. Brock, M. Holmgren. Dynamic Interaction of Societies and Ecosystem—Linking Theories from Ecology, Economy, and Sociology. In [Gunderson and Holling, 2002, pp. 195-239].

[Simberloff, 2004] D. Simberloff. Community Ecology: Is It Time to Move On? *The American Naturalist*, 163(6), 787-99, 2004.

[Solé and Bascompte, 2006] R. V. Solé, J. Bascompte. *Self-Organization in Complex Ecosystems*. Princeton University Press, Princeton, 2006.

[Stewart, 1989] I. Stewart. Warning—Handle with Care. *Nature*, 355(2), 16-7, 1989.

[Swenson, 1989] R. Swenson. Emergent Attractors and the Law of Maximum Entropy Production: Foundations to a Theory of General Evolution. *Systems Research*, 6(3), 187-197, 1989.

[Taylor, 2005] P. J. Taylor. *Unruly Complexity: Ecology, Interpretation, and Engagement*. Chicago University Press, Chicago, 2005.

[Toulmin, 1961] S. Toulmin. *Foresight and Understanding: An Enquiry into the Aims of Science*. Harper, New York, 1961.

[Ulanowicz, 1997] R. E. Ulanowicz. *Ecology: the Ascendent Perspective*. Columbia University Press, New York, 1997.

[Ulanowicz, 1986] R. E. Ulanowicz. *Growth and Development: Ecosystems Phenomenology*. Springer-Verlag, New York, 1986.

[Velarde and Normande, 1980] M. G. Velarde ad C. Normande. Convection. *Scientific American*, 243(1), 78-93, 1980.

[Wittgenstein, 1969] L. Wittgenstein. *On Certainty*. G. E. M. Anscombe, G. H. von Wright (eds.). Blackwell, Oxford, 1969.

COMPLEX ECOLOGICAL SYSTEMS

Jay Odenbaugh

1 INTRODUCTION

In this essay, I consider some of the ways in which nonlinear dynamics is changing the science of ecology. Specifically, I first consider a bit of history; namely, debates over the stability of populations and communities in population and community ecology. Here we explore arguments over population regulation in population ecology and the debate over diversity-complexity-stability in community ecology. This serves to highlight how ecological systems have been evaluated with the tools and assumptions of linear dynamical systems. Second, I turn to some conceptual issues. That is, what different concepts of *stability* are at work in ecology. I provide a taxonomy of these concepts and how they are related to one another. Unfortunately, many of these concepts are mostly applicable when thinking of linear systems. As an example of nonlinear dynamics in ecology, I consider the case of deterministic chaos. Using very simple discrete difference equations suited for insect populations, for example, we see the characteristic of sensitivity to initial conditions. Finally, I consider the impact of complex systems analysis on issues in and around ecology. Specifically, I examine the rise of "resilience thinking" and debates over ecological laws as examples of how nonlinear dynamics is challenging ecological theory and practice.

2 THE BALANCE OF NATURE?

One of the most common metaphors in environmental thought is that of a "balance of nature." Put simplistically, the metaphor suggests that ecological systems after a disturbance will return to some preferred or natural state by the actions of an intentional agent (i.e., God) or some non-intentional mechanism. Scientific creationism has largely and rightfully been rejected and thus there has been a continual search for causal mechanisms which produce this supposed balance. Moreover, as the science of ecology developed, the metaphor of a balance of nature has been transformed (or eliminated) using more technical notions of "ecological stability" which can be used in population, community, and ecosystem ecology. As examples, we think of bird populations as hovering around certain numbers, given their food resources, forest communities developing to some climax stand after, say, a fire, or nutrient inputs and outputs as balancing one another over

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

time in a ecosystem. So, informally, when a system returns to the salient state, we think of that state as *stable* and if it does not so return then it is *unstable*. This dichotomous way of thinking of course hides much complexity; nevertheless, these fundamental contrasts were of crucial significance in ecology's history. In the remaining part of this section, I want to sketch just how this was so in population and community ecology respectively. This will allow us to see how nonlinear dynamics is changing ecology.

One of the most fundamental conceptual and empirical issues in population ecology is whether populations of plants and animals are "regulated." As biologists noted for a long time, populations persist over enough time so as to require an explanation. From a statistical point of view, the mean values of population abundance have moderate amounts of variation around that mean over time.

Early in population ecology's history, there were two schools of thought over explaining (or sometimes rejecting) population regulation. The *biotic school* [Howard, Fiske, 1911; Nicholson, 1933; Smith, 1935] contended that populations are caused to vary moderately around their respective means due to a dependence on their own density or abundance. These density-dependent feedback loops will increase the population size when low and decrease it when it is high. Thus, average population densities remain relatively stationary. On the other hand, the *climatic school* [Bodenheimer, 1931; Andrewartha and Birch, 1954] claimed that populations are driven by changes in the abiotic environment by factors like weather and though they are causally independent of population size, they are bounded too. As these factors change over time, the abiotic environment can deny density-dependent factors the opportunity to take effect. However, what is this notion of "stability" at work in the population regulation debate and how can it be quantified?

Single population. In order to understand this notion, it is very useful to consider a remarkably important mathematical model, the logistic growth model. This model represents density-dependent population growth and regulation. To begin suppose we have a species with an abundance N and a per capita change rate of r, so that rN is the fractional change in N per unit time as a result of births and deaths. Then we can describe this population's growth over time with the following equation.

$$\frac{dN}{dt} = rN$$

Given that $N > 0$, the population's rate of change is determined by r and the direction of change is determined by whether r is positive (exponential growth), negative (exponential decline), or zero (stasis). Real world populations of course do not have an unending supply of resources and compete for the supplies they have. In short, they do not grow exponentially forever. To roughly capture this we can build a feedback term into our model giving us the famous logistic model:

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right)$$

Here K is the carrying capacity of the population. Suppose our population is not growing; i.e., $dN/dt = 0$. If we further suppose that each parameter, r, K, and our variable N, take non-zero values, then $dN/dt = 0$ just in case $N/K = 1$ or $N = K$. Is this equilibrium "stable"? Yes, because (i) at $dN/dt = 0, N$ will remain constant, (ii) if $N > K$, then $N/K > 1$ and $dN/dt < 0$ and will be so until $N = K$, while (iii) if $N < K$, then $N/K < 1$ and $dN/dt > 0$ and will be so until $N = K$.

This first-order differential equation can be solved analytically. If $N = N_0$ at $t = 0$, then:

$$n(t) = \frac{K}{1 + \left[\frac{K - N_0}{N_0}\right] e^{-rt}}$$

Thus, $N(t) \to N^* = K$ as $t \to \infty$. Put differently, any initial condition such that $N_0 > 0$ will eventually converge to the carrying capacity $N^* = K$. In contemporary terms, we say that N^* is an *asymptotically stable point attractor* for any initial condition $N_0 > 0$. This model and the way in which stability was conceptualized were incredibly important in the history of ecology. It expressed clearly the notion of a single population's or species' abundance being "balanced" through the mechanism of intraspecific competition for resources and would be the locus of initial debates over the value of quantitative models in ecology [Kingsland, 1995].

Community of populations. Ecologists are not only interested in populations of course, but in interacting populations or communities. That is, we are interested in what dynamical behaviors result when species prey on, compete with, and cooperatively depend on each other. Community ecology has been exploring conceptually and empirically what are called *diversity-complexity-stability hypotheses*.[1] In the 1950s, Charles Elton [2000] and Robert MacArthur [1955] each suggested diversity-complexity-stability hypotheses. Elton provided a handful of empirical arguments for such a hypothesis involving mathematical models, bottle experiments, flora and fauna on islands, insect outbreaks in monocultures and tropical communities, and invasions on islands. These arguments were not entirely well received (see [Goodman, 1975] for example). MacArthur on the other hand looked at the dynamics of food webs of a very simple sort and determined a measure of stability; namely, "[t]he amount of choice which the energy has in following the paths up through the food web is a measure of the stability of the community" [1955, p. 534]. He argued that in order to maintain the species in the community, there must be a large number of paths in the food web.

Theoretically, everything changed regarding the community diversity-complexity-stability hypothesis in the 1970s with the appearance of the ex-particle physicist

[1]Strictly speaking, a diversity-stability hypothesis has the following form: increasing (decreasing) species richness increases (decreases) stability. However, a complexity-stability hypothesis involves increasing variables other than simply the number of species.

Robert May. The classic [May, 1973] examined relationships between complexity and stability in the context of dynamical systems theory.[2] Using local stability analyses, he demonstrated that small perturbations to systems of differential equations describing species abundances around their equilibrium could take the community away from this equilibrium when complexity increased. Thus, increasing the complexity of a community decreases its "local asymptotic stability." Let us consider May's work in more detail.

May assumed there are m species whose abundances or densities can be represented by the following nonlinear first-order differential equations:

$$\frac{dN_i(t)}{dt} = F_i[N_1(t), N_2(t), \ldots, N_m(t)]$$

The equilibrium for a species i (where $i = 1, 2, \ldots, m$), denoted as N_i^*, occurs when $dN_i/dt = 0$, as above with the logistic model. Unlike the logistic model above however, models with many species have to take into account the effects of other species on a given perturbed population and vice versa, and typically do not have analytic solutions. Instead, May proceeded to evaluate their stability in "neighborhoods" around these equilibria by asking whether the community would return to its previous equilibrium after an extremely small perturbation.

To answer this question, he wrote an equation for the perturbation itself. Let $x_i(t)$ denote an arbitrarily small change to the community equilibrium N_i^* at time t, yielding a new abundance of population $N_i(t)$ at t, whence $x_i(t) = N_i(t) - N_i^*$. If we assume that the perturbation is infinitesimally small, we can linearly approximate the function that describes $x_i(t)$ through time. May derived a set of m linear first-order differential equations which describe the perturbation's dynamics.

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{m} a_{ji} x_j(t), \quad a_{ji} = \frac{\partial F_i}{\partial N_j} \mid_{N^*}$$

a_{ij} is the interaction coefficient between species i and j, representing the per capita effect of the species j on species i; that is, the effect on individuals of species i of adding an individual of species j:

- If i is a predator and j is prey, then $a_{ij} > 0$ and $a_{ji} < 0$.
- If i and j are competitors, then $a_{ij} < 0$ and $a_{ji} < 0$.
- If i and j are mutualists, then $a_{ij} > 0$ and $a_{ji} > 0$.
- If i and j do not interact, $a_{ij} = 0$.

Thus May abstractly considered the interactions amongst species in terms of just the signs of a_{ij}. Using linear algebra, we can rewrite the above equation as

[2]Ecologist Richard Levins [1968] recognized that community interactions could be described by what he called a "community matrix." One should also see [Gardner and Ashby, 1970] for a more general exploration of complexity and stability in complex systems.

$d\mathbf{x}/dt = \mathbf{Ax}$, where $\mathbf{x}$ is the $x \times 1$ column matrix formed by the x_i and $\mathbf{A}$ is the $m \times m$ "community matrix" formed by the $a_{ij} = \frac{\partial F_i}{\partial N_j} |_{N^*}$. Finally, he assumed that the resulting perturbations to the population numbers can also be represented by a linear sum:

$$n_j(t) = \sum_{j=1}^{m} \xi_j e^{\lambda_j t}$$

where $j = 1, 2, \ldots, m$. Here ξ_j are constant eigenvectors and λ_j are the eigenvalues of the Jacobian matrix $\mathbf{A}$. If there is a j such that the real part of $\lambda_j > 0$, then the difference between $N_i(t)$ and N_i^* increases with time and the perturbations are unstable; otherwise, they are stable.[3]

May assigned signs to the off-diagonal interaction coefficients $a_{ij}, i \neq j$, at random, and negative signs to the diagonal coefficients a_{ii} to express density dependence for each species. He defined the *connectance* C of a community to be the proportion of non-zero elements. Finally, the intensity s of the interspecific interactionwas a random variable with a mean of zero and a variance of s^2. May proved that a model community is almost certainly qualitatively stable just in case

$$s(cm)^{\frac{1}{2}} < 1$$

This had a surprising consequence. Suppose $P(m, s, C)$ is the probability that a model community is unstable. Then $P(m, s, C) \to 1$ as $m \to \infty$ and $P(m, s, C) \to 0$ as $m \to 0$. Therefore, all else being equal, and contrary to the drift of MacArthur's food web principle, an increase in the number of species m will lead to a decrease in the stability of a community.

May's theoretical analysis and model assumptions have been challenged. Donald DeAngelis [1975] argued that the results would change if donor-dependence was added to the models which he contended was more biologically realistic. If a species j is eaten by a species i, then donor-dependence occurs if $\partial f_{ij}/\partial N_j > \partial f_{ij}/\partial N_i$; in other words, the predator's dynamics are determined more by changes in the prey's density than in the predator's density. DeAngelis argued then that donor-dependence can generate stable communities if combined with other assumptions. Larry Lawlor argued that May's randomly constructed communities contain predators with no prey and prey without predators. Similarly, May's randomly constructed food webs have food web loops in which species i feeds on species j, j feeds on species k, and k feeds on i [Lawlor, 1978]. He insisted that all of these phenomena were biologically unrealistic.

In the end, ecologists such as C. S. "Buzz" Holling [1973], Gordon Orians [1975], and Stuart Pimm [1984] recognized that actually there are many different *diversity*, *complexity*, and *stability* concepts at issue, which in turn leads to many distinct hypotheses. Hence, May's result was by no means the final word (as he himself un-

[3]If the real part of $\lambda_j = 0$ then the perturbation is constant, otherwise $N_i(t)$ returns to N_i. If the imaginary part of λ_j is non-zero then $N_i(t)$ oscillates around N_i.

derstood).[4] Consider then how "stability" has been conceptualized by ecologists.

3 THE VARIOUS *STABILITIES* AND THEIR ANALYSIS

As theoretical ecology has matured, ecologists have arrived at a variety of stability concepts. One useful source for understanding these different concepts comes from the work of Stuart Pimm ([1992]). He distinguishes between definitions of *complexity*, *variables of interest* and *stability.* The complexity of a community or ecosystem can be characterized in terms of species richness, their connectance, and interaction strength, variants of May's parameters above. Species richness is simply the number of species in a community possibly coupled with their distributional evenness. The connectance of a community is the number of interspecific interactions out of those possible. Interaction strength is the average value of the interspecific interactions in the community. The variables of interest are usually species abundances or species biomass but can also include species taxonomic composition and trophic level abundance. Given these definitions we can now define several different stability concepts for a system responding to a perturbation.

- *Basic Stability:* a system is stable just in case all the variables return to their initial equilibrium values following a perturbation.
- *Resilience:* how fast the variables return to their equilibrium following a perturbation.
- *Persistence:* how long the value of a variable lasts before it changes to a new value.
- *Resistance:* the degree to which a variable is changed following a perturbation.
- *Variability:* the degree to which a variable varies over time, its volatility.

The concepts of *persistence*, *resistance*, and *variability* unlike *basic stability* and *resilience* are defined by Pimm without reference to an initial equilibrium state; however we could revise them in the following way:

- *Persistence**:how long the value of a variable lasts after a perturbation before it returns to its initial equilibrium value.

[4]For example, in recent work Tilman and colleagues (see [Kinzig, *et al.*, 2002]) have been conducting large-scale experiments at Cedar Creek Natural History Area in Minnesota. These experiments have shown that species richness is correlated with plant community stability. As the variability in plant community biomass decreases the number of species in the community increases and vice versa. There is of course debate as to whether increasing diversity causes increasing plant community stability.

- *Resistance**: the degree to which a variable is changed from its initial equilibrium value following a perturbation.
- *Variability**: the degree to which a variable varies from its initial equilibrium over time due to a perturbation.

Of course, these three revised definitions could continue to be revised for more than one variable and all of the definitions to include alternative stable states.

As ecologists have studied these properties they have used the analytic tools available to them. Fortunately for them, there has been a rich tradition of dynamical systems theory which has been used successfully in the physical sciences. As a result, many of these tools have been imported into ecology. One such tool we have seen is local stability analysis and more specifically the notion of Lyapunov stability.[5] A system is Lyapunov stable at a point $\mathbf{x}^*$ in its dynamical state space to a perturbation p (is at equilibrium at $\mathbf{x}^*$ relative to p) if its trajectory following p stays within some neighbourhood of $\mathbf{x}^*$ and is asymptotically stable there if it eventually remains at $\mathbf{x}^*$. The strength of this conception is that it can be applied to any kind of system and perturbations, irrespective of the dynamics involved; its weakness is that it does not specify how to go about identifying such equilibria in practice. The direct approach would be to simply examine the manifold of trajectories and detect the local attractors. But the differential equations that describe these systems are typically non-linear and coupled and so it is rare for there to be analytic solutions to them and testing for stability through computational construction of trajectories is typically resource intensive, uncertain in prospect and, where bifurcations are involved (including strange attractors), impossible. Instead an indirect method is generally used to determine whether a given equilibrium point of a system is Lyapunov stable, namely, linearise around the equilibrium point x^*, and extract the exponential growth parameters λ_j as discussed above.

Though local stability analysis provides a useful mathematical framework for thinking about stability, in ecology there are many limitations to such methods. First, identifying local stability through Lyapunov coefficients requires applying local linear approximation to departures from equilibrium and hence is restricted to infinitesimal perturbations. Ecological perturbations like a flood, wildfire, or species invasion are obviously not infinitesimally small. Hence, this technique can tell us very little about the dynamical behavior of ecological systems subject to these perturbations – even though species may become specifically and successfully adapted to their occurrence. Second, even where there are such small joslings of ecological systems, there is often no way to measure them; in practice we can usually only measure much larger, coarse-grained changes. Third, Lyapunov stability only applies to perturbations of the state variables. However, many ecological perturbations in fact change the parameters themselves such as the intrinsic rates of growth, carrying capacities and interaction coefficients. Hence, many of the relevant perturbations simply cannot be captured in this classical framework. Fourth,

[5] In this discussion of Lyapunov stability, I am indebted to James Justus' ([2008]) succinct explanation of these techniques and properties and their limitations.

nonlinear systems and their dynamics can exhibit dynamics more complex than those captured in the above framework. As mentioned, the concepts of *basic stability* and *resilience* presupposes that ecological systems have stable equilibria in the sense discussed and nonlinear systems need not.

4 NONLINEAR DYNAMICS IN ECOLOGY.

One of the first demonstrations of the possible applicability of nonlinear dynamics in ecology comes from the now familiar work of Robert May [1973; 1974; 1975; 1976; May and Oster, 1976]. Using very simple models, May was able to show that ecological systems could in principle produce chaotic behavior. Roughly put, a dynamical system exhibits chaos when it is extremely sensitive to initial conditions and is aperiodic.[6] A dynamical system exhibits *sensitive dependence on initial conditions* if, and only if, for arbitrarily close initial conditions the subsequent behaviors of the dynamical system diverge exponentially. (This is sometimes called the "butterfly effect.") A system is aperiodic if its trajectories never repeat their state sequences. However, chaotic orbits are bounded in state space, forming what are called strange dynamical attractors (see Figure 1). In consequence, chaotic systems are Lyapunov stable (stay within a neighbourhood) but not basically stable (do not return to their initial states).

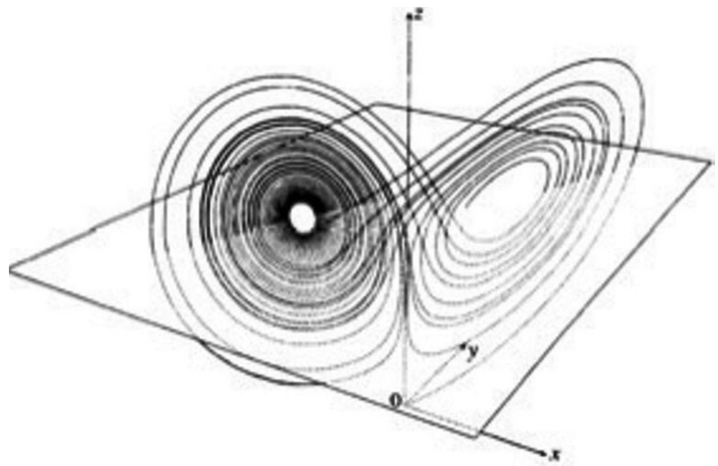


Figure 1. A strange attractor in a state space. (In this case the Lorenz attractor, discovered by Lorenz when studying dynamical models of weather and the first such attractor discovered.)

[6]See Kellert [1993] and Smith [1998] for philosophical discussions as to how *chaos* should be defined.

Let's now consider May's ecological model. Suppose there is a seasonally breeding population of insects whose generations do not overlap. Modelling their population dynamics requires finding an expression that relates the population at $t+ 1$ to that of $t, x_{t+1} = F(x_t)$. The simplest useful expression is a nonlinear difference equation called the "logistic map": $x_{t+1} = \mu x_{t+1}(1 - x_{t+1})$. In this model, x_t is normalized; it is scaled by dividing population size by the maximum value and so $0 < \mu < 1$. If $\mu < 1$, then the population will decrease to extinction; however, if $\mu > 1$ the dynamics of our model can become very complex. As the value of μ increases toward 3.5 it exhibits first basically stable behaviour as a point attractor, then, beginning at $\mu = 3.3$, periodic behaviour in limit-cycle attractors, with the number of available attractors doubling as μ increases, and finally chaotic behaviour in a strange attractor above approximately $\mu = 3.5$. Each transition between attractors is a bifurcation, that is, a change in the dynamical form of state space and not just in the dynamical state. Let us say $\mu = 2.5$ and let us suppose our initial population value $x_0 = 0.25$. With these values, the equilibrium population is $x^* = 0.6$ and it arrives at this point attractor independent of its initial population value. If $\mu = 3.3$, then our population arrives at a period two cycle. That is, $x_n = x_{n+2}$. If we set $\mu = 3.5$ there is a cycle of period four or $x_n = x_{n+4}$. Amazingly, as we increase $\mu > 3.5$, mathematical chaos sets in. For example, if $\mu = 3.7$, then there seems to be no order in the time series as the population oscillates in a seemingly random manner without any periodicity.

Determining in practice whether a given ecological system is chaotic is challenging and especially so in ecology [Hastings, *et al.*, 1987; Cushing *et al.*, 2003]. Attempting to predict the system trajectory to test for its characteristics breaks down in a strange attractor since even small differences may lead to widely divergent predictions, but system variables (let alone initial states) can only ever be determined to some finite accuracy. The alternative is to look at an entire time series of finitely-accurate measurements to see if some characteristic feature of the whole marking chaos can be determined, for example, a Lyapunov exponent.

An example arguing that chaos is present in real ecological populations is provided by by Constantino *et al.* [1997] who tested a model which exhibits chaos against the controlled dynamics of *Tribolium* (flour beetle) populations. They wanted to see if "... specific sequences of transitions among qualitatively different dynamical regimes occur in response to changing biological parameters" [1997, p. 389]. The model they employed involves three stochastic difference equations representing the dynamic interactions among the larval, pupal, and adult stage of an insect population.[7] With one unmanipulated control treatment, these ecologists

[7]The equations are $L_{t+1} b A_t \exp([(-)] c_{el} - c_{ea} A_t) + E_{1t}, P_{t+1} = L_t(1 - \mu_t)\exp(E_{2t})$, and $A_{t+1} = [P_t \exp[(-c_{pa} A_t) + A_t(1 = \mu_a)]\exp(E_{et})]$ where L_t is the number of feeding larvae at time $t; P_t$ is the number of large larvae, non-feeding larvae, pupae, and callow adults at time t and A_t is the number of sexually mature adults at time t. The parameter b is the number of larval recruits per adult per unit time in the absence of cannibalism. The parameters μ_l and μ_a are the larval and adult rates of mortality in one unit of time. The expressions $\exp[(-c_{el} L_t)]$ and $\exp[(-c_{ea} A_t)]$ are the probabilities that an egg is not eaten in the presence of L_t larvae and A_t adults in one time unit. The expression $\exp[(-c_{pa} A_t)]$ is the survival probability of a pupa

had twenty-four cultures of *Tribolium castaneum* initiated with 250 small larvae, 5 pupae, and 100 young adults where three populations were randomly assigned to each of the eight treatments. Constantino et al. manipulated the recruitment rate into adult stage at values predicted to produce a sequence of dynamical behaviors including a stable equilibrium, quasi-periodic, periodic cycles, and chaos. Every two weeks the L-, P-, and A-stages were tallied and dead adults were counted and removed with procedure continuing for eighty weeks.

The LPA model was fit to time series data using maximum likelihood parameter estimation. Using the model and estimates, they determined a bifurcation diagram and Liapunov exponents – see Figure 2. In these graphs, solid bands indicate quasi-periodicity ($\lambda = 0$) or chaos ($\lambda > 0$) and open areas containing lines indicate the result of periodic cycles ($\lambda < 0$).

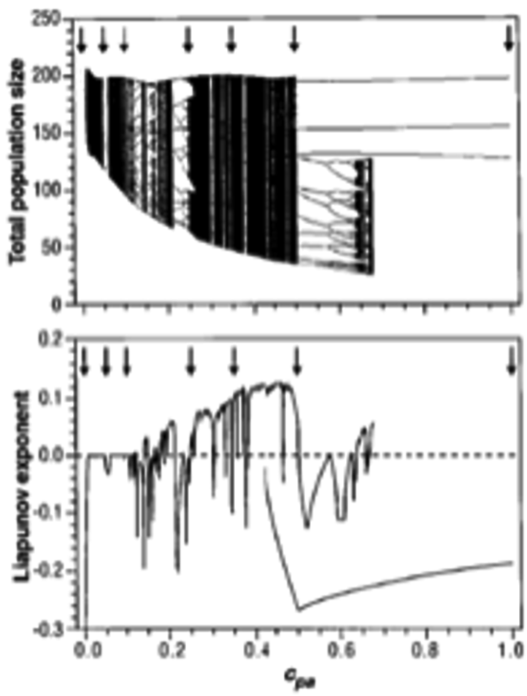


Figure 2. Bifurcation and Lyapunov exponent diagram. The recruitment rates actually studied and their respective predicted dynamics were (i) 0.00/stable equilibrium; (ii) 0.05/8-cycle attractor; (iii) 0.10/invariant loop; (iv) 0.25/chaotic attractor; (v) 0.35/chaotic attractor; (vi) 0.50/chaotic attractor and 3-cycle attractor; (vii) 1.00/3-cycle attractor.

in the presence of A_t adults in one time unit. Finally, the terms E_{1t}, E_{2t}, and E_{3t} are random noise variables which have a joint multivariate normal distribution with means of zero.

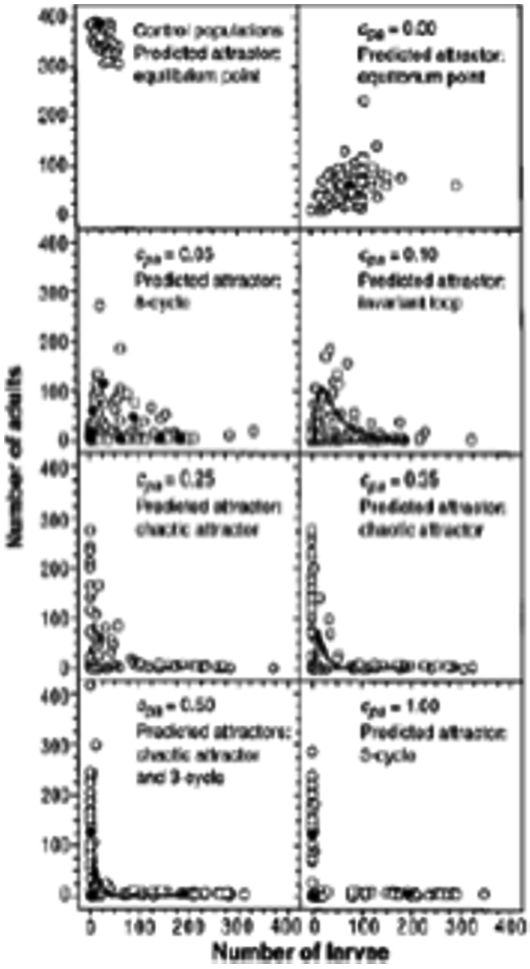


Figure 3. LPA model in comparison with data

Thus, the LPA model is highly confirmed and is one of the few relatively uncontroversial confirmations of chaos in a biological population.

> Ecologists Elisabeth Crone and Jane Molofsky have criticized these results.
>
> This [experiment] is one step beyond computer simulations of interactions, in that there is some possibility that flour beetles could never have been forced to mimic theoretical models. Nonetheless, in these experiments, it would be more surprising if the experimental populations did not display the predicted dynamics. [1999, p. 210]

Similarly, ecologist John Lawton writes,

> Critics argue, in contrast, that despite these (and other advantages), the artificial nature of laboratory system, the restricted nature of their inhabitants, the general absence of environmental disturbances, and the small scale mean that experiments carried out in them are at best a harmless game and at worst a waste of time and money. [1998, p. 178]

Even if a model can accurately represent the dynamics of such an artificial system, one might argue that this has next to no significance for the status of such models in non-laboratory systems because there humans typically cannot intervene to control all the relevant variables and parameters. This might be so, in which case we may not often expect to see chaotic dynamics outside the laboratory. However, note, first, that this would not show that they were not part of the in-principle dynamical repertoire of natural systems and, second, that we can expect to find relevantly similar conditions outside the laboratory whenever the dynamics is insensitive to the uncontrolled variables (whether generally or for certain naturally occurring values). In any event, laboratory experiments of this sort are the first step in testing the accuracy of chaotic ecological models.

There are two more powerful objections to the generality of chaotic behavior in ecological populations. First, the growth rates which generate the bifurcations needed to get chaos are unreasonably high given measurements of non-laboratory populations. Second, in the sorts of models and bifurcation diagrams we have considered, population abundances will be low in certain years and coupled with environmental fluctuations this would raise considerably the probability of extinction. Some have thus theorized that chaotic dynamics would be selected against by natural selection though the mechanism for such selection regimes is remains unclear [Solé and Bascompte, 2006, pp. 58–59].

5 THE IMPACT OF COMPLEX SYSTEMS ANALYSIS.

While the discussion of section 4 underlines the practical complexities involved in applying complex dynamical systems models to real world ecologies, the relevance of non-linear dynamics generally, e.g. of cycling or oscillatory behaviours, is not in doubt and the importance of these models to understanding ecological behaviour is only strengthened. From this perspective, it is clear that this work is of crucial importance to ecology. First, it advises ecologists to pay attention to nonlinear, unstable, and aperiodic behavior. Moreover, it has provided ecologists with a host of new tools that take seriously the characteristic features of nonlinear systems – a shift of especial note given that theoretical ecologists had hitherto assumed that ecological systems exhibit a balance of nature which can be studied with models utilizing only linear stability analysis. Using the tools of bifurcation diagrams, Lyapunov exponents, etc. we can represent and understand a whole host of dynamical behaviors that had hitherto been ignored. Further, this shift includes changing the very concept of *stability* used, and introducing new concepts, for example, of *resilience*, which has come to play an important role in ecological thought.

In this respect, very important conceptual progress in the context of nonlinear dynamics was provided by ecologist C. S. "Buzz" Holling [1973]. As we noted, an equilibrium is *basically stable* if the state of the system returns to that equilibrium after a perturbation. Following Stuart Pimm, *resilience* is the amount of time required for a perturbed system to return to its stable equilibrium. Holling recognized that for many nonlinear systems, these notions were insufficient to describe their dynamics. Suppose we have a system which possesses *alternative stable states*. For each of those stable states, there will be a basin of attraction such that for system states in that basin a system will be disposed to return to that equilibrium. The second sort of resilience he noted emphasizes behavior away from an equilibrium and concerns thresholds which can flip a system to another "regime" (that is, another dynamical form and hence change the attractor landscape itself). Holling termed the first sense of resilience "engineering resilience" and the second "ecological resilience." Ecological resilience is "measured by the magnitude of the disturbance that can be absorbed before the system redefines its structure by changing the variables and processes that control behaviour" [Gunderson, *et al.*, 1997]. For any system with a single stable attractor the measure of its ecological resilience is (roughly) the radius of the attractor basin, whereas its engineering resilience is measured (roughly) by its time to traverse the basin. For any system which has alternative stable states, the notion of engineering resilience will be insufficient to describe its dynamics far away from equilibrium and we will need a more complex conceptual repertoire including the notion of ecological resilience.

As an example of how the notion of ecological resilience works, consider the Spruce budworm (*Choristoneura fumiferana*) which is a serious pest in eastern Canada and northern Minnesota. The budworm crawls upon and consumes the leaf buds of coniferous trees. Excessive consumption can damage and kill the host. The budworms themselves are eaten primarily by birds, who eat many other insects as well. A key factor in determining the spruce budworm population is the leaf surface area per tree. Larger trees have larger leaf surface areas, resulting in larger spruce budworm populations. It had been observed that the spruce budworm population underwent irruptions approximately every 40 years. For unknown reasons the budworm population would explode, devastating forests, and then return to their previous manageable levels. In an effort to understand the cycles of budworm populations, several ecologists and mathematicians at the University of British Columbia including C. S. Holling studied the problem and produced a series of mathematical models [Ludwig, *et al.*, 1978].

Without going into the mathematical details, they found in their models that when the forest is young, the average leaf surface area is small and there is only one small positive equilibrium point, which is a sink or "refuge" (representing a small refuge population of budworms). When the forest grows, it passes a threshold and there are three positive equilibrium points, two "refuges" and a much larger "outbreak" one. Once the forest is past this threshold, the budworm population has a sudden increase in a short time. When an outbreak occurs, the budworm population is high and the forest's growth cannot keep up with the budworm. The budworm population size does decrease though not back to the refuge level, which is termed a "hysteresis effect." Eventually, the fir trees are defoliated by budworm, and the forest is taken over by birch trees. Here we see complex nonlinear dynamics including the crossing of thresholds into new regimes.

Second, work on nonlinear dynamics changes the nature of ecological debates. Consider the following passage,

That simple models can do complicated things is not a new idea: Poincaré expressed despair of ever completely understanding the motions of even trivial mechanical systems. Nevertheless, there is still a tendency on the part of most ecologists to interpret apparently erratic data as either stochastic "noise" or random experimental error. There is, however, a third alternative, namely, that wide classes of deterministic models can give rise to apparently chaotic dynamical behavior. [May and Oster, 1976, p. 573].

May is arguing that ecologists when seeing an exceptionally complex time series have been unjustified in assuming that it is the result of either stochasticity ("noise") or measurement error alone. Rather, it is clearly possible (and maybe actual) that such time series are the result of deterministic chaos. This challenged a common presupposition of the debates among population ecologists over population regulation. The "biotic" and "abiotic" schools differed vehemently, presupposing something like the following: if population densities are stationary, then the biotic school is correct; if population densities are wildly fluctuating, then the abiotic school is correct. May argues that this is fundamentally wrongheaded.

These studies of the Logistic Map revolutionized ecologists' understanding of the fluctuations of animal populations... With the insights of the Logistic Map, it was clear that the Nicholson-Birch controversy was misconceived. Both parties missed the point: population-density effects can, if sufficiently strong..., look identical to the effect of external disturbances. The problem is not to decide whether populations are regulated by density-dependent effects (and therefore steady) or whether they are governed by external noise (and therefore fluctuate). It's not a question of either/or. Rather, when ecologists observe a fluctuating population, they have to find out whether the fluctuations are caused by external environmental events (for example, erratic changes in temperature or rainfall), or by its own inherent chaotic dynamics, as expressed in the underlying deterministic equation that governs the population's development [May, 2002, p. 39-40]

Whether chaos is present in ecological populations, the possibility of its presence radically challenges traditional debates.

Lastly, nonlinear dynamics raises many interesting philosophical issues and let's consider just one. There is an ongoing controversy amongst biologists and philosophers of biology as to whether biology has laws. As a starting point, it is assumed that a law of nature is a universal generalization which supports counterfactuals. The positions staked out are these:

- There are biological laws so defined.

- There are no biological laws so defined.
- There are biological laws but the notion of lawhood must be redefined.

Let consider each of these views.

First, many biologists and philosophers of biology grant there are true biological generalizations which support counterfactuals [Brandon 1997; Colvyan and Ginsberg, 2006].[8] However, one worry is that these generalizations admit of exceptions. If biological generalizations admit of exceptions, then one might argue they are not laws since they are false generalizations. There are a variety of strategies for dealing with this problem (i.e., they are conjoined with ceteris paribus clauses, they are statistical, etc.); however, those who defend traditional laws in biology argue that the problem of exceptions exists with respect to laws in physics and chemistry as well. Thus, if they are not a problem in physics and chemistry, then they should not be a problem in biology [Colvyan and Ginsberg, 2006]. Still, these considerations have lead some philosophers of physics to be equally skeptical of physical laws [Cartwright, 1983; van Fraassen, 1989].

Second, some have argued that there are no biological laws since these generalizations are the product of evolutionary history and their truth values change over time [Beatty, 1992]. Consider a biological generalization of the form, "All members of species S have trait T." Claims of this sort fail to be laws since for any trait, evolutionary processes like selection, mutation, and drift can eliminate such character traits from the species. Still, even if there are no laws concerning specific taxa this does not entail there are no biological laws. There might very well be laws concerning ecological natural kinds such as decomposer, primary producer, and secondary producer, host and parasitoid, and predator and prey. These categories are defined by causal/functional roles and not historical properties.

Finally, some argue there are biological laws but they are not laws traditionally understood. Some philosophers argue that biological generalizations are mathematical truisms which can be known true a priori [Brandon, 1997; Sober, 1997]. Consider Ronald Fisher's three generation sex ratio model. Suppose there are N individuals in grand-offspring generation, the offspring generation contains m males and f females, there is random mating, and the parents differ in the number of sons and daughter they produce. The average son has N/m offspring and the average daughter will have N/f daughters. Thus, parents in the offspring generation who produce the minority sex will have more offspring on average. Therefore, the sex ratio evolves to a stable equilibrium of 1:1. Elliott Sober claims that this model represents a law, though a mathematical law which can be known independently of sense experience [2000, p. 72]. This account of biological laws faces an obvious worry: insofar as a law is mathematical and known a priori, it is unclear how it could be biological.

Nonlinear dynamics arguably provides new insights into the laws of biology debate. Least controversially, nonlinear dynamics provides a set of concepts which

[8]In ecology, here are a few: Liebeg's law of the minimum, the 3/2 thinning law, and the species area law.

allow scientists to classify the patterns or regularities that they discover. Consider again the logistic map. Suppose that the model accurately represents the dynamics of some insect population. As the growth rate increases, we have a stable equilibrium and thereafter there is period-doubling with 2-point, 4-point, 8-point, 16-point, 32-point cycles, and ultimately chaos. The notions of sensitive-dependence on initial conditions, bifurcations, chaos, path-dependence, hysteresis, etc. provide further tools for describing the dynamical patterns that ecologists and biologists more generally find. In fact, one might argue that the difficulty in finding robust ecological generalizations has been in part due to ignoring nonlinear dynamical behavior.

One might also argue that nonlinear dynamics allows biologists to describe law-like mechanisms that they heretofore have not noticed. The "period-doubling path to chaos" may be one such mechanism. However, Kellert is skeptical that chaos theory can provide such causal explanations. He offers several arguments for his view but there are two which are of special interest here [Kellert, 1993, pp. 77–118]. First, Kellert assumes that causal explanations are reductive — the whole is explained on the basis of the parts and their interactions. Second, in order to explain some phenomena by law, you must be able to derive the phenomena from the law in some suitable way. On his view, chaos theory is "holistic" and "experimental"; that is, it is not reductive and involves "inductive" computer simulations. However, neither of these reasons is particularly persuasive. First, philosophers of biology have long recognized that many explanations in the biological sciences involve supervenient properties or functional kinds as we saw above. Second, they have also long denied the "symmetry thesis" that every explanation of a phenomenon is simply a prediction in different clothes. Moreover, were Kellert's position accepted it would return us to an earlier problem: if what explains chaotic phenomena are "geometric mechanisms," then these appear to be mathematical in nature, or at least they are not uniquely biological but physical in some general sense. Lastly, this issue is part of a larger debate concerning whether complex system dynamics – for example, processes of self-organization — can generate biological forms or patterns independently of natural selection [Burian and Richardson, 1991; Kauffman, 1993].

6 CONCLUSION

In this essay, I first sketched some of the debates over the stability of ecological systems in population and community ecology. We found that ecologists have long assumed that populations and communities exhibit a "balance of nature." Second, we examined the work of theoretical ecologists who conceptualized this balance through various concepts of *stability* and *resilience* and utilized techniques like linear stability analysis. Third, we considered the work of pioneers like Robert May and C. S. Holling who recognized the very possibility of chaos and the importance of nonlinear dynamics more generally. Nonlinear dynamics has provided a suite of mathematical and conceptual tools for understanding many different sorts of dynamical behavior and challenged many of the assumptions of previous ecologists. Moreover, it raises interesting philosophical questions including ones concerning laws of nature in ecology and biology more generally.

Nonlinear dynamics raises other important questions both philosophical and practical which we have not been able to explore here. For example, given the difficulties in comparing models and data over long stretches of time how can nonlinear models best be confirmed? Likewise, chaotic models involve strange attractors and "infinite self-similarity" – these are certainly well-defined mathematical forms, but can these models be realistic representations of their target systems? Finally, ecological systems which are far away from an equilibrium can abruptly change from regime to another. How can we manage such systems? One plausible suggestion involves "adaptive management"; that is, feedbacks between monitoring and policy and where there is an explicit characterization of the uncertainties involved [Holling, 1978; Holling and Gunderson, 2002]. In the end, nonlinear dynamics challenges our understanding and ultimately our management of ecological systems.

BIBLIOGRAPHY

[Andrewartha and Birch, 1954] H. Andrewartha and C. Birch. *The Distribution of and Abundance of Animals*. University of Chicago Press, Chicago, 1954.

[Beatty, 1997] J. Beatty. Why Do Biologists Argue Like They Do? *Philosophy of Science*, 64, 231 – 242, 1997.

[Bodenheimer, 1928] F. S. Bodenheimer. Welche Faktoren regulieren die Individuenzahl einer Insekternart in der Natur? *Biologisches Zentralblatt*, 48, 714-739, 1928.

[Brandon, 1997] R. N. Brandon. Does biology have laws? The experimental evidence. *PSA 1996*, vol. 2, 444—457, 1997.

[Burian and Richardson, 1991] R. Burian and R. Richardson. Form and order in evolutionary biology: Stuart Kauffman's transformation of theoretical biology. *PSA 1990*, vol.2, 267—287, 1991.

[Cartwright, 1983] N. Cartwright. *How the Laws of Physics Lie*. Oxford University Press, Oxford, 1983.

[Colyvan and Ginsberg, 2006] M. Colyvan and L. Ginsberg. *Ecological Orbits*: *How Planets Move and Populations Grow*. Oxford University Press., Oxford, 2006.

[Costantino *et al.*, 1997] R. F. Costantino, R. A. Desharnais, J. M. Cushing, and B. Dennis. Chaotic dynamics in an insect population. *Science*, 275, 389-391, 1997.

[Crone and Molofsky, 1999] E. Crone and J. Molofsky. Message in a Bottle? Utility and limitations of
ecological bottle experiments. *Integrative Biology*, 1, 209-214, 1999.

[Cushing *et al.*, 2003] J. M. Cushing, B. Dennis, R. Constantino, R. Besharnais, and S. M. Henson. *Chaos in Ecology*: *Experimental Nonlinear Dynamics*. Academic Press, London, 2003.

[DeAngelis, 1975] D. L. DeAngelis. Stability and Connectance in Food Web Models. *Ecology*, 55, 238 – 243, 1975.

[Elton, 2000] C. Elton. *The Ecology of Invasions by Animals and Plants*. University of Chicago, Chicago, 2000.

[Gardner and Ashby, 1970] M. R. Gardner and W. R. Ashby. Connectance of Large, Dynamical (Cybernetic) Systems: Critical Values for Stability. *Nature*, 228, 784, 1970.

[Goodman, 1975] D. Goodman. The Theory of Diversity-Stability Relationships in Ecology. *Quarterly Review of Biology*, 50, 237–266, 1975.

[Gunderson *et al.*, 1997] L. H. Gunderson, C. S. Holling, L. Pritchard, and G. D. Peterson. *Resilience in Ecosystems, Institutions and Societies.* Beijer International Institute of Ecological Economics, Beijer Discussion Paper Series No. 95, 1997.

[Hastings, 1987] A. Hastings. Chaos in Ecology: Is Mother Nature a Strange Attractor? *Annual Review of Ecology and Systematics*, 24, 1-33, 1987.

[Holling, 1973] C. S. Holling. Resilience and Stability of Ecological Ecosystems. *Annual Review of Ecology and Systematics*, 4, 1-23, 1973.

[Holling, 1978] C. S. Holling. ed. *Adaptive Environmental Assessment and Management.* John Wiley & Sons, London, 1978.

[Holling and Gunderson, 2002] C. S. Holling and L. Gunderson, eds. *Panarchy: Understanding Transformations in Human and Natural Systems.* Island Press, Washington DC, 2002.

[Howard and Fiske, 1911] L. O. Howard and W. F. Fiske. The Importation into the United States of the Parasites of the Gypsy-Moth and the Brown-Tail Moth. *U.S. Department of Agriculture, Bureau of Entomology Bulletin*, 91, 1911.

[Justus, 2008] J. Justus. Complexity, Diversity, Stability. In S. Sarkar, A. Plutynski (eds.), *A Companion to the Philosophy of Biology*, Blackwell, Malden MA, 2008.

[Kauffman, 1993] S. Kauffman. *Origins of Order: Self-Organization and Selection in Evolution.* Cambridge University Press, New York, 1993.

[Kellert, 1993] S. Kellert. *In the Wake of Chaos.* University of Chicago Press, Chicago, 1993.

[Kingsland, 1995] S. Kingsland. *Modeling Nature*, 2^{nd} edn. University of Chicago Press, Chicago, 1995.

[Kinzig *et al.*, 2002] A. P. Kinzig, S. Pacala, and G. D. Tilman. *The Functional Consequences of Biodiversity: Empirical Progress and Theoretical Extensions.* Princeton University Press, Princeton, 2002.

[Lawlor, 1978] L, Lawlor. A Comment on Randomly Constructed Model Ecosystems. *American Naturalist*, 112, 445-47, 1978.

[Lawton, 1998] J. Lawton. Ecological Experiments with Model Systems: The Ecotron Facility in Context. In W. J. Resetarits, J. Bernardo (eds.), *Experimental Ecology: Issues and Perspectives*, Oxford University Press, Oxford, 1998.

[Levins, 1968] R. Levins. *Evolution in Changing Environments.* Princeton University Press, Princeton, 1968.

[Ludwig *et al.*, 1978] D. Ludwig, D. D. Jones, and C. S. Holling. Qualitative Analysis of Insect Outbreak Systems: The Spruce Budworm and Forest. *The Journal of Animal Ecology*, 47, 315–332, 1978.

[MacArthur, 1955] R. H. MacArthur. Fluctuations of Animal Populations and a Measure of Community Stability. *Ecology*, 36, 533–536, 1955.

[May, 1973] R. M. May. *The Stability and Complexity of Model Ecosystems.* Princeton University Press, Princeton, 1973.

[May, 1974] R. M. May. Biological Populations with Non-Overlapping Generations: Stable Points, Stable Cycles, and Chaos. *Science*, 186, 645-647, 1974.

[May, 1975] R. M. May. Biological Populations Obeying Difference: Stable Points, Stable Cycles, and Chaos. *Journal of Theoretical Biology*, 49, 511-524, 1975.

[May, 1976] R. M. May. Simple Models with Very Complicated Dynamics. *Nature*, 261, 459-467, 1976.

[May, 2002] R. M. May. The Best Possible Time to be Alive: The Logistic Map. In G. Farmelo (ed.), *It Must be Beautiful: Great Equations of Modern Science.* Granta Books, London, 2002.

[May and Oster, 1976] R. M. May and G. Oster. Bifurcations and Dynamic Complexity in Simple Ecological models. *American Naturalist*, 110, 573-599, 1976.

[Nicholson and Bailey, 1935] A. J. Nicholson and V. A. Bailey. "The Balance of Animal Populations, Part I," Proceedings of the Zoological Society, London, 3: 551-98, 1935.

[Orians, 1975] G. Orians. Diversity, Stability, and Maturity in Natural Ecosystems. In W. H. van Dobben, R. H. Lowe-McConnel, (eds.), *Unifying Concepts in Ecology.* W. Junk, The Hague, 139 – 50, 1975.

[Pimm, 1984] S. Pimm. The Complexity and Stability of Ecosystems. *Nature*, 307, 321 – 26, 1984.

[Pimm, 1992] S. Pimm. *The Balance of Nature.* University of Chicago Press, Chicago, 1992.

[Smith, 1935] H. S. Smith. The Role of Biotic Factors in the Determination of Population Densities. *Journal of Economic Entomology*, 28, 873–898, 1935.

[Smith, 1998] P. Smith. *Explaining Chaos.* Cambridge University Press, Cambridge, 1998.

[Sober, 1997] E. Sober. Two Outbreaks of Lawlessness in Recent Philosophy of Biology. *Philosophy of Science*, 64, 458–467, 1997.

[Sober, 2000] E. Sober. *Philosophy of Biology*. Westview Press, Boulder and Oxford, 2000.

[Sole and Bascompte, 2006] R. Sole, J. Bascompte. *Self-Organization in Complex Ecosystems*. Princeton University Press, Princeton, 2006.

[van Fraassen, 1989] B. van Fraassen. *Laws and Symmetry*. Oxford University Press, Oxford, 1989.

Part IV

Engineering

BEHAVIOR AND COGNITION AS A COMPLEX ADAPTIVE SYSTEM: INSIGHTS FROM ROBOTIC EXPERIMENTS

Stefano Nolfi

1 INTRODUCTION

Recent advances in different disciplines, ranging from cognitive sciences and robotics, biology and neurosciences, to social sciences and philosophy are clarifying that intelligence resides in the circular relationship between the brain of an individual organism, its body, and the environment (including the social environment).

In this chapter we will focus our attention on the evidence collected in robotic research with particular reference to results obtained in experiments in which the robots develop their skill autonomously while they interact with the external environment through an adaptive process. In particular we will demonstrate how the behavioural and cognitive skills developed by the robots can be properly characterized as complex adaptive systems which: (a) arise from the interactions between the brain of the robots, their body, and the environment and eventually between the dynamical process occurring within the robot and within the environment, (b) display a multi-level and a multi-scale organization in which the interaction between behavioural and cognitive properties at a certain level of organization lead to higher-level properties and in which higher-level properties later affect the lower-level properties.

The complex system nature of behaviour and cognition has important consequences both from an engineering and a modelling point of view. From the point of view of developing effective robotic artefacts it implies the need to rely on "design for emergence" techniques, i.e. techniques allowing the development of robots which are able to exploit useful emergent properties.

From the point of view of modelling biological systems, it implies the need to conceptualize behaviour and cognition as dynamical processes which unfold in time while the organism interacts with the environment.

Viewing behaviour and cognition as a complex adaptive system represents a new conceptual framework that can have profound consequences for cognitive science. In particular, as we will discuss in the concluding section, it might allow us to clarify better the notion and the role of embodiment and situatedness.

In section 2, we will demonstrate how behavioural and cognitive skills developed by robots which adapt to their task-environment through an adaptive process can

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

be properly characterised as a complex system. In section 3, we will discuss the relation between adaptation and the complex system nature of behaviour. Finally in section 4, we will discuss the implications of the complex adaptive systems of behavioural and cognitive skills.

2 BEHAVIOR AND COGNITION AS COMPLEX SYSTEMS

In agents which are embodied (i.e. have a physical body) and are situated (i.e. are located in a physical environment with which they interact) behavioural and cognitive skills are dynamical properties which unfold in time and which arise from the interaction between agents' nervous system, body, and the environment [Beer, 1995; Chiel and Beer, 1997; Keijzer, 2001; Nolfi and Floreano, 2000; Nolfi, 2005] and from the interaction between dynamical processes occurring within the agents' control system, the agents' body, and within the environment [Beer, 2003; Tani and Fukumura, 1997; Gigliotta and Nolfi, 2008]. Moreover, behavioural and cognitive skills typically display a multi-level and multi-scale organization involving bottom-up and top-down influences between entities at different levels of organization. These properties imply that behavioural and cognitive skills in embodied and situated agents can be properly characterized as complex systems [Nolfi, 2005].

These aspects and the complex system nature of behaviour and cognition will be illustrated in more detail in the next section with the help of examples involving robotic experiments.

2.1 Behavior and cognition as emergent dynamical properties

Behaviour and cognition are dynamical properties which unfold in time and which emerge from high-frequency non-linear interactions between the agent, its body, and the external environment [Chiel, Beer, 1997].

At any time step, the environmental and the agent/environmental relation co-determine the body and the motor reaction of the agent which, in turn, co-determines how the environment and/or the agent/environmental relation vary. Sequences of these interactions, occurring at a fast time rate, lead to a dynamical process — behaviour — which extends over significantly larger time spans than the interactions (Figure 1).

Since interactions between the agent's control system, the agent's body, and the external environment are nonlinear (i.e. small variations in sensory states might lead to significantly different motor actions) and dynamical (i.e. small variations in the action performed at time t might significantly impact later interactions at time t_{+x}) the relation between the rules that govern the interactions and the behavioural and cognitive skills originating from the interactions tend to be very indirect. Behavioural and cognitive skills thus emerge from the interactions between the three foundational elements and cannot be traced back to any of the three elements taken in isolation. Indeed, the behaviour displayed by an embodied and situated agent can hardly be predicted or inferred by an external observer even on the basis of a complete knowledge of the interacting elements and of the rules governing the interactions.

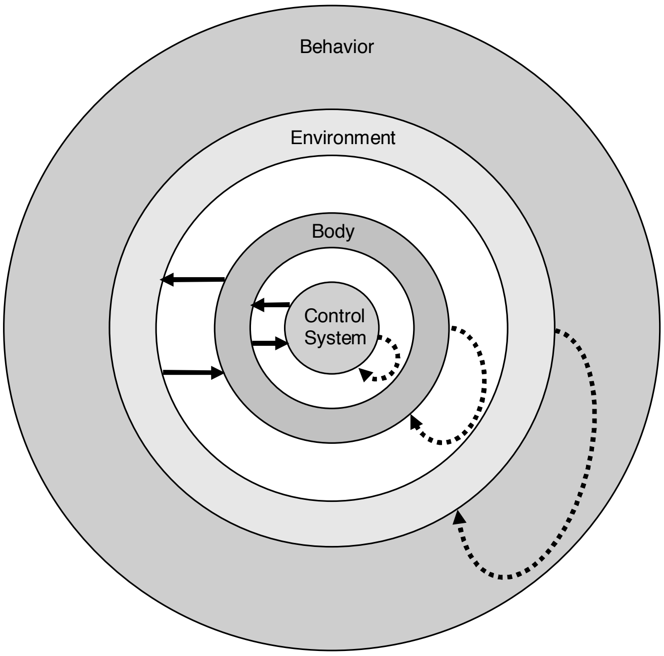


Figure 1. A schematic representation of the relation between agent's control system, agent's body, and the environment. The behavioural and cognitive skills displayed by the agent are the emergent result of the bi-directional interactions (represented with full arrows) between the three constituting elements – agents' control system, agent's body, and environment. The dotted arrows indicate that the three constituting elements might be dynamical systems on their own. In this case, agents' behavioural and cognitive skills result from the dynamics originating from the agent/body/environmental interactions but also from the combination and the interaction between dynamical processes occurring within the agents' body, within the agents' control system, and within the environment (see section 3)

A clear example of how behavioural skill might emerge from the interaction between the agents' body and the environment is constituted by the passive walking machines developed in simulation by McGeer [1990] — two-dimensional bipedal machines able to walk down a four-degree slope with no motors and no control system (Figure 2). The walking behaviour arises from the fact that the physical forces resulting from gravity and from the collision between the machine and the slope produce a movement of the robot and the fact that the robot's movements produce a variation of the agent-environmental relation which in turn produces a modification of the physical forces to which the machine will be subjected in the next time step. The sequence of bi-directional effects between the robot's body and the environment can lead to a stable dynamical process — the walking behaviour.

The type of behaviour which arises from the robot/environmental interaction depends on the characteristics of the environment, the physics laws which regulate the interaction between the body and the environment, and the characteristics of the body. The first two factors can be considered as fixed but the third factor, the body structure, can be adapted to achieve a given function. Indeed, in the case of this biped robot, the author carefully selected the leg length, the leg mass, and the foot size to obtain the desired walking behaviour. In more general terms, this example shows how the role of regulating the interaction between the robot and the environment in the appropriate way can be played not only by the control system but also by the body itself, providing that the characteristics of the body have been shaped to favour the exhibition of the desired behaviour. This property, i.e. the ability of the body to control its interaction with the environment, has been named 'morphological computation' [Pfeifer *et al.*, 2006]. For related work which demonstrates how effective walking machines can be obtained by integrating passive walking techniques with simple control mechanisms, see [Bongard and Paul, 2001; Endo, 2002; Vaughan *et al.*, 2004]. For related work which shows the role of elastic material and elastic actuators for morphological computing see [Schmitz *et al.*, 2007; Massera *et al.*, 2007].

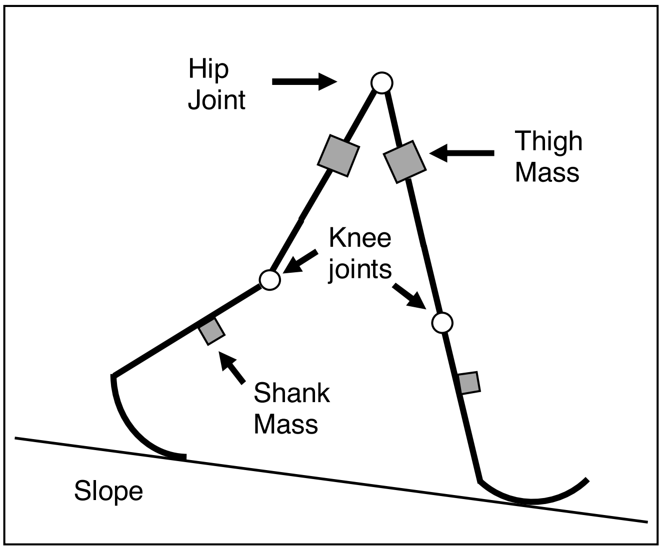


Figure 2. A schematization of the passive walking machine developed by McGear [1990]. The machine includes two passive knee joints and a passive hip joint.

To illustrate how behavioural and cognitive skills might emerge from agent's body, agent's control system, and environmental interactions we describe a simple experiment in which a small wheeled robot situated in an arena surrounded by walls has been evolved to find and to remain close to a cylindrical object [Nolfi, 2002]. The Khepera robot [Mondada *et al.*, 1993] is provided with eight infrared sensors and two motors controlling the two corresponding wheels (Figure 3).

From the point of view of an external observer, solving this problem requires robots able to: (a) explore the environment until an obstacle is detected, (b) discriminate whether the obstacle detected is a wall or a cylindrical object, and (c) approach or avoid objects depending on the object type. Some of these behaviours (e.g. the wall-avoidance behaviour) can be obtained through simple control mechanisms but others require non trivial control mechanisms. Indeed, a detailed analysis of the sensory patterns experienced by the robot indicated that the task of discriminating the two objects is far from trivial since the two classes of sensory patterns experienced by robots close to a wall and close to cylindrical objects largely overlap.

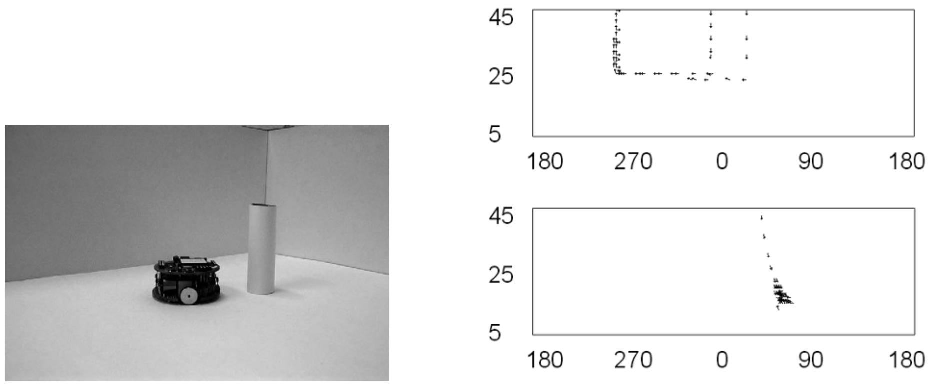


Figure 3. **Left**: The agent situated in the environment. The agent is a Khepera robot [Mondada *et al.*, 1993]. The environment consists of an arena of 60×35 cm containing a cylindrical object placed in a randomly selected location. **Right:** Angular trajectories of an evolved robot close to a wall (top graph) and to a cylinder (bottom graph). The picture was obtained by placing the robot at a random position in the environment, leaving it free to move for 500 time steps each lasting 100ms, and recording its relative movements with respect to the two types of objects for distances smaller than 45 mm. The x-axis and the y-axis indicate the relative angle (in degrees) and distance (in mm) between the robot and the corresponding object. For sake of clarity, arrows are used to indicate the relative direction, but not the amplitude of movements.

The attempt to solve this problem through an evolutionary adaptive method (see section 4) in which the free parameters (i.e. the parameters which regulate the fine-grained interaction between the robot and the environment) are varied randomly and in which variations are retained or discarded on the basis on an evaluation of the overall ability of the robot (i.e. on the basis of the time spent by the robot close to the cylindrical object) demonstrated how adaptive robots can find solutions which are robust and parsimonious in terms of control mechanisms [Nolfi, 2002]. Indeed, in all replications of these experiments, evolved robots solve the problem by moving forward, by avoiding walls, and by oscillating back and fourth and left and right close to cylindrical objects (Figure 3, right). All these behaviours result from sequences of interactions between the robot and the environment mediated by four types of simple control rules which consist in: turning left when the right infrared sensors are activated, turning right when the left infrared sensors are activated, moving back when the frontal infrared sensors are activated, and moving forward when the frontal infrared sensors are not activated.

To understand how these simple control rules can produce the required behaviours and the required arbitration between behaviours we should consider that the same motor responses produce different effects on different agent/environmental situations. For example, the execution of a left-turning action close to a cylindrical object and the subsequent modification of the robot/object relative position produce a new sensory state that triggers a right-turning action. Then, the execution of the latter action and the subsequent modification of the robot/object relative position produce a new sensory state that triggers a left-turning action. The combination and the alternation of these left and right-turning actions over time produce an attractor in the agent/environmental dynamics (Figure 3, right, bottom graph) that allows the robot to remain close to the cylindrical object. On the other hand the execution of a left-turning behaviour close to a wall object and the subsequent modification of the robot/wall position produce a new sensory state that triggers the reiteration of the same motor action. The execution of a sequence of left-turning actions then leads to the avoidance of the object and to a modification of the robot/environmental relation that finally leads to a sensory state that triggers a move-forward behaviour (Figure 4, right, top graph).

Before concluding the description of this experiment, it is important to notice that, although the rough classification of the robot motor responses into four different types of actions is useful to describe the strategy with which these robots solve the problem qualitatively, the quantitative aspects which characterize the robot motor reactions (e.g. how sharply a robot turns given a certain pattern of activation of the infrared sensors) are crucial for determining whether the robot will be able to solve the problem or not. Indeed, small differences in the robot's motor response tend to accumulate in time and might prevent the robot from producing successful behaviour (e.g. might prevent the robot producing a behavioural attractor close to cylindrical objects).

This experiment clearly exemplifies some important aspects which characterize all adaptive behavioural system, i.e. systems which are embodied and situated and which have been designed or adapted so to exploit the properties that emerge from the interaction between their control system, their body, and the external environment. In particular, it demonstrates how required behavioural and cognitive skills (i.e. object categorization skills) might emerge from the fine-grained interaction between the robot's control system, body, and the external environment without the need for dedicated control mechanisms. Moreover, it demonstrates how the relation between the control rules which mediate the interaction between the robot body and the environment and the behavioural skills exhibited by the agents are rather indirect. This means, for example, that an external human observer can hardly predict the behaviours that will be produced by the robot, before observing the robot interacting with the environment, even on the basis of a complete description of the characteristics of the body, of the control rules, and of the environment.

2.2 Behaviour and cognition as phenomena originating from the interaction between coupled dynamical processes

Up to this point we restricted our analysis to the dynamics originating from the agent's control system, agents' body, and environmental interactions. However, the body of an agent, its control system, and the environment might have their own dynamics (dotted arrows in Figure 1). For the sake of clarity, we will refer to the dynamical processes occurring within the agent control system, within the agent body, or within the environment as *internal dynamics* and to the dynamics originating from the agent/body/environmental interaction as *external dynamics*. In cases in which agents' body, agents' control system, or the environment have their own dynamics, behaviour should be characterized as a property emerging from the combination of several coupled dynamical processes.

The existence of several concurrent dynamical processes represents an important opportunity for the possibility to exploit emergent features. Indeed, behavioural and cognitive skills might emerge not only from the external dynamics, as we showed in the previous section, but also from the internal dynamical processes or from the interaction between different dynamical processes.

As an example which illustrates how complex cognitive skills can emerge from the interaction between a simple agent/body/environmental dynamic and a simple agent's internal dynamic consider the case of a wheeled robot placed in a maze environment (Figure 4) which has been trained to: (a) produce a wall-following behaviour which allows the robot to periodically visit and re-visit all environmental areas, (b) identify a target object constituted by a black disk placed in a randomly selected position in the environment for a limited time duration, and (c) recognize the location in which the target object was previously found every time the robot re-visits the corresponding location [Gigliotta, Nolfi, 2008].

The robot has infrared sensors (which provide information about nearby obstacles), light sensors (which provide information about the light gradient generated by the light bulb placed in the central corridor), ground sensors (which detect the colour of the ground), two motors (which control the desired speed of the two corresponding wheels), and one additional output unit which should be turned on

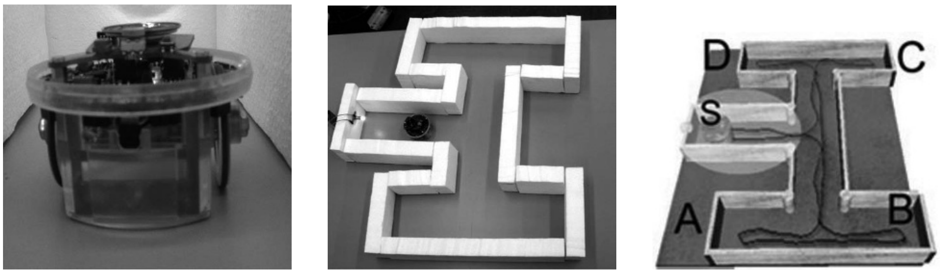


Figure 4. **Left**: The e-puck robot developed at EPFL, Switzerland (http://www.e-puck.org/). **Centre**: The environment which has a size of 52cm by 60cm. The light produced by the light bulb located on the left side of the central corridor cannot be perceived from the other two corridors. **Right**: The motor trajectory produced by the robot during a complete lap of the environment.

when the robot re-visit the environmental area in which the black disk was previously found. The robot's controller consists of a three-layer neural network which includes a layer of sensory neurons (which encode the state of the corresponding sensors), a layer of motor neurons which encode the state of the actuators, and a layer of internal neurons which consist of leaky integrators operating at tuneable time scale [Beer, 1995; Gigliotta, Nolfi, 2008]. The free parameters of the robot's neural controllers (i.e. the connection weights, and the time constant of the internal neurons which regulate the time rate at which these neurons change their state over time) were adapted through an evolutionary technique [Nolfi, Floreano, 2000].

By analysing the evolved robot the authors observed how it is able to generate a spatial representation of the environment and of its location in the environment while it is situated in the environment itself. Indeed, while the robot travel by performing different laps of the environment (see Figure 4, right), the states of the two internal neurons converge on a periodic limit cycle dynamic in which different states correspond to different locations of the robot in the environment (Figure 5).

As we mentioned above, the ability to generate this form of representation, that allows the robot to solve its adaptive problem, originates from the coupling between a simple robot's internal dynamics and a simple robot/body/environmental dynamics. The former dynamics is characterized by the fact that the state of the two internal neurons tends to move slowly toward different fixed point attractors, in the robot's internal dynamics, which correspond to different types of sensory states exemplified in Figure 5. The latter dynamics originate from the fact that different types of sensory states last for different time durations and alternate with a given order while the robot moves in the environment. The interaction between these two dynamical processes leads to a transient dynamics of an agent's internal state that moves slowly toward the current fixed point attractor without never fully reaching it (thus preserving information about previously experienced sensory states, the time duration of these states, and the order with which they have been experienced). The coupling between the two dynamical processes originates from the fact that the free parameters which regulate the agent/environmental dynamics (e.g. the trajectory and the speed with which the robot moves in the environment) and the agent internal dynamics (e.g. the direction and the speed with which the internal neurons change their state) have been co-adapted and co-shaped during the adaptive process.

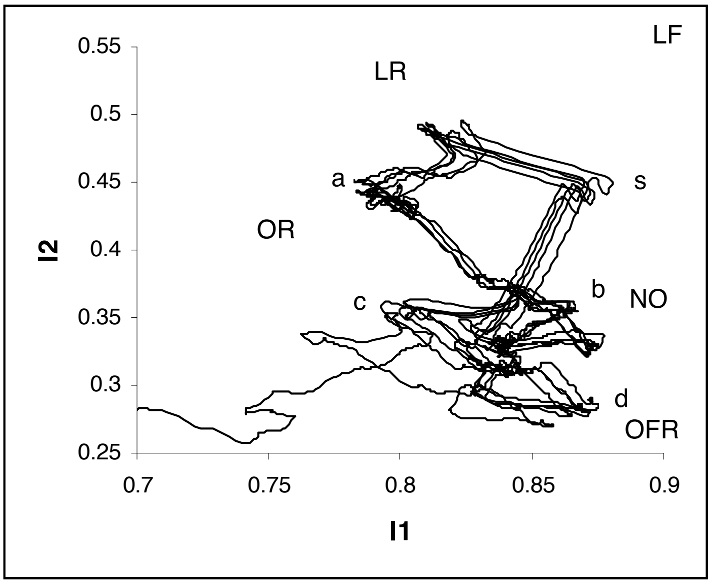


Figure 5. The state of the two internal neurons (i1 and i2) of the robot recorded for 330s while the robot performs about 5 laps of the environment. The *s*, *a*, *b*, *c*, and *d* labels indicate the internal states corresponding to five different positions of the robot in the environment shown in Figure 4. The other labels indicate the position of the fixed point attractors in the robot's internal dynamics corresponding to five types of sensory states experienced by the robot when it detects: a light in its frontal side (LF), a light on its rear side (LR), an obstacle on its right and frontal side (OFR), an obstacle on its right side (OR), no obstacles and no lights (NO).

For related works which show how navigation and localization skills might emerge from the coupling between agent's internal and external dynamics, see [Tani, Fukumura, 1997]. For other works addressing other behavioural/cognitive capabilities see Beer [2003] for what concerns categorization, [Goldenberg *et al.*, 2004; Slocum *et al.*, 2000] for what concerns selective attention, Sugita and Tani [2005] for what concern language and compositionality.

2.3 Behaviour and cognition as phenomena with a multi-level and multi-scale organization

Another fundamental feature that characterizes behaviour is the fact that it is a multi-layer system with different levels of organizations extending at different time scales [Keijzer, 2001; Nolfi, 2005]. More precisely, as exemplified in Figure 6, the behaviour of an agent or of a group of agents involve both lower- and higher- level behaviours that extend for shorter or longer time spans, respectively. Lower-level behaviours arise from few agent/environmental interactions and short term internal dynamical processes. Higher-level behaviours, instead, arise from the combination and interaction of lower-level behaviours and/or from long term internal dynamical processes.

The multi-level and multi-scale organization of agents' behaviour plays important roles: it is one of the factors which allow agents to produce functionally useful behaviour without necessarily developing dedicated control mechanisms [Brooks, 1991; Nolfi, 2005], it might favour the development of new behavioural and/or cognitive skills thanks to the recruitment of pre-existing capabilities [Marocco, Nolfi, 2007], it allow agents to generalize their skills in new task/environmental conditions [Nolfi, 2005].

An exemplification of how the multi-level and multi-scale organization of behaviour allow agents to generalize their skill in new environmental conditions is represented by the experiments carried our by Baldassarre *et al.* [2006] in which the authors evolved the control system of a group of robots assembled into a linear structure (Figure 7) for the ability to move in a coordinated manner and for the ability to display a coordinated light approaching behaviour.

Each robot [Mondada *et al.*, 2004] consists of a mobile base (chassis) and a main body (turret) that can rotate with respect to the chassis around a vertical axis. The chassis has two drive mechanisms that control the two corresponding tracks and teethed wheels. The turret has one gripper, which allows robots to assemble together and to grasp objects, and a motor controlling the rotation of the turret with respect to the chassis. Robots are provided with a traction sensor, placed at the turret-chassis junction, that detects the intensity and the direction of the force that the turret exerts on the chassis (along the plane orthogonal to the vertical axis) and light sensors. Given that the orientations of individual robots might vary and given that the target light might be out of sight, robots need to coordinate to choose a common direction of movement and to change their direction as soon as one or few robots start to detect a light gradient.

Evolved individuals show the ability to negotiate a common direction of movement and by approaching light targets as soon as a light gradient is detected. By testing evolved robots in different conditions the authors observed that they are able to generalize their skills in new conditions and also to spontaneously produce new behaviours which have not been rewarded during the evolutionary process. More precisely, groups of assembled robots display a capacity to generalize their skills with respect to the number of robots which are assembled together and to the

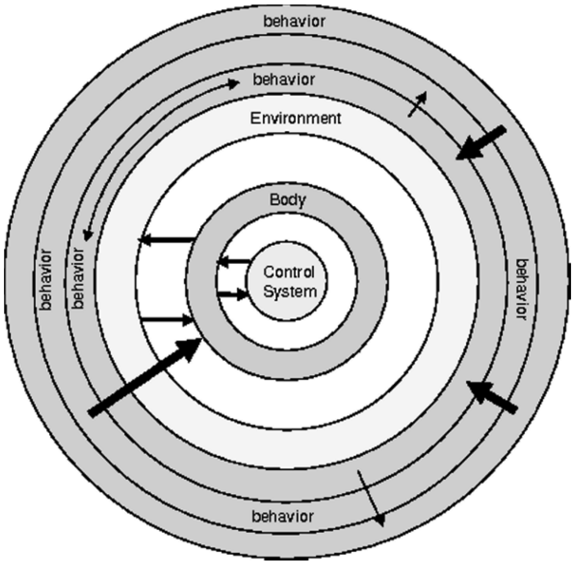


Figure 6. A schematic representation of multi-level and multi-scale organization of behaviour. The behaviours represented in the inner circles represent elementary behaviours which arise from fine-grained interactions between the control system, the body, and the environment, and which extend over limited time spans. The behaviours represented in the external circles represent higher-level behaviours which arise from the combination and interaction between lower-level behaviours and which extend over longer time spans. The arrows which go from higher-level behaviour toward lower levels indicate the fact that the behaviours currently exhibited by the agents later affect the lower-level behaviours and/or the fine-grained interaction between the constituting elements (agent's control system, agent's body, and the environment).

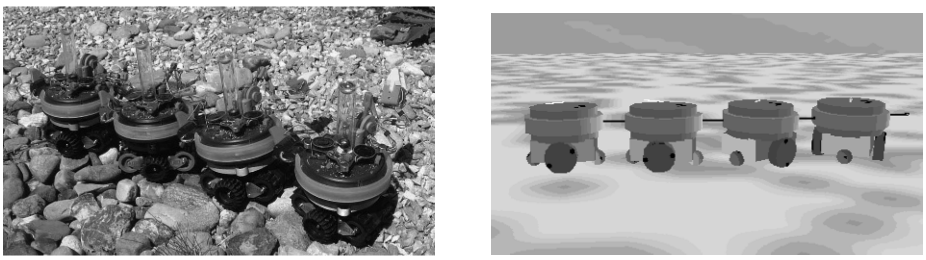


Figure 7. **Left:** Four robots assembled into a linear structure. **Right**: A simulation of the robots shown in the left part of the figure.

shape formed by the assembled robots. Moreover, when the evolved controllers are embodied in eight robots assembled so as to form a circular structure and situated in the maze environment shown in Figure 8, the robots display an ability to collectively avoid obstacles, to rearrange their shape so to pass through narrow passages, and to explore the environment. The ability to display all these behavioural skills allow the robots to reach the light target even in large maze environments, i.e. even in environmental conditions which are rather different from the conditions that they experienced during the training process (Figure 8).

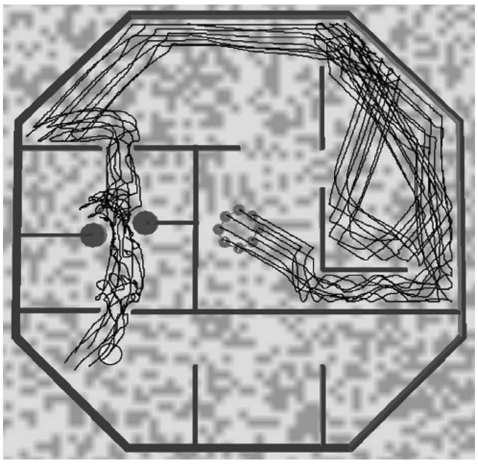


Figure 8. The behaviour produced by eight robots assembled into a circular structure in a maze environment including walls and cylindrical objects (represented with grey lines and circles). The robots start in the central portion of the maze and reach the light target located in the bottom-left side of the environment (represented with an empty circle) by exhibiting a combination of coordinated-movement behaviours, collective obstacle-avoidance, and collective light-approaching behaviours. The irregular lines, that represent the trajectories of the individual robots, show how the shape of the assembled robots changes during motion by adapting to the local structure of the environment.

By analysing the behaviour displayed by the evolved robots tested in the maze environment, a complex multi-level organization can be observed. The simpler behaviours that can be identified consist of low level individual behaviours which extend over short time spans:

1. A *move-forward behaviour* which consists of the individuals' ability to move forward when the robot is coordinated with the rest of the team, is oriented toward the direction of the light gradient (if any), and does not collide with obstacles. This behaviour results from the combination of: (a) a control rule which produces a move forward action when the perceived traction has a low

intensity and when difference between the intensity of the light perceived on the left and the right side of the robot is low, and (b) the sensory effects of the execution of the move forward action selected mediated by the external environment (which does not produce a variation of the state of the sensors until the conditions that should be satisfied to produce this behaviour hold).

2. A *conformistic behaviour* which consists of the individuals' ability to conform its orientation with that of the rest of the team when the two orientations differ significantly. This behaviour results from the combination of: (a) a control rule that makes the robot turns toward the direction of the traction when its intensity is significant, and (b) the sensory effects produced by the execution of this action mediated by the external environment, that lead to a progressive reduction of the intensity of the traction until the orientation of the robot conforms with the orientation of the rest of the group.

3. A *phototaxis behaviour* which consists of the individuals' ability to orient toward the direction of the light target. This behaviour results from the combination of: (a) a control rule that makes the robot turns toward the direction in which the intensity of the light gradient is higher, and (b) the sensory effects produced by the execution of this action mediated by the external environment, that lead to a progressive reduction of the difference in the light intensity detected on the two sides of the robot until the orientation of the robot conforms with the direction of the light gradient.

4. An *obstacle-avoidance behaviour* which consists of the individuals' ability to change direction of motion when the execution of a motor action produced a collision with an obstacle. This behaviour results from the combination of: (a) the same control rule that led to behaviour #2 and that makes the robot turn toward the direction of the perceived traction (which in this case is caused by the collision with the obstacle, while in the case of behaviour #2 it is caused by the forces exerted by the other assembled robots), and (b) the sensory effects produced by the execution of the turning action mediated by the external environment, that make the robot turn until collisions no longer prevent the execution of a moving forward behaviour.

The combination and the interaction between these four behaviours produce the following higher-level collective behaviours that extend over a longer time span:

5. A *coordinated-motion behaviour* that consists in the ability of the robots to negotiate a common direction of movement and to keep moving along such direction by compensating further misalignments originating during motion. This behaviour emerges from the combination and the interaction of the conformistic behaviour (which plays the main role when robots are misaligned) and the move-forward behaviour (which plays the main role when robots are aligned).

6. A *coordinated-light-approaching behaviour* which consists in the ability of the robots to co-ordinately move toward a light target. This behaviour emerges from the combination of the conformistic, the move-forward, and the phototaxis behaviours (which is triggered when the robots detect a light gradient). The relative importance of the three control rules which lead to the three corresponding behaviours depends both on the strength of the corresponding triggering condition (i.e. the extent of lack of traction forces, the intensity of traction forces, and the intensity of the light gradient, respectively) and on a priority relation among behaviours (i.e. the fact that the conformistic behaviour tends to play a stronger role than the phototaxis behaviour).

7. A *coordinated-obstacle-avoidance behaviour* which consists in the ability of the robots to co-ordinately turn to avoid nearby obstacles. This behaviour arises as the result of the combination of the obstacle avoidance, the conformistic and the move-forward behaviours.

The combination and the interaction between these behaviours leads in turn to the following higher-level collective behaviours that extend over still longer time spans:

8. A *collective-exploration-behaviour* that consists in the ability of the robots to visit different areas of the environment when the light target cannot be detected. This behaviour emerges from the combination of the coordinated-motion behaviour and the coordinated-obstacle-avoidance behaviour that ensures that the assembled robots can move in the environment without getting stuck and without entering into limit cycle trajectories.

9. A *shape-re-arrangement behaviour* which consists in the ability of the assembled robots to dynamically adapt their shape to the current structure of the environment so to pass through narrow passages especially when the passages to be negotiated are in the direction of the light gradient. This behaviour emerges from the combination and the interaction between coordinated motion and coordinated-light-approaching behaviours mediated by the effects produced by relative differences in motion between robots resulting from the execution of different motor actions and/or from differences in the collisions. The fact that the shape of the assembled robots adapt to the current environmental structure so as to facilitate the overcoming of narrow passages can be explained by considering that collisions produce a modification of the shape that affects the relative positions of the colliding robots, in particular with respect to the axis of the narrow passage.

The combination and the interaction of all these behaviours leads to a still higher-level behaviour:

10. A *collective-navigation-behaviour* which consists in the ability of the assembled robots to navigate toward the light target by producing coordinated

movements, exploring the environment, passing through narrow passages, and producing a coordinated-light-approaching behaviour (Figure 8).

This analysis illustrates two important mechanisms that explain the remarkable generalization abilities of these robots. The first mechanism consists in the fact that the control rules that regulate the interaction between the agents and the environment so as to produce certain behavioural skills in certain environmental conditions, will produce different but related behavioural skills in other environmental conditions. In particular, the control rules that generate the behaviours #5 and #6, for which evolving robots have been evolved in an environment without obstacles, also produce behaviour #7 in an environment with obstacles. The second mechanism consists in the fact that the development of certain behaviours at a given level of organization that extend for a given time span will, through their combination and interaction, automatically lead to the exhibition of related higher-level behaviours extending over longer time spans (even if these higher-level behaviours have not been rewarded during the adaptation process). In particular, the combination and the interaction of behaviours #5, #6 and #7 (that have been rewarded during the evolutionary process or that arise from the same control rules that lead to the generation of rewarded behaviours) automatically lead to the production of behaviours #8, #9, and #10 (that have not been rewarded). Obviously, there is no warranty that the new behaviours obtained as a result of these generalization processes will play useful functions. However, the fact that these behaviours are related to the other functional behavioural skills implies that the probabilities that these new behaviours will play useful functions are significant.

In principle, these generalization mechanisms can also be exploited by agents during their adaptive process to generate behavioural skills which play new functionalities and which emerge from the combination and the interaction between pre-existing behavioural skills playing different functions.

2.4 On the top-down effect from higher to lower levels of organization

In the previous sections we have discussed how the interactions between the agents' body, the agents' control system, and the environment lead to behavioural and cognitive skills and how such skills have a multi-level and multi-scale organization in which the interaction between lower-level skills leads to the emergence of higher-level skills. However, higher-level skills also affect lower-level skills up to fine-grained interactions between the constituting elements (agents' body, agents' control system, and environment). More precisely, the behaviours that originate from the interaction between the agent and the environment and from the interaction between lower-level behaviours, later affect the lower-level behaviours and the interaction from which they originate. These bottom-up and top-down influences between different levels of organization can lead to circular causality [Kelso, 1995] where high-level processes act as independent entities that constraint the lower-level processes from which they originate.

One of the most important effects of this top-down influence consists in the fact that the behaviour exhibited by an agent constrains the type of sensory patterns that the agent will experience later on (i.e. constrains the fine-grained agent/environmental interactions that determine the behaviour that will be later exhibited by the agent). Since the complexity of the problem faced by an agent depends on the sensory information experienced by the agent itself, this top-down influence can be exploited in order to turn hard problems into simple ones.

One neat demonstration of this type of phenomena is given by the experiments conducted by Marocco and Nolfi [2002] in which a simulated finger robot with six degree of freedom provided with sensors of its joint positions and with rough touch sensors is asked to discriminate between cubic and spherical objects varying in size. The problem is not trivial since, in general terms, the sensory patterns experienced by the robot do not provide clear regularities for discriminating between the two types of objects. However, the type of sensory states which are experienced by the agent also depend on the behaviour previously exhibited by the agent itself — agents exhibiting different behaviour might face simpler or harder problems. By evolving the robots in simulation for the ability to solve this problem and by analyzing the complexity of the problem faced by robots of successive generations, the authors observed that the evolved robots manage to solve their adaptive problem on the basis of simple control rules which allow the robot to approach the object and to move following the surface of the object from left to right, independently of the object shape. The exhibition of this behaviour in interaction with objects characterized by a smooth or irregular surface (in the case of spherical or cubic objects, respectively) ensures that the same control rules lead to two types of behaviours depending on the type of the object. These behaviours consist in following the surface of the object and then moving away from the object in the case of spherical objects, and in following the surface of the object by getting stuck in a corner in the case of cubic objects. The exhibition of these two behaviours allows the agent to experience rather different proprioceptors states as a consequence of having interacted with spherical or cubic objects that nicely encode the regularities that are necessary to differentiate the two types of object.

For other examples which show how adaptive agents can exploit the fact that behavioural and cognitive processes arising from the interaction between lower-level behaviours or between the constituting elements later affect these lower-level processes, see [Scheier *et al.* 1998; Nolfi, 2002; Beer, 2003].

3 BEHAVIOUR AND COGNITION AS ADAPTIVE SYSTEMS

The complex system nature of behaviour and cognition can also help us to understand why embodied and situated agents are difficult to handcraft by a human designer. The fact that the relation between the characteristics of the robot and of the environment and the behavioural and cognitive skills that emerge from robot/environmental interactions is so complex and indirect, in fact, implies that a human observer cannot predict the behaviour that will be exhibited by the robot (without having had the chance to observe the behaviour exhibited by the robot in interaction with the environment) even on the basis of a detailed description of the characteristics of the robot and of the environment and of the rules that regulate the robot/environmental interaction. This in turn implies that the task faced by the robot designer which consists in inferring the fine-grained characteristics that the robot and the robot's control system should have in order to produce a given desired behaviour is extremely hard if not completely hopeless [Funes *et al.*, 2003; Nolfi, 2005].

On the other hand, as we have shown in the examples illustrated in the previous section, the fine-grained characteristics of the robot can be effectively shaped through an adaptive process. In particular, robots able to display the required behavioural skills can be developed through an evolutionary technique [Nolfi and Floreano, 2000] in which the fine-grained characteristics of the robots' control system (and eventually of the robot body) are varied randomly and variations are retained or discarded on the basis of the their effect at the level of the global behaviour exhibited by the robot (or by the robots). This type of adaptive process in fact is able to discover and retain the characteristics of the robot or robots that, in interaction with their other characteristics and with the environment, can lead to useful emergent properties through a trial and error process that does not require computing the relation between the fine-grained characteristics and the higher-level properties that emerge from their interactions.

From an engineering point of view these considerations imply that the development of effective embodied agents (i.e. agents able to exploit the properties emerging from the interactions) require new methods. One possibility consists in *artificial adaptive* methods, as the evolutionary method illustrated in section 2, in which the fine-grained characteristics of the agents are modified through a trial and error process and in which variations are retained or discarded on the basis of their effects at the level of the global behaviour exhibited by the agents in the environment. A second possibility might consist in the development of new *design for emergence* methods that could provide a way to handcraft agents able to exploit emergent properties. Whether effective methods of this kind can be developed or not represents an open question at the moment. For a preliminary attempt to formulate design for emergence principles see [Pfeifer and Scheier, 1999].

The complex system nature of behaviour and cognition and the consequent difficulty of identifying the characteristics that the fine-grained interaction between the robot and the environment should have in order to produce a desired global behaviour do not imply that the behavioural and cognitive skills exhibited by adaptive agents are intrinsically unpredictable. It only implies that solutions analogous to those that can be generated through an adaptive process cannot be generated by human designers. In other words, adaptive solutions tend to be qualitatively different from hand-crafted solutions. Moreover, the complex system nature of behaviour and cognition does not imply that adapted solutions are inscrutable or inexplicable as we have also demonstrated in the example reported above. Indeed, the detailed analysis of adapted natural and artificial solutions represents the main instrument that we have to identify the general principles which characterize embodied intelligence.

From a modelling point of view, the complex system nature of behaviour and cognition implies that the characteristics of living organisms crucially depend on the characteristics of the processes that determine how they change, phylogenetically and ontogenetically, as they adapt to their environment. In other words these considerations suggest that adaptivity should be considered an equally fundamental property of behavioural systems as their embodied and situated nature.

4 DISCUSSION AND CONCLUSION

In this paper we illustrated how the behavioural and cognitive skills displayed by embodied and situated agents can be properly characterized as a complex system with multi-level and multi-scale organization and involving both bottom-up and top-down influences that emerge from several fine-grained interactions between the agent (or the agents) and the environment.

The complex systems nature of adaptive agents that are embodied and situated has important implications that constrain the organization of these systems and the dynamics of the adaptive process through which they develop their skills.

With respect to the organization of these systems, this complexity implies that agents' behavioural and/or cognitive skills (at any stage of the adaptive process) cannot be traced back to anyone of the three foundational elements (i.e. the body of the agents, the control system of the agents, and the environment) in isolation but should rather be characterized as properties which emerge from the interactions between these three elements and the interaction between behavioural and cognitive properties emerging from the former interactions at different levels of organizations. Moreover, it implies that 'complex' behavioural or cognitive skills might emerge from the interaction between simple properties or processes. A similar argument can be made for what concerns the course of the adaptive process, which cannot be traced back to the three foundational elements in isolation but rather depends on the interactions between these elements and on the interaction between the higher-level behavioural and cognitive processes that emerge from lower-level interactions. Indeed new behavioural skills might originate during the adaptive process both as a result of variations of the characteristics of the robots and as a result of variations of the physical or social environment.

With respect to agents' adaptive process, the development of new 'complex' skills does not necessarily require the development of new morphological features or new dedicated control mechanisms. Indeed, behavioural or cognitive skills might arise spontaneously as a result of the interaction between properties serving different functions and/or as a result of simple additional characteristics thanks to the possibility to exploit the emergent results of the interaction between these new characteristics with the other pre-existing characteristics and skills.

The study of adaptive behaviour in artificial agents that has been reviewed in this paper has important implications both from an engineering point of view (i.e.

for progressing in our ability to develop effective machines) and from a modelling point of view (i.e. for understanding the characteristics of biological organisms).

In particular, from an engineering point of view, progress in our ability to develop complex adaptive behavioural and cognitive systems can lead to development of new artefacts playing useful functionalities.

From a modelling point of view, progress in our ability to model and analyze behavioural and cognitive skills in adaptive embodied agents can improve our understanding of the general mechanisms behind animal and human intelligence.

More specifically, the analysis of these systems can allow us to identify which are the fundamental properties of natural intelligence that characterize a large variety of species independently from their specific morphological, neuro-physiological, and ecological characteristics.

Moreover, this type of research can help us to formulate new theoretical concepts and terms which can allow us to model and describe the key characteristics of natural intelligence. For example, it can contribute to better define two fundamental aspects of natural (and artificial) intelligence: *morphological computation* and *sensory-motor coordination*. The former concept refers to the fact that the characteristics of the body of an agent (from the overall morphological structure to the fine-grained characteristics of the body such us the exact position of the receptors or the degree of elasticity of different body parts) strongly determine the skills that an agent might exhibit and the complexity of the control mechanisms which are required to produce such skills. The latter concept refers to the fact that the sensory states experienced by an agent are determined not only by the characteristics of the environment and by the agent/environmental relation but also by the motor actions previously performed by the agent itself. Indeed, behavioural and cognitive skills might emerge from the dynamical process arising from the agent/environmental interactions without the need of dedicated control mechanisms provided that the rules that regulate how the agent reacts to sensory states have been shaped to appropriately exploit the properties emerging from the agent/environmental interactions.

Finally, the comprehension of the complex system nature of behavioural and cognitive skills illustrated in this paper can allow us to better define the notion of embodiment and situatedness, that are universally recognized as central aspects in the study of natural and artificial intelligence but that still lack a clear and uncontroversial definition. Possessing a body and being in a physical environment certainly represent a pre-requisite for considering an agent embodied and situated. However, a more useful definition of embodiment (or of true degree of embodiment) can be given in terms of the extent to which a given agent exploits its body characteristics to solve its adaptive problem (i.e. the extent to which its body structure can be adapted to the problems to be solved, or in other words, the extent to which its body performs morphological computation). Similarly, a more useful definition of situatedness (or of true degree of situatedness) can be given in terms of the extent to which an agent exploits its interaction with the physical and social environment and the properties originating from this interaction to solve its adaptive problems. For the sake of clarity we can refer to the former definition of the terms (i.e. merely possessing a physical body and being situated in a physical environment) as embodiment and situatedness in the weak sense, and to the latter definitions as embodiment and situatedness in a strong sense.

BIBLIOGRAPHY

[Baldassarre *et al.*,, 2006] G. Baldassarre, D. Parisi, and S. Nolfi. Distributed coordination of simulated robots based on self-organisation. *Artificial Life*, 12(3), 289-311, 2006.

[Beer, 1995] R. D. Beer. A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence,* 72, 173-215, 1995.

[Beer, 2003] R. D. Beer. The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, 11, 209–243, 2003.

[Bongard and Paul, 2001] J. C. Bongard and C. Paul. Making evolution an offer it can't refuse: Morphology and the extradimensional bypass. In J. Keleman, P. Sosik (eds.), *Proceedings of the Sixth European Conference on Artificial Life*, Lecture Notes in Artificial Intelligence, 2159. Springer-Verlag, Berlin, 2001.

[Brooks, 1991] R. A. Brooks. Intelligence without reason. In J. Mylopoulos, R. Reiter (eds.), *Proceedings of 12th International Joint Conference on Artificial Intelligence*, San Mateo, CA, Morgan Kaufmann, 1991.

[Chiel and Beer, 1997,] H. J. Chiel and R. D. Beer. The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences*, 20, 553-557, 1997.

[Endo *et al.*, 2002] I. Endo, F. Yamasaki, T. Maeno, and H. Kitano. A method for co-evolving morphology and walking patterns of biped humanoid robot. In *Proceedings of the IEEE Conference on Robotics and Automation*, 2002.

[Funes *et al.*, 2003] P. Funes, B. Orme, and E. Bonabeau. Evolving emergent group behaviours for simple humans agents. In P. Dittrich, J. T. Kim (eds.), *Proceedings of the Seven European Conference on Artificial Life*, 2003.

[Gigliotta and Nolfi, 2008] O. Gigliotta and S. Nolfi. On the coupling between agent internal and agent/environmental dynamics: Development of spatial representations in evolving autonomous robots. *Connection Science*, 16, 148 – 165, 2008.

[Goldenberg *et al.*, 2004] E. Goldenberg, J. Garcowski, and R. D. Beer. May we have your attention: Analysis of a selective attention task. In S. Schaal, A. Ijspeert, A. Billard, S. Vijayakumar, J. Hallam, J.-A. Meyer (eds.), *From Animals to Animats 8*, Proceedings of the Eighth International Conference on the Simulation of Adaptive Behavior. Cambridge, MA, MIT Press, 2004.

[Keijzer, 2001] F. Keijzer. *Representation and behavior.* MIT Press, London, 2001.

[Marocco and Nolfi, 2007] D. Marocco and S. Nolfi. Emergence of communication in embodied agents evolved for the ability to solve a collective navigation problem. *Connection Science*, 19(1), 53-74, 2007.

[McGeer, 1990] T. McGeer. Passive walking with knees. In *Proceedings of the IEEE Conference on Robotics and Automation*, volume 2, 1640–1645, 1990.

[Mondada *et al.*, 1993] F. Mondada, E. Franzi, and P. Ienne. Mobile Robot miniaturisation: A tool for investigation in control algorithms. In *Proceedings of the Third International Symposium on Experimental Robotics*, Kyoto, Japan, 1993.

[Mondada *et al.*, 2004] F. Mondada, G. Pettinaro, A. Guigrard, I. Kwee, D. Floreano, J-L. Denebourg, S. Nolfi, L. M. Gambardella, and M. Dorigo. Swarm-bot: A new distributed robotic concept. *Autonomous Robots*, 17(2-3), 193-221, 2004.

[Nolfi and Floreano, 2000] S. Nolfi and D. Floreano. Evolutionary Robotics: The biology, intelligence, and technology of self-organizing machines. MIT Press/Bradford Books, Cambridge MA, 2000.

[Nolfi, 2002] S. Nolfi. Power and limits of reactive agents. *Neurocomputing*, 49, 119-145, 2002.

[Nolfi and Marocco, 2002] S. Nolfi and D. Marocco. Active perception: A sensorimotor account of object categorization. In B. Hallam, D. Floreano, J. Hallam, G. Hayes, J-A. Meyer (eds.), *From Animals to Animats* 7, Proceedings of the VII International Conference on Simulation of Adaptive Behavior, MIT Press, Cambridge, MA, 266-271, 2002.

[Nolfi, 2005] S. Nolfi. Behaviour as a complex adaptive system: On the role of self-organization in the development of individual and collective behaviour. *Complexus*, 2(3-4), 195-203, 2005.

[Pfeifer *et al.*, 2006] R. Pfeifer, F. Iida, and G. Gómez. Morphological computation for adaptive behavior and cognition. In *International Congress Series*, 1291, 22-29, Elsevier, 2006. Available at `http://www.sciencedirect.com/{doi:10.1016/j.ics.2005.12.080}`.

[Pfeifer and Scheier, 1999] R. Pfeifer and C. Scheier. *Understanding Intelligence.* MIT Press, Cambridge MA, 1999.

[Scheier *et al.*, 1998] C. Scheier, R. Pfeifer, and Y. Kunyioshi. Embedded neural networks: exploiting constraints. *Neural Networks*, 11, 1551-1596, 1998.

[Schmitz *et al.*, 2007] A. Schmitz, G. Gómez, F. Iida, and R. Pfeifer. On the robustness of simple speed control for a quadruped robot. In *Proceeding of the International Conference on Morphological Computation*, CD-ROM, 2007.

[Slocum *et al.*, 2000] A. C. Slocum, D. C. Downey, and R. D. Beer. Further experiments in the evolution of minimally cognitive behavior: From perceiving affordances to selective attention. In J. Meyer, A. Berthoz, D. Floreano, H. Roitblat and S. Wilson (eds.), *From Animals to Animats 6*, Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior. MIT Press, Cambridge MA, 2000.

[Sugita and Tani, 2005] Y. Sugita and J. Tani. Learning semantic combinatoriality from the interaction between linguistic and behavioral processes. *Adaptive Behavior*, 13(1), 33-52, 2005.

[Tani and Fukumura, 1997] J. Tani and N. Fukumura. Self-organizing internal representation in learning of navigation: A physical experiment by the mobile robot Yamabico. *Neural Networks*, 10(1), 153-159, 1997.

[Vaughan *et al.*, 2004] E. Vaughan, E. A. Di Paolo, and I. Harvey. The evolution of control and adaptation in a 3D powered passive dynamic walker. In J. Pollack, M. Bedau, P. Husband, T. Ikegami, R. Watson (eds.), *Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems.* MIT Press, Cambridge MA, 2004.

Part V

Climatology

THE COMPLEX DYANMICS OF THE CLIMATE SYSTEM: CONSTRAINTS ON OUR KNOWLEDGE, POLICY IMPLICATIONS AND THE NECESSITY OF SYSTEMS THINKING

Carolyn W. Snyder, Michael D. Mastrandrea, and Stephen H. Schneider

1 INTRODUCTION

Now is a critical time to understand the climate system. Humans continue to burn fossil fuels, dumping heat-trapping greenhouse gases (GHGs) into the Earth's atmosphere. In the last 150 years, the atmospheric concentration of carbon dioxide (CO_2) has increased over 35% from 280 to 385 parts per million. CO_2 is currently at a concentration higher than at any time in the past 800,000 years [Lüthi *et al.*, 2008], and likely higher than any in the past 45 million years [Pagani *et al.*, 2005]. The influences of warming temperatures can already be seen around the world: melting mountain glaciers and polar ice; rising and increasingly acidic seas; increasing severity of droughts, heat waves, fires, and hurricanes; changes in the lifecycles and ranges of plants and animals; spreading infectious diseases; and increasing heat-related deaths [Solomon *et al.*, 2007; Parry *et al.*, 2007]. The most recent of the global assessments by the Intergovernmental Panel on Climate Change (IPCC)[1] concluded that "the warming of the climate system is unequivocal" and that "most of the observed increase in global average temperatures since the mid-20th century is *very likely* due to the observed increase in anthropogenic greenhouse gas concentrations" [Solomon *et al.*, 2007].

Anthropogenic (human-caused) climate change involves interactions among complex global geophysical, biological, social, and economic systems. Systems concepts, principles, and methods are essential to understand the climate system and the dynamics of climate change. The Earth's climate system hosts a myriad of nested and coupled sub-systems, functioning at various temporal and spatial

[1]The Intergovernmental Panel on Climate Change (IPCC) is a scientific intergovernmental body established in 1988 by the World Meteorological Organization (WMO) and the United Nations Environment Program (UNEP). The most recent 2007 IPCC assessment was the product of experts from more than 130 countries, with more than 450 lead authors, 800 contributing authors, and an additional 2,500 experts reviewing the draft documents. The IPCC shared the 2007 Nobel Peace Prize with former Vice President of the United States Al Gore.

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

scales. The climate system also provided the first inspiration for **chaos theory** and **complexity** science, through Edward Lorenz's investigations of weather prediction in 1961. Lorenz discovered that regardless of computing power, it is impossible to accurately predict the time-evolving details of the weather beyond about ten days due to the rapid growth rate of even tiny errors in **initial conditions** — this **extreme sensitivity to initial conditions** is the famous "butterfly effect." However, the climate system can behave very differently than weather: at some scales climate system behavior is **chaotic**, whereas at other scales it is **stochastic** or even **deterministic** — illustrating the complex behavioral landscape of the dynamics of the climate system.

The various scientific disciplines studying components of the climate system often neglected to focus on the **complexity** of this multi-component system, and instead modeled the sub-components of the climate system in isolation and along distinct disciplinary lines. Gradualism and linearity usually were assumed, and most models produced internally stable and predictable behavior. Even when **nonlinearity** was explicitly modeled, the models have been based on simplified representations of small-scale phenomena. The lack of coupling between climate subsystems in most early studies mirrored the lack of coupling within the diverse, disciplinary-oriented scientific community. **Emergent properties**, **multiple equilibria**, **path dependence**, and **nonlinearities** inherent in the climate system were often overlooked, and when discovered, were sidelined as exceptions rather than fundamental properties [Rial, *et al.*, 2004]. There are, however, an increasing number of researchers who apply the tools of complexity analysis to the climate system, and in turn, climate science has expanded and enriched complexity theory.

Chapter outline

The goal of this chapter is to describe the contribution of **complexity science** to our understanding of the climate system and the unique challenges its **complex properties** pose to climate predictions and policy analysis.

We first present a brief exploration of the Earth's climate system through the lens of complexity science. We then introduce the data sources and modeling strategies that climate science uses to understand past behavior, to fingerprint causes of current climate changes, and to project future climate. The complex dynamics of the climate system constrain our ability to gain knowledge about the climate system and add uncertainty to predictions of the impacts of human-induced climate change. We investigate six case studies that illustrate the importance and development of key complexity themes in climate science: glacial-interglacial cycles, thermohaline ocean circulation, ice sheets, vegetation cover changes, extinction, and overshoot scenarios.

We also investigate the implications of the complexity of the Earth system for climate policy analysis. Assessments of the impacts of climate change are often disciplinary-based and not sufficiently integrative across important disciplinary

DETERMINISTIC: A deterministic system has no randomness in the development of future states of the system. The future dynamics of a deterministic system are fully defined by their initial conditions.

STOCHASTIC: A stochastic system is the opposite of a deterministic system. In stochastic or random dynamics there is indeterminacy in the future evolution of the system, which can be described with probability distributions. Even if the initial conditions were known, there are many possible states the system could reach, and some states can be more probable than others.

CHAOS: Chaos theory describes the behavior of certain dynamical systems that are highly sensitive to initial conditions, while having dynamical trajectories that are bounded (a strange attractor) and non-repeating (aperiodic). Although the behavior of chaotic systems can appear random due to the exponential growth of perturbations in the initial conditions, they are actually deterministic. For all practical purposes, however, chaotic systems are unpredictable despite being deterministic, due to the high sensitivity to initial conditions.

NONLINEARITY: Linear systems obey the principle of superposition: that the net response at a given place and time caused by two or more stimuli is the sum of the responses which would have been caused by each stimulus individually. In contrast, a nonlinear system is one where the principle of superposition fails. Thus, the behavior of nonlinear systems cannot be expressed as a sum of the behavior of its parts (or their multiples).

COMPLEXITY: Complexity has many definitions, but one perspective is that the complexity of a phenomenon is a measure of how difficult it is to describe. Both a simple linear system and a simple, purely stochastic system can be fully described with little information, and thus complexity is a characteristic independent of the stochastic/deterministic spectrum.

COMPLEX SYSTEM: A complex system is composed of interconnected parts that as a whole exhibit one or more properties not present in the individual parts alone. Examples of some potential features of complex systems include unpredictability, emergence, interactions between system components, simultaneous order and disorder, heterogeneity, chaos, nonlinearity, feedback loops, and hysteresis. These features will be defined and explored in more detail in Section 2.

subcomponents, producing misleading results that have potentially dangerous environmental consequences. The current framework of cost-benefit optimization is particularly flawed. We discuss how one should restructure climate policy analysis as an integrated assessment process, combining data and relationships from the physical, biological and social sciences, that includes robust assessments of potential risks within a vulnerability framework.

2 THE EARTH'S CLIMATE AS A COMPLEX SYSTEM

We now briefly explore the Earth's climate system through the lens of **complexity science**: **complex system** structure, **feedbacks**, **transient responses**, **emergent properties, multiple equilibria**, and **path dependence**. This is not intended as an introduction to climate science, and the reader is referred to the first chapter of Solomon *et al.* [2007] for an overview of the climate system and to other chapters in the report for more in-depth discussions of particular system components and interactions. Alternatively, Schneider *et al.* [2009] provide a broad and less technical overview.

2.1 Structure

The climate system consists of **nested and interlinked subsystems**, often divided (along disciplinary lines) into the following large-scale components: the atmosphere, the ocean, terrestrial ecosystems, and the cryosphere (see Figure 1 for a summary illustration). Each of these subsystems is itself a **complex system**. An example of nested systems can be seen in moving from: a single plant cell; to a whole plant; to the interactions within a forest community of plants, atmosphere, and soil; to the larger ecosystem of herbivores, plant competition, nutrient cycling, and microclimates; to the continental scale of terrestrial sinks and sources of GHG and changes in regional energy fluxes and water circulation; and finally, to the aggregate global impact on temperature and precipitation on the scale of hundreds to thousands of years. At different scales, we find different combinations of stabilizing and destabilizing **feedbacks**.

2.2 Feedbacks

> **FEEDBACKS:** Feedback is a circular causal process in which some portion of a system's output is returned (fed back) into the system's input. **Negative feedbacks** act to stabilize the system, decreasing a new output change. In contrast, **positive feedbacks** act to destabilize the system, increasing or amplifying a new output change.

Some of the connections between system components create **positive or destabilizing feedbacks** and some **negative or stabilizing feedbacks**. An example of a **positive feedback** is changes in ice and snow cover. The reflectivity of the Earth's surface is called albedo, and snow and ice have very high albedo and thus

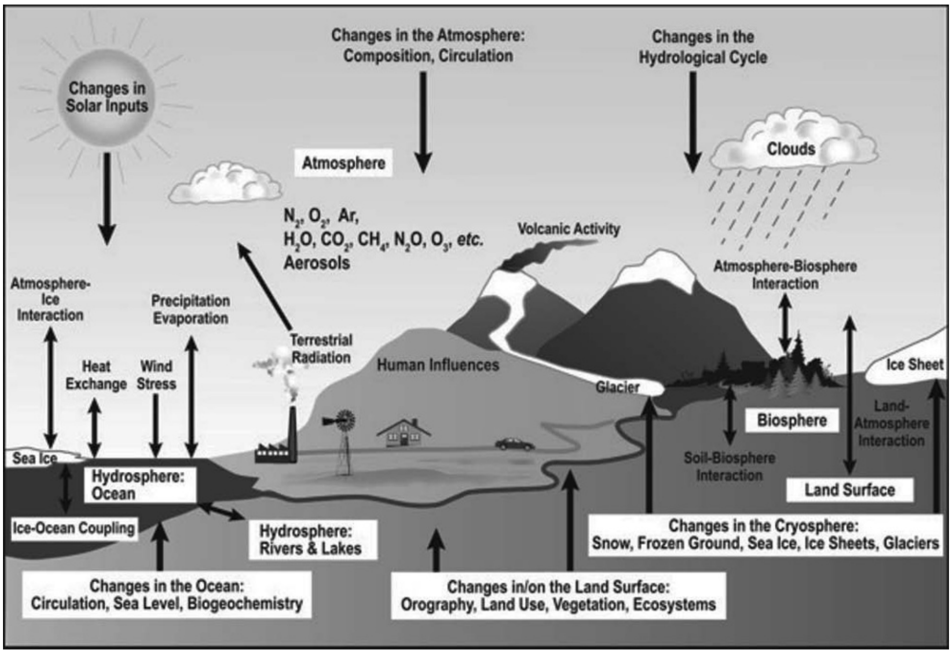


Figure 1. Summary schematic view of the components of the climate system, their processes, and their interactions. From [Solomon *et al.*, 2007, FAQ 1.2, Figure 1].

reflect more incident solar radiation, cooling the Earth's surface. An increase in snow and ice cover increases the Earth's albedo, causing the Earth's surface to cool, and leading in turn to increased ice and snow cover, further increasing surface albedo, which causes further cooling. Conversely, a decrease in snow and ice cover decreases the Earth's albedo, causing the Earth's surface to warm, leading to further melting of snow and ice, exposing the darker land, vegetation, and/or water beneath and decreasing the surface albedo, which causes further warming. An example of a **negative feedback** is the uptake of CO_2 from the atmosphere by the oceans. The higher the CO_2 concentration in the atmosphere, the more CO_2 the ocean absorbs, thus reducing the net change of the atmospheric composition and helping to stabilize the climate.

Terrestrial ecosystems can provide both **positive and negative feedbacks**, sometimes from the same event and often at different rates and spatial scales. Warming temperatures and milder winters can cause the northern expansion of woody vegetation. Such expansion can store CO_2 from the atmosphere and slow further global warming. However, vegetation expansion also can decrease surface albedo, especially when growing over snow-covered ground, thus potentially causing regional warming. Current models suggest that the net effect for such increased forest cover in boreal latitudes is warming, but effects depend on the complexities of cloud feedbacks and might be highly scale and location dependent [Bala *et al.*, 2007; Gibbard *et al.*, 2005; Betts, 2000; Bonan *et al.*, 1992]. An important and highly debated **negative feedback** of vegetation is the potential for increased storage of carbon in response to higher CO_2 concentrations (so-called CO_2 fertilization).

Water vapor causes another complex **feedback** in the climate system. Increased temperatures cause more water to evaporate from the ocean, producing higher amounts of water vapor in the atmosphere, enhanced by the ability of warmer air to hold more water vapor — a well-understood, **nonlinear**, but deterministic, process. Water vapor is itself a GHG, and thus results in further increased temperatures. Increased water vapor also may cause more widespread cloudiness, and clouds can act as both a **negative and a positive feedback**. Clouds reflect incident solar radiation, causing cooling. However, increases in the height of the tops of clouds can cause warming [Schneider, 1972]. The net feedback from water vapor and clouds is one of the most uncertain elements of the climate system. Variations in how clouds are described in current climate models causes the majority of the variation in the projected warming from a specific GHG atmospheric concentration change, with upper estimates of 80% of the variation in model projections coming from their representations of clouds [Soden and Held, 2006; Webb *et al.*, 2006; Knight *et al.*, 2007; Solomon *et al.*, 2007].

Lastly, a more fundamental understanding of the climate system necessitates full consideration of the human component of the system, as discussed further in Section 5. The social system is also a set of **nested systems** with **complex interactions** and **feedbacks**. For example, rising temperatures might increase the use of air conditioning, which will in turn cause increases in electricity use and thus GHG emissions (if the electricity is generated by burning fossil fuels), **a positive feedback**. In contrast, experiencing the early negative consequences of climate change might cause increased policy and behavior change, **a negative feedback** [Schneider, 1997; Liu *et al.*, 2007].

2.3 Transient Responses

TRANSIENT: A transient is the dynamical trajectory of a system not in equilibrium. Transient responses vary over time. Often in the climate system, different components respond differently over time, creating unique patterns as a system evolves over time toward equilibrium.

Different system components respond to external or internal changes at different rates and in spatially heterogeneous patterns, resulting in a **nonlinear** and **transient** response. For example, changes in GHG concentrations change the energy balance of the planet, which in turn causes changes in water cycles and clouds within days, circulation patterns within days to years, terrestrial ecosystems within years to centuries, oceans within decades to a millennium, and ice sheets within centuries to many millennia. The aggregate effect on global temperature is thus only fully realized after thousands of years. These differences in timing result in complicated leads, lags, delays, and inertia in the climate system. Moreover, many processes are often highly sensitive to the time features of the perturbation, such as the rate of change, timing of change, and the location in history. These transient responses pose serious challenges to understanding the climate system and identifying the key dynamical causes at work, because it is hard to know if a transient path represents movement toward a stable condition, a movement toward a radically different dynamical form (a **bifurcation**), or simply unending transient change.

2.4 Emergent Properties

> **EMERGENCE:** Emergence is the phenomenon where a complex system's behavior pattern arises out of a multiplicity of relatively simple interactions in ways that fundamentally change the system dynamics (a bifurcation), and the system is said to self-organize. Some use the terms 'emergence' and 'self-organization' more broadly for any sufficiently novel behavior. Emergent properties could not be predicted from knowing the interactions of components one level below the phenomenon, and emergence can only be studied at a level higher than the constituent components.

At some scales, climate behavior is simple, deterministic, and relatively straightforward to model and predict—a truly remarkable example of emergence, considering the complex nested structure described previously. It is important to understand that climate is the average behavior of the system over decades or longer as well as over regional spatial scales, not changes in the day-to-day local weather or year-to-year variability. Examples of emergent properties of the climate system include seasonal cycles of temperature, some ocean circulation patterns, and the glacial-interglacial cycles.

For example, as the Earth went in and out of ice ages and other periods of rapid temperature change, the climate followed an asymmetrical pattern of slow cooling and steep warming [Rial *et al.*, 2004]. Over the past 350,000 years, there seems to be a correlation between CO_2 concentrations and Antarctic temperature with a substantial R^2 (least squares correlation) of $\sim$0.8, meaning that 80% of the total variation in Antarctic temperature over the past 350,000 years can be explained solely from variations in CO_2 concentrations [Cuffey and Vimeux, 2001]. This strong correlation is striking given that these are times when many conditions have changed drastically: ice sheets have expanded and contracted over thousands of kilometers growing to heights of several km, sea levels have changed over 100 meters, and ecosystems have migrated across continents. Several similar patterns over different time scales are observed in the Earth's climatic history, and they vary in their spatial extent, the coupling or decoupling with different system components (such as GHG), and the relative timing of changes in different subsystems, e.g., [Zachos *et al.*, 2001]. Section 3 discusses the complications of inferring causality from the climate record and the imperfections in the available data and proxy methods. Section 4 discusses examples of using complexity theory to explain and model these patterns.

2.5 Multiple Equilibria

MULTIPLE EQUILIBRIA: Multiple equilibria occur when several different local regions of the same phase space are dynamical attractors. The same system dynamics can lead to more than one stable state depending on the location in phase space of the system, and thus on initial conditions. Knowing the system dynamics is not enough to know the long-term, stable outcome of the system. Minor perturbations can cause the system to shift between different equilibria or dynamical attractors, causing abrupt and dramatic changes in the system (though short of bifurcations).

THRESHOLDS: Thresholds can mark the borders between different equilibria, and thus crossing thresholds may cause dramatic changes in the system. Thresholds can also mark the transition between different system dynamics entirely, governed by a different set of differential equations (also called bifurcations), always a dramatic change. The term threshold is also used more broadly to define the minimum change before impacts are recognized as important or dangerous.

Nonlinear dynamics can create **multiple equilibria** in the Earth system. Lorenz termed this behavior "almost-intransitivity": more than one climate regime can be consistent with a given set of boundary conditions, and the observed climate may switch between regimes in a rather abrupt manner [Lorenz, 1968; 1976]. The system can have embedded **thresholds** or conditions under which abrupt and rapid changes occur, which are often irreversible on timescales relevant to humans.

There is no single **threshold** in the climate system, but rather many **interdependent thresholds** for different processes under different conditions and rates. **Thresholds** are difficult to predict due to dependence on **initial conditions**, couplings with other system components, and rapid change between **multiple equilibria**. Examples of abrupt climate change observed over Earth history are summarized in [Overpeck and Cole, 2006; Alley *et al.*, 2003; Schneider, 2004; Higgins *et al.*, 2002]. Examples of such effectively irreversible changes (on civilization timescales) include ocean circulation, species extinction, and vegetation cover and are discussed further in Section 4.

Lenton *et al.* [2008] brought together an international group of scientists to critically evaluate potential policy-relevant "tipping elements." They define tipping elements as large-scale components of the Earth system currently in danger of passing a tipping point (**threshold**), where a small perturbation at a critical point qualitatively alters the future fate of the Earth system with large-scale impacts on human and ecological systems. Their article provides a review of the current literature on the various tipping elements in the Earth System. Their top nine most policy-relevant tipping elements are depicted in Figure 2.

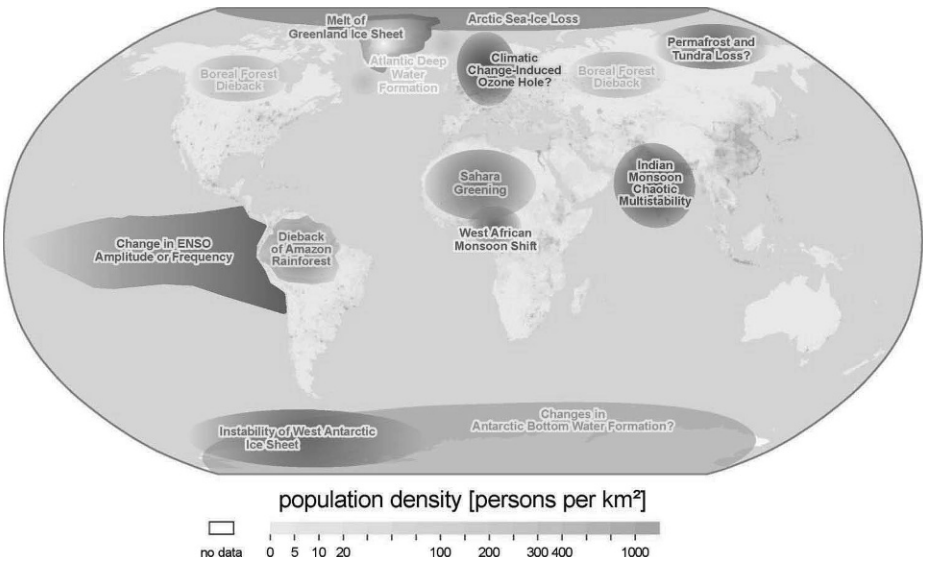


Figure 2. Map of potential policy-relevant tipping elements in the climate system, overlain on global population density. Subsystems indicated could exhibit threshold-type behavior in response to anthropogenic climate forcing, where a small perturbation at a critical point qualitatively alters the future fate of the system. They could be triggered this century and would undergo a qualitative change within this millennium. Excluded from the map are systems in which any threshold appears inaccessible this century (e.g., East Antarctic Ice Sheet) or the qualitative change would appear beyond this millennium (e.g., marine methane hydrates). Question marks indicate systems whose status as tipping elements is particularly uncertain. From [Lenton *et al.*, 2008, Figure 1].

2.6 Path Dependence

PATH DEPENDENCE & HYSTERESIS: Path dependence and hysteresis both describe the phenomenon of system memory, where the system state depends not only on the system dynamics and input, but also on the previous states of the system, such as initial conditions. The path a system takes through time impacts later system dynamics independent from new stimuli. Hysteresis occurs if the removal of a stimulus does not result in a system returning to its initial conditions. In this sense the system behaves irreversibly.

The climate system also exhibits memory or path dependence: the current state of the system cannot be explained simply from the current conditions alone. An example of path dependence in the climate system is vegetation cover. In some parts of the world, both dry grassland and wet rainforest are possible, despite having the same climate boundary conditions. The system's past determines the state in which it is stabilized. A fire or deforestation by humans can cause the rainforest to irreversibly become grassland even though the climate boundary conditions remain the same. This is because each vegetation type modifies its local climate and creates stable local conditions for its own existence.

Arctic sea ice provides another example of **path dependence**. Once lost, sea ice is very hard to regrow sufficiently to be able to subsist through the summer melt, even though thick sea ice could stably persist in the same climate conditions. The sea ice loss of 2007 was dramatic, shattering previous records and breaking from the existing gradual trend of reduced summer sea ice extent. The melt in 2007 caused 2008 to have thinner ice and be more vulnerable to melting. Despite 2008 being a mild year climatically for sea ice loss, there occurred a record year of sea ice loss almost as extreme as in 2007, see Figure 3. The dramatic difference in sea ice melt between 2008 and previous years is not explained by the climatic conditions, but by the historical legacy of the 2007 ice loss. The events of 2007 may have permanently (on the scale of tens to hundreds of years) shifted the Arctic sea ice regime from what it would have been due to average climate changes alone.

In conclusion, the complexity of the climate system is fundamental to understanding and predicting its dynamics. Simplifications to make sub-systems tractable to model are often necessary, but models of the climate system must be designed to represent the nonlinear and complex characteristics that control the climate system's behavior at its core.

3 CLIMATE SCIENCE

The study of climate science is limited by the inability to perform controlled, repeated experiments with our planet. Research must rely on historical data of responses to anthropogenic forcings over the past century, paleoclimatic data from the Earth's history, and modeling exercises. Observations provide evidence of correlations, but there will always be the limitation that correlations cannot by themselves prove causation, as they do not represent a controlled experiment. Thus, climate science focuses on detection and attribution: a rigorous process of establishing the most likely causes for a detected change in the climate system. We first summarize the key data constraints that climate science faces, then discuss climate models and their challenges, and conclude with a discussion of advances in detection and attribution studies that support strong conclusions about the impact of humans on the climate system.

3.1 Climate Data Constraints

Due to the inability to perform repeated experiments with our planet, data of past climate behavior are especially important in climate science, serving as our only observations of the system dynamics. Thus, the limitations of available climate

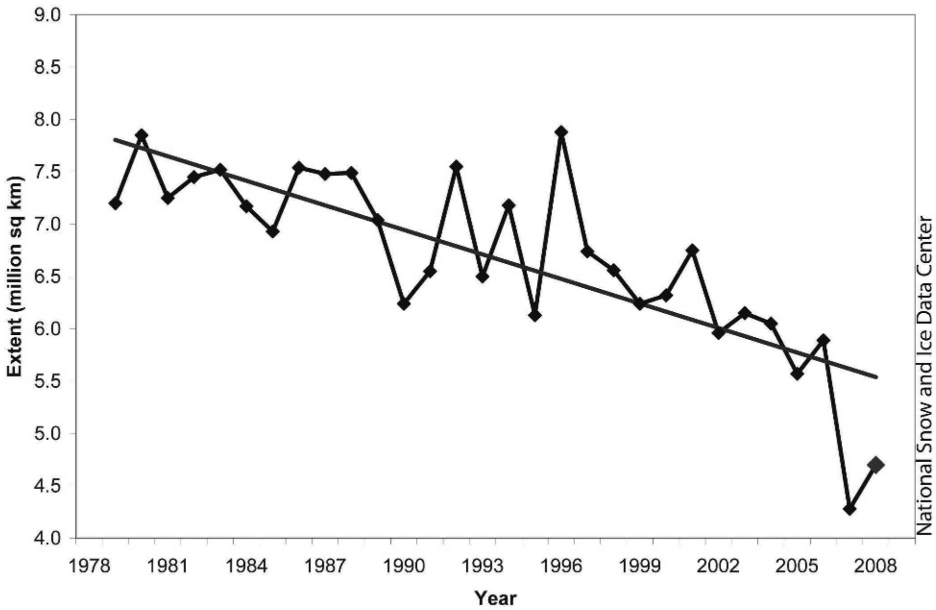


Figure 3. Observed decline in September sea ice extent from 1979 to 2008. The x-axis is years (1978–2008). The y-axis is September sea ice extent in millions of square km. The blue line shows the average rate of sea ice decline 1979–2005. (National Snow and Ice Data Center, `http://nsidc.org/arcticseaicenews/index.html`)

data pose serious limitations to climate science as a whole. Data with high accuracy and precision are needed to help characterize the climate system dynamics, especially due to the potential exponential growth of data uncertainty in complex dynamics. However, the available data on the climate system are incomplete, and many records contain considerable uncertainty. Many variables are changing at the same time, and the observed values of important forcings, like aerosol effects in clouds or solar variations before the satellite era, are often only poorly known. Without sufficient data, multiple stressors and different potential dynamics are hard to distinguish.

Modern Data. The most extensive and reliable data are those from modern measurements, such as weather stations and satellite observations — though translation of satellite radiance data into meaningful climatic variables like atmospheric temperatures is a difficult modeling exercise, e.g., [Santer *et al.*, 2003]. Historical quantitative data only cover a maximum of a few hundred years and have patchy spatial distribution. Natural variation limits the generality of studies of short duration, and sophisticated statistics are needed to investigate potential trends.

With short time series and small perturbations, it is hard to discern the signal from within the natural variability.

Paleoclimate Data. Before modern data, historical records and paleoclimatic records are exclusively based on proxies for most quantities of interest, because direct measurements are unavailable. A proxy is an observable entity that is thought to vary in a deterministic way with the quantity of interest. Potential proxies include tree rings, bore hole temperatures, coral reef chemistry, fossil leaves, ocean sediments, ice layers, layering in lakebeds, fossil pollen, and the shells of marine microorganisms. For example, the ratio of magnesium and calcium (Mg/Ca) in shells of marine microorganisms varies with the temperature of the water during the creation of the shell [Anand *et al.*, 2003]. Thus, the Mg/Ca ratio can serve as a record of the local ocean temperature.

The Challenges of Proxy Records. The paleoclimate record is attractive because it has large perturbations and long time series, but there is also increased uncertainty in the data. Proxies involve many complications, such as measuring the proxy itself, determining the age of the sample of interest, and accounting for potential post-depositional processes that might have altered the proxy or mixed samples across different ages. The main source of proxy uncertainty comes from the proxy calibration, where modern observational data or experiments are used to estimate the relationship between the proxy and the variable of interest. In our shell example, the Mg/Ca ratio should only vary due to temperature and not other environmental variables, and should vary in a way that can be determined from data from the modern oceans and be applied across long periods of time.

The proxy calibration process assumes that the relationship found today will hold for the different climatic or chemical conditions of the past, and also that the only variable that would cause changes in the proxy value is the particular variable of interest. Sometimes questionable assumptions of linearity and uniformity are inherent in the calibration process. To continue our Mg/Ca example, what if other system elements were different, such as the concentration of the ions in the seawater, the seawater pH, or the metabolism of the species? Multi-proxy studies can help reduce such uncertainties by combining estimates from several independent proxies to identify more robust conclusions. A famous example of this is how Michael Mann and his colleagues used over 100 proxy indicators to reconstruct the global mean surface temperature over the past two millennia [Mann *et al.*, 1998; Mann and Jones, 2003].

Another challenge with paleoclimatic records is that they often have inherent time averaging, from sediment mixing, bulk measurement processes, and dating uncertainties. Recent improvements in dating and chemical analyses have enabled finer and finer slices of time to be investigated in the Earth's history. Such work has revealed climate change events far more rapid than expected. Indeed, there had been a widespread assumption that the climate changed gradually and smoothly over time. Recent research has challenged such assumptions, exposing the complex nature of the behavior of the climate system and the capacity for abrupt climate change, see Section 4 for specific examples of abrupt climate change and [Overpeck and Cole, 2006] for a recent review. The modern view of abrupt climate variation as an emergent property of a complex system owes much to this advance in higher resolution paleoclimatology proxy records.

3.2 Climate Model Challenges

For some scales and processes of the climate system, climate science is dependent upon numerical computational modeling. Numerical approximations and computer simulations are essential, because the system of differential equations rarely can be solved analytically. Given the inability to perform controlled experiments with the Earth and the data constraints, model simulations (linked with available data) often provide the most useful tool to explore the potential dynamics of the climate system. Large complex climate models are a crucial contributor to recent conclusions about the attribution of anthropogenic climate change, and to the generation of future climate predictions needed for climate impact analysis and to inform policy discussions. Additionally, models have provided a key medium for interaction between disparate disciplinary communities with different sets of data and methods, providing the core connection for the "epistemic community" of global change science and serving as heuristic guides to complex phenomena for both researchers and policymakers [Edwards, 1996].

Background on Climate Models. A climate model is a set of mathematical statements describing the physical, biological, and chemical processes that determine climate. The ideal climate model would include all the processes known to have climatological significance and would involve spatial and temporal detail sufficient to model phenomena occurring over small geographic regions and over short time periods. Ideally, models would include all the relevant scales of interactions and processes within the system, including the various nested subsystems, their nonlinear behavior, and the couplings throughout the system. Today's best models try to approach this ideal, but still entail many compromises and approximations because of computational limits and our lack of understanding of many small-scale processes and how they scale up to drive larger-scale behavior. Computational limits impose trade-offs between spatial and temporal scales and between processes included or neglected.

Types of Climate Models. Climate models vary in complexity from zero-dimensional models to models so complex only a few supercomputers in the world are capable of running them. What must go into a climate model depends on what one wants to learn from it, and thus the scale and prioritization of system components depends on the scale and nature of the questions of interest. A model designed to probe glacial-interglacial cycles (scales of 100,000 years) would be very different from one investigating changes over the next century. Simple models have the advantage that their predictions are easily understood based on well-known physical, chemical, or ecological principles. They produce results quickly, and therefore, can be used to test a wide range of assumptions by varying structural parameters or boundary conditions in the model. Such simple models often focus on the emergent properties of the climate system that are easily modeled without direct representations of subsystem processes.

The simplest of models treats the Earth as a single point, with no atmosphere, no distinction between land and oceans, and merely a consideration of balancing ingoing and outgoing radiation. More advanced are "multi-box" models that treat land, ocean, and atmosphere as separate "boxes," and include flows of energy and matter between those boxes. The most sophisticated models are general circulation models (GCMs) that typically divide the world into high-resolution, three-dimensional grid cells (currently roughly 50-100km horizontally and 1km vertically). Earth system models of intermediate complexity (EMICs) represent a compromise: EMICs sacrifice resolution and "more correct" physical, biological, and chemical structural functions for computational efficiency. EMICs enable easier and faster experimentation than with GCMs, but retain more complexity and explicit representation of subsystem processes than simple models. Which level in the modeling hierarchy to choose depends on the problems under consideration and validation exercises to test the applicability of each approach, e.g., [Schneider, 1992].

Recent Trends in Climate Models. The trend over the past decade has been to improve and sequentially add coupling between the major subsystems of the climate. Early GCMs focused on the atmosphere, then oceans were coupled with the atmosphere, and most recently dynamic vegetation and carbon-cycle models are being coupled to sophisticated climate models. Each additional coupling causes significant changes to the model outputs. For example, Solomon *et al.* [2007] included a $\sim 1^{o}$C greater uncertainty in the upper range estimates for warming per emissions scenario by 2100 due to the recently included positive feedback from coupling the climate with the carbon cycle [Friedlingstein *et al.*, 2006]. Many processes remain to be better incorporated into climate models, including permafrost, ice sheets, ocean circulation, fire, and human land use [Field, *et al.*, 2007]. For example, humans currently modify 33-50% of the Earth's vegetated surface [Vitousek, *et al.*, 1997; Goldewijk, 2001], yet most current climate models assume 100% natural vegetation. There has also been a trend to focus more on transient climate changes and not just on equilibrium results. Transient analyses enable the investigation of potential rapid events, thresholds, and other highly nonlinear processes, many of which are imperative to an assessment of potential impacts from climate change, e.g., [Mastrandrea and Schneider, 2004].

Limitations of Modeling Sub-grid-scale Processes. Models are unable to explicitly model processes that occur at a scale smaller than the model's own grid cell. Such **sub-grid-scale processes** are treated implicitly in models through parametric representations that connect such processes via semi-empirical rules to grid box average values. Parameterization is the modeler's attempt to address the complexity challenge of connecting processes at different scales, including phenomena below the "grain size" of the smallest resolved element in the models. The process of developing and testing parameterizations that reliably incorporate sub-grid-scale processes and interactions is one of the most important, time consuming, and controversial tasks of climate modelers. Clouds are one of the most challenging processes to model, as even the highest resolution model still has grid cells far larger than clouds. For example, variations in the treatment of how to represent cloud patchiness within a grid cell in climate models can cause significant differences in model predictions [Schneider, Dickinson, 1976].

Climate Model Validation. How many levels of resolution are sufficient in a model? A common modeling principle in complexity science is the three-level model, where the model includes the focal level at which the main phenomena are located and then includes one level above and below the focal level. However, the climate system often necessitates systems and subsystems at far more than three levels, as discussed in Section 2. That the actual dynamics should dictate the number of levels in the model is straightforward, but how to know when a climate model with unintuitive and complex dynamics has sufficient levels is a very difficult challenge. Parameterization is a common practice in modeling simple systems as well, but in a complex system, how can we know that the parameterization is sufficient to capture the complex interactions that might manifest themselves several levels higher in the model? As stated above, addressing this question is one of the most important, time consuming, and controversial tasks for climate modelers. A variety of model validation techniques have been developed, including both comparisons with available data and comparisons among independent models. Data constraints yet again pose a challenge to the advancement of climate science.

Model validation based off available climate data usually attempts to reproduce known climatic conditions in response to known forcings, such as volcanic eruptions, seasonal variations, and past climates. Today's best GCMs do a good job reproducing global temperature records, as well as geographic and seasonal patterns of temperature. Models are less accurate in representing climatic variations involving precipitation and other aspects of the hydrologic cycle. The wide-ranging validation methods give considerable confidence that these models are treating the essential climate-determining processes with reasonable accuracy, but many areas still need improvements. A truism, likely to persist for decades more, is that some aspects of climate model projections are well established, others demonstrate competing explanations, and yet others are likely to continue to remain speculative.

3.3 Detection and Attribution

Given all these uncertainties and complications, how did the international community come to the consensus summarized in the introduction: "the warming of the climate system is unequivocal" and "most of the observed increase in global average temperatures since the mid-20th century is *very likely* due to the observed increase in anthropogenic greenhouse gas concentrations" [Solomon *et al.*, 2007]? It is amazing that despite all the data uncertainties and modeling challenges, the impacts of humans' GHG emissions can be defined and strong conclusions can be drawn. To do so, climate science uses a method called detection and attribution.

Human-induced changes to the climate occur against a backdrop of natural internal and externally forced climate variability that all can occur on similar temporal and spatial scales. Detection is the process of demonstrating that the climate has changed in some defined statistical sense, without providing a reason for that change. Attribution is the process of establishing the most likely causes for the detected change with some defined level of confidence. The attribution of observed climate changes to a given combination of human activities and natural influences requires consideration of multiple lines of evidence to demonstrate within a specified margin of error that the observed changes meet three criteria: 1) unlikely to be due entirely to internal variability; 2) consistent with estimated (modeled) response to the given combination of human and natural forcing; and 3) not consistent with alternative, physically plausible explanations of recent climate change that exclude some forcings. Unequivocal attribution would require controlled experimentation with the climate system [Mitchell *et al.*, 2001; Hegerl *et al.*, 2007].

The community works carefully to disentangle three potential explanations of observed changes in global mean surface air temperature: 1) natural internal variability; 2) natural external forcings; and 3) anthropogenic forcings. The first step is to estimate the statistical properties and spatial distribution of natural internal variability, also known as "climatic noise." Next is to determine if there is a significant trend discernible above the natural internal variability. Sophisticated statistical analysis of the historical and proxy data reveals a $\sim$0.75^{o}C "unequivocal" warming trend since 1850 [Solomon *et al.*, 2007].

To investigate the causation of this trend, scientists use the detection process outlined above, searching for "human fingerprints" or unique patterns that would be expected in response to anthropogenic forcings. For example, if the observed climate change was from variations in solar radiation, then all layers of the atmosphere should be warming. However, if GHG emissions are causing the warming, then the surface and lower atmosphere should have warmed while the upper atmosphere should have cooled. Observations support the latter pattern of warming and cooling, and thus this fingerprint strongly suggests that humans, not solar variations, are the cause of the recent decades of observed warming.

Climate models can perform another fingerprint analysis by testing for causes of observed climate trends, such as those observed over the second half of the 20th century. Models are run with only observed natural forcings (e.g., volcanic dust veils and solar radiant energy changes), with only observed anthropogenic forcings (e.g., GHG and aerosol emissions), and then with both natural and anthropogenic forcings. The models produce "surrogate" data for each of the three categories that are tested against real data observations. Only the combined natural and anthropogenic forcings provide the best fit for regional and global records, and the anthropogenic forcing alone performs significantly better than the natural forcings alone. Although correlation does not necessarily mean causation, together these investigations explicitly separate the various climate forcings and provide strong evidence for causation. A similar fingerprint analysis for the timing of events of plants, such as bloom dates, and of animals, such as return dates from migration, found the same results: natural and anthropogenic forcings together produced the highest correlations [Root *et al.*, 2005]. These and many other fingerprint analyses motivate the international consensus that humans "very likely" have caused most of the warming since the second half of the 20th century [Solomon, *et al.*, 2007].

4 CASE STUDIES OF COMPLEXITY IN THE CLIMATE SYSTEM

We investigate six examples that illustrate the importance and development of key complexity themes in climate science: glacial-interglacial cycles, North Atlantic thermohaline ocean circulation, ice sheets, vegetation cover changes, extinction, and overshoot scenarios. These case studies explore how the complex dynamics of the climate system limit our understanding and modeling of the climate system and add uncertainty to our ability to predict the impacts of climate change.

4.1 Glacial-Interglacial Cycles

The glacial-interglacial cycles illustrate a striking **emergent property** of the climate system, but we have yet to understand the system dynamics that create such behavior. Limitations from data, process understanding, and modeling ability over such long timescales have prevented us from understanding their true causes. With limited data it is impossible to distinguish between **transient** unique changes, cycles, bifurcations, and changes in boundary conditions. The long timescales require modelers to specify too many parameters and interaction coefficients without sufficient data, and thus most research has used simple models with a few key dynamics parameterized and fit to the limited available data. Glacial-interglacial cycles over the past million years are depicted from proxy data from Antarctic ice cores in Figure 4.

The ice ages of the late Pleistocene are remarkable for their quasi-periodic nature and repeated pattern of long slow cooling and abrupt warming (see Figure 4). The predominant theory in paleoclimatology explaining the timing of these repeated "saw-toothed" cycles is changes in the geometry of the Earth's orbit [Milankovitch, 1941; Hays *et al.*, 1976; Imbrie *et al.*, 1993]. Variations in the Earth's

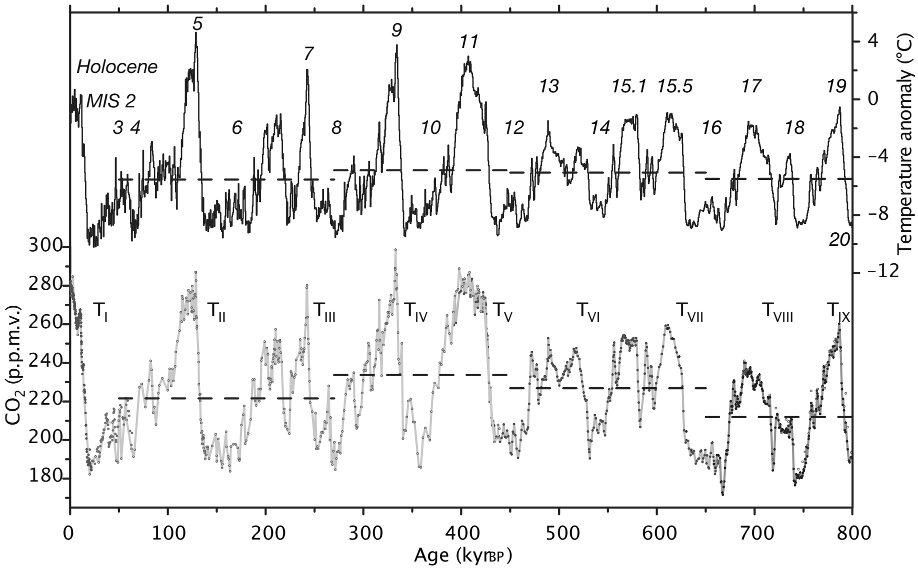


Figure 4. Glacial-Interglacial cycles over the last 800,000 years. The top curve is the Antarctic temperature anomaly record with respect to the mean temperature of the last millennium, based on original deuterium isotope data from the EPICA Dome C ice core and plotted on the EDC3 timescale [Jouzel *et al.*, 2007]. The bottom curve is a composite record for global CO_2 obtained from air bubbles from three different Antarctic ice cores (EPICA in blue and black lines, Vostok in green lines, and Taylor Dome in brown). Horizontal lines are the mean values of temperature and CO_2 for the time periods 799–650, 650–450, 450–270 and 270–50 thousands of years before present. Glacial terminations are indicated using Roman numerals in subscript (for example T_{IV}). Marine Isotope Stages (MIS) are given in italic Arabic numerals. From [Lüthi *et al.*, 2008, Figure 2].

eccentricity (the shape of the Earth's orbit around the sun), obliquity (the inclination of the Earth's rotational axis in relation to its plane of orbit around the sun), and precession (the change in direction of the Earth's rotational axis relative to fixed stars) comprise the three dominant cycles, collectively known as the "Milankovitch cycles." Each of the three effects has different quasi-periodicity: eccentricity at around 100 and 400 thousand years, obliquity at around 41 thousand years, and precession at around 19 and 23 thousand years.

However, theories regarding Milankovitch cycles are merely a statistical association, and not a physical explanation. The orbital changes in the Milankovitch cycles cause a seasonal and latitudinal redistribution of the solar radiation received at the Earth's surface, but a negligible change in the annual global average of radiation received. How do such redistributions of energy cause such large climatic changes? If the orbital changes are a deterministic driver of ice age cycles, then mechanisms must explain the **highly nonlinear** response of the climate system to the changes in the geometry of the Earth's orbit and the differential response to different orbital cycles [Schneider and Thompson, 1979]. High-confidence proposals for such mechanisms continue to evade researchers.

About one million years ago, the frequency of the ice ages switched from a glacial quasi-period 41,000 years long to a quasi-period of about 100,000 years with higher amplitude—a phenomenon called the mid-Pleistocene transition. The mid-Pleistocene transition occurred without a corresponding change in the frequencies of orbital changes. This abrupt shift in glacial frequency and amplitude illustrates the nonlinear potential of the climate system to abruptly shift between equilibrium states, here of different dominant frequencies and amplitudes. There is no accepted explanation for the cause of the mid-Pleistocene transition, but theories include a global cooling due to decreased atmospheric CO_2 [Berger *et al.*, 1999], an increased temperature contrast across the equatorial Pacific Ocean [de Garidel-Thoron *et al.*, 2005], and a switch to thicker ice sheets due to high-friction bedrock being exposed [Clark *et al.*, 2006].

Another enigma that remains unresolved in paleoclimatology is an explanation for why the quasi-periodic 100,000 year eccentricity cycle, the weakest of the orbital cycles, dominates the frequency of the ice ages of the past million years [Imbrie *et al.*, 1993]. Also in need of explanation: why is there a near absence of response at the strongest eccentricity period (about 400,000 years), and why has the timing between glacial periods been steadily increasing over the last 500,000 years [Raymo, 1997; Petit *et al.*, 1999; EPICA, 2004]? Many different mechanisms have been proposed to be the driver of the 100,000-years quasi-cycle: some involve internal dynamics of the climate system, such as ice sheet dynamics, ocean processes, and greenhouse gases, and others involve external drivers, such as changes in the geometry of the Earth's orbit. All require **high nonlinearity** in the climate system to create the large glacial cycles. Schneider and Thompson [1979] conclude that it is likely that no single physical process can be identified as predominant in this amplification, but rather it is due to the interactions of a number of processes on a variety of time scales.

Rial [1999] recently proposed a new description of the recent glacial cycles consistent with the classic theory of Milankovitch cycles: the glacial cycles represent **frequency modulation** of the 100,000-year eccentricity quasi-cycle by the 400,000-year eccentricity quasi-cycle. The frequency modulation is proposed to account for the variable duration of the ice ages, the multiple-peak character of the time series spectra, and the notorious absence of the significant 400,000-years quasi-cycle. Rial's [1999] theory provides a new description but still does not provide a mechanism for the **nonlinear amplification** of such external forcing beyond suggesting that the ice sheets may be acting as a "resonant oscillator." A mechanistic explanation of the Earth's emergent pattern of glacial-interglacial cycles remains an open debate.

4.2 North Atlantic Thermohaline Ocean Circulation

Abrupt change in the North Atlantic thermohaline ocean circulation (THC) is the most dramatic example in the climate system of **multiple equilibria** and **thresholds** that are **duration and path dependent**. Paleoclimate evidence of past abrupt changes in the North Atlantic THC revolutionized the field and changed the paradigm from gradual simple systems to abrupt complex systems. Moreover, the North Atlantic THC is of particular interest because it might be significantly altered due to global warming, causing large impacts to the world. Computational models have been useful tools in exploring potential complex system dynamics of the North Atlantic THC. However, our understanding of and ability to model the North Atlantic THC continues to be limited by the complexity of the system.

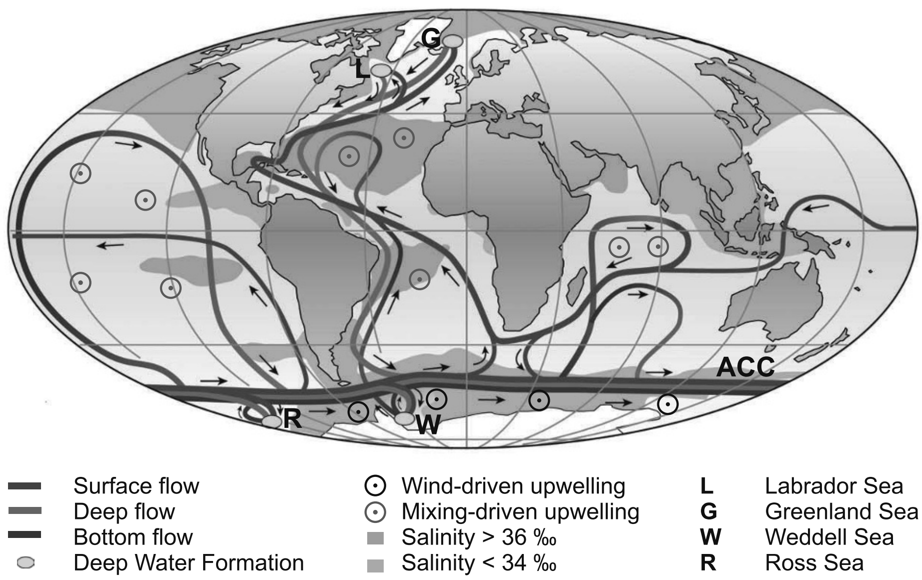


Figure 5. Schematic representation of the global thermohaline ocean circulation (THC). Surface currents are shown in red, deep waters in light blue and bottom waters in dark blue. The main deep-water formation sites are shown in orange. Average ocean salinity is 35‰ (‰ represents "parts per thousand"). The color-coding for water above 36‰ and below 34‰ indicates water significantly different from average. From [Rahmstorf, 2006, Figure 1].

The thermohaline ocean circulation (THC) is a density-driven circulation in the ocean, caused by differences in temperature and salinity, see Figure 5. In the North Atlantic, the THC brings warm tropical water northward, raising sea surface temperatures about 4^oC relative to temperatures at comparable latitudes in the Pacific Ocean [Higgins *et al.*, 2002]. The warmer waters of the North Atlantic warm and moisten the atmosphere, making Greenland and Western Europe roughly 5-8^{o}C warmer than they otherwise would be and increasing precipitation throughout the region [Stocker and Marchal, 2000; Broecker, 1997]. The THC is density-driven: warm, salty water flows into the North Atlantic, becomes more dense as it cools, and then sinks and flows to the southern hemisphere via the deep western boundary current, creating a "conveyor belt" of fast moving seawater [Rahmstorf, 2006; Broecker *et al.*, 1990]. Disruption of the density gradient through surface freshening or surface warming could cause a failure of sinking in the North Atlantic that would cause the North Atlantic THC to slow or even stop [Broecker, 1997]. The convective mixing process also is a self-sustaining, but **highly nonlinear**, process [von Deimling *et al.*, 2006]. There is strong evidence that the North Atlantic THC slowed or completely stopped in the Earth's history, causing rapid regional and global climatic changes [Broecker, 1997; Bond *et al.*, 1997; Rahmstorf, 2000b].

Paleoclimate reconstructions and model simulations suggest that there are **multiple equilibria** for the THC in the North Atlantic: rapid and repeated switching between equilibria over a period of years to decades is associated with dramatic changes in regional, and potentially global, climate [Alley *et al.*, 2003]. These multiple equilibria constitute an **emergent property** of the coupled ocean-atmosphere system. Henry Stommel was the first to describe the bi-stability of the North Atlantic THC system, using only a simple two-box model of the ocean [Stommel, 1961]. Figure 6 presents a summary of the three possible North Atlantic THC **equilibrium states** and the theoretical mechanisms for switching between them. In addition, the model predicts response **hysteresis** in which an earlier transition due to fresh water forcing (arrows "a" or "b") results in a system requiring a larger stimulus to return to its previous state (arrow "d"). A recent intercomparison of 11 different climate models of intermediate complexity found that all models show this North Atlantic THC hysteresis response to freshwater forcing [Rahmstorf *et al.*, 2005]. Both the total amount of freshwater forcing and the rate of forcing are important, because they can determine which of the opposing feedbacks — stabilization from local cooling and destabilization from reduced saltwater input — will dominate, and thus whether a rapid change will occur. Model experiments also find **importance for the duration and location** of freshwater inputs, in addition to the rate and magnitude of changes in freshwater [Rahmstorf, 2000b].

Global warming from anthropogenic GHG emissions could potentially impact the North Atlantic THC due to changes in temperature, precipitation, melt water, and circulation patterns. The existence of abrupt past climate changes from changes in the North Atlantic THC has fuelled concern over the possibility of setting off similar changes in the future. Models vary in their predictions of the weakening or shutdown of the North Atlantic THC with future anthropogenic climate change [Solomon *et al.*, 2007]. Disagreement between models on the proximity of the present-day climate to the Stommel **bifurcation point**[2] and on the

[2]The Stommel bifurcation is the point beyond which no North Atlantic Deep Water (NADW)

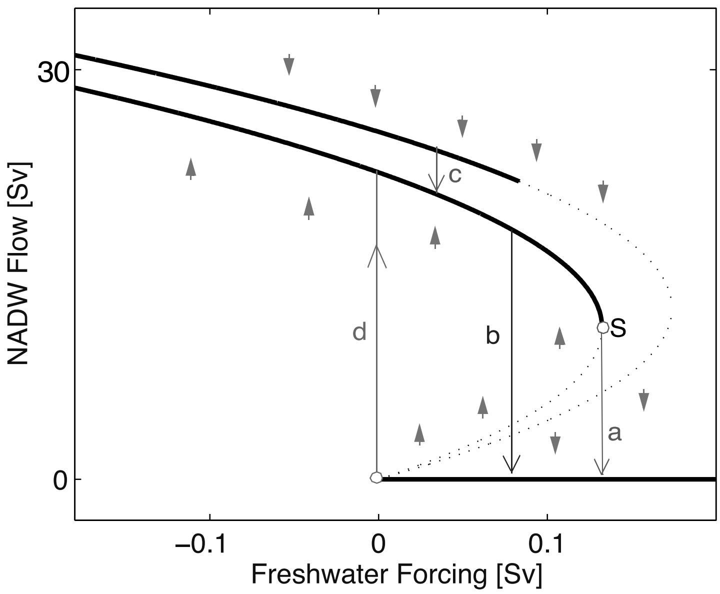


Figure 6. Schematic stability diagram for the North Atlantic Thermohaline Ocean Circulation (THC), with solid black lines indicating stable equilibrium states and dotted black lines unstable states. The rate of ocean flow is measured in Sverdrups (Sv): a flow rate of one Sv = one million cubic meters per second. The x-axis shows the amount of freshwater forcing in Sv relative to present. The y-axis shows the strength of the North Atlantic Deep Water (NADW) formation, also in Sv. Transitions are indicated by arrows: (a) advective spindown, (b) convective shutdown, (c) transition between different convection patterns, and (d) restart of convection. "S" marks the Stommel bifurcation beyond which no North Atlantic Deep Water (NADW) formation can be sustained. Adapted from [Rahmstorf, 2000a, Figure 2].

relative strengths of various **feedback processes** makes it difficult to compare model results. It is often unclear what model features are causing some models to not produce North Atlantic THC collapse when other models are [Schneider, 2004]. These model disagreements make it difficult to assign confident probabilities to the occurrence of a North Atlantic THC collapse and impossible to rule out global warming-induced North Atlantic THC collapse at a high level of confidence [Rahmstorf *et al.*, 2005; Schneider and Thompson, 2000). Models do agree that the future of the North Atlantic THC is dependent upon which mode it currently is in and the **amount and rate of change** to the system both from warming and

formation can be sustained, and it is marked as "S" in Figure 4.

from freshwater increases that result from global warming [Stocker and Schmittner, 1997; Stocker and Marchal, 2000; Higgins *et al.*, 2002].

4.3 Ice Sheets

Our ability to understand the complex dynamics of ice sheets is challenged by the need to model **sub-grid-scale processes** and differentiate between general mechanistic processes and specific geologic **path dependence**. Our knowledge of ice sheets is limited by the lack of knowledge of the small-scale physical processes driving change in the ice sheets and by the lack of models that include that level of small-scale complexity.

Ice sheets tend to gradually accumulate ice, but melt and disintegrate in abrupt and catastrophic processes. There are many internal feedbacks within ice sheets, including the gravity-driven flows of ice, the raising of the top of the sheet to colder altitudes, and the isostatic sinking of the continent under the weight of the ice sheet. The following discussion focuses on the debate regarding the West Antarctic ice sheet (WAIS).

The WAIS accounts for 10% of the volume of the entire Antarctic ice sheet, and complete disintegration of the WAIS would raise sea levels by 4 to 6 meters. For a given climate warming, there is a large range of proposed probabilities and rates of disintegration of the WAIS and there is no scientific consensus [Oppenheimer, 1998; Oppenheimer and Alley, 2005]. Estimates for the rate of disintegration range between 5 and 50 centuries, but shorter time scales have also been proposed for a significant fraction of WAIS disintegration. Limited proxy data suggest that the WAIS might have disintegrated during previous periods that had a global mean temperature only 2-3^{o}C warmer than today [Oppenheimer and Alley, 2004].

A significant challenge in predicting future changes in ice sheets is that current models do not adequately capture the dynamics of ice sheets. No ice sheet model has accurately reproduced the existing ice streams, which are fast-moving rivers of ice and are the main dynamic features of WAIS [Oppenheimer and Alley, 2004]. Further complex interactions in ice sheets include basal lubrication by melt water, ice shelf buttressing of glaciers, and melt-water pools increasing warming from lower albedo. The collapse of one 2,500-km^2 section of the Larsen B Ice Shelf in 2002 was a surprise to many in the community, especially due to its rapidity: the shelf shattered within days [Oppenheimer and Alley, 2004]. It is unclear how the rapidity of the ice shelf collapse translates to the potential speed of collapse of the entire WAIS. **Sub-grid-scale processes** of surface melt-water pools and crevasses were potential drivers in the collapse [Scambos *et al.*, 2000; Oppenheimer and Alley, 2004].

Lack of consideration of the **complex dynamics** of ice sheets (as in all current climate models) seems to result in a significant underestimation of the possible rates of change and loss of ice sheets. Rahmstorf *et al.* [2007] find that observed sea level rise has been faster than predicted by the range of models in the most recent IPCC global assessment [Solomon *et al.*, 2007]. As with the North Atlantic

THC, improving our understanding and ability to model ice sheets is crucial to our ability to project potential impacts of global warming.

4.4 Vegetation Cover Change

Similar to the North Atlantic THC, vegetation cover can also have **multiple equilibria** and **hysteresis**, which result from couplings between the atmosphere and biosphere. Regions with **multiple equilibria** can change state rapidly following even relatively mild or short-lived perturbations. The dynamics of such systems are **highly dependent upon initial conditions**. Our ability to model such **multiple equilibria** systems is limited by our understanding of the complex atmosphere-biosphere interactions and by our ability to model dynamics from **sub-grid-scale processes**. We discuss a few examples of the coupled climate-vegetation system, focusing at the broadest scales of ecosystem structure and function. Different processes and characteristics at other biological scales can also have **multiple equilibria**.

Most regions of the world reach a single equilibrium regardless of initial conditions. Boreal forest-tundra boundary ecosystems provide an example of land cover with only a single stable equilibrium under current climate conditions. As discussed previously, through changes in surface albedo, increases in forests at high latitudes probably cause net regional warming while increases in tundra or grasslands can cause regional net cooling [Bala *et al.*, 2007; Bonan *et al.*, 1992].[3] These two states potentially could be self-stabilizing once initiated. However, historical observations and model simulations suggest that the boreal system converges to a single stable equilibrium, and does not have multiple stable equilibria [Levis *et al.*, 1999].

Some subtropical regions of the world reach a **different equilibrium depending upon the initial vegetation distributions**, such as in the Amazon Basin and West Africa [Kleidon *et al.*, 2000; Kleidon and Heimann, 1999; Claussen, 1998; Siegel *et al.*, 1995]. Model simulations find dry-desert or wet-vegetation self-reinforcing stable states in the subtropics. Six thousand years ago, the Sahara was heavily vegetated but the region then abruptly underwent desertification as a consequence of a small change in the Earth's orbital geometry [Claussen *et al.*, 1999; deMenocal *et al.*, 2000]. This suggests that during the mid-Holocene[4] (about 5 to 8 thousand years ago) the Sahara used to have the potential for **abrupt and irreversible changes**, even though it currently is in a single, stable equilibrium desert state [Higgins *et al.*, 2002].

In West Africa, the strength of the tropical monsoon influences the vegetation distribution but the monsoon itself depends upon that vegetation [Eltahir, 1996; Wang and Eltahir, 2000b]. Historical evidence in the Sahel region of West Africa

[3]It is important to note that these models still have many limitations, such as in their parameterizations of clouds. In addition, the differences in soil moisture between vegetation types could also change surface albedo.

[4]The Holocene is a geologic time period that began approximately 11,500 years ago after the end of the last glacial period and continues to the present.

suggests two stable equilibria and that vegetation is partly responsible for the low-frequency variability in regional climate and the transitions between equilibrium states [Wang and Eltahir, 2000a; 2000b]. Some models of West Africa suggest that monsoon circulation is sensitive to deforestation, and that such sensitivity is **highly dependent upon the specific location of the change** in vegetation [Zheng and Eltahir, 1997; 1998]. These modeling experiments suggest that relatively small areas of land cover can determine the equilibrium state of the atmosphere-biosphere system of an entire region [Higgins *et al.*, 2002].

It is important to consider that the results from models, including those referenced above, are sensitive to how the model aggregates and parameterizes processes that occur at smaller scales than the grid cells of the simulation. The reader is referred to the discussion in Section 3 of the model parameterization process and the challenges of representing clouds in models. In addition to **sub-grid-scale processes**, the models are also limited in their representation of natural variability and both natural and human ecosystem disturbance. Furthermore, natural ecosystems rarely, if ever, exist at equilibrium for the particular spatial and temporal scale of interest due to **transient responses** [Higgins *et al.*, 2002]. Because of these limitations, it is necessary to test results across a hierarchy of models incorporating different processes at different scales, iteratively comparing models with each other and with available data from different scales [Root and Schneider, 2003].

4.5 Species Extinction

Species extinctions and ecosystem composition changes provide important examples of the challenges of **multiple stressors** interacting with **multiple thresholds**, **path dependence**, and **sub-grid-scale dynamics**. Natural ecosystems often respond **nonlinearly** to external stresses, where a small loss of resilience can cascade into large and surprising changes that can be **difficult or impossible to reverse** [Liu *et al.*, 2007]. Walker and Meyers [2004] analyzed 64 examples of ecosystem changes and found that 40% of the regime shifts were **irreversible**. Communities are subjected to **multiple stressors** in addition to climate change, including habitat fragmentation, atmospheric deposition, invasive species, pests and diseases, and modified fire regimes. The **history** of previous stresses that have strained the resilience of an ecosystem and the **synergy of multiple stresses** both cause nonlinear responses.

Migrations and changes in ecosystem composition during the transitions between glacial and interglacial cycles show changing assemblages of species that do not repeat each cycle, because of random processes, path dependencies, interactions between communities and climate, and different rates of migration and expansion for different species [Bennett, 1997]. Such novel community assemblages have been observed extensively in the pollen records of past plant communities during the most recent deglaciation [Overpeck *et al.*, 1992]. The result for some species is extinction, even though they might have thrived in the same climatic conditions during the previous cycle. The ability to model the dynamics of ecosystem change and species extinction is essential for understanding the impacts of global warming and to prepare adaptation strategies.

4.6 Overshoot Scenarios

A prominent theme in our discussion has been the importance of sub-system **sensitivity to perturbation features**. It is not only the total size of the perturbation that matters, but also the perturbation's rate of change, maximum change, duration of change, and location of change. Moreover, the relative importance of these different features of the perturbation can depend upon the **initial conditions** of the system.

Only a few studies have investigated how the rate and pattern of emissions and the potential change over time can have important impacts on future climate states, such as for highly nonlinear events. Considerable attention has been devoted to **path-dependent** mitigation costs, but few studies have investigated the different climate impacts from various emissions trajectories [O'Neill and Oppenheimer, 2004; Schneider *et al.*, 2007]. Most studies focus almost exclusively on the final stabilization concentrations and not the emissions pathway. O'Neill and Oppenheimer [2004] investigated a range of emissions pathways that reached the same stabilization concentration. They find that some physical and ecological systems (such as coral reefs, the North Atlantic THC, ice sheets, and sea level rise) appear to be **highly sensitive to the specific transient approach** to stabilization. Trajectories that delay emissions reductions or overshoot the final concentration both increase the likelihood of dangerous climatic events.

Schneider and Mastrandrea [2005] investigated the characteristics and impacts of "overshoot" or "peaking" emission scenarios where concentrations peak and then decline, with temperatures following. The relative importance of **different perturbation characteristics** (such as maximum, cumulative amount, and rate of change) depends upon the specific system, its history, and the external conditions. For example, some species cannot survive above a threshold temperature for more than a short time, and thus are highly sensitive to the maximum temperature. Ice sheet dynamics, in contrast, seem to be driven more by cumulative warming over time. The vulnerability of most systems is from a combination of these characteristics, as shown in the North Atlantic THC example, and thus it is not possible to specify a single, isolated threshold value. Schneider and Mastrandrea [2005] find that overshoot profiles can significantly increase the likelihood of exceeding dangerous climate impact **thresholds**, despite having identical equilibrium warming. More research needs to explore further the impact of **perturbation features** on climate systems and system dynamics. Such impact studies will be crucial to determine what emissions pathways avoid dangerous climate change.

5 IMPLICATIONS FOR CLIMATE POLICY ANALYSIS

Thus far, this chapter has focused on how complex dynamics in the climate system limit our understanding of the Earth system and increase the uncertainty in predictions about the impacts of climate change. We now investigate the implications of the complexity of the coupled human and natural systems for climate policy analysis. Current decision-making paradigms that are often used in climate policy analysis have many limitations and produce misleading results that can lead to potentially dangerous environmental consequences. We propose a new decision analysis framework that includes the fundamentally complex nature of the climate system through robust risk assessment.

5.1 Uncertainty Amplified

Even if future GHG emissions were certain, there remains significant uncertainty in our current ability to predict climate change. One representation of such uncertainty is the most recent consensus range for climate sensitivity values: there is a likely (66-90%) chance that the warming for a doubling of atmospheric CO_2 will be between 2^oC and 4.5^oC, but a significant chance (5-17%) of warming above that range [Solomon *et al.*, 2007]. In addition, research has found that there will always be more uncertainty in predictions of abrupt climate change than of gradual climate change [Alley *et al.*, 2003]. However, as discussed previously, despite these uncertainties some well-defined experiments can be defined, such as for the detection and attribution of the impact of human GHG emissions on the current climate.

Climate change policy decisions require consideration not only of climate science but also of the entire coupled human and natural systems. Climate science is one part of a larger integrated assessment process: an end-to-end analysis, with human society and GHG emissions at one end and projected climate change impacts on human and natural systems at the other, which in turn affect society and emissions in a feedback process [Root and Schneider, 2003]. There is a **cascade of uncertainties** in predictions of climate impacts in the coupled human and natural systems, making impacts even harder to predict than climate dynamics alone, as illustrated in Figure 7. The human system contributes to uncertainty through predictions of human GHG emissions and in the impacts of specific climate changes on society.

Coupled human and natural systems have intricate organizational, spatial, and temporal **couplings at a variety of scales** [Liu *et al.*, 2007]. An analysis of the consequences of anthropogenic climate change must include the interactions of the various subsystems of the natural climate system with the various social and economic **subsystems** of human society. Interdisciplinary collaborations and systems thinking are required, as research in isolation might miss key system behaviors [Liu *et al.*, 2007]. For example, there likely will be a negative climate **feedback** of human behavior change in response to observed impacts of climate change, such as

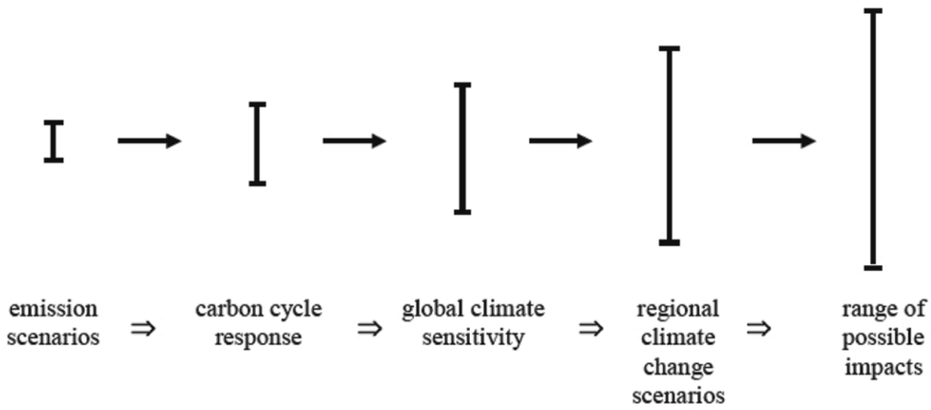


Figure 7. "Cascade" of uncertainties, building from emissions, to concentrations, to global and regional climate change, to climate impacts. Each layer adds an additional source of uncertainty, represented by the widening vertical band. Modified after [Jones, 2000] and the "cascading pyramid of uncertainties" in [Schneider, 1983].

emissions reductions or adaptation planning. However, the strength of such a **feedback** is highly limited by the **inertia** in the climate system that will delay perceived impacts, and the **inertia** in the social and economic systems that will delay the effects of new decisions [Gunderson and Holling, 2001].

In addition to adding uncertainty to predictions of the impacts of climate change, the **complex dynamics** of the coupled human and natural systems also add uncertainty to the predictions of the costs of reducing GHG emissions to avoid climate change (mitigation). The behavior of the world's economic and energy systems are also **complex systems**, with **inertia**, **nonlinearity**, **path dependence**, and much uncertainty.

The complex dynamics of the coupled human and natural systems also limit the ability to plan adaptation policy. For example, climate models are especially uncertain at the specific local and temporal scales necessary for local impact prediction and adaptation planning. As with many **complex dynamics**, the **rate of change** in the climate system is important for the adaptive capacity of natural and human systems. Rapid climate changes reduce the ability of adaptive agents to have knowledge about future changes and the time to marshal the resources to adapt, making faster and less-anticipated climate changes much more costly [Parry *et al.*, 2007; Schneider *et al.*, 2000a; Reilly and Schimmelpfennig, 2000; Yohe *et al.*, 2004].

The lesson of **complexity** is humility in our interactions, predictions, and modifications of the climate system — as well as in claims about high confidence skill in projections beyond a handful of important but still limited examples.

6 CURRENT DECISION-MAKING PARADIGMS

Decisions need to be made within and with respect to all these uncertainties in the coupled human and natural systems, but current methods fail to do so. Most current analyses use integrated assessment models that insufficiently incorporate the **complex dynamics** of the earth system, producing misleading results that have potentially dangerous environmental consequences. Integrated assessment models (IAMs) are coupled system models, where uncertainties in each sub-system cascade to create uncertainty in the final conclusions, as depicted in Figure 8. Most IAMs currently consider only a simple and deterministic climate system. Additionally, most IAMs use smoothly varying climate change and thus have overestimated the human capacity to adapt to rapid climate change [Schneider, Thompson, 2000].

The economic costs of GHG emissions are **highly dependent upon the rates of change** in emissions (which determines the rate of change in the climate system), especially when possible rapid changes are considered in the impact analysis [O'Neill and Oppenheimer, 2002]. Yet few IAMs include this in their analyses, because they mainly focus on aggregate emissions rather than comparing different emission paths, assuming a linear system that we know to be false. The costs of emissions reductions are also highly **path dependent** [O'Neill *et al.*, 2006b].

The climate system includes many phenomena that display fat upper tails of their probability distributions, where high consequence, rare events are more likely than from a symmetrical or normal distribution. Many argue that climate scientists need to further investigate these low probability, high consequence events in the upper tails of the skewed distributions, e.g., [Weitzman, 2007; 2008; Stern *et al.*, 2006; Schellnhuber *et al.*, 2006; Schneider, 2004]. Current IAMs usually do not include these dangerous events, and thus they underestimate the risks of climate change.

Mastrandrea and Schneider [2001] provide a quantitative example of the changes to the conclusions from policy analysis with IAMs that result from including the **complex dynamics** of the climate change. Mastrandrea and Schneider [2001] extended the conventional, smooth IAM analysis from a simple energy economy model (the Dynamic Integrated Climate Economy model, abbreviated DICE[5]) by adding a climate model capable of one type of abrupt change, changes in the strength and possible collapse of the North Atlantic THC. They find a significant impact on the optimal policy calculation, where near-term abatement is increased to reduce the risk of abrupt climate changes. Most previous IAM work does not consider potential abrupt changes, and tends to find economic optimality of negligible near-term abatement. Mastrandrea and Schneider [2001] also found that their results were **highly sensitive** to the discount rate, because there is a threshold in discount rate above which the present value of future damages is so low that even very large enhanced damages in the twenty-second century, when a significant abrupt change such as a THC collapse would be most likely to

[5]For more about the DICE model see: `http://www.econ.yale.edu/~nordhaus/homepage/DICE2007.htm`

occur, do not increase optimal control levels sufficiently to prevent such a collapse. Their research supports the importance of transparency and sensitivity analysis of discount rates in decision-making.

The current dominant framework for climate policy analysis is optimization of IAMs. Mitigation policy is set where the net present value of the costs of reducing emissions is equal to the net present value of the benefits of avoiding climate change. However, the limited predictive capacity of current climate models alone suggests that an optimizing calculation is not feasible in climate change policy analysis. Moreover, the limitations and arbitrary assumptions of IAMs preclude the meaningful use of IAMs in optimization analysis for climate policy decisions and illustrate the need for probabilistic analysis [Roughgarden and Schneider, 1999; Rial, 2004; Mastrandrea and Schneider, 2004]. As discussed previously, IAMs are limited by flawed assumptions, oversimplification of the dynamics of the climate system, and severe uncertainties that are not adequately incorporated into policy analysis.

Given the uncertainty in climate projections, there are some who claim that we ought to wait and observe more climate change to refine our scientific understanding before acting, supported by simplified IAMs that assume mitigation costs are path independent. More rigorous analysis finds that the "wait & see" strategy is more costly and that the prospect of learning does not support the postponement of emissions reductions today [O'Neill *et al.*, 2006a; Yohe *et al.*, 2004]. It will be more expensive to meet the same emissions targets by delaying now and reducing more later. There are three main reasons for this. The first is the **inertia** in the economic and energy system that causes a significant delay before options become available after policy action: action now will open more options later. The second is that the high rates of reductions in the future required under delayed action are estimated to be exponentially higher than a more gradual reduction (path dependent). Lastly, when analyses include the potential low probability, high consequence impact events, the potential costs of delay become significantly higher [O'Neill and Oppenheimer, 2002]. "Wait & see" is a dangerous strategy that implicitly assumes linearity, reversibility, and path-independence — assumptions that do not hold in the climate system.[6]

6.1 Improved Decision Frameworks

A new decision analysis framework is necessary that includes the fundamentally complex nature of the climate system through a probabilistic risk assessment of the full range of outcomes using transparent impact metrics. We discuss potential improvements to the current decision-making framework used in climate policy analysis.

[6]Moreover, due to the complex nature of the climate system, we might not learn much more by having observations until 2050. For example, most of the IPCC climate projections do not diverge substantially until the second half of the century, see [Solomon *et al.*, 2007].

The intrinsic uncertainty and skewed distributions in the climate system necessitate that climate prediction be approached in a probabilistic way, e.g., [Giorgi, 2005; Schneider, 2001; 2002]. Probabilistic risk analysis (PRA) investigates the likelihood of the full range of potential outcomes [Paté-Cornell, 1996]. Current models and emission scenarios represent a limited sample of potential outcomes, hindered by the similarities, historical connections, and overlapping assumptions between different models and scenarios. In addition, experts generally underestimate uncertainty [Tversky and Kahneman, 1974]. Thus, the full range of potential outcomes would be much larger than the current limited set of models and scenarios [Schneider and Kuntz-Duriseti, 2002].

In traditional PRA, engineering systems are broken into independent subsystems (or elementary events) that are then analyzed in isolation, using a combination of historical data, test data, and/or expert opinions. The subsystems are combined probabilistically using a logical function (whether the components are connected in series or parallel) to compute the risk of the entire system [Paté-Cornell, 1996]. **Path-dependency** and **interconnections between numerous sub-systems** would cause the climate system to need an exponentially increasing number of conditional probabilities for each system component given the other components and given the path history. In isolation, specific events or **emergent behavior** can be approximated with a set of probabilities, but an analysis that included several different abrupt events and numerous couplings would not be able to be completed in such a manner. Traditional PRA is hard to apply given all the probability distributions that need to be specified but are unknown in the complex dynamics of the climate system, especially for low probability, high consequence events. The suitability of traditional risk analysis to this situation, as well as how best to approximate the process to enable feasible calculations, are both questions that remain open.

Given the limitations of applying rigorous risk analysis to climate change policy, some advocate that we should use the precautionary principle in this situation, e.g., [Oppenheimer, 2005]. The precautionary principle generally states that if an action or policy might cause severe or irreversible harm to the public or to the environment and there is no scientific consensus that harm would not ensue, then the burden of proof falls on those who would advocate taking the action [Raffensberger and Tickner, 1999]. The precautionary principle prevents the use of the lack of full scientific certainty as a reason for not acting to avoid potential dangerous environmental harm. However, others argue that, although appealing in theory, the precautionary principle is impractical to implement, e.g., [Sunstein, 2002].

Another option is to perform a risk analysis as best as is possible, utilizing a diverse set of metrics in a vulnerability framework and not merely a single damage function. This helps free the decision making from the single optimization constraint by separating the impacts into different metrics, leaving the analysis transparent and enabling decision makers to apply their own personal value system. Examples include the five "reasons for concern" from [McCarthy *et al.*, 2001] or the five "numeraires" from Schneider *et al.* [2000b]. These assessments of key vulnerabilities include the **complex behavior** of the climate system through certain metrics, such as through direct consideration of "large-scale singular events," "extreme weather events," and "risks to unique and threatened systems" [Parry *et al.*, 2007]. Integrated assessments within the framework of vulnerability are the new preferred method, where risk assessment and disaster prevention replace an emphasis on prediction and optimization [Parry *et al.*, 2007; Rial *et al.*, 2004].

7 CONCLUSIONS

1. Complexity is key to understanding the climate system. Complexity is the fundamental nature of the system, not the rare exception.

2. The complexity of the climate system introduces unique challenges to climate science, many of which remain to be addressed. The complex dynamics of the climate system limit our understanding and ability to model the climate system. The complexity of the coupled human and natural systems further amplifies the uncertainty in our predictions of the impacts of human-induced climate change.

3. The climate science community is itself a complex system that, as it develops, is increasingly linking subsystems and breaking corresponding disciplinary boundaries that have long prevented the full exploration of the complexity of the climate system.

4. The study of climate science is limited by the inability to perform controlled, repeated experiments with our planet. Thus, data of past climate behavior are especially important in climate science, serving as our only observations of the system dynamics. However, given data constraints, model simulations (linked with available data) often provide the most useful tool to explore the potential dynamics of the climate system.

5. Despite these many uncertainties and challenges, some well-defined experiments can be constructed. Using the methods of detection and attribution, the international scientific community has come to the consensus that: "the warming of the climate system is unequivocal" and "most of the observed increase in global average temperatures since the mid-20th century is *very likely* due to the observed increase in anthropogenic greenhouse gas concentrations" [Solomon *et al.*, 2007]. Using a variety of historical and proxy data in combination with a diversity of modeling experiments, climate scientists have been able to project the potential impacts of humans' greenhouse gas emissions [Parry *et al.*, 2007]. A truism, likely to persist for decades more, is that some aspects of climate model projections are well established, others demonstrate competing explanations, and yet others are likely to continue to remain speculative.

6. Systems thinking is essential in policy analysis of anthropogenic climate change. Current integrated assessment regimes insufficiently incorporate the complexity of the earth system, producing misleading results that have potentially dangerous environmental consequences. Robust risk assessments must include the potential for surprises, irreversibility, thresholds, rapid change, and low probability, high consequence events. Decisions need to be made within and with respect to all these uncertainties in the coupled human and natural systems, but current methods fail to do so.

7. A new decision analysis framework is necessary that includes the fundamentally complex nature of the climate system through robust risk assessment that utilizes diverse and transparent impact metrics within a broader vulnerability framework. The current framework of cost-benefit optimization is particularly flawed, and has potentially dangerous consequences for climate policy.

The complexity of the climate system produces formative challenges and powerful implications for climate modeling and climate policy analysis. Further interdisciplinary collaboration and the teaching and application of systems thinking are needed for all those working on climate issues — whether they be scientists or politicians. Global change research has made immense progress over the past decades, but much remains to be addressed. Climate science requires systems thinking, improved risk analysis, and, above all, humility. At this critical time of potential dangerous climate change, it is imperative that these lessons are learned and implemented.

BIBLIOGRAPHY

[Alley *et al.*, 2003] R. B. Alley, J. Marotzke, W. D. Nordhaus, J. T. Overpeck, D. M. Peteet, R. A. Pielke, R. T. Pierrehumbert, P. B. Rhines, T. F. Stocker, L. D., Talley, and J. M. Wallace. Abrupt climate change. *Science,* 299, 2005-10, 2003.

[Anand *et al.*, 2003] P. Anand, H. Elderfield, and M. H. Conte. Calibration of Mg/Ca thermometry in planktonic foraminifera from a sediment trap time series. *Paleoceanography,* 18, 1050, 2003.

[Bala *et al.*, 2007] G. Bala, K. Caldeira, M. Wickett, T. J. Phillips, D. B., Lobell, C. Delire, and A. Mirin. Combined climate and carbon-cycle effects of large-scale deforestation. *Proceedings of the National Academy of Sciences of the United States of America,* 104, 6550-6555, 2007.

[Bennett, 1997] K. D. Bennett. *Evolution and Ecology: The Pace of Life.* Cambridge University Press, Cambridge, 1997.

[Berger *et al.*, 1999] A. Berger, X. S. Li, and M. F. Loutre. Modelling northern hemisphere ice volume over the last 3 Ma. *Quaternary Science Reviews,* 18, 1-11, 1999.

[Betts, 2000] R. A. Betts. Offset of the potential carbon sink from boreal forestation by decreases in surface albedo. *Nature*, 408, 187-190, 2000.

[Bonan *et al.*, 1992] G. B. Bonan, D. Pollard, and S. L. Thompson. Effects of boreal forest vegetation on global climate. *Nature,* 359, 716-18, 1992.

[Bond *et al.*, 1997] G. Bond, W. Showers, M. Cheseby, R. Lotti, P. Almasi, P. Demenocal, P. Priore, H. Cullen, I. Hajdas, and G. Bonani. A pervasive millennial-scale cycle in North Atlantic Holocene and glacial climates. *Science,* 278, 1257-66, 1997.

[Broecker, 1997] W. S. Broecker. Thermohaline circulation the Achilles heel of our climate system: will man made CO_2 upset the current balance? *Science,* 278, 1582-8, 1997.

[Broecker *et al.*, 1990] W. S. Broecker, G. Bond, M. Klas, G. Bonani, and W. Wolfli. A salt oscillator in the glacial Atlantic? 1. The concept. *Paleoceanography,* 5, 469-477, 1990.

[Clark *et al.*, 2006] P. U. Clark, D. Archer, D. Pollard, J. D. Blum, J. A. Rial, V. Brovkin, A. C. Mix, N. G. Pisias, and M. Roy. The middle Pleistocene transition: characteristics, mechanisms, and implications for long-term changes in atmospheric pCO_2. *Quaternary Science Reviews,* 25, 3150-3184, 2006.

[Claussen, 1998] M. Claussen. On multiple solutions of the atmosphere-vegetation system in present-day climate. *Global Change Biology,* 4, 549-559, 1998.

[Claussen *et al.*, 1999] M. Claussen, C. Kubatzki, V. Brovkin, A. Ganopolski, P. Hoelzmann, and H. J. Pachur. Simulation of an abrupt change in Saharan vegetation in the mid-Holocene. *Geophysical Research Letters,* 26, 2037-40, 1999.

[Cuffey and Vimeux, 2001] K. M. Cuffey and F. Vimeux. Covariation of carbon dioxide and temperature from the Vostok ice core after deuterium-excess correction. *Nature,* 412, 523-7, 2001.

[De Garidel-Thoron *et al.*, 2005] T. De Garidel-Thoron, Y. Rosenthal, F. Bassinot, and L. Beaufort. Stable sea surface temperatures in the western Pacific warm pool over the past 1.75 million years. *Nature,* 433, 294-8, 2005.

[Demenocal *et al.*, 2000] P. Demenocal, J. Ortiz, T. Guilderson, J. Adkins, M. Sarnthein, L. Baker, and M. Yarusinsky. Abrupt onset and termination of the African Humid Period: rapid climate responses to gradual insolation forcing. *Quaternary Science Reviews,* 19, 347-61, 2000.

[Edwards, 1996] P. N. Edwards. Global comprehensive models in politics and policymaking. *Climatic Changes,* 32, 149-161, 1996.

[Eltahir, 1996] E. A. B. Eltahir. Role of vegetation in sustaining large-scale atmospheric circulations in the tropics. *Journal of Geophysical Research D: Atmospheres,* 101, 4255-4268, 1996.

[EPICA, 2004] EPICA. Eight glacial cycles from an Antarctic ice core. *Nature,* 429, 623-628, 2004.

[Field *et al.*, 2007] C. B. Field, D. B. Lobell, H. A. Peters, and N. R. Chiariello. Feedbacks of Terrestrial Ecosystems to Climate Change. *Annual Review of Environment and Resources,* 32, 7.1–7.29, 2007.

[Friedlingstein *et al.*, 2006] P. Friedlingstein, P. Cox, R. Betts, L. Bopp, W. von Bloh, V. Brovkin, P. Cadule, S. Doney, M. Eby, I. Fung, G. Bala, J. John, C. Jones, F. Joos, T. Kato, M. Kawamiya, W. Knorr, K. Lindsay, H. D. Matthews, T. Raddatz, P. Rayner, C. Reick, E. Roeckner, K. G. Schnitzler, R. Schnur, K. Strassmann, A. J. Weaver, C. Yoshikawa, and N. Zeng. Climate-carbon cycle feedback analysis: Results from the (CMIP)-M-4 model intercomparison. *Journal of Climate,* 19, 3337-3353, 2006.

[Gibbard *et al.*, 2005] S. Gibbard, K. Caldeira, G. Bala, T. J. Phillips, and M. Wickett. Climate effects of global land cover change. *Geophysical Research Letters,* 32, L23705, 2005.

[Giorgi, 2005] F. Giorgi. Climate change prediction. *Climatic Change,* 73, 239-65, 2005.

[Goldewijk, 2001] K. K. Goldewijk. Estimating global land use change over the past 300 years: The HYDE Database. *Global Biogeochemical Cycles,* 15, 417-433, 2001.

[Gunderson and Holling, 2001,] L. H. Gunderson and C. S. Holling, eds. *Panarchy: Understanding Transformation in Human and Natural Systems,* Island Press, Washington D.C., 2001.

[Hays *et al.*, 1976] J. D. Hays, J. Imbrie, and N. J. Shackleton. Variations in the earth's orbit: Pacemaker of the ice ages. *Science,* 194, 1121-1132, 1976.

[Hegerl *et al.*, 2007] G. C. Hegerl, F. W. Zwiers, P. Braconnot, N. P. Gillett, Y. Luo, J. A. Marengo Orsini, N. Nicholls, J. E. Penner, and P. A. Stott, Understanding And Attributing Climate Change. In [Solomon, *et al.*, 2007].

[Higgins *et al.*, 2002] P. A. T. Higgins, M. D. Mastrandrea, and S. H. Schneider. Dynamics of climate and ecosystem coupling: Abrupt changes and multiple equilibria. *Philosophical Transactions of the Royal Society of London Series B Biological Sciences,* 357, 647-655, 2002.

[Imbrie *et al.*, 1993] J. Imbrie, A. Berger, E. A. Boyle, S. C. Clemens, A. Duffy, W. R. Howard, G. Kukla, J. Kutzbach, D. G. Martinson, A. Mcintyre, A. C. Mix, B. Molfino, J. J. Morley, L. C. Peterson, N. G. Pisias, W. L. Prell, M. E. Raymo, N. J. Shackleton, and J. R. Toggweiler. On the structure and origin of major glaciation cycles. 2. The 100,000-year cycle. *Paleoceanography,* 8, 699-735, 1993.

[Jones, 2000] R. N. Jones. Managing uncertainty in climate change projections - issues for impact assessment: An editorial comment. *Climatic Change,* 45, 403-419, 2000.

[Jouzel *et al.*, 2007] J. Jouzel, V. Masson-Delmotte, O. Cattani, G. Dreyfus, S. Falourd, G. Hoffmann, B. Minster, J. Nouet, J. M. Barnola, J. Chappellaz, H. Fischer, J. C. Gallet, S. Johnsen, M. Leuenberger, L. Loulergue, D. Luethi, H. Oerter, F. Parrenin, G. Raisbeck, D. Raynaud, A. Schilt, J. Schwander, E. Selmo, R. Souchez, R. Spahni, B. Stauffer, J. P. Steffensen, B. Stenni, T. F. Stocker, J. L. Tison, M. Werner, and E. W. Wolff. Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science,* 317, 793-796, 2007.

[Kleidon *et al.*, 2000] A. Kleidon, K. Fraedrich, and M. Heimann. A green planet versus a desert world: estimating the maximum effect of vegetation on the land surface climate. *Climatic Change,* 44, 471-93, 2000.

[Kleidon and Heimann, 1999] A. Kleidon and M. Heimann. Deep-rooted vegetation, Amazonian deforestation, and climate: results from a modelling study. *Global Ecology and Biogeography,* 8, 397-405, 1999.

[Knight *et al.*, 2007] C. G. Knight, S. H. E. Knight, N. Massey, T. Aina, C. Christensen, D. J. Frame, J. A. Kettleborough, A. Martin, S. Pascoe, B. Sanderson, D. Stainforth, and M. Allen. Association of parameter, software, and hardware variation with large-scale behavior across 57,000 climate models. In *Proceedings of the National Academy of Sciences of the United States of America,* 104, 12259-12264, 2007.

[Lenton *et al.*, 2008] T. M. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber. Tipping elements in the Earth's climate system. *Proceedings of the National Academy of Sciences,* 105, 1786-1793, 2008.

[Levis *et al.*, 1999] S. Levis, J. A. Foley, and D. Pollard.Potential high-latitude vegetation feedbacks on CO_2-induced climate change. *Geophysical Research Letters,* 26, 747-50, 1999.

[Liu *et al.*, 2007] J. Liu, T. Dietz, S. R. Carpenter, C. Folke, M. Alberti, C. L. Redman, S. H. Schneider, E. Ostrom, A. N. Pell, J. Lubchenco, W. W. Taylor, Z. Ouyang.,P. Deadman, T. Kratz, and W. Provencher. Complexity of Coupled Human and Natural Systems, *Science*, 317, 1513-1516, 2007.

[Lorenz, 1968] E. N. Lorenz. Climate determinism. *Meteorological Monograph,* 8, 1-3, 1968.

[Lorenz, 1976] E. N. Lorenz. Nondeterministic theories of climatic change. *Quaternary Research,* 6, 495-506, 1976.

[Lüthi *et al.*, 2008] D. Luthi, M. Le Floch, B. Bereiter, T. Blunier, J. M. Barnola, U. Siegenthaler, D. Raynaud, J. Jouzel, H. Fischer, K. Kawamura, and T. F. Stocker. High-resolution carbon dioxide concentration record 650,000-800,000 years before present. *Nature,* 453, 379-382, 2008.

[Mann *et al.*, 1998] M. E. Mann, R. S. Bradley, and M. K. Hughes. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature,* 392, 779-87, 1998.

[Mann and Jones, 2003] M. E. Mann and P. D. Jones.Global surface temperatures over the past two millennia. *Geophysical Research Letters,* 30, 2003.

[Mastrandrea and Schneider, 2001,] M. D. Mastrandrea and S. H. Schneider. Integrated assessment of abrupt climatic changes. *Climate Policy,* 1, 433-449, 2001.

[Mastrandrea and Schneider, 2004] M. D. Mastrandrea and S. H. Schneider. Probabilistic Integrated Assessment of "Dangerous" Climate Change. *Science,* 304, 571-575, 2004.

[McCarthy *et al.*, 2001] J. J. McCarthy, O. F. Canziani, N. A. Leary, D. J. Dokken, and K. S. White, eds. *Climate Change 2001: Impacts, Adaptation, and Vulnerability, Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change,* Cambridge University Press, Cambridge, 2001.

[Milankovitch, 1941] M. M. Milankovitch. Canon of Insolation and the Ice-Age Problem. Royal Serbian Academy, Belgrade, Translated by the Israel Program for Scientific Translation, Jerusalem, 1969.

[Mitchell *et al.*, 2001] J. F. B. Mitchell, D. J. Karoly, G. C. Hegerl, F. W. Zwiers, M. R. Allen, and J. Marengo. Detection Of Climate Change And Attribution Of Causes. In J. T. Houghton, Y. Ding, Y., D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, C. A. Johnson (eds.). *Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridg, 2001.

[O'Neill and Oppenheimer, 2002] B. C. O'Neill and M. Oppenheimer. Dangerous climate impacts and the Kyoto Protocol. *Science,* 296, 1971-1972, 2002.

[O'Neill and Oppenheimer, 2004] B. C. O'Neill and M. Oppenheimer. Climate change impacts are sensitive to the concentration stabilization path. *Proceedings of the National Academy of Sciences of the United States of America,* 101, 16411-16416, 2004.

[O'Neill *et al.*, 2006a] B. C. O'Neill, P. Crutzen, A. Grübler, M. H. Duong, K. Keller, C. Kolstad, J. Koomey, A. Lange, M. Obersteiner, M. Oppenheimer, W. Pepper, W. Sanderson, M. Schlesinger, N. Treich, A. Ulph, M. Webster, and C. Wilson. Learning and climate change. *Climate Policy,* 6, 585-589, 2006.

[O'Neill *et al.*, 2006b] B. C. O'Neill, M. Oppenheimer, and A. Petsonk. Interim targest and the climate treaty regime. *Climate Policy,* 5, 639-645, 2006.

[Oppenheimer, 1998] M. Oppenheimer. Global warming and the stability of the West Antarctic Ice Sheet. *Nature,* 393, 325-32, 1998.

[Oppenheimer, 2005] M. Oppenheimer. Defining dangerous anthropogenic interference: The role of science, the limits of science. *Risk Analysis,* 25, 1399-1407, 2005.

[Oppenheimer and Alley, 2004] M. Oppenheimer and R. B. Alley. The West Antarctic ice sheet and long term climate policy: An editorial comment. *Climatic Change,* 64, 1-10, 2004.

[Oppenheimer and Alley, 2005] M. Oppenheimer and R. B. Alley. Ice sheets, global warming, and Article 2 of the UNFCCC: An editorial essay. *Climatic Change,* 68, 257-267, 2005.

[Overpeck and Cole, 2006] J. T. Overpeck, J. E. Cole. Abrupt change in earth's climate system. *Annual Review of Environment and Resources,* 31, 1-31, 2006.

[Overpeck *et al.*, 1992] J. T. Overpeck, R. S. Webb, and T. Webb. Mapping Eastern North-American vegetation change of the past 18ka - No analogs and the future. *Geology,* 20, 1071-1074, 1992.

[Pagani *et al.*, 2005] M. Pagani, J. C. Zachos, K. H. Freeman, B. Tipple, and S. Bohaty. Marked decline in atmospheric carbon dioxide concentrations during the Paleogene. *Science,* 309, 600-3, 2005.

[Parry *et al.*, 2007] M. Parry, O. Canziani, J. Palutikof, P. van der Linden, and C. Hanson, eds. *Climate Change 2007: Impacts, Adaptation, and Vulnerability, Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change,* Cambridge University Press, Cambridge, 2007.

[Paté-Cornell, 1996] M. E. Paté-Cornell. Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering and System Safety,* 54, 95-111, 1996.

[Petit *et al.*, 1999] J. R. Petit, J. Jouzel, D. Raynaud, N. I. Barkov, J. M. Barnola, I. Basile, M. Bender, J. Chappellaz, M. Davis, G. Delaygue, M. Delmotte, V. M. Kotlyakov, M. Legrand, V. Y. Lipenkov, C. Lorius, L. Pepin, C. Ritz, E. Saltzman, and M. Stievenard. Climate and atmospheric history of the past 420000 years from the Vostok ice core, Antarctica. *Nature,* 399, 429-36, 1999.

[Raffensberger and Tickner, 1999] C. Raffensberger and J. Tickner, eds. *Protecting Public Health and the Environment: Implementing the Precautionary Principle,* Island Press, Washington D.C., 1999.

[Rahmstorf, 2000a] S. Rahmstorf. The thermohaline ocean circulation: A system with dangerous thresholds? *Climatic Change,* 46, 247-256, 2000.

[Rahmstorf, 2000b] S. Rahmstorf. The thermohaline ocean circulation: A system with dangerous thresholds? An editorial comment. *Climatic Change,* 46, 247-256, 2000.

[Rahmstorf, 2006] S. Rahmstorf. Thermohaline Ocean Circulation. IN ELIAS, S. A. (Ed.) *Encyclopedia of Quaternary Sciences.* Elsevier, Amsterdam, 2006.

[Rahmstorf *et al.*, 2005] S. Rahmstorf, M. Crucifix, A. Ganopolski, H. Goosse, I. Kamenkovich, R. Knutti, G. Lohmann, R. Marsh, L. A. Mysak, Z. M. Wang, and A. J. Weaver. Thermohaline circulation hysteresis: A model intercomparison. *Geophysical Research Letters,* 32, L23605, 2005.

[Rahmstorf *et al.*, 2007] S. Rahmstorf, A. Cazenave, J. A. Church, J. E. Hansen, R. F. Keeling, D. E. Parker, and R. C. J. Somerville. Recent Climate Observations Compared to Projections. *Science,* 316, 709-*, 2007.

[Raymo, 1997] M. E. Raymo. The timing of major climate terminations. *Paleoceanography,* 12, 577-585, 1997.

[Reilly and Schimmelpfennig, 2000] J. Reilly and D. Schimmelpfennig. Irreversibility, uncertainty, and learning: Portraits of adaptation to long-term climate change. *Climatic Change,* 45, 253-278, 2000.

[Rial, 1999] J. A. Rial. Pacemaking the ice ages by frequency modulation of Earth's orbital eccentricity. *Science,* 285, 564-568, 1999.

[Rial, 2004] J. A. Rial. Abrupt climate change: Chaos and order at orbital and millennial scales. *Global and Planetary Change,* 41, 95-109,2004.

[Rial *et al.*, 2004] J. A. Rial, R. A. Pielke, M. Beniston, M. Claussen, J. Canadell, P. Cox, H. Held, N. de Noblet-Ducoudre, R. Prinn, J. F. Reynolds, and J. D. Salas. Nonlinearities, feedbacks and critical thresholds within the Earth's climate system. *Climatic Change,* 65, 11-38, 2004.

[Root *et al.*, 2005] T. L. Root, D. P. Macmynowski, M. D. Mastrandrea, and S. H. Schneider. Human-modified temperatures induce species changes: Joint attribution. *Proceedings of the National Academy of Sciences of the United States of America,* 102, 7465-7469, 2005.

[Root and Schneider, 2003] T. L. Root, and S. H. Schneider. Strategic Cyclical Scaling: Bridging Five Orders of Magnitude Scale Gaps in Climatic and Ecological Studies. In J. Rotmans, D. S. Rothman (eds.). *Scaling Issues in Integrated Assessment.* Swets and Zeitlinger Publishers, Lisse The Netherlands, 2003.

[Roughgarden and Schneider, 1999] T. Roughgarden and S. H. Schneider. Climate change policy: Quantifying uncertainties for damages and optimal carbon taxes. *Energy Policy,* 27, 415-429, 1999.

[Santer *et al.*, 2003] B. D. Santer, T. M. L. Wigley, G. A. Meehl, M. F. Wehner, C. Mears, M. Schabel, M., F. J. Wentz, C. Ammann, J. Arblaster, T. Bettge, T., W. M. Washington, K. E. Taylor, J. S. Boyle, W. Bruggemann, and C. Doutriaux. Influence of Satellite Data Uncertainties on the Detection of Externally Forced Climate Change. *Science,* 300, 1280-1284, 2003.

[Scambos *et al.*, 2000] T. A. Scambos, C. Hulbe, M. Fahnestock, and J. Bohlander. The link between climate warming and break-up of ice shelves in the Antarctic Peninsula. *Journal of Glaciology,* 46, 516-530, 2000.

[Schellnhuber *et al.*, 2006] H. J. Schellnhuber, W. Cramer, N. Nakicenovic, T. Wigley, and G. Yohe, eds. *Avoiding Dangerous Climate Change* Cambridge University Press, Cambridge, 2006.

[Schneider, 1972] S. H. Schneider. Cloudiness as a global climatic feedback mechanism: The effects on the radiation balance and surface temperature of variations in cloudiness. *Journal of the Atmospheric Sciences,* 29, 1413-22, 1972.

[Schneider, 1983] S. H. Schneider. CO_2, climate and society: A brief overview. In R. S. CHEN, E. M. BOULDING, S. H. SCHNEIDER (eds.), *Social Science Research and Climate Change: An Interdisciplinary Appraisal.* D. Reidel Publishing, Dordrecht, 1983.

[Schneider, 1992] S. H. Schneider. Introduction to Climate Modeling. In K. E. TRENBERTH (ed.), *Climate System Modeling.* Cambridge University Press, Cambridge, 1992.

[Schneider, 1997] S. H. Schneider. Integrated assessment modeling of global climate change: Transparent rational tool for policy making or opaque screen hiding value-laden assumptions? *Environmental Modeling and Assessment,* 2, 229-249, 1997.

[Schneider, 2001] S. H. Schneider. What is 'dangerous' climate change? *Nature,* 411, 17-19, 2001.

[Schneider, 2002] S. H. Schneider. Can we estimate the likelihood of climatic changes at 2100?: An editorial comment: Why this editorial? *Climatic Change,* 52, 441-451, 2002.

[Schneider, 2004] S. H. Schneider. Abrupt non-linear climate change, irreversibility and surprise. *Global Environmental Change,* 14, 245-258, 2004.

[Schneider and Dickinson, 1976] S. H. Schneider, R. E. Dickinson. Parameterization of fractional cloud amounts in climate models: The importance of modeling multiple reflections. *Journal of Applied Meteorology,* 15, 1050-1056, 1976.

[Schneider *et al.*, 2000a] S. H. Schneider, W. E. Easterling, and L. O. Mearns. Adaptation: Sensitivity to natural variability, agent assumptions and dynamic climate changes. *Climatic Change,* 45, 203-221, 2000.

[Schneider *et al.*, 2000b] S. H. Schneider, K. Kuntz-Duriseti, and C. Azar. Costing nonlinearities, surprises, and irreversible events. *Pacific and Asian Journal of Energy,* 10, 81-106, 2000.

[Schneider and Kuntz-Duriseti, 2002] S. H. Schneider and K. Kuntz-Duriseti. Uncertainty and Climate Change Policy. In S. H. Schneider, A. Rosencranz, J. O. Niles (eds.) *Climate Change Policy: A Survey.* Island Press, Washington D.C., 2002.

[Schneider and Mastrandrea, 2005] S. H. Schneider ad M. D. Mastrandrea. Probabilistic assessment "dangerous" climate change and emissions pathways. *Proceedings of the National Academy of Sciences of the United States of America,* 102, 15728-15735, 2005.

[Schneider *et al.*, 2007] S. H. Schneider, S. Semenov, A. Patwardhan, I. Burton, C. H. D. Magadza, M. Oppenheimer, A. B. Pittock, A. Rahman, J. B. Smith, A. Suarez, F. Yamin, J. Corfee-Morlot, A. Finkel, H. M. Füssel, K. Keller, K., D. P. Macmynowski, M. D. Mastrandrea, and A. Todorov. Assessing Key Vulnerabilities And The Risk From Climate Change. In M. Parry, O. Canziani, J. Palutikof, P. van der Linden, C. Hanson (eds.), *Climate Change 2007: Impacts, Adaptation, and Vulnerability, Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, 2007.

[Schneider *et al.*, 2009] S. H. Schneider, A. Rosencranz, and A. M. D. Mastrandrea. *Climate Change Science and Policy,* Island Press, Washington D.C., 2009.

[Schneider and Thompson, 1979] S. H. Schneider and S. L. Thompson. Ice ages and orbital variations: some simple theory and modeling. *Quaternary Research,* 12, 188-203, 1979.

[Schneider and Thompson, 2000] S. H. Schneider and S. L. Thompson. A Simple Climate Model Used In Economic Studies Of Global Change. In S. J. Decanio, R. B. Howarth, A. H. Sanstad, S. H. Schneider, S. L. Thompson (eds.), *New Directions in the Economics and Integrated Assessment of Global Climate Change.* The Pew Center on Global Climate Change, Washington D.C., 2000.

[Siegel *et al.*, 1995] E. Siegel, H. Dowlatabadi, and M. J. Small. A probabilistic model of ecosystem prevalence. *Journal of Biogeography,* 22, 875-879, 1995.

[Soden and Held, 2006] B. J. Soden and I. M. Held. An assessment of climate feedbacks in coupled ocean-atmosphere models. *Journal of Climate,* 19, 3354-3360, 2006.

[Solomon *et al.*, 2007] S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. M. B. Tignor, and H. L. Miller, eds. *Climate Change 2007: The Scientific Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change,* Cambridge, UK, Cambridge University Press, 2007.

[Stern *et al.*, 2006] N. Stern,, S. Peters, V. Bakhshi, A. Bowen, C. Cameron, S. Catovsky, D. Crane, S. Cruickshank, S. Dietz, N. Edmonson, S. Garbett, L. Hamid, G. Hoffman, D. Ingram, B. Jones, N. Patmore, H. Radcliffe, R. Sathiyarajah, M. Stock, C. Taylor, T. Vernon, H. Wanjie, and D. Zenghelis. Stern Review: The Economics of Climate Change. London, HM Treasury, 2006.

[Stocker and Marchal, 2000] T. F. Stocker and O. Marchal. Abrupt climate change in the computer: Is it real? *Proceedings of the National Academy of Sciences of the United States of America,* 97, 1362-1365, 2000.

[Stocker and Schmittner, 1997] T. F. Stocker and A. Schmittner. Influence of CO_2 emission rates on the stability of the thermohaline circulation. *Nature,* 388, 862-5, 1997.

[Stommel, 1961] H. Stommel. Thermohaline convection with two stable regimes of flow. *Tellus,* 13, 224-230, 1961.

[Sunstein, 2002] C. Sunstein. *Risk and Reason,* Cambridge, Cambridge University Press, 2002.

[Tversky and Kahneman, 1974] A. Tversky and D. Kahneman. Judgment under uncertainty: heuristics and biases. Biases in judgments reveal some heuristics of thinking under uncertainty. *Science,* 185, 1124-1131, 1974.

[Vitousek *et al.*, 1997] P. M. Vitousek, H. A. Mooney, J. Lubchenco, and J. M. Melillo. Human domination of Earth's ecosystems. *Science,* 277, 494-499, 1997.

[Von Deimling *et al.*, 2006] T. S. Von Deimling, H. Held, A. Ganopolski, and S. Rahmstorf. Climate sensitivity estimated from ensemble simulations of glacial climate. *Climate Dynamics,* 27, 149-163, 2006.

[Walker and Meyers, 2004] B. Walker and J. A. Meyers. Thresholds in ecological and social-ecological systems: A developing database. *Ecology and Society,* 9, 2004.

[Wang and Eltahir, 2000a] G. L. Wang and E. A. B. Eltahir. Ecosystem dynamics and the Sahel drought. *Geophysical Research Letters,* 27, 795-798, 2000.

[Wang and Eltahir, 2000b] G. L. Wang and E. A. B. Eltahir. Role of vegetation dynamics in enhancing the low-frequency variability of the Sahel rainfall. *Water Resources Research,* 36, 1013-1021, 2000.

[Webb *et al.*, 2006] M. J. Webb, C. A. Senior, D. M. H. Sexton, W. J. Ingram, K. D. Williams, M. A. Ringer, B. J. Mcavaney, R. Colman, B. J. Soden, R. Gudgel, T. Knutson, S. Emori, T. Ogura, Y. Tsushima, N. Andronova, B. Li, I. Musat, S. Bony, and K. E. Taylor. On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Climate Dynamics,* 27, 17-38, 2006.

[Weitzman, 2007] M. L. Weitzman. A review of the Stern Review on the Economics of Climate Change. *Journal of Economic Literature,* 45, 703-724, 2007.

[Weitzman, 2008] M. L. Weitzman. On Modeling and Interpreting the Economics of Catastrophic Climate Change. Available at `http://www.economics.harvard.edu/faculty/weitzman/files/REStatModeling.pdf`.

[Yohe *et al.*, 2004] G. Yohe, N. Andronova, M. Schlesinger. To hedge or not against an uncertain climate future? *Science,* 306, 416-417, 2004.

[Zachos *et al.*, 2001] J. Zachos, H. Pagani, L. Sloan, E. Thomas, and K. Billups. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science,* 292, 686-693, 2001.

[Zheng and Eltahir, 1997] X. Zheng and E. A. B. Eltahir. The response to deforestation and desertification in a model of West African monsoons. *Geophysical Research Letters,* 24, 155-158, 1997.

[Zheng and Eltahir, 1998] X. Zheng and E. A. B. Eltahir. The role of vegetation in the dynamics of West African monsoons. *Journal of Climate,* 11, 2078-2096, 1998.

Part VI

Economics

ECONOMIC SYSTEMS

John Foster

1 INTRODUCTION

The core of conventional economics lies in the mathematics of constrained optimisation where a decision-maker is presumed to be faced by a measurable set of opportunities and constraints that are amenable to the discovery of a stable equilibrium solution. In recent years, the presence of other decision-makers has been acknowledged through the application of game theory to discover stable strategic equilibrium solutions. Again, decision-makers are presumed to be able to reach such solutions given available information. The economic literature stemming from these theoretical foundations is now vast. However, it is very striking how little acknowledgement there is of the fact that decision-makers are operating in complex systems which have features that strongly contradict the assumptions made in strategic optimisation analysis. So we have a situation where there exists a body of literature which is true in an abstract, logical sense but false in a concrete sense [Foster, 2005]. So, at the end of the day, instead of an economic science that can help us to understand our economic history and geography and to provide our policy-makers with reliable advice, what we have is complicated mathematics that is impenetrable by the lay person who is, instead, patronised by 'just so' stories designed to 'explain' what the mathematics is all about. Unfortunately these are significantly less entertaining and enlightening than those beautifully crafted by Rudyard Kipling!

In this chapter, the goal is to explain why it is essential to conduct all economic analysis within a complex (adaptive) systems framework and to show why this makes a very significant difference. It is explained what the original purpose of constrained optimisation analysis was when it was imported into economics over a century ago and why what was a sensible and pragmatic application of logic in special circumstances became a general approach to dealing with economic phenomena. It will be argued that acknowledgement that economic systems are both complex and adaptive means that we have to begin by viewing them as *network structures of elements and connections* that are, at base, energetically-driven and, thus, dissipative in nature. As such, they are subject to the laws of thermodynamics and survive and grow only if energy throughput can be increased. However, with economic systems, we have to travel well beyond Ilya Prigogine's [1945] seminal representation of a complex physiochemical system as a free energy processor and even beyond Daniel Brooks and Edward Wiley's [1985] controversial, but very

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

insightful, representation of biological systems as self-organising information processors, relying on a store of knowledge acquired from experience. Sophisticated economic systems are different because they actively seek new knowledge that is artificially constructed in the human imagination and is not just a by-product of experience.

Such thinking in economics was relatively rare prior to the 1990s. Friedrich Hayek [1967], who finally gave up on the company of conventional economists in favour of Chicago lawyers in the 1950s, is a good example. So is Nicholas Georgescu-Roegen [1971], who set aside his interest in high economic theory in order to stress the importance of dealing with economic systems as energy intensive and knowledge-based systems. Both understood, intuitively, that economists have to accept that they are dealing with complex adaptive systems. However, there was little direct connection between their writings and the emerging complex systems literature in the natural sciences. Broadly speaking, what is often referred to as 'heterodox economics' is rich with insights that we can now classify in terms of complex adaptive economic systems behaviour. But the problem with heterodox economics has always been that it could never challenge conventional economics because it is a fragmented collection of critiques and insights which lacks any analytical unity. It is here that the theory of complex adaptive systems can make a real contribution by providing a consistent analytical framework within which many of the best contributions of heterodox economists can be given expression.

An issue that arises immediately is related to what we actually mean by 'complexity'. Some economists have seen this as about mathematical complexity. So many attempts have been made to use complicated dynamical equations, in both differential and discrete form, to capture nonlinearity and temporal delay in economic relationships [Day, 1994]. However, it doesn't take long to realise that mathematical deduction of this kind cannot be used to obtain realistic representations of structure and process in an evolving complex economic system. The nature and role of knowledge precludes the use of such analytical devices in any scientific sense, irrespective of whether they are specified in linear or nonlinear terms. This is, in fact, well understood by many conventional economic theorists and their response has been to make unrealistically strong assumptions concerning the universality of knowledge and the human capacity to process information in order to make their resultant models computable [Velupillai, 2000]. Criticisms of such assumptions, even when made by Herbert Simon, a Nobel Laureate, have gone largely unheeded as economists have tried to create a discipline that has the form of a science, but not the content [Simon, 1972]. Over the past half century, economic theorists have stubbornly held on to their view that economics should be a branch of decision theory that involves optimizing choices along artificially smooth and conveniently specified production and utility functions, subject to constraints.

When we depart from this analytically convenient, but illusory, context we do not enter a scientific domain that is very easy to deal with — it is clearly much more difficult terrain than that confronted in physics, chemistry and biology. We are immediately confronted with very significant philosophical and methodological questions:

1. What *should the fundamental unit of analysis be in economics — the individual or some collective, networked entity or should this be variable depending on context?* The complex systems perspective leads us quickly to the conclusion that a sharp distinction between economics and sociology cannot be maintained, either philosophically or scientifically.

2. *How do economic systems emerge*? The conventional economist's inability to deal with emergence means that an understanding of the actual process of economic evolution and the associated growth trajectory is not possible. From a complex systems perspective, emergence can be predictable in the presence of appropriate energy sources and a cognitive capacity to acquire and apply knowledge in creative ways.

3. *If we cannot use conventional mathematics how do we proceed analytically?* Economic evolution involves constantly changing network structures. Sometimes this happens in a slow and steady way while, at other times, it is abrupt and destabilizing. So, economic science has to be done in a fundamentally different way to the conventional approach. Theories cannot just be drawn from the air through armchair contemplation, they have to be discovered in a rigorous way through the careful and systematic acquisition of knowledge. An alternative to mathematical logic has to be used in order to discover theoretical knowledge that is firmly connected to real, observed experience. The availability of large amounts of computing power can now allow us to simulate the behaviour of interacting, interdependent agents in exhaustive ways and then examine the resultant trajectories of artificially generated economic variables to see if they correspond with what we observe in reality. Thus, the generation of theoretical knowledge becomes an *algorithmic* exercise and a range of methodological issues concerning calibration and other ways of verifying the validity of theories arise.

4. *What are the boundary constraints on economic evolution and how can they be captured theoretically?* Brooks and Wiley [1986] stressed that, in biological evolution, the most important boundaries on self-organisational development are historical. This is also true in economic evolution. Well-functioning, evolving economic structures depend upon reliable internal networks that are durable in history. Much of economic development consists of expanding these networks in ever more complex ways. However, network structure, once created, can begin to place severe limitations on the adaptiveness of economic entities, such as firms. The formation of external network connections, through trade and contracting, is also important and this depends, crucially, on the existence of reliable institutional rules. So both internal (organisational) rules that determine network structure and the external (institutional) rules adhered to by the economic entity are crucial. But both kinds

of rules can have a Jekyll and Hyde character, transforming from life-giving facilitators to strangling constraints in the course of time.

5. *How can we capture historicalness in theory?* An essential component of theorizing is to understand the historical context in which a theoretical proposition is being made. Without controlling for historical uniqueness, no general theoretical principles can be identified and isolated. So, if a theory developed by agent-based simulation is to be calibrated on historical data, it is essential that the key external and internal rules that prevail over the period of interest are identified and built into the artificial structure within which simulations are conducted. This can only be done through in depth historical and statistical studies that can lead to the discovery of key rules (which may well be found to be emergent or in the process of abandonment) in operation.

6. *Can economic theory and economic history be separated?* Conventional economists often pretend that they can be separated and, indeed, argue that they must be separated for economic science to be conducted. However, as noted, useful theory cannot be developed in a historical vacuum, as is the case with most of conventional economic theorising. Useful theory offers a general representation of historical reality, not something independent of that reality. And, because the latter involves interacting complex systems, the only theories that are admissible are those that do not breach energetic laws and observed cognitive capabilities. In theorizing, we must understand fully the physical, psychological and sociological sciences relevant to the economic decision-making that we are trying to understand. Philosophical and methodological conundrums arise when we explore this terrain - these are rarely addressed in conventional economics.

In what follows, these questions will be looked at in turn and answers given that demonstrate that economic complexity, although homologous with biological and physiochemical complexity, is different in some crucial respects. It is only by appreciating these differences that we can obtain a full understanding of why the remarkable economic achievements of the human race led to the emergence of serious ecological and environmental difficulties over the past two centuries. Equally, it is only by appreciating these differences that effective solutions to these problems can be obtained. In answering these questions, there will be no attempt to offer a comprehensive review of the literature on the application of complex systems science in economics. Readers are referred to Markose [2005] and Beinhocker [2006] for reviews.

2 WHAT SHOULD THE FUNDAMENTAL UNIT OF ANALYSIS BE IN ECONOMICS — THE INDIVIDUAL OR SOME COLLECTIVE, NETWORKED ENTITY OR SHOULD THIS BE VARIABLE DEPENDING ON CONTEXT?

On the face of it, this seems a rather odd question to start with, yet it is absolutely fundamental. Modern economics has been primarily concerned with the analytics of constrained optimisation of individual economic agents. Recently, this has become somewhat more sophisticated in the sense that micro-economists have conducted such optimization in strategic contexts using game theory. However, the interactions dealt with remain very limited: usually there are two or, with some strong simplifying assumptions, three agents. It remains the case that equilibrium solutions are sought: but those of Nash, rather than Pareto. Because analytical mathematics can only be done when structure remains fixed, i.e., when there is no emergence, no structural transformation and no terminal decline, such an approach cannot deal with economic evolution in any realistic manner. In order to obtain point equilibrium solutions, systems are usually presumed to be closed, static and linear. Relaxation of any one of these presumptions immediately results in no analytical solution that can be represented by a point equilibrium. At best, multiple equilibria exist or there is an equilibrium region, characterized by chaotic dynamics, that is only an equilibrium in the sense that boundaries on possible outcomes can be discerned. So it is frequently the case that there is no deterministic solution that is interpretable analytically.

Because it is individuals that are presumed to optimize in neoclassical economics, macroeconomics can only be conducted in a consistent way by presuming it is a summation of microeconomic behaviour. However, this can only be done by presuming the existence of homogenous (identical) agents in order that aggregation of behaviour can occur. So we have the remarkable presumption that the behaviour of the economic system is a scaled up version of the behaviour of one 'representative' micro-agent. Although neoclassical economic theory is, ostensibly, about individual behaviour, this kind of macroeconomics is much more reminiscent of the old socialist planning model, *circa* 1950, so strongly critiqued by Friedrich Hayek, involving one imagined collective being. The patent absurdity of such an approach is a by-product of an unerring quest to find mathematical models with equilibrium solutions, at both the microeconomic and macroeconomic levels of inquiry. The 'existence of equilibrium' becomes of greater concern than the usefulness of such analysis in economic science. And it is not very useful because the real world is characterized by non-equilibrium (historical) processes which can never be captured, in any general sense, by equilibria computed by constrained optimization algorithms.

This is an unbridgeable divide; nonetheless, the pretence that it can be crossed is contained in *ad hoc* 'disequilibrium adjustment' hypotheses, loosely associated with 'transaction costs' and/or 'information costs,' that are used to translate timeless equilibrium hypotheses, drawn from optimization theory, into specifications that can address historical data. This is a serious methodological error but it is widely accepted in applied economics. It is clear that the introduction of 'transactions costs' and 'information costs' is deliberately *ad hoc* because any formal treatment of computational costliness in the face of complexity [Albin, 1998] and undecidability [Arthur, 1999] makes it clear that the mathematics of constrained optimization cannot be used to capture economic behaviour except as a very special approximation. Furthermore, Epstein and Axtell [1996] used agent-based simulation modelling to show that optimising economic agents, who face incomplete information, generally fail to attain any equilibrium, never mind an optimal one.

Despite all this, the notion that people try to be logical in their behaviour is still a good starting point for the analysis of human behaviour in the social sciences and one that the famous English economist Alfred Marshall suggested a century ago would be a good basis for economic analysis since economics is about doing things better and finding new ways to accumulate income and wealth. But Marshall recognized fully that optimization is not a way to represent human behaviour in any general sense (see [Foster, 1993]). Human experience is principally about emotional, learning-by-experience and routine behaviour, punctuated by the application of logic and related experimentation. So it is truly remarkable, from a history of science perspective, that economists could come to think that behaviour can be understood *generally* as a constrained optimization problem. This has been noted a number of times, but perhaps most provocatively by Nicholas Kaldor [1972]. He stressed that neoclassical economics, although its rhetoric suggests that it is constructed from quite reasonable axioms concerning individual behaviour, does not, and cannot, specify any particular unit of analysis. Since constrained optimization is a theory of outcomes, it says nothing about processes in historical time, so the unit can be anything (or nothing) you like: an individual, a firm an economy. . . .take your pick, it doesn't matter when we are in a timeless world where everything can change notionally but nothing can change really.[1]

Evolutionary economists, particularly since the seminal work of Nelson and Winter [1982], have formed a different view of what the appropriate unit of analysis should be in economics [Dopfer, 2005]. Organisations, particularly firms, are viewed as the key decision units in the economy and the systemic connections within and between organizations are also viewed as important. Such units do not optimize across all their activities because it is infeasible for them to do so, instead, they tend to 'satisfice' in the manner suggested by Herbert Simon, given their cognitive limitations in the face of the dual challenges of uncertainty with regard to the future and the historical lock-in of networked systems. Efficiency improvement and productivity growth are viewed as being achieved, not through universal rational choice but through processes of competitive selection whereby

[1] At the Santa Fe Institute, which has been at the forefront of developments in complexity science, there has been a reluctance to admit that neoclassical economics has to be rejected in this fundamental sense. For example, Blume and Durlauf [2005], in dedicating their book to Kenneth Arrow, state that "the models here presented do no represent any sort of rejection of neoclassical economics" when, in fact, some of the authors in the volume do just this.

the firms that generate most value-added, either by luck or good judgment, come to dominate. This process, in turn, cannot occur unless there is considerable variety in the knowledge and skills that exist and can be drawn upon by entrepreneurs and innovators [Metcalfe, 2002].

Sometimes, evolutionary economists use analogies with Darwinian (or Lamarkian) evolutionary biology to explain how competition works.[2] However, it is mistaken to take this too far because there is no meaningful way of distinguishing genotypes and phenotypes in the economic domain nor is there a unit of selection as firmly structured as a gene. But this does not mean that competitive selection does not occur, nor does it mean that we cannot model it. When innovation occurs, entrepreneurs create new processes, products and organizational arrangements. These are not random processes: there are goals that are aspired to and strategies that are planned to get there, but these are rarely confined to an individual, they are mostly the product of groups of individuals connected within some kind of organisational network, ranging from a few relatives in a family to thousands of employees in a global corporation. So the variety upon which selection works involves self-organised islands of cooperative behaviour and shared knowledge. Of course, because it is not possible to fully understand the strategies and plans of those on other islands, miscalculations are common and these can result in both unanticipated obstacles and opportunities.

In the 1990s, evolutionary economists came to realize that competitive selection could not be the whole story of economic evolution and began to consider the nature of the self-organisational processes that underpin the formation of the variety upon which selection works. By this time, this had already become a much discussed and controversial topic in both the physical and biological sciences [Depew and Weber, 1996]; Foster [1997] and Witt [1997] argued that self-organisation was relatively more important in economic processes because knowledge is actively acquired through the novel patterning of perceptual information, it is not just a product of learning by doing and trial and error, as is the case with other mammals.[3] Explicit knowledge sharing and team work are essential before economic structures can be created that yield products of economic value. Potts [2000] and Foster and Metcalfe [2001] argued that what was being proposed had to be set, explicitly, within a complex economic systems framework and this meant that the network structures that they embody have to be analysed. Thus, the unit of analysis became firmly focused on the 'system' which was viewed as any network structure capable of absorbing free energy and exporting entropy in order to service a knowledge structure, embodied in groups of skilled people and in non-human repositories, that could absorb information and materials and export goods and services with economic value.[4]

[2] See [Aldrich, *et al.*, 2008] for an extreme 'Universal Darwinist' position along these lines.

[3] Cordes [2007] provides a useful historical critique of the over zealous use of biological analogies in economics, arguing for the 'continuity hypothesis' which acknowledges the biological determinants of economic behavior but insists upon the inclusion of additional cultural and institutional dimensions to properly understand such behavior and the system within which it occurs.

[4] This view of complex systems focuses upon connective relations and has a direct lineage back

However, when the unit of analysis is as broad as a 'system' it becomes difficult to engage in analytical or empirical economics beyond case studies. Dopfer, Foster and Potts [2004] tried to address this by providing some general principles that can be applied by thinking of complex systems as composed of bundles of interconnected rules. A particular kind of rule, a 'meso' rule (an institutions, law, norm, convention, etc) is viewed as the core unit in complex economic systems, somewhat like Richard Dawkins' 'meme' but from a quite different self-organisational perspective. Such rules facilitate a vast diversity of microeconomic behaviour and the outcome of all the consequent economic activity results in aggregations of value at the macro level. The goal of this approach is to provide an analytical perspective that recognises, explicitly, that we are dealing with interconnected complex systems as incomplete networks of rules which facilitate individual creativity, imagination and logic in the production, distribution and consumption of goods and services [Dopfer, Potts, 2007].

So, from this perspective, the fundamental unit of analysis is the meso rule, not an individual, firm or other organizational entity. To identify these rules and the extent to which their population of adherents is growing or declining requires detailed historical study over whatever period of time is of interest. It is often the case that the spatial dimension will also be important (see [Martin, Sunley, 2007]). Such study is unusual in mainstream economics but it is prevalent in evolutionary economics, where institutions have always received much more attention. However, this raised a number of difficult issues concerning the most appropriate methodology to adopt. It is fair to say that this debate continues and has some distance to run (see [Foster and Potts, 2009]).

3 HOW DO ECONOMIC SYSTEMS EMERGE?

How economic systems emerge is both a fundamentally important question and a very difficult one to answer. Conventional economists cannot even ask this question within their closed, fixed rule-structure framework of analysis. However, neo-Schumpeterian evolutionary economists do not fare much better because their models are largely concerned with the way in which innovation diffusion and the forces of competition affect the process of economic development. This has been pointed out most strongly by neo-Austrian evolutionary economists, such as Witt [2007]. They argue that new ideas come from the imagination and are implemented by entrepreneurs who also use imagination to conjure up the possible profits that might be earned by translating an invention into an innovation with a commercial application. Such activity mostly lies in the domain of uncertainty, where the number of outcomes is unknown, as are the probabilities that can be assigned to those that are known. So the 'rational calculation in the face of risk' approach favoured by neoclassical economists is rarely relevant beyond hypothetical strategic planning exercises. This is a major difficulty for economics because all wealth

to Hayek [1967] and his vision of 'spontaneous order'.

and income ultimately derives from decisions taken in conditions of uncertainty, even though they are amplified through an innovation diffusion process and the efficiency of their application is enhanced, wherever possible, by rational decision-making.

There is a temptation to argue that the analysis of emergence lies outside economics, in the encompassing domains of psychology, sociology and cultural studies. However, this tacitly externalist position is itself undermined by the historical emergence of the economic sector: from ancient times, where little separable economic activity occurred, human societies have been differentiating their activities until today there have emerged huge economic organisations on a global scale that, ironically, threaten the reverse absorption of social and cultural life into the economic. Thus, although an understanding of these encompassing domains is clearly important, at the least the economic system and the behavioural rules that operate in it clearly feed back into behaviour in these non-economic domains. So we have interconnected economic and non-economic systems with nonlinear features. There was a time in human history when cultural and social barriers made it very difficult to innovate and profit from it. Today, entrepreneurship is an accepted practice because of the emergence of an understanding that such behaviour is crucial for the operation of the economic system.

The behavioural rules adhered to in the economic system have been proven to be of value and there has been feedback into social and cultural attitudes. When, a century ago, Joseph Schumpeter referred to entrepreneurs as 'heroes' he was opposing a culture which regarded them as social outsiders engaged in opportunism and greed. For example, much of anti-Semitism was driven by the fact that the Jewish community, lacking longstanding claims to land and property, had become adept entrepreneurs and were resented for their success. Similarly, in Nineteenth Century England the aristocracy looked down on those that relied on 'commerce' for their accumulated wealth and, in turn, the working classes viewed them as exploiters when they enjoyed the fruits of their entrepreneurial talents. Today, inequality in remuneration is widely accepted and entrepreneurs are admired and imitated.

When we look at modern society we see a massive transformation in our relations with others: we deal predominantly with anonymous individuals in trading and contracting. The tight networks of kinship and community in the medieval village have been replaced with open networks governed by widely accepted meso rules. The emergence of more and more complex systems has involved the expansion of network structures that facilitate the increased use of human and non-human energy to produce increasing numbers of goods and services. We live in vast multi-layered city communities that function because there is enough value added to make adherence to meso rules worthwhile.

The core meso rules of family and immediate community have not died out but have become secondary to adherence to rules which have large scale acceptance and enable profitable connections to be made with previously unidentified individuals and groups. Economic success brings freedom to divorce a disliked partner, to live independently from a disliked family, to change communities, etc. Old network links can be broken and reformed. Thus, new meso rules and associated forms of value adding can emerge without resistance from a conservative family or community. The habitual and routine character of behaviour in small closed networks is replaced by open networks within which innovation and entrepreneurship can flourish. This is what constitutes the secular modern society, uninhibited by moral and spiritual dogma yet adherent to a set of ethical rules that makes the functioning of civil society possible.

These rules are built upon mutually understood economic principles: respect for property rights, respect for contractual agreements, reliability in trade and exchange. When the majority of the population adheres to these rules and the institutions that reflect them, we get value adding and, in turn, these successful rules start to become applied in the social sphere as well. The emergence of successful economic systems leads to the emergence of new kinds of social systems. However, history does matter and many of the rules of secular society still reflect some of the old religious commandments. 'Thou shalt not kill' and 'thou shalt not steal' remain so in the secular world while 'thou shalt not commit adultery' is also upheld but this no longer involves the death penalty, instead becoming a matter of breach of contract and subject to compensation claims.

The emergence of the sophisticated and highly complex economic systems that we now live within has involved socioeconomic feedback whereby creative ideas concerning new technologies and new organisational structures have been sanctioned by the adoption of new socio-cultural meso rules. Socio-cultural meso-rules emerge when there are collective benefits from their widespread adoption but, inevitably, this is opposed by vested interests adhering to existing meso rules. Thus, as Joseph Schumpeter explained, change is a matter of "creative destruction," whereby the emergent displaces the obsolete. These changes have often been characterised by political conflicts and the emergence of democratic systems has been fundamentally important in breaking the grip of conservative forces. No one has written more vividly about the emergence of 'spontaneous order' in the economic system and the importance of political freedom for meso rule adaptation than Friedrich Hayek. However, despite the award of a Nobel Prize in Economics, his treatment of emergence in a complex economic system (see [Hayek, 1967]) has been largely ignored by the economics profession because it cannot be formalised in mathematics.

It is clear that economic growth and development stem from emergent economic systems and that these arise because of the emotional and aspirational dispositions of individuals and groups, not narrowly construed constrained optimisation. Because this must be the case, economists cannot separate their analysis from sociology, psychology and politics. A viable economics is one that fully integrates the relevant aspects of these disciplines. The meso rules that constitute the core of the economic system arise outside the province of both optimising and routine economic behaviour. They begin in the quest for novelty which is a subject of psychology, they spread to become collectively adopted through sociological interactions and their ultimate dominance is challenged by political forces. Again, it was Joseph Schumpeter in *Capitalism, Socialism and Democracy* that made this case strongly, only to be ignored by economists anxious to distance themselves from the other social sciences that were deemed to be 'unscientific'. And, indeed, this poses a dilemma. How do we proceed to be 'scientific' in doing interdisciplinary social science? American institutional economists have been trying to answer this question for at least a century following Thorstein Veblen's famous article in 1898 asking "[why] is economics not an evolutionary science?" but no alternative science has emerged. So, armed with complex system theory, can we make scientific progress?

4 IF WE CANNOT USE CONVENTIONAL MATHEMATICS HOW DO WE PROCEED ANALYTICALLY?

It has been argued that constrained optimisation is at the core of conventional economics and this has been expressed in analytical mathematics which is inappropriate to understand how and why economic systems develop and evolve. In the unconventional field of evolutionary economics, there has never been a strong tradition of using mathematics and econometrics, beyond the investigation of special questions such as the parametric structure of innovation diffusion curves. The reason for this is clear: when significant structural change is present, conventional mathematics and associated econometric methods are mostly unsuitable for empirical research. Much of evolutionary economics has been focused upon the behaviour of the firm and the industries that they populate. In this context, inductive theorizing has been undertaken using simulation/calibration techniques to explore the outcomes of economic process within firms, between firms and between firms and consumers. Agent-based modelling (ABM) has been an important tool for providing support to analytical propositions in modern evolutionary economics ever since the seminal contribution of Nelson and Winter [1982]. Perhaps the best recent example of research in this tradition is that of Malerba *et al.* [2001] who offer a 'history friendly' methodology. The goal of this methodology is to conduct ABM in contexts that, as accurately as possible, reflect the historical and institutional conditions that existed in the period of time under consideration. Once these are controlled for, a clearer picture of economic behaviour can be discerned.

However, as Werker and Brenner [2004] point out, it is possible to generate a wide range of ABM models that can calibrate on a given set of time series data since there are no formal restrictions that can be placed on all of the chosen parameters. This led them to argue for a critical realist methodology that, somewhat like the history friendly approach, involves considerable historical and case study investigation prior to simulation. This results in models that are quite specific to the firm or industry in question. The stylized representations of complicated historical processes that are obtained are then used for counterfactual experiments, but it is uncommon for researchers in this tradition to draw out general theoretical principles from these simulation exercises.

History friendly modelling is, essentially, about the existence and adoption of rules in economic behaviour. In this sense, it belongs to what Nelson and Winter [1982] referred to as "appreciative theory" concerning the pivotal rules that are observed to operate in economic organizations, particularly firms. Formal theory, of the conventional kind is not used. Instead, replicator dynamics, drawn from evolutionary biology, are applied to model the competitive process. From a complex systems perspective, the problem with this is that it only looks at one side of the evolutionary economic process, namely, selection. How the variety upon which competition works is generated receives too little attention. Related to this is the fact that the supply side is emphasized rather than the demand side [Foster, Potts, 2006].

Furthermore, although Nelson and Winter [1982] emphasized the importance of routines in firms and, thus, the core role that rules play, too little attention is given to the generic rules that facilitate coordination in the wider economy [Dopfer, Potts, 2007]. As previously noted, it was not until the 1990s, that a literature began to develop in evolutionary economics where it was argued that, because the process of variety generation is often markedly non-random, this implies that the generic rules involved are of a quite different character to those envisaged in classical genetics. Dopfer, Foster and Potts [2004] went on to exposit a 'micro-meso-macro' approach in which rule systems (and how they adapt) are the building blocks in complex economic systems. A key implication of this approach is that the variety generation process and associated learning and innovative processes are of prior importance over competitive selection mechanisms in understanding the process of economic evolution, echoing older and less formal institutionalist and neo-Austrian perspectives, but within a modern systems perspective.

The micro-meso-macro framework places rules at the centre of economic analysis. The economic system is viewed as being made of cognitive, emotional, socio-cultural, organizational, technical and institutional rules. The analytic concept of a meso unit is a rule and its population of carriers and, in this sense, the economy is made of meso units. For instance, laptop computers manifest a common collection of meso rules: central programmable processing unit plus separate memory storage, add-on I/O devices (BSB, DVD, ...) and so on, and users of these rules include all the manufacturers of laptops, I/O devices, those creating media and media content for them (movies, games, simulations, ...) plus all the laptop owners who purchase and use all this. Microeconomic analysis is the study of the individual carriers of the rule and their local operations, and macroeconomic analysis is the study of the effects of coordination and change in the meso structure of the whole economy. From this perspective, economic evolution involves the origination, adoption and retention of a novel meso rule in the micro and macro structure of the economy.

The evolutionary micro domain contains meso rule adopters engaged in a myriad of activities (or operations) resulting in a heterogeneous range of processes and products that have economic value. These can be aggregated up to a defined macroeconomic level as the aggregate of operational value yielded by a given set of meso rules. Some meso rules have long lives and operate at the core of the economic structures, such as the rules of markets, property rights, hierarchy or other forms of organization. Others are short lived, coming and going in fads and fashions. At a point in time, the economic activity of an economy is determined by its generic structure of meso rules, some of which are deep and stable, others of which are shallow and passing. Economic evolution is a process that is associated with change in that generic structure. This is abstractly conceived as a three-phase rule-trajectory consisting of: (1) the origination of the novel rule as an innovation, (2) the adoption of that rule into a population of micro agents to form a meso unit, and (3) the ongoing retention and use of the meso rule. New meso rules come into being when an idiosyncratic rule becomes the basis of an innovation and is adopted by others because its application yields significant economic value. Equally, meso rules can also die out when their applications cease to be of value and their populations diminish.

The behavioural heterogeneity that we observe in real systems is due to the presence of a myriad of idiosyncratic ways in which specific agents apply meso rules. A meso rule may be widely adopted in a population with high fidelity and efficacy, but the environments faced by adopters may vary considerably. This results in micro variety that, in addition to providing a very heterogeneous set of goods and services, can yield meso rule adaptations through a process of learning and selection. However, as has been noted, theorizing about the emergence of new and adapted meso rules cannot involve formal mathematical analysis, nor can the resultant economic outcomes be described in terms of formal mathematical solutions. All formal deduction requires structure to be invariant, i.e., all chosen elements and connections must be fixed. As we have noted, in evolutionary economics, this has given rise to a simulation/calibration methodology that allows us to study how heterogeneous agents apply meso rules and shift from one meso rule to another when circumstances dictate that this is worthwhile.

However, once we think in terms of the micro-meso-macro framework, we cannot restrict our simulations to processes of competitive selection (generally represented by replicator dynamics). Evolution also involves expansions and integrations of connected networks, i.e., parallel expansions of order and complexity. This is a process of self-organization which can only be understood by exploring how learning by doing, innovation and adoption occur as the population of meso rule carriers increases. The field of innovation research is already replete with case studies of this kind, see [Rogers, 2003], but simulations, using a general meso rule perspective of the kind proposed by Dopfer and Potts [2007], remain rare. Part of the difficulty lies in the fact that self-organisation processes involve a different set of constraints to those faced by selection processes. Selection processes are constrained by the range of variety upon which they operate — they stop when variety has been eliminated. This is a process that is relatively easy to simulate. Self organisation processes, on the other hand, create variety by forging new connections, resulting in more ordered and complex structures. The main constraint on such a process is historical [Brooks, Wiley, 1986] in that there is a limit to the extent that expansion of order and complexity can occur, given a meso rule, because structures that result from the application of the rule are, necessarily, irreversible to some degree in historical time. So, because self-organisation involves structural change that merges and connects elements to form larger elements, simulation is difficult because the historical constraint itself evolves

Furthermore, as noted, the consequent flows of microeconomic value yield aggregate income/expenditure flows at the macroeconomic level and, in turn, these flows feed back into the decision making processes of individuals. So we have a two-way value flow interaction between the microeconomic and the macroeconomic that has, at its base, an interconnected set of meso rules. Recognition of this two-way process raises important questions concerning the nature of boundary constraints on economic behaviour at high levels of aggregation. In standard Keynesian macroeconomics there exist income-expenditure interconnections that move boundaries endogenously so that, for example, an economy can get stuck in a state of high unemployment of resources that cannot be easily reversed by the action of market mechanisms. Modern macroeconomics, however, is rarely discussed in a complex systems framework but, instead, a simple 'market failure' approach is generally been applied. This abstracts from system complexity and tries to connect with the constrained optimising behaviour of decision-makers, i.e., a familiar theoretical apparatus is preferred. The end result has been a macroeconomics that delivers very few insights that economic policymakers can use in their deliberations.

5 WHAT ARE THE BOUNDARY CONSTRAINTS ON ECONOMIC EVOLUTION AND HOW CAN THEY BE CAPTURED THEORETICALLY?

Economic systems are defined by their boundaries. This is true even though these boundaries are open, as they must be to absorb new energy and information and to export goods and services and entropic wastes. Without boundaries we cannot identify a network structure as a unit of analysis. Although we can examine systems on a case-by-case basis, we can only make general propositions about systems that have similar boundaries. We have, for example, a class of systems that we call 'firms' which have boundaries that are defined by: legal definitions about how they are organized, how they keep financial records, how they treat their employees, how they are taxed, how they are owned; the technological rules that are embodied in their capital and their labour; the organizational rules that are adopted in the productive network; the availability of both human and non-human energy of particular kinds; the availability of information of all kinds relevant to production and distribution.

We can consider sub-groups of firms that organize production and distribution in particular ways that set boundaries on their growth. The ultimate boundary faced by an economic system is the set of meso rules which it has adopted by choice, convention or law. The meso rule provides the continuity and the common interface which is fundamental in an economic system: it constitutes both a boundary and an opportunity. Individual members of an identifiable group of systems can still differ in many ways. They can produce products with very different characteristics and with different degrees of efficiency. For example, there are hundreds of different models of notebook computers but they all share a set of technological meso rules that will only change when some radical technological breakthrough occurs. Particular models of notebooks laptops have very short lives, determined by fashion and changing consumer preferences, while the notebook, as a product category, built upon a set of technological meso rules, has lasted for a considerable period of time. Meso rules determine the general character of the network structure of an economic system, but its specific character depends upon unique micro-rules that arise because of innovative and entrepreneurial behaviour. Very successful micro-rules are adopted by others and they become meso-rules. So, unlike genes in biology, meso-rules can come and go even though they offer the durability necessary for an economic system to survive. Our units of analysis are necessarily provisional in a historical sense. Very few exist in the same form over very long periods.

So, for example, a firm in the mid Nineteenth century, which was most likely to be relatively small and partnership or family based, bears little relation to the giant global corporations of the 21^{st} century. It is a serious mistake to think that same unit of analysis should be used to understand the behaviour of both just because we call them 'firms'. There is, however, a historical connection between them: there has been an evolutionary process at work which has resulted in an expansion of organized complexity and a pushing out of boundaries. Why this occurs remains something of a mystery; it's 'progress' which seems to be about cumulative improvement in the material conditions of humanity. But its unintended consequences have been to increase human population by unparalleled amounts and to throughput energy in quantities that are so vast that the ecological system has been affected in fundamental ways. The boundaries have been pushed out in ways that may not necessarily be 'progressive' in the long term.

Shifting units of analysis, as structural change proceeds, make theorizing difficult. Trying to generalize about the behaviour of a system that is mutating into a different system, either quickly or slowly, is not amenable to mathematical formalization in any deductive sense. Physicists interested in structural transitions, for example, in the shift from normal to laser light as energy input is increased, use synergetics to provide formal representations of the process. However, these are not operationalisable in economics for a range of reasons (see [Foster and Wild, 1996]). Dynamic mathematical models can be promulgated that can track observable macro-variables in transitions, e.g., Benard cell formation is determined by matching up the two different macro fluid dynamical d.e. models operating before and after convective boiling. But such models do not capture the actual processes at work and, thus, are neither useful in a scientific sense nor for forecasting. It follows that we have to think of 'theory' in a different way.

6 HOW CAN WE CAPTURE HISTORICALNESS IN THEORY?

Nelson and Winter [1982] distinguished 'formal' theory from 'appreciative' theory. The latter is built out of careful historical and institutional study rather than from an axiomatic base. As noted, this was developed over two decades to give rise to the 'history-friendly' methodology that has been applied in several industrial settings. Now, although such studies are very useful in understanding the evolution of a complex economic system, the findings are not strictly generalisable, even within the industry in question. In other words, appreciative theorizing seems to be something ontologically different from standard theorizing in economics. It does not offer a skeletal representation of the essential structure and behaviour of a complex system that can be applied to other systems that qualify to be similar units of analysis. However, it has to be said immediately that conventional economic theory does not do this either — it is not a skeletal representation of reality. It is not a simplification of the complex but, rather, a 'simplistic', unworldly representation of some axioms that can be argued to have some relevance to actual behaviour. This offers an illusion of connection with reality, no general representation of actual behaviour, despite assurances to the contrary [Foster, 2005].

The history friendly approach digs too deep to ever produce any theory that can be generalized over many other economic systems. This can only be done by going up to a higher level of inquiry where the unit of analysis is much more expansive. The appropriate level is where there are meso rules that span a significant number of systems. A hypothesis as to how a meso rule affects behaviour can be the basis of a theory that can be applied in a general manner [Dopfer, Potts, 2007]. Theorising is only feasible when structure remains relatively unchanged over time. Thus, we can theorise when meso rules are identified as stable since they constitute the core from which complex behaviour arises. But, as has been noted, meso rules are not like genes, they are much more malleable. There is, at any point, a significant stock of them but they also come and go. Their arrival is associated with a growth in their application which shows up in diffusion curves for products. These can be fairly easy to discover and, as noted, the literature on innovation studies contains a large number of examples. A key theoretical proposition is that micro rules which are of value will spread until they can be classified as meso rules. More difficult to deal with are the meso rules that are abandoned because they are no longer useful in value generation. This can be an abrupt and nonlinear process, with the economic structures that are tied to obsolete meso rules crashing dramatically.

A good example is a 'bubble'. As a successful rule is adopted, a micro-rule becomes a meso rule. We see expansion of its application as more and more people adopt the rule. However, this process need not stop when the rule has reached the limit of its usefulness. If people keep adopting, the rule will continue to operate on, for example, an asset price. The price will drift inexorably beyond what is justified by the 'fundamentals'. Similarly, a firm which has expanded as a product has risen up a diffusion curve will, when faced with saturation, market the product in a way that persuades those who don't really need it to buy. Marketing encourages the adoption of a meso rule. So, in the ill-fated US sub-prime property market of 2007, hard selling property agents persuaded people to buy houses they couldn't afford. Meso rules are comfortable to adopt because you know that many others also adopt them.

Typically, major 'crashes' occur when meso rules have become seriously inconsistent with reality which can happen easily when decision-makers do not understand the complex implications and real value of the new connections that have been forged in an evolving network. So firms invest in new capacity in boom economies and, in so doing, restrict consumption spending. The boom ends unexpectedly and firms are left with excess capacity. In times when monetary growth is permissive, growth can occur for much longer. But, when monetary growth finances wage rises, we get profit constraints, and boosts in spending that result in inflation. In recent times, monetary growth has financed asset price bubbles. Rises in asset prices, in turn, have strengthened consumption. But when the meso rule that 'property can never be a bad investment' breaks down, then the resultant negative wealth effect lowers consumption in dramatic ways. Meso rules drive the growth of systems that exhibit behavioural feedback connections and nonlinear macroeconomic trajectories can only be understood from a complex system perspective.

Econometrics can be useful in understanding diffusional phases of economic evolution but is of little help in understanding why crashes occur. Crashes involve rapid structural change, driven by the fast abandonment of a meso rule. Econometrics cannot be used to test the kinds of theories relevant to such fluctuations of historical data. Such theories have to be embedded in history and, as such, have to be inductively obtained. The most effective form of theorising about history currently available is through the application of agent-based simulations. Foster and Potts [2009] argue for a methodology that begins with careful historical investigation to identify key meso rules and the extent to which their adoption varies over the historical period under investigation. They argue for comprehensive investigations of the statistical properties of all time series variables and the associations between them. These are then matched with identified meso rules and a parsimonious statistical model obtained and estimated over a sub-period of the data that does not contain obvious discontinuous change. The estimated parameters obtained are then used as limiting constraints on an agent-based simulation. A simulation search is conducted until one or more combinations are found that calibrate upon the discontinuous data in the sample. If there is more than one of these, each is inspected to see which matches the qualitative historical evidence concerning meso rules and their movements. The outcome is an inductively obtained theoretical hypothesis concerning an aspect of economic evolution which may or may not be generalisable to other cases.

7 CAN ECONOMIC THEORY AND ECONOMIC HISTORY BE SEPARATED?

In conventional economics, theory and history are regarded as separate domains and yet, oddly, the former is expected to enlighten us about the latter. And, as discussed above, applied economists use unscientific methods to try to connect them. It has been argued here that, given that we are dealing with complex systems, a theory that is not embedded in history is a theory that can tell us little about economic phenomena. The conventional economist is prone to thinking in the 'old physics' way: provided that you can identify the underlying enduring laws, then you can understand the complexity that is observed. Per Bak's [1997] sand pile experiments show that avalanches are no more than the operation of physical laws in a frictional context. Some living systems can also be understood in this way. Complex termite mounds can be understood as the outcome of a few simple rules [Langton, 1989]. Economic systems are different because they are reflexive. For example, physical laws need not be relevant because humans can create artificial structures that negate the effect of a physical law, e.g., by making use of the Bernoulli Law for moving fluids, planes counteract the effect of the Law of Gravity that would otherwise confine them to the ground. In the biological domain, natural selection can also lead to a similar outcome, e.g., birds can also negate the Law of Gravity. But natural selection takes a very long time to work, whereas human imagination and design can create structures very rapidly. However, in both of these cases, the Law of Gravity presented a physical constraint that was overcome through the operation of a historical process whereby the organised complexity of dissipative structures evolved.

What this means is that the formation (and demise) of structures over historical time is as much a product of history as it is the outcome of physical laws. Defying such laws requires the creative use of energy to build physical and organisational structures. But, equally, physical laws can operate with a vengeance when there is system breakdown — birds and planes fail to get off the ground and return there very rapidly without sufficient energy. It follows that, to understand economic phenomena unfolding over time, fundamental laws can be secondary to the particular pseudo-laws (meso rules) that emerge in history, but may re-emerge as dominant when these rules fail. It is, then, meso rules — their emergence, operation and failure - that are the appropriate subject of evolutionary economics. The axioms of neoclassical economics are pseudo-laws because they cannot explain behaviour in the face of historical constraints and the uncertainty of the future. The behaviour of the entrepreneur and the innovator cannot be explained using these axioms. Their behaviour is governed by meso-rules that may or may not be enduring in history. In order to understand our economy and where it is going, we must understand what these meso rules are and their relative importance. In this regard, there is little doubt that the writings of a good modern economic historian contain more useful insights into an economy than the vast bulk of conventional economic analysis contained in journal articles. But, paradoxically, economic historians are a fast disappearing species and interesting questions arise as to why this is so.

Complex systems analysis tells us that we cannot keep theory and history separate. To do so at the economic level of inquiry is a negation of scientific principles. Yet, many conventional economists think precisely the opposite. And, in many ways, this delusion is the source of our economic problems. If the trained economist is blind to the historical trajectory of a complex economic system and advises policymakers, such a system is headed for trouble. The conventional economic viewpoint is itself a meso rule which, reflexively, impacts upon the economy. This is not a new perspective, however. It was upheld most notably by John Maynard Keynes who rejected all claims by conventional economists to be able to advise beyond a short period of time and only in local, well known circumstances. In taking this position, he followed the line of his famous teacher, Alfred Marshall who, paradoxically, was also a founding father of neoclassical economics. He saw it as useful a body of theory but not as the general analytical core of all economic thinking. So some economists have known for a long time that economic systems are complex and adaptive and, as such, only understandable from a historical perspective, but considerations of ideology and scientific respectability have pushed the discipline in the opposite direction back towards an imitation of Nineteenth Century physics with a bit of Machiavelli thrown in.

8 CONCLUSION

Thinking about economics in terms of complex systems science remains in its infancy. Progress has certainly been made in the sense that phenomena that cannot be dealt with effectively in conventional economic analysis, such as increasing returns to scale, satisficing behaviour, routine behaviour and a willingness to seek profits in states of uncertainty, can all be better understood from a complex systems perspective. It has led to a much better understanding of how firms behave and why industries grow and decline. However, eschewing mathematical formalism has meant that a predictive science has not been offered and, thus, there has been limited assistance provided to policymakers keen to forecast future economic events. However, the availability of large amounts of computing power has enabled agent-based simulation to become an alternative way of developing theories of economic evolution but there is still a lack of discriminating tests of significance of the kind that that are employed routinely by conventional economists in testing their simple linear econometric models. Complex systems science tells us that we can only understand the behaviour of economic systems from a historical perspective but historical data of a high enough frequency is seriously lacking in most cases.

But all this may be asking too much of economists. Modern economists have sought to attain the status of natural scientists, applying mathematics and seeking models that forecast accurately but, a century ago, economics was seen more as a discipline seeking to understand economic history not a source of forecasting models. It was true that theorising using mathematics existed at that time but this was not regarded as a substitute for historical inquiry. Instead, mathematics was merely a tool to take a stylised fact concerning the observed relationship between phenomena to its logical conclusion. So modern economists should display some of the modesty of their ancestors in forming their scientific aspirations and accept that they are dealing with complex, inter-connected economic systems. There is, no doubt, an important role for economists in providing us with an understanding of how our economic systems work and, in so doing, enabling us to modify or change the rules that drive them for collective benefits. But this is a role that must begin in careful historical study, not in clever mathematics. It must also be acknowledged from the outset that economic systems are network structures and that network connections, or a lack of them, with non-economic systems are crucial. In particular, spatial connections in and between systems cannot be just ignored or assumed away. So economists need acquire a renewed respect for economic geographers and, in this regard, it was very heartening when Paul Krugman won the Nobel Prize in Economics in 2008.

So it is not the role of economists to do science in the manner of physicists, their true scientific role is to explain how complex economic systems work and how their workings have evolved over time. That is something that hardly anyone living within such systems can begin to understand. Ignorance of the system in which you are a component is a serious disadvantage. Understanding a complex system is not possible using the conventional economics of an independent optimising individual operating in a sea of perfect knowledge. The key challenge faced by economists prepared to apply complex systems science is to find rhetorical devices, pedagogical techniques and tools of discourse that can demonstrate that a superior understanding can be gained by approaching economics from an entirely different perspective. Agent bases simulations abound, but this more important work has barely begun.

BIBLIOGRAPHY

[Albin, 1998] P. Albin. *Barriers and Bounds to Rationality: Essays on Economic Complexity and Dynamics in Interactive Systems.* Princeton University Press, Princeton, 1998.

[Aldrich *et al.*, 2008] H. E. Aldrich, G. M. Hodgson, D. L. Hull, T. Knusden, J. Mokyr, and V. J. Vanberg. In defence of generalized Darwinism. *Journal of Evolutionary Economics,* 18, 577-596, 2008.

[Arthur, 1999] B.W. Arthur. Complexity and the economy. *Science,* 284, 107-109, 1999.

[Bak, 1997] P. Bak. *How Nature Works.* Oxford University Press, Oxford, 1997.

[Beinhocker, 2006] E. D. Beinhocker. *Origin of Wealth: Evolution, Complexity, and the Radical Remaking of Economics.* Harvard Business School Press, Cambridge, Mass., 2006.

[Blume and Durlauf, 2005] L. Blume, S. N. Durlauf. *The Economy as an Evolving Complex System III.* Oxford University Press, Oxford, 2005.

[Brooks and Wiley, 1986] D. R. Brooks and E. O. Wiley. *Evolution as Entropy: Towards a Unified Theory of Biology.* Chicago University Press, Chicago, 1986.

[Cordes, 2007] C. Cordes. Turning economics into an evolutionary science: Veblen, the selection metaphor and analogical thinking. *Journal of Economic Issues,* XLI, 135-154, 2007.

[Day, 1994] R. H. Day. *Complex Economic Dynamics.* Cambridge University Press, Cambridge, 1994.

[Depew and Weber, 1995] D. J. Depew and B. H. Weber. *Darwinism evolving: System Dynamics and the Geneology of Natural Selection.* MIT Press, Cambridge, Mass.,1995.

[Dopfer, 2005] K. Dopfer. *The Evolutionary Foundations of Economics.* Cambridge University Press, Cambridge, 2005.

[Dopfer *et al.*, 2004] K. Dopfer, J. Foster, and J. Potts. Micro-meso-macro. *Journal of Evolutionary Economics,* 14, 263-79, 2004.

[Dopfer and Potts, 2007] K. Dopfer and J. Potts. *The General Theory of Economic Evolution.* Routledge, London, 2007.

[Epstein and Axtell, 1996] J. M. Epstein and R. Axtell. *Growing Artificial Societies – Social Science from the Bottom Up.* Cambridge University Press, Cambridge, 1996.

[Foster, 1993] J. Foster. Economics and the self-organisation approach: Alfred Marshall revisited? *The Economic Journal*, 103, 975-991, 1993.

.

[Foster, 1997] J. Foster. The analytical foundations of evolutionary economics: from biological analogy to economic self-organization. *Structural Change and Economic Dynamics,* 8, 427-451, 1997.

[Foster, 2005] J. Foster. From simplistic to complex adaptive systems in economics. *Cambridge Journal of Economics*, 29, 873-892, 2005.

[Foster and Metcalfe, 2001] J. Foster and S. Metcalfe. *Frontiers of Evolutionary Economics: Competition, Self-Organization and Innovation Policy*, Edward Elgar, Cheltenham, 2001.

[Foster and Potts, 2006] J. Foster and J. Potts. Complexity, evolution and the structure of demand. In M. McKelvey, M. Holmén (eds.). *Flexibility & Stability in the Innovating Economy.* Oxford University Press, Oxford, 99-120, 2006.

[Foster and Potts, 2009] J. Foster and J. Potts. A micro-meso-macro perspective on the methodology of evolutionary economics: integrating history, simulation and econometrics. In U. Cantner, J-L. Gaffard, and Nesta, eds. *Schumpeterian Perspectives on Innovation, Competition and Growth* , Springer: Berlin, 2009.

[Foster and Wild, 1996] J. Foster and P. Wild. Economic evolution and the science of synergetics. *Journal of Evolutionary Economics*, 6, 239-260, 1996.

[Hayek, 1967] F.A. Hayek. The theory of complex phenomena. In F.A. Hayek (ed.), *Studies in Philosophy, Politics and Economics*, University of Chicago Press, Chicago, 1967.

[Kaldor, 1972] N. Kaldor. The Irrelevance of Equilibrium Economics. *The Economic Journal*, 82, 1237-1255, 1972..

[Langton, 1972] C.G. Langton. *Artificial Life,* Addison Wesley, New York, 1972.

[Malerba *et al.*, 2001] F. Malerba, R. Nelson, L. Orsenigo, and S. Winter. History-friendly models: An overview of the case of the computer. *Journal of Artificial Societies and Social Simulation*, 4/3/6, 2001.

[Markose, 2005] S. Markose. Computability and evolutionary complexity: markets as a complex adaptive system, *The Economic Journal,* 115, 159-192, 2005.

[Martin and Sunley, 2007] R. Martin and P. Sunley. Complexity thinking and evolutionary economic geography. *Journal of Economic Geography*, 7, 573-601, 2007.

[Metcalfe, 2002] J. S. Metcalfe. Knowledge growth and the growth of knowledge. *Journal of Evolutionary Economics,* 12, 3-15, 2002.

[Nelson and Winter, 1982] R. Nelson and S. Winter. *An Evolutionary Theory of Economic Change.* Belknap/Harvard University Press, Cambridge, Mass., 1982.

[Potts, 2000] J. Potts. *The New Evolutionary Microeconomics: Complexity, Competence and Adaptive Behaviour.* Edward Elgar, Cheltenham, 2000.

[Rogers, 2003] E.M. Rogers. *Diffusion of Innovations. 5^{th} edn.* Free Press, New York, 2003.

[Schumpeter, 1942] J.A. Schumpeter. *Capitalism, Socialism and Democracy.* Harper & Row, New York, 1942

[Simon, 1972] H. Simon. The architecture of complexity. *Proceedings of the American Philosophical Society,* 106, 467-482, 1972.

[Veblen, 1898] , T. B. Veblen. Why is economics not an evolutionary science? *Quarterly Journal of Economics*, 12, 373-426, 1898.

[Velupillai, 2000] V. Velupillai. *Computable Economics, the Arne Ryde Memorial Lectures.* Oxford University Press, Oxford, 2000.

[Werker and Brenner, 2004] C. Werker and T. Brenner. Empirical calibration of simulation models. *Papers on Economics and Evolution* No. 0410, Max Plank Institute of Economics, Jena, 2004.

[Witt, 1997] U. Witt. Self-organization and economics – what is new? *Structural Change and Economic Dynamics*, 8, 489-507, 1997.

[Witt, 2007] U. Witt. Heuristic twists and ontological creeds – a road map for evolutionary economics. *Papers on Economics and Evolution,* No. 0701, Max Plank Institute of Economics, Jena, 2007.

ECONOPHYSICS AND THE COMPLEXITY OF FINANCIAL MARKETS

Dean Rickles

1 INTRODUCTION

Within the 'complexity science' camp there are two broadly distinct ways of modeling the properties and behavior of socioeconomic systems[1]:

- 'Econobiology' ('evolutionary economics') perspective: uses the lessons of evolutionary biology to explain economic phenomena — economic complexity is viewed as analogous to, or grounded in, *biological* complexity.

- 'Econophysics' perspective: applies to economic phenomena various models and concepts associated with the *physics* of complex systems — e.g. statistical mechanics, condensed matter theory, self-organized criticality, microsimulation, etc.

Both of these approaches are 'population-level' ones (*cf.* [Mayr, 1970; Sober, 1980]): they seek to account for 'global' or 'collective' phenomena. They both do so in a 'bottom-up', 'generative' manner: collective ('macroscopic') properties are viewed as the result of interactions at the level of the ('microscopic') constituents. However, the aggregate and its elements are deemed to be of different *kinds* with causal lives of their own, the former (minimally) being supervenient on the latter. In interesting cases (i.e. where there is complexity) the aggregate system's properties (and dynamics, laws, etc.) are said to be 'emergent' in the sense that they are not reducible to some particular configuration of the constituents (and their properties) despite the fact that some such configurations will be *sufficient* for the generation of said properties — hence, the particular configuration will be sufficient but not necessary for the production of the emergent property. In other

[1]Some seem to think that the subsumption of complex socioeconomic behaviour under 'Self-Organized Criticality' counts a distinct third way. However, self-organized criticality can be easily accommodated within both of the economic complexity frameworks I mention and, hence, should not really be considered a separate enterprise. It is, strictly speaking, a part of (non-equilibrium) statistical physics.

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

words, the properties of the complex system are 'multiply-realizable' by distinct configurations (physicists refer to this latter property as 'universality').[2]

Here I restrict my attention solely to the physics-based approach[3] (i.e. econophysics), an approach generally couched in the language of *statistical* physics.[4] Statistical physics is a framework that allows systems consisting of many (possibly heterogeneous) particles to be rigorously analyzed. In econophysics these techniques are applied to 'economic particles', namely investors, traders, consumers, and so on. Markets are then viewed as (macroscopic) complex systems with an internal (microscopic) structure consisting of many of these 'particles' interacting so as to generate the systemic properties (the microstructural components being 'reactive' in this case, as mentioned already, resulting in an *adaptive* complex system relative to another system, or the environment).

I further restrict my attention to financial markets since that is where most work in econophysics has been conducted, on account of the availability of copious amounts of high-frequency data. Indeed, at the root of most of the work carried out in econophysics is a family of 'stylized facts' (empirically observable universal generalizations) that are to be found in this economic data — see §6. Econophysicists seek to find new instances of such facts and to explain these and previously known stylized facts using physics-inspired techniques and models, with the ultimate aim of providing these facts with a theoretical basis.

[2]The concepts of 'supervenience,' 'multiple-realization,' and, more so, 'emergence' are still very slippery and I shall avoid them for the most part in this chapter (with the exception of §2). However, economic systems do raise interesting and potentially novel issues *vis-á-vis* these concepts: for example, the 'subvenience basis' of elements responsible for the (supervenient) economic properties and behaviour — that is, the economic agents and their properties — have strategy and foresight, and therefore *respond* to the unitary properties and behaviour they create together. This highlights quite starkly one of the reasons why a simple ('macro-to-micro' or 'micro-to-macro') causal story cannot be told about events involving economic systems and economic agents (and complex systems and their parts more generally): the two form a *co-evolving* pair, updating their behaviour in the light of changes in the others' properties. In this way the 'micro-macro' disconnect of traditional economic theory is overcome.

[3]The econophysics approach cannot be the whole story in and of itself; it cannot (and should not) be considered as completely distinct from other approaches. The underlying *behaviour* that generates the economic data that econophysicists deal with is, after all, generated by sociobiological systems (of a rather special sort, as mentioned in the previous footnote). No doubt there will, at some level, have to be a union of the two perspectives ('physical' and 'sociobiological') — some early progress in this regard has been made in behavioural finance, including 'herding models' [Cont and Bouchaud, 2000] and 'minority game' models [Challet *et al.*, 2005]. My thanks to Clifford Hooker for raising my awareness of the difficult 'integrative' issue (private communication).

[4]There are other physics-inspired approaches to economics that do not utilize this analogy to statistical physics, using an analogy to some other branch of physics — interesting examples are gauge theory [Ilinski, 2001] and quantum field theory [Baaquie, 2004]. However, these approaches, though often referred to as examples of econophysics, do not match what most econophysicists have in mind (nor what I have in mind); namely, an approach that seeks to build *physically realistic* models and theories of economic phenomena from the actual empirically observed features of economic systems. Statistical physics is a many-body theory as, in general, is economics. One doesn't get the same intuitive connection with models based on quantum field theory and gauge theory — though, it has to be said, they do surprisingly well at *reproducing* economic data.

In fact, economists (primarily econometrists and those working in empirical finance) have been well aware, for quite some time, of most of the phenomena that econophysicists have 'discovered'. This has led to some impatience with econophysicists amongst economists — see, for example, [Gallegati *et al.*, 2006; Lux and Ausloos, 2002]. However, the econophysicists differ from the economists in that they aim to *explain* the various phenomena catalogued in the stylized facts by providing physically realistic ('microscopic') models and underlying theories. Also, the econophysicists tend to view the stylized facts more robustly, as genuine laws (on a par with those of fundamental physics) rather than lesser cousins as economists seem to. Hence, a claim often made by econophysicists is that their models are more 'realistic' than those offered up by economists and econometricians (see, for example, [Stanley *et al.*, 2006, p. 330]. This realism is supposed to be a consequence of the physics-based methodology which is more empirical: 'data first, then model' — I consider this claim in §5.2.

My aim in this chapter is simply to present the central ideas of econophysics: to show where they come from (their motivations), and to show how it all fits in with complex systems science. Since the notion of complexity is, to a large extent, still 'up in the air', I shall begin in §2 by getting straight on what I mean by this term within the confines of this chapter and as applied to (financial) economics. In §3 I introduce some elementary facts from statistics that will be used in subsequent sections. The main features of NeoClassical economics and finance are presented in §4. I then present some of the background to econophysics, and introduce the basic idea behind it in §5. This is followed in §6 by a look at the statistical puzzles in economic data (known as the stylized facts). Econophysics is a reaction to the standard model and a response to the problems faced by the standard model: we see how this is so in §7. I then consider some more overtly conceptual issues: in §8 I consider the issue of laws and invariance in econophysics and, finally, in §9 I present, and rebut, some recent objections to the econophysicists' inferences to complexity underlying the stylized facts.

2 COMPLEXITY AND COMPLEX SYSTEMS

'Complexity' is a notoriously slippery concept. Usually, precise definitions — in the sense of necessary and sufficient conditions — are avoided. However, whatever complexity may be, complex *systems* are supposed to possess it, so we can reframe the discussion so as to refer to these rather than complexity *per se*. Herbert Simon gives a rough characterization of a complex system as follows:

> by a complex system I mean one made up of a large number of parts that interact in a nonsimple way. In such systems, the whole is more than the sum of its parts, not in an ultimate, metaphysical sense, but in the important pragmatic sense that, given the properties of the parts and the laws of their interaction, it is not a trivial matter to infer the properties of the whole [Simon, 1981, p. 4].

Economic systems are an obvious candidate for the 'complexity treatment': they contain multiple agents, of different types (producers and consumers; risk averse and risk takers; firms and individuals, etc.), all competing for finite resources of some kind or another, and interacting in such a way as to generate the properties and dynamics of economic systems and subsystems. Econophysicists (and a small but growing number of economists) agree that these properties and the dynamics fit the 'complex system' bill: one finds, for example, scaling and universality, criticality, fractal patterns, and (candidates for) emergent properties. All attributes that a good complex system should possess.

2.1 Characteristics of complex systems

It is unfortunate that a more precise definition of a 'complex system' is still not agreed upon: there are almost as many definitions as there are discussions — indeed, the difficulty of the problem of definition points, I think, to the fact that we should avoid 'unificatory' approaches to complexity. However, it is reasonably safe to assume a *kernel* that these diverse accounts share. This kernel involves a triplet of characteristics (I hesitate to call them *necessary* conditions):

- A (unit) complex system must contain *many* subunits (the exact number being left vague).
- These subunits must be *interdependent* (at least *some* of the time).
- The interactions between the subunits must be nonlinear (at least *some* of the time).

The properties of the (unit) complex system are understood to be *generated by* or *supervenient on* the properties and interactions of the subunits that constitute it: there is no difference in the unit system without a difference in the subunits (though it is possible that a difference in the subunits does not manifest itself at the unit level). These properties are said to be 'emergent' when they amount to new complex ('systemic') structure that, as Kim puts it, "in some sense transcend[s] the simpler properties of [its] constituent parts" [2003, p. 556]. The subunits need not be identical, and the introduction of heterogeneity can also result in the emergence of higher-order properties of the unit system.[5]

[5]The canonical example in social science is Thomas Schelling's study of segregation [1971]; [1978, p. 147–55]. Here, slight differences in the (microscopic) preferences of individuals lead to massive, unexpected (i.e. emergent) macroscopic differences: very slight individual preferences to have neighbours 'like themselves' (i.e. in terms of colour, wealth, or in fact any property one cares to choose) can, despite a preference for integration, lead an initially well-integrated population into total segregation with respect to the chosen property. In other words, heterogeneity (with respect to the chosen property) coupled to a preference to be near others like oneself in that property (however weak that preference might be) provides a (self-organizing) clustering mechanism serving to partition the population. A more formal (and fundamental) model of these characteristics of a complex system is the Ising model, with spin components $s = +1$ or $s = -1$ and interaction Hamiltonian $H = -J \sum_{\langle i,j \rangle} s_i s_j$ (with coupling constant 'J').

If we are talking about an *adaptive* complex system then we should add the following condition:

- The individual subunits modify their properties and behaviour with respect to a changing environment resulting in the generation of new systemic properties that 'reflect' the change that the environment has undergone.

If we are talking about a *self-organizing adaptive* complex system then we should also add:

- The individual subunits modify their own properties and behaviour with respect to the properties and behaviour of the unit system they jointly determine — in other words, there is 'downward causation' operating from the systemic properties to the subunits' properties.[6]

These characteristics certainly seem to be in tune with most contemporary discussions of complex systems. However, as Latora and Marchiori [2004, p. 377] point out, these characteristics (and, indeed, most such characterizations) miss out on what they take to be an essential aspect of complex systems: the *network* structure of the subunits. Much recent work, especially on the *modeling* of complex systems and the *reproduction* of 'real-world' economic phenomena such as price dynamics, has focused on the structural features of such networks, rather than on the specific form of the nonlinear interactions between individual subunits — see [Amaral and Ottino, 2004] for further details on the relevance of networks to complex systems science. It is highly likely that future econophysics research will include complex networks as a major component, and this may function as the ontological glue that sticks together econophysics' models and the underlying sociobiological mechanisms responsible for the economic reality these models are intended to represent — Vega-Redondo has done some interesting work on this subject: see [Vega-Redondo, 2007].

2.2 Extreme events as an indication of complexity

There are additional features of complex systems that are involved in economics — in large part these can be derived from the aforementioned features. For example, complex systems often exhibit large and surprising changes that appear

[6]There are problems with the notion of downward causation, notably that any causal chain from an emergent property E to some property P of the subunits (which subunits, you will recall, form the 'base' B 'generating' E) is underdetermined by the base itself. I.e. whenever there is E there is B (or something resembling B in terms of its causal powers and its ability to generate E — E being multiply realizable), so whenever we say that E 'downwardly causes' P we might just as well say B causes P and dispense with the notion of downward causation altogether. This is a general problem with 'population thinking' according to which aggregate-level phenomena have a causal life of their own. However, I agree with O'Connor and Wong [2005] that this problem evaporates once we realize that emergence is not a synchronic relationship between the subunits and the unit but a dynamical process (and, hence, a diachronic relationship). This is indeed borne out by many branches of complexity science where we see that it is *local iterations of processes* that lead to emergent (global) phenomena. For a similar argument see [Hooker, 2004].

not to have an outside cause, instead arising endogenously. The corresponding economic phenomenon here is, of course, the stock market crash (see [Sornette, 2003]), where *post hoc* examination reveals no sign that the arrival of news caused the crash, nor do the dynamics of the financial fundamentals appear to be involved. These speculative bubbles then correspond to a self-organization process (see [Lux, 1996]). Econophysicists argue that stock market crashes, and other economic phenomena (that are often *puzzling* from the perspective of standard economic theory, 'outliers' in fact) are an entirely natural consequence of the view that economic systems, such as financial markets, are complex. Extreme events, involving collective phenomena (resulting from the iteration of nonlinear interactions), such as herding or alignment (as seems to occur in bubbles and crashes), are an integral part of scaling theory (itself a part of statistical physics). They correspond to critical phenomena in which there is long-range dependence between the elements (i.e. diverging correlation length) so that small changes in certain parameter values can result in massive systemic changes. More generally, criticality involves fluctuations of the 'order parameter' (say the returns[7] on some asset) and power law behaviour. Hence, extreme behaviour in a system is a strong indication that complexity is involved. Before we turn to these puzzles and issues, let us first briefly present some basic facts from probability theory and financial economics.

3 PROBABILITY DISTRIBUTIONS

Probability distributions are of vital importance in complex systems research, especially in the investigation of the properties of financial markets. They are what allow us to ascertain the inner workings of complex systems, to uncover their regularities and aspects of their structure.

Given an experiment (or process) with outcome sample space S, a random variable X is a map from outcomes to real numbers — we assume that the map is exhaustive in that every point of S is assigned some (not necessarily distinct) value. Given such a random variable X, a probability density function $\mathcal{P}(x)$ provides information concerning the way the variable is distributed. To work out the probability that the value of X is in between the values a and b one simply computes the integral $\int_b^a \mathcal{P}(x)dx$. Of course, the most well-known example of such a distribution is the Gaussian ('normal') case, with distribution function:

$$(1) \quad \mathcal{P}_{\text{Gauss}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

Here, μ is the mean ($= \Sigma_{i=1}^n x_i/n$) and σ^2 is the variance. This distribution is ubiquitous in the natural (and social) world because it is linked to the central

[7]Returns are defined as follows: Let $p(t)$ be the price of some financial asset at time t. The return $R_\tau(t)$ from the asset, at time t for scale factor τ (giving the frequency of returns), is the relative variation of it's price from t to $t+\tau$, or: $R_\tau(t) = \frac{p(t+\tau)-p(t)}{p(t)}$. According to the standard model of finance these returns are uncorrelated IID (independent and identically-distributed random) variables; a feature flatly contradicted by the empirical data from real markets.

limit theorem which tells us, roughly, that any stochastic process (understood as the aggregated result of a complex mixture of (independent) random factors) will be characterized by a Gaussian distribution.

A more appropriate distribution for finance is the lognormal distribution which simply involves the (natural) logarithm of x being normally distributed:

$$(2) \quad \mathcal{P}_{\mathrm{LogNorm}}(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left[-\frac{(\log x - \mu)^2/(2\sigma^2)}{x\sigma}\right]$$

Of course, $\ln(x)$ for $x < 0$ is undefined, which matches the non-negativity of most assets. However, variables that are normally distributed tend to exhibit rather mild fluctuations. They are clearly not capable of dealing with the kinds of large-scale extreme fluctuations that correspond to stock market crashes for example — fine for human weight, not for human wealth.[8] Despite this the (log-) normal distribution is a central component of the standard model of finance, as we see in the next section — recall that the central limit theorem plays a role here too, only with a multiplication of factors replacing the additivity of factors in the normal distribution.

Hence, much of the action in econophysics research tends to focus on probability distributions for variables representing financial observables where it is argued, on the basis of empirical evidence, that certain financial and economic observables do not fit a Gaussian curve, but fit instead some other distribution. The evidence depicts a system with observables that frequently take on 'extreme' values, values that would be counted as incredibly rare (impossible *for all practical purposes*) according to a Gaussian distributed variable.[9]

Examples of alternatives to the normal distribution are the exponential, stretched exponential, and the Lévy distributions (with 'fat' power law tails: see note 9) — there are *very* many more. Power law distributions are especially important in the context of complex systems research since they are believed to point to the underlying complexity of the generating process (they are viewed as 'signatures' of complexity). For cases in which x is large (in which case the Lévy distribution possesses 'Pareto tails'), the power law tail distribution function is:

$$(3) \quad \mathcal{P}_{\mathrm{Power}}(x) \propto \frac{\alpha A_{\pm}^{\alpha}}{\|x\|^{1+\alpha}}\,, \quad \|x\| \to \infty$$

[8] I've heard of extraordinary cases of people weighing around seven times my own body weight, and that is truly exceptional: it must mark some natural boundary on possible weight sustainable by the human form. However, there are very many people who earn many orders of magnitude more money than I do: Bill Gates (as of 2007) earns around 750000 times more money per year than me!

[9] The lognormal distribution fares slightly better than the normal distribution by having more probability in the tails — that is, having higher values of integrals $\int_b^a \mathcal{P}(x)dx$ for $a - b$ ranges covering larger x values. However, as I mentioned above, it still radically underestimates the probabilities of extreme events.

Here the (constant) exponent α is the 'tail amplitude' (or 'tail index') which provides information about the tail (it sets the slope of the graph, for example[10]) — $\alpha = 2$ corresponds to a Gaussian; as α approaches zero the center becomes more peaked and the tails fatter. The constant $A_\pm$ determines the size of the fluctuations of the relevant observable, x.

Power law distributions are characterized by the slow decay of probability along the tails of the distribution — it is for this reason that they are known as 'fat tailed distributions'. The thickness of the tails is one of the stylized facts that the standard model of finance has trouble explaining, since that model is based on a Gaussian (or log-normal) distribution which decays much more rapidly in the tails. I should point out that the log-normal distribution provides a very good fit of the data from a great many situations in economics and finance — indeed, it is often difficult to distinguish log-normal from power law distributions. The problems arise when one considers the extremes of the distribution (e.g. for big earners or for very large fluctuations in stocks prices — bubbles and crashes, that is). Note also that the size of the slope (determined by the power law exponent) varies according to the financial instrument involved: commodities and stocks, for example, appear to demand significantly different values (thus altering the kind of distribution). It is possible, then, that there are different mechanisms generating the behaviour of financial observables for a range of financial systems.

Notably, from the point of view of complexity research, these power law distributions are *scale invariant* (like fractals only in function-space): events (or phenomena) of all magnitudes can occur, with no characteristic scale. What this means is that the (relative) probability of observing an event of magnitude $\|x\| = 1000$ and observing one of $\|x'\| = 100$ does not depend on the standard of measurement (i.e. on the reference units). The ratio between these probabilities will be the same as that for $\|x\| = 1000$ and $\|x''\| = 10000$. Hence, there is no fundamental difference between extreme events and events of small magnitude: they are described by the same law (that is, the distribution *scales*).

Scale invariance of this sort is a feature of the so-called *critical* phenomena (at phase transitions) studied in statistical physics where one has the simultaneous involvement of many (widely) different length scales.[11] Given that this problem

[10]When we plot the distribution of a power law, $\text{Prob}(x) \sim x^\alpha$, on a $\log(\text{Prob}(x))$ versus $\log(\|x\|)$ plot (where $\|x\|$ is the size of some event or magnitude of some phenomenon (a stock price fluctuation, say) and $\text{Prob}(x)$ is its occurrence probability) we find a straight line of slope α, suggesting (if not quite implying) that the distribution is scale-invariant (the ratio of $\|x\|$ to its occurrence probability — that is, the number of fluctuations of a given magnitude — is invariant under rescaling). That a power law distribution would show up linear on log-log paper follows from the fact that given a power law $f(x) = x^\alpha$, $\log f(x) = \alpha \log x$. Note that one often sees the (complementary) *cumulative* distribution according to which one considers not the probability of an $\|x\|$-event, but of events greater than or equal to $\|x\|$.

[11]Specifically: near a critical point, fluctuations of the (macroscopic) order parameter will appear at all possible scales. In the case of the liquid-gas phase transition one will have liquid drops and gas bubbles ranging from the molecular level to the volume of the entire system. At the critical point these fluctuations become infinite. The analogous situation in the financial context would be, for example, fluctuations in asset returns at all possible scales. An excellent

(of physics at many scales) has been solved, one might expect that the methods used therein can be usefully transferred to the economic case. This explains the interest of statistical physicists, and the *raison d'être* of econophysics.

4 NEOCLASSICAL ECONOMICS

NeoClassical Economics depicts markets as efficient machines, automatically seeking out the configuration that is best (or optimal) for all economic agents. This configuration is an equilibrium state, the one that maximizes utility. The theoretical framework involves several extreme idealizations, not least of which are the assumptions of perfect rationality and omniscience (including unlimited foresight) on the part of the economic agents! Indeed, having rationality is the same thing as maximizing expected utility. Specific economic theories are constructed from this framework by applying the basic postulates to various economic situations. Financial economics is no exception. Neoclassical finance is based on a random walk model which states that sequences of measurements made to determine the value of some financial observable (returns for example) are such that the present (discounted) value is the best estimate (prediction) one can give for future values — or, in general, prices are Martingales (see note 13 for more details). However, this assumption is empirically inadequate and leads to underestimation with respect to the frequency of large price changes, as we see in §6.1.

4.1 The standard model of finance

Johannes Voit [2005] calls "the standard model of finance" the view that stock prices exhibit *geometric Brownian motion* — i.e. the *logarithm* of a stock's price performs a random walk.[12] Assuming the random walk property, we can roughly set up the standard model using three simple ideas: (1) the best estimation of an asset's future price is its current price[13], (2) the distribution of price changes forms a bell-curve ('mesokurtic' or Gaussian condition), and (3) buys balance sales. In the context of finance, these principles of the standard model are encoded in the

introduction to the theory of critical phenomena is [Binney *et al.*, 1992].

[12]This central idea of the standard model, though with an *arithmetic* Brownian motion, can be traced back to the doctoral thesis of Louis Bachelier, a student of Poincaré — this thesis, from 1900, is in print again in an English translation: [Davis and Etheridge, 2006]. This model was, more or less, later rediscovered (independently) by the physicist M. F. M. Osborne [1959]. For those with no knowledge whatsoever of mathematical finance who wish to learn more, I recommend Ross [2002].

[13]This is known as the *Martingale condition* defined by the conditional probability $E[X_{n+1} \mid x_1, ..., x_n] = x_n$ (where E is the *average* or *expected* value of what is enclosed in square brackets and X_i is a random variable conditioned on outcomes x_j). Strictly speaking the fundamental features of modern finance (as contained in the efficient market hypothesis: on which see below) can be derived from the statement that prices are Martingales (or sub-Martingales). The geometric Brownian motion model is really a special case of the family of Martingale models: a random walk is just a zero-mean Martingale — my thanks to Enrico Scalas for pointing this out to me (private communication).

central tool for pricing options[14]: the 'Black-Scholes-Merton model' [BSM] [Black and Scholes, 1973; Merton, 1973]. This is often viewed as a piece of early econophysics, though not along the lines of todays econophysics which is concerned with 'out of equilibrium' aspects. Moreover, the model is derived from the postulates of the neoclassical theory, rather than from the data (independently of an *a priori* theory of how markets ought to behave).

The central idea that underpins BSM is that one views markets as many-body dynamical systems. This insight is then used to draw an analogy with concepts from thermodynamics. In particular, the BSM equation brings over the concept of *thermodynamic equilibrium* into finance. This is defined in the financial context as a steady state reached when the underlying stock and the stock option are balanced in terms of the payoff they yield compared to the risk they entail.[15] The BSM equation describes this relationship[16]:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{4}$$

The solution, $C(S,t)$, of this equation then gives us the cost of constructing an option from the specified stock (or the 'rational value' of the option). Assuming constant r and σ, this is:

$$C(S,t) = SN(d_1) - Le^{-r(T-t)}N(d_2) \tag{5}$$

Here, L is the option's 'strike price' and T is its time to maturity. $N(\,)$ is the cumulative probability distribution function for a (standard) normal random variable. The arguments of the function are:

$$d_1 = \frac{\log(S/L) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}} \tag{6}$$

$$d_2 = \frac{\log(S/L) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}} = d_1 - \sigma\sqrt{(T-t)} \tag{7}$$

The problem of finding the best price for options is reformulated as a diffusion equation from which one gets the prices of various option-types by imposing various appropriate boundary conditions on the possible solutions.

[14]In brief, options are contracts that give the owner the right but not the obligation to buy (= 'call option') or sell (= 'put option') some asset (= 'the underlying') for a pre-specified price (= the 'strike price') at some pre-specified time in the future. Hence, the 'payoff' of an option is a function of the future price of the asset (or a group of such) — for this reason they are part of the family of financial instruments known as 'derivatives'.

[15]There is an immediate problem with the idea of a system reaching equilibrium in diffusional model like BSMs. The model has no equilibrium distribution: time evolution progressively flattens the probability density $\rho(x,t)$, but it does not converge to some time-independent function.

[16]The various terms in this partial differential equation are interpreted as follows: V is the value of some specified option (the details change depending on the type of option involved: in this case we consider European call options), σ is the stock's implied volatility (standard deviation of stock returns), S is the current price of the underlying stock, and r is the (risk-free) interest rate.

In terms of the probability distributions from the previous section, then, the relevant function is clearly the log-normal: this is required as a postulate.[17] Hence, modern finance is very much a NeoClassical theory. However, as mentioned, normal distributions cover only fairly mild fluctuations around some central value. Used as a model for ascertaining the riskiness of certain options, the BSM equation will assign vanishingly small probabilities to extreme fluctuations that are, in reality, not all that rare.

4.2 Market efficiency

This formal framework of BSM is given conceptual foundation *via* the efficient market hypothesis which states that prices always reflect all available information *in actual markets* [Fama, 1970] — the prices themselves emerge (aggregatively) through the consensus amongst a group of perfectly rational agents (the prices themselves are therefore rational). Price changes occur as a result of the exogenous intervention on the market by a piece of news, itself an unpredictable event. It follows that price changes are themselves unpredictable. Or, as Joseph McCauley expresses it: "there are no patterns/correlations in the market that can be exploited for profit" [2004, p. 88]. Let us spell this out in more detail.

The efficient market hypothesis is an inference from (NeoClassical) rational expectations principles: traders will wish to maximize their utility. This implies that they will look to exploit the market. The way to do this would be to spot patterns in price movements and then buy when they expect the price to give higher (than average) returns and sell when they expect lower (than average) returns. The trouble is, in doing this they will change the very patterns they are attempting to exploit: buying increases the price and selling drives the price down. This equalizes the market so that all financial instruments give the same return (modulo risk). In other words, the information that this arbitageur trader had about the market (the patterns) becomes reflected in the market prices. An endgenous process balances the market out. So, while there can be patterns that can be exploited, this is short lived (for the specific pattern), since acting on the information affects the prices and patterns get erased (by a process analogous to Walras' *tâtonnement*). This means that the best estimate for the future is the present price because that price reflects all known information (modulo short term discrepancies). One is left with the problem of what causes the price changes: the answer has to be external factors, and given the vast number and unpredictability of these, they are best modeled as random processes. Therefore, prices changes follow a random walk, and this gives us the foundation of modern (academic) finance and financial risk evaluation.

From the efficient market hypothesis we can quite clearly derive a testable prediction about the behaviour of financial observables (such as prices, returns, etc.): they should follow random walks in time — experience leads one to suggest a *biased*

[17]Further postulates required by the model are: the efficient market hypothesis (see below), constant interest rate, zero commission charges, and no dividend payouts.

random walk to account for the steady growth over long time scales. However, as we see in the next section, returns don't appear to behave in this 'random walk' way in real markets (*cf.* [LeBaron, 2006, p. 222-4]).

5 THE ROUGH GUIDE TO ECONOPHYSICS

The term 'econophysics' was chosen with some care to follow the path of such mergers as 'astrophysics', 'biophysics', and 'geophysics'. The reason for this was to keep the kind of work carried out by econophysicists within physics departments [Stanley *et al.*, 2006, p. 337] — note that it was H. E. Stanley who thus christened the field (in print) in [Stanley *et al.*, 1996a]. Minimally, econophysics is based on the observation of similarities between economic systems and concepts and those from physics. For example, Bertrand Roehner defines it simply as "the investigation of economic problems by physicists" [2005, p. 3]. A slightly less general definition comes from Mantegna and Stanley [2000]: "The word econophysics describes the present attempts of a number of physicists to model financial and economic systems using paradigms and tools borrowed from theoretical and statistical physics" (p. 355). However, as they go on to say, a "characteristic difference [from traditional approaches to economics and mathematical finance — DR] is the emphasis that physicists put on the empirical analysis of economic data" (*ibid.*). This latter factor is supposed to constitute the 'added value' of econophysics.

5.1 Some econophysics pre-history

The existence of a close relationship between physics and economics is nothing new, of course: many of the great economists did their original training in physics, and the influence of physics is clearly evident in many of economic theory's models — see [Mirowski, 1989; Ingrao and Israel, 1990; Cohen, 1994; Schabas, 2006]. I already mentioned too how the centerpiece of modern finance, the Black-Scholes-Merton model, is directly derived from physics. There are, moreover, many instances of physicists who have applied 'the physicist's method' to social phenomena. For example, Daniel Bernoulli found that there were statistical regularities in what are *prima facie* unpredictable events — e.g. the number of letters in the Paris dead-letter office (see [Farmer *et al.* , 2005, p. 37]). The enigmatic theoretical physicist Ettore Majorana outlined and defended the application of statistical physics to social phenomena [Mantegna, 2005]. There have been attempts to apply statistical mechanics to economics as far back as 1959 when M. F. M. Osborne developed his Brownian motion model of a stock market [Osbourne, 1959]. Duncan Foley has done extensive work in this area [Foley, 1994]. Farjoun and Machover also develop this analogy [Machover and Farjoun, 1983]. In each case, however, the model is equilibrium statistical mechanics, and it is precisely the equilibrium condition that is thought to be at fault by econophysics. One cannot forget, either, Mandelbrot's discovery of scaling behaviour of cotton prices: [Mandelbrot, 1963]. The 'statistical physics connection' also underpins much of this earlier work: the idea is that one can ignore microscopic detail (the individual social entities) in favour of the coarse-grained macro-description (the groups of individuals bound together by social forces), with the aim of deriving macro-level laws. Econophysics is, at bottom, this same thought played out again, only now with the benefit of an underlying theory of multi-scale systems (i.e. renormalization group theory) giving principled reasons to ignore microscopic details in favour of a few choice parameters.

5.2 The methodology of econophysics

Econophysics gets itself off the ground as a separate enterprise from economics because, unlike the former, the latter supposedly has an unscientific 'non-empirical' (or 'axiomatic') methodology.[18] As Zhang [1998] puts it, "as a physicist, one may get the strange feeling that the theory [the standard model of economics — DR] is detached from the experiment" (p. 51). Likewise, Challett *et al.* write that

> physicists generally feel uneasy about several pillars of mainstream economic theory, such as rational expectations, the efficient market hypothesis and the notion of equilibria to name a few. This approach looks too axiomatic and formal to deal with complex systems as, for example, financial markets. ... [E]conophysicists deny the very rules of the game on which mainstream academic research in economics is based. [Challet *et al.*, 2005, p. 14]

As this quotation makes plain, econophysics is viewed (by most of its practitioners) as a *revolutionary* reaction to standard economic theory that threatens to enforce a paradigm shift in thinking about economic systems and phenomena.

We have here, I think, something roughly related to the 'principle-theory' versus 'constructive theory' distinction that Einstein made in regard to his 1905 formulation of special relativity, as compared with Lorentz's approach [Einstein, 2002] — the distinction was based on thermodynamics (on the 'principle' side) and statistical mechanics (on the 'constructive' side). As Einstein himself explains,

> [Constructive theories] attempt to build up a picture of the more complex phenomenon out of the materials of a relatively simple formal scheme from which they start out. Thus the kinetic theory of gases seeks to reduce mechanical, thermal and diffusional processes to movements of molecules - i.e. to build them up out of the hypothesis of

[18]This brings us to a more general point concerning what Thomas Gieryn calls "Boundary Work" (see, for example, [Gieryn, 1999]). Gieryn argues that when a new discipline comes along, it must strive to separate itself off from other ongoing endeavours and, in particular, aim to demonstrate that it, unlike its competitors, is truly scientific. It seems that econophysicists are doing just this in opposing the axiomatic style of NeoClassical economics. However, there's more to economic theory than the axiomatic approach, and in focusing too heavily on this aspect (to draw the boundaries) econophysicists are ignoring many important details. I think in this case we can agree with Gieryn that these objections are rhetorical devices employed to create an illusion of importance and originality for the new field of econophysics.

> molecular motion. When we say we have succeeded in understanding a group of natural processes, we invariably mean that a constructive theory has been found which covers the processes in question.
>
> Along with this most important class of theories there exists a second, which I will call "principle theories." These employ the analytic, rather than synthetic, method. The elements which form their basis and starting-point are not hypothetically constructed but empirically discovered ones, general characteristics of natural processes, principles that give rise to mathematically formulated criteria which the separate processes or the theoretical representations of them have to satisfy. Thus the science of thermodynamics seeks by analytical means to deduce necessary connections, which separate events have to satisfy, from the universally experienced fact that perpetual motion is impossible. [Einstein, 2002, pp. 100–101]

Though admittedly not a perfect fit, the approach of Black, Scholes, and Merton involves something more like a principle-theory-type approach in that the various principles going into their model possess the status of postulates of universal empirical generality, concerning the behaviour of economic agents — crucially, no underlying mechanisms for the phenomena are elucidated (*à la* thermodynamics). By contrast, econophysicists, making use of the statistical physics (rather than thermodynamical) analogy, adopt more of a constructive-theory-type approach. As Johnson *et al.* [2003] state: "[a]s physicists, our tendency is to put our trust in models which are microscopically realistic, and where the model parameters hold some physical meaning" (p. 251).

Stanley *et al.* [1999] claim that econophysics thus approaches economic systems "in the spirit of experimental physics" (p. 157): in contrast to standard methods in economics, econophysicists "begin empirically, with real data that one can analyze in some detail, but without prior models" (ibid.). While almost any philosopher of science would disagree with the details of this statement, the point is well-taken: data first, then model (whether the 'raw data' is itself encoded in a data model or not we can ignore here). As Bouchaud and Potters [2003] put it: "no theoretical model can ever supersede empirical data [in physics]" ... [p]hysicists insist on a detailed comparison between 'theory' and 'experiments' (i.e. empirical results, whenever available)" (p. xvi). However, it is absurd to think that economists would disagree with this in principle: there is an entire field (empirical finance) that adopts this same methodology (often employing 'model free' nonparametric statistics to analyze financial data). Moreover, to return to a point that was shelved earlier, the notion of data (or measurement) without theory was the subject of years of intense debate in the early days econometrics. Tjalling Koopmans [1947], for example, argued strongly for the view that measurement (of economic variables) without theory simply doesn't make sense — this was, broadly, the view adopted by the Cowles Commission (see [Hildreth, 1986]). Indeed, the early issues of the Econometric Society's journal, *Econometrica*, are littered with debate over the relationship and relative priority of data and theory.

The simplistic view adopted by econophysicists seem to be more in line with early business-cycle theorists, such as Burns and Mitchell [1946], which Koopmans characterises as involving the observation of economic phenomena "made with a minimum of assistance from theoretical conceptions of hypotheses regarding the nature of the economic processes by which the variables studied are generated" [Koopmans, 1947, p. 161]. We could easily criticize the work of econophysicists on the same grounds as this "measurement without theory"-oriented work on business cycles was criticized; namely, that there is no such notion of a basic theoretically unladen 'fact' on which to build. Moreover, without theoretical guidance, the generated facts (laden or not) are impotent.

It seems that the specific target for econophysicists' animosity is some form of the (rationalist) 'Austrian-type' economic theory, with its rejection of an empirical approach in favour of the logical derivation of economics from axioms of (individual) human action. That is, the problem is with "theory without measurement"-style methodology. This view is, however, not at all mainstream (indeed, it is sometimes labeled 'heterodox'!). The problem is, there are plenty of economists who are equally (if not *more*) uneasy about rational expectations models, utility maximization, the efficient market hypothesis, and general equilibrium — Austrian economics is a case in point. There are plenty of examples of *empirical* economics too: experimental economics being one obvious example in which NeoClassical ideas are put to the test and found to be empirically inadequate — see [Guala, 2005]. Moreover, there are plenty of physicists who appear to be unperturbed about working in a manner detached from experiment: quantum gravity, for example. Here, the characteristic scales are utterly inaccessible, there is no experimental basis, and yet the problem occupies the finest minds in physics.

I don't think we can agree, then, that econophysics adopts a different 'physics based' methodology and that this is what distinguishes it from plain vanilla economics. Let us put this aside, though, and consider why complex systems might be difficult for axiomatic approaches (as the above quote from Challett *et al.* suggests). One can well imagine axioms governing the behaviour of complex systems, with emergent laws and so on. Surely it is the *particular* axioms that NeoClassical economics is based on that are problematic from the perspective of complex systems, not the axiomatic approach *per se*? The axioms do not do justice to the reality of markets, involving, as econophysicists claim, out of equilibrium behaviour, self-organization, phase transitions, clustering, and a host of other phenomena not adequately captured by the equilibrium theory. If this is what is meant, then it is a fair point: the axioms make the wrong predictions about real economic systems. This is hardly an original point, but it points to a *genuine* problem with NeoClassical economics *vis-à-vis* complex systems science.

If we are talking not of econophysics *per se* but the complexity approach in general then we do witness a significant difference between this approach and mainstream economic theory. Moreover, in this case it does turn on a methodological issue: methodological individualism to be exact. NeoClassical economics is based on the idea that the way to understand complex socioeconomic phenomena is to examine the individuals. By synthesizing one's knowledge of the individual level, one can deduce the various phenomena. This is purely a mechanical picture, along the analytical lines of classical mechanics. To understand a complicated entity, decompose it into its parts: the whole is nothing but the individual parts that compose it. In other words, NeoClassical economic theory does not treat economic systems as complex systems where network, structure, interaction and emergence play a central explanatory role and individual details are largely irrelevant. And in its avoidance of the importance of non-individualistic matters and interactions, NeoClassical theory fails to be empirically adequate. We highlight these flaws in the next section.

6 STATISTICAL PUZZLES (AKA 'THE STYLIZED FACTS')

Financial time series display some *prima facie* puzzling empirical (statistical) regularities that make their modeling a tricky business. These are called "stylized facts". As Cont [2001] explains, a "stylized fact" is a "set of properties, common across many instruments, markets and time periods" (p. 223). In other words, stylized facts are *universal* regularities, independent of time, place, and many specific compositional details. Coolen [2004] refers to these regularities as "benchmarks, to be met by any theory claiming to explain aspects of financial time series" (p. 234).

This is a puzzle: why should stocks in, say, pork bellies look the same (statistically) as stocks in technology companies? The curious statistical properties of the data are fairly well-known amongst economists, but they *remain* a puzzle for economic theory.[19] This is where physicists come in: whereas some economists had attempted to recover the stylized facts in their models, the models had no empirical grounding; their sole purpose was to *replicate* the statistical properties by any means (admittedly, no small feat in itself given the constraints these stylized facts impose). Econophysicists, by contrast, use the statistical properties as their starting point; the basis from which to construct realistic models: the universality of the statistical properties — i.e. the fact that they reappear across many and diverse financial markets — suggests (to physicists at least) a common origin at work behind the scenes and points towards the theory of critical phenomena (with its notion of *universality*). Many econophysicists view their task as searching for

[19]The school known as 'behavioural economics' has made some progress along a different ('internal') route to both standard economic theory and econophysics (see [Shefrin, 2002; Shleifer, 2000]) — Herbert Simon [1955] did some excellent early work on behavioural models of rational choice, into which he tried to inject some realism concerning actual decision making behaviour (*vis-á-vis* the actual computational powers of humans and their limitations in terms of access to information). There are some overlaps between the behavioural models and the statistical physics models used by econophysicists: in particular, there are analogies between the cooperative or collective phenomena of the physics of critical phenomena and the imitative models used by behavioural economists. Sornette [2003] offers a nice integration of behavioural models and statistical physics.

and elucidating this common mechanism. This is also where the connection to contemporary research on complexity comes in: the stylized facts are understood to be emergent properties of complex economic systems. Here, then, we have a genuinely novel and potentially important feature of econophysics: the search for mechanisms underlying the economic phenomena utilizing the direct intuitive link between these phenomena and aspects of statistical physics.[20] Let us finally present these stylized facts on which so much hangs.

6.1 The facts of the matter

Here I mention only those stylized facts those most relevant to complexity issues[21]:

Fat Tails: the returns of various assets (evaluated at high frequencies: e.g. a month and less) exhibit fourth moments (kurtosis levels) that are anomalously high when superimposed over a Gaussian distribution. The distributions are roughly bell-shaped but assign greater (than normal) probability to events in the center (i.e. they are more peaked) and at the extremes (i.e. they exhibit heavy tails). In other words, the time series for returns display a significantly larger number of extreme events than a Gaussian process would generate.

- The standard model of finance involves the idea that price changes obey a lognormal probability distribution. This implies that massive fluctuations (crashes or 'financial earthquakes') are assigned a vanishingly small probability: if the world really were like this, then we should not be seeing the kinds of crashes we *do* see.[22]

Volatility Clustering: periods of intense fluctuations and mild fluctuations tend to cluster together: big price changes, of either sign, follow big price changes and little ones, of either sign, follow little ones.

---

[20]Again, however, I don't see how this can be sufficient as it stands. Statistical physics denatures the economic agents as severely as any NeoClassical theory, and yet surely the nature of the agents has to play a role in generating the behaviour. True, there might well be emergent effects that enable us to ignore these details when it comes to modeling, but if it is *understanding* we seek, then we cannot ignore the behaviour of the agents. Certainly, if we are to *do* anything of practical importance with econophysics then we need some way of translating the statistical physics talk into talk about real individuals and economic reality. (I am grateful to Allan Walstad for impressing this problem on me — email communication).

[21]There are very many more than I present here, but those presented will suffice to show how econophysics is supposed to score over the standard model and why the data is believed to point towards a complex systems approach. See [Cont, 2001] for more examples.

[22]In statisticians' terms, if the price changes — or, strictly speaking, their logarithms since we a dealing with a log-normal distribution in the standard model — behaved according to the standard model, the probability distribution would have a kurtosis of around 0 (0 is the value for data that fit the bell curve exactly): such distributions are called "mesokurtic". Distributions of data from *real* markets, with their characteristic 'fat tails,' exhibit positive kurtosis (giving "leptokurtic" probability distributions).

- If the process generating a time series were Gaussian, then we would expect to see a very uniform time distribution of large and small fluctuations. Instead what we see are sequences of periods of large fluctuations and periods of small fluctuations (high and low volatility).[23]

Volatility Persistence ('Long Memory'): there is a dependency between stock market returns at different times. Technically, the volatility has slowly decaying autocorrelations.[24]

- The autocorrelation of returns decays very quickly to zero, providing support for the efficient market hypothesis and for the Brownian motion model of the standard model of finance.

  (8) Linear Returns $\equiv C(\tau) = \mathbf{Cor}(r(t, \Delta t), r(t + \tau, \Delta t))$

  However, the autocorrelation for squared returns decays much more slowly, and can remain positive for as long as a month. Hence, there exists nonlinear dependence.

  (9) Squared Returns $\equiv C_2(\tau) = \mathbf{Cor}(|r(t + \tau, \Delta t)|^2, |r(t, \Delta t)|^2)$

Relations: This persistence is obviously related to the above volatility clustering, and is essentially just what one computes to gain a numerical purchase on the clustering phenomenon. The clustering itself generates excess volatility (fat tails). Hence, explaining the clustering and long memory will most likely constitute an explanation of the fat tails. One would like and expect an integrated account of the stylized facts, according to which the same mechanism is responsible for generating multiple stylized facts in a unified manner. This is what econophysicists aim to provide.

In short: price changes change by too much, too often, and with too much 'order' to fit the geometric Brownian motion model that the standard model is based on. There would not be the quantity of large crashes that have been witnessed if that model were true.[25] If we plot the size of price changes against time, we see

[23] Again, in statisticians' lingo we find significant *conditional heteroskedasticity* of returns. These sudden switches bring to mind phase transitions, and indeed this analogy is pressed by many econophysicists — attempts are made to connect this to reality by considering the phenomena to be the result of cooperative ('herding') and competitive effects amongst agents. Again, [Sornette, 2003] is the best place to learn about this.

[24] One characterizes the dependence between points of a time series *via* the Hurst exponent $H = 1 - \alpha/2$ (where α is the tail exponent for a power law distribution): $H = 0.5$ indicates a Brownian motion; $0.5 < H < 1$ indicates positive long range correlations (and an underlying long memory process) — the corresponding data set is known as a fractional Brownian motion. See [Clegg, 2006] for an elementary guide to the Hurst exponent.

[25] Recall that many crashes have no apparent cause, which signals that we might have a complex system. The idea of the efficient market hypothesis, that exogenous factors must be responsible for any price changes (including the massive ones in crashes), does not seem to fit the fact that often no such cause (e.g. in the fundamentals) can be found in 'postmortems' of real cases of crashes. Further, changes, contributing to the volatility, cannot always be due to the impact of some relevant piece of news since the changes are far more frequent than the arrival of such news — see Bouchaud *et al.* [2004, p. 176].

that there are far more large changes than the standard model suggests. There is too much predictability to the time series for it to be a random walk process generating the data, on account of the clustering and dependence. Mainstream financial economics does not fit the empirical facts.

The probability distribution is best matched by power-law tails rather than a Gaussian distribution. The existence of a power-law distribution often points to some underlying complexity in the system that generates it. It is on this basis that many tools from the theory of critical phenomena and condensed matter physics have been brought over. Often, these are used to provide the *physical* underpinnings of various economic phenomena, in addition to providing the mathematical concepts. It is this feature that makes econophysics so interesting from a philosophical point of view.[26]

6.2 Significance of the stylized facts

What these features demonstrate is that the underlying mechanism responsible for generating time series data is not one that produces a normal distribution (nor log-normal). The latter simply does not fit the observed statistical properties. It follows that a model based on a normal distribution would assign lower probabilities to extreme events than it ought to to be empirically successful. Given that the standard model of finance involves returns that are log-normally distributed, there is a clear conflict here between theory and evidence. There is every reason to attempt alternative approaches: econophysics is one such, but there are others that do well. It is notable that the other approaches that tend to do well with the stylized facts are complex systems oriented — agent-based computational economics, for example (see [Tesfatsion and Judd, 2006] for a good overview). The idea is to view the stylized facts as emergent properties of a complex system.

There are diverse attempts to reproduce these stylized facts with various methods and models: broadly one can replicate (instrumentalism) or one can explain (realism). Econophysics uses models based on statistical physics and is an attempt along the latter lines. However, this is not the only option: as mentioned, econophysics does not have a monopoly on the stylized facts. For starters, these stylized facts were isolated well before the emergence of econophysics. In fact they go back at least as far as Wesley Mitchell [1913], who identified such properties in the context of his work on business cycles. Large fluctuations in the economy were found to occur roughly every ten years. These fluctuations were thought to be due

[26]However, one should be careful in extrapolating too much from the existence of power law behaviour. While complexity may be at work 'behind the scenes', power laws can spring from rather more innocuous sources. Take Zipf's Law for example. This says that the frequency of the nth most common word in some text is inversely proportional to n. Surely we cannot expect this phenomena and the stylized facts of financial markets to have a common origin — though Zipf [1949] came close to saying something like this! See Newman [2005] for a very clear-headed review of power laws and their interpretation, including the various mechanisms by which they can be generated. In practice it is also often difficult to distinguish power-laws from other distributions: see [Laherrère and Sornette, 1998; Pisarenko and Sornette, 2006].

to external factors. Stanley Jevons [Jevons, 1884] too drew attention to similar regularities in his work on commercial fluctuations. Jevons famously argued (on the basis of a statistical correlation) that the fluctuations were the result of similar fluctuations in the weather (the cycles of sunspots).

Microsimulation models (such as the 'agent-based' models mentioned above) seem to fall somewhere between instrumentalism and realism. For example, the model of Lux (an economist) and Marchesi (an engineer) [1999] involves heterogeneous trading strategies: there are 'noise traders' or 'chartists' on the one hand (whose decisions are based on the price histories) and 'fundamentalists' on the other (whose decisions are based on the notion that there is a *fundamentally* correct price, namely the discounted sum of future earnings). Switches between strategies are also possible. From this setup they are able to recover statistical aspects of real markets. (For a methodologically similar approach see [Bak *et al.*, 1997].) Hence, micro-simulations constitute an attempt to provide an explanation of what were "hitherto mysterious statistical findings like the fat tails and clustered volatility of financial markets" [Lux and Heitger, 2001, p. 123]. Are these econophysics models? It seems not, though they are clearly *inspired* by statistical mechanics. Given the availability of other effective models we ought, then, to view econophysics as one of several ways of making sense of the complexity of financial systems. That is not to say that these ways are disparate: microsimulations are obviously used in statistical physics and the conceptual connections are readily apparent — i.e. both involve the idea of *generating* macrophenomena from microbehaviour.

The stylized facts hold further significance on a more conceptual level: the stylized facts encode non-trivial social regularities. It is the *universality* of these regularities that most interests econophysicists — after all, physicists are in the business of laws and invariances. The stylized facts appear to be invariances that reappear over apparently unrelated systems suggesting that some common underlying mechanism is responsible. The case for comparison here is with the theory of critical phenomena and phase transitions. During a phase transition a system will shift from a relatively disordered global state to a more ordered global state. Or, in other words, parts go from not imitating one another to imitating one another, so that everything depends on everything else (infinite correlation length): a shift in one part propagates (thanks to massive connectivity) to every other part. In such 'critical' conditions the system is said to be 'scale free' so that events of any size can occur, corresponding, of course, to the fat tails which can exhibit themselves as (not infrequent) stock market bubbles and crashes.[27] We turn to the statistical physics explanation in the next section. We consider the issue of the lawhood of the stylized facts in the subsequent section.

[27]The phase transition suggestion has been taken further with specific applications of spin-system models. The idea here is that stock prices respond to demand in the same way that the magnetization of an interacting spin-system responds to changes in the magnetic field (see [Plerou *et al.*, 2002]).

7 STYLIZED FACTS ACCORDING TO ECONOPHYSICS

In contrast to the standard model, which views the stylized facts as the result of exogenous factors, the econophysics approach views them as emergent properties resulting from the internal dynamics (the interactions that occur between individual traders). In a nutshell: the traders' local interactions (the market's microstructure) determines a global pattern (the market's macrophenomena), which feeds back to the traders' future behaviour. This perspective is developed from parts of statistical physics, where one deals with many interdependent parts that 'imitate' each others' behaviour (collective phenomena). In such cases, systems of *prima facie* extremely diverse nature and constitution are found to behave in the same way (the feature physicists call *universality*). According to econophysics, financial markets are part of this class of systems too.

Thus, the ground is set to apply a host of techniques from statistical physics: simply view financial markets as complex systems whose 'particles' are the traders and apply the techniques as usual. As a further step one then constructs 'microsimulations' (or 'agent based models') to test the models and to study the mechanism whereby the macrophenomena emerge from the microstructure ([Levy *et al.*, 2000] offers a comprehensive treatment). These simulations attempt to generate properties of financial markets, such as the stylized facts, from the interactions and properties of traders, thus mimicking statistical physics' models in which particles' behaviour generates properties of a unit complex system — see, again, Lux and Marchesi [1999] for a good example and [Tesfatsion and Judd, 2006] for a more comprehensive treatment.

7.1 The statistical physics analogy

Very often in systems with interacting parts, and whose interacting parts generate properties of the unit system, one finds that the thus generated properties obey scaling laws. Scaling laws tell us about statistical relationships in a system that are invariant with respect to transformations of scale (once certain transformations have been carried out on various parameter values). In statistical physics these scaling laws are viewed as emergent properties generated by the interactions of the microscopic subunits. Scaling laws are explained, then, *via* collective behaviour amongst a large number of mutually interacting components. The components in this financial case would simply be the market's 'agents' (traders, speculators, hedgers, etc...). These laws are 'universal laws', independent of microscopic details, and dependent on just a few macroscopic parameters (e.g. symmetries and spatial dimensions). Econophysicists surmise that since economic systems consist of large numbers of interacting parts too, perhaps scaling theory can be applied to financial markets; perhaps the stylized facts can be represented by the universal laws arising in scaling theory. This analogy is the motivation behind a considerable chunk of work in econophysics; it is through this analogy, then, that the stylized facts receive their explanation — though, as I have said, presumably not their *ultimate* explanation which will involve such things as the agents' psychology, the institutions in which the agents operate, and so on (*cf.* [Lux, 2001, p. 562]).[28]

The power-law behaviour of financial instruments can be explained in terms of the scaling laws that arise in the theory of critical phenomena: complex phenomena involving the collective behaviour of a family of subunits produce such power-laws. Power-laws have a particularly simple form, as we have already seen (but we repeat in a modified form for simplicity). For an event (e.g. an earthquake, stock market crash or price fluctuation, etc...) of magnitude (energy, size, etc...) $\|x\|$ (or greater) the probability $\mathrm{Prob}(x)$ that x will occur is given by:

(10) $\mathrm{Prob}(x) \sim \|x\|^{-\alpha}$

Here, the exponent $-\alpha$ — now called the *critical* exponent in the context of the theory of critical phenomena — is set by the empirically observed behaviour of the system in question. Systems with identical critical exponents belong to the same 'universality class' and will exhibit similar statistical properties (near their critical points). Such systems are interesting because they do not have an 'average' $\overline{x}$ or a 'width' σ. What this means is that there is a greater chance for complex systems having massive events than there is for systems that fit a normal distribution. Hence, financial markets are complex systems because they exhibit such massive events more often than normally distributed systems in a manner consistent with power law behaviour (often viewed as a 'signature' of complexity).

7.2 Scaling, universality and criticality

Most of the research conducted in econophyics is based on an analogy between financial markets and scaling theory and theory of critical phenomena. We present the argument explicitly in §7.3, first we spell out some of the details.

Given some observable $\mathcal{O}$ and driving parameter x, a scale-invariant function can be defined by the functional equation (where $x \to \lambda x$ is an arbitrary rescaling):

(11) $\mathcal{O}(x) = \mu\mathcal{O}(\lambda x)$

The solution of this equation is the functional relation (a power law, in fact):

(12) $\mathcal{O}(x) = x^{\alpha}$

[28]Note that there have been 'ultimate' explanations for these facts that do not involve the sociobiological characteristics of the economic agents. The general approach to power-law distributions and complex behaviour in nature (including the biological and social realm) given in the theory of 'self-organized criticality' [Bak, 1996] is supposed to accomplish such a general explanation. The idea is that complex systems spontaneously tune themselves (their parameters) to the critical values required for power law (scale-free) behaviour to emerge. However, this approach is controversial: it is a moot point, to say the least, whether the mechanism could be the same across social, biological, and more fundamental systems. See [Frigg, 2003] for a critical discussion.

That is, equations of this functional form are scale-invariant. They are of interest in statistical physics (and econophysics) because many-body systems that are close to critical (or bifurcation) points obey such power laws. Renormalization group theory analysis shows that there are universal properties in such systems (systems in the same universality class), meaning that diverse systems share the same critical exponents (and scaling behaviour) and so display qualitatively identical macroscopic properties (when approaching criticality), for a certain class of 'fluctuation-based' properties. Of course, stock markets are not going to always be poised at a critical point, so one expects to see different regimes separated by phase transitions. Kiyno *et al.* [2006] show that such behaviour can be found in S&P500 market data. When financial markets occupy a regime near to a critical point the behaviour corresponds to the volatility clustering, akin to the spins in a magnet aligning versus pointing in the same direction — otherwise there is disordered behaviour well approximated by (geometric) Brownian motion. The large fluctuations correspond to the scale invariance of systems near to critical points. One sees in this way how a unified account of the three stylized facts emerges from a statistical physics based account.

7.3 Unpacking the econophysics argument

It is an uncontested fact that financial market time series display statistical regularities (whether they constitute *laws* or not *is* contested). These regularities have similar characteristics to those obeyed by other complex systems in the physics of critical phenomena. In particular, one can interpret the stylized facts as a consequence of scaling laws. I think we can discern at the root of a great deal of econophysics research the following argument from analogy:

> (P1) Financial markets are made up of a large number of interacting agents
> (P2) According to statistical physics, physical (natural) systems that are composed of large numbers of interacting individuals, and are near critical points, follow scaling laws that are universal and issue in statistical features characteristic of power law distributions
> (P3) Financial markets do exhibit universal regularities characteristic of power law distributions and that show up as stylized facts in their time series
>
> ---
>
> (C) The stylized facts are the consequences of scaling laws of the kind found in statistical physics

In other words, given that financial markets have a physical composition relevantly like that of systems dealt with in statistical physics (large numbers of interacting individuals) and given, furthermore, that the time series exhibit statistical regularities similar to that of systems dealt with in statistical physics, it follows that a good modeling strategy is to apply statistical physics to financial markets. It is clear that the argument is unsound: there are both markets and microsimulation models that would falsify P1. However, the argument is clearly presented as a 'plausibility argument':

> Simply put, statistical physicists have determined that physical systems which consist of a large number of interacting particles obey universal laws that are independent of the microscopic details. This progress was mainly due to the development of scaling theory. Since economic systems also consist of a large number of interacting units, it is plausible that scaling theory can be applied to economics. [Stanley *et al.*, 1996c, p. 415]

A further step is to take statistical physics as providing 'physically realistic' models capable of *explaining* financial phenomena — something Stanley and his group seem to endorse. But clearly the argument is not deductively valid; nor is it intended to be. It is intended to increase our confidence in the application of statistical physics to financial markets and our confidence in the complexity of financial markets. Given that scaling theory and the theory of critical phenomena are associated with complex systems, it would further follow that financial markets are complex systems, in the sense that they undergo phase transitions and, at least some of the time, exist near criticality (between order and chaos, if you like that way of speaking). Since this behaviour is seen too, then the position is further bolstered.

8 STYLIZED FACTS AS LAWS OF NATURE

An old theoretical headache for economists is the apparent inability of economic theory to predict specific, quantitative outcomes. This seems to stem from a lack of *laws* of the 'universal' kind found in physics and chemistry (and certain parts of biology). Inasmuch as economic theory does permit the generation of predictions they are either heavily localized (to some specific region or period) or else severely hedged (e.g. on the continued stability of some inherently unstable process), or both. Otto Neurath made this point well in an early discussion on sociological predictions:[29]

> [M]ost sociological regularities that support the deduction of predictions are formulated in such a way that they are valid only for relative complex structures of certain geographical regions and historical

[29]Lionel Robbins is rather more blunt on the matter: 'not one single "law" deserving of the name, not one quantitative generalisation of permanent validity has emerged from their efforts [the economists — DR]" [Robbins, 1984, p. 104]. However, he is quick to qualify this: economic laws are just as necessary as any other laws in the sense that if the conditions postulated by the law are given, the consequence will follow. Clearly, much rests on these "conditions" that must be "given" (and excluded) — of course, this difficulty is usually summarized under the banner of *ceteris paribus* conditions.

> periods. If one has established, say, constant quantitative relations between changes in illiteracy and criminality in some towns of certain region during a certain period, then one may suspect that the same quantitative relations will also be valid to some extent for the other towns of the area, but one would be hesitant to assume that they will hold for other regions and other periods, let alone towns that would be discovered on Mars. [Neurath, 2004, p. 506]

In other words, though there are laws, they are strongly sensitive on boundary conditions, unlike many laws in physics that are more robust — the universal law of gravitation being such an example. In fact, Alfred Marshall argued that the comparison to exact laws such as the gravitational law was not appropriate.[30] Instead, he argued that the "laws of economics are to be compared with the laws of the tides" [Marshall, 1961, p. 32]. Given the complexity of the details that must go into the initial conditions that combine with the law of gravitation (including other laws governing the stretching and squeezing of the Earth, and so on) to generate predictions about the tides, the predictive accuracy is severely weakened. Sufficient weakening will render the predictions at best qualitative and probabilistic. This seems to be what happens in the case of economic predictions and can account for their apparent localisation — or, in the jargon of experimental design theory, their low external validity.

Johnson *et al.* [2003] claim that "[o]ne of the significant contributions of Econophysics has been to establish that the price statistics across a wide range of supposedly different financial markets worldwide, exhibit certain universal properties" (p. 57). The existence of these stylized facts is then taken to point to a shared structure or process generating the statistics in each case (chance, self-organized criticality, etc.). What distinguishes econophysics from other approaches in economics is the *interpretation* of the stylized facts. In addition to viewing them as emergent properties of a complex system, it also includes a greater commitment to the stylized facts, treating them as genuine laws rather 'mere' local regularities. Part and parcel of this view of the stylized facts as genuine laws is, as Coolen [2004] puts it, that "[i]t is irrelevant...whether the microscopic variables at hand represent coordinates of colliding particles, microscopic magnets, replicating polymers, or (as in the present case) the actions of decision-making agents" (pp. 1-2). That is, the laws themselves are emergent in the sense that they don't depend on the microscopic details of the system.

However, this stance is at odds with (econophysicist) Joseph McCauley [2004], who views the addition of decision-making agents (with free will) as very relevant indeed,[31] even in econophysics: such agents are incompatible with the possibility

[30]In any case, as I point out in fn. 8, even the gravitational law is hedged in a sense, since it depends on a specific value of Newton's constant which it clearly does not possess of necessity.

[31]Similarly, in his Nobel Prize lecture, Trygve Haavelmo notes that "the economics of a society is in fact governed by rules that are themselves a product of man" [Haavelmo, 1997, p. 14]. Haavelmo likes this conclusion since it leaves open the possibility to make society *better* by finding *new* rules. I don't think that this possibility follows only if the rules of economics are

of invariances, and without invariances there are no symmetries, and without symmetries there are no laws. McCauley argues, therefore, that, like economics, econophysics can at best be a *descriptive* historical science (see also [Rosenberg, 1992]), analyzing what happened in some particular economic situation. Other econophysicists (most, in fact) believe that they can find some laws for market dynamics, albeit statistical ones of course; but, says McCauley, "[t]here are no known socioeconomic invariances to support that hope" [2004, p. 4]. Whereas laws of nature are independent of initial conditions, socioeconomic behaviour is not: "socioeconomic behaviour is not necessarily universal but may vary from country to country" (ibid.).

It is of course perfectly true that there will be many differences in socioeconomic behaviour between countries, but that does not mean that there are no invariants. As empirical work by econophysicists and econometrists (and sociologists) has revealed, there *are* some surprising commonalities. It is these very commonalities that kick-started econophysics (at least, most of it). There do seem to be examples of statistical laws that *do not* vary from country to country: Pareto's Law, for example. Pareto investigated the statistical properties concerning the distribution of wealth over the individuals in a population. The model involved is the scale-invariant (cumulative) probability distribution function $p(w \geq x) \sim x^{-\alpha}$. This says that the number of people with income w greater than or equal to x is given by x raised to some power $-\alpha$ (which, for a variety of data sets, Pareto found to be around 1.5). Pareto found that the relationship applies to very different nations, of different sizes and composition: "as different as those of England, of Ireland, of germany, and even of Peru" ([Pareto, 1897], §958; quoted in [Stanley, 2003]). In other words, it is *universal*: multiple distinct nations satisfy the same power-law — i.e. the universality class of Pareto's law includes England, Ireland, Germany, and Peru. Closer, and more recent scrutiny finds universal power-law behaviour too, but only in the tails, for the most extreme (i.e. the richest) 5% or so of the distribution: the first 95% or so is characterized by a more traditional 'log-normal' or exponential curve — see the papers in [Chatterjee *et al.*, 2005] for more details.[32]

But there is something right about McCauley's objection; and he is well aware of the example just given of course. His point is that "*markets merely reflect what we are doing economically*, and the apparent rules of behaviour of markets, whatever they may appear to be temporarily, can change rapidly with time" (*ibid.*, p. 200 — emphasis in original). This non-stationarity is, of course, a general feature of adaptive complex systems: "the empirical distribution is not fixed once and for all by any law of nature [but] is also subject to change with agents' collective

man-made and contingent. For example, the laws of evolution by natural selection are hardly man made, yet we can use them to intervene in Man's condition in a variety of ways.

[32]It is rather interesting to note that Schumpeter, writing of Pareto and his work on 'social invariants', suggested the approach (or one like it) that econophysics now seems to be pursuing: "nobody seems to have realized that the hunt for, and the interpretation of, invariants of this type [Pareto's Law — DR] might lay the foundations for an entirely novel type of theory" [1951, p. 121].

behaviour" (ibid., p. 185). *Prima facie* the problem appears to be more serious in human systems: as Zhou and Sornette put it, "human beings are not spins, they can learn, that is, adapt the nature and strength of their interactions with others, based on past experience" [2006, p. 1].[33]

It is true that the macroscopic (emergent) distribution would be altered if agents altered their (microscopic) patterns *enough.* However, the issue here is whether there is a lawlike relation between certain patterns of (microlevel) behaviour and the distribution: do we find the same distributions appearing in cases where the patterns are a certain way, and in particular when we have systems behaving according to the characteristics of complex systems given in §2? Here the econophysicists who believe they are discovering laws would appear to have some supporting evidence in the form of the scaling laws that Pareto first investigated (but that have been found in a much wider variety of economic observables). In truth, the evidence is not conclusive.

An alternative response to McCauley might be to point to the fact that it is the norms and institutions that circumscribe economic behaviour (and these too can arise *à la* complex systems theory). Since there are common norms and institutions in many and varied countries we should expect to find the statistical regularities we do find. These norms and institutions *could* be altered, resulting in the regularities vanishing but while there is some common socioeconomic system in place there will be common emergent properties (these are the stylized facts). Of course, this ultimately concedes the point to McCauley. It seems that at best only a weakened account of laws can be sustained: Laws relative to the socioeconomic system that is in place. This can perhaps be rendered universal in the physicist's sense since we are at liberty to say that whenever the system is set up in *this* way (with large numbers of parts interacting so as to generate a complex system) it will exhibit *these* properties (i.e. the stylized facts). Indeed, this doesn't seem to be at odds with laws in physics since even the most fundamental are hedged in *some* way: Einstein's law describing the relationship between gravity and mass-energy, for example, is not scale-invariant; it gets modified at the Planck scale (of the order $10^{-33}cm$). Likewise, the field equations of the standard model of particle physics can only be to be taken to apply in a universe without gravity. What I am suggesting here is similar: in a socioeconomic system where there is free-trade, and the trading is done in such and such a way, then we get a set of corresponding distributions of events for the observables — my thanks to Clifford Hooker for pressing this point on me.[34]

[33]An interesting connection can perhaps be made here with MacKenzie's study of financial markets and their relationship to financial models [2006]: as theories and models about the workings of markets change, the way trading is done changes in response and so, as a consequence, do the financial observables and their distributions. (See also [Johnson *et al.*, 2003, pp. 223–4] for a similar point.)

[34]Indeed, I think we can go further: the laws of all fundamental physical theories depend on the values of some fundamental constants and various free parameters. If we adopt a 'multiverse view' in which the assignment of values instantiated in our universe constitute just one amongst many possible ways, then (barring the development of 'multiversal laws') we must also view even

9 ARE FINANCIAL MARKETS *REALLY* COMPLEX?

Do financial markets possess the characteristics of complex systems? Econophysicists (and, indeed, many economists) certainly view financial markets as particularly fine examples of complex systems, as cases of 'complexity in action.' For example, Lisa Borland [2005] writes that "[p]erhaps one of the most vivid and richest examples of the dynamics of a complex system at work is the behaviour of financial markets" (p. 228). Stanley *et al.* [2001] write that "[t]he economy is perhaps one of the most complex of all complex systems" (p. 3). Johnson *et al.* [2003] claim that "it would be hard to find anyone who disagreed with the statement that a financial market is indeed a 'complex' system" (p. 2) — however, there *are* dissenting voices, or at least those who think that the evidence isn't as good as the previous quotations suggest.

Durlauf [2005] argues that the evidence from the econophysicists research and empirical work is not conclusive evidence in favour of economic complexity since the evidence is underdetermined by alternative approaches. Recent work by Pisarenko and Sornette [2006] also pours water on the flames for similar reasons: they show that the power law model at best provides an approximation of the behaviour of market returns. The real story is much more complex. The lesson they draw is that we shouldn't assume that power law behaviour will extend into the unobserved regions of the distribution's tail. In particular, there are plenty more fat-tailed distributions[35] that offer as good an approximation to many data sets in finance (and social science in general).

Likewise, Brock [1999] is concerned with the kinds of process that are capable of generating the data that enter financial time series. He points to an underdetermination of stochastic processes by scaling laws.[36] In other words, one and the same scaling law is compatible with multiple distributions. This is an *identification problem*: "uncovering and estimating the underlying causal data generating mechanism" [Brock, 1999, p. 411]. The mere isolation of scaling laws in the data is not sufficient to enable the making of inferences about the nature of the data generation process.

He argues that scaling laws are useful in a limited way: they function as constraints on the underlying causal data generating process. They do not serve as a decisive guide to the kind of process responsible for the data. Likewise, the

---

the most fundamental laws as hedged in exactly the way suggested above. Neurath seems to have realised this too (though not invoking the multiverse, of course), writing that "physicists rarely mention that all of their laws can only be stated with certain cosmological reservations; it is more likely that astronomers consider whether the rate of decay of uranium is not dependent upon a cosmological totality that might be different several million years hence" [Neurath, 2004, p. 507].

[35]They give the example of the 'stretched exponential' family of distributions defined by: $\mathcal{P}_{SE}(x)_{\geq u} = 1 - \exp\left[-(x/d)^c + (u/d)^c\right]$ (where $x \geq u$). One can 'fatten' the tails of the distribution by playing with the various constants in this expression.

[36]Hence, we have two levels of underdetermination here: (1) concerning whether the correct distribution really is a power law. (2) concerning the mechanism that generated the distribution (assuming it really is a power law).

scaling laws can restrict the class of distributions — Gaussian distributions are rendered impotent, for example. So the scaling laws can falsify but they cannot provide decisive confirmation.[37] The scaling law research is, as Durlauf puts it, "consistent with complex systems models" but "[the] evidence is far from decisive and is amenable to alternative interpretations" [Durlauf, 2005, p. F226]. Hence, both Brock and Durlauf are not convinced that econophysics has demonstrated economic complexity.

However, this does not negate the claim that financial markets are complex systems — it just means that the claim that they are should be treated with more care and attention than it has received so far within the econophysics community. It has to be said that for most econophysics researchers the complexity of economic systems is an assumption taken for granted rather than something they seek to prove. It simply seems obvious to them that economic systems are complex systems. However, I think that Brock and Durlauf put too much weight on underdetermination: I think we can defend the complexity view from these objections.

The scaling laws are but one of the stylized facts to be met by an approach. One has to explain the closely-related clustering and long-range dependence too. And there are a great many more statistical facts that have to be accommodated by an approach. In principal I think we agree with Brock that the stylized facts function as constraints, but, taken together, this constrains the possible stochastic processes a great deal. It is very difficult to get *any* models that reproduce all of the features. As mentioned in note 37, the statistical physics (econophysics) approach involving scaling laws does an exceptional job at covering a great many stylized facts in a unified way. We can use this achievement to support an 'inference to the best explanation' kind of argument: this seems consistent with the way econophysics talk about complexity.

Still, it seems that the ability to definitely demonstrate the existence of a unique mechanism responsible for the fit to the statistical model involving a power law distribution evades us. Brock and Durlauf are quite right to insist on the application of better statistical testing to see what kind of distribution is implied by a data set.[38]

[37]Philosophers of science are well acquainted with this kind of problem. It is simply the curve-fitting problem (a variant of the problem of underdetermination of theory by data; itself a special case of the problem of induction) in a different guise. We can argue that, for a given data set, *any* general theory can only ever be consistent with it (rather than proven from it). However, just because multiple theories are equally well-cnfirmed by the data does not imply that they are equal *simpliciter*. For example, one theory may have better unifying power (it might cover a larger range of phenomena): econophysics seems to have this virtue; it can accommodate multiple stylized facts across many different markets using the *same* concepts. Other possible distinguishing marks are simplicity, elegance, cohesion with background knowledge, and so on.

[38]Aaron Clauset, Cosma Shalizi, and M. E. J. Newman have developed some impressive (open source) R and Matlab code that is capable of discriminating between a large number of closely related relevant distributions, and this should put an end to the debate on first level of underdetermination (namely, whether a data set does indeed obey a power law distribution). This is available at `http://www.santafe.edu/~aaronc/powerlaws/`. A paper [Clauset *et al.*, 2007] corresponding to the code, dealing with the identification of power laws, can also be found on this

10 CONCLUSION

Intuitively, financial markets appear to be complex and they pass many tests for what we expect a complex system to be like. There is still some doubt as to whether they are *truly* complex (whatever that might mean): when it comes to more rigorous statistical tests of complexity the jury is still out. Be that as it may, the profusion of data collected about financial markets makes them an ideal specimen from the point of view of complexity research: nowhere else do we possess as much data recorded so accurately and at so many time scales. Econophysics has proven itself to be an excellent way of probing this conjectured complexity, treating the socially-based data as if it were generated by a purely natural physics experiment. Moreover, the concepts and methods of econophysics have already begun to spill over into other social sciences (including mathematical sociology and political science), offering a fresh perspective in these fields.

The controversial claims made by econophysicists, *vis-à-vis* traditional economic theory and social laws offer the potential to invigorate old debates in the philosophy of science. Seemingly bringing the 'natural' and 'social' worlds together, it is clearly another chapter (and a novel one) in the saga started by Adolphe Quetelet in his bid to found a *physique sociale*. It is also another instance of the strange coupling between physics and economics: one wonders, for example, whether econophysics is just economics 'catching up' with physics. These themes conspire to make econophysics eminently fit for philosophical consumption. It ought, therefore, to play a central role in future discussions of the philosophy of complex systems science and, I would submit, philosophy of science more generally.

RESOURCES AND FURTHER READING

Articles on econophysics appear in a variety of journals. The primary ones are: *Physica A*[39], *Quantitative Finance*, *Physical Review E*, and *Europhysics Letters*. One can also find pertinent articles in *The European Physical Journal B*, *Fractals*, *Advances in Complex Systems*, *International Journal of Theoretical and Applied Finance*, *Physical Review Letters*, and *Nature*.

Internet resources:

- The Econophysics Forum: `http://www.unifr.ch/econophysics/`.
- Econophysics.org: `http://www.ge.infm.it/~ecph/library/index.php`.
- The Econophysics Blog: `http://econophysics.blogspot.com/`.

web page.

[39]Volume 324, Issue 1-2, 2003, contains the proceedings of the international econophysics conference held in Bali in August 2002.

- The economist J. Barkley Rosser has some excellent econophysics articles freely downloadable from his website: `http://cob.jmu.edu/rosserjb/`.

- A great many articles from H. E. Stanley's group can be found at the website for the Center for Polymer Studies at Boston University: `http://polymer.bu.edu/~hes/econophysics/`.

- A fairly comprehensive bibliography of relevant material can be found at: `http://www.ge.infm.it/~ecph/bibliography/bibliography.html`.

A more general website, which occasionally features econophysics-related news, is *The Complexity Digest*: `http://www.comdig.com/`.

Textbooks:

There are now many econophysics textbooks on the market. I mention just six of the best here:

- The first textbook on econophysics is Rosario Mantegna and Eugene Stanley's *Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, 1999). This is still an excellent book to quickly gain a feel for the nature of econophysics.

- An in-depth guide, primarily focused on risk management, is Bouchard and Potters' *Theory of Financial Risk and Derivative Pricing* (Cambridge University Press, 2003).

- An excellent general guide to the entire field is Voit's *The Statistical Mechanics of Financial Markets* (Springer, 2005).

- A more controversial text covering plenty of philosophical issues is McCauley's *Dynamics of Markets* (Cambridge University Press, 2004).

- In terms of the 'complexity connection' the best books are *Financial Market Complexity* by Johnson *et al.* (Oxford University Press, 2003) and *Minority Games* by Challett *et al.* (Oxford University Press, 2004).

ACKNOWLEDGMENTS

My thanks to Enrico Scalas for many useful comments on this and a related paper. My thanks also to the editor, Clifford Hooker, who gave this chapter a thorough shake-up, resulting in many improvements.

BIBLIOGRAPHY

[Amaral and Ottino, 2004] L. A. N. Amaral and J. M. Ottino. Complex networks: Augmenting the framework for the study of complex systems. *Euro Physics Journal B*, 38:147–162, 2004.

[Baaquie, 2004] B. E. Baaquie. *Quantum Finance: Path Integrals and Hamiltonians for Options and Interest Rates*. Cambridge University Press, 2004.

[Bak *et al.*, 1997] P. Bak, M. Paczuski, and M. Shubik. Price variations in a stock market with many agents. *Physica A*, 246:430–440, 1997.

[Bak, 1996] P. Bak. *How Nature Works: The Science of Self-Organized Criticality*. Copernicus, New York, 1992.

[Binney *et al.*, 1992] J. J. Binney, N. J. Dowrick,A. J. Fisher, and M. E. J. Newmann. *The Theory of Critical Phenomena: An Introduction to the Renormalization Group*. Oxford University Press, New York, 1996.

[Black and Scholes, 1973] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–659, 1973.

[Borland, 2005] L. Borland. Long-range memory and nonextensivity in financial markets. *Europhysics News*, 36:228–231, 2005.

[Bouchaud and Potters, 2003] J-P Bouchaud and M. Potters. *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*. Cambridge University Press, Cambridge, 2003.

[Bouchaud *et al.*, 2004] J-P. Bouchaud, Y. Gefen, M. Potters, and M. Wyart. Fluctuations and response in financial markets: The subtle nature of 'random' price changes. *Quantitative Finance*, 4:176–190, 2004.

[Burns and Mitchell, 1946] A. F. Burns and W. C. Mitchell. *Measuring Business Cycles*. National Bureau of Economic Research, Studies in Business Cycles, No. 2: New York, 1946.

[Challet *et al.*, 2005] D. Challet, M. Marsili, and Yi-Cheng Zhang. *Minority Games: Interacting Agents in Financial Markets*. Oxford University Press, Oxford, 2005.

[Chatterjee *et al.*, 2005] A. Chatterjee, S. Yarlagadda, and B. K. Chakrabarti, editors. *Econophysics of Wealth Distributions*. New Economic Windows. Springer, 2005.

[Cohen, 1994] I. Bernard Cohen. *Interactions: Some Contacts between the Natural Sciences and the Social Sciences*. M.I.T. Press, 1994.

[Brock, 1999] W. A. Brock. Scaling in Economics: A Reader's Guide. *Industrial and Corporate Change*, 8(3):409–446, 1999.

[Clauset *et al.*, 2007] A. Clauset, C. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. arXiv:physics/0706.1062, 2007.

[Clegg, 2006] R. G. Clegg. A Practical Guide to Measuring the Hurst Parameter. arXiv:math.ST/0610756, 2006.

[Cont and Bouchaud, 2000] R. Cont and J-P. Bouchaud. Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics*, 4:170–196, 2000.

[Cont, 2001] R. Cont. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.

[Coolen, 2004] A. C. C. Coolen. *The Mathematical Theory of Minority Games: Statistical Mechanics of Interacting Agents*. Oxford University Press, Oxford, 2004.

[Davis and Etheridge, 2006] M. Davis and A. Etheridge. *Louis Bachelier's Theory of Speculation: The Origins of Modern Finance*. Princeton University Press, 2006.

[Durlauf, 2005] S. Durlauf. Complexity and empirical economics. *The Economic Journal*, 115(504):F225–F243, 2005.

[Einstein, 2002] A. Einstein. What is the theory of relativity? In *The Collected Papers of Albert Einstein, Vol.7: The Berlin Years*, pages 100–105. Princeton University Press, Princeton, 2002.

[Fama, 1970] E. Fama. Efficient capital market: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417, 1970.

[Machover and Farjoun, 1983] E. Farjoun and M. Machover. *Laws of Chaos: Probabilistic Approach to Political Economy* . Verso Books, 1983.

[Farmer *et al.* , 2005] J. Doyne Farmer, M. Shubik, and E. Smith. Is economics the next physical science? *Physics Today*, 58(9):37–42, 2005.

[Feigenbaum, 2003] J. Feigenbaum. Financial physics. *Reports on Progress in Physics*, 66:1611–1649, 2003.

[Foley, 1994] D. K. Foley. A statistical equilibrium theory of markets. *Journal of Economic Theory*, 62(2):321–345, 1994.

[Frigg, 2003] R. Frigg. Self-organised criticality - what it is and what it isn't. *Studies in History and Philosophy of Science*, 34(3):613–632, 2003.

[Gallegati *et al.*, 2006] M. Gallegati, S. T. Keen, Lux, and P. Ormerod. Worrying trends in econophysics. *Physica A*, 370:1–6, 2006.

[Gieryn, 1999] T. F. Gieryn. *Cultural Boundaries of Science: Credibility on the Line*. University Of Chicago Press, 1999.

[Gligor and Ignat, 2001] M. Gligor and M. Ignat Econophysics: a new field for statistical physics? *Interdisciplinary Science Reviews*, 26(3):183–190, 2001.

[Guala, 2005] F. Guala. *The Methodology of Experimental Economics*. Cambridge University Press, 2005.

[Haavelmo, 1997] T. Haavelmo. "Econometrics and the Welfare State." *The American Economic Review* **87**(6): 13-15.

[Hildreth, 1986] C. Hildreth. *The Cowles Commission in Chicago, 1939-1955*. Springer, 1986.

[Hooker, 2004] C. Hooker. Asymptotics, reduction, and emergence. *British Journal for the Philosophy of Science*, 55:435–479, 2004.

[Ilinski, 2001] K. Ilinski. *Physics of Finance: Gauge Modelling in Non-equilibrium Pricing*. John Wiley & Sons, LTD, 2001.

[Ingrao and Israel, 1990] B. Ingrao and G. Israel. *The Invisible Hand: Economic Equilibrium in the History of Science*. M.I.T. Press, 1990.

[Jevons, 1884] W. S. Jevons. On the Study of Periodic Commercial Fluctuations. Reprinted in D. F. Hendry and M. S. Morgan, editors, *The Foundations of Econometric Analysis*, pages 113–121. Cambridge University Press, Cambridge, 1995.

[Johnson *et al.*, 2003] N. F. Johnson, P. Jefferies, and Pak Ming Hui. *Financial Market Complexity: What Physicists can Tell us About Market Behaviour*. Oxford University Press, Oxford, 2003.

[Kim, 2003] J. Kim. Supervenience, emergence, realization, reduction. In M. Loux and D. Zimmerman, editors, *The Oxford Handbook of Metaphysics*, pages 556–584. Oxford University Press, 2003.

[Kiyono, 2006] K. Kiyono, Z. R. Struzik, and Y. Yamamoto. Criticality and Phase Transition in Stock-Price Fluctuations. *Physical Review Letters*, 96:068701-1–068701-4, 2006.

[Koopmans, 1947] T. Koopmans. Measurement Without Theory. *The Review of Economics and Statistics*, 29(3): 161–172, 1947.

[Laherrère and Sornette, 1998] J. Laherrère and D. Sornette. Stretched exponential distributions in nature and economy: Fat tails with characteristic scales. *Euro Physics Journal B*, 2:525–539, 1998.

[Latora and Marchiori, 2004] V. Latora and M. Marchiori. The architecture of complex systems. In M. Gell-Mann and C. Tsallis, editors, *Nonextensive Entropy — Interdisciplinary Applications.* Oxford University Press, Oxford, 2004.

[LeBaron, 2006] B. LeBaron. Agent-based financial markets: Matching stylized facts with style. In D. Colander, editor, *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, pages 221–235. Cambridge University Press, Cambridge, 2006.

[Levy *et al.*, 2000] M. Levy, H. Levy, and S. Solomon. *Microscopic Simulation of Financial Markets. From Investor Behavior to Market Phenomena.* Academic Press, 2000.

[Lux and Ausloos, 2002] T. Lux and M. Ausloos. Market fluctuations i: Scaling, multiscaling, and their possible origins. In A. Bunde, J. Kropp, and H. J. Schnellnhuber, editors, *The science of disasters: Climate disruptions, heart attacks, and market crashes*, pages 373–409. Springer-Verlag, Berlin, 2002.

[Lux and Heitger, 2001] T. Lux and F. Heitger. Micro-simulations of financial markets and the stylized facts. In H. Takayasu, editor, *Empirical Science of Financial Fluctuations: The Advent of Econophysics*, pages 123–134. Springer-Verlag, 2001.

[Lux and Marchesi, 1999] T. Lux and M. Marchesi. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 397:498–500, 1999.

[Lux, 1996] T. Lux. The stable paretian hypothesis and the frequency of large stock returns: An examination of major german stocks. *Applied Financial Economics*, 6(6):463–475, 1996.

[Lux, 2001] T. Lux. Power laws and long memory. *Quantitative Finance*, 1(6):560–562, 2001.

[MacKenzie, 2006] D. MacKenzie. *An Engine, Not a Camera: How Financial Models Shape Financial Markets.* M.I.T. Press, 2006.

[Mandelbrot, 1963] B. Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 36:394–419, 1963.

[Mantegna and Stanley, 2000] R. N. Mantegna and H.E. Stanley. *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, 2000.

[Mantegna, 2005] R. N. Mantegna. Presentation of the english translation of Ettore Majorana's paper: The value of statistical laws in physics and social sciences. *Quantitative Finance*, 5(2):133–140, 2005.

[Marshall, 1961] A. Marshall. *Principles of Economics (9th Edition)*. Macmillan, 1961.

[Mayr, 1970] E. Mayr. *Populations, Species, and Evolution*. Harvard University Press, Cambridge, Mass, 1970.

[McCauley, 2004] J. L. McCauley. *Dynamics of Markets: Econophysics and Finance*. Cambridge University Press, Cambridge, 2004.

[Merton, 1973] R. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[Mirowski, 1989] P. Mirowski. *More Heat than Light: Economics as Social Physics, Physics as Nature's Economics*. Cambridge University Press, Cambridge, 1989.

[Mitchell, 1913] W. N. Mitchell. *Business Cycles*. Berkeley, 1913.

[Neurath, 2004] O. Neurath. *Economic Writings: Selections 1904-1945* (Edited by T. E. Uebel and R. S. Cohen). Springer, 2004.

[Newman, 2005] M. E. J. Newman. Power laws, Pareto distributions, and Zipf's law. *Contemporary Physics*, 46(5):323–351, 2005.

[O'Connor and Wong, 2005] T. O'Connor and H. Y. Wong. The metaphysics of emergence. *Noûs*, 39(4):658–678, 2005.

[Osbourne, 1959] M. F. M. Osborne. Brownian motion in the stock market. *Operations Research*, 7(2):145–173, 1959.

[Pareto, 1897] V. Pareto. *Cours d'Economie Politique*. Rouge, Lausanne, Switzerland, 1897.

[Pisarenko and Sornette, 2006] V. Pisarenko and D. Sornette. New statistic for financial return distributions: Power law or exponential? *Physica A*, 366:387–400, 2006.

[Plerou *et al.*, 2002] V. Plerou, P. Gopikrishnan, X. Gabaix, and H. E. Stanley. Quantifying stock-price response to demand fluctuations. *Physical review E*, 66:027104, 2002.

[Rickles, 2007] D. Rickles. Econophysics for philosophers. *Studies in History and Philosophy of Modern Physics*, 38(4):948–978, 2007.

[Robbins, 1984] L. Robbins. *The Nature and Significance of Economic Science* . New York University Press, 1984.

[Roehner, 2005] B. M. Roehner. *Patterns of Speculation: A Study in Observational Econophysics*. Cambridge University Press, Cambridge, 2005.

[Rosenberg, 1992] A. Rosenberg. *Economics — Mathematical Politics or Science of Diminishing Returns?* University of Chicago Press, 1992.

[Ross, 2002] S. Ross. *An Elementary Introduction to Mathematical Finance: Options and Other Topics*. Cambridge University Press, Cambridge, 2002.

[Schabas, 2006] M. Schabas. *The Natural Origins of Economics*. The University of Chicago Press, 2006.

[Schelling, 1971] T. C. Schelling. Dynamic models of segregation. *Journal of Mathematical Sociology*, 1:143–186, 1971.

[Schelling, 1978] T. C. Schelling. *Micromotives and Macrobehavior*. W. W. Norton and Co., 1978.

[Schumpeter, 1951] J. Schumpeter. *Ten Great Economists: From Marx to Kenyes*. Oxford University Press, New York, 1951.

[Shefrin, 2002] H. Shefrin. *Beyond Greed and Fear: Understanding Behavioral Finance and the Psychology of Investing*. Oxford University Press, Oxford, 2002.

[Shleifer, 2000] A. Shleifer. *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford University Press, Oxford, 2000.

[Simon, 1955] H. Simon. A Behavioural Model of Rational Choice. *The Quarterly Journal of Economics*, 69(1):99–118.

[Simon, 1981] H. Simon. *The Sciences of the Artificial*. M.I.T. Press, 2 edition, 1981.

[Sober, 1980] E. Sober. Evolution, population thinking, and essentialism. *Philosophy of Science*, 47(3):350–383, 1980.

[Sornette, 2003] D. Sornette. *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton University Press, 2003.

[Stanley *et al.*, 1996a] H. E. Stanley, V. Afanasyev, L. A. N. Amaral, S. V. Buldyrev, A. L. Goldberger, S. Havlin, H. Leschhorn, P. Maass, R. N. Mantegna, C. K. Peng, P. A. Prince, M. A. Salinger, M. H. R. Stanley, and G. M. Viswanathan. Anomalous fluctuations in the dynamics of complex systems: From DNA and physiology to econophysics. *Physica A*, 224:302–321, 1996.

[Stanley *et al.*, 1996c] M. H. R. Stanley, L. A. N. Amaral, S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger, H. E. Stanley. Can statistical physics contribute to the science of economics? *Fractals*, 4(3):415–425, 1996.

[Stanley *et al.*, 1999] H. E. Stanley, L. A. N. Amaral, D. Canning, P. Gopikrishnan, Y. Lee, and Y. Liu. Econophysics: Can physicists contribute to the science of economics? *Physica A*, 269:156–169, 1999.

[Stanley *et al.*, 2001] H. E. Stanley, L. A. N. Amaral, P. Gopikrishnan, V. Plerou, and B. Rosenow. Quantifying empirical economic fluctuations using the organizing principles of scale invariance and universality. In H. Takayasu, editor, *Empirical Science of Financial Fluctuations: The Advent of Econophysics.*, pages 3–11. Springer-Verlag, Berlin, 2001.

[Stanley *et al.*, 2006] H. E. Stanley, X. Gabaix, P. Gopikrishnan, and V. Plerou. Economic fluctuations and statistical physics: The puzzle of large fluctuations. *Nonlinear Dynamics*, 44:329–340, 2006.

[Stanley, 2003] H. E. Stanley. Statistical physics and economic fluctuations: Do outliers exist? *Physica A*, 318:279–292, 2003.

[Tesfatsion and Judd, 2006] L. Tesfatsion and K. L. Judd. *Handbook of Computational Economics, Volume 2: Agent-Based Computational Economics*. North-Holland Publishing Company, 2006.

[Voit, 2005] J. Voit. *The Statistical Mechanics of Financial Markets*. Springer-Verlag, 2005.

[Vega-Redondo, 2007] F. Vega-Redondo. *Complex Social Networks*. Cambridge University Press, 2007.

[Zhang, 1998] Y-C Zhang. Evolving models of financial markets. *Europhysics News*, pages 51–54, March/April 1998.

[Zhou and Sornette, 2006] W. X. Zhou and D. Sornette. Self-organizing ising model of financial markets. *The European Physics Journal B*, DOI: 10.1140/epjb/e2006-00391-6:1–6, 2006.

[Zipf, 1949] G. K. Zipf. *Human Behaviour and the Principle of Least Effort*. Addison-Wesley Press, Cambridge, MA, 1949.

Part VII

Anthropology

COMPLEXITY AND ANTHROPOLOGY

J. Stephen Lansing and Sean S. Downey

Anthropology has historically embraced a rich diversity of questions and methods. For the purposes of this essay it is convenient to group them into the conventional categories of *Geisteswissenschaften* and *Naturwissenschaften*; humanistic and scientific studies. The coexistence of these two approaches within anthropology has itself been an enduring source of controversy; humanistic anthropologists tend to be dubious about the application of scientific methods to anthropological subjects, while the scientists vary in their tolerance for humanistic methods in proportion to their adherence to a Popperian or positivist belief in the unity of scientific method. The entry of the nascent field of "complexity" into anthropology promises to complicate this picture. For our purposes, a simple definition of "complexity" as the study of nonlinear processes will be adequate. Nonlinearities abound in anthropology, and as awareness of their properties spreads it seems inevitable that few topics in anthropology will remain untouched.

This essay begins with a descriptive tour of examples of the application of ideas from complexity to anthropological questions, roughly divided into the Geistes- and Naturwissenschaften. Themes to be touched on, which will be defined below, include emergence, agency, network dynamics, multi-scale interactions, path dependence, evolvability and robustness. In each case, after outlining an anthropological question we consider the extension of simple mathematical models to some interesting empirical examples. The essay concludes with a brief summary of major themes.

1.1 Complexity and the Geisteswissenschaften: structuralism

For roughly the past half century, humanistic approaches to sociocultural anthropology have been dominated by the structural anthropology of Claude Lévi-Strauss, and the "post-structuralism" of his many successors, among them the Tel Quel group in Paris (1960-1982) which included Roland Barthes, Georges Bataille, Maurice Blanchot, Jacques Derrida, Michel Foucault and Julia Kristeva. Structuralism posed a profound challenge to the earlier humanistic tradition in anthropology, which sought to uncover the subjective meaning of cultural symbols and practices. Structuralists did away with the question of the subject's awareness of meaning, replacing it with an account of how language produces meanings that define subjects. The prominent structuralist Roland Barthes (1915-80) argued that the implications of this epistemological reversal could hardly be exaggerated, predicting that the "infinity of language" would replace the Kantian-Husserlian "infinity of consciousness." The ascendancy of structuralism in anthropology in

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

the 1960s created an ongoing philosophical crisis with respect to the nature of the anthropological subject, which continues today.indexmorphemes

Interestingly, it is probably easier to give a coherent account of the structuralist program from the perspective of complexity, than from that of humanistic anthropology. Structuralism defines various components of language, such as phonemes and morphemes, in terms of logical operations on trees or networks. This marked a radical departure from traditional interpretive approaches to language and culture. A century ago, the Swiss linguist Ferdinand de Saussure (1857-1913) defined the linguistic sign as comprised of two elements, the sensible sound-image (signifier) and the intelligible concept (signified). Saussure argued that linguistic signs are unmotivated and acquire their meaning only through differential relations with other signs.[1] He suggested that the same dynamics occur at the level of phonology: the boundaries of phonemes are defined by paired contrasts with the other phonemes that they most closely resemble. Thus in English the slight difference between [p] and [b] marks the boundary between two phonemes, creating a meaningful distinction between, for example, "pit" and "bit." In this way, binary contrasts or antonymy define signifiers: the written phonetic symbol [b] points to a particular sound (or range of sounds produced by different speakers). Roland Barthes later described this as first-order signification; i.e. the denotative meaning of the signifier. Barthes developed a concept of higher-order signifiers which enabled him to extend the structuralist approach from language to cultural phenomena. For example, the denotative or first-order meaning of the English signifier "blue" depends on the other color terms with which it can be contrasted. Barthes argued that second-order meanings are also defined by binary contrasts. Thus blue is traditionally associated with male infants, and pink with female infants, in American hospitals. This association is an example of metonymy: blue is to pink as male is to female. Barthes argued that such metonymic associations are ubiquitous, generating symbolic classificatory systems for cultural objects.

This idea was further developed by anthropologists such as Marshall Sahlins, who used it to analyze the systemic properties of cultural symbols. For example, Sahlins argued that the Fijian words for "sea" and "land" are first-order signifiers defined by their binary opposition: that which is sea is not land [Sahlins, 1976]. This contrast is extended by metonymic chaining: in Fiji, men are associated with the sea and women with the land; further, chiefs are also associated with the sea and commoners with the land. The seaward side of a Fijian house thus is associated with male and chiefly power. Similarly, the sea itself is subclassed into the lagoon (landward sea) and the outer or seawards sea. Fishing is a male occupation, but if women fish, they do so in the lagoon.

In this example, a relationship of binary opposition between two first-order signifiers "sea" and "land", forms the root of a tree of symbolic associations (Fig. 1) in which the initial defining contrast is repeated with other paired oppositions, like seawards land and inland land.

[1] "The linguistic sign unites, not a thing and a name, but a concept and a sound-image." [Saussure, 1983, chapter one].

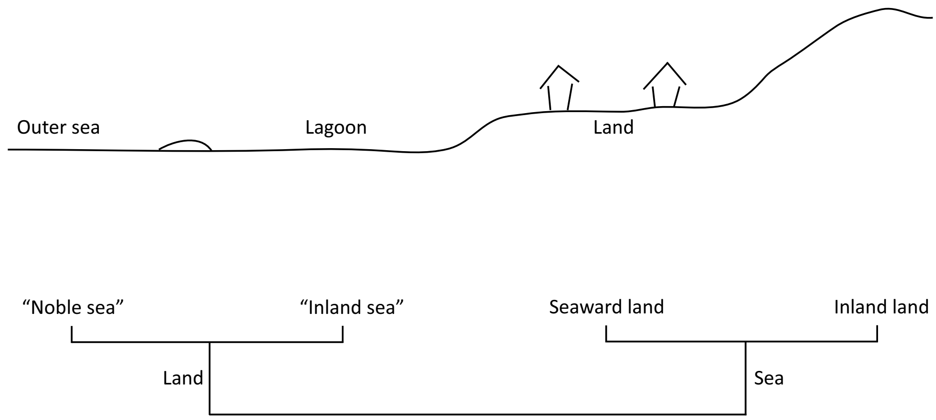


Figure 1. A structuralist analysis of Fijian classification.
Figure redrawn from: Sahlins, M. D. (1976). Culture and practical reason. University of Chicago Press.

The tree model was criticized by post-structuralists, who argued that there are no privileged first-order signifiers which unambiguously root trees of symbolic associations (thus the Sea/Land opposition in Sahlin's example would not be accepted as foundational). According to this argument, signification is not fully defined by any single oppositional pair in the chain of signifiers, but rather by metonymic associations. The post-structuralist psychoanalyst Jacques Lacan argued that the mind glides like a butterfly through networks of signifiers, each of which points beyond itself to other signifiers. Hence the correct model is not a rooted tree, but rather a network of signifiers: the chain of differences spreads throughout semantic space and never comes to rest in an ultimate 'signified' [Sandwell, 1996, 365-6]. This argument was elaborated by Jacques Derrida, who drew attention to the 'free play' of signifiers: they are not fixed to their signifieds but point beyond themselves in an 'indefinite referral of signifier to signified' [Derrida, 1978, 25]. Hence both Derrida and Lacan portray relationships among signifiers as networks, not trees. While for Saussure the meaning of signs derived from how they differ from each other, Derrida coined the term *différance* to allude to the ways in which meaning is endlessly deferred. He concluded that there is no 'transcendent signified' [Derrida, 1978, 278-280; 1976, 20].

Derrida and other post-structuralists famously developed these ideas into a relativistic epistemology, arguing that the meaning of texts can never be fixed. This conclusion echoed that of many Anglo-American analytic philosophers, who at about the same time (1970s) had begun to acknowledge that their quest for an unambiguous observation language had failed. Meanwhile in linguistics, the structuralist program sputtered to an end as it became clear that networks defined by binary oppositions are not very informative for linguistic phenomena more complex than phonemes and morphemes.

1.2 Language networks and the topology of the possible

When Derrida, Barthes and Lacan began their studies of networks of signifiers, not much was known about the mathematical properties of networks. But subsequently this topic moved to the forefront of research in complex systems. In the past decade a number of complexity researchers have begun to pick up where the structuralists left off, exploring the application of network models to language. In the original structuralist models, binary opposition was the sole logical operator, and the problem of logical closure was unresolved. (Thus while phonemes or color terms may form small closed networks in which the boundary of each signifier depends to some extent on the others, this is not obviously true for other signifiers). Like the post-structuralists, complexity researchers investigate the properties of networks of signifiers. But their most recent work extends network analysis to syntax, which the structuralists never attempted, and considers other semantic relationships besides binary opposition, or antonymy.

As with much research in complexity, the study of language networks is often motivated by the physicist's passion for discovering universals. The study of networks and graph theory has revealed many common features in such diverse phenomena as food webs, social networks, software maps, power grids, genomes and neural connections. That languages might also display such regularities was suggested by the early work of George Zipf [1902-1950], who showed that if all the words in a text are ordered by rank, from the most common to the rarest, their frequency (number of appearances) decays inversely with their rank [Zipf, 1949]. Most words are rare, whereas a few (such as *the*, *of*, *and*, *to*, *I*, etc.) are very common. Zipf observed that this relationship appears to hold for all natural (spoken) languages.

A recent study by Ferrer i Cancho *et al.* builds on Zipf's work, using network theory to develop a simple model for the emergence and structure of syntax. Syntax can be thought of as a set of rules to combine words in such a way as to make sentences meaningful. Some authors, such as Derek Bickerton, have argued that syntax required a pre-adaptation of the human brain. Ferrer i Cancho et al. propose a simpler explanation. In their model, words are associated with objects. In accordance with Zipf's law, a few words are very polysemous (refer to many objects), while most denote only one or two. Figure 2a depicts this variation: word 1 denotes object 2, while word 3 denotes two objects, 3 and 4. In Figure 2b, this data is reorganized into a network displaying the linkages of words via their objects of reference. This simple network has many words with few links, but a few, like word 11, act as multi-linked or polysemous hubs. Ferrer i Cancho et al. conclude that connectedness arises naturally from Zipf's law, independently of the details of the linguistic setting [Cancho, *et al.*, 2005].

What is interesting about this network is what is called the "small world" phenomenon: despite its potentially very large size and sparseness (i.e. most words have few links), it is usually easy to reach one node (word) from another along a few links, thanks to the scattered presence of some very well-connected hubs.

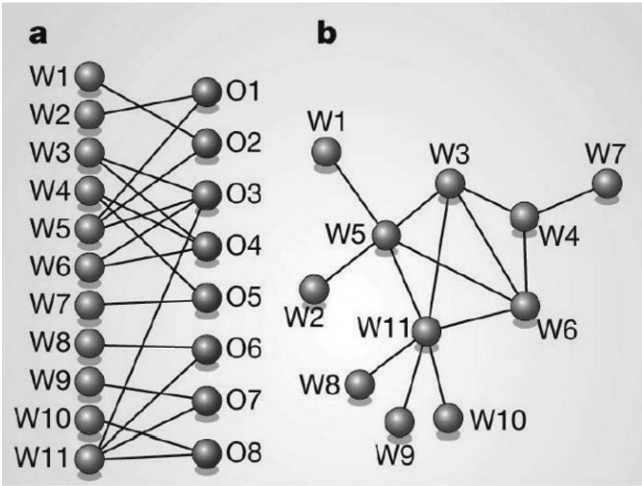


Figure 2. Building a protolanguage network. A) A bipartite set of connections can be built by linking signals (words, red) to objects (blue). Most words are specific, referring to only one or two objects, whereas a few of them have many links. B) A new network is obtained by linking words that share at least one object of reference. The resulting network has many words with few links, but some acting as hubs. Ferrer i Cancho et al. believe that syntactic rules might have emerged from such a scale-free network architecture
Here degree distributions are defined as the frequency $P(k)$ of having a word with k links. Most complex networks are characterized by highly heterogeneous distributions, following a power law (scale-free) shape $P(k) \sim k^{\gamma}$, with $2 < \gamma < 3$.
Figure redrawn with permission from: Sole, R. (2005). Language syntax for free?. Nature, 434(7031), 289.

Ferrer i Cancho et al. suggest that the sometimes illogical and quirky appearance of syntactic rules in natural languages might be merely a by-product of this scale-free network structure, and that Zipf's law may be a necessary precondition for the development of syntax and grammar.

Another physicist, Ricard Solé, has suggested that languages exhibit well-defined network properties at all levels: phonetic, lexical, syntactic, semantic [Solé, 2005, 289]. Like Ferrer i Cancho, he emphasizes the functional significance of the small-world phenomenon. Figure 3 illustrates the patterns formed by various semantic relationships, including binary opposition (antonymy) and hypernymy (words or phrases whose semantic range is included within that of another word). While hypernymy produces tree-like networks, other relationships create shortcuts through the entire semantic network, producing an overall small-world effect.

Another example, also taken from Solé, shows how syntactic networks can be built up from textual data, using a passage from Virginia Woolf's *A Room of One's*

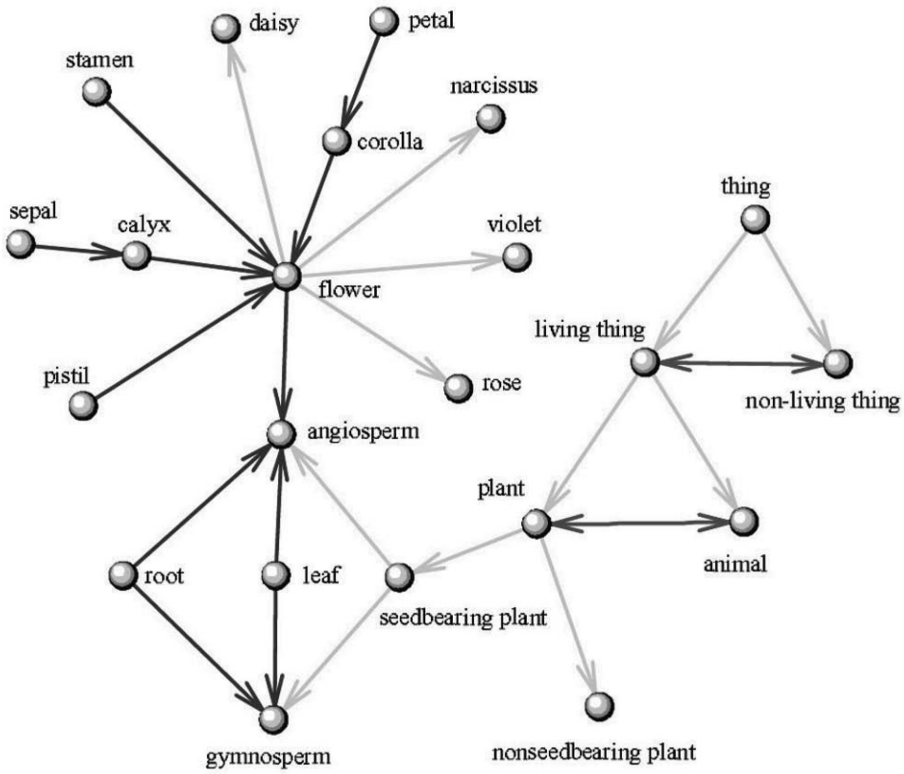


Figure 3. Semantic webs can be defined in different ways. The figure shows a simple network of semantic relations among lexicalized concepts. Nodes are concepts and links semantic relations between concepts. Links are coloured to highlight the different nature of the relations. Yellow arcs define relations of hypernymy (Flower $\rightarrow$ Rose implies that Flower is a hypernym of Rose). Two concepts are related by a blue arc if there is a part-whole relation (metonymy) between them. Relations of binary opposition (antonymy) are bidirectional and coloured violet. Hypernymy defines a tree-structured network and other relations produce shortcuts that leads the network to exhibit a small world pattern, making navigation through the network more easy and effective.

Figures 3-6 redrawn with permission from: R. V. Sol, B. C. Murtra, S. Valverde, L. Steels. Language Networks: their structure, function and evolution. Working Papers of the Santa Fe Institute 05-12-042, 2005.

Own. Here is the passage:

> But, you may say, we asked you to speak about women and fiction — what has that got to do with a room of one's own. I will try to explain. When you asked me to speak about women and fiction I sat down on the banks of a river and began to wonder what the words meant. They might mean simply a few remarks about Fanny Burney; a few more about Jane Austen; a tribute to the Brontës and a Sketch of Haworth Parsonage under snow; some witticism if possible about Miss Mitford; a respectful allusion to George Eliot; a reference to Mrs. Gaskell and one would have done. But at second sight the words seemed not so simple.

Starting from this text several types of language networks can be created based on different types of relationships among words. Figure 4 shows a co-occurrence network; Figure 5 creates the corresponding syntactic network, taking as a descriptive framework dependency syntax [Melčuk, 1988].

The network shown in Figure 5 takes verbs as the nucleus of well-formed sentences, and builds arcs that begin in complements and end in the nucleus of the phrase. Like Ferrer i Cancho's toy model, the resulting structure is a small world network.

The network properties of natural languages define what Walter Fontana has called the topology of the possible,[2] a concept that would surely have pleased Jacques Lacan. The discovery of shared topological features, such as the emergent "small world" properties of both semantic and syntactic networks, promise new insights into the cognitive and social processes that underlie the creation, maintenance and transmission of human languages. Like the structuralists before them, complexity researchers are also beginning to explore the epistemological implications of language topologies. An interesting recent example is a simulation of the emergence of linguistic categories by Puglisi *et al.* [2008]. These researchers begin with the question of how linguistic categories, which are culture-dependent conventions, come to be accepted at a global level without any central coordination. For example, the few "basic" color terms that are present in natural languages are a remarkably consistent subset of an almost infinite number of perceivable different colors. Extensive simulations showed that a simple negotiation scheme, based on memory and feedback, is sufficient to guarantee the emergence of a self-organized communication system that is able to discriminate objects in the world, requiring only a small set of words [Puglisi *et al.*, 2008, 7939].

These initial explorations of the network properties of language suggest several tentative conclusions. Clearly, there are statistical universals in language networks,

[2]Fontana developed this concept in the context of evolutionary biology, with reference to the genotype-phenotype map: "...what is needed is a criterion of *accessibility* of one phenotype from another by means of mutations on their underlying genetic representation. Such a notion of accessibility can then be used to define a concept of *neighborhood* which generates the structure of phenotype space in the absence of a distance notion..." [Fontana, 1993, 15].

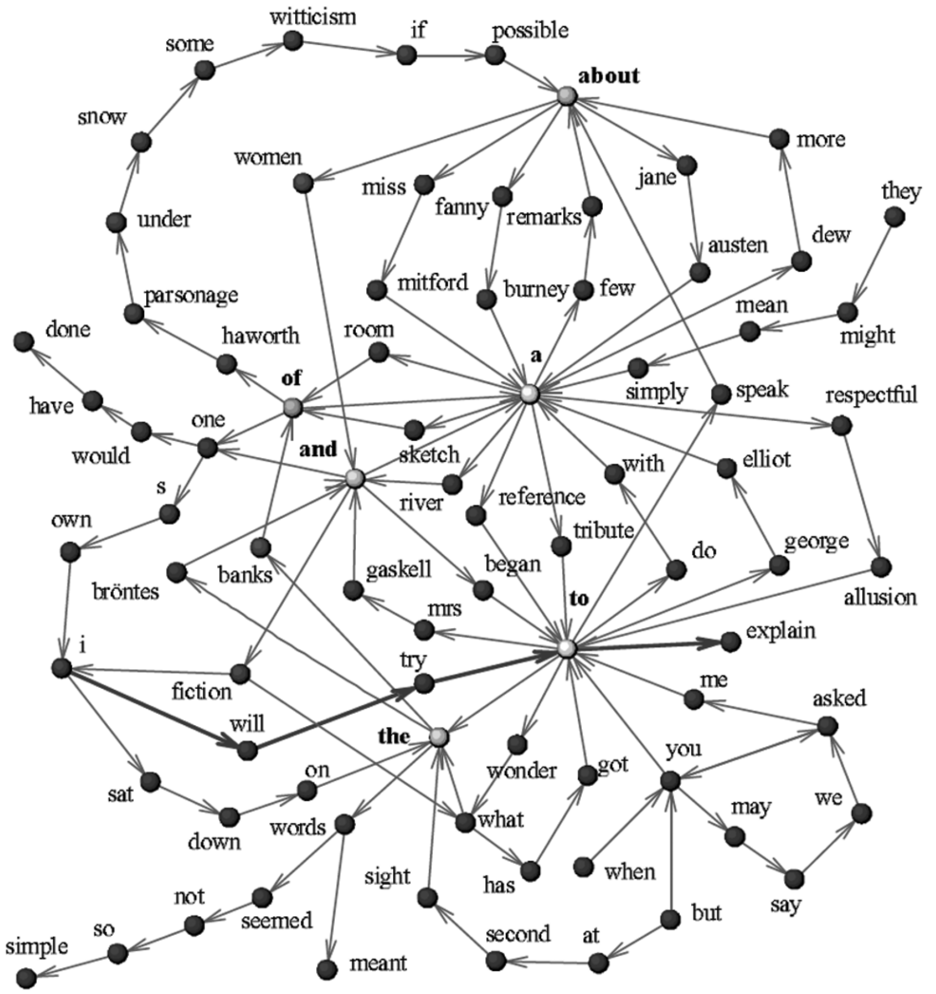


Figure 4. A co-occurrence network. Words are depicted as nodes; the lighter their color, the higher their degree. Paths on this network can be understood as the potential universe of sentences that could be constructed with this lexicon. An example of such path is the sentence shown in red.

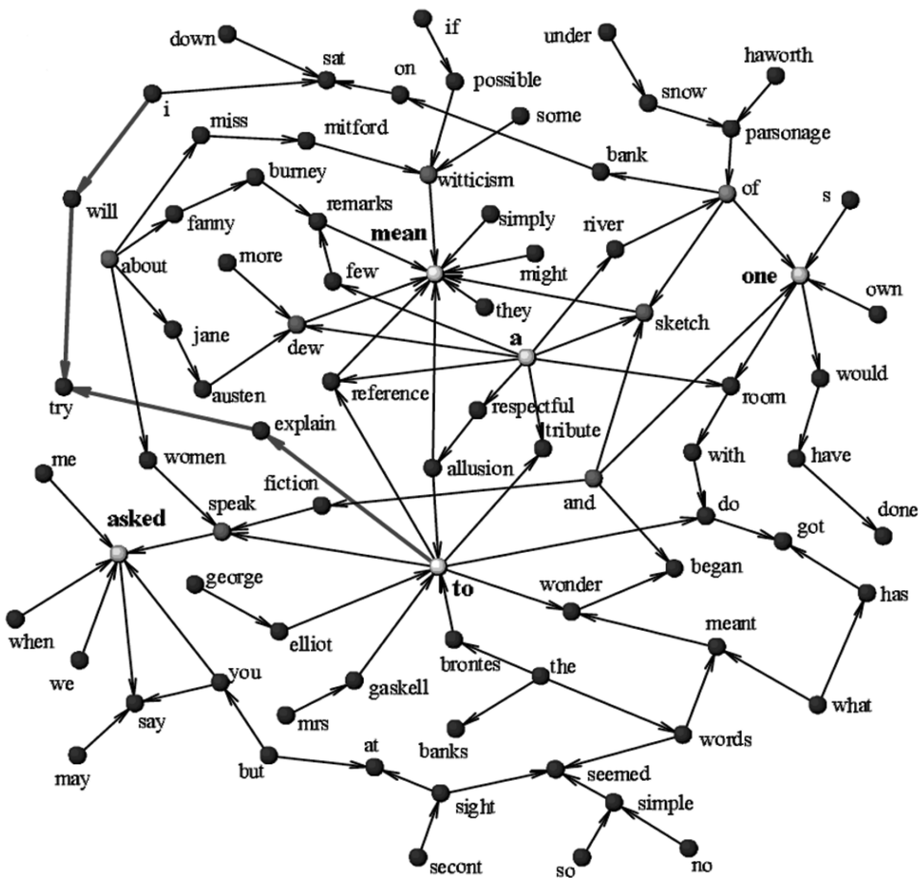


Figure 5. The corresponding syntactic network. The previous sentence appears now dissected into two different paths converging towards “try.”

which are similar to the features found in other 'scale free' networks arising in physics, biology and the social sciences. As Solé *et al.* observe, [2005], this points to new types of universal features of language, which do not focus on properties of the elements in language inventories as the traditional study of universals (e.g. phoneme inventories or word-order patterns in sentences) but rather on statistical properties. Second, the pervasive nature of these network features suggests that language may be subject to the same kinds of self-organization dynamics as other natural and social systems.

1.3 Agency and social networks

Social networks are a venerable topic in the social sciences. But recent work from a "complexity" perspective shifts the analytical focus from descriptive enumerations of social relationships, to abstract models of the effects of dynamical rules on the historical trajectories of networks. As in the studies of language networks described above, these studies explore the topologies of the possible: the consequences of network structures, in terms of the kinds of interactions they facilitate.

In the social sciences, the concept of "agency" has two meanings. One pertains to the capacity of persons to depart from their assigned social roles, to choose not to follow norms; in short, to swim against the tide. Alternatively, agency can also mean a person's social efficacy; their ability to mobilize resources, influence others, or take effective action. Social networks provide a way to carry out comparative, empirical studies of the latter form of agency (e.g. social efficacy). Structural network measures are traditionally used to identify influential actors, based on the assumption that such measures offer a robust way to identify influential individuals in a community [Wasserman and Faust, 1994]. This can become the starting-point for more detailed explorations of topics such as the situatedness of environmental knowledge.

For example, Atran *et al.* used network models to explore differences in rates of deforestation in the vicinity of three communities located in the Petén region of Guatemala: native Itza' Maya, Spanish-speaking immigrant Ladinos, and immigrant Q'eqchi' Maya [Atran *et al.*, 2005]. Petén's forests are a common-pool resource that is rapidly being depleted, and Q'eqchi' forest-clearance rates are more than five times greater than those for Itza'. Measurements of soils, biodiversity, and canopy cover indicate that Itza' promote forest replenishment, while Q'eqchi' foster rapid forest depletion, and Ladinos fall somewhere in between. To discover the reasons for these differences, Atran et al. asked questions designed to elicit both farmers' personal knowledge of ecology, and whom they turn to when making decisions about resource use. They found significant differences in the structure of both types of networks among the three communities, which helped to explain the variation in deforestation rates. Over time, these rates are sensitive to the depth and accessibility of ecological knowledge available to farmers. A similar study by Bodin et al. constructed network models for social relations and ecological knowledge among fishermen in a coastal Kenyan village [Bodin and Crona, 2008]. Like

Atran *et al.*, they found large differences in local ecological knowledge between different occupational groups, such as inshore and offshore fishers.

In these studies, the connectedness of individuals in network diagrams is often interpreted as an index of their social capital (e.g. [Borgatti and Foster, 2003]). Using survey data, researchers can also investigate variation in the "social embeddedness" of knowledge and more specific forms of social capital [Granovetter, 1985]. But as informative as the resulting networks may be, they are essentially snapshots of conditions that existed at the time the survey was done. Recently, several scholars have begun to investigate changes in network structures over time. This added temporal dimension has encouraged the development of new analytical methods. Two of the most interesting examples will be briefly discussed here.

John Padgett and Paul McLean used network analysis to investigate the emergence of financial capitalism in Renaissance Florence at the Medici bank, and sister institutions [Padgett and McLean, 2006]. They argue that "the poisedness of a system to reconfiguration by an invention is as much a part of the phenomenon to be explained as is the system's production of the invention itself" [ibid., 1464]. The particular invention they trace is the discovery, in the late 1300s, of a new organizational form: a set of legally autonomous companies linked through one person or a small set of controlling partners, called a cambio bank. This new "network-star" ownership structure largely displaced earlier legally unitary companies, often built collectively by patrilineal kinsmen, which were common in the early 1300s.

Using archival sources, Padgett and McLean were able to trace multiple networks of relationships (partnerships, kinship, dowries, etc) for nearly complete lists of cambio bankers in four time periods: 1348–58, 1369, 1385–99, and 1427. In 1350 Florentine companies were extremely specialized by industry, but by 1427 they had become much more diversified [ibid., 1482]. The desire to interpret these changing patterns of relationships led Padgett and McLean to expand their concept of agency, to include collective institutions like patrilineages and banks as well as individuals:

> Actors are clusters of relational ties. In the activity plane of economics, for example, collective actors called companies are composed of partnership ties. These companies trade with each other. In the domain of kinship, for another example, collective actors called patrilineages are composed of genealogy ties. These patrilineages marry each other. And in the domain of politics, collective actors called factions are composed of clientage ties. These factions do political deals with each other. [ibid., 1468]

This perspective might seem to eliminate agency in the usual sense altogether. Not so, however, because institutions like banks are treated not as unitary actors, but rather as comprised of multiple networks linking specific individuals. The focus is on "careers and biographies as these wend their way across organizations and domains...". In other words, both organizations and people are shaped, through network coevolution, by the history of each flowing through the other" [ibid., 1470-1]. Changes in one network, such as dowries or marital ties, have consequences for other networks, such as partnerships. As these patterns slowly ripple across social networks, the new forms of banking emerge.

A similar approach was developed by David Stark and Balázs Vedres to analyze the reformation of the Hungarian economy during the transition from communism to capitalism, another period of rapid and profound economic change [Stark and Vedres, 2006]. This study traced changes in the ownership structure of 1,696 of the largest Hungarian enterprises from 1987 to 2001, with particular attention to the role of foreign investment. They argue that "the transformation of a national economy is not a unitary process obeying a single logic but is formed out of the interweaving of multiple processes with distinct temporalities. . . .thus, in place of the properties of the global network we focus on variation in local properties" [ibid., 1399].

To apply network analysis to the ownership data, Stark and Vedres defined ties between firms as equity ownership of at least 1%, among the largest Hungarian firms. Types of ownership were classified into four categories: State, Hungarian firm, Hungarian person, and foreign owner. This enabled them to pose historical questions, such as: when communism ended, did domestic networks grow unchecked? Or did the economy segregate into two domains, one controlled by foreign capital and the other by Hungarians?

Figure 6 shows the history of a single firm's ownership patterns. It begins as an isolate under State ownership. After three years, it becomes the periphery of a small star. In 1992 the topography of the firm's local network is a cohesive cluster, and after three years, these network ties are transformed into a strongly cohesive group. Eventually, it shrinks back into a dyad.

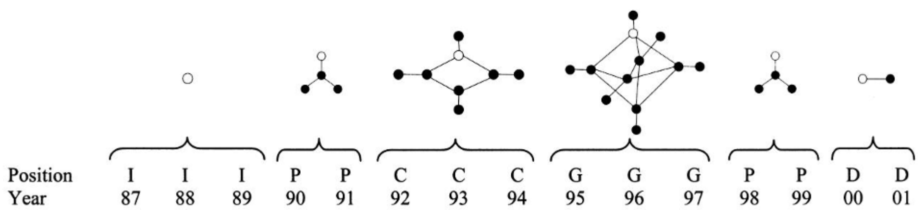


Figure 6. A Hungarian firm's ownership pattern can be represented as a network which changes structure though time.
Redrawn from: Stark, D., & Vedres, B. (2006). Social times of network spaces: Network sequences and foreign investment in Hungary. American Journal of Sociology, 111(5), 1367-1411.

Using this approach 1,696 network histories were constructed. Analytical techniques originally developed for understanding genetic sequences were adapted to search for the most common trajectories of change. Ultimately, 12 pathways accounted for 59% of the variance in inter-sequence distances. Thus, a small number of pathways explain a great deal of the change that occurred in the ownership of most of the largest firms in Hungary during this period. The authors make the further point that these shifting patterns of relationships were not merely the residue of external events, but may have had adaptive significance in themselves: "Hungary's transformation from state socialism to an emerging market economy with sizable foreign investment did not occur despite its inter-organizational property networks but, in part, because of and through these networks" [ibid., 1399]. The authors conclude that developing economies do not necessarily face a forced choice between networks of global reach and those of local embeddedness.

Historians are not always able to acquire the necessary volume of information to quantify network relationships. Nevertheless, Elisabeth Wood has shown that social processes that occur during civil wars can leverage pre-existing social networks and may fundamentally change them after the war ends [Wood, 2008]. Political mobilization usually occurs before armed conflict begins, as state and non-state actors seek to gain resources and power. Local elites access kinship and clientelist networks to protect their interests from insurgents or opposing local elites. Sometimes an insurgent group, such as the Revolutionary United Front in Sierra Leone's Civil War, forcibly recruited members without drawing on local social networks, which led to a heterogeneous and weakly organized internal structure. State militaries may purposely recruit from a wide range of social groups, ignoring local networks in order to build national unity within the armed force. Civilians who remain neutral during conflict often become isolated from kinship and labor networks. Thus, pre-existing social networks shape the nature and outcome of the conflict. Further, as the conflicts continue the networks themselves may be reshaped. Thus Balcells Ventura observed how patterns of lethal violence that occurred during the Spanish Civil war affected voting patterns forty years later [Ventura, 2007].

Any historical analysis needs to address the question of what gets the ball rolling. In these three studies, lots of balls receive a nudge here and a twitch there, until some of them find themselves in unfamiliar territory. A new "structure" or pattern of relationships emerges, in the shape of a Medici bank, or evolving clusters of Hungarian firms. Social ties at the level of the individual give rise to patterns or structures, like banks or patrilineages, which in turn influence social life at the level of the individual. This view of social process is not entirely new; Anthony Giddens' influential theory of "structuration" gives a similar account of agency:

> Human social activities, like some self-reproducing items in nature, are recursive. That is to say, they are not brought into being by social actors but continually recreated by them via the very means whereby they express themselves as actors. In and through their activities agents reproduce the conditions that make these activities possible. [Giddens, 1984]

Giddens' ideas about "structuration" appeared in the early 1980s, a time when many social theorists were troubled by absence of agency in Lévi-Strauss' structuralism. By depicting social actors as the creators of structure as well as its instruments, Giddens sought both to restore agency and to emphasize temporal processes of change. But his theory was pitched at a very general level, a description of the human condition rather than a methodology for investigating specific processes of change. In contrast, the network studies described above offer insights into the topology of the possible for particular historical settings. Evolving networks are inevitably path dependent; future states are constrained by the past. But as these studies show, agency in the more powerful sense of the ability to shape genuine innovations, like Medici banks, can arise from the ordinary form of agency exhibited by people going about their daily business of commerce, marriage and politics.

2.1 Complexity and the Natuurwissenschaften: evolutionary dynamics

The examples we have just considered address two fundamental questions in anthropology: the role of symbols in social life, and the relationship between agency and social structure. Here we take up another venerable question on which contemporary anthropology is sharply divided: are Darwinian processes at work in the social or cultural realms? As before, we begin by defining the issues from the perspective of complexity theory; sketch the outlines of a relevant mathematical model, and briefly consider some empirical examples.

Darwinian models in anthropology generally assume that behaviors, like genes, are constantly being scrutinized by selection. Thus "analyses typically take the form of the following question: in what environmental circumstances are the costs and benefits of behavior X such that selection would favor its evolution?" [Smith and Winterhalder, 1992, 23]. To investigate selection at the level of individuals, we compare their fecundity. The "fittest" individuals are those that produce the most offspring. Because there is always variation in fecundity, at this level it would appear that selection is always at work. But one of the key insights in the mathematical analysis of complex systems is the concept of emergent properties. In this context, one can ask, what would be the population-level consequences of selection at the level of the individual? Figure 7 illustrates three alternatives. In the first case (Fig 7a), selection is not present so the make-up of the population as time goes forward depends on drift (about which more will be said below). If selection occurs (Fig. 7b) the effect at the population level is to reduce the diversity of the population. But the evolutionary consequences of selection depend on how long it persists. Fluctuating dominance, in which some individuals in each generation attain higher reproductive fitness, but do not pass this on to their descendants, produces Red Queen dynamics (Figure 7c). The Red Queen forestalls evolution by preventing any individual or group from gaining a lasting advantage. As she explained to Alice, sometimes "it takes all the running you can do to keep in the same place."

What about the neutral case? Until the 1960s, biologists assumed that almost

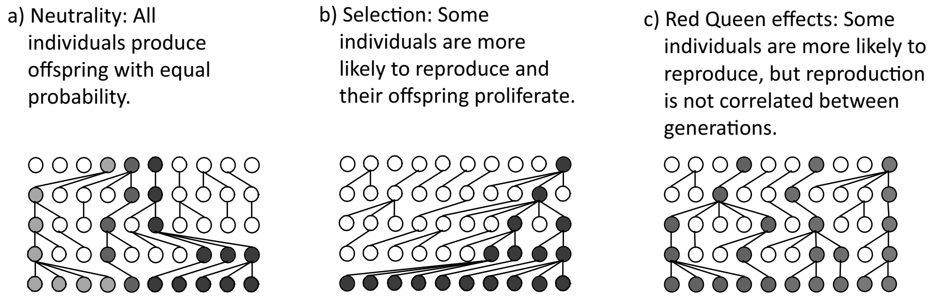


Figure 7. Selection, neutrality and Red Queen dynamics. Here each color represents a single entity with the capacity for reproduction (such as a person, a behavior or a strategy). (a) A population at neutral equilibrium; (b) another undergoing directional selection due to differential fecundity with a high inheritance coefficient; (c) Red Queen dynamics, in which higher fecundity is not strongly inherited.

all mutations are under selection. But in 1963 geneticist Motoo Kimura asked what genetic variation would look like if selection were not present. Even in the absence of selection, Kimura reasoned, evolutionary change will occur as a result of chance, and this can be analyzed with tools from probability theory. As Kimura wrote in 1983, "It is easy to invent a selectionist explanation for almost any specific observation; proving it is another story. Such facile explanatory excesses can be avoided by being more quantitative" [Kimura, 1983].

The processes that lead to neutral equilibrium can be explained with the statistician's favorite example, a bag of colored marbles. To model the effects of drift, the experimenter reaches into the bag and grabs two marbles. One is randomly tossed aside and the other is magically duplicated; the latter (identical) pair of marbles is put back into the bag. Starting with a bag of ten marbles, each with a different color, all the marbles in the bag will be the same color after a few replacements. This process will take much longer with bags of 100 or 1000 marbles. Thus drift reduces the number of colors in the bag. Mimicking the effects of mutation can counteract this process: from time to time a marble with a new color is added to the bag as a replacement for a discarded marble. Neutral equilibrium is reached when the introduction of new colored marbles by mutation (or migration) matches the rate at which existing colors are removed by drift.

A charming example that demonstrates the application of the neutral theory to culture was recently published by Hahn and Bentley, who investigated the changing frequency of baby names in the United States. One can easily imagine selectionist explanations for the prevalence of names; for example, in each generation parents might preferentially choose the names of culturally dominant or prestigious individuals for their children. The alternative, neutral hypothesis, predicts a distribution of names governed solely by chance. While any number of selectionist models could be proposed for the frequency distribution of baby names, there is only one neutral distribution for any given dataset. This neutral distribution depends solely on the total population size and the rate at which new names appear.

In 2002, the Social Security Administration published the thousand most common baby names in each decade of the twentieth century, based on a sample of 5% of all social security cards issued to Americans. Most parents chose a pre-existing name for their infant, but occasionally a new name was introduced. Hahn and Bentley found that the distribution of names from one decade to the next fits a power-law distribution with an r^2 value above 0.97. A very few names were extremely popular, while others persisted at decreasing frequencies. However, the prevalence of certain names also changed as the century progressed.

To explain this stable distribution of name frequencies despite changes in the popularity of particular names, the researchers created a simulation based on Kimura's neutral theory, and compared these results with the observed data. In a neutral model, random drift causes changes in the frequencies of names as they are repeatedly sampled; some are lost, other arise de novo, while still others drift up or down in frequency. To simulate this process, N new babies appear at each time-step and are named by copying the name of a randomly chosen baby from the previous time-step. A small fraction, m, of the N babies receive a name that was not present earlier. The neutral theory predicts that, at equilibrium, the number of variants (baby names) with frequency x at a single moment is given by θx-1(1-x)θ-1 where $\theta = 4\mathrm{Ne}\mu$ [Slatkin, 1996]. A regression between the logs of the average values of the model and the data yielded $r^2 = 0.993$ for boys' names and $r^2 = 0.994$ for girls' names. Power law distributions can result from many causes. In this case, the neutral theory predicts not only the power law distribution, but also its slope, with near-perfect accuracy. Chance, not selection, determines the frequencies at which baby names occur in the US population.

In genetics, the neutral theory was hotly debated for decades. As Kimura observed in his 1968 paper, the prevalent view in the 1960s held that almost all mutations are under selection, and this opinion was slow to change. But as Stephen J. Gould wrote in 1989, "These equations give us for the first time a baseline criterion for assessing any kind of genetic change. If neutralism holds, then actual outcomes will fit the equations. If selection predominates, then results will depart from predictions" [Gould, 1989]. This led to a fundamental reformulation of how selection was viewed in molecular biology: geneticists now infer selection only when it can be shown that the assumption of neutrality has been violated. As E.G. Leigh observed in a recent retrospective about the neutral theory, "no population geneticist, not even Kimura, sought to deny the importance of adaptive evolution. Instead, all major workers "were interested, at least to some degree, in how neutral processes affected adaptive evolution" [Leigh, 2007, p. 2076]. In ecology, as Leigh further noted [ibid, p. 2087], everyone, even the advocates of the neutral theory, recognize that neutral theory is wrong when taken to extremes: adaptive processes clearly do matter. In genetics, the question of precisely which regions of the genome are under selection is being revisited using neutral theory [Hey, 1999].

But in anthropology, Darwinian models of cultural evolution continue to focus on selective processes occurring at the level of the individual, rather than the population-level consequences [Richerson and Boyd, 2006]. Most research in human behavioral ecology is explicitly pan-selectionist, asking "what are the fitness effects of different strategies in particular environments?" [Clarke and Low, 2001] rather than "are the behaviors we observe actually under selection?" In "The Spandrels of San Marco and the Panglossian paradigm", their well-known critique of pan-selectionism in biology, Gould and Lewontin commented on the need for an explicit test for 'adaptationist' explanations:

> We would not object so strenuously to the adaptationist programme if its invocation, in any particular case, could lead in principle to its rejection for want of evidence. We might still view it as restrictive and object to its status as an argument of first choice. But if it could be dismissed after failing some explicit test, then alternatives would get their chance. [Gould and Lewontin, 1979]

The neutral theory provides such a test, for cultural evolution as well as genetics and ecology. It shifts the analytical focus from selection at the level of the individual, to the population-level consequences of both selection and neutral processes over the course of multiple generations. In place of the pan-selectionist assumptions of evolutionary game theory and behavioral ecology, it provides a mathematically explicit null model. For example, a recent study investigated the magnitude of selection effects stemming from reproductive competition among men in 41 Indonesian villages [Lansing *et al.*, 2008]. Many studies have argued that reproductive skew biased toward dominant or high-ranking men is very common in human communities: "In more than one hundred well studied societies, clear formal reproductive rewards for men are associated with status: high-ranking men have the right to more wives" [Clarke and Low, 2001]. Demographic statistics collected over short time scales support these claims [Winterhalder and Smith, 2000]. Although variation in male fitness is known to occur, an important unanswered question is whether such differences are heritable and persist long enough to have evolutionary consequences at the population level.

In this study, genetic data showed that dominance effects generally do not persist over multiple generations. The lack of evidence of reproductive skew in these communities means that heritable traits or behaviors that are passed paternally, be they genetic or cultural, are unlikely to be under strong selection. The discovery that neutral processes can explain most haplotype frequency distributions in these communities parallels earlier results from the development of neutral theory in genetics and ecology. As Kimura observed in his original article, the prevalent opinion in the 1960s held that almost all mutations are under selection [Kimura, 1968]. This opinion was slow to change. More recently, ecologists similarly have suggested that a neutral model, in which species in the same trophic level are functionally equivalent or neutral with respect to each other, might adequately explain species-abundance distributions in ecological communities [Hubbell, 2001]. In anthropology, the recent availability of appropriately sampled community-level polymorphism data now enables us to distinguish both genetic and cultural selection from neutral demographic processes with surprising precision. In these Indonesian communities, male dominance seldom translates into increased fertility among descendents over evolutionary timescales.

In both genetics and ecology, the neutral theory played an important role in introducing a systems-level dynamical perspective to evolutionary theory. One advantage for anthropology as a relative latecomer to this perspective is that anthropologists are in a position to benefit from several decades of theoretical work, including a substantial body of elegant mathematics. A particularly salient lesson may be the identical outcome of the debates that occupied both genetics and ecology for years, both of which pitted pan-selection against pan-neutrality. In both fields, this debate was largely resolved by adopting a view of neutrality as a null model, rather than as a strict alternative to Darwinism [Alonso *et al.*, 2006; Hu *et al.*, 2006].

2.2 Coupled human and natural systems

Our final topic is the interaction of societies with local environments. A key theoretical issue that arises from the study of such interactions, from the perspective of complex systems, is how patterns may emerge from multiple processes occurring at different scales and how spatio-temporally variegated outcomes may optimally resolve conflicts among conflicting goals irreconcilable in logic. In one of the most cited papers on this topic, Simon Levin observed that patterns are often generated by the collective behavior of smaller scale units, which "operate at different scales than those on which the patterns are observed" [Levin, 1992]. Ecological journals are filled with examples of such processes, with a growing emphasis on global-scale phenomena such as climate change. But these ideas have been slow to spread to the social sciences. Karl Marx famously dismissed the peasants as a "sack of potatoes", and for most social scientists, it is probably still true that one piece of countryside looks much like the next. Even anthropologists are seldom inclined to search for the kinds of pattern-and-scale interactions that Levin describes.

The examples to be considered here have to do with the management of rice paddies on the island of Bali, and dipterocarp forests on the island of Borneo. We suggest that the failure of planners to appreciate the role of emergent patterns in both of these cases led to disastrous errors. It's particularly interesting to note that the dynamical processes that were vital to both cases are quite similar, even though the underlying ecologies are very different. To highlight this similarity, we begin with a model that captures the key insight. The model is "Daisyworld", a thought experiment created by chemist James Lovelock [Lovelock, 1992]. Daisyworld has several useful features: the biology is as simple as Lovelock could make it; the model shows precisely how small-scale local adaptations can produce an emergent global structure; and it also shows why such global structures can easily fade from view, becoming noticeable only when the system as a whole has been pushed near its limits.

Daisyworld is an imaginary planet orbiting a star like the Sun and at the same orbital distance as the Earth. The surface of Daisyworld is fertile earth sown uniformly with daisy seeds. The daisies vary in color, and daisies of similar color grow together in patches. As sunshine falls on Daisyworld, the model tracks changes in the growth rate of each variety of daisy, and changes in the amount of the planet's surface covered by different-colored daisies. The simplest version of this model contains only two varieties of daisies, white and black.

Black daisies absorb more heat than bare earth, while whites reflect sunshine. Clumps of same-colored daisies create a local microclimate for themselves, slightly warmer (if they are black) or cooler (if white) than the mean temperature of the planet. Both black and white daisies grow fastest and at the same rate when their local effective temperature (the temperature within their microclimate) is 22.5 $^\circ$ C, and they respond identically, with a decline in growth rate, as the temperature deviates from this ideal. Consequently, at given average planetary temperatures, black and white daisies experience different microclimates and therefore different growth rates.

If the daisies cover a sufficiently large area of the surface of Daisyworld, their color affects not only their own microclimate but also the albedo or reflectance of the planet as a whole. Like our own sun, the luminosity of Daisyworld's star is assumed to have gradually increased. A simulation of life on Daisyworld begins in the past with a cooler sun. This enables the black daisies to spread until they warm the planet. Later on, as the sun grows hotter, the white daisies grow faster than black ones, cooling the planet. So over the history of Daisyworld, the warming sun gradually changes the proportion of white and black daisies, creating the global phenomenon of temperature regulation: the planet's temperature is held near the optimum for the daisies, as shown in Fig. 8.

Imagine that a team of astronauts and planners is sent to investigate Daisyworld. They would have plenty of time to study the only living things on the planet, and they would almost certainly conclude that the daisies had evolved to grow best at the normal temperature of the planet, 22.5 $^\circ$ C. But this conclusion would invert the actual state of affairs. The daisies did not adapt to the temperature of the planet; instead they adapted the planet to suit themselves [Saunders, 1994]. A Daisyworld without daisies would track the increase in the sun's luminance (line 2), rather than stabilizing near the ideal temperature for daisies (line 1). Only when the sun's luminosity becomes too hot for the daisies to control ($\sim$1.4) will the daisy's former role in temperature stabilization become apparent.

Lacking this understanding, planners hoping to exploit Daisyworld's economic potential as an interstellar flower supplier would fail to appreciate the possible consequences of different harvesting techniques. While selective flower harvests would cause small, probably unnoticeable tremors in planetary temperature, clear-

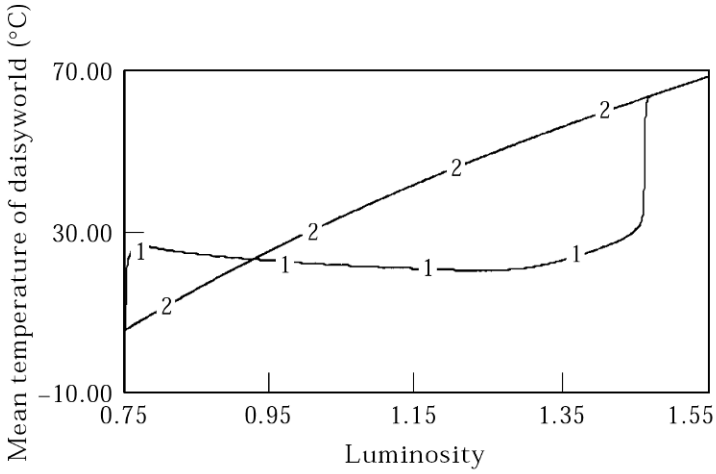


Figure 8. Results of a simulation of temperature regulation on Daisyworld. As the sun ages and its luminosity increases from 0.75 to 1.5 times the present value (1.0), the temperature of a bare planet would steadily rise (line 2). In contrast, with daisies present, the temperature stabilizes close to 22.5 C (line 1).

cutting large contiguous patches of daisies would create momentary changes in the planet's albedo that could quickly become permanent, causing temperature regulation to fail and daisy populations to crash.

2.2.1 Emergence in Balinese water temple networks

The Daisyworld model offers insight into the emergence of a complex adaptive system based on the role of water temples in managing wet-rice cultivation on the Indonesian island of Bali [Lansing *et al.*, 1998]. In Bali, rice is grown in paddy fields on steep hillsides fed by elaborate irrigation systems dependent on seasonal rivers and ground water flows, dominated by an elevated volcanic crater lake. Gravity-fed irrigation works route the water to the various fields. The rugged topography and interconnections among the fields create a very interdependent system that can, at times, be quite fragile and subject to major disruptions. Decisions about irrigation are made by groups of farmers who share a common water source, in a Balinese institution called a *subak* [Lansing *et al.*, 2009].

Water performs a variety of complex biological processes in the rice paddy ecosystem. Careful control of the flow of water into the fields creates pulses in several important biochemical cycles necessary for growing rice. Water cycles have a direct influence on soil PH, temperature, nutrient circulation, aerobic conditions, microorganism growth, weed suppression, etc. In general, irrigation demands are highest at the start of a new planting cycle, since the dry fields must first be

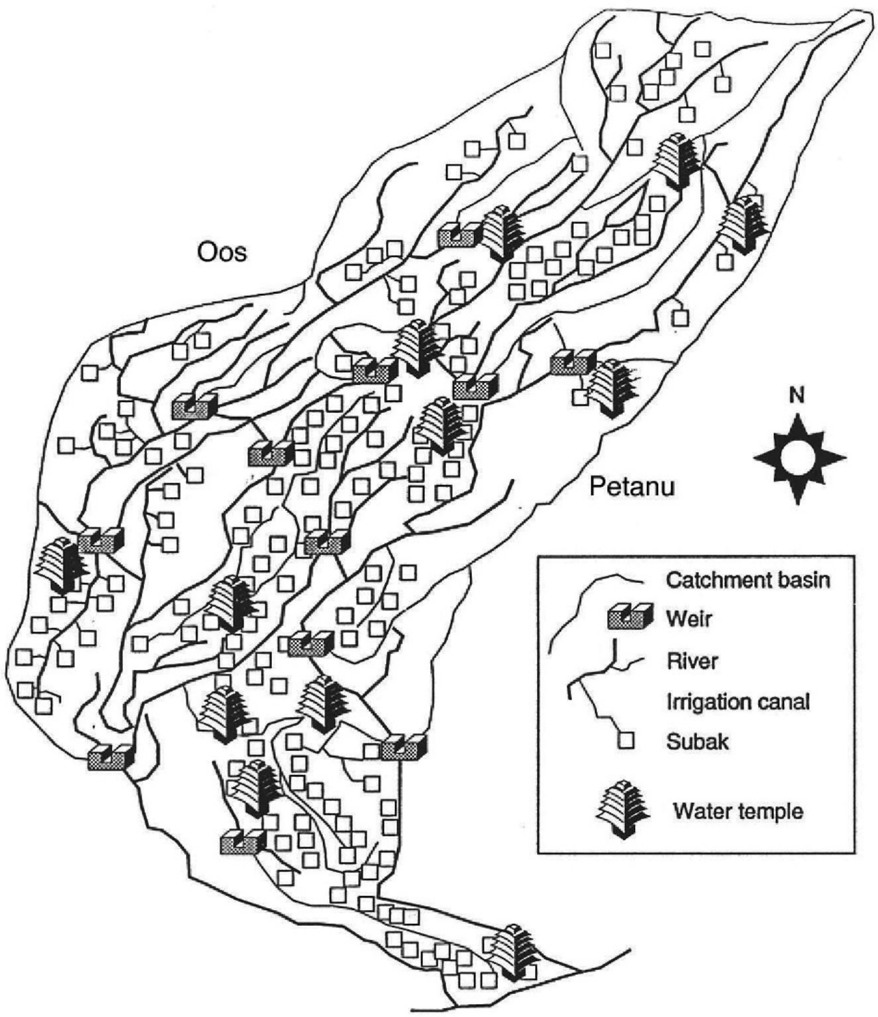


Figure 9. The Bali model.

saturated with water.

The flooding and draining of blocks of terraces also has important effects on pests (including insects, rodents, and bacterial and viral diseases). If farmers with adjacent fields can synchronize their cropping patterns to create a uniform fallow period over a sufficiently large area, rice pests are temporarily deprived of their habitat and their populations can be sharply reduced. Field data indicate that synchronized harvests result in pest losses of around 1% compared to losses upwards of 50% during continual cropping. How large an area must be fallow, and for how long, depends on specific pest characteristics. If too many subaks follow an identical cropping pattern in an effort to control pests, then peak water demands will coincide. The existing watershed seldom provides sufficient water to meet the full needs of all subaks in such a case.

Paralleling the physical system of terraces and irrigation works, the Balinese have also constructed intricate networks of shrines and temples dedicated to agricultural deities and the Goddess of the Lake. At meetings held in the temples, subaks decide on their irrigation schedules and what crops they will plant. These meetings provide a way for the subaks to coordinate cropping patterns with their neighbors, using the framework of ties that exists between the water temples. But is this system of bottom-up control effective? The key question is not whether flooding and fallowing can control pests, but rather whether the entire collection of temples in a watershed can strike an optimal balance between water sharing and pest control.

Using empirical data on the location, size and field conditions of 172 subaks in the watershed of the Oos and Petanu rivers in southern Bali in 1987-8, Lansing and Kremer modeled changes in the flow of irrigation water and the growth of rice and pests as subaks decided whether to cooperate with their neighbors. The "Bali Model" shown in Figure 9 simulates the flow of water from the headwaters of the two rivers to the sea, at monthly intervals. The amount of water available for any given subak depends on seasonal patterns of rainfall and ground water flow, and the amount of water diverted by upstream subaks for their own needs. As a new year begins, each of the 172 subaks is given a planting schedule which determines which crops it will grow, and when they will be planted. As the months go by, water flows, crops grow, and pests migrate across the landscape. When a subak harvests its crop, the model tabulates losses due to water shortages or pests. At the end of the year, aggregate harvest yields are calculated for the subaks. Subsequently, each subak checks to see whether any of its closest neighbors got higher yields. If so, the target subak copies the cropping schedule of its (best) neighbor. If none of the neighbors got better yields, the target subak retains its existing schedule. When all the subaks have made their decisions, the model cycles through another year. These simulations begin with a random distribution of cropping patterns (a typical example is shown in Fig. 10). After a year the subaks in the model begin to aggregate into patches following identical cropping patterns, which helps to reduce pest losses. As time goes on these patches grow until they overshoot, causing water stress and patch sizes to become smaller. Yields fluctuate but gradually rise. The

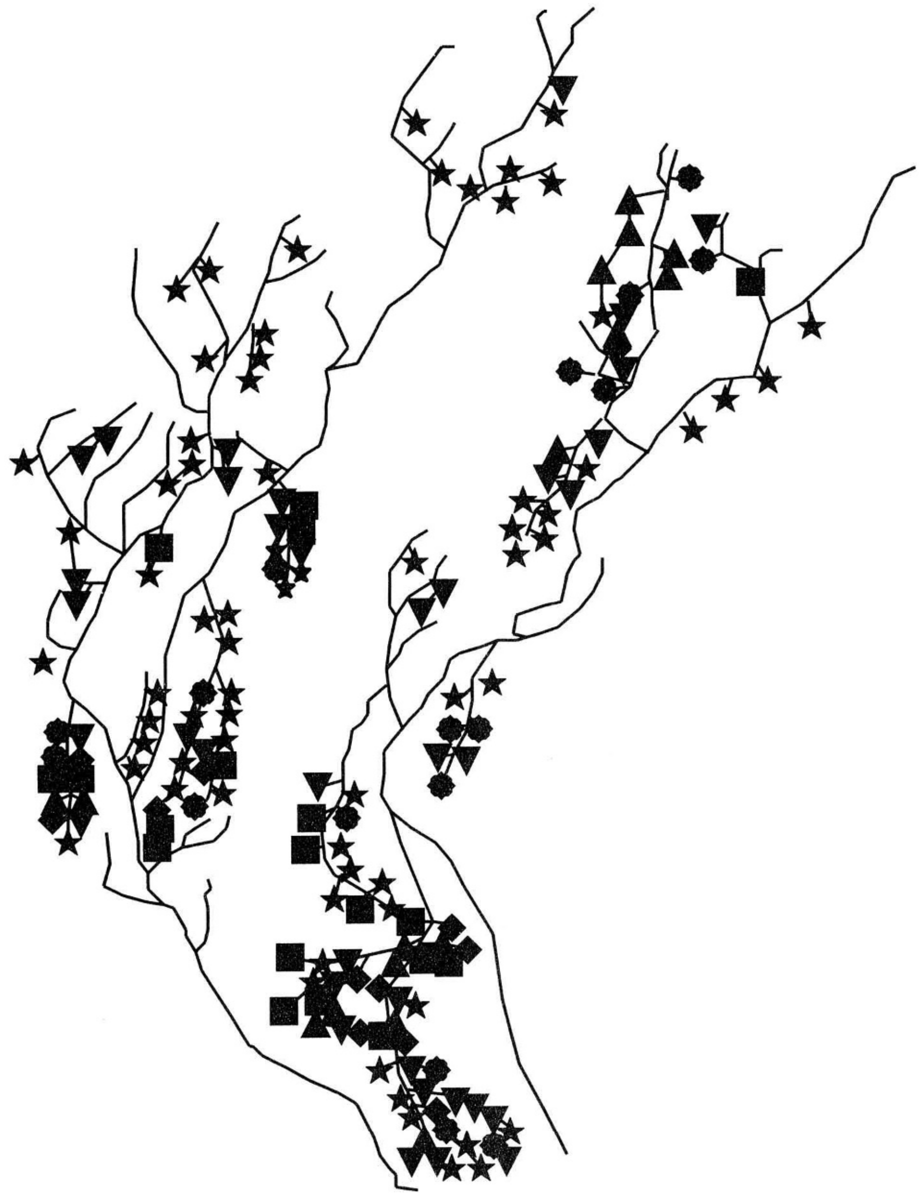


Figure 10. Initial conditions for a simulation model of irrigation flows and rice and pest growth for 172 subaks. Differences in cropping patterns are indicated by different symbols (subaks with the same symbols have identical cropping patterns).

program continues until most subaks have discovered an optimal cropping pattern, meaning that they cannot do better by imitating one of their neighbors.

Experiments with this model indicate that the entire collection of subaks quickly settles down into a stable pattern of synchronized cropping schedules that optimizes both pest control and water sharing. The close relationship between this pattern as calculated in the model (Fig. 11), and the actual pattern of synchronized planting units (Fig. 12) is apparent. In the model, as patterns of coordination resembling the water temple networks emerge, both the mean harvest yield and the highest yield increase, while variance in yield across subaks declines (Fig. 13). In other words, after just a few years of local experimentation, yields rise for everyone while variation in yields declines. Subsequent simulations showed that if the environment is perturbed, either by decreasing rainfall or by increasing the virulence of pests, a few subaks change their cropping patterns, but within a few years a new equilibrium is achieved [Lansing, 2006, 67-88].

These results helped explain the decline in rice production in Bali that began in the 1970s, after farmers were ordered to stop planting native Balinese rice according to the temple calendars, and instead plant hybrid "Green Revolution" rice as fast as possible [Lansing, 1991]. While the "Green Revolution" rice could grow faster and produce more grain than the native plants, these potential gains were offset by the disruptions caused by abandoning the temple calendars. This effect can be demonstrated in the simulation model by running it in reverse: beginning with a patchwork of synchronized multi-subak groups, and breaking up synchrony by encouraging the subaks to compete with one another.

2.2.2 Emergence in the dipterocarp forests of Borneo

Are the water temples of Bali a unique case? This question came up soon after the Bali Model analyses were published. Perhaps the Maya or the ancient Khmer had invented something like the Balinese water temples? But so far, the most interesting comparison comes from a site much closer to Bali. And it has nothing to do with irrigation, temples or rice.

In 1967, the year the Green Revolution in rice began in most of Indonesia, another government program opened the forests of the Outer Islands to logging for export. Like the Green Revolution, this policy inadvertently set in motion an experimental test of the resilience of a tropical ecosystem. And like the Green Revolution, it produced immediate, spectacular results. By the early 1970s, logging exports were generating annual export earnings of over US$1.5 billion, eventually rising to as much as $6 billion.[3] As the Ministry of Forestry proclaimed in 1990,

> The logging industry is a champion of sorts. It opens up inaccessible areas to development; it employs people, it evolves whole communities; it supports related industries. . . It creates the necessary conditions for

[3] "Over the past two decades, the volume of dipterocarp timber exports (in cubic meters) from Borneo (Kalimantan, Sarawak and Sabah) exceeded all tropical wood exports from tropical Africa and Latin America combined" [Curran *et al.*, 2004].

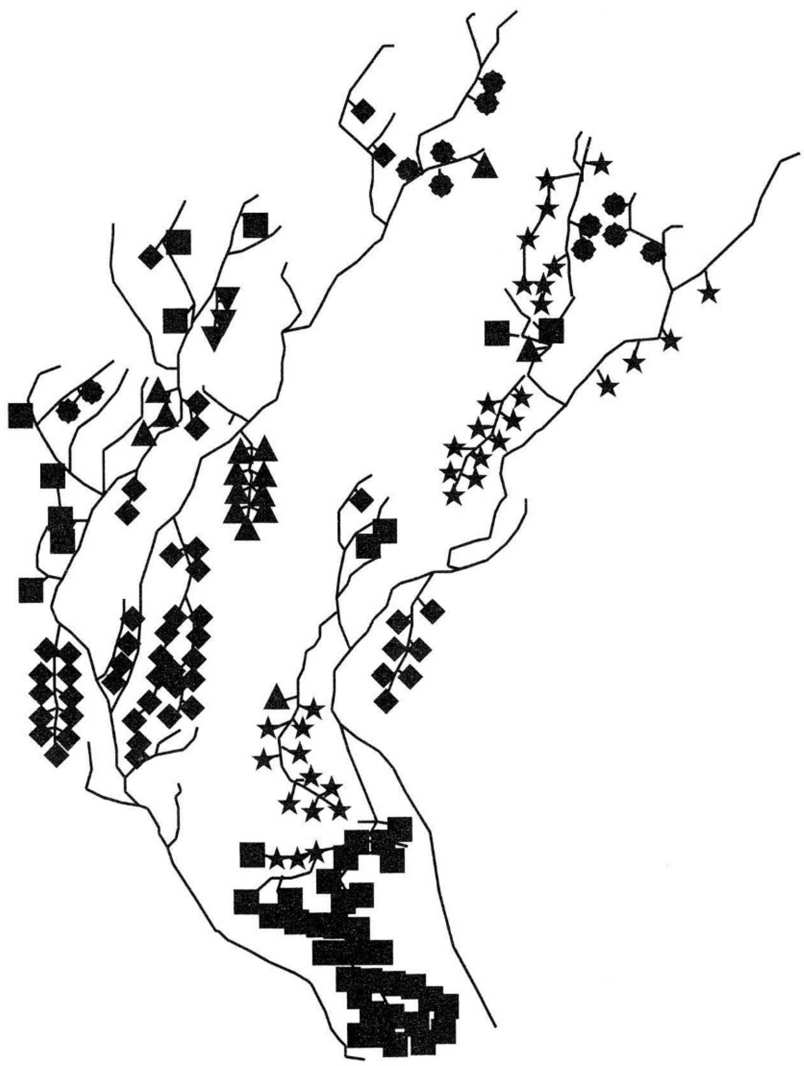


Figure 11. Model cropping patterns after 11 years

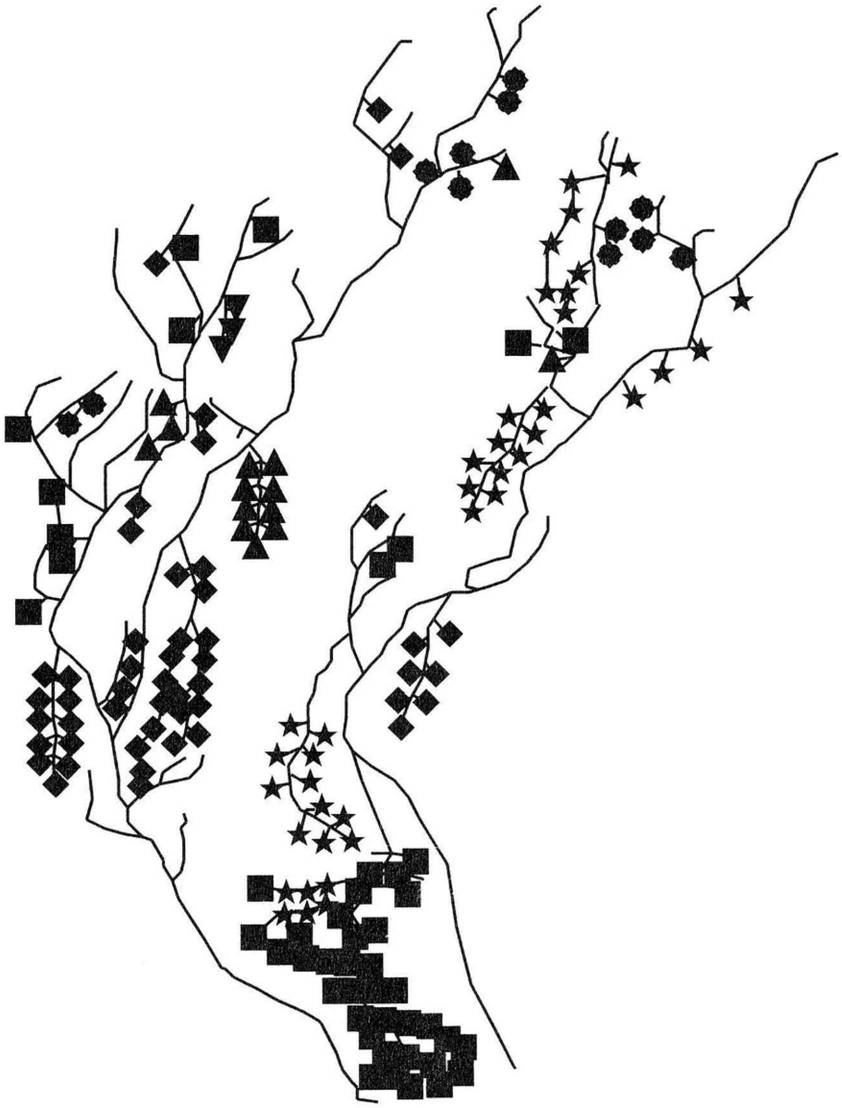


Figure 12. Actual observed cropping patterns (1987)

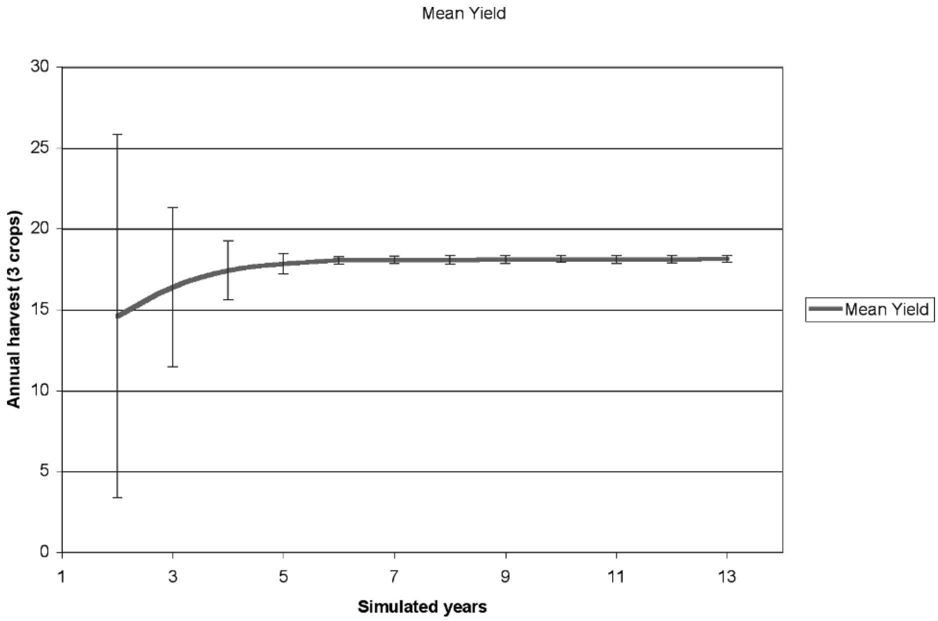


Figure 13. Reduction in variance of harvest yields as cooperation spreads in the simulation model of the Oos and Petanu watersheds. Rainfall, rice and pest parameters based on data collected in 1988.

> social and economic development. Without forest concessions most of the Outer Islands would still be underdeveloped. [Government of Indonesia, 1990]

By the 1980s, in response to indications of forest degradation from logging, the Ministry began to promote industrial tree plantations for the pulp and paper industry, supported by interest-free loans from the "Reforestation Fund" and international investment. Along with reforestation, the government also encouraged the creation of palm oil plantations on logged land. Sawmills, logging roads and palm plantations proliferated in the 1990s, and exports of pulp, paper and palm oil boomed. In 2002, export taxes on raw logs were eliminated and Indonesian firms were permitted to sell logs to anyone. Plans for biodiversity conservation were based on selective logging and reforestation, and the creation of national parks [Gellert, 2005].

The dominant canopy tree family in Borneo and Sumatra is the *dipterocarpaceae*, which consists of $\sim$267 tree species that make up over 85% of Indonesia's tree exports. The sustainability of the timber industry thus depends on the regenerative capacity of dipterocarp forests. In 1999, ecologist Lisa Curran and her colleagues reported the results of a comprehensive 14 year investigation of the ability of the dipterocarps to reproduce. Regrowth depends on the survival of sufficient quantities of seedlings. Many forest animals and birds are seed predators, so the trees are engaged in a continuous race to produce more seeds than the predators can consume. Curran found that long ago, the trees evolved essentially the same solution to the problem of controlling predation that was later discovered by the Balinese farmers: reproductive synchrony. Dipterocarp forests produce nearly all of their seeds and fruits within a very small window in time, in a phenomenon known to ecologists as "mast fruiting." For seed predators, this means that large quantities of dipterocarp fruits and seeds only become available in short irregular bursts that occur every 3-6 years, triggered by the El Nino Southern Oscillation (ENSO). ENSO is a global climatic cycle that causes an extreme reduction in rainfall in Borneo from June to September. The ENSO dry spell is used by the trees as a signal to initiate flowering and reproduction. Seed predators respond by synchronizing their own reproductive cycles to ENSO years, and by moving across the landscape, far from their usual ranges, to feed on dipterocarp seeds.

Over the past three decades, the harvesting of timber caused widespread fragmentation of what had formerly been a vast contiguous expanse of dipterocarp forest in Borneo, disrupting regional reproductive synchrony. Once synchrony was lost, small-scale local masts could not produce enough seedlings to escape being eaten by predators. Curran found that "Seed escape, and thus regeneration, only occurred in major mast events when all dipterocarp species across large areas participated" [Curran and Leighton, 2000]. This closely parallels the Balinese case. In the rice terraces of Bali, the "Green Revolution" caused disruption of the synchronized planting schedules formerly organized by water temple networks. This led to crop losses, as migrating pests moved across the landscape consuming one harvest after the next. Similarly, in Borneo the mast synchrony of canopy trees in ENSO years was triggered by signals transmitted through the root system. When the forests became fragmented, it was no longer possible to overwhelm predators with a vast synchronized mast.

We now know that in both Bali and Borneo, large-scale reproductive synchrony emerged as a solution to the problem of controlling seed predators. But in both cases, this cyclical pattern was invisible to planners. In Bali, the farmers were able to respond in the nick of time and restore control to the temple networks. But the trees were not so fortunate. The latest research by Curran and her colleagues shows that the lowland forests of Indonesian Borneo have lost the capacity to regenerate, probably beyond hope of recovery. As a consequence, ENSO — formerly a great forest regenerator — has become a destructive regional phenomenon, triggering droughts and wildfires with increasing intensity. By the 1990s, much of Indonesian Borneo had been deforested, leaving logging debris in place of canopy trees. When ENSO arrived in 1998, forest fires raged across the island and four hundred million metric tons of carbon were released into the atmosphere. Even peat swamps caught fire, adding another two hundred million tons of carbon. (For comparison, the Kyoto target for reduction in carbon emission for the whole earth was five hundred million tons).[4]

[4]Curran estimated that in 2005, less than 35% of the lowland forests (<500 m a.s.l.) were

Thus in both Borneo and Bali, synchronized growing cycles emerged as a solution to the problem of controlling predator populations in the winterless tropics, imposing a clockwork pattern on the life cycles of many species. At least in this respect, the water temple networks of Bali are not unique. Might other, similar systems exist elsewhere? If so, would they always be driven by the need for predator control? How much of the functional structure of the water temple networks is directly tied to the ecology of Bali or the biology of pests? These questions remain open. The temple networks came into view partly as a result of the Green Revolution, which exposed their ecological role, and partly through our expanding familiarity with the properties of complex adaptive systems like Daisyworld. Indeed, the enduring message of this case may be how easy it was to miss the significance of the temple networks, just as planners failed to appreciate the functional significance of the "forest clock" in Borneo.

3 CONCLUSION: COMPLEXITY AND ANTHROPOLOGY

In one of the foundational articles that launched complexity studies, physicist P.W. Anderson quoted Marx's observation that quantitative differences become qualitative differences. "More is different", Anderson observes, because at each new level of complexity entirely new properties appear [Anderson, 1972]. By tracing patterns of interaction among the elements of a system, one can sometimes discover emergent properties at a higher level. This phenomenon is common to all of the examples we have considered here, from semiotics to agency, innovation, cultural evolution and human-environmental interactions. But until recently our mathematical tools were not well suited to investigate emergence, or other properties of out-of-equilibrium dynamical systems. As recently as 1990, Karl Popper argued that social scientists who wish to take advantage of mathematics have the choice of only two approaches [Popper, 1990, 18-19]. The first is essentially Newtonian and is best represented by general equilibrium theories, for example in economics. Such theories take the form of systems of differential equations describing the behavior of simple homogeneous social actors. Change occurs as a result of perturbations and merely leads from one equilibrium state to another. The second type of theory is statistical. If one cannot write the equations to define a dynamical system, it may yet be possible to observe statistical regularities in social phenomena. Both approaches have obvious weaknesses: the assumption of equilibrium is forced by the mathematics, not by the observation of social behavior, and sifting for patterns with descriptive statistics is at best an indirect method for discovering causal or developmental relationships.

The landscape looks very different today. Out-of-equilibrium systems are the main topic of complexity research, and there has been a proliferation of new analytical methods to investigate them. Of particular interest to anthropologists are the models of "artificial societies", of which the "Bali Model" described above is

still standing, most of them already degraded (*personal communication).*

an example. Artificial societies are computational models comprised of populations of social agents that flourish in artificial environments. While the behavior of the agents is governed by explicit rules, their behavior can change in response to changes in the environment or as a consequence of learning and memory. In this way the behavior of agents can become heterogeneous in ways that cannot occur in conventional equilibrium models. Further, agents can retain this local variability, which becomes dynamically relevant to future possibilities. Repeated simulations enable the investigator to study how such model systems evolve over time, and to investigate their global properties. These features have helped make artificial societies the most common route by which anthropologists and archaeologists have become interested in the study of complex systems [Lansing, 2002].

But today, forward-in-time simulations of artificial societies are just one of many approaches to complexity research in anthropology. Thus in molecular anthropology, a maximum likelihood approach is used to assess backward-in-time models of stochastic processes [Majoram and Tavaré, 2006]. These methods are now being adopted in historical linguistics, triggering something of a methodological revolution as coalescent models are adapted to linguistic data [Pagel, Mace, 2004]. Maximum likelihood methods are also being introduced to network analysis, supplementing descriptive statistics with forward-in-time network simulations [Robins and Morris, 2007]. Along with these probabilistic methods, anthropologists have also begun to study formal models of nonlinear dynamical systems and complex adaptive systems. Robustness and resilience are among the themes that have recently emerged. The notion of robustness captures our intuitive sense of one of the key determinants of long-term success or failure. This topic comes to us from physics and engineering, and focuses on the ability of a system to maintain specified features when subject to assemblages of perturbations, either internal or external [Jen, 2005]. In contrast, the concept of resilience originated in ecology with C.S. Holling, in his studies of the properties of adaptive cycles [Holling, 2001]. Both approaches encourage the investigator to think about tradeoffs between robustness (or resilience) and evolvability. Stepping further back, it is clear that many anthropological questions involve the analysis of systems that possess structures, topologies, networks and adaptive dynamics. The study of these phenomena by anthropologists has just begun.

We conclude with a final thought about the relationship of complexity studies to the historic divide in anthropology between the Geistes- and Natuurwissenschaften. It is interesting to speculate whether the concepts and methods we have considered may have the potential to render this distinction obsolete. In the famous debate between Karl Popper and the "Critical theorists" of the Frankfurt School, Theodor Adorno offered an interesting rationale for his critique of the equilibrium theories and descriptive statistics of positivist science. Adorno argued that social theory must be able to conceive of an alternative to contemporary society: "only through what it is not will it disclose itself as it is..." [Adorno *et al.*, 1976, 296]. This led him to a critique of descriptive statistics as the primary tool for social inquiry. He observed that "A social science that is both atomistic, and ascends through classification from the atoms to generalities, is the Medusan mirror to a society which is both atomized and organized according to abstract classificatory principles. . . ." Adorno's point was that a purely descriptive, statistical analysis of society at a given historical moment is just "scientific mirroring" that " . . . remains a mere duplication." To break the seal of reification on the existing social order, he argued, it is necessary to go beyond descriptive statistics or equilibrium models to explore historical contingency. But the mathematical tools that might assist this kind of investigation did not yet exist. Today, they do.

BIBLIOGRAPHY

[Adorno *et al.*, 1976] T. W. Adorno, R. Dahrendorf, H. Pilot, H. Albert, J. Habermas, and K. Popper. *The positivist dispute in German sociology* (transl. G. Adey, D. Frisby). Heinemann, London, 1976.

[Alonso *et al.*, 2006] D. Alonso, R. Etienne, and A. McKane. The merits of neutral theory. *Trends in Ecol. Evol.* 21, 451–457, 2006.

[Anderson, 1972] P. W. Anderson. More is Different. *Science*, 177(4047), 393-6, 1972.

[Atran *et al.*, 2005] S. Atran, D. L. Medin, and N. O. Ross. The Cultural Mind: Environmental Decision Making and Cultural Modeling Within and Across Populations. *Psychological Review* 112 (4), 744-776, 2005.

[Bodin and Crona, 2008] Ö. Bodin and B. I. Crona. Management of natural resources at the community level: exploring the role of social capital and leadership in a rural fishing community. *World Development*, 36(12), 2763-2779, 2008.

[Borgatti and Foster, 2003] S. Borgatti and P. C. Foster. The network paradigm in organizational research: a review and typology. *Journal of Management*, 29(6), 991-1013, 2003.

[Cancho *et al.*, 2005] R. F. i Cancho, O. Riordan and B. Bollobás. The consequences of Zipf's law for syntax and symbolic reference. *Proc. R. Soc. B.,* 272, 561-565, 2005.

[Clarke and Low, 2001] A. L. Clarke and B. S. Low. Testing evolutionary hypotheses with demographic data. *Popul. Dev. Rev.* 27, 633–660, 2001.

[Curran and Leighton, 2000] L. M. Curran and M. Leighton. Vertebrate response to spatiotemporal variation in seed production of mast-fruiting dipterocarpacae. *Ecological Monographs*, 70(1), 101-128, 2000.

[Curran *et al.*, 2004] L. M. Curran, S. N. Trigg, A. K. McDonald, D. Astiani, Y. M. Hardiono, P. Siregar, *et al.* Lowland forest loss in protected areas of Indonesian Borneo. *Science*, 303(5660), 1000-1003, 2004.

[Derrida, 1976] J. Derrida. *Of Grammatology* (transl. G. C. Spivak). Johns Hopkins Univ. Press, Baltimore, 1976.

[Derrida, 1978] J. Derrida. *Writing and difference*. Routledge & Kegan Paul, London, 1978.

[Fontana, 1993] W. Fontana, The Topology of the Possible. Working Papers of the Santa Fe Institute 03-03-017, 1993.

[Gellert, 2005] P. K. Gellert, The shifting natures of "development": growth, crisis and recovery in Indonesia's forests. *World Development*, 33(8),1345-1364, 2005.

[Giddens, 1984] A. Giddens. *The constitution of society: outline of the theory of structuration.* University of California Press, Berkeley, 1984.

[Gould, 1989] S. J. Gould. Through a lens, darkly: do species change by random molecular shifts or natural selection? *Natural History*, 98 (9), 16-24, 1989.

[Gould and Lewontin, 1979] S. J. Gould and R. C. Lewontin. The spandrels of San Marco and the panglossian paradigm: a critique of the adaptationist programme. *Proceedings Of The Royal Society of London, Series B*, 205(1161), 581-598, 1979.

[Government of Indonesia, 1990] Situation and outlook for the forestry sector in Indonesia. Jakarta: Food and Agriculture Organization and Directorate General of Forest Utilization, Government of Indonesia, 1990.

[Granovetter, 1985] M. S. Granovetter. Economic action and social structure: the problem of embeddness. *American Journal of Sociology*, 91(3), 481-510, 1985.

[Hey, 1999] J. Hey. The neutralist, the fly, and the selectionist. *Trends in Ecology and Evolution*, 14, 35-38, 1999.

[Holling, 2001] C.S. Holling. Understanding the complexity of economic, ecological and social systems. *Ecosystems*, 4, 390-405, 2001.

[Hu *et al.*, 2006] X-S. Hu, F. He, and S. P. Hubbell. Neutral theory in macroecology and population genetics. *Oikos*, 113, 548–556, 2006.

[Hubbell, 2001] S. P. Hubbell. *The unified neutral theory of biodiversity and biogeography.* Princeton Univ. Press, Princeton, N. J. 2001.

[Jen, 2005] E. Jen, ed. *Robust design: a repertoire of biological, ecological and engineering case studies.* Oxford University Press, Oxford, 2005.

[Kimura, 1968] M. Kimura. Evolutionary rate at the molecular level. *Nature*, 217, 624–626, 1968.

[Kimura, 1983] M. Kimura. *The neutral theory of molecular evolution.* Cambridge University Press, Cambridge, 1983.

[Lansing, 1991] J. S. Lansing. *Priests and programmers: technologies of power in the engineered landscape of Bali.* Princeton University Press, Princeton, N.J., 1991.

[Lansing, 2002] J. S. Lansing. Artificial societies and the social sciences. *Artificial Life*, 8, 279–292, 2002.

[Lansing, 2006] J. S. Lansing. *Perfect order: recognizing complexity in Bali.* Princeton University Press, Princeton, N.J., 2006.

[Lansing *et al.*, 1998] J. S. Lansing, J. N. Kremer, and B. B. Smuts. System-dependent selection, ecological feedback and the emergence of functional structure in ecosystems. *J. Theor. Biol.*, 192, 377-391, 1998.

[Lansing *et al.*, 2008] J. S. Lansing, J. C. Watkins, B. Hallmark, M. P. Cox, T. M. Karafet, H. Sudoyo, M. F. Hammer. Male dominance rarely skews the frequency distribution of Y chromosome haplotypes in human populations. *Proc. Natl. Acad. Sci. USA*, 105, 11645–11650, 2008.

[Lansing *et al.*, 2009] J. S. Lansing, M. P. Cox, S. S. Downey, M. A. Janssen, and J. W. Schoenfelder. A robust budding model of Balinese water temple networks. *World Archaeology*, 41(1), 112 – 133, 2009.

[Leigh, 2007] E.G. Leigh Jr. Neutral theory: a historical perspective. *J. Evol. Biol.,* 20, 2075-2091, 2007.

[Levin, 1992] S. A. Levin. The problem of pattern and scale in ecology. *Ecology*, 73, 1943-1967, 1992.

[Lovelock, 1992] J. E. Lovelock. A numerical model for biodiversity. *Phil. Trans. R. Soc. Lond. B*, 338, 365-373, 1992.

[Majoram and Tavaré, 2006,] P. Majoram and S. Tavaré. 2006. Modern computational approaches for analysing molecular genetic variation data. *Nature Reviews: Genetics.*, 7, 759-770, 2006.

[Melčuk, 1988] I. A. Melčuk. *Dependency syntax: Theory and practice.* State University of New York Press, Albany, 1988.

[Padgett and McLean, 2006] J. F. Padgett and P. D. McLean. Organizational Invention and Elite Transformation: The Birth of Partnership Systems in Renaissance Florence. *Am. J. Sociology*, 111(5), 1463-1568, 2006.

[Pagel and Mace, 2004] M. Pagel and R. Mace. The cultural wealth of nations. *Nature*, 428, 275-278, 2004.

[Popper, 1990] K. Popper. *A world of propensities.* Thoemmes, Bristol, 1990.

[Puglisi *et al.*, 2008] A. Puglisi, A. Baroncghelli, and V. Loreto. Cultural route to the emergence of linguistic categories. *Proc. Natl. Acad. Sci.* 105(23), 7936–7940, 2008.

[Richerson and Boyd, 2006] P. Richerson and R. Boyd. *Not by genes alone: how culture transformed human evolution.* University of Chicago Press, Chicago, 2006.

[Robins and Morris, 2007] G. Robins and M. Morris. Advances in exponential random graph (p*) models. *Social Networks*, 29(2), 169-172, 2007.

[Sahlins, 1976] M. Sahlins. *Culture and practical reason.* University of Chicago Press, Chicago, 1976

[Sandwell, 1996] B. Sandwell. *Reflexivity and the crisis of western reason.* Routledge, New York, 1996.

[Saunders, 1994] P. T. Saunders. Evolution without natural selection: further implications of the daisyworld parable. *J. Theor. Biol.* 166, 365-373, 1994.

[Saussure, 1983] F. de Saussure. *Course in general linguistics.* (transl. Roy Harris). Duckworth, London, 1983.

[Slatkin, 1996] M. Slatkin. A correction to the exact test based on the Ewens sampling distribution. *Genet. Res.*, 68, 259-260, 1996.

[Smith and Winterhalder, 1992] E.A. Smith and B. Winterhalder, eds. *Evolutionary ecology and the social sciences.* Aldine/Transaction, New York, 1992.

[Solé, 2005] R. V. Solé. Syntax for free? *Nature*, 434(17 March), 289, 2005.

[Solé *et al.*, 2005] R. V. Solé, B. C. Murtra, S. Valverde, and L. Steels. Language Networks: their structure, function and evolution. Working Papers of the Santa Fe Institute 05-12-042, 2005.

[Stark and Vedres, 2006] D. Stark and B. Vedres. Social Times of Network Spaces: Network Sequences and Foreign Investment in Hungary. *Am. J. Sociology*, 111(5), 1367-1411, 2006.

[Ventura, 2007] L. B. Ventura. The effects of wars on political identities: ideological change and continuity after the Spanish Civil War (1936–39). Presented at the 103rd Annual Meeting, Am. Polit. Sci. Assoc., Chicago, IL. 2007.

[Wasserman and Faust, 1994] S. Wasserman and K. Faust. *Social network analysis — methods and applications.* Cambridge University Press, Cambridge, 1994.

[Winterhalder and Smith, 2000] B. Winterhalder and E. A. Smith. Analyzing adaptive strategies: human behavioral ecology at twenty-five. *Evol. Anthropol.* 9, 51–72, 2000.

[Wood, 2008] E. J. Wood. The social processes of civil war: the wartime transformation of social networks. *Annual Review of Political Science*, 11, 539-561, 2008.

[Zipf, 1949] G. K. Zipf. *Human Behavior and the Principle of Least-Effort.* Addison-Wesley, Reading, MA., 1949.

Part VIII

Psychology

DYNAMICS OF THE PROCESS OF DEVELOPMENT

Adam Sheya and Linda B. Smith

One of the central problems studied by mankind is the problem of the succession of form ... the universe we see is a ceaseless creation, evolution, and destruction of forms and the purpose of science is..., if possible, [to] explain it.

Rene Thom

The origin of new forms is the core question in all the sciences. What are the processes that create new structures that did not exist before? The question of new forms is an especially daunting one when we ask it about ourselves and about mind. What is human intelligence made out of? What makes our rich understanding of causality and intention, of number and objects, of space and time, of language? These are the questions that define cognitive developmental psychology.

The question is so hard that there are some who would deny it. There is always a tension in developmental theory between constructivist approaches that take the whole endeavor as trying to explain how one can get something more from something less and preformationist approaches (in all their guises including current notions of biological determinism) which essentially deny the creative power of development. We will not review this tension. Instead, we take as our starting point that development is constructive. Human babies begin life unable to control their body in even the most rudimentary ways. Yet new patterns emerge as steady and unrelenting progress—from not reaching to reaching, from not walking to walking, from no words to language, to deception, to mathematical reasoning. The classic ideas of developmental theory are all ideas of construction (Piaget, Vygotsky, Baldwin, Kuo, Gottlieb, Oyama, etc.), of something much more from something much less.

These classic approaches do not always use the words and formal theories of dynamic or complex systems. For example, the contemporary tradition know as developmental systems pursues the constructive nature of developmental process as a self-organizing process emergent in the interactions of the organism with the environment. Theorists in this tradition who have explored the probabilistic, epigenetic nature of ontogeny include Kuo [1967], Oyama [1985], Lickliter

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

[2000], Blumberg [2005], Karmiloff-Smith [1992] and Gottlieb [2007]. Other contemporary theorists who have more explicitly embraced dynamic systems include Thelen and Smith [1994], Elman [2003], Lewis and Granic [2000], and Spencer and Schoner [2003]. Both developmental systems and dynamic systems approaches to development have their historical origins with biologists and psychologists such as Waddington [1957], von Bertalanffy [1933], Lewin [1946], and Gesell [1939], who envisioned behavior and development as morphogenetic fields that unify multiple, underlying components. We will not consider the fine points of these various traditions, but rather focus on the central ideas shared across these approaches.

The core ideas to be considered in this essay are all illustrated in one simple example of behavior change, first described by Piaget [1952] in his book, The Origins of Intelligence. The case we will describe concerns an infant's first experience with a rattle, but it is a general example of what Piaget called a secondary circular reaction. Imagine a 3-month-old who has never held a rattle (and indeed cannot grasp an object yet on their own). If we place the rattle in the infant's hand, the infant — because they are always moving — will shake the rattle causing noise. Indeed, as the infant moves the rattle, it will both come into sight and make noise. This will arouse and agitate the infant more, causing more body motions, and thus, causing the rattle to move into and out of sight and to make more noise. Infants at this age have very little organized control over hand and eye. They cannot yet reach for a rattle and if given one, they do not necessarily shake it. But if the infant accidentally moves it, and sees and hears the consequences, the infant will become captured by the activity — moving and shaking, looking and listening. And, here is the important part, the infant in this one continuous experience will, incrementally, through repeated action gain the "intentional goal" of shaking and controlling the rattle. Piaget thought that this pattern of activity — an accidental action that leads to an interesting and arousing outcome, and thus, more activity and the re-experience of the outcome — to be foundational to development itself. At one level, circular reactions are perception-action loops that create opportunities for learning. Each moment of activity, changes the learner and the learning environment and creates opportunities for learning. In the case of the rattle example, the repeated activity teaches how to control one's body, which actions bring held objects into view, and how sights, sounds and actions correspond. At another level, secondary circular reactions are a clear case of creating something more from something less: From its own activity, an infant with no knowledge of rattles, with no obvious goal — no intention — to shake to make sound, creates a new level of organization, what we as theorists (or parents) observing the behavior would call means, goals, and outcomes.

This example also presents the core ideas to be emphasized in this chapter:

1. Coupled heterogeneous components
2. Non-trivial causal spread
3. No predefined endpoint

4. Nested timescales

In what follows, we present examples of each of these aspects of development. We end up with a view of process of development that is complex, multi-layered, and with nontrivial causal spread, a view of process that requires a complex systems approach. In the final section of the essay, we consider the challenges all this presents to current standards of explanation and experimentation.

The power of coupled heterogeneous components in creating change

In his book, Neural Darwinism, Edelman [1987] used Piaget's secondary circular reaction to emphasize the creative power of coupled heterogeneous sensory-motor systems in the creation of cognition. In the context of a specific task, distinct processes are time-locked to each other and to the world, and it is this that creates change. To illustrate this idea, Reeke and Edelman [1984] built a simple computational device. The device's task was to learn to recognize all varieties of the letter A, from the mere experience of looking at As. Figure 1 provides a schematic illustration of the device. The device is composed of two coupled subsystems. The feature-analysis subsystem consists of line detectors excited by corresponding patterns of stimulation. The tracing subsystem gathers information about shape through "eye-movements" as the letter is scanned. The developmental power arises because the activation patterns in these two subsystems are time-locked to each other and to the same physical world enabling straightforward Hebbian learning to create systematic and adaptive change.

That is, at the same time that the feature analyzer is analyzing features, the shape tracer is extracting a global description of shape. The outputs of these two heterogeneous processes, at every step in time, are mapped to each other. There are actually seven mappings being accomplished simultaneously in real time. One mapping, the feature analysis map, maps an input letter to a list of features. The second mapping, the tracing map, maps the input letter to the action sequences of scanning. The third map — from the tracing process to the physical world — selects moment by moment the input (the specific letter part) to both subsystems. The fourth and fifth maps are the recurrent activity within each subsystem: at any moment in time, the activity in the feature analysis subsystem, for example, depends not only on the current input but also on its just preceding state. The sixth and seventh maps are what Edelman calls re-entrant maps; they map the activities of the two subsystems to each other. Thus, two independent mappings of the stimulus to internal activity take qualitatively different glosses on the perceptual information and through their re-entrant connections, by being correlated in real time, by being coupled to the same physical world, educate each other. Reeke and Edelman's simulation successfully taught itself to recognize all varieties of A, generalizing to novel fonts and handwriting, merely from actively "looking" at As.

The human sensory-motor (i.e., cognitive) system is far more complex than the model system shown in Figure 1. There are many more component subsystems and

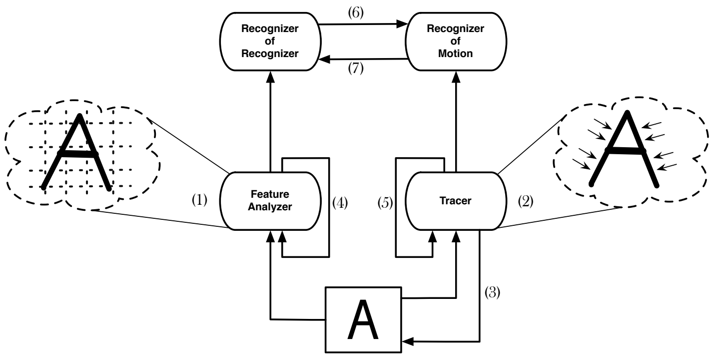


Figure 1. Shows a schematic of Reeke and Edelman's [1984] network model of letter recognition. The letter A at the bottom of the figure depicts the 2-dimensional input array. This input is connected to both a feature analysis system and a tracing system. The recurrent connection for each of these systems represents the system's dependence not only on input but also on its own history. The feature analysis system is composed of feature detectors which track the local structure of the input array, like an oriented line segment. This system outputs to a more abstract detector that integrates information across the local detectors capturing the global structure of the input. The tracing system scans the input array and detects the contour of objects. This system, like the feature analysis system, outputs to a higher-level network that captures shared characteristics of related input arrays. The two higher-level networks are connected to each other enabling the two subsystems (feature analysis and tracing) to work together to classify letters.

variable and complex patterns of connectivity among them. For example, many cortical brain regions associated with specific modalities are comprised of densely connected subregions that capture specific feature states within that modality and also systems of connections among those feature areas. There are also integrations across modalities [e.g., Martin and Chao, 2001; Rogers, Patterson and Graham, 2007; Pulvermuller et al., 2005]. Rapidly advancing work in systems neuroscience and in cognitive neuroscience increasingly documents the pervasiveness of these cross-area integrations and indeed the greater interconnectivity of different sensory systems early in development (see, [Stiles, 2008]). The extreme complexity of so many coupled heterogeneous systems is a powerful creative force.

Nontrivial causal spread

Multimodality may also be key because of what Edelman [1987] calls degeneracy. Originally from the mathematical sense of the word, degeneracy in neural structure means that any single function can be carried out by more than one configuration of neural signals and that different neural clusters also participate in a number of different functions, an idea also know in some literatures as functionality [McIntosh, 1999; 2000; Sporns and Tononi, 2007]. Figure 2 illustrates these ideas in schematic form. The nodes in the figures indicate component processes. These might be imagined to each be a process or sensory system as in Figure 1. What the figure shows is different ways these component processes may be connected. Degeneracy is an intermediate level of complexity in a network between modularity and complete (or random) connections. A modular system is one in which units are portioned into fixed and separate patterns of connectivity. For example, inputs to one module are unaffected by the activity in other modules, as illustrated by the parallel and distinct triplets in Figure 2a. This kind of organization yields highly stable but also highly limited patterns of activity. In a completely connected or random network (Fig. 2c), every component is connected to every other component. The consequence is limited stability and considerable variability. Inbetween is a degenerate network (Fig. 2b), with partially overlapping patterns of connectivity that include dense local connections and sparser longer pathways. The consequence is considerable complexity and many dynamically stable states [Sporns, 2002]. The human brain is generally understood to be degenerate in this sense.

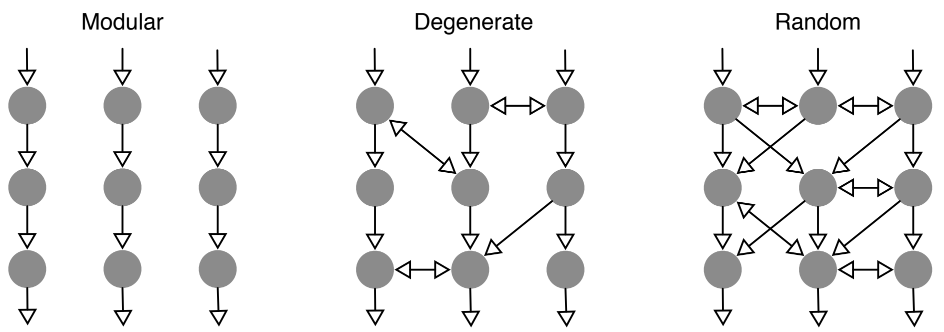


Figure 2. The nodes of the networks depict tightly, densely coupled subnetworks and the edges depict inter-subnetwork connections. The Modular network consists of independent subnetworks. The Degenerate network has a sparse pattern of interconnected subnetworks. The Random network consists of densely connected sub-networks.

One consequence of degeneracy is cascading change. The partial redundancy in functional connectivity means that different tasks can recruit different but overlapping consortiums of subsystems, a fact of considerable developmental importance. In one suggestive experiment, Needham, Barrett and Peterman [2002] fit 2- to 5-month-old infants with Velcro-covered 'sticky mittens'. These mittens enabled the infants to grab objects merely by swiping at them, enabling them to precociously coordinate vision and reaching. Infants who were given 2 weeks of experiences with 'sticky' mittens subsequently showed more sophisticated object exploration even with the mittens off. They looked at objects more, made more visually coordinated swipes at objects than did control infants who had no exploratory experiences with 'sticky mittens'. Needham et al., found that the sticky-mitten task not only facilitated the development of reaching for objects but also visual-oral exploration. That is, infants who had experience with sticky mittens looked at objects more — even in non-reaching tasks — and also mouthed and orally explored objects in more advanced ways.

Figure 3 provides a schematic illustration of what may be the profound significance of these results. Two subsystems — reaching and looking — are coordinated in the sticky mitten task and in so doing educate each other. But these components are also involved in other coordinations, in other tasks that recruit other coalitions of subsystems. Thus, extra experience in the coordination of reaching and looking with sticky mittens ends up not being just about looking and reaching but potentially about other developments, other coordinations, generating cascading developmental consequences in tasks in which some of the same subsystems are involved.

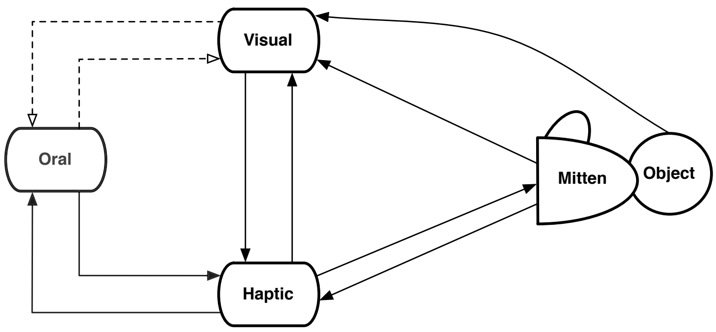


Figure 3. Presents schematic illustration of the affect of 'sticky' mittens on the visual, haptic, and oral systems. The use of 'sticky' mittens during manual exploration reorganizes the coordination of the visual and haptic systems. Although the oral system, grayed in the figure, is not directly involved in this activity, it is connected to the haptic system (infants manually and orally explore objects) and through this connection is potentially influenced by the visual-haptic reorganization.

Infants' learning about transparent surfaces presents another example of the complex causal relations that exist in developmental processes. Learning about transparent surfaces has been interesting to developmental psychologists precisely because transparency violates the usual hand-eye correlations in the world. In most cases, one can directly reach to a seen object by following the line of sight. In contrast to this usual case, transparent yet solid surfaces block direct line of sight reaching paths. Notice, there can be no evolutionary blueprint for learning about transparency since transparent and solid surfaces are not common in nature but are artifacts, invented by people.

Babies do not do well with this violation of usual expectations. The experimental demonstration of infants' difficulties is provided by Diamond [1990] who presented 9-month-old infants with toys hidden under boxes. The boxes were either opaque — hiding the toy — or transparent enabling the infants to see the toy under the box. As illustrated in Figure 4, the boxes were open on the side, so that infants, by reaching to that side, could retrieve the object. Diamond found that infants were able to retrieve the toy from the opaque container, reaching around to the side opening. However, they frustratingly failed in the same task with a transparent container. The problem with the transparent container is that infants attempt to reach for the toy in the usual way (trying to put their hand through the transparent surface following the line of sight), and thus, fail to search for and find the opening.

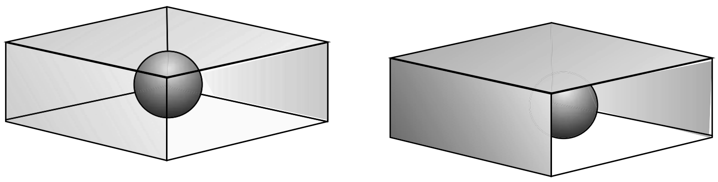


Figure 4. Depicts the stimuli used in Diamond [1990] and Titzer *et al.*'s [2003] experiments on transparency. The picture on the left depicts the transparent box and the picture on the right depicts the opaque box. Both boxes have openings on the right side allowing infants to retrieve contained objects.

But infants are not stuck with what is a maladaptive approach in the modern artifact-filled world. Infants easily learn to solve this problem through their ordinary interactions with transparent containers. In a demonstration of this point, Titzer [1997, see also Smith and Gasser, 2005] conducted a microgenetic study in which 8-month-old infants were given either a set of opaque or transparent containers to play with at home. Parents were given no instructions other than to put these containers in the toy box, making them available to the infants during play. When the infants were 9 months old, they were tested in Diamond's task. The babies who had played at home with opaque containers failed to retrieve objects from transparent ones just as in the original Diamond study. However, infants who had played at home with the transparent containers sought out and rapidly found the openings and retrieved the object from the transparent boxes.

Infants' at home explorations of the transparent containers did not include the specific task of sideways retrieval of objects, although it seems likely that in their spontaneous play, objects were both put into and retrieved from the openings of the containers. Titzer [1997] proposed that during play — through the coordination of seeing and touching as they put objects in and out of the containers — infants learned to recognize the subtle visual cues that distinguish solid transparent surfaces from openings and had learned that surfaces with the visual properties of transparency are, in fact, solid and thus to reach around to the openings. In the framework of Figure 3, the haptic cues from touching the transparent surfaces educated vision, and vision educated reaching and touching, enabling infants when subsequently tested in Diamond's task to find the openings in transparent containers.

Critically, these coordinations of touch and sight had broader cascading consequences beyond retrieving objects from small transparent effects that some might want to summarize under the umbrella of a "concept of solid and supporting surface". These consequences were observed through an additional transfer task, the visual cliff. The visual cliff was originally designed to study infant depth perception [Gibson and Walk, 1960]. It consists of a transparent but solid surface placed over a visual "drop off." In the classic version of the task, 8- and 9-month-old infants are placed on the "safe" side of the surface. Infants this age typically avoid the visual drop, not moving onto the transparent surface that extends over the vertical drop.

Titzer [1997] tested the babies who had participated in their training study with transparent and opaque containers on the visual cliff and found that the infants who played with the transparent containers at home did not avoid the visual cliff. Instead they happily crawled onto the transparent surface over the drop off, showing no apprehension whatsoever. The babies who had played with opaque containers, in contrast, avoided the "edge" refusing — even when called by their parent — to approach the visually (but not tactually) apparent cliff. The infants, who had extensive play with transparent containers, were apparently both sensitive to the subtle visual cues that specify the solidity of a transparent surface and to the cues felt from their hands as they felt the surface, and thus, were confident of its support. Again, two subsystems — seeing and touching — are coordinated when playing with transparent containers, each system educating the other in the discovery of relevant regularities to that coupling. The changes in these component subsystems — the regularities found in one task will be transported to other tasks that also recruit these same subsystems [for another compelling example, see Bertenthal and Campos, 1990]. In this way, the coordination of multimodal subsystems in specific tasks may create abstract and transportable ideas.

Development creates with no predefined endpoint

Knowing how to shake rattles, or catch things with sticky mittens, or reach for things in transparent containers cannot be pre-specified by evolution. The solutions infants find must be created. These are new tasks, requiring new solutions. Human beings continually learn and do things that have never been done before. Even young children exhibit cognitive skills that were not imaginable generations ago, learning about transparency, pictures, and videos, about drag and click, how to program, and the rhythm and syntax of texting, as well as those older intellectual artifacts of reading and multiplication. There seems no limit to what we can potentially incorporate, make into everyday cognition, even into child's play. In this way, human cognition is decidedly not predetermined and not preset, but is rather open-ended. This open-endedness begins (but does not end) with our ability to discover, though our own activity, goals.

The lesson from rattle-shaking, sticky mittens, and transparency is this: Goals are not prior to the task but are themselves emergent in the infants' engagement with the world. Prior to shaking the rattle, or catching a toy with the sticky mittens, infants can have no specific goal to shake to make noise or to swat to snatch an object. One example of goal discovery through action is "infant conjugate reinforcement" [Rovee-Collier and Hayne, 1987; Angulo-Kinzer, Ulrich and Thelen, 2002]. Infants (as young as 3 months) are placed on their backs and their ankles are attached by a ribbon to a mobile which is suspended overhead. The mobile, which produces interesting sights and sounds, provides the infant with many time-locked patterns of correlations. More importantly, infants themselves discover these relations through their own movement patterns. The faster and harder infants kick, the more the mobile moves fueling the infants' kicking. This is a highly engaging task for infants; they smile and laugh, and often become angry when the contingency is removed. This experimental procedure, like the world, provides complex, diverse, and never exactly repeating events yet all perfectly time-locked with the infant's own actions. It is spontaneous non-task-related movement that starts the process off by creating the opportunity for the coordination of the infant's action with the mobile's. It is this coordination that ultimately defines the task, and thus, becomes the "goal". In other words, the term goal is just a descriptor of this coordination. That is, in every context there are many potential actions, although, only one action can be performed at any given moment. What this experiment illustrates is that feedback from spontaneously generated action reshapes the landscape of possible actions, making some more stable, or likely, in a given context.

Because the goals that drive "intentional" action (and learning) are not pregiven but are discovered by the individual through action, development itself is highly specific to the individual. Different children will follow different developmental paths that depend on the specific tasks they discover and the intrinsic dynamics of their own system. One elegant demonstration of the individual nature of developmental trajectories is Thelen, Corbetta, Kamm, Spencer, Schneider, and Zernicke's [1993] week-by-week study of the transition from not reaching to reaching for visually presented objects. Thelen et al. studied four babies and found four different patterns of activity, and thus, four different patterns of development. The basic developmental pattern was: The presentation of an enticing toy is arousing and elicits all sorts of nonproductive actions, and very different actions in individual babies. These actions are first, quite literally, all over the place with no clear coherence in form or direction. But by acting, each baby in its own unique fashion, sooner or later makes contact with the toy — banging into or brushing against it or swiping it. These moments of contact select some movements, carving out patterns that are then repeated with increasing frequency. Over weeks, the cycle repeats — arousal by the sight of some toy, action, and occasional contact. Over cycles, increasingly stable, more efficient and more effective forms of reaching emerge.

As infants produce different movements — in their uncontrolled actions initiated by the arousing sight of the toy — they each discover initially different patterns and different developmental tasks to be solved. Some babies in the non-reaching period hardly lift their arms at all. Other babies flail and flap and are always moving. These different babies must solve different problems to grasp an object. The flailer needs to become less active lowering the hands to bring them to midline and create balance. The placid baby needs to be more active, to raise her hands and to lift them up.

What is remarkable in the developmental patterns observed by Thelen and collaborators is that each infant found a solution by following individual developmental pathways that eventually converged to highly similar outcomes. Because action defines the task and because action — through the coordination of heterogeneous sensory systems — finds the solution, development is very much an individual and context-dependent matter, and not pre-defined prior to action itself. Given the constraints of the world, of human bodies, and of the heterogeneous and multimodal system out of which intelligence is made, different individuals will develop broadly similar systems (what one might summarize as "universals") but at its core, development (like evolution) is opportunistic, individualistic, and local in its causes.

Nested timescales

People do new things when faced with new challenges, often changing their behavior flexibly in the moment, when faced with some new input from the world. People learn new skills after hours, days and months of practice. Developmental change — the emergence of reaching, or walking or talking — occurs over weeks, months and years. If we are to explain how development creates we must understand processes of change over multiple timescales. Accordingly, we present one more task first introduced by Piaget [1954] to illustrate the importance of nested times scales of change in understanding the creative force of development.

Piaget [1954] invented the A-not-B task to answer the question of "when do infants acquire the concept of object permanence?" He devised a simple object-hiding task that works as follows: The experimenter hides a tantalizing toy under a lid at location A and the infant reaches for the toy. This A-location trial is repeated several times. Then, there is the crucial switch trial: the experimenter hides the object at new location, B. At this point, 8- to 10-month-old infants make a curious 'error'. If there is a short delay between hiding and reaching, they reach not to where they saw the object disappear, but back to A, where they found the object previously. This 'A-not-B' error is especially interesting because it is tightly linked to a highly circumscribed developmental period: infants older than 12 months of age search correctly on the crucial B trials. Thelen, Smith, Schöner, Spencer and colleagues[1] have offered a formal theory, the dynamic field model to explain the A-not-B error as the emergent product of multiple causes interacting over nested timescales. Here we use the ideas from this formal theory to highlight the importance of nested times scales.

What might cause an infant to reach — at any one moment — to A or to B? Figure 5 provides the key task analysis: One influence on behavior is the continually present box and hiding wells. A second influence is the specific or transient input (top row) that occurs when the experimenter hides the toy. This is a momentary event that will be remembered in the system for some time, the endurance of this memory is, of course, critical to infants' success in the task. Then the infant reaches, an act that will create its own memories. Moreover, the infant sees the transient event and reaches repeatedly to A before the shift trial at B. These are events and memories at another timescale. The dynamic field theory account of the error is at its core an account of the dynamics of how all these processes at different times scales change over the course of the trials in the tasks and how they integrate to yield a decision to reach to A or to B.

Figure 6a illustrates the evolution of activation on the very first A trial. Before the infant has seen any object hidden, there is activation in the field at both the A and B locations from the two covers. As the experimenter directs attention to the A location by hiding the toy, it produces a high, transient activation at A. Then a decision emerges in the field over time. When the activation at a location stabilizes, the infant reaches to that location. Most crucial for this account is that once infants reach, a memory of that reach becomes another input to the next trial. Thus, at the second A trial, there is some increased activation at site A because of the previous activity there. This combines with the hiding cue to produce a second reach to A. Over many trials to A, a strong memory of previous actions builds up. Each trial embeds the history of previous trials. Now consider the crucial B trial (Fig. 6b). The experimenter provides a strong cue to B. But as that cue decays, the lingering memory of the actions at A begins to dominate the field, and indeed, over time, to shift the decision back to the habitual, A side.

The model clearly predicts that the error is time dependent: there is a brief period immediately after the hiding event when infants should search correctly, and indeed they do [Smith, Thelen, Titzer and McLin, 1999; Bremner and Bryant, 2001]. Using this model as a guide, experimenters can experimentally make the error come and go, almost at will. This is achieved by changing the delay, by heightening the attention-grabbing properties of the covers or the hiding event, and by increasing and decreasing the number of prior reaches to A [Smith et al.,

[1]Smith, Thelen, Titzer and McLin, 1999; Diedrich et al., 2001; Thelen, Schöner, Scheier, and Smith, 2001; Spencer, Smith and Thelen, 2001; Schutte and Spencer, 2002; Spencer and Hund, 2003; Schutte, Spencer and Schöner, 2003; Spencer and Schutte, 2004; Clearfield, Diedrich, Smith and Thelen, 2006; Feng et al., 2007

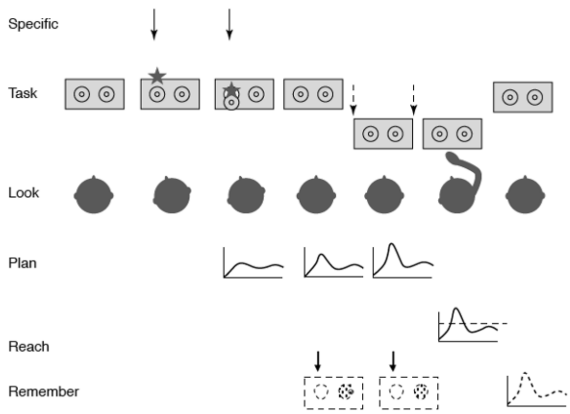
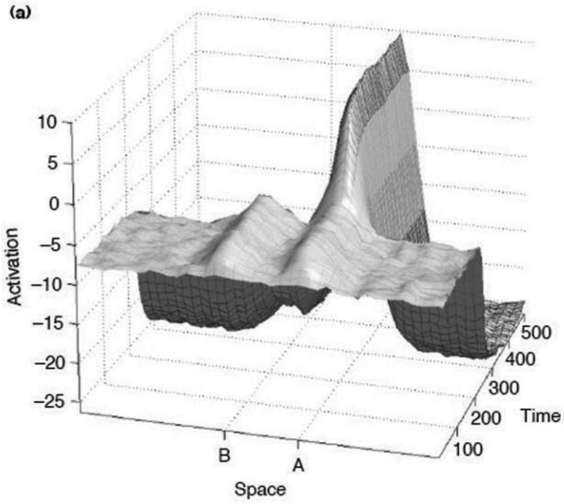


Figure 5. Depicts a task analysis of the A-not-B error for a typical A-location hiding event. The Task input consists of the box and hiding wells that are a continually present visual input. The specific or transient input (top row) consists of the hiding event in the 'A' well (on the left here). After a delay the infant is allowed to search for the toy. During these events, the infant looks at the objects in view, remembers the cued location and undertakes a planning process leading to the activation of reach parameters, followed by reaching itself. Finally, the infant remembers the parameters of the current reach.

1999; Marcovitch, Zelazo and Schmuckler, 2002; Marcovitch and Zelazo, 2006]. Researchers have made the error occur (and not occur!) even when there is no toy to be hidden [Smith et al., 1999]. Directing attention to an in-view object (A) heightens activation at the location and, in the experiment, infants reach to that continually in-view object. Subsequently, when the experimenter directs attention to a different nearby in-view object (B), infants watch, but then reach back to the original object (A).

Experimenters have also made the error vanish by making the reaches on the B trials different in some way from the A trial reaches. In the model, these differences decrease the influence of the A trial memories on the activations in the field. One experiment achieved this by shifting the posture of the infant [Smith et al., 1999]. An infant who sat during the A trials would then be stood up, as shown in Fig. 7, to watch the hiding event at B, during the delay and during the search. This posture shift causes even 8- and 10-month-old infants to search correctly, just like 12-month-olds. These results underscore the highly decentralized nature of error: the relevant causes include the covers on the table, the hiding event, the delay, and the past activity of the infant and the feel of the body of the infant. This multicausality demands a rethinking of what is meant by knowledge and development. Intelligence is creative — as a unified system — over multiple timescales, including the timescales of individual action as well as the timescales of



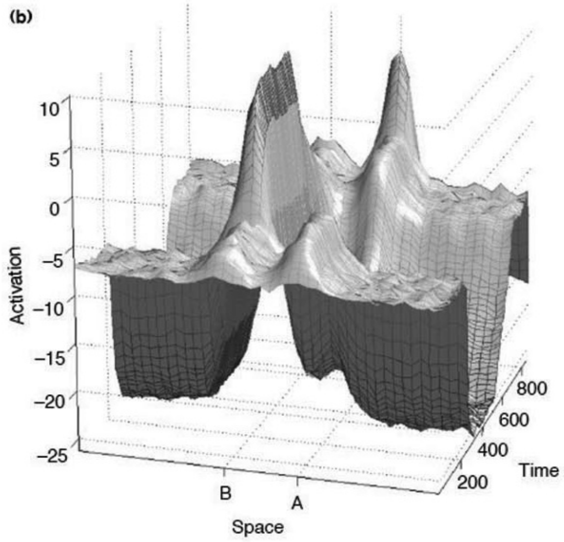


Figure 6. (a) The time evolution of activation in the planning field on the first A trial. First the activation rises as the object is hidden and then, owing to self-organizing properties in the field, is sustained during the delay. (b) The time evolution of activation in the planning field on the first B trial. Before the hiding event, there is heightened activation at A, owing to memory for prior reaches. Activation rises at B as the object is hidden there, but as this transient event ends, owing to the memory properties of the field, activation at B declines and at A rises.

learning and developmental change. The A-not-B error has been of most interest to developmental theorists, however, because of the changes over developmental time. This is an error tightly linked to a few months in infancy. Or is it?

Spencer, Smith and Thelen [2001] invented an A-not-B task that was suitable for 2-year-olds by hiding toys in a sandbox. The surface of the sand presents a uniform field so there are no markers to indicate the two possible hiding locations. Experimenters gave toddlers many trials at location A, and then hid the toy at location B. With a delay of 10s, the toddlers, having watched the toy being hidden at location B, still returned to the A location to dig in the sand for the toy. Indeed there are many other situations in which both children and adults fall back on a habit despite new information [Butler, Berthier and Clifton, 2002]. Nonetheless, in the standard A-not-B task, infants change their behavior over 2 months. In the field model increasing the resting activation of the field simulates this by making it easier for the input from the hiding cue to form a self-sustaining peak at B and thus compete with the A memory.

Figure 7. Shows how the infant's posture was shifted from a sitting position an A trial (left) to a standing position for a B trial (right).

If self-sustaining memories drive the successes of older children, then we must ask where they come from: What are infants doing everyday that improves their location memory? One possibility is their self-locomotion. Crawling and self-locomotion appears to improve the spatial memories of infants and to predict success in the A-not-B task [Bertenthal and Campos, 1990]. Indeed, teaching babies to self-locomote (in walkers) causes them to succeed at a younger age than usual in the A-not-B task [Kermoian and Campos, 1988]. Like sticky-mittens and visual cliffs, the causes of developmental change may lie in seemingly disparate achievements that involve overlapping component processes.

Dynamic systems approaches to cognition have been criticized as being suitable for the study of perception and action, but not for higher level human cognition [Abrahamsen and Bechtel, 2006; Dietrich and Markman, 2003] — the cognition that, in these views, underlies such abstract aspects of human thought as mathematics, science, allegory, and poetry. Accordingly, in the next section we outline how the coupling of sensory-motor processes to the perceivable consequences of those processes may make higher-level cognition.

Doing with images

Alan Kay, one of the founding fathers of object-oriented programming and graphical user interfaces (the idea behind the original MacIntosh windows and drag and click we now all use) is a visionary computer scientist who gave a talk in 1987 with a cult-like following. The talk's title was "Doing with images makes symbols". Inspired by Piaget [1952], Bruner and colleagues [1956] and Vygotsky [1978], Kay proposed that abstract ideas (and symbolic thought) were built out of real-time sensory-motor interactions with images, that is, with the stable *perceivable consequences* of our own actions. This is like the idea of a closed-loop in active vision in which every action creates perceivable consequences that also guide action at the next step. But the key idea here is that some perceivable consequences are special in being *image-like,* in that they are stable and enduring, a perceivable constant that is coupled to the messier context-specific and continuous dynamics of perception and action. These external stabilities — "artifacts" that may be initially produced without plan or goal — are a profound force on human cognition.

The development of spatial classification provides an interesting case. Between their first and third birthdays, children begin to use space to represent similarity, putting like things close together [Sugarman, 1983]. Indeed, during this period they become almost compulsive spatial sorters. Confronted with an array of 4 identical cars and 4 identical dolls, they physically group them — moving all the cars spatially close to each other and spatially apart from the groups of dolls even though there is no explicit task to do so. They are so reliable at doing this that many developmental psychologists use the task as a way to measure young children's knowledge of categories [e.g., Nelson, 1973; Mandler, Bauer and McDonough, 1991; Rakison and Butterworth, 1998]. Their reasoning is that if a 2-year-old knows that two objects are the same kind of thing, she should spatially group them together. A perhaps just as interesting question is why the child spatially groups objects at all.

The developmental evidence suggests a progressive discovery that starts around 7 months with the onset of stable visually-directed reaching and culminates with exhaustive sorting around 36 months of age. When presented with sets of objects of like kinds children first bimanually grasp objects presented at their midline, then bring objects together, one in each hand, by holding them at midline, then bang objects together, then stack them, and finally place objects adjacent to each other [Forman, 1982]. Although this trend is not sequential and linear, in that these behaviors overlap in developmental time and only really change in degree, through out this progression of behaviors there is a clear trend from first bringing any two objects into spatial proximity, either by holding, banging, stacking, or placing, to more consistently bringing like objects into spatial proximity. Making it seem that object similarity is scaffolded by spatial similarity.

Around 12 months of age infants begin to produce a different kind of manipulation in which like objects are manipulated in sequence [Sugarman, 1983]. For example, if given 4 cars and 4 dolls, a 12-month-old may systematically push each of the four cars. Around 18 months of age, children will not only manipulate objects from one category in sequence but also systematically manipulate in different ways objects from two different categories, for example, first pushing each of four cars, one after another and then touching each of four dolls in turn. Sometime after 24 months, the sorting seems more purposeful with all of one kind gathered to form one group and the other kind left unorganized.

Sheya and Smith [Sheya, 2005; Sheya and Smith, in press] propose that this developmental pattern emerges though the child's own actions, actions that at first have no goal of creating a classification. Four behavioral tendencies are proposed to drive the process. The first is that infants reach to objects in which they are interested. The second is that infants have a tendency to repeat just performed motor acts, and in particular to repeat reaches to nearby locations (e.g., [Smith *et al.*, 1999]). The third is that perceptually similar objects are similarly enticing to infants. The fourth is that infants notice the outcomes of their own actions. These four tendencies can be understood in terms of a dynamic salience map, a map that determines where infants look to and reach next.

Imagine an array of 8 toys, 5 of one kind and 3 of another as illustrated in Figure 8. Attention to and the touching of one toy alters the salience map, by activating the spatial location of that toy. This activation can spread along two potential dimensions — physical space or a feature space according to the feature similarity of the objects. In their behavioral experiments, Sheya and Smith showed that activation in this salience space (as measured by the next toy touched) spreads mostly by space for younger infants (12-month-olds) but by feature similarity for older infants (18-month-olds). However, for both age groups closeness and similarity interacted, with *close* and *similar* things being more salient than both close and dissimilar things and distant and similar things.

These tendencies can create the sequential touching of like objects. As children are drawn to nearby and similar things, they are likely — through just these processes alone — to drop similar things near others, with the interactive effects of spatial proximity and physical similarity increasing the salience of reaching, again and again, to like and near things. A system whose activity is biased to both reach to similar locations and to reach to similar objects will, as consequence of reaching and dropping those things, end up with similar things near each other. It is here that Alan Kay's idea enters in. This unplanned consequence of similar things ending up near each other creates *an image*, a stable array of like things in proximity and apart from different things.

Namy, Smith, and Gershkoff-Stowe [1997] conducted a microgenetic study with the goal of encouraging the development of spatial classification in toddlers who did not yet spatially group like objects. The children's "training" was a fun task of putting objects into a shape sorter. As illustrated in Figure 9, the shape sorter was

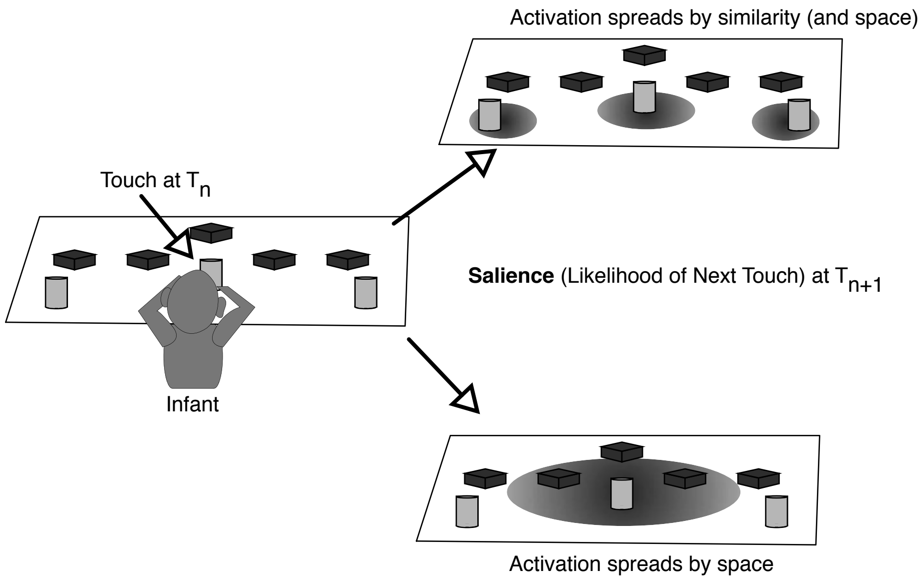


Figure 8. The board on the left depicts a reaching event in which the infant reaches to the center object. The upper right board shows a change in salience when activation spreads along dimensions of similarity and the lower board depicts how salience increases when activation spreads along physical spatial dimensions.

a transparent container structured so that children could see the objects once they had been dropped inside. Children were given two different kinds of objects (e.g., blocks and dolls) that might be put into the container. The opening on the top of the shape container only allowed one type of object to fit inside the hole. Children at this age have strong perseverative tendencies to repeat the same action, and so they (quite happily) attempted to put all the objects into the container — the kind that fit and the kind that did not. But their actions led to only one kind actually being in the container together and thus in spatial proximity. Their actions thus produced a stable image of like things being near each other and apart from different things.

This experience turned these children in spatial classifiers, advancing them several months in this developmental progression. In the transfer task, children were given sets of 8 objects — 4 of one kind or 4 of another and no shape sorter. Children who had previously sorted with the transparent shape sorter — that enabled the children to see the product of their activity — were much more likely to sort the objects into spatially organized groups than children in the control conditions. Namy *et al.* [1997] suggest that their experimental procedure intensified the experiences characteristic of children's every day activities. Children are attracted to objects that are physically similar, they tend to repeat similar acts and this leads

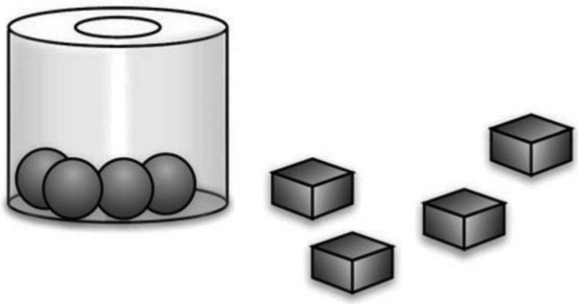


Figure 9. Depicts the transparent sorter used in Namy *et al.* [1998]. The sorter enabled young children to spatial segregate objects of different kinds.

to similar things often ending up close in space. As children play with things in their world, a world in which physically like things have similar propensities, like things will end up near each other. Moreover, as children interact with a world structured by adults with a similar psychology, they will encounter cups near cups on the shelves and socks all together in the sock drawer. In these ways, spatial proximity becomes *the* foundational metaphor for both informal ideas and formal mathematical theories about similarity.

The processes outlined in Figure 8 for how children's sequential touching of like and near objects changes are fundamentally the same as those posited by the dynamic field account of the A-not-B error. Infants' interactions with objects in the shape sorter task lead to new goals and new constructs just as in secondary circular reactions. The coupling of heterogeneous processes to events in the world through perceiving and acting in particular tasks creates change in the sensory-motor system on the whole. The problem for a theory of development is to explain all this.

Explaining development

Any theory of development must explain how to get something more from something less. Thus, any theory must resolve the fundamental contradiction of development: open and constructive process leading to seemingly universal outcomes. Despite their differing personalities, including differing capacities and learning styles, all intact human infants learn to walk, to progress from making the A-not-B error to not making it, to speak their native language and to form intense social relationships. This may make it seem that development is shaped by universal and direct causal mechanisms. But when one looks at the details of development, the picture seems far less simple and direct. Children from the same family grow up to be amazingly different from one another. Children with social and economic advantages sometimes fail in life, whereas those from impoverished backgrounds sometimes overcome them. As this delicate dance of sameness and difference illustrates, there is considerable indeterminacy within development, making it unrealistic to reduce the process of development to a simple causal model of change. Rather what is needed is an understanding of the process of development itself, and in particular of what are the sources of indeterminacy and how they are related to sources of structure within the process of developmental change.

The sources of indeterminacy can be found at the microlevel of development, the moment-to-moment, real-time production of behavior. At this level the picture is messier, allowing the idiosyncratic real-time activities of the child to dynamically generate the trajectory of change [Muchisky, Gershkoff-Stowe, Cole and Thelen, 1996]. Moment-to-moment, the state of the system and the task at hand, create change and moment-to-moment the developmental trajectory. At the same time, each new coordination enables new possible assemblies of subsystems that generate new actions that create new tasks (opportunities for reorganization) that create new organizations. This cascading process generates its own sources of order through the coordination and re-coordination of the sensory-motor systems that are sensitive to the constraints of a physical world. Because of both these features — because on the one hand the mechanism of change is the individual's momentary task, making space for variability, and, on the other hand, the rich interrelationships within individuals mean that many differently detailed coordinations can nonetheless underwrite the same overall functional capacities (like walking and talking) — development is open to multiple outcomes and multiple paths to the same ends. Simply put, the key to explaining development lies in understanding the process by which the everyday activities of children create change — both the universal attainments and the individual pathways.

The starting point for any such understanding is the dynamics of development. However, multiple levels of analysis at multiple timescales are needed to examine how many components open to influence from the external world interact yielding coherent higher-order behavioral forms that then feedback on the system, and change it. In this complex interaction, every neural event, every reach, every smile and every social encounter sets the stage for the next and is the real-time causal force behind change. The hard problem of explaining development is constructing a characterization that accounts for these details of development and that also facilitates novel predictions.

Solving this hard problem might mean identifying a class of dynamical systems that captures the emergence properties of naturally developing systems. Identifying potential classes is not a trivial task in itself. A complex systems perspective facilitates this process of discovery by providing a language of change through which insights gained from simulation and experimental observation can be combined. Simulation based studies provide a unique testing ground for theoretical ideas and models because they allow full control over the brain-body-environment system. In addition, because they require the implementation of intuitions, that is real working agents, potentially novel theoretical insights to the fundamental nature of cognition can be gained.

One style of simulation study is to analyze simplified artificial neural-body-environment systems that reflect what we know about the real dynamics of bodies. For example, Kuniyoshi and Sangawa [2006] present a simulation that uses real world physics to examine how spontaneous exploration by a infant can entrain motor dynamics and produce new coherent patterns of activity, like crawling. This type of simulation provides both a powerful test for models of developmental change and a demonstration of the rich information generated from spontaneous action in a physical world. Combining this kind of simulation with experimental study of human infants will provide insight into the dynamics of the feedback mechanism that drives the emergence of new behavioral forms.

Beer [1996] provides another approach to simulation that focuses on the nature of cognition. Beer's work in this area has primarily used evolutionary algorithms to evolve dynamical neural networks that control the behavior of model agents. Analyzing the dynamics of the resulting agent-environment systems provides insight into the minimal cognitive apparatus needed to perform basic tasks, like discrimination. Instead of assuming that an agent has a richly structured cognitive system, Beer treats cognition as a dynamical system that in conjunction with the dynamics of the agent's body and environment is tuned to produce effective behavior. Because he uses evolutionary algorithms instead of building his intuitions about cognition into the agents, he is able to find minimal requirements, given a body and environment, of a task that are in the language of dynamics and not in terms of gross descriptions of cognitive processes, like "intention".

A complex system perspective provides a means to deal directly with the intricate details of developmental change and a language of change in which to construct an explanation of development in terms far removed from vague intuitive constructs, like intention. Thus, moving theoretical thinking in psychology beyond simple gross causal models to theories that not only explain development but potentially provide the means to build artificial developmental systems.

BIBLIOGRAPHY

[Abrahamsen and Bechtel, 2006] A. Abrahamsen and W. Bechtel. Phenomena and mechanisms: Putting the symbolic, connectionist, and dynamical systems debate in broader perspective. In R. J. Stainton (Ed.), *Contemporary debates in cognitive science: Contemporary debates in philosophy*, (pp. 159-185). Malden, MA: Blackwell Publishing, 2006.

[Angulo-Kinzler, Ulrich, and Thelen, 2002] R. M. Angulo-Kinzler, B. Ulrich, and E. Thelen. Three-month-old infants can select specific leg motor solutions. Motor Control, 6(1), 52-68, 2002.

[Beer, 1996] R. D. Beer. Toward the evolution of dynamical neural networks for minimally cognitive behavior. In P. Maes, M. Mataric, J. Meyer, J. Pollack and S. Wilson (Eds.), *From animals to animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior* (pp. 421-429). MIT Press, 1996.

[Bertenthal and Campos, 1990] B. I. Bertenthal and J. J. Campos. A systems approach to the organizing effects of self-produced locomotion during infancy. *Advances in Infancy Research*, 6, 1-60, 1990.

[Blumberg, 2005] M. S. Blumberg. *Basic Instinct: The Genesis of Behavior.* New York: Thunder's Mouth Press, 2005.

[Bremner and Bryant, 2001] A. Bremner and P. Bryant. The effect of spatial cues on infants' responses in the AB task, with and without a hidden object. *Developmental Science*, 4(4), 408-415, 2001.

[Bruner, Goodnow, and Austin, 1956] J. S. Bruner, J. J. Goodnow, and G. A. Austin. *A study of thinking*. Oxford, England: Wiley, 1956.

[Clearfield *et al.*, 2006] M. W. Clearfield, F. J. Diedrich, L. B. Smith, and E. Thelen. Young infants reach correctly in A-not-B tasks: On the development of stability and perseveration. *Infant Behavior and Development*, 29(3), 435-444, 2006.

[Diamond, 1990] A. Diamond. Developmental time course in human infants and infant monkeys and the neural bases of inhibitory control in reaching. *Annals of the New York Academy of Science*, 608, 637- 676, 1990.

[Diedrich *et al.*, 2001] F. J. Diedrich, T. M. Highlands, K. A. Spahr, E. Thelen, and L. B. Smith. The role of target distinctiveness in infant perseverative reaching. *Journal of Experimental Child Psychology*, 78(3), 263-290, 2001.

[Dietrich and Markman, 2003] E. Dietrich and A. B. Markman. Discrete thoughts: Why cognition must use discrete representations. *Mind and Language*, 18(1), 95-119, 2003.

[Edelman, 1987] G. M. Edelman. *Neural Darwinism: The theory of neuronal group selection.* Basic Books: New York, 1987.

[Elman, 2003] J. Elman. Development: It's about time. *Developmental Science*, 6(4), 430-433, 2003.

[Forman, 1982] G. E. Forman. A search for the origins of equivalence concepts through a microanalysis of block play. In G. E. Forman (Ed.), *Action and thought: From sensorimotor schemes to symbolic thought*, (pp. 97-134). New York: Academic Press, 1982.

[Feng *et al.*, 2007] J. Feng, J. Spencer, G. Schöner, M. W. Clearfield and E. Thelen. A novel technology for investigating the dynamics of infant perseverative reaching. *Behavior Research Methods*, 39(4), 911-919, 2007.

[Gesell, 1939] A. Gesell. Reciprocal interweaving in neuromotor development. *Journal of Comparative Neurology*, 70, 161–180, 1939.

[Gibson and Walk, 1960] E. J. Gibson and R. D. Walk. The "visual cliff". *Scientific American*, 202(4), 64-71, 1960.

[Gottlieb, 2007] G. Gottlieb. Probabilistic epigenesis. *Developmental Science*, 10, 1-11, 2007.

[Karmiloff-Smith, 1992] A. Karmiloff-Smith. *Beyond Modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press, 1992.

[Kermoian and Campos, 1988] R. Kermoian and J. J. Campos. Locomotor experience: A facilitator of spatial-cognitive development. *Child Development*, 59, 908-917, 1988.

[Kuniyoshi and Sangawa, 2006] Y. Kuniyoshi and S. Sangawa. Early motor development from partially ordered neural-body dynamics: Experiments with a cortico-spinal-musculo-skeletal model. *Biological Cybernetics*, 95, 589-605, 2006.

[Kuo, 1967] Z. Y. Kuo. *The dynamics of behavior development: An epigenetic view*. New York: Random House, 1967.

[Lewin, 1946] K. Lewin. Behavior and development as a function of the total situation. In L. Carmichael (Ed.), *Manual of child psychology* (pp. 791–844). New York: Wiley, 1946.

[Lewis and Granic, 2000] M. D. Lewis and I. Granic, eds. *Emotion, development, and self-organization: Dynamic systems approaches to emotional development*. New York: Cambridge University Press, 2000.

[Lickliter, 2000] R. Lickliter. An ecological approach to behavioral development: Insights from comparative psychology. *Ecological Psychology*, 12, 319-334, 2000.

[Mandler *et al.*, 1991] J. M. Mandler, P. J. Bauer, and L. McDonough. Separating the sheep from the goats: Differentiating global categories. *Cognitive Psychology*, 23(2), 263-298, 1991.

[Marcovitch and Zelazo, 2006] S. Marcovitch and P. D. Zelazo. The influence of number of A trials on 2-year-olds' behavior in two A-not-B-type search tasks: A test of the hierarchical competing systems model. *Journal of Cognition and Development*, 7(4), 477-501, 2006.

[Marcovitch *et al.*, 2002] S. Marcovitch, P. D. Zelazo, and M. A. Schmuckler. The effect of the number of A trials on performance on the A-not-B task. *Infancy*, 3(4), 519-529, 2002.

[Martin and Chao, 2001] A. Martin and L. L. Chao. Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology Special Issue: Cognitive neuroscience*, 11(2), 194-201, 2001.

[McIntosh, 1999] A. R. McIntosh. Mapping cognition to the brain through neural interactions. *Memory*, 7, 523–548, 1999.

[McIntosh, 2000] A. R. McIntosh. Towards a network theory of cognition. *Neural Network*, 13, 861–870, 2000.

[Muchisky *et al.*, 1996] M. Muchisky, L. Gershkoff-Stowe, E. Cole, and E. Thelen. The epigenetic landscape revisited: A dynamic interpretation. In C. Rovee-Collier, and L. P. Lipsitt (Eds.), *Advances in infancy research, vol. 10.* (pp. 121-159). Westport, CT: Ablex Publishing 1996.

[Namy *et al.*, 1997] L. L. Namy, L. B. Smith, and L. Gershkoff-Stowe. Young children discovery of spatial classification. *Cognitive Development*, 12(2), 163-184, 1997.

[Needham *et al.*, 2002] A. Needham, T. Barrett, and K. Peterman. A pick me up for infants' exploratory skills: Early simulated experiences reaching for objects using 'sticky' mittens enhances young infants' object exploration skills. *Infant Behavior and Development*, 25(3), 279-295, 2002.

[Nelson, 1973] K. Nelson. Some evidence for the cognitive primacy of categorization and its functional basis. *Merrill Palmer Quarterly*, 19(1), 21-39, 1973.

[Oyama, 1985] S. Oyama. *The ontogeny of information: Developmental systems and evolution.* Cambridge, England: Cambridge University Press, 1985.

[Piaget, 1952] J. Piaget. *The origins of intelligence in the children.* New York: W. W. Norton, 1952.

[Piaget, 1954] J. Piaget. *The construction of reality in the children.* New York: Basic Books, 1954.

[Pulvermüller *et al.*, 2005] F. Pulvermüller, O. Hauk, V. V. Nikulin, and R. J. llmoniemi. Functional links between motor and language systems. *European Journal of Neuroscience*, 21(3), 793-797, 2005.

[Rakison and Butterworth, 1998] D. H. Rakison and G. E. Butterworth. Infants' attention to object structure in early categorization. *Developmental Psychology*, 34(6), 1310-1325, 1998.

[Reeke and Edelman, 1984] G. N. Reeke and G. M. Edelman. Selective networks and recognition automata. *Annals of the New York Academy of Science*, 426, 181-201, 1984.

[Rogers *et al.*, 2007] T. T. Rogers, K. Patterson, and K. Graham. Colour knowledge in semantic dementia: It is not all black and white. *Neuropsychologia*, 45(14), 3285-3298, 2007.

[Rovee-Collier and Hayne, 1987] C. Rovee-Collier and H. Hayne. Reaction of infant memory: Implications for cognitive development. In H. Reese (Ed), *Parallel distributed processing: Explorations in the microstructure of cognition, Volume 2.* Cambridge, MA: MIT Press, 1987.

[Schutte and Spencer, 2002] A. R. Schutte and J. P. Spencer. Generalizing the dynamic field theory of the A-not-B error beyond infancy: Three-year-olds' delay- and experience-dependent location memory biases. *Child Development*, 73(2), 377-404, 2002.

[Schutte *et al.*, 2003] A. R. Schutte, J. P. Spencer, and G. Schöner. Testing the dynamic field theory: Working memory for locations becomes more spatially precise over development. *Child Development*, 74(5), 1393-1417, 2003.

[Sheya, 2005] A. Sheya. The Origins of Classification: Proximity, Perseveration, and Similarity. To be presented at the biennial meeting of the Society for Research on Child Development, Atlanta, GA, 2005.

[Sheya and Smith, in press] A. Sheya and L. B. Smith. Changing priority maps in 12- to 18-month-olds: An emerging role for object properties. *Psychonomic Bulletin & Review.* In press.

[Smith and Gasser, 2005] L. Smith and M. Gasser. The Development of Embodied Cognition: Six Lessons from Babies. *Artificial Life*, 11, 13-30, 2005.

[Smith *et al.*, 1999] L. B. Smith, E. Thelen, R. Titzer, and D. McLin. Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychological Review*, 106(2), 235-260, 1999.

[Spencer and Hund, 2003] J. P. Spencer and A. M. Hund. Developmental continuity in the processes that underlie spatial recall. *Cognitive Psychology*, 47(4), 432-480, 2003.

[Spencer and Schöner , 2003] J. P. Spencer and G. Schöner. Bridging the representational gap in the dynamic systems approach to development. *Developmental Science*, 6, 392-412, 2003.

[Spencer, and Schutte, 2004] J. P. Spencer and A. R. Schutte. Unifying representations and responses: Perseverative biases arise from a single behavioral system. *Psychological Science*, 15(3), 187-193, 2004.

[Spencer *et al.*, 2001] J. P. Spencer, L. B. Smith, and E. Thelen. Tests of a dynamic systems account of the A-not-B error: The influence of prior experience on the spatial memory abilities of two-year-olds. *Child Development*, 72(5), 1327-1346, 2001.

[Sporns, 2002] O. Sporns. Network analysis, complexity and brain function. *Complexity*, 8, 56-60, 2002.

[Sporns and Tononi, 2007] O. Sporns and G. Tononi. Structural Determinants of Functional Brain Dynamics. In V.K. Jirsa and A. R. McIntosh (Eds.), *Handbook of Brain Connectivity.* (pp. 117-148). Berlin: Springer, 2007.

[Stiles, 2008] J. Stiles. *The fundamentals of brain development: Integrating nature and nurture.* Cambridge, MA: Harvard University Press, 2008.

[Sugarman, 1983] S. Sugarman. *Children's early thought : developments in classification.* New York: Cambridge University Press, 1983.

[Thelen *et al.*, 1993] E. Thelen, D. Corbetta, K. Kamm, and J. P. Spencer. The transition to reaching: Mapping intention and intrinsic dynamics. *Child Development*, 64(4), 1058-1098, 1993.

[Thelen *et al.*, 2001] E. Thelen, G. Schöner, C. Scheier, and L. B. Smith. The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioural and Brain Sciences*, 24, 1–86, 2001.

[Thelen and Smith, 1994] E. Thelen and L. B. Smith. *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press, 1994.

[Thom, 1989] R. Thom. *Structural Stability and Morphogenesis.* New York: Westview Press, 1989.

[Titzer, 1997] R. C. Titzer. Infants' understanding of transparency: A reinterpretation of studies using the object retrieval task and visual cliff. (Doctoral dissertation, Indiana University, 2005) *Dissertation Abstracts International*, 58, 1570, 1997.

[von Bertalanffy, 1933] L. von Bertalanffy. *Modern theories of development.* London: Oxford University Press, 1933.

[Vygotsky, 1978] L. Vygotsky. *Mind in Society: The development of higher mental functions.* Cambridge, MA: Harvard University Press, 1978.

[Waddington, 1957] C. H. Waddington . *The strategy of the genes.* London: Allen and Unwin, 1957.

LIVING IN THE PINK: INTENTIONALITY, WELLBEING, AND COMPLEXITY

Guy C. Van Orden, Heidi Kloos and Sebastian Wallot

A barrel racer trains many years to move in sync with her horse. To keep her seat, she rises and falls in phase with powerful centripetal lunges around each barrel. The result is a skilled coordination in which rider and horse gallop together. Consider another example. A teacher steps forward while speaking. Every step pre-engages coordinative structures, flexing and extending muscles across torso, arms, and neck to guarantee balance in a continuous anticipatory flow. The racer needs only to race, and the teacher to teach, but what happens is a vastly complicated coordination of minds and bodies with their environments. Coordination is essential to cognition and behavior, yet except in motor coordination it has not been a prominent topic of cognitive science.

In this essay, we discuss how complexity science has filled this gap. We begin with problems inherited from conventional cognitive science, for example the question of intentionality. We then discuss conceptual building blocks of complexity with respect to brains, bodies, and behavior. These include constraints, phase transitions, interdependence, and self-organized criticality – concepts that address emergent coordination among system components. From there we go on to discuss ubiquitous pink noise in human performance. Pink noise is a fundamentally complex phenomenon that reflects an optimal coordination among the components of person and task environment. Departures from this optimum occur in advanced aging and dynamical disease, including Parkinson's disease, as we will discuss. We conclude this essay with a survey of present challenges and opportunities for complexity and cognitive science.

1 INTENTIONALITY AND OTHER DILEMMAS

Intentionality is central to subjective experience and permeates all human activities. It plays an equally prominent role in cognitive experiments, with special significance for cognitive science. Before meaningful data can be collected, intentions must be invoked in the participant to perform as instructed. Data — the foundation of what scientists know about cognition – depend fundamentally on the will, purpose, and goals of the participant. Yet the role of intentions in data collection and laboratory experiments is usually ignored [Vollmer, 2001]. Indeed a *Science Watch* forum concluded that experiments tap involuntary, automatic, or unconscious processes exclusively [Science Watch, 1999].

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

Intentionality suggests a capacity to bring behavior into existence, to cause behavior. The intention to step forward to teach, for example, might cause the right leg to move forward. Yet intentions cannot be ordinary causes and still make sense scientifically. This is because the causal viewpoint ignores the question of what causes the intention in the first place. Maybe the intention to step was caused by the intention to teach, and the intention to teach was caused by the intention to remain employed. Still what caused the intention to remain employed? Either the intention to remain employed has a magical status, as a prime mover homunculus, or we enter the logical regress of seeking the cause of the cause of the intention to behave [Juarrero, 1999].

Intentions also require that cognition stays open to outside factors to promote intended goals, while at the same time ignoring irrelevant factors that might derail them. Once instructed to pay close attention to ball handling in a basketball game, for example, the observer will fail to notice the man in the gorilla suit who stops and pounds his chest while walking through the scene [Simons and Chabris, 1999]. How does the mind stay connected to the outside world, but only selectively, in pursuit of its goals? The question is to the crux of selective attention, the capacity to turn a blind eye to aspects of the environment that are irrelevant to purposes at hand [Mack and Rock, 2000]. A conventional solution might be a decision device that could select relevant factors and purposes. Yet which homunculus decides whether things capture involuntary attention?

Of course dilemmas in conventional cognitive science are not limited to questions of intention [Hollis *et al.*, 2009]. Take for example the coming into existence of a completely novel insight or novel behavior. The dilemma stems from equating cognition with information processing, either as *mentalese* by analogy to language, *computation* by analogy to computer software, or *activation* by analogy to neurons, synapses, and action potentials. For information processing, novelty becomes either a simple combination of existing structures, juxtaposed or added together in representations, or novelty must preexist in some way before the novel behavior is realized. The latter solution yields another logical regress: If the cause of novelty preexists, then what caused the preexisting cause of novelty?

Another dilemma in conventional cognitive science is presented by the protracted failure to connect mind to body. Conventional theories have failed to bridge the gap that separates mind and body. This failure to naturalize mental constructs stems from the causal gamble that functional components of cognitive activities, perception, and memory can be isolated and explained [Bechtel, 2009]. Yet the gamble has led to a hodgepodge of conflicting mechanisms, with little agreement about details such as boundaries or number of mechanisms, the ontological status of mechanisms, the relation between cognitive mechanisms and brains, or the developmental basis of cognitive mechanisms (e.g., [Dreyfus, 1992; Harley, 2004; Searle, 1980; Stanovich, 2004; Thelen and Smith, 1994; Uttal, 2001; 2007; Watkins, 1990; Weldon, 1999]). Lacking clearly worked out cause and effect relations, mind and body appear to lack common currency for interaction. And regarding the results of neuroimaging research: "How do we say something is somewhere if we do not exactly know what that something is?" [Greenberg, 2002, p. 111].

The often-voiced hope of the larger research community is the possibility that converging data and theory may themselves sort out the existing hodgepodge. Perhaps converging operations remain to be discovered, maybe through a triangulation of mutually acceptable results about brain, behavior, and conscious experience [Roepstorff and Jack, 2004]. However, this hope continues to rest on the assumed causal and methodological transparency among brain, behavior and consciousness. Transparency requires concatenated effects, meaning effects follow one from the other like dominoes tipping one into the next down a line. Consequently, interaction effects must be additive in proportion to factorial manipulations, but linearity and additivity are scarcely evident or nonexistent [Van Orden and Paap, 1997]. Each cognitive factor appears to interact multiplicatively with every other, and each interaction changes in the context of every new additional factor. Consequently, the sum of evidence across the vast empirical literature of cognitive science yields an equally vast higher-order multiplicative interaction [Van Orden *et al.*, 2001]. The unsupported assumption of transparency and concatenated domino effects has resulted in a crisis for measurement, which is seldom discussed [Michell, 1999].

Complexity theory circumvents these dilemmas by emphasizing emergent coordination, temporary dynamical structure, and the creation of information in behavior. Complexity science is not concerned with cause and effect primarily, so it averts dilemmas that arose from seeking causes of behavior as information processors or homunculi. Those efforts are replaced by a search for strategic reductions to laws, principles, and mechanisms of emergent coordination. Such strategic reductions find the same principles at work across different systems and at all levels of a system. In the next section we describe ideas from complexity science that introduce these principles.

2 CONCEPTUAL BUILDING BLOCKS

The view of human behavior as emergent coordination offers a new and theory-constitutive metaphor for cognitive behavior, a complete reconstitution of method, theory, and assumptions. In this section, we define theoretical terms of complexity science that have proven useful in cognitive and behavioral science. They culminate in the ideas of self-organized criticality and soft-assembly: Living systems are attracted to optimal temporary states of flexible coordination, which best guarantees contextually appropriate behavior and the wellbeing of the actor.

2.1 Constraints and Control Parameters

Constraints arise in relations among a system's components, and they reduce the degrees of freedom for change. Consider the constraints that limit the range of motion of an arm or a leg. Relations among joints, muscles, fasciae, and the nervous system allow limbs to move some ways, but not others. They reduce the degrees of freedom for change in limb motion. An expanded example of constraints, less tangible perhaps, is the indefinite sea of constraints among living beings and their worlds [Shanon, 1993]. These include relations with artifacts and the environment, the myopic limits on attention and stream of consciousness, the constraints that arise from idiosyncratic details of each actor's previous history, and in relations to other living beings.

Even with limiting constraints, however, a body in motion retains far too many degrees of freedom to be explicitly or mindfully controlled. For example, estimating the parts to be coordinated, a human body has something like 10^2 joints, 10^3 muscles and 10^{14} cells (Turvey, 1990). For each part that must be causally controlled, a conventional model must accord one controlling structure to each degree of freedom. Given that behavior is highly variable, causal resources are quickly overwhelmed, historically well known as the degrees-of-freedom problem in on-line kinematics of behavior [Bernstein, 1967].

In contrast to causal control, complexity science emphasizes constraints as temporary structures, not unlike the temporary coordination among molecules in a convection cell. They are conceived as emerging from the temporary coupling among embodied components and among components and the environment [Van Orden *et al.,* 2003]. Like a newly formed convection cell controls the fluid molecules of which it is composed, constraints reduce degrees of freedom in coordination. Emergent constraints have the capacity to further self-organize into still higher-order emergent structures. That is to say, first-order emergent structures may combine iteratively into second-order and still higher-order temporary dynamical structures. This iterative capacity has been observed in brain data, for example [Ito *et al.,* 2005]: First-order emergent patterns of coordination, visible in coordination among signals of separate EEG leads, were themselves part of the second-order coordination across time. Iterative higher-order emergence is bounded only by material, temporal, metabolic, and informational limits of the system.

Constraints that control behavior are summarized mathematically in control parameters. To explain, consider the stepping behavior in infants: Soon after birth, and long before learning to walk, a young infant, held above the ground with feet touching the floor, will move legs and feet as though stepping. This early stepping behavior then disappears and remains absent until later in the first year, when it reappears. A conventional causal story sees two different causes behind the two instances of stepping behavior, with no connection in between: Initial stepping behavior is attributed to primitive reflexes that quickly disappear as the baby matures; and the later stepping behavior is attributed to the maturation of a motor schema for walking [McGraw, 1945].

The constraint account, in contrast, focuses on a single control parameter to capture the developmental sequence of stepping behavior. In particular, there are two main constraints that determine the availability of stepping behavior: (1) the strength of the baby's leg, and (2) the weight of the leg. The relevant control parameter is a ratio that pits leg strength against leg weight. Early in development, the baby's legs are relatively light in weight compared to how strong they are, making initial stepping possible. As the baby gains weight, however, gravity's pull on the heavy legs exceeds the strength of the legs, and stepping behavior disappears. In turn, as the baby builds more strength during the first year, stepping behavior reappears [Thelen and Smith, 1994].

The control parameter for stepping behavior captures two salient relations between the infant actor and her environment. Specifically, the numerator in the example (leg weight) summarizes *embedding* constraints in the infant's relation to the environment. This type of constraint delimits *affordances*, the dispositions of the surrounding environment directly relevant for action [Gibson, 1979]. Conversely, the denominator in the example (leg strength) concerns *embodied* constraints of the actor. This second type of constraint refers to *effectivities*, the capacities and capabilities of the actor to exploit the available affordances [Shaw *et al.*, 1982].

Explaining change in behavior through changes in control parameters has several advantages compared to traditional accounts. First, control parameters give a more inclusive account of development because they can account for individual differences across participants. Imagine, for example, an infant with very strong (or very skinny) legs. Such a child is likely to retain the capacity for stepping behavior throughout the first year. The changing control parameter for this particular child can be measured precisely. Conventional accounts, on the other hand, require exceptional assumptions to account for idiosyncratic differences. The initial stepping reflex might be stronger in this child than in another, inhibition of the reflex might be delayed, or the motor schema might mature earlier than predicted — or some combination of these possibilities. Further problems arise in having to determine normative development in this case. Yet, movement and its development is hardly uniform [Adolph, 2009].

Second, control-parameters can account for ubiquitous context effects. In stepping behavior, context changes in holding a non-stepping baby upright in a shallow pool, as opposed to outside of the pool, and previously nonexistent stepping will now appear. Or the context can be changed by placing weights on the legs of a baby who can step — and existing stepping behavior will disappear [Thelen *et al.*, 2002]. Conventional accounts assume that successful performance reflects the presence of an underlying cognitive structure, while unsuccessful performance reflects its absence. Such accounts are quickly overwhelmed by the sheer number of context effects, often found in the same person and after only trivial changes of context [Kloos *et al.*, 2009; Van Orden *et al.*, 1999].

At the minimum, context sensitivity requires that performance reflect some form of interaction between the cognitive structures of the actor and the context of the environment. Yet complexity goes well beyond a mere interaction. In each different context, a different mesh of available constraints reduces degrees of freedom to favor kinematics suitable for that context or task protocol, e.g. [Balasubramaniam *et al.*, 2000; Flach, 1990; Riley, 2007; cf. Glenberg, 1997]. No two situations yield identical constraints, so a laboratory's situated mesh of constraints specifies a unique niche for performance [Flach *et al.*, 2007]. In a similar vein, no two persons embody identical constraints because no two persons have identical histories. Consequently, behavior in the same task will differ in quality as well as quantity [Ashby *et al.*, 1993; Balakrishnan and Ashby, 1991; Holden, 2002; Holden *et al.*, 2009; Luce, 1986; Maddox *et al.*, 1998; Molenaar, 2008].

2.2 Critical States

As behavior changes across development, say from the presence of early stepping behavior to its absence, the relevant control parameter passes through a critical value, a value that defines a critical state of the system. In the stepping example, the critical value (and therefore the critical state) is reached when the pull of gravity exactly equals leg strength. Now the two opposing actions, stepping and not stepping, are in precise balance, and therefore equally possible. In this critical state, even tiny changes in control parameters may tip the balance and break the symmetry of the poised alternatives. That is to say, even tiny changes in the environment-infant system can be relevant contingencies that break symmetry.

Given that relevant contingencies are necessary to enact behavior, and they suffice to enact behavior, they can be conceived as causes. For example, a hungry dieter who comes across a candy bar will likely eat it, though he might prefer to have made a healthier choice. The simple contingency of first coming across the candy bar enacts behavior consistent with the need for food. The mere sight of the candy bar therefore causes the dieter's lapse in healthy eating. Laboratory findings sometimes discover nothing but effects of contingencies. This might explain why scientists take the prevalence of reported contingency effects to imply the lack of intentionality in laboratory behavior, e.g. [Science Watch, 1999]. Conscious will might be nothing more than the illusion of causality after all, e.g. [Wegner, 2002]. Yet these conclusions are misguided. Before a contingency can enact behavior, the body must already be in a critical state. Available constraints must first specify propensities to act. Only then do mere contingencies have the power to cause behavior.

Critical states exist until relevant contingencies occur. Importantly, critical states are not perturbed by irrelevant factors, factors that do not favor a particular action over any other. Change in a baby's arm weight or finding a toy candy bar while hungry are not sufficiently relevant to the specified critical states. Only relevant events can favor a relevant propensity. So, in a sense, the critical state can "filter out" irrelevant contingencies, and explain selective attention. It is the critical state that allows the actor to stay open to outside events, without being derailed by irrelevant factors.

The prominent role of critical states, susceptible to relevant contingencies, may also explain why mindful, forbidding self-control is notoriously difficult to put into action. It is well known, for example, that a dieter forbidding himself to eat candy, or telling himself to "eat healthily," are ineffective diet solutions [Baumeister and Heatherton, 1996; Rachlin, 2000]. A focus on healthy or unhealthy edible things has the side effect of instantiating propensities to eat that remain susceptible to accidental candy bars. In an effective solution the dieter concentrates on the abstract end-goals of dieting, such as facilitating connectedness to others, or a change in personal wellbeing [Fujita and Han, 2009]. The more abstract goal is less likely to include propensities for kinematics to grab up the first food available. The abstract focus makes candy bars less salient as food and more salient as diet busters.

Critical states are not only relevant to understanding selective attention and the relevance of contingencies. Far beyond, they are proposed to be the center of coordination. Rather than coming into existence passively, as control parameters change, complex systems are drawn toward critical states, they self-organize critical states [Bak, 1997; Bak *et al.*, 1987]. Note the superficial paradox of self-organized criticality: Critical states are by nature unstable, given that the smallest relevant contingency can collapse the system into one action or another, so critical states must be repellers, boundaries between basins of attraction. However, critical states can also be attractors [Chialvo, 2008].

2.3 Phase Transitions

As the system passes through a critical state, a phase transition takes place. The term phase transition comes from thermodynamics and describes how phase relations among molecules change suddenly and qualitatively to more efficiently dissipate heat. As a system passes through a critical state (and a control parameter passes through the critical value), the system components suddenly and spontaneously reorganize to produce a different kind of behavior, together at almost the same time. Immediately before a phase transition, disorder will increase in the system. This increase coincides with the break up of existing structure prior to the reorganization. After the phase transition, the level of disorder drops to a lower level than the level it was originally. This drop is called negentropy, and it stands for the difference between the entropy before the start of the phase transition and the entropy immediately after the phase transition. Negentropy coincides with the emergence of new thermodynamically advantaged structure due to an increase in how quickly the system can export entropy.

Changes in entropy have been observed for phase transitions that occur during problem solving [Stephen *et al.*, 2009]. Given the turning direction of the first gear in a chain of gears, the problem to be solved was the turning direction of the last gear. Typical participants transition from tracing the direction of each gear to a parity strategy, after the insight that every other gear turns in the same direction [Dixon and Kelley, 2006; 2007; Schwartz and Black, 1996; Dixon and Bangert, 2004]. Angular velocities of finger movements were densely sampled across trials of separate gear problems. As expected, entropy in angular velocity increased just prior to the phase transition, while negentropy was observed immediately after. The pattern was replicated in densely sampled eye movements in the gear-turning task [Stephen *et al.*, 2009], and it was found in a balance-beam problem-solving task [Cox and Hasselman, 2009]. Negentropy results are compelling. They suggest that new problem solutions are thermodynamically advantaged, a profound similarity between phase transitions in problem solving and phase transitions in nonliving physical systems.

Bifurcation theory provides a mathematical account of phase transitions in nonlinear dynamical systems. The change from the absence to the presence of a behavior, say, from absent stepping behavior to stepping behavior, is one kind of bifurcation, while a change from one type of behavior to a different type of behavior is a different kind of bifurcation. Through reliable mathematical accounts, the nature of phase transitions can be understood [Meillassoux, 2008]. Moreover, if bifurcation theory should fail to illuminate changes among coordinative structures in human behavior, we would lack any other alternative in which qualitative changes generalize across instances.

Phase transitions occur in many living and nonliving systems. A mix of chemicals forms qualitatively different patterns when the petri dish is tipped; amoebas lacking sufficient food resources transition from single-celled organisms to a multi-cell spoor-bearing slime mold [Nicolis, 1989], coordination between human behavior and a metronome-beat transitions from syncopation to synchrony as the metronome speed increases [Kelso, 1995], just to name a few. Despite differences in types of systems, these phase transitions share common diagnostic patterns, called catastrophe flags, with common theoretical underpinnings.

Examples of catastrophe flags include critical fluctuations and critical slowing, both of which were observed in phase transitions of brain and behavior [Kelso *et al.*, 1992]. The behavioral task was to flip a switch repeatedly between the beats of a metronome, in syncopation with the metronome beats. This task was chosen because syncopation behavior loses stability at a critical value of metronome frequency, and then transitions to synchrony, flipping the switch on the beat [Kelso, 1995]. To test for catastrophe flags in this phase transition, metronome frequency was increased incrementally to perturb the coupling between participant and metronome. SQUID brain images, EEGs, and behavioral measures were recorded continuously. Indeed, just before the phase transition, the perturbation produced critical fluctuations and critical slowing. That is to say, in all measures there was a nonlinear increase both in the variability in the phase relation between beat and behavior (demonstrating critical fluctuations) and in the recovery time after perturbation to regain syncopation (demonstrating critical slowing). It was as though a protracted struggle occurred in brain and body to decide which propensity would be expressed, syncopation or synchrony.

The crucial finding, however, was that brain and behavior reorganize together, at the same time, too close in time to allow information processing. In particular, the lag in reorganization of brain and behavior was no more than 170 msec, not enough time for information processing, though sufficient time for the creation of information in the collapse of a critical state. The virtually simultaneous reorganization of brain and body agrees with reports of ultra-fast cognition, reliable perception after impossibly brief visual displays, for example, and reliable cognitive performance with electric speed. Perception and action occur too fast to allow for information processing to take place. Sometimes the body appears to make do with one-way activation, traveling at speed, from eye to hand [Fabre-Thorpe *et al.*, 2001; 1996; Greene and Olivia, 2009; Grill-Spector and Kanwisher, 2006; Thorpe, 2002; VanRullen and Thorpe, 2002].

2.4 Interdependence and Soft-Assembly

Is it surprising that finger movements reveal the same changes in entropy as eye movements during the phase transition in a problem-solving task? Neither finger movements nor eye movements have an obvious causal connection to the participant's reasoning or to the novel insight. Yet they both show characteristic signatures of a phase transition. Complexity theory anticipates such coupling. This is because components of a complex system are interdependent, one with another; they change each other's dynamics as they interact with each other. Interdependence allows soft assembly of behavior, meaning that behavior emerges and cannot be parsed further, or reduced, into component functions that would exist in a dormant state, even when their behavior is not present [Hollis *et al.*, 2009; Kloos and Van Orden, 2009; Turvey and Carello, 1981].

Interaction-dominant dynamics are the basis of interdependence and emergence; interactions among components dominate the intrinsic dynamics of the components themselves [Jensen, 1998]. Interaction-dominant dynamics originate in multiplicative interactions and feedback among the interacting components. As a result, they predict non-additive, strongly nonlinear effects [Holden *et al.*, 2009], and emergent properties that cannot be deduced from causal properties of components [Boogerd, *et al.*, 2005]. In contrast, *component-dominant dynamics* underlie the expectation of additive effects embedded in Gaussian random variability [Van Orden *et al.*, 2003]. Gaussian variability, for example, is the variability of independent perturbations that sum up as measurement noise.

A consequence of interdependence is to allow a system's phase space to be reconstructed from a well-chosen one-dimensional data series of repeated measurements. In essence, if every part affects every other part then coordinated changes can be recovered from measured values kept in the time-ordered sequence in which they were collected. The reconstructed phase space is a rearrangement of data points as neighbors, which means they are close together in the phase space and products of the dynamics in that neighborhood. Phase space reconstruction requires the right tools of course, and elegant mathematical theorems, now taught in undergraduate mathematics classes, prove that higher-dimensional neighborhood structures can be unfolded and made available for additional analysis [Mañé, 1981; Takens, 1981].

If each component's dynamics is entangled with the dynamics of every other component, it can become impossible to isolate components and study them separately. So how do we determine which components are involved in a particular cognitive activity? This concern reflects the strategy of seeking isolated components, typical of conventional information-processing accounts. It is motivated by the idea that the parts of a system have distinct functions that are preserved or encapsulated through component-dominant dynamics. Component-dominant dynamics underlie the expectation that behavioral effects result from interaction among components that do not change their intrinsic properties [Van Orden *et al.*, 2003]. An arch is an example of a component-dominant system. While blocks interact to form an arch, they are not interdependent in their function. Supportive properties of a particular arch can be deduced from the material composition and arrangement of the component blocks.

How do we know if a system is driven by component-dominant or interaction-dominant dynamics? The crux is whether the system shows strongly emergent properties. Component-dominant systems have only weakly emergent properties and their behaviors can be deduced from causal properties of components and their arrangement, see also [Boogerd *et al.*, 2005]. Conversely, interaction-dominant systems have strongly emergent properties, visible in catastrophe flags discussed earlier. They are also expressed in scaling relations across repeated measurements. Such scaling relations are now so commonly observed in cognitive science that they are claimed to be *universal* [Gilden, 2001; Kello and Van Orden, 2009; Riley and Turvey, 2002]. They are even found in subjective evaluations of wellbeing, such as repeated self-esteem ratings over the course of a year [Delignières *et al.*, 2004], or changes in mood over the course of a day [Isenhower *et al.*, 2009]. They provide strong evidence that human behavior soft assembles in interaction-dominant dynamics.

2.5 Homeorhesis

At one time, medicine, biology, and the behavioral sciences embraced homeostasis as the guiding dynamic of wellbeing. Figure 1 illustrates how repeated measurements would appear in homeostatic dynamics. Homeostasis assumes mean values come from set points of a system, and random noise around the mean values comes from external perturbations. Absent external perturbations, homeostasis predicts that systems come to rest at their average values.

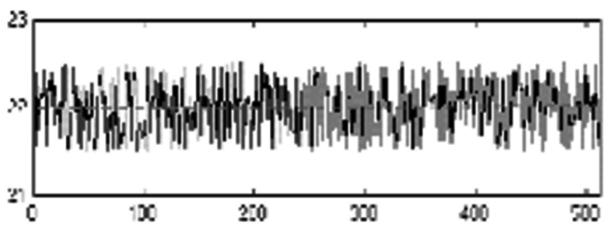


Figure 1. A random noise data series centred on a mean value indicated by the red line, to illustrate homeostatic behavior. The random variation comes from peturbations to the static mean.

In line with this hypothesis, the body was thought to sustain an average heartbeat, for example, to satisfy the average needs of cells for nutrients and oxygen. Organisms were thought to find sufficient food to maintain an average nutrient base. And medicine acquired the goal of returning systems to their capacity to sustain homeostasis, sometimes recruiting artificial devices to do the same job [West, 2006]. Although homeostasis was intuitive, it did not correctly anticipate the ubiquitous cycles in living systems. The heart does not have a reliable average time between beats and cycles of nutrient intake, energy liberation, and waste expulsion, essential for life, recur on the multiple scales of cells, organs and the body as a whole. Homeostasis was therefore challenged by the homeokinesis hypothesis, in biophysics and physiology.

Homeokinesis is the idea that a body and its relations to the environment can be broken down into distinct cycles of nonlinearly stable dynamics. Homeokinetic systems repeat their behavior in limit cycles [Iberall, 1970; Iberall and McCulloch, 1969]. Figure 2 illustrates the predicted pattern of repeated measurements governed by homeokinetic dynamics, a limit cycle plus random noise. Proponents of homeokinesis assembled most of the conceptual pieces necessary for a robust account of variability in living systems. They could, in their time, with their tools, demonstrate component limit cycles (plus random noise) in physiology and all the way out into behavior, e.g. [Kay, 1988]. However, evidence against homeokinesis existed even as it was proposed. This is because homeokinesis posited a distinction between dynamics on different timescales. For example, limit cycles of cell dynamics were thought to be independent of limit cycle dynamics of organs and organ systems, and between organisms and environments.

Homeokinesis allowed interactions between cyclic processes, in nutrient and oxygen transport for example, but not among their cyclic dynamics. Consequently, a change in an organism's circadian rhythm with the environment should not change the cycle frequencies among organs or among cells. Yet, it is now widely accepted that organism-environment cycles are linked to the cycles within organisms. For example, a feckless chicken kept in constant red light (to break the entrainment with the environment's circadian rhythm) suffers a break down of healthy coordination among heart rate, locomotor activity, and deep body temperature [Winget *et al.*, 1968]. Healthy intrinsic dynamics of the chicken's body require entrainment to the circadian rhythm to remain in order.

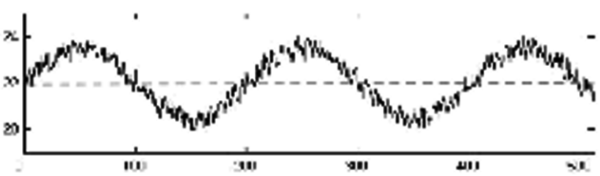


Figure 2. A sine-wave data series with added random noise to illustrate a homeokinetic procss. The red line indicates the mean of a data series around which the limit cycle fluctuates.

More slowly changing cycles were eventually recognized as supplying supportive constraints to sustain faster changing cycles [Simon, 1973; Newell, 1990]. The hypothesis predicts that more slowly changing dynamics can constrain faster changing dynamics, but not vice versa. The prediction leads to a nonsensical conclusion, however, considering current knowledge. Timescales of behavior do not overlap much with timescales of the brain. Measured changes in overt behavior happen on the time scales of years, months, days, hours, minutes, and seconds. Yet the brain's slowest delta waves index changes with a period of about 2.5 seconds [Buzsáki, 2006]. Therefore, timescales of behavior are mostly too slow to be controlled by the brain. Even conscious self-control, by some estimates, occurs more slowly than cycles in the brain, e.g. [Iberall, 1992]. So how could the brain function in control of behavior? A logical conclusion might be that the brain functions to smooth out the kinematics of behavior, in a kind of dithering function, like the high-frequency dithering that makes digital music sound more like analog.

Eventually, with the development of new tools, scientists could reliably distinguish chaotic oscillators from limit cycles with random noise, e.g. [Mitra *et al.*, 1997]. As a result, limit cycles were rejected as the basis of cycles in physiology and behavior. In their place, a hypothesis of homeorhesis was proposed. Homeorhesis is the idea that the dynamics of living systems reflect flexible entrainment to changes in their environments. It predicts a kind of flow of behavior through the environment that negotiates constraints, reflecting previous as well as present relations with the environment, e.g. [Warren, 2006]. Homeorhesis hinges on the idea that the brain, body, and environment soft assemble behavior. It is therefore a direct analog to the idea of interdependence and soft-assembly discussed above.

To summarize Section 2, conceptual building blocks from complexity science enhance our understanding of cognitive behavior. Embedding and embodied constraints combine in control parameters whose critical values define critical states of phase transitions. As a system passes through a critical state, the system undergoes a phase transition, a qualitative change in its organization to soft assemble qualitatively different behavior. Phase transitions are identified using catastrophe flags like critical fluctuations and critical slowing. Phase transitions are shaped by temporary dynamical structures as constraints, which allow flexibly situated soft assembly of cognition and behavior. In the next sections, we build upon and expand these ideas to discuss nontrivial changes in how to understand cognition and behavior.

3 THE THIRD KIND OF BEHAVIOR

Before complexity science, variation in repeatedly measured values was divided into two categories: regular changes from one measured value to another, or random changes. Regular changes were thought to be the explainable variance, while random variance was equated with measurement error. In cognitive science, explainable variance was conscripted to reveal component mechanisms of *memory*, *reasoning*, *syntax*, *semantics,* and so on. The empirical variance we describe in

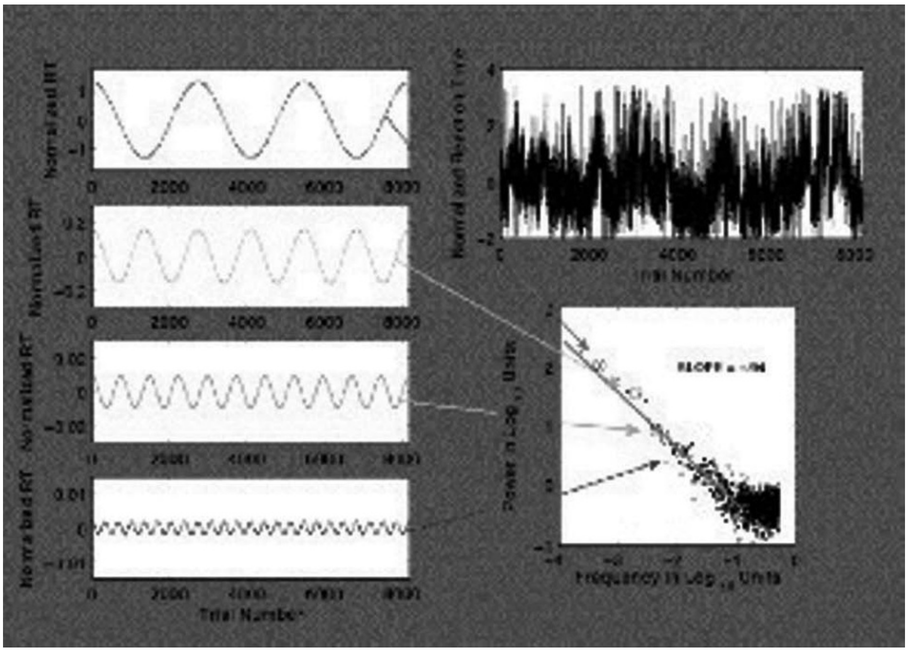


Figure 3. One person's response time data. Specific frequencies and amplitudes of change to approximate the rough graph of the data in the upper right of the figure, plus the outcome of the spectral anaysis below. The spectral slope = −.94, which is approximately $\alpha \approx 1$. Note that the Y-axes in the illustrations have been adjusted to make smaller amplitude sine waves visible.

this section is neither regular nor random. It constitutes a third kind of variability, one that is captured in scaling relations and that cannot be categorized by conventional approaches.

3.1 Pink Noise

The data series on the right in Figure 3 is decomposed into sine waves of different amplitudes, shown on the left. Slow changes in the data series are captured by low frequency, high-amplitude sine waves (top left), and fast changes in the data series are captured by high-frequency, low-amplitude waves (bottom left). Amplitude reflects the size of change S(f) between values across the data series and appears on the Y-axis of the power spectrum, plotted against the frequency (f) of changes at that size. The relation between size and frequency of change is the scaling relation estimated by the slope of the line in the spectral plot.

In the scaling relation illustrated in Figure 3, the size of change S(f) is inversely proportional to its frequency (f): $\mathrm{S(f)} = 1/\mathrm{f}^{\,\alpha} = \mathrm{f}^{\,-\alpha}$, with scaling exponent $\alpha \approx$

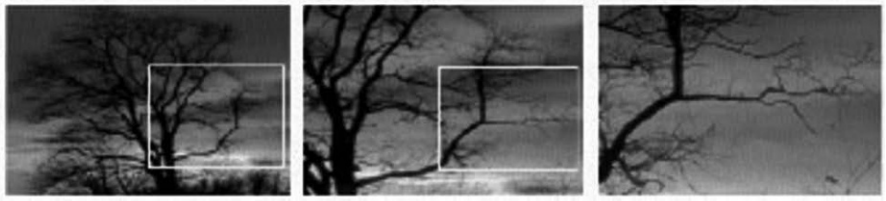


Figure 4. Fractal branching of a tree

1. It is this value of the scaling exponent that reflects the third kind of behavior. It is called *pink noise* because visible light with the same spectral slope has a pinkish cast from power concentrated in lower, redder frequencies. We use the phrase pink noise throughout, due to its accidental association with old-fashioned phrases about wellbeing, like *in the pink* and *pink of health*. However, depending on discipline, the phenomenon may be called *flicker noise, 1/f noise, 1/f scaling, intermittency, multiplicative noise, edge of chaos, fractal time, long-range correlations, red noise, self-affinity*, or something else. Similarly, there are many ways to portray this behavior in numerical and geometric analyses, each with its own vulnerabilities and caveats [Holden, 2005]. The many different names give credence to a core thesis of complexity science that common dynamical organizations will appear in systems of different material construction, even in living and nonliving matter.

What is the meaning of pink noise? Debates about this question have taken place in every discipline that has confronted complexity, including cognitive science [Chen *et al.*, 2001; Dale, 2008; Delignières *et al.*, 2008; Ding *et al.*, 2002; Diniz *et al.*, in press; Edelman, 2008; Farrell *et al.*, 2006; Gilden, 2001; Kello *et al.*, 2007; 2008; Kello and Van Orden, 2009; Newell and Slifkin, 1998; Riley and Turvey, 2002; Thornton and Gilden, 2005; Torre *et al.*, 2007; Torre and Wagenmakers, 2009; Wagenmakers *et al.*, 2004; 2005; Ward, 2002; Van Orden, 2008; Van Orden and Holden, 2002; Van Orden *et al.*, 2003; 2005; 1997]. The difficulty comes from the dual nature of pink noise, namely that it can appear as either a regular or an irregular phenomenon. The regularity is in the scaling relation, whether the basis of the scaling relations is a sine wave, square waves, V-waves, or irregularly spaced waves with different average frequencies. Yet pink noise appears irregular and unstructured in a data graph where it is an aperiodic waveform like random Gaussian noise or chaos. In truth it is neither regular nor random but a strongly nonlinear pattern that exists between these two extremes [Nicolis and Rouvas-Nicolis, 2007; Sporns, 2007; Tsonis, 2008].

The crux of pink noise is self-similar structure. Mathematical pink noise expresses formal self-similarity, and empirical pink noise expresses statistical self-similarity, not unlike the branching structure of a tree. From the bottom to the top of a tree, branches become thinner in diameter as they become more numerous. The same relation holds even when a window on the tree is decreased and

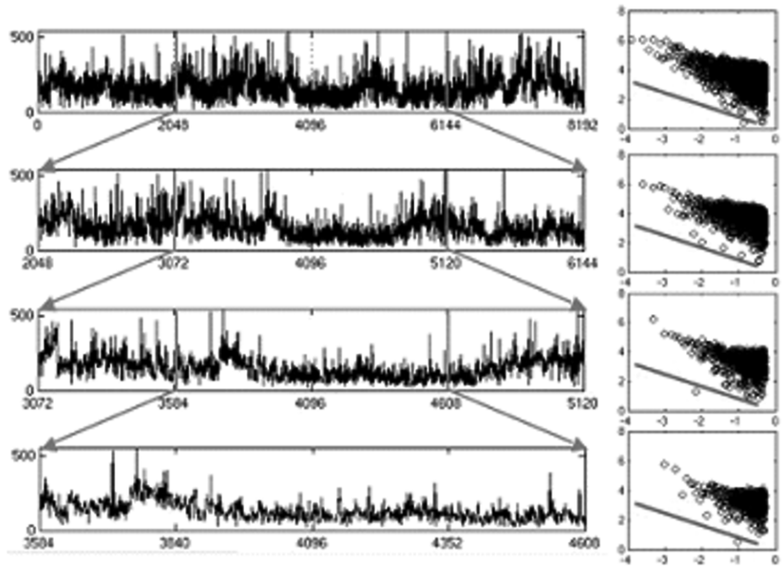


Figure 5. Trial-ordered series of reaction-time trials (left) and the resulting spectral plot (right). The top panel includes 8192 trials in the data series, while all other panels are a subset of the original data series. The first and last quarter are consecutively cut off to eventually yield a series with 1024 reaction times (bottom). The scaling relation remains very similar for each nested series.

one considers only a part of a tree, as in Figure 4. In particular, the relation between branch diameter, S(f), is inversely proportional with how often branches of that size occur (f). The resulting scaling exponent stays within a narrow range of values. Fractal structure makes it appear that every scale of measurement is stitched together with every other scale of measurement (e.g., the decreasing scale diameters of tree branches), in a nested pattern.

Comparable statistical self-similarity in fractal patterns can be seen in repeated measures of human performance, say when a participant produces simple reaction times, trial after trial (see Figure 5). A spectral plot across the entire data series of about 8000 reaction times results in pink noise. Importantly, when the data series is cropped at both ends, such that only half of the length of the original data series is considered for the spectral plot, a similar spectral slope is obtained. Again, when the shortened data series is cropped further, the slope stays within a small range. Just as for tree branches, each repeatedly measured value of brain or behavior appears stitched to every other in the fractal wave.

Finally, while pink noise has statistical self-similarity, variance within a data set does not stay the same. Note in Figure 5, as the data series get shorter, values in the spectral plots shrink along the Y-axis (magnitude of changes), as well as along the X-axis (frequency of changes).

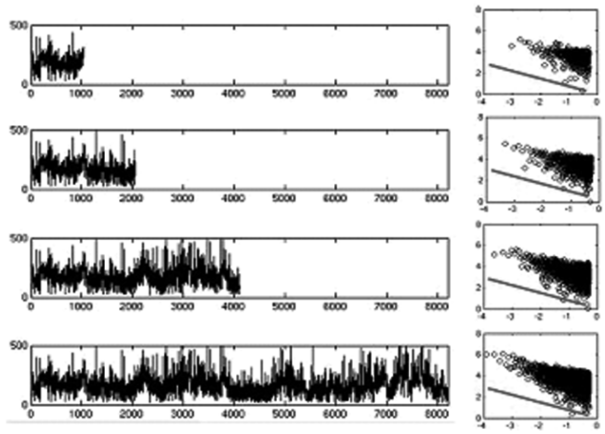


Figure 6. Trial-ordered series of reaction-time trials (left) and the resulting spectral plot (right). The top panel includes the first 1024 trials, while all other panels increase the length of the data series. The scaling relation remains virtually the same for each increasingly longer data series.

In other words, large rare oscillations disappear as the data series shrinks in length. The inverse is seen as more-and-more data are collected. Figure 6 portrays changes in the magnitude of variability as a data series gets longer. Variation grows by orders-of-magnitude as we gain access to rare but much larger amplitude changes in longer data series.

Conventional theories have difficulty accounting for the fact that more data equal more extreme variability. Conventional methods assume the opposite, namely that larger data sets yield more reliably stable estimates of average performance, meaning that error variance should not increase as more data points are collected. This is a false assumption as we have tried to illustrate. Longer data series include more extreme values, which destabilize the mean value of the data. No reliable mean value exists. This fact undermines the very foundation of conventional approaches, namely that variances can be ignored because data, at heart, are equal to their mean.

How does complexity science explain the nested fractal structure of pink-noise? The self-similarity of a mature tree, for example, is produced by the iterative growth processes of the living tree. An iterative process takes its present status, or output, as input in the next time step. In the tree example, the same growth processes of branching and thickening produce all the branches at all the different scales of the tree, and so the tree grows to resemble itself on large and small scales, and in the scaling relation between size and frequency of branches. In human behavior the present status of a person is input to embodied interaction-dominant dynamics, which produce the status in the next time step, and so behavior unfolds to resemble itself across time in the scaling relation between the size of change and its frequency.

Given these considerations, the following things appear true: Pink noise is neither regular nor random. Irregular, aperiodic data points are woven as an exotic fractal pattern. At present time, each repeated measurement of brains and behaviors appears to be sewn together in this fractal pattern. Within the pattern, every measured value is long-range correlated with every other value to span the experiment. Complexity science first recognized the aperiodic, fractal pattern as a third kind of behavior.

3.2 Soft Assembly of Performance Devices

Complexity science suggests that we view performance as a soft-assembled coordination or coupling between task and participant. Given that every task entails a different set of constraints, a new coordination should emerge every time we change tasks. This was indeed found in a simple key-pressing experiment in which adults had to press a key in response to a signal on a computer screen [Kello *et al.*, 2007]. Two measures were taken: (1) the time it took the participant to press the key upon seeing the signal (i.e., key-press response time), and (2) the time it took the participant to release the key to return to the ready position for the next trial (e.g., key-release response time). The two resulting data series (key press and key release) were subjected to spectral analyses, which revealed pink noise in each separate data series. Importantly, however, the two streams of data were uncorrelated. Although each measured key-press time was long-range correlated with every other key-press time, and each measured key-release time was correlated with every other key-release time, they were not correlated with each other.

A conventional explanation would posit two distinct and independent decision mechanisms, one for key-pressing and one for key-releasing. Of course positing new decision mechanisms for every dissociated effect quickly loses the elegance of parsimony, given that a myriad of trivial changes in task demands of very simple tasks yield similar dissociations, e.g. [Durgin and Sternberg, 2002]. A claim of separate decision mechanisms for separate effects also undermines generality, given that a key-press decision or a key-release decision has to be closely associated with the specifics of the task. Finally, it is not clear why a decision about pressing a key would require a different cognitive mechanism than deciding to release the key.

Complexity, on the other hand, explicitly predicts such dissociations, because performance is the becoming of a performance device entrained to the specific constraints of task demands. In some sense, task couplings create new ‘devices’ of the participant with even subtle changes in task demands. Pressing down a key entails different constraints than releasing the same key and, while the two movements are interleaved in time, their respective sources of constraint may very well change independently. To test these claims more directly, another key-press experiment was conducted, with one crucial manipulation: Instead of a predictable signal about which key to press, signals were alternated unsystematically which introduced uncertainty about which key to press until the signal appeared. Again, two data series were collected, one for the time it took a participant to press a key and one for the time it took to release the key. The results showed that uncertainty about which key to press affected the key-press data series, but not the key-release data series. More specifically, while the key-release data series retained their pink noise pattern, observed before, the key-press data series were de-correlated by the injected uncertainty and appeared closer to random noise [Kello *et al.*, 2007].

Task coupling gives a simple and sensible account of the key-press response data. Unpredictable signals injected uncertainty as an unsystematic perturbation of the entrainment to each trial's signal to respond by pressing the key. The unsystematic perturbation resulted in less systematic coupling which de-correlated the otherwise long-range correlated data series. Key release durations were unaffected because the coupling of the key-release response was the same across all trials. The participant was always at a key contact point, at the bottom of a key-press, before the key-release response was initiated, irrespective of which key was pushed down. At the bottom of a key-press response no uncertainty exists about which key-press to key-release.

Taken together, these key-press results support the idea that the body coordinates itself into temporary performance devices to fit the specifics of the tasks. The apparent devices are soft-assembled coordinative structures. Even when tasks differ merely in uncertainty about which key to press or the direction of the finger's motion in key pressing versus key releasing, the body will appear to create specialized devices to accommodate the different demands. Devices do not refer to permanent mental functions or components, but instead exist only so long as the specific task demands are present and performance continues.

3.3 Attraction to Complexity

As discussed in Section 2, the quality of task-person coupling reflects the extent to which the effectivities of the participant (embodied constraints) match the affordances of the task (embedding constraints). Pink noise might reflect such an ideal match between embodied and embedding constraints. Consistent with this prediction, pink noise is the central tendency of variability in skilled healthy behavior [Kello *et al.*, 2008]. Participants were asked to say the same word ('bucket') over and over. Each instance of the spoken word was then parsed identically into dozens of frequency bins and the amplitude of each frequency-bin was tracked across all spoken instances of the word 'bucket.' This resulted in dozens of separate data-series, and each data-series yielded a spectral exponent. Aggregating all the estimated scaling exponents in a histogram yielded normal Gaussian distribution around the scaling exponent of 1. In other words, the coupling of healthy skilled participants to a repetitive speech task reveals evidence of attraction to pink noise.

If pink noise reflects an optimal coupling for performance, then what are the less-than-optimal types of coupling that the system is moving away from? As we mentioned earlier, pink noise lies between regular and random behavior. Still, how do regular or random behavior appear in this complex system? Little would be gained by positing hard-assembled causes of regular behavior plus different causes of random behavior. Instead, a single control parameter may serve to produce regular and random behavior, as well as behavior in between. The critical value of the parameter should yield pink noise, bracketed by attraction to over-random and over-regular behaviors. So what is the ratio of this control parameter?

Clues came from failed attempts in physics to corroborate self-organized criticality. The designated model system consisted of grains of sand, dropped one at a time to build a pile in which eventually, a dropped grain of sand triggers an avalanche. Volume and time between avalanches were measured repeatedly but, contrary to expectations, sand pile avalanches never became sufficiently large to fill out a scaling relation between size S(f) and frequency (f). They appeared instead to be over-random inertia-driven avalanches, exclusively irregular avalanche behavior. (For a review see [Jensen, 1998].) Self-organized criticality was found only after grains of sand were replaced with kernels of rice [Frette *et al.*, 1996]. The rice kernels varied in their aspect ratio of kernel length to kernel width. Lower aspect-ratio kernels behaved like sand, while rice with higher aspect ratios yielded critical behavior. Higher aspect ratios imply greater surface area to create more friction between kernels, sufficient to build small piles of rice, distributed throughout the larger pile, at or near their threshold for toppling. With so much rice poised to topple, the rice pile could produce the large rare avalanches necessary to fill out an inverse scaling relation between size of avalanche S(f) and frequency (f) of avalanches of that size.

The control parameter of success and failure is the ratio of inertia to friction. The inertia-numerator is a source of over-random behavior, and the friction-denominator is a source of over-regular behavior, cf. [Kinouchi and Copelli, 2006]. Their ratio is the external control parameter of avalanche behavior. The specific ratio is anticipated in the ratio of inertia to viscosity of Reynolds numbers in fluid dynamics and heat transfer [Iberall, 1970], so both ratios are nominated as external control parameters of complexity.

Piles with 'too much' friction or 'too little' inertia are too coherent and rule-bound, like a mud pile for instance. Piles with 'too little' friction or 'too much' inertia are too random, like a sand pile. Critical behavior is found in the balance between regular and random and the same kind of control parameter can be envisioned for the coupling of task and participant. A control parameter that emphasizes over-regular tendencies yields the over-regular behavior that brackets pink noise; but if the parameter is changed to emphasize over-random tendencies, the coupling between task and person yields an unsystematic relation between size and frequency of variation across repeatedly measured behavior, the over-random bracket.

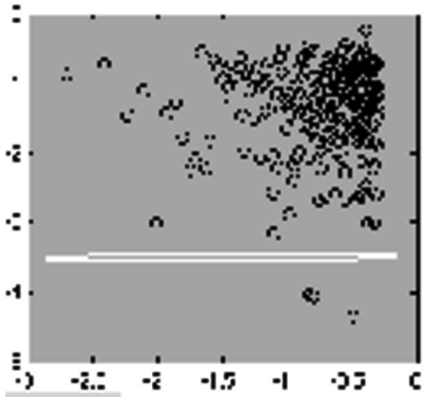


Figure 7. Spectral portrait of a random noise data series. The white line illustrates the slope of a regression line to fit to the data: the slope of zero indicates the unsystematic realtion between power and frequency.

Loss of structure, due to the over-random tendencies, is indicated in data by a white noise scaling exponent. The spectral portrait of behavior dominated by unsystematic sources of variation is illustrated in Figure 7, mapping out again a relation between size of change S(f) and frequency of change (f). Size is on the Y-axis and frequency on the X-axis and their relation is the flat slope of the white line in the figure. White noise is disorderly, irregular, random noise. Changes of every size are equally likely, as though sizes and frequencies were shuffled and dealt like cards into meaningless pairs. Any particular magnitude of variation is equally likely to be paired with any particular frequency of variation. This is represented in Figure 7 by the flat white line with a slope of zero ($\alpha \approx 0$), the spectral slope of white noise.

The other bracket must be over-regular behavior. However, even the most regularly structured behavior in a living system will appear somewhat irregular, as illustrated in the data graph of Figure 8, from an over-regular heartbeat of a person with congestive heart disease. A spectral plot of the data series resembles *brown noise*, irregular behavior that is dominated by changes on slow time scales. The spectral slope of size S(f) against frequency (f) is shown in the Figure, very close to an idealized $\alpha \approx 2$. This slope is steeper than the spectral slope of pink noise due to over-regular oscillations in behavior. The steep slope of the line in the spectral plot suggests that large over-regular changes will occur, and that still larger changes quickly become improbable. It emphasizes high-amplitude and low frequency in a relatively narrower range. All three categories of noise — *white*, *pink*, and *brown* — appear together in Figure 9, each with their characteristic ideal slopes.

Self-organized criticality predicts that performance will be drawn toward pink noise and attraction toward pink noise and away from white noise was observed as adults gained practice with a Fitt's tracing task [Wijnants, *et al.*, 2009]. Adult participants produced pinker data after practice. Participants were asked to trace between two dots on an electronic tablet as the trace-time from dot-to-dot was measured. After several blocks of practice, 5500 trace-trials total, the central tendency of the spectral plot had moved to $\alpha \approx 1$ of pink noise. The results are portrayed in Figure 10 to illustrate the statistical character of the phenomenon in developed healthy adults.

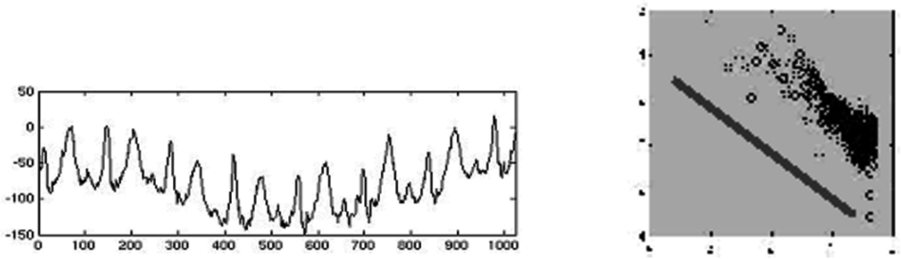


Figure 8. Heart beat data of a patient with congestive heart failure in the graph on the left and a spectral portrait of this data series on the right. The brown line illustrates the slope of a regression line to fit to the data: the slope is close to -2, indicating a scaling relation close to that of brown noise with $\alpha = 2$.

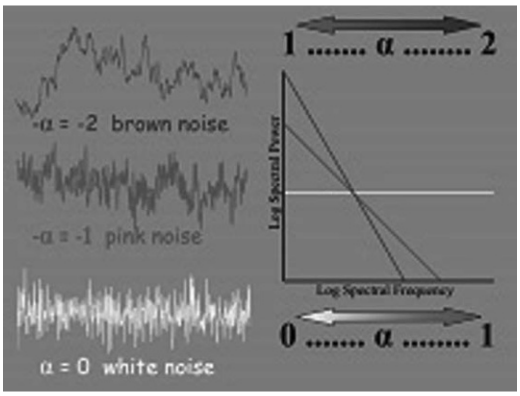


Figure 9. Summary characteristics of brown, pink and white noises. Data series appear on the left (together with their characteristic alpha values), and spectral slopes appear on the right.

In development, performance is drawn toward pink noise from two directions of change. One direction of change was observed in development of gait in walking and the other in cognitive performance of time estimation. In the cognitive task, preschool children and adults were asked to estimate a short time interval over-and-over, pressing a button each time it had passed. Spectral slopes of the variation in

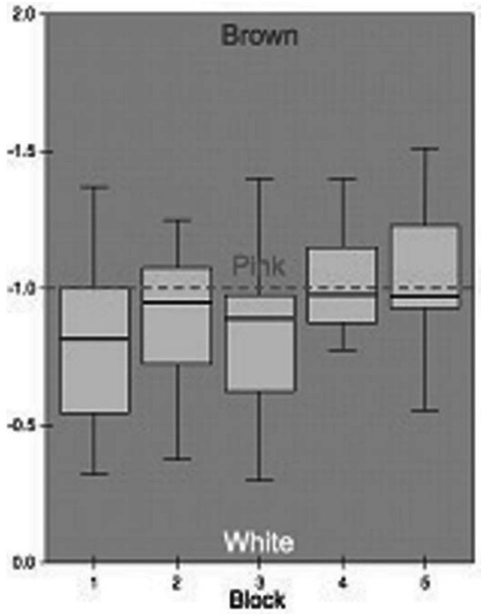


Figure 10. Change in spectral slopes of data species across five consecutive blocks of practice in a Fitt's tracing task.

their estimates showed a developmental progression toward pink noise, and away from white noise, across age. The attraction was clearly visible in dynamics, while the average estimates only marginally distinguished the youngest children from all other children and adults [Kloos *et al.*, 2009].

In the walking task children and adults walked on a treadmill while stride interval times were measured [Hausdorff *et al.*, 1999]. Like in the time estimation task, spectral slopes of stride variation showed an attraction toward pink noise as age increased, but this time slopes moved away from brown noise. Spectral exponents of 4 and 5 year-olds' gaits were heavily skewed toward the $\alpha \approx 2$ of brown noise, while exponents for adults were distributed narrowly and closer to pink noise (on the white side of pink noise).

What accounts for the changes in development, and the differences between the two tasks? The plausible hypothesis for development overall is that embodied constraints and sensitivity to embedding constraints are not optimally tuned for a child (or for an unskilled adult). While both children and adults could do both tasks, the task-child system was not coordinated optimally. Components that are not well coordinated show more independent variation, which perturbs the task-system coupling of repeated measurement. Across development, children accrue sufficient constraints to better coordinate their bodies with the cognitive task. They better accommodate arbitrary and idiosyncratic task constraints, and they can better sustain constraints of intentions that follow from a scientist's instructions.

A plausible hypothesis for task differences is simply different task demands. In walking on the treadmill, the task-child system shows evidence of over-rigid control. Not unexpectedly, when learning to walk children initially lock out degrees of freedom in legs and body to avoid falling. This over-rigid control yields over-regular behavior and brown noise variation in measured gait. With practice and development the child comes to embody flexible constraints among legs and body to negotiate the varieties of terrain in the world. Fluid constraints allow less rigid control as the body flexibly adjusts degrees of freedom to negotiate the varieties of terrain with smooth gaits.

Taken together, both practice and development reveal attraction toward criticality as pink noise. These patterns provide evidence that critical states are self-organized, meaning that living systems are drawn toward states of flexible coupling in which multiple propensities for action are available. They furthermore mark the endpoint of ideal coordination between body and environment. The next issue we explore then pertains to how the pattern changes as coordination deteriorates.

3.4 Departure from Complexity

Pink noise is most prominent in simple tasks that repeat identical trials, e.g. [Gilden, 1997]. The pattern changes however as tasks get more complicated. For example, the spectral slope is whitened when trial response decisions are made more difficult [Correll, 2008; Clayton and Frey, 1997; Kello *et al.*, 2007; Ward 2002]. Likewise, in a dual task experiment, walking on a treadmill while repeatedly estimating short time intervals whitened the spectral slope of time estimation [Kiefer *et al.*, 2009]. Fractal patterns of gait in the dual task produced pink noise, probably because walking has greater priority than time estimation. Both tasks produced pink noise as single tasks and the change away from pink noise was only found in the dual-task scenario, and only for the time-estimation task of lower priority.

In principle, one could also imagine a departure from pink noise in the direction of brown noise, as task constraints increased or participants adopted a strategy of over-rigid control. This was the case for toddlers, for example, who locked down degrees of freedom needed for flexible control of gait. Provisional evidence was found in data from a driving-simulator in which lane positions are over-constraining (Geoff Hollis, personal communication, October 6, 2008). Car position data resembled brown noise, but no condition was included that produced pinker data for comparison.

Similar departures from complexity are found in advanced aging and dynamical diseases [Glass and Mackey, 1988]. With advanced age, posture and gait show departure toward white noise in spectral plots, while heartbeat, body temperature, and neural activity (resting fMRI) show a departure toward brown noise. Figure

11 summarizes age related changes. In atrial fibrillation, a rare form of heart disease, heartbeats depart from pink noise in the direction of white noise [West, 2006]. In Huntington's disease, gait departs toward white noise [Hausdorff *et al.*, 1997], and in Parkinson's disease, gait, arm movements, and speech all depart in the direction of brown noise. What's more the degree of departure from pink noise toward brown noise reliably predicts the severity of other Parkinson's symptoms [Pan *et al.*, 2007], and the degree of departure toward white noise predicts severity of other symptoms in Huntington's disease [Hausdorff *et al.*, 1997].

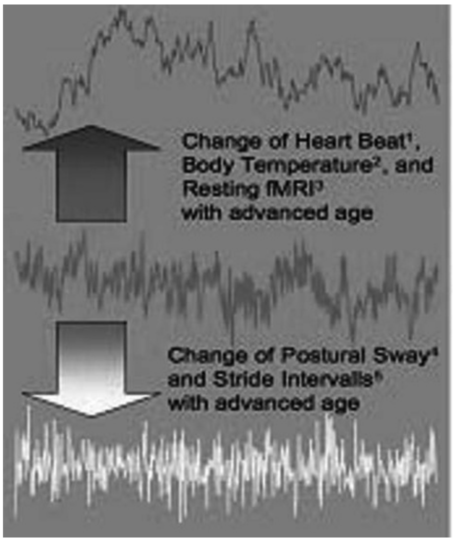


Figure 11. Departures from complexity due to advanced age. References: [1]Beckers *et al.*, 2006; [2]Varela *et al.*, 2003; [3]Wink *et al.*, 2006; [4]Duarte and Sternag, 2006; Lin *et al.*, 2008; Norris *et al.*, 2006; Thurner *et al.*, 2002; [5]Hausdorff *et al.*, 1997.

Why does performance deviate from complexity and pink noise in much the same way for task changes, aging, and dynamical diseases? Deviations toward white noise suggest loss of structure in dynamics or sources of unsystematic perturbations to the coupling of task and person, or between organ systems. Changes from pink to brown noise as health deteriorates suggest loss of flexibility in dynamics or sources of over-rigid control. Parkinson's is typified by a loss of flexibility and over-regular movements: Patients can no longer produce smooth kinematics in response to rapid changes in the environment and they have difficulty initiating and controlling motion. Figure 12 organizes Parkinson's symptoms as they might appear in a complexity account and we discuss Parkinson's symptoms next in more detail.

Parkinson's symptoms originate in damage to areas of the brain that produce the neurotransmitter dopamine. Indeed, a conventional causal story might propose that the reduction in dopamine production disrupts a causal chain from stimulus to response, or intention to action. In line with this reasoning, dopamine has been marketed as the causal basis of the mind, the brain-within-the-brain, so to speak [Previc, 1999]. However, most prominent Parkinson's symptoms, including reduced dopamine, have not yet found their place in a causal account. How do gradual changes in dopamine availability produce qualitative changes in perception, action and cognition? Why does Parkinson's erode cognition along with mobility; and why do cognitive symptoms appear idiopathic? Why are fine-grain capacities most vulnerable early in Parkinson's? Basic neural conduction among modules is intact in Parkinson's, and the conduction rate across neurons is plenty fast to move fast-changing information through the nervous system. Why then do early Parkinson's symptoms include disruptions in fast-changing perception-action cycles? The complexity explanation is subtle, speculative, but compelling. The emphasis shifts from a faulty isolated component (such as a faulty dopamine-uptake system) to faulty coupling among components. It is the erosion of system capacities to coordinate mind, body and environment that lead to loss of flexibility in behavior, e.g. [Edwards and Beuter, 1999; Goldberger *et al.*, 2002a; 2002b].

Dopamine bridges synaptic gaps between neurons to perpetuate electrochemical waves of action potentials, like any neurotransmitter. Action potentials create feedback loops of neuronal activity that self-organize into larger traveling waves. Traveling waves are an observable realization of emergent constraints in motor coordination, perception, and cognition [Davia, 2005; Freeman; 2000; Hollis *et al.*, 2009; Kelso, 1995]. Damage that reduces dopamine in the brain reduces the capacity for traveling waves to coordinate, which in turn affects cognitive functions, motor coordination, and the dynamics of physiology. Parkinson's is systematically progressive. The first constraints to erode are those that change on the fastest timescales — they are necessary for detecting subtle changes in emotional tone or social alliances, for making fine-grained perceptual distinctions, and for initiating sudden or rapid movements. In other words, Parkinson's first destabilizes the capacity to rapidly organize or reorganize perception and action.

Erosion of constraints on fast timescales explains the unwelcome palsy in Parkinson's. In a sense, the palsy originates in less refined, less well-coordinated control. Palsy is a kind of overshoot phenomenon, like oscillations in room temperature around a thermostat setting. The relatively preserved capacities for constraint that change on intermediate timescales lack the finer-grain, faster-changing, automatic dithering control of constraints from faster timescales that insure smooth and precise movements. Parkinson's eventually erodes intermediate and slow timescale capacities as well, such that late-stage Parkinson's sufferers appear to express frozen postures and gaits, although in truth they are moving on the very slow timescales of the last remaining capacities to constrain and change behavior.

The protracted unraveling of constraints from faster to slower timescales erodes capacities to coordinate brain, body and world, including the coordination of cognitive capacities. The subsequent deficit or lost cognitive capacities appear to be idiopathic symptoms because cognition expresses the idiosyncratic contingencies of a patient's mental and physical history. Idiosyncratic histories of education,

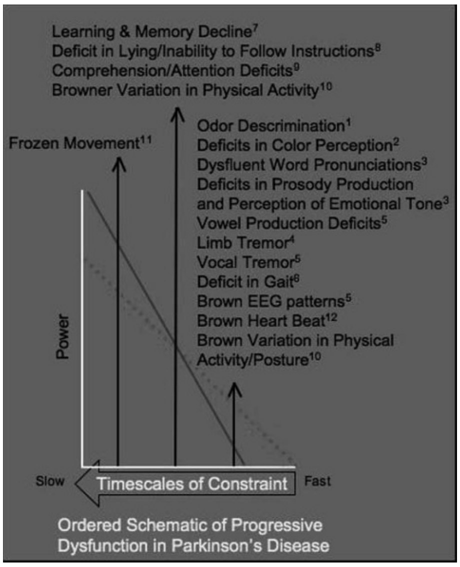


Figure 12. Approximate progression of Parkinson's disease, estimated from the cited descriptions of patients, plus pink and brown noise scaling relations as a backdrop. Parkinson's first erodes constraints changing on fastest timescales and then intermediate and slow changing constraints. Eventually sufferers appear frozen in time although they continue to move on the very slow timescales of very slowly changing constraints. References: [1]Double *et al.*, 2003; [2]Diederich *et al.*, 2002; [3]Ariatti *et al.*, 2008, Lloyed, 1999; Goerman *et al.*, 2008; [4]Aly *et al.*, 2007; Jankovic *et al.*, 1999; [5]Hertrich *et al.*, 1997; Zhang and Jiang, 2008; [6]Blin *et al.*, 1998; Hausdorff *et al.*, 1995; [7]Allain, 1995; Howard and Binks, 2000; Price and Shin, 2009; [8]Abe *et al.*, 2009; [9]Peron *et al.*, 2007; Grosman *et al.*, 2000; [10]Pan *et al.*, 2007; Schmit *et al.*, 2006; [11]Hausdorff *et al.*, 2003; [12]Haapanierni *et al.*, 2001

language, work, hobbies, travel, and health shaped the idiosyncratic strengths and stabilities of cognitive capacities well before the Parkinson's began. They in turn shape the expressed cognitive deficits seen in an individual patient. Almost all healthy people walk and manipulate things with their hands — and indeed most Parkinson's patients show similar deficits in gait and hand-eye coordination. In some cases, as capacities for constraint and change erode, however, control parameters of coordination cross their critical values. Consequently patients express additional idiopathy as idiosyncratic changes to qualitatively different functioning, qualitative reorganizations of mind and body into tragically dysfunctional modes.

In sum, the accumulated evidence nominates pink noise as the signature of complexity — its third kind of behavior — as in variability that is neither too regular nor too random. Pink noise reflects an optimal flexible coordination that a system is drawn toward as it develops or practices. Such optimal coordination can be obtained in accrued constraints (to move performance from white toward pink noise) or by loosening up over-rigid constraints (to move performance from brown toward pink noise). Similarly, pink noise reflects an ideal from which a system departs as coordination deteriorates.

4 CHALLENGES AND OPPORTUNITIES

In this final section, we discuss challenges and opportunities that complexity presents. They include issues pertaining to the interpretation of scaling exponents, the naturalization of intentionality in principles that apply to nature generally, piecewise determinism, and emergent coordination among multiple actors. We discuss each in turn.

4.1 The Scaling Exponent Dilemma

An ideal coordination between task and person reveals itself in pink noise, a fractal pattern with a scaling exponent of about 1. A reasonable conclusion then would be that any scaling exponent reliably above or below 1 reflects a less-than-ideal coordination. More specifically, a scaling exponent closer to zero should reveal a coordination that is over-random, and a scaling exponent closer to 2 should reveal a coordination that is over-regular. Consistent with this interpretation, pink noise is characteristic of healthy adults performing a comfortable task, while white noise was found when task difficulty was increased, expertise of participants was reduced, or participants suffered dynamical diseases.

However, the simple mapping of scaling exponent to kind of coordination does not fit with all the evidence. Take for example continuation tapping, a task in which participants tap from memory after a metronome is turned off. Continuation tapping yields clear pink noise behavioral signals. However, a task in which participants tap in synch to the beat of a metronome produces whiter signals than continuation tapping [Chen *et al.*, 2001]. Why so?

Intuitively entrainment in synch with a metronome should reveal over-regular coordination, because the beats of the metronome are so regular. Continuation tapping, without the metronome, should then yield less regular coordination by the same intuition. Going from tapping with a metronome to tapping without should decrease the scaling exponent (reflecting change from over-regular to less regular). Yet, this is not what was found: Variability during entrainment to a metronome yields whiter scaling exponents farther from pink noise.

Another example comes from a time-estimation task, much like continuation tapping, in which participants were either provided with accuracy feedback or not [Nikita Kuznetsov, personal communication, August 23, 2009]. Accuracy feedback is another source of external control — like the entraining metronome beat — and should therefore promote over-regular structure in performance variability. However, while time estimation without feedback yielded a pink-noise signal, trial by trial accuracy feedback whitened the signal. Despite external sources for over-regular control, the structure of variability in both examples showed over-random tendencies. How can these findings be reconciled with the idea that scaling exponents predict the type of coupling between person and task?

The dilemma stems in part from the duality of pink noise, the fact that pink noise is simultaneously regular (it obeys an orderly scaling relation) and irregular (it is aperiodic nonetheless). In every estimate of pink noise, order and disorder trade off in the repeated measurements. A scaling exponent by itself is therefore inherently ambiguous. To understand particular tradeoffs of order and disorder in performance, it is necessary to put the system in motion to examine changes in scaling exponents, rather than a static value. Nonetheless, these facts alone do not remedy the challenge to understand the whiter signals that results from entrainment and feedback.

To address the challenge, we revisit the numerator and denominator of the control parameter we have relied on until now. Recall that the numerator, on the one hand, comes from affordances delimited by embedding constraints of the environment or task. Affordances define the degrees of freedom available to the actor within the task. The denominator, on the other hand, comes from effectivities, which determine which degrees of freedom can be brought under control. In task performance, the degrees of freedom required in a successful performance must correspond to controllable degrees of freedom of the person's effectivities.

As for the control parameter in the example of entrainment, external sources of constraint increase when the metronome is running. This changes both the numerator and denominator of the control parameter. The numerator-source of over-random behavior is reduced as the available degrees of freedom are reduced, compared to no-metronome conditions. An environment that supplies more constraint affords fewer degrees of freedom. At the same time, however, the metronome minimizes the denominator, the source of over-regular tendencies. Entrainment to the metronome minimizes degrees of freedom that must be controlled for successful continuation tapping.

The minimum value of the denominator means minimal sources of over-regular variation. Also, the previously person-controlled degrees of freedom, for task success, become available during entrainment as uncontrolled degrees of freedom, adding sources of over-random behavior and increasing the numerator. Altogether, these changes favor over-random sources of variability. A similar argument can be made for trial feedback. Accuracy feedback supplies constraints that reduce available degrees of freedom and therefore reduce requirements for successful performance from the effectivity denominator. This releases previously person-controlled degrees of freedom. Constraints when present imply fewer degrees of freedom, so constraints when absent imply greater degrees of freedom. Whiter behavioral signals result.

However, consider another piece of evidence from the posture of elite ballet dancers [Schmit *et al.*, 2005]: A dancer's torso remains upright, while she is in motion, over her body's center of balance. This allows the visibly unique gait in which a dancer can appear to glide across the stage. It controls for the ordinary tendency of torsos to move past the body's tipping and falling point in each step. The over-trained posture ingrains constraints and controls degrees of freedom in posture. These constraints count among the effectivities the dancer brings to the dance. Effectivities are the sources of over-regular variation in measured performance. Nonetheless, a whiter signal is observed in dancers' posture compared to posture of ordinary adults or different elite athletes. The control parameter that accounted for whiter scaling exponents in entrainment and feedback fails to explain the dancer's whiter posture. To address this challenge, we must address a second challenge, namely that of voluntary control and intentions.

4.2 Naturalizing Intentionality

As discussed in Section 1, intentionality has constituted a major stumbling block for conventional approaches. How then does complexity science solve the problem of intentionality? We have proposed that intentions affect behavior as constraints, not causes. Intentions as constraints are temporary dynamical structures, soft assembled from interdependent components to function in control parameters to create critical states [Riley and Turvey, 2001; Van Orden and Holden, 2002]. Constraints circumvent dilemmas that arose from viewing intentions as causes [Juarrero, 1999]. Constraints are therefore no less natural or no more magical than causes, or the convectional cells of the atmosphere that change the weather. Thus the complexity account makes progress toward naturalizing intentionality.

Intentions are of the same nature as other constraints and should have the same consequences. In nature, constraints dampen vibration and oscillation for example. Intentions also dampen oscillations in the voluntary actions of Parkinson's sufferers. And the intention to move can eliminate palsy during movements early in the disease, and partly dampen it in later stages. For example, in Parkinson's, the palsy appears in unintentional involuntary movement, but intentional voluntary movement dampens the palsy, so long as voluntary movement exists.

The challenge from ballet dancers' posture still remains, however. We remain stuck with a control parameter that predicts pinker or browner noise in dancers' posture and elite dancers whose posture reveals whiter noise. Otherwise, this control parameter predicted the direction of change for every task and performer we have reviewed, within a plausible account of intentionality. Yet we have not successfully naturalized intentionality, due to contradictory evidence. But to meet the challenge, we can look for what is common across the three exceptions here considered. What is it that is common to: (1) entrainment, (2) accuracy feedback, and (3) over-trained posture? Each example includes a prominent source of constraint, and each source of constraint functions to reduce or minimize the demands for voluntary control in task coupling.

The dancer requires less voluntary control to sustain erect posture. She has over-trained posture to stand upright, even balanced on a force plate that measures variation in posture. In contrast, a Parkinson's sufferer exhibits over-rigid control, to not fall down. Over-rigid control in Parkinson's shows up as large deviations around the center of pressure of the force plate [Schmit *et al.*, 2006]. Parkinson's patients produce a browner pattern of variation in posture, compared to healthy control participants who produce pinker variation. Thus reducing the need for voluntary control is associated with whiter signals, and exaggerated purposeful control with browner signals. This pattern motivates a revision to the control parameter. The key evidence motivating a revised control parameter is that reduced demands for voluntary control in the coupling between task and person yield performance dynamics that depart from pink noise toward white noise. If this fact proves reliable, then reduced voluntary control is reliably distinguished in empirical contrasts by whiter noise, all other things equal. We may combine affordances and effectivities in the numerator to define available degrees of freedom, which has been the role of the numerator all along. The numerator now equals the difference between degrees of freedom, afforded, versus degrees of freedom that can be controlled (reduced) by effectivities, as embodied constraints. Effectivities have been moved from the denominator to the numerator. What then is the denominator? We suggest that the denominator is volition, itself. Volition picks up the slack, so to speak, the remaining degrees of freedom, and reduces white noise in performance variation. The proposal presents a historical opportunity. Since Freud, the distinction has been made between consciously controlled, strategic, voluntary behavior versus automatic, unconscious, involuntary behavior. However, no empirical evidence for reduced voluntary control has yet stood the test of time [Fearing, 1970; Goldstein, 1939; Van Orden and Holden, 2002]. Each source of evidence, in its turn, has been found to be equivocal [Bauer and Besner, 1997; Besner and Stolz, 1999a; 1999b; Besner *et al.*, 1997; Kanne *et al.*, 1998; Pansky and Algom, 1999; 2002; Prochazka *et al.*, 2000; Tzelgov, 1997]. Presently, the distinction is supported by intuition alone but if whiter noise in task coupling (departing from pink) is a reliable consequence of reduced voluntary control, then we have naturalized intentionality.

Other challenges remain, however. Intentions satisfy needs and goals of the actor, and in this service, shape critical states that include propensities for serviceable actions. If purposeful behaviors originate in critical states, then it should be possible to connect more dots in analogies with thermodynamic systems, at least that is the challenge. Enacted behavior creates information and reduces the entropy of the critical state. In comparison, physical critical states and phase transitions concern energy and entropy. Thermodynamics creates structure and constrains molecules to better transport energy to more efficiently produce entropy. How does creation of information in behavior coincide, or does it? So far, the answer to this question has not been forthcoming [Nicolis and Nicolis, 2007]. Perhaps soft assembly of action also more effectively dissipates energy, compared to relatively hard-wired behaviors. If so then less probable, creative, and rare actions may most effectively dissipate energy — or maybe we have it exactly backwards. Or perhaps the debt to entropy is only fully paid by social systems or ecosystems and not by individuals alone, cf. [Ulanowicz, 2000].

Another challenge also stems from critical states of propensities to act. Propensities, in some fashion, anticipate the behavior that will be enacted. Critical states concern the future because they contain anticipated propensities-to-act. However we have not yet discussed a mechanism to acquire information about anticipated activities. To meet this challenge is important as, arguably, anticipation is the quintessential cognitive activity [Changizi *et al.*, 2008; Jacob, 1982; Jordan, 2008; Jordan and Hunsinger, 2008].

The opportunity to meet this challenge comes from a recent simulation of anticipation [Stepp and Turvey, 2009]. The simulation used time-delayed coupling. Imagine an environment *leader* and an organism *follower*. The organism is coupled by a time-delay to the environment. Present states of the environment are coupled to past states of the organism. The coupling term is the difference between the *current* state of the environment minus the *previous* time-delayed state of the organism. In the model, the time-delayed organism comes to minimize the difference between its current state and possible future states of the environment. In minimizing the difference, the organism successfully entrains to future environments, see also [Dubois, 2003].

The simulation also shows Pavlovian learning, perhaps the most well known example of anticipation. Imagine now the food served to Pavlov's dog, the current state of the *leader* environment, which co-occurs in delay-coupling with a past state, a sounded bell. The sounded bell's relation to the food is captured as a regularity by which to anticipate the future. The drooling dog's unconditioned-response thus becomes a means to better contend with uncertainty, to anticipate the arrival of food [Stepp and Turvey, 2009]. Indeed, the delay-coupling model shares formal parallels with a contemporary model of conditioned regulation [Dworkin, 1993].

The model predicts, necessarily, that anticipation is based on statistical regularities between past and future. Pavlov's sounded bell preceding dog food might have been 100% reliable, but most future events are much less certain and can only be known in their statistical broad strokes. Low-frequency large-magnitude oscillations in pink or brown noise are examples of broad statistical regularities. Similar regularities occur in chaos which was used to corroborate the prediction. Anticipatory tapping of college student participants successfully distinguished long-range statistical structures of different chaotic signals in metronome beats [Stephen *et al.*, 2008]. The simulated model plus its empirical support suggest a near term opportunity to integrate anticipation of the future with anticipatory propensities to act.

4.3 Piecewise Determinism

We have relied throughout on a control parameter of task coupling that takes on different values based on task and participant. If this parameter changes its values midstream, so to speak, during performance of the task, it may also explain piecewise determinism. Piecewise determinism is behavior that changes abruptly and discontinuously [Riley and Turvey, 2002]. For example, a task coupling may change with lapsed attention or vigilance, a change in strategy, or some other reorganizing change. Task performance may even change contingent on where the previous trial's performance leaves the performer, regarding the next trial's task demands.

Piecewise determinism illustrates a challenge for measurement that stems from blind spots inherent in spectral analyses and other linear methods to estimate scaling exponents. Spectral analyses assume that data series express idealized dynamics, smoothly continuous over time. The assumptions are called Lipschitz conditions of equations that are everywhere differentiable [Strogatz, 1994; Zak, 1993]. Analyses that assume Lipschitz conditions are blind to piecewise determinism. Nevertheless, these violations of Lipschitz conditions have empirical consequences, which are realized in both quantum mechanics and are also mundane features of behavior [Zbilut, 2004]. In a key-press experiment, for example, the finger approaches a singular solution, the contact point, in finite time (response time). 'Singular solutions in finite time' are a predicted 'pathology' of systems that violate Lipschitz conditions [Strogatz, 1994]. Across trials, successive 'intersecting singular solutions' occur when the finger presses the same key repeatedly, another pathology confirmed.

Violations of Lipschitz conditions are found in system behaviors that start and stop and repeat themselves in piecewise determinism. The fact that piecewise behaviors have explanations in quantum mechanics presents an opportunity to broach piecewise determinism in human behavior, cf. [Giuliani, *et al.*, 1996]. Formal analogies can bootstrap studies of piecewise human behavior, an opportunity also recommended by a growing menagerie of recognizably quantum-like phenomena in cognitive science [Atmanspacher, *et al.*, 2008; Atmanspacher, *et al.*, 2006; Bruza and Cole, 2005; Bruza, *et al.*, 2009; Kelso and Tognoli, 2007; Nelson and McEvoy, 2007; Turvey and Moreno, 2006; Van Orden *et al.*, 2010].

Piecewise determinism has also been discovered in task coupling data. Reanalyses of data series from Wagenmakers *et al.* [2004] and Van Orden *et al.* [2003] found piecewise determinism, where none had been reported previously [Ihlen and Vereijken, in press]. In these data, the evidence for piecewise-determinism is like abrupt changes in spectral slopes and scaling exponents during data collection. The abrupt changes are also equivalent to abrupt changes in fractal dimension. Thus piecewise data series divide into pieces with different fractal dimensions. Data with multiple fractal dimensions are called multifractals and it was advances in multifractal analysis that made possible the detection of piecewise-determinism. Previous multifractal methods required much more data than these series contained [Van Orden *et al.*, 2003].

Contemporary wavelet methods are multifractal analyses to analyze shorter data series. Wavelet methods detect abrupt local changes in fractal dimension using a moving cone of wavelets; the tip of the cone hits each data point in its turn to examine local task coupling. Wavelet analysis yields a second measured aspect or dimension of data series along with a scaling exponent. Task coupling varies along two outcome measures: a center value and a spectrum of values around the center. The center value is approximately equal to the value got from a monofractal analysis, so center values can be expected to corroborate changes toward pink noise or departing from pink noise. In addition however the width of the spectrum, around the center value, varies from wide, to narrow, to virtually no dispersion at all, and the width varies independently of the center value and gives independent information about task coupling (Espen Ihlen & Beatrix Vereijken, personal communication, August 12, 2009).

The extra outcome measure allows that different task-person couplings may be more-or-less multifractal along with being more-or-less pink, a kind of more-or-less *piecewise homeorhesis*. It remains to be discovered whether optimal coupling will turn out to be multifractal pink, so to speak, or monofractal pink, or sometimes one, sometimes the other. By comparison, a protracted debate about heart dynamics concludes that cardiovascular wellness is associated with healthy multifractal dynamics [Baillie *et al.*, 2009] versus unhealthy monofractal brown noise in congestive heart disease [Ivanov *et al.*, 1999].

4.4 Joint Action

We began this essay with two examples, a barrel racer racing and a teacher teaching, to introduce the central ideas of coupling and coordination. The focus throughout, though, has been the task performance of individuals, and not the coupling between multiple actors as in the joint action of horse and barrel racer. The rider's skill, to move jointly with her horse, and the horse's skill, to move jointly with the rider, are the basis for their expert coupling to the race course. With skill, coordinative structures emerge in joint action between these members of different species. Joint actions include many opportunities for complexity science, and we remedy the omission in this last section.

Recent efforts on joint action promise a synthesis or rapprochement between conventional science and complexity science. The opportunity came into being with the recognition of language as joint action [Clark, 1996] and an emphasis on the role of language to facilitate coordination [Brennan and Hanna, 2009]. Notice the implicit feedback loop from action-to-language and language-to-action. This feedback loop allows joint attention to reduce demands on language communication in a joint task, for example [Clark and Krych, 2004]. Coordination cannot be encapsulated in a task-person coupling; it emerges across actors.

The capacity for joint action is present within the first year of life [Carpenter, 2009] and constraints that emerge in joint interaction affect the architecture of cognition [Sebanz *et al.*, 2006]. Notice another feedback loop, joint action supports cognitive development that makes more and new joint actions possible. Similar but much slower feedback processes have been posited in the prehistory of human evolution, and the posited feedback loops are tested in experimental semiotics to see whether similar joint actions among contemporary participants bootstrap modes of communication [Galantucci, 2009].

Joint action studies discovered coordinative structures that emerge across individuals. In the classic demonstration, human participants swung their legs together as paired volunteers [Schmidt, 1989]. One of two coordinated patterns emerged: in-phase or anti-phase leg movements between the pairs. Phase dynamics of paired leg swinging revealed phase transitions from one pattern to the other, with concomitant catastrophe flags [Schmidt *et al.*, 1990]; see also [Richardson *et al.*, 2007].

Again, the central tenet of complexity science is that common principles of emergence operate at multiple levels of organization in complex systems — individuals, dyads, groups, society — though each level may also bring into existence new possibilities for action [Marsh *et al.*, 2009]. For example, both conventional and complexity studies of joint action suggest that coordination is the basis of social affiliation, and social affiliation is crucial for individual health and emotional wellbeing, another feedback loop.

Conventional studies discovered a predictive relation between social affiliation and non-conscious mimicry [Lakin and Chartrand, 2003]. The details of coordination dynamics greatly expanded this finding. A variety of manipulations affect the capacity of two individuals to entrain and the same manipulations determine how positive, friendly, and harmonious volunteers rate the experience, and each other as possible teammates [Marsh *et al.*, 2006; 2007]. The degree of entrainment determines whether volunteers like each other [Ouillier *et al.*, 2008].

The strength of entrainment falls off depending on whether individuals can fully focus attention on each other [Richardson *et al.*, 2007]. The basis for entrainment need not be visual however [Richardson *et al.*, 2005]. Two people can perform as one, although they receive only indirect auditory feedback about each other's actions in their separate roles within a shared eye-hand coordination task [Knoblich and Jordan, 2003]. Subtle cues organize dynamics across the two people to perform together as well as a single person with all the information. Does this imply emergent joint intentionality?

Joint action has also become a focal area to introduce new nonlinear methods. Cross-Recurrence analysis, a nonlinear counterpart to correlation, was developed specifically to study shared movements, as in conversations [Shockley, 2005] and was anticipated in Recurrence Quantification Analysis, a nonlinear analog of autocorrelation [Weber and Zbilut, 2005]. These methods were built upon the mathematical theorems of phase space reconstruction, mentioned much earlier. Cross-Recurrence analysis was used first in cognitive science to quantify emergent, coordinative structures between persons in conversation. Joint gaze and joint body posture show spontaneous coordination that predicts mutual understanding. (For a review see [Shockley *et al.*, 2009].)

Nonlinear methods have become more common and existing nonlinear tools, plus new tools in development, promise a truly fresh understanding of behavioral data, e.g. [Marwan, 2009; Riley and Van Orden, 2005; Zbilut and Marwan, 2008]. In retrospect, earlier accomplishments in cognitive science, though inspired by new theoretical ideas from complexity science, were also restricted by caveats on data tools from linear analysis. Future discoveries will benefit from new nonlinear tools that minimize or dispense with such caveats. We stand now, surrounded by opportunities, at a cusp linking theory with new rigorous methods for this science of complexity.

ACKNOWLEDGEMENTS

Preparation of this essay was greatly aided by Anna Haussmann and supported by grants to Guy Van Orden (NSF DHB #0728743; NSF BCS #0642716) and Heidi Kloos (NSF DRL #723638; NICHD HD055324).

BIBLIOGRAPHY

[Abe *et al.*, 2009] N. Abe, T. Fujii, K. Hirayama, A. Takeda, Y. Hosokai, T. Ishioka, Y. Nishio, K. Suzuki, Y. Itoyama, S. Takahashi, H. Fukuda, E. Mori. Do parkinsonian patients have trouble telling lies? The neurobiological basis of deceptive behaviour. *Brain*, 132, 1386-1395, 2009.

[Adolph, 2009] K. Adolph. *Emerging actions: Discovery and selection of new means.* International Conference on Perception-Action, Minneapolis, 2009.

[Allain *et al.*, 1995] H. Allain, A. Lieury, V. Thomas, J. M. Reymann, J. M. Gandon, S. Belliard. Explicit and procedural memory in Parkinson's disease. *Biomedicine & Pharmacotherapy*, 49, 179-186, 1995.

[Aly *et al.*, 2007] N. M. Aly, J. R. Playfer, S. L. Smith, D. M. Halliday. A novel computer-based technique for the assessment of tremor in Parkinson's disease. *Age and Ageing*, 36, 395-399, 2007.

[Ariatti *et al.*, 2008] A. Ariatti, F. Benuzzi, P. Nichelli. Recognition of emotions from visual and prosodic cues in Parkinson's disease. *Neurological Science*, 29, 219-227, 2008.

[Ashby *et al.*, 1993] F. G. Ashby, J.-Y. Tien, J. D. Balikrishnan. Response time distributions in memory scanning. *Journal of Mathematical Psychology*, 37, 526–555, 1993.

[Atmanspacher *et al.*, 2008] H. Atmanspacher, T. Filk, H. Römer. Complementarity in bistable perception. In H. Atmanspacher, H. Primas, eds., *Recasting Reality. Wolfgang Pauli's Philosophical Ideas and Contemporary Science* (pp. 135-150). Berlin: Springer, 2008.

[Atmanspacher *et al.*, 2006] H. Atmanspacher, T. Filk, H. Römer. Weak quantum theory: Formal framework and selected applications.In G. Adenier, A. Khrennikov, Th. M. Nieuwenhuizen. eds., *Quantum Theory: Reconsideration of Foundations* (pp. 34-46). New York: American Institute of Physics, 2006.

[Baillie *et al.*, 2009] R. T. Baillie, A. A. Cecen, C. Erkal. Normal heartbeat series are nonchaotic, nonlinear, and multifractal: New evidence from semiparametric and parametric tests. *Chaos, 19,* 2009.

[Bak, 1997] P. Bak. *How nature works: The science of self-organized criticality.* Oxford University Press, Oxford, 1997.

[Bak *et al.*, 1987] P. Bak, C. Tang, K. Wiesenfeld. Self-organized criticality: An explanation of 1/f noise. *Physical Review Letters*, 59, 381-384, 1987.

[Balakrishnan and Ashby, 1991] J. D. Balakrishnan and F. G. Ashby. Is subitizing a unique numerical ability? *Perception & Psychophysics*, 50, 555-564, 1991.

[Balasubramaniam *et al.*, 2000] R. Balasubramaniam, M. A. Riley M. T. Turvey. Specificity of postural sway to the demands of a precision task. *Gait & Posture*, 11, 12-24, 2000.

[Baumeister and Heatherton, 1996] R. F. Baumeister and T. F. Heatherton. Self-regulation failure: An overview. *Psychological Inquiry*, 7, 1–15, 1996.

[Bauer and Besner, 1997] B. Bauer and D. Besner. Mental set as a determinant of processing in the Stroop task. *Canadian Journal of Experimental Psychology, 51,* 61–68, 1997.

[Bechtel, 2009] W. Bechtel. Constructing a philosophy of science of cognitive science. *Topics in Cognitive Science, 1*, 548-569, 2009.

[Beckers *et al.*, 2006] F. Beckers, B. Verheyden, A. E. Aubert. Aging and nonlinear hear rate control in a healthy population. *American Journal of Physiology*, 290, 2560-2570, 2006.

[Bernstein, 1967] N. A. Bernstein. *The Coordination and Regulation of Movements.* Pergamon Press, London, 1967.

[Besner and Stolz, 1999a] D. Besner and J. A. Stolz. What kind of attention modulates the Stroop Effect? *Psychonomic Bulletin and Review*, 6, 99-104, 1999a.

[Besner and Stolz, 1999b] D. Besner and J. A. Stolz. Unconsciously controlled processing: The Stroop effect reconsidered. *Psychonomic Bulletin & Review, 6,* 449–455, 1999b.

[Besner *et al.*, 1997] D. Besner, J. A. Stolz, C. Boutilier. The Stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review, 4,* 221–225, 1997.

[Blin *et al.*, 1990] O. Blin, A. M. Ferrandez, G. Serratrice. Quantitative analysis of gait in Parkinson patients: Increased variability of stride length. *Journal of the Neurological Sciences*, 98, 91–97, 1990.

[Boogerd *et al.*, 2005] F. C. Boogerd, F. J. Bruggeman, R. C. Richardson, A. Stephan, H. V. Westerhoff. Emergence and its Place in Nature: A Case Study of Biochemical Networks. *Synthese*, 145, 131–164, 2005.

[Brennan and Hanna, 2009] S. E. Brennan and J. E. Hanna. Partner-specific adaptation in dialog. *Topics in Cognitive Science*, 1, 274-291, 2009.

[Bruza and Cole, 2005] P. Bruza and R. Cole. Quantum Logic of Semantic Space: An Exploratory Investigation of Context Effects in Practical Reasoning. In S. Artemov, H. Barringer, A. S. d'Avila Garcez, J. H. Woods. eds., *We will show them: essays in honour of Dov Gabbay* (pp. 339-361). London: College Publications, 2005.

[Bruza *et al.*, 2009] P. Bruza, K. Kitto, D. Nelson, C. McEvoy. Extracting Spooky-Activation-at-a-Distance from Considerations of Entanglement. *Lecture Notes in Artificial Intelligence, Vol. 5494*, 2009.

[Buzsáki, 2006] G. Buzsáki. *Rhythms of the brain.* Oxford University Press, New York, 2006.

[Carpenter, 2009] M. Carpenter. Just how joint is joint action in infancy? *Topics in Cognitive Science*, 1, 380-392, 2009.

[Changizi *et al.*, 2008] M. A. Changizi, A. Hsieh, R. Nijhawan, R. Kanai, S. Shimojo. Perceiving-the-present and a systematization of illusions. *Cognitive Science*, 32, 459-503, 2008.

[Chen *et al.*, 2001] Y. Chen, M. Ding, J. A. S. Kelso. Origins of timing errors in human sensorimotor coordination. *Journal of Motor Behavior*, 33, 3-8, 2001.

[Chialvo, 2008] D. R. Chialvo. Emergent complexity: What uphill analysis or downhill invention cannot do. *New Ideas in Psychology*, 26, 158-173, 2008.

[Christensen and Moloney, 2005] K. Christensen, N. R. Moloney. *Complexity and criticality.* London: Imperial College Press, 2005.

[Clark, 1996] H. H. Clark. *Using language.* Cambridge: Cambridge University Press, 1996.

[Clark and Krych, 2004] H. H. Clark and M. A. Krych. Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50, 62-81, 2004.

[Clayton and Frey, 1997] K. Clayton and B. B. Frey. Studies of mental "noise". *Nonlinear Dynamics, Psychology, and Life Sciences*, 1, 173-180, 1997.

[Correll, 2008] J. Correll. $1/f$ noise and effort on implicit measures of bias. *Journal of Personality and Social Psychology*, 94, 48-59, 2008.

[Cox and Hasselman, 2009] R. F. A. Cox and F. Hasselman. Insight in problem solving: Recurrence analysis of structural changes in cognition. *Talk presented at the 3rd International International Symposium on Recurrence Plots*, Montreal, Canada, 2009.

[Dale, 2008] R. Dale. The possibility of a pluralist cognitive science. *Journal of Experimental and Theoretical Artificial Intelligence*, 20, 155-179, 2008.

[Davia, 2005] C. J. Davia. Life, catalysis and excitable media: A dynamic systems approach to metabolism and cognition. In J. Tuszynski (ed.), *The physical basis for consciousness* (pp. 229-260). Heidelberg: Springer, 2005.

[Diederich *et al.*, 2002] N. J. Diederich, R. Raman, S. Leurgans, C. G. Goetz. Progressive worsening of spatial and chromatic processing deficits in Parinson Disease. *Archieves of Neurology*, 59, 1249-1252, 2002.

[Delignières *et al.*, 2004] D. Delignières, M. Fortes, G. Ninot. The fractal dynamics of self-esteem and physical self. *Nonlinear Dynamics in Psychology and Life Sciences*, 8, 479-510, 2004.

[Delignières *et al.*, 2008] D. Delignières, K. Torre, L. Lemoine. Fractal models for event-based and dynamical timers. *Acta Psychologica*, 127, 382-397, 2008.

[Ding *et al.*, 2002] M. Ding, Y. Chen, J. A. S. Kelso. Statistical analysis of timing errors. *Brain and Cognition*, 48, 98-106, 2002.

[Diniz *et al.*, in press] A. Diniz, M. L. Wijnants, K. Torre, J. Barreiros, N. Crato, A. M. T. Bosman, F. Hasselman, R. F. A. Cox, G. C. Van Orden, D. Delignières. Re-assessing classical motor control theories through serial correlation analysis. *Human Movement Science*, in press.

[Dixon and Bangert, 2004] J. A. Dixon and A. S. Bangert. On the spontaneous discovery of a mathematical relation during problem solving. *Cognitive Science*, 28, 433-449, 2004.

[Dixon and Kelley, 2007] J. A. Dixon and E. Kelley. Theory revision and redescription: Complementary processes in knowledge acquisition. *Current Directions in Psychological Science*, 16, 111-115, 2007.

[Dixon and Kelley, 2006] J. A. Dixon and E. A. Kelley. The probabilistic epigenesis of knowledge. In R. V. Kail (ed.), *Advances in child development and behavior* (Vol. 34, pp. 323-362). New York: Academic Press, 2006.

[Double *et al.*, 2003] K. L. Double, D. B. Rowe, M. Hayes, D. K. Y. Chan, J. Blackie, A. Corbett, R. Joffe, V. S. Fung, J. Morris, G. M. Halliday. Identifying the pattern of olfactory deficits in Parkinson Disease using the brief smell identification test. *Archives of Neurology*, 60, 545-549, 2003.

[Dreyfus, 1992] H. L. Dreyfus. *What computers still can't do: A critique of artificial reason.* MIT Press, Cambridge Mass.,1992.

[Duarte and Sternad, 2008] M. Duarte and D. Sternad. Complexity of human postural control in young and older adults during prolonged standing. *Experimental Brain Research*, 191, 265-276, 2008.

[Dubois, 2003] D. Dubois. Mathematical foundations of discrete and functional systems with strong and weak anticipations. *Lecture Notes in Computer Science*, *Vol. 2684*, 110-132, 2003.

[Durgin and Sternberg, 2002] F. H. Durgin and S. Sternberg. The time of consciousness and vice versa. *Consciousness and Cognition*, 11, 284-290, 2002.

[Dworkin, 1993] B. Dworkin. *Learning and physiological regulation.* University of Chicago Press, Chicago, 1993.

[Edelman, 2008] S. Edelman. A swan, a pike, and a crawfish walk into a bar. *Journal of Experimental and Theoretical Artificial Intelligence*, 20, 257-264, 2008.

[Edwards and Beuter, 1999] R. Edwards and A. Beuter. Parkinsonian Tremor and Simplification in Network Dynamics. *Bulletin of Mathematical Biology*, 51, 157-177, 1999.

[Fabre-Thorpe *et al.*, 2001] M. Fabre-Thorpe, A. Delorme, C. Marlot, S. Thorpe. A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience*, 13, 1-10, 2001.

[Fabre-Thorpe *et al.*, 1996] S. Fabre-Thorpe, D. Fize, C. Marlot. Speed of processing in the human visual system. *Nature*, 381, 520-522, 1996.

[Farrell *et al.*, 2006] S. Farrell, E.-J. Wagenmakers, R. Ratcliff. 1/f noise in human cognition: Is it ubiquitous, and what does it mean? *Psychonomic Bulletin & Review*, 13, 737-741, 2006.

[Fearing, 1970] F. Fearing. *Reflex Action: A Study in the History of Physiological Psychology.* MIT Press, Cambridge Mass., 1970.

[Flach, 1990] J. M. Flach. Control with an eye for perception: Precursors to an active psychophysics. *Ecological Psychology*, 2, 83-111, 1990.

[Flach *et al.*, 2007] J. M. Flach, S. Dekkers, P. J. Stappers. Playing twenty questions with nature (the surprise version): Reflections on the dynamics of experience. *Theoretical Issues in Ergonomic Science*, 9, 125–154, 2007.

[Frenkel-Toledo *et al.*, 2005] S. Frenkel-Toledo, N. Giladi, C. Peretz, T. Herman, L. Gruendlinger, J. M. Hausdor?. Treadmill walking as an external pacemaker to improve gait rhythm and stability in Parkinson's disease. *Movement Disorders*, 20, 1109–1114, 2005.

[Freeman, 2000] W. J. Freeman. *How brains make up their minds.* Cambridge University Press, New York, 2000.

[Frette *et al.*, 1996] V. Frette, K. Christensen, A. Malthe-Sorenssen, J. Feder, T. Jossang, P. Meakin. Avalance dynamics in a pile of rice. *Nature*, 379, 49-52, 1996.

[Fujita and Han, 2009] K. Fujita and H. A. Han. Control of Impulses: The Effect of Construal Levels on Evaluative Associations in Self-Control Conflicts. *Psychological Science*, 20, 799-804, 2009.

[Galantucci, 2009] B. Galantucci. Experimental semiotics: A new approach for studying communication as a form of joint action. *Topics in Cognitive Science*, 1, 393-410, 2009.

[Gibson, 1979] J. J. Gibson. *The ecological approach to visual perception*. Houghton Mifflin, Boston, 1979.

[Gilden, 2001] D. L. Gilden. Cognitive emissions of 1/f noise. *Psychological Review*, 108, 33-56, 2001.

[Gilden, 1997] D. L. Gilden. Fluctuations in the time required for elementary decisions. *Psychological Science*, 8, 296-301, 1997.

[Giuliani *et al.*, 1996] A. Giuliani, P. Lo Giudice, A. M. Mancini, G. Quatrini, L. Pacifici, C. L. Jr. Webber, M. Zak, J. P. Zbilut. A Markovian formalization of heart rate dynamics evinces a quantum-like hypothesis. *Biological Cybernetics*, 74, 181-187, 1996.

[Glass and Mackey, 1988] L. Glass, M. C. Mackey. *From clocks to chaos: The rhythms of life*. Princeton University Press, Princeton, 1988.

[Glenberg, 1997] A. M. Glenberg. What memory is for. *Behavioral and Brain Sciences*, 20, 1-55, 1997.

[Goberman *et al.*, 2008] A. M. Goberman, M. Blomgren, E. Metzger. Characteristics of speech disfluency in Parkinson disease. *Journal of Neurolinguistics*, 2008.

[Goldberger, 1996] A. L. Goldberger. Non-linear dynamics for clinicians: Chaos theory, fractals, and complexity at the bedside. *The Lancet*, 347, 1312-1214, 1996.

[Goldberger *et al.*, 2002a] A. L. Goldberger, L. A. N. Amaral, J. M. Hausdorff, P. Ch. Ivanov, C.-K. Peng, H. E. Stanley. Fractal dynamics in physiology: Alterations with disease and aging. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 2466-2472, 2002.

[Goldberger *et al.*, 2002b] A. L. Goldberger, C.-K. Peng, L. A. Lipsitz. What is physiologic compexity and how does it change with aging and disease? *Neurobiology of Aging*, 23, 23-26, 2002.

[Goldstein, 1939] K. Goldstein. *The Organism, A Holistic Approach to Biology Derived from Pathological Data in Man*. American Book Company, Cincinnati, 1939.

[Greenberg, 2002] G. Greenberg. After the decade of the brain: now what? Review of: *The new phrenology: The limits of localizing cognitive processes in the brain*, by William Uttal. *Behavioral Processes*, 58, 111-114, 2002.

[Greene and Olivia, 2009] M. R. Greene and A. Olivia. The briefest of glances. The time course of natural scene understanding. *Psychological Science*, 20, 464-472, 2009.

[Grill-Spector and Kanwisher, 2006] K. Grill-Spector and N. Kanwisher. Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science*, 16, 152-160, 2006.

[Grossman *et al.*, 2000] M. Grossman, J. Kalmanson, N. Bernhardt, J. Morris, M. B. Stern, H. I. Hurtig. Cognitive Resource Limitations during Sentence Comprehension in Parkinson's Disease. *Brain and Language*, 73, 1-16, 2000.

[Haapaniemi *et al.*, 2001] T. H. Haapaniemi, V. Pursianinen, J. T. Korpelainen, H. V. Huikuri, K. A. Sotaniemi, V.V. Myllylä. Ambulatory ECG and analysis of heart rate variability in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, 70, 305-310, 2001.

[Harley, 2004] T. A. Harley. Promises, promises. *Cognitive Neuropsychology*, 21, 51–56, 2004.

[Hausdorff *et al.*, 2003] J. M. Hausdorff, Y. Balash, N. Giladi. Time series analysis of leg movements during freezing of gait in Parkinson's disease: akinesia, rhyme or reason? *Physica A*, 321, 565-570, 2003.

[Hausdorff *et al.*, 1998] J. M. Hausdor?, M. E. Cudkowicz, R. Firtion, J. Y. Wei, A. L. Goldberger. Gait variability and basal ganglia disorders: Stride-to-stride variations of gait cycle timing in Parkinson's disease and Huntington's disease. *Movement Disorders*, 13, 428–437, 1998.

[Hausdorff *et al.*, 1997] J. M. Hausdorff, S. L. Mitchell, R. Firtion, C.-K. Peng, M. E. Cudkowicz, J. Y. Wei, A. L. Goldberger. Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease. *Journal of Applied Physiology*, 82, 262-269, 1997.

[Hausdorff *et al.*, 1995] J. M. Hausdor?, C. K. Peng, Z. Ladin, J. Y. Wei, A. L. Goldberger. Is walking a random walk? Evidence for long-range correlations in stride interval of human gait. *Journal of Applied Physiology*, 78, 349–358, 1995.

[Hausdorff *et al.*, 1999] J. M. Hausdorff, L. Zemany, C.-K. Peng, A. L. Goldberger. Maturation of gait dynamics: stride-to-stride variability and ist temporal organization in children. *Journal of Applied Physiology*, 86, 1040-1047, 1999.

[Hertrich *et al.*, 1997] I. Hertrich, W. Lutzenberger, S. Spieker, H. Ackermann. Fractal dimension of sustained vowel productions in neurological dysphonias: An acoustic and electroglottographic analysis. *Acoustical society of America*, 102, 652-654, 1997.

[Holden, 2002] J. G. Holden. Fractal characteristics of response time variability. *Ecological Psychology, 14,* 53-86, 2002.

[Holden, 2005] J. G. Holden. Gauging the fractal dimension of response times from cognitive tasks. In M. A. Riley, G. C. Van Orden, eds., *Contemporary nonlinear methods for behavioral scientists: A webbook tutorial* (pp. 267-318). Retrieved August 23, 2009, from `http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp`, 2005.

[Holden *et al.*, 2009] J. G. Holden, G. C. Van Orden, M. T. Turvey. Dispersion of response times reveals cognitive dynamics. *Psychological Review*, 116, 318, 342, 2009.

[Hollis *et al.*, 2009] G. Hollis, H. Kloos, G. C. Van Orden. Origins of Order in Cognitive Activity. In: S. J. Guastello, M Koopmans, D. Pincus, eds., *Chaos and Complexity in Psychology*, 206-241. Cambridge: Cambridge University Press, 2009.

[Howard and Blinks, 2000] L. A. Howard and M. G. B. Binks. The contribution of apraxic speech to working memory deficits in Parkinson's disease. *Brain and Language*, 74, 269-288, 2000.

[Iberall, 1970] A. S. Iberall. On the general dynamics of systems. *General Systems*, *XV*, 7-13, 1970.

[Iberall, 1992] A. S. Iberall. Does intention have a characteristic fast time scale? *Ecological Psychology*, 4, 39-61, 1992.

[Iberall and McCulloch, 1969] A. S. Iberall and W. McCulloch. The organizing principle of complex living systems. *Transactions of the ASME Series D, Journal of Basic Engineering*, 91, 290, 1969.

[Ihlen and Vereijken, in press] E. A. F. Ihlen and B. Vereijken. Beyond $1/f^{\alpha}$ fluctuation in cognitive performance. *Journal of Experimental Psychology: General*, in press.

[Isenhower *et al.*, 2009] R. W. Isenhower, T. D. Frank, B. A. Kay, C. Carello. *Dynamics of valenced events.* Paper presented at the 15^{th} International Conference on Perception and Action, Minneapolis, 2009.

[Ito *et al.*, 2005] J. Ito, A. R. Nikolaev, C. van Leeuwen. Spatial and temporal structure of phase synchronization of spontaneous alpha EEG activity. *Biological Cybernetics*, 92, 54-60, 2005.

[Ivanov *et al.*, 1999] P. C. Ivanov, L. A. N. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, H. E. Stanley. Multifractality in human heartbeat dynamics. *Nature, 399*, 461-465, 1999.

[Jacob, 1982] Jacob, F. (1982). *The possible and the actual.* University of Washington Press, Seattle, 1982.

[Jankovic *et al.*, 1999] J. Jankovic, K. S. Schwartz, W. Ondo. Re-emergent tremor of Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, 67, 646-650, 1999.

[Jensen, 1998] H. J. Jensen. *Self-Organized Criticality. Emergent Complex Behavior in Physical and Biological Systems.* Cambridge: Cambridge University Press, 1998.

[Jordan, 2008] J. S. Jordan. Wild-agency: Nested intentionalities in neuroscience and archeology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363, 1981-1991, 2008.

[Jordan and Hunsinger, 2008] J. S. Jordan, M. Hunsinger. Learned patterns of action-effect extrapolation contribute to the spatial displacement of continuously moving stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 113–124, 2008.

[Juarrero, 1999] A. Juarrero. *Dynamics in action: Intentional behavior as a complex system.* MIT Press, Cambridge Mass., 1999.

[Kanne *et al.*, 1998] S. M. Kanne, D. A. Balota, D. H. Spieler, M. E. Faust. Explorations of Cohen, Dunbar, and McClelland's (1990) connectionist model of Stroop performance. *Psychological Review, 105* (1), 174-187, 1998.

[Kay, 1988] B. A. Kay. The dimensionality of movement trajectories and the degrees of freedom problem: A tutorial. *Human Movement Science*, 7, 343-364, 1988.

[Kello *et al.*, 2008] C. T. Kello, G. G. Anderson, J. G. Holden, G. C. Van Orden. The pervasiveness of 1/f scaling in speech reflects the metastable basis of cognition. *Cognitive Science*, 32, 1217-1231, 2008.

[Kello *et al.*, 2007] C. T. Kello, B. C. Beltz, J. G. Holden, G. C. Van Orden. The emergent coordination of cognitive function. *Journal of Experimental Psychology: General*, 136, 551-568, 2007.

[Kello and Van Orden, 2009] C. T. Kello and G. C. Van Orden. Soft-assembly of sensorimotor function. *Nonlinear Dynamics, Psychology, and Life Sciences*, 13, 57-78, 2009.

[Kelso, 1995] J. A. S. Kelso. *Dynamic Patterns: The Self-Organization of Brain and Behavior.* Cambridge: MIT Press, Cambridge Mass., 1995.

[Kelso *et al.*, 1992] J. A. S. Kelso, S. L. Bressler, S. Buchanan, G. C. Deguzman, M. Ding, A. Fuchs, T. Holroyd. A phase-transition in human brain and behavior. *Physics Letters A*, 169, 134-144, 1992.

[Kelso and Tognoli, 2007] J. A. S. Kelso and E. Tognoli. Toward a complementary neuroscience: Metastable coordination dynamics of the brain. In R. Kozma, L. Perlovsky, eds., *Neurodynamics of Cognition and Consciousness* (pp. 39-60). Heidelberg: Springer, 2007.

[Kiefer *et al.*, 2009] A. W. Kiefer, M. A. Riley, K. Shockley, S. Villard, G. C. Van Orden. Walking changes the dynamics of cognitive estimates of time intervals. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 1532-1541, 2009.

[Kinouchi and Copelli, 2006] O. Kinouchi and M. Copelli. Optimal dynamical range of excitable networks at criticality. *Nature Physics*, 2, 348–352, 2006.

[Kloos *et al.*, in press] H. Kloos, A. Fisher, G. C. Van Orden, G. C. Situated Naive Physics: Task Constraints Decide what Children Know about Density, *Journal of Experimental Psychology: General*, in press.

[Kloos *et al.*, 2009] H. Kloos, A. W. Kiefer, L. Gresham, K. Shockley, M. A. Riley, G. C. Van Orden,. Response time dynamics of children and adults. *Talk presented at the 15th International Conference on Perception and Action*, Minneapolis, 2009.

[Kloos and Van Orden, 2009] H. Kloos and G. C. Van Orden. Soft-assembled mechanisms for the unified theory. In J.P. Spencer, M. Thomas, J. McClelland, eds., *Toward a Unified Theory of Development: Connectionism and Dynamics Systems Theory Reconsidered*, 253-267. New York: Oxford University Press, 2009.

[Knoblich and Jordan, 2003] G. Knoblich and S. Jordan. Action coordination in groups and individuals: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 29, 1006-1016, 2003.

[Lakin and Chartrand, 2003] J. L. Lakin and T. L. Chartrand. Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, 14, 334-339, 2003.

[Lin *et al.*, 2008] D. Lin, H. Deol, M. A. Nussbaum, M. L. Madigan. Reliability of COP-based postural sway measures and age-related differences. *Gait & Posture*, 28, 337-342, 2008.

[Lloyd, 1999] A. J. Lloyd. Comprehension of prosody in Parkinson's disease. *Cortex*, 35, 389-402, 1999.

[Luce, 1986] R. D. Luce. *Response Times.* New York: Oxford University Press, 1986.

[Mack and Rock, 2000] A. Mack and I. Rock, *Inattentional Blindness.* MIT Press, Cambridge, Mass., 2000.

[Maddox *et al.*, 1998] W. T. Maddox, F. G. Ashby, L. R. Gottlob. Response time distributions in multidimensional perceptual categorization. *Perception & Psychophysics*, 60, 620-637, 1998.

[Mañé, 1981] R. Mañé. On the dimension of the compact invariant sets of certain non-linear maps (pp. 230–242). In D. Rand, L.-S. Young, eds., *Lecture Notes in Mathematics*, *Vol. 898*, 1981.

[Marsh *et al.*, 2009] K. L. Marsh, M. J. Richardson, R. C. Schmidt. Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, 1, 320-339, 2009.

[Marsh *et al.*, 2006] K. L. Marsh, M. J. Richardson, R. C. Schmidt. Emergence of sociality from the dynamics of bodily synchrony. In E. Smith (Chair), *Social embodiment. Invited symposium presented at the meeting of the Society of Experimental Social Psychology*, Philadelphia, 2006.

[Marsh *et al.*, 2007] K. L. Marsh, M. J. Richardson, R. C. Schmidt. Social connection through joint action and interpersonal coordination. *Paper presented at the 2nd Joint Action Meeting*, Newark, 2007.

[Marwan, 2009] N. Marwan. Hot topics in the recurrence plot field. *Paper presented at the 3rd International Symposium on Recurrence Plots*, Montreal, 2009.

[McGraw, 1945] M. B. McGraw. *The neuromuscular maturation of the human infant.* Columbia University Press, New York, 1945.

[Meillassoux, 2008] Q. Meillassoux. *After Finitude: An Essay on the Necessity of Contingency.* Continuum, New York, 2008.

[Michell, 1999] J. Michell. *Measurement in Psychology. A Critical History of a Methodological Concept.* Cambridge University Press, Cambridge, 1999.

[Mitra *et al.*, 1997] S. Mitra, M. Riley, M. T. Turvey. Chaos in human rhythmic movement. *Journal of Motor Behavior*, 29, 195-198, 1997.

[Molenaar, 2008] P. C. M. Molenaar. On the implications of the classical ergodic theorems: Analysis of developmental processes has to focus on intraindividual variation. *Developmental Psychobiology*, 50, 60–69, 2008.

[Nelson and McEvoy, 2007] D. Nelson and C. McEvoy. Entangled associative structures and context. In P. Bruza, W. Lawless, C. v. Rijsbergen, D. Sofge, eds., *Proceedings of the AAAI Spring Symposium on Quantum Interaction*. Menlo Park: AAAI Press, 2009.

[Newell, 1990] A. Newell. *Unified Theories of Cognition*. Harvard University Press, Cambridge Mass., 1990.

[Newell and Slifkin, 1998] K. M. Newell and A. B. Slifkin. The nature of movement variability. In J. P. Piek (ed.), *Motor behavior and human skill: A multidisciplinary approach* (pp. 143-160). Champaign, IL: Human Kinetics, 1998.

[Nicolis, 1989] G. Nicolis. Physics of far-from-equilibrium systems and self-organisation. In P. Davies (ed.), *The New Physics* (pp. 316-347). New York: Cambridge University Press, 1989.

[Nicolis and Nicolis, 2007] G. Nicolis and C. Nicolis. *Foundations of Complex Systems*. World Scientific, Hackensack NJ, 2007.

[Nicolis and Rouvas-Nicolis, 2007] G. Nicolis and C. Rouvas-Nicolis. Complex systems. *Scholarpedia*, 2, 1473. Retrieved April 24, 2009, from `http://www.scholarpedia.org/article/Complex_systems`, 2009.

[Norris *et al.*, 2005] J. A. Norris. A. P. March, I. J. Smith, R. I. Kohut, M. E. Miller. Ability of static and statistical mechanics posturographic measures to distinguish between age and fall rist. *Journal of Biomechanics*, 38, 1263-1272, 2005.

[Oullier *et al.*, 2008] O. Oullier, G. C. de Guzman, K. J. Jantzen, J. Lagarde, J. A. S. Kelso. Social coordination dynamics: Measuring human bonding. *Social Neuroscience*, 3, 178-192, 2008.

[Pan *et al.*, 2007] W. Pan, K. Ohashi, Y. Yamamoto, S. Kwak. Power-law temporal autocorrelation of activity refelcts severity of Parkinsonism. *Movement Disorders*, 22, 1308-1313, 2007.

[Pansky and Algom, 1999] A. Pansky, D. Algom. Stroop and Garner effects in comparative judgment of numerals: The role of attention. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 39-58, 1999.

[Pansky and Algom, 2002] A. Pansky and D. Algom. Comparative judgment of numerosity and numerical magnitude: Attention preempts automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*(2), 259–274, 2002.

[Peron *et al.*, 2009] J. Peron, S. Vicente, E. Leray, S. Drapier, D. Drapier, R. Cohen, I. Biseul, T. Rouaud, F. LeJeune, P. Sauleau, M. Verin. Are dopaminergic pathwaqys involved in theory of mind? A study in Parkinson's disease. *Neuropsychologia*, 47, 406-414, 2009.

[Previc, 1999] F. H. Previc. Dopamine and the Origins of Human Intelligence. *Brain and Cognition*, 41, 299-350, 1999.

[Price and Shin, 2009] A. Price and J. C. Shin. The impact of Parkinson's disease on sequence learning: Perceptual pattern learning and executive function. *Brain and Cognition*, 69, 252-261, 2009.

[Prochazka *et al.*, 2000] A. Prochazka, J. R. Wolpaw, F. Clarac, G. E. Loeb, J. C. Rothwell. What do reflex and voluntary mean? Modern views on an ancient debate. *Experimental Brain Research*, 130, 417-432, 2000.

[Rachlin, 2000] H. Rachlin. *The Science of Self-Control*. Harvard University Press, Cambridge Mass., 2000.

[Richardson *et al.*, 2007] M. J. Richardson, K. L. Marsh, R. M. Baron. Judging and Actualizing Intrapersonal and Interpersonal Affordances. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 845-859, 2007.

[Richardson *et al.*, 2007] M. J. Richardson, K. L. Marsh, R. Isenhower, J. Goodman, R. C. Schmidt. Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, 26, 867-891, 2007.

[Richardson *et al.*, 2005] M. J. Richardson, K. L. Marsh, R. C. Schmidt. Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology: Human Perception & Performance*, 31, 62-79, 2005.

[Riley, 2007] M. A. Riley. Postural sway dynamics and the task-specific, soft assembly of postural control. *Workshop on Behavioral and Cognitive Dynamical Systems*, Ithaca, NY, 2007.

[Riley and Turvey, 2002] M. A. Riley and M. T. Turvey. Variability and determinism in motor behavior. *Journal of Motor Behavior*, 34, 99-125, 2002.

[Riley and Van Orden, 2005] M. A. Riley, G. C. Van Orden, eds. *Tutorials in contemporary nonlinear methods for the behavioral sciences* (pp. 353-400). Retrieved May 1, 2009, from `http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp,` 2005.

[Riley and Turvey, 2001] M. A. Riley and M. T. Turvey. The self-organizing dynamics of intentions and actions. *American Journal of Psychology, 114,* 160-169, 2001.

[Roepstorff and Jack, 2004] A. Roepstorff and A. I. Jack. Trust or Interaction?: Editorial Introduction. *Journal of consciousness Studies*, 11, v-xxii, 2004.

[Saltzman and Holt, 2009] E. Saltzman and K. Holt. *Affordances, effectivities, coordinative structures: A graph- and task-dynamic framework*. Paper presented at the 15th International Conference on Perception and Action, Minneapolis, 2009.

[Schmidt, 1989] R. C. Schmidt. Dynamical constraints on the coordination of rhythmic limb movements between two people. Doctoral dissertation, University of Connecticut, 1988. *Dissertation Abstracts International*, 50, 2128, 1989.

[Schmidt *et al.*, 1990] R. C. Schmidt, C. Carello, M. T. Turvey. Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 227-247, 1990.

[Schmit *et al.*, 2005] J. M. Schmit, D. Regis, M. A. Riley. Dynamic patterns of postural sway in ballet dancers and track athletes. *Experimental Brain Research*, 163, 370-378, 2005.

[Schmit *et al.*, 2006] J. M. Schmit, M. A. Riley, A. Dalvi, A. Sahay, P. K. Shear, K. D. Shockley, R. Y. K. Pun. Deterministic center of pressure patterns characterize postural instability in Parkinsons's disease. *Experimental Brain Research*, 168, 357-367, 2006.

[Schwartz and Black, 1996] D. L. Schwartz and J. B. Black. Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science*, 20, 457-497, 1996.

[Science Watch, 1999] Science Watch. *American Psychologist*, 54, 461–515, 1999.

[Searle, 1980] J. R. Searle. Minds, brains, and programs. *Behavioral and Brain Sciences, 3*, 417-424, 1980.

[Sebanz *et al.*, 2006] N. Sebanz, H. Bekkering, G. Knoblich. Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, 10, 70-76, 2006.

[Simon, 1973] H. A. Simon. The organization of complex systems. In H. H. Pattee (ed.) *Hierarchy theory: The challenge of complex systems* (pp. 1–27). New York: George Braziller, 1973.

[Simons and Chabris, 1999] D. J. Simons and C. F. Chabris. Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, 28, 1059-1074, 1999.

[Shanon, 1993] B. Shanon. *The Representational and the Presentational: An Essay on Cognition and the Study of Mind*. Harvester-Wheatsheaf, London, 1993.

[Shaw and Turvey, 1999] R. E. Shaw, M. T. Turvey. Ecological foundations of cognition: II. Degrees of freedom and conserved quantities in animal-environment systems. *Journal of Consciousness Studies*, 6, 111-123, 1999.

[Shaw *et al.*, 1982] R. E. Shaw, M. T. Turvey, W. M. Mace. Ecological psychology: The consequence of a commitment to realism. In W. Weimer, D. Palermo, eds., *Cognition and the symbolic processes* (Vol. 2, pp. 159–226). Hillsdale: Erlbaum, 1982.

[Shockley, 2005] K. Shockley. Cross recurrence quantification of interpersonal postural activity. In M. A. Riley, G. Van Orden, eds., *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences* (pp. 142-177). Retrieved May 1, 2009, from `http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp`, 2005.

[Shockley *et al.*, 2009] K. Shockley, D. C. Richardson, R. Dale. Conversation and coordinative structures. *Topics in Cognitive Science*, 1, 305-319, 2009.

[Sporns, 2007] O. Sporns. Complexity. *Scholarpedia*, 2, 1623. Retrieved April 24, 2009, from `http://www.scholarpedia.org/article/Complexity`, 2007.

[Stanovich, 2004] K. E. Stanovich. *The robot's rebellion: Finding meaning in the age of Darwin*. University of Chicago Press, Chicago, 2004.

[Stephen *et al.*, 2009a] D. G. Stephen, R. Boncoddo, J. S. Magnuson, and J. A. Dixon. The dynamics of insight: Mathematical discovery as a phase transition. *Memory and Cognition*, 37, 1132-1149, 2009.

[Stephen *et al.*, 2009b] D. G. Stephen, J. A. Dixon, and R. W. Isenhower. Dynamics of representational change: Entropy, action, and cognition. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 1811-1822, 2009.

[Stephen *et al.*, 2008] D. G. Stephen, N. Stepp, J. Dixon, M. T. Turvey. Strong anticipation: Sensitivity to long-range correlations in synchronization behavior. *Physica A*, 387, 5271-5278, 2008.

[Stepp and Turvey, 2009] N. Stepp and M. T. Turvey. On strong anticipation. *Cognitive Systems Research*, doi: 10.1016/j.cogsys.2009.03.003, 2009.

[Strogatz, 1994] S. H. Strogatz. *Nonlinear dynamics and chaos.* Addison Wesley, Reading, 1994.

[Takens, 1981] F. Takens. Detecting strange attractors in turbulence. In D. Rand and L.-S. Young, eds., *Dynamical Systems and Turbulence, Warwick 1980.* Lecture Notes in Mathematics, Vol 898, 366-381, Springer, Berlin, 1981.

[Thelen *et al.*, 2002] E. Thelen D. M. Fisher, R. Ridley-Johnson. The relationship between physical growth and a newborn reflex. *Infant Behavior & Development*, 25, 72-85, 2002.

[Thelen and Smith, 1994] E. Thelen and L. B. Smith. *A dynamic systems approach to the development of cognition and action.* Bradford Books/MIT Press, Cambridge Mass., 1994.

[Thornton and Gilden, 2005] T. L. Thornton and D. L. Gilden. Provenance of correlations in psychological data. *Psychonomic Bulletin & Review*, 12, 409-441, 2005.

[Thorpe, 2002] S. J. Thorpe. Ultra-Rapid Scene Categorization with a Wave of Spikes. *Lecture Notes In Computer Science Vol. 2525*, 1-15, 2002.

[Thurner *et al.*, 2002] S. Thurner, C. Mittermaier, K. Ehrenberger. Change of Complexity Patterns in Human Posture during Aging. *Audiology & Neurotology*, 7, 240-248, 2002.

[Torre *et al.*, 2007] K. Torre, D. Delignières, L. Lemoine. $1/f\beta$ fluctuations in bimanual coordination: An additional challenge for modeling. Experimental Brain Research, 183, 225-234, 2007.

[Torre and Wagenmakers, 2009] K. Torre and E. J. Wagenmakers. Theories and models for $1/f^{\beta}$ noise in human movement science. *Human Movement Science*, 28, 297-318, 2009.

[Tsonis, 2008] A. A. Tsonis. *Randomicity: Rules and Randomness in the Realm of the Infinite.* Imperial College Press, London, 2008.

[Turvey, 1990] M. T. Turvey. Coordination. *American Psychologist*, 45, 938-953, 1990.

[Turvey and Carello, 1981] M. T. Turvey and C. Carello. Cognition: The view from ecological realism. *Cognition*, 10, 313-321, 1981.

[Turvey and Moreno, 2006] M. T. Turvey and M. A. Moreno. Physical metaphors for the mental lexicon. *The Mental Lexicon*, 1, 7-33, 2006.

[Tzelgov, 1997] J. Tzelgov. Specifying the relations between automaticity and consciousness: A theoretical note. *Consciousness and Cognition, 6*, 441–451, 1997.

[Ulanowicz, 2000] R. E Ulanowicz. *Growth and Development: Ecosystems Phenomenology.* toExcel Press, San Jose CA, 2000.

[Uttal, 2001] W. R. Uttal. The new phrenology: *The limits of localizing cognitive functions in the brain.* MIT Press, Cambridge Mass., 2001.

[Uttal, 2007] W. Uttal. *Neural Theories of Mind: Why the Mind-Brain Problem may Never be Solved.* Erlbaum, Mawah, 2007.

[Van Orden, 2008] G. C. Van Orden. Commentary: Peirce's abduction and cognition as we know it. *Journal of Experimental & Theoretical Artificial Intelligence*, 20, 219-229, 2007.

[Van Orden, 2007] G. C. Van Orden. *Prospects at Hand for Cognitive Dynamics.* Talk given at the Brain Network Dynamics Conference, 2007. University of California at Berkeley. Berkeley, CA. Retrieved May 25, 2009, from `http://www.archive.org/details/Brain_Network_Dynamics_2007-29-Guy_Van_Orden`, 2007.

[Van Orden and Holden, 2002] G. C. Van Orden and J. G. Holden. Intentional contents and self-control. *Ecological Psychology*, 14, 87-109, 2002.

[Van Orden *et al.*, 1999] G. C. Van Orden, J. G. Holden, M. N. Podgornik, C. S. Aitchison. What swimming says about reading: coordination, context, and homophone errors. *Ecological Psychology*, 11, 45-79, 1999.

[Van Orden *et al.*, 2003] G. C. Van Orden, J. G. Holden, M. T. Turvey. Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, 132, 331-350, 2003.

[Van Orden *et al.*, 2005] G. C. Van Orden, J. G. Holden, M. T. Turvey. Human cognition and 1/f scaling. *Journal of Experimental Psychology: General*, 134, 117-123, 2005.

[Van Orden *et al.*, 2010] G. C. Van Orden, C. T. Kello, J. G. Holden. Situated behavior and the place of measurement in psychological theory. *Ecological Psychology*, 22, 24-43, 2010.

[Van Orden and Paap, 1997] G. C. Van Orden and K. R. Paap. Functional neuroimages fail to discover pieces of mind in the parts of the brain. *Philosophy of Science*, 64, 85-94, 1997.

[Van Orden *et al.*, 2001] G. C. Van Orden, B. F. Pennington, G. O. Stone. What do Double Dissociations Prove? *Cognitive Science*, 25, 111-172, 2001.

[Van Orden *et al.*, 1997] G. C. Van Orden, M. A. Jansen op de Haar, A. M. T. Bosman. Complex dynamic systems also predict dissociations, but they do not reduce to autonomous components. *Cognitive Neuropsychology*, 14, 131-165, 1997.

[VanRullen and Thorpe, 2002] R. VanRullen and S. J. Thorpe. Surfing a spike wave down the ventral stream. *Vision Research*, 42, 2593-2615, 2002.

[Vollmer, 2001] F. Vollmer. The control of everyday behaviour. *Theory & Psychology*, 11, 637–654, 2001.

[Varela *et al.*, 2003] M. Varela, L. Jimenez, R. Farina. *European Journal of Applied Physiology*, 89, 230-237, 2003.

[Wagenmakers *et al.*, 2004] E.-J. Wagenmakers, S. Farrell, R. Ratcliff. Estimation and interpretation of 1/f noise in human cognition. *Psychonomic Bulletin & Review*, 11, 579–615, 2004.

[Wagenmakers *et al.*, 2005] E.-J. Wagenmakers, S. Farrell, R. Ratcliff. Human cognition and a pile of sand: A discussion on serial correlations and self-organized criticality. *Journal of Experimental Psychology: General*, 134, 108-116, 2005.

[Ward, 2002] L. Ward. *Dynamical cognitive science*. MIT Press, Cambridge Mass., 2002.

[Warren, 2006] W. Warren. The dynamics of perception and action. *Psychological Review*, 113, 358-389, 2006.

[Watkins, 1990] M. J. Watkins. Mediationism and the obfuscation of memory. *American Psychologist*, 45, 328–335, 1990.

[Webber and Zbilut, 2005] C. L. Jr. Webber, J. P. Zbilut. Recurrence quantification analysis of nonlinear dynamical systems. In M. A. Riley, G. C. Van Orden, eds., *Tutorials in contemporary nonlinear methods for the behavioral sciences* (pp. 26-96). Retrieved May 1, 2009, from `http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp`, 2005.

[Wegner, 2002] D. M. Wegner. *The Illusion of Conscious Will*. MIT Press, Cambridge Mass., 2002.

[Weldon, 1999] M. S. Weldon. The memory chop shop: Issues in the search for memory systems. In J. K. Forster, M. Jelicic, eds., *Memory: Systems, process, or function?* (pp. 162–204). New York: Oxford University Press, 1999.

[West, 2006] B. J. West. *Where medicine went wrong. Rediscovering the path to complexity.* World Scientific, London, 2006.

[Winget *et al.*, 1968] C. Winget, D. Card, J. Pope. Circadian oscillations of three parameters at defined light intensities and color. *Journal of Applied Physiology*, 24, 401-406, 1968.

[Wijnants *et al.*, 2009] M. L. Wijnants, A. M. T. Bosman, F. Hasselman, R. F. A. Cox, G. C. Van Orden. $1/f$ Scaling in movement time changes with practice in precision aiming. *Nonlinear Dynamics, Psychology, and the Life Sciences*, 13, 79-98, 2009.

[Wink *et al.*, 2006] A. M. Wink, F. Bernard, R. Salvador, E. Bullmore, J. Suckling. Age and cholinergic effects on hemodynamics and functional coherence of human hippocampus. *Neurobiology of Aging*, 27, 1395-1404, 2006.

[Zak, 1993] M. Zak. Introduction to terminal dynamics *Complex Systems*, 7, 59-87, 1993.

[Zbilut, 2004] J. P. Zbilut. *Unstable singularities and randomness: their importance in the complexity of physical, biological and social sciences*. Elsevier, Amsterdam, 2009.

[Zbilut and Marwan, 2008] J. P. Zbilut and N. Marwan. The Wiener-Khinchin theorem and recurrence quantification. *Physics Letters A*, 372, 6622–6626, 2008.

[Zhang and Jiang, 2008] Y. Zhang and J. J. Jiang. Nonlinear dynamic mechanism of vocal tremor from voice analysis and model simulations. *Journal of Sound and Vibration*, 316, 248-262, 2008.

Part IX

Medicine

CHINESE MEDICINE AND COMPLEX SYSTEMS DYNAMICS

W. E. Herfel, Y. Gao and D. J. Rodrigues

The concepts, principles and methods of complex systems dynamics have a profound impact on our appreciation of and engagement with Chinese medicine. Firstly, they offer a framework for understanding the fundamental approach of Chinese medicine that for the first time renders it systematically intelligible in the terms of Western science whilst respecting its own salient commitments. In the process they underline the ways in which the orthodox Western medical world-view contrasts with that of Chinese medicine, and thereby the ways that complex systems approaches within Western medicine itself likewise differ from the orthodoxy. Secondly, they provide an articulation of Chinese medical aims and practices which emanate from its conceptions of health and disease permitting engagement with those of Western medicine. Thirdly, they highlight the differences between two distinct approaches in medical treatment and treatment validation: One, stemming from complex dynamics, employs methods focussed on tuning the patient to harmony. The other, deriving from classical science, aims at restoring homeostatic balance. The former emphasises the context dependence — especially patient-dependence — of diagnosis and treatment, whilst the latter seeks to abstract away from details of context, which from its point of view are "confounding variables" by employing traditional logical-causal methods. Fourthly and finally, their application to Chinese medicine raises important epistemological questions about the roles of evidence, judgment, specificity and individuality in objective scientific medical research. These final issues are best analysed by illuminating the situated cognition and embodied knowledge of Chinese medicine in light of a complex systems dynamics perspective. This analysis shows that evidence, judgment and practice are more subtly interconnected in both medical traditions than one gets the impression from examining the literature generated by advocates of "evidence-based medicine". In the next section the groundwork for the enquiry are laid by contrasting the ancient cosmologies that underwrite Chinese and Western medical practice and theory.

1 CONTRASTING COSMOLOGIES: THE ELEATIC AND DAOIST TRADITIONS

The dominant tradition in Western philosophy gives primacy to the notion of eternal truth. On this view the most significant form of knowledge is of things

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

that do not change. This tradition dates back to the pre-Socratic philosophers of Elea, some of whom provided elegant logical proof that change was impossible, and is articulated clearly in the work of Plato. In the *Republic* [Plato, 1961, p. 813] it states:

> Which do you think more truly *is,* that which clings to what is ever like itself and immortal and to the truth, and that which is itself of such a nature and is born in a thing of that nature, or that which clings to what is mortal and never the same and is itself such and is born in such a thing?
>
> That which cleaves to what is ever the same far surpasses, he said.
>
> Does the essence of that which never abides the same partake of real essence any more than of knowledge?
>
> By no means.
>
> Or of truth and reality?
>
> Not of that either.
>
> And if a thing has less of truth has it not also less of real essence or existence?
>
> Necessarily.

According to Greek scholar, G. M. A. Grube [1980, p. 5], Plato's argument rests on the following:

> If every science fulfils its function by having some one thing as its object, there must be such a single thing which is object of that science, it must be unchanging and eternal, an eternal model beyond the particular sensible things, for these cannot be the objects of knowledge in any proper sense.

Today Western science retains this Platonic legacy. Physics has progressed by explaining change in terms of that which remains constant. The mathematics of classical dynamics is a mathematics of certainty. The conservation laws are the fixed frame of reference within which dynamics is intelligible. Of course in the Twentieth Century all the great theories in physics have called these foundational certainties into question. Special Relativity rejects the fixed frame of reference. Quantum Mechanics, the most rigourously and thoroughly tested scientific theory of all time, posits irreducible uncertainty at the bottom of reality. Far-from–Equilibrium Thermodynamics seeks to articulate the energetics of the emergence of genuine novelty in complex systems. It is only in the early Twenty-first Century, with the advent of *complex systems dynamics* that we arrive at a physics of change, a true science of contingency. This vehicle is an ideal medium through which to articulate that most elusive of ancient Chinese practices, the traditional healing arts whose theory rests on a cosmology presupposing change, not constancy, as ontologically primary. In this way the ancient Chinese worldview is the antithesis of that of the ancient Eleatics.

The ancient cosmology upon which Chinese medicine rests articulates reality as a system of complexly interacting processes. This cosmology sees the world not as primarily mysterious, but as an intelligible continuum. To get an idea of how this worldview is organised we can reflect on how the ancient Chinese might view the structure of a river. In Chinese cosmology the world is like a river. The river flows because of the energy differential between its source and estuary. In other words the dissipation of energy enables the characteristic features of the river to emerge. As the river flows waves, eddies, currents, rapids, pools, etc. emerge. We identify such features, but they do not have existence independent of the river. They exist only as the result of river flow yet they are entities in their own right. The river is the material and energetic substrate that enables the existence of such phenomena as waves and eddies. The river is a process of flow that is differentiated into various sub-processes underlying such features as waves, eddies, etc. It can be compelling to view the riverbed as the *permanent* structure that gives *form* to these phenomena. However, whilst at any given time the riverbed plays a constraining, shaping role, the process of flow, including such factors as the rate of flow, are also crucial to determining the location, shape and size of such features as eddies and waves. Furthermore, change and permanence are functions of time scale. For instance, over time, water flow gives rise to such processes as erosion that, reverse-wise, impact on the shape of the riverbed. Various aspects of the structure of the river are thus dynamic over different time-scales. The flow of the river and the structure of the riverbed interact over time to give rise to the shape of the dynamics of the river.

Eddy, river and river bed are all parts of the same interconnected processes. In Chinese cosmology the entire world is like our river example. Chinese cosmology is holistic in viewpoint recognizing the interconnection of energetic processes occurring in the vicinity of the earth. Of course, for some purposes it may be useful to isolate individual phenomena. But this analysis is only a useful fiction. One could model the eddy ignoring the river that sustains it. In other words, one could black-box the energy source that maintains the eddy, but this does not grant the eddy independent existence in any material sense: when the river ceases to flow the eddy ceases to exist. Equally important is the fact that its dependence on the river does not make the eddy any less real.

Recent research by Chinese scholars shows that there was a multitude of philosophic traditions contributing to the theoretical framework of classical Chinese medicine as represented in its classic text the *Huang Di Nei Jing Su Wen*.[1] Such traditions include the *Yin-Yang* School, the Five Phase School, the *Yi Jing*[2],

[1]This is the *Yellow Emperor's Inner Classic: Plain Questions*. It will be referred to as *Su Wen* throughout. All translations are our own; page numbers in parentheses refer to NCMCCG [1997]. We have also consulted Lu [1978] in the course of our translation.

[2]This is referred to in English as the *Book of Changes*. Translations are our own but references to English translations of this work, as well as to the Chinese originals that we consulted, appear in the References. There are many English translations, but the Baynes' rendering into English

numerology and astronomy, all of which prevailed during the formative time of Chinese medicine. Although it is hard to clarify which tradition contributed what retrospectively, it is not difficult to articulate the philosophic framework presented in the *Su Wen* from the text and identify the philosophical assumptions shared with the other traditions, including:

- Constant change (normal and abnormal) is natural
- Recognition and prediction of patterns of change guide the course of human action
- The cosmos and human body are both perceived as processes sharing the same characteristic dynamics
- Integrative and differentiating diagnostic strategies are preferred over seeking causal relationships between events
- Individualized treatment presupposes an adaptation strategy
- Evaluation ought to be based on improved physiological signs, e.g. pulse quality and complexion, which change according to seasonal influences and medical intervention

The assumption of an ever changing universe was clearly expressed in the *Yi Jing* the most influential text during the formative age of Chinese philosophy. It is understandable that it has shaped the Chinese conception of the natural as well as human world. As the title of the text suggests, the world in the *Book of Changes* is one in constant transformation. It is stated clearly in the *Ta Chuan* (*Yi Jing*, Ch.1 Pt. I):

> The sky is high, the earth is low; the *yang* and *yin* are thus distinguished; high and low are displayed, superior and inferior are established. Movement and rest have their regularity, hard and soft differ. ... [Such changes] complete the phenomena in the sky, and forms on earth, making the changes and transformations apparent. The hard and the soft rub each other; the eight processes wash across each other; drumming them with thunder and lightning; moistening them with wind and rain; the sun and moon move in cycles, [making things] cold and then hot.

It is quite clear from the *Yi Jing* (Ch. 1 Pt. V) that in the world of traditional Chinese cosmology transformation follows a loosely cyclical pattern with rhythms across a variety of temporal scales:

---

of Wilhelm's translation in to German has become the canonical English version (*The I Ching*, 1968).

> The alternation of *yin* and *yang* is called the *Dao* (the ongoing dynamic process that give rise to all other things and beings in this universe); .... That y*in* gives way to *yang* and then *yang* gives way to *yin* is called change; ....the unpredictability of *yin yang* alternation is called beyond description.

Such assumptions are also explicitly expressed in *Dao De Jing*[3] (Ch. 25) when describing *Dao*, the ongoing process that gives birth to everything in this universe:

> There was a spontaneously formed process,
>
> Emerging prior to the sky and the earth,
>
> Silent and empty,
>
> Standing alone, its substance unchanging.
>
> Moving in cycles, never pausing,
>
> It is the mother of sky and earth.

The constant movement of *qi*[4] is the generative process that created the environment inhabited by human beings. The human body therefore corresponds to the changes of the climate (categorized by the six *qi*, i.e., wind, cold, heat, humidity, drought, and fire) and the changes in the processes on earth (categorized by the five phase transformation sequence), and should change with similar patterns. The *Su Wen* (Ch. 68 (503-4)) makes this clear:

> *Qi Bo* said: When discussing change in climate, one examines the six *qi*; when discussing the change in the processes on earth, one examines their positions in the five phase sequence; when discussing changes in human bodies, one examines the interactions of *Qi Jiao*.[5]
>
> The Yellow Emperor asked: Where is this *Qi Jiao*?
>
> *Qi Bo* said: Above (six *qi*) and below (five phases) interact. Human beings reside within this exchange of *qi*. Therefore, it is said that above the natural pivot point the six *qi* patterns govern; below the

[3]The title of the *Dao De Jing*, the classic text of Taoism legendarily attributed to *Lao Zi*, is usually left untranslated, but it can be rendered in to English as "The Way and its Power". Again we have provided our own translations, but references to English translations of this work, as well as to the Chinese originals that we consulted appear in the References.

[4]*Qi* is generally rendered into English as 'energy' or 'energy flow'. It refers to a range of energy and material flows in nature, including those within the human body. *Qi* sustains the material structure of life, and it refers to the whole range of different forms of energy and nutrition, ranging from such subtle forms as air or breath to such substantial forms as food and drink. Because the term '*qi*' fails to map exactly onto the term 'energy' in the scientific sense we generally employ the term '*qi*'. The idea of 'flow' is often contained in the term "*qi*", thus denoting either "energy" or "energy flow," There would be some redundancy in the phrase '*qi* flow';" however, we sometimes employ '*qi* (flow)' in our paper when it is important to emphasise this aspect of *qi*.

[5]The *Qi Jiao* is the meeting point where sky *qi*, ie, the climate change patterns, interact with earth *qi*, i.e., the five phase transformation pattern; we leave the term untranslated.

> natural pivot point the five phase patterns govern; in the middle of this exchange, the human *qi* follows [the changes from above and below], and the ten thousand things transform accordingly. This is what is called *Qi Jiao*.
>
> The Yellow Emperor asked: What is the rising and descending movement of *qi* like?
>
> *Qi Bo* said: The rising and descending of *qi* is the alternating function [oscillation] of sky and earth.
>
> The Yellow Emperor asked: I would like to hear about how it works.
>
> *Qi Bo* said: When *qi* rises to the top, it descends; descending *qi* is heaven (*tian*). When *qi* descends, it rises again; rising *qi* is earth. Heaven *qi* descends, flowing to the ground; earth *qi* rises, flowing to the sky. *Qi* from above and below resonate; rising *qi* and descending *qi* cause each other's movement, therefore bringing about changes.

Health and disease is closely related to the changes of the six *qi* patterns and the five phase transformation. In the *Su Wen*, it is taken for granted that the human body transforms itself through the five phases during circadian cycles, annual cycles and the twelve and sixty year cycles; as well the six *qi* impact on the human body as they do on everything else in the world (the ten thousand things). Health requires the seasonal patterns of *qi* transformation between earthly processes and human bodies to occur harmoniously. When sudden changes occur in either earthly or body processes, e.g. weather change or emotional shock, then there will be disorders:

> The Yellow Emperor asked: When cold and humidity meet; drought and heat come together; wind and fire occur at the same time, is there a gap between patterns?
>
> *Qi Bo* said: One type of *qi* pattern might override another; and a weak pattern might become strong.

The oscillation of different patterns leads to the emergence of different types of transformation (*hua*), different functions and sudden changes (*bian*). Disharmony (*xie*) resides with sudden changes (*bian*) (*Su Wen* Ch. 68 (504)). Chinese medicine conceives of health and disease in terms of complex dynamics, with a diseased state being one of disharmony [Herfel, *et al.*, 2007]. The strategies to treat disease, like the strategies in a military campaign, are focused on discerning the dynamic configuration of various forces interacting within a highly adaptive system.

In ancient Chinese cosmology, within its constantly changing world, some changes are beyond human comprehension, whilst others follow various cyclical patterns. Thus, it is taken for granted that the natural world, including the human body, changes in only partially predictable patterns. Sometimes, it is possible for talented persons who are cultivated in following the natural course of things to recognize such patterns and act accordingly. The importance of this talent is stated in both *Dao De Jing* and *Zhuang Zi*[6]:

> To know the workings of nature, to know the workings of human beings, this is the ultimate knowledge. Knowing the workings of nature, one will be sustained by nature; knowing the workings of human beings, one can rely on what one knows to predict what one does not know. Then one could live to one's natural life span without being cut short in the middle of one's life. This is the vitality brought by knowledge. (*Zhuang Zi,* Ch.6)

The ability to discern and take advantage of such changes is a skill one cultivates through engaging with the rhythmic activity of the world. This is discussed in *Zhuang Zi* with such stories as "Carving up the Ox". As indicated in such stories, skills and knowledge are embodied in the practitioners of each craft. Usually, such knowledge is not possible to put into words, let alone develop into theories. Take for example, the following story from *Zhuang Zi* (Ch. 13):

> Duke *Huan* sat in his hall reading a book. The wheelwright *Bian* was making a wheel in the courtyard below the hall. Putting down his hammer and chisel, *Bian* went up the steps and said, "I venture to ask your Grace what works you are reading."
>
> "The words of the sages," replied the duke.
>
> "Are those sages alive?"
>
> "They are long dead."
>
> "Then what you, my Ruler, are reading are only the dregs and sediments of those old men."
>
> "How should you, a wheelwright, have anything to say about the book I am reading? If you can explain yourself, very well; if not, you shall die!"
>
> The wheelwright replied, "Your servant will look at it from the point of view of his own work. In making a wheel, if I proceed lightly, it is pleasant, but the work is not sturdy. If I proceed vigorously, it is exhausting and the joints do not fit. If the movements of my hand are neither too light nor too vigorous, the idea in my mind is realised. I cannot tell you how to do this in words: there is an art to it. I cannot teach this to my son, nor can my son learn it from me. So I am in my seventieth year, and I am still making wheels in my old age. But the ancients and what was not possible to convey are dead and have dispersed. Therefore what you, my Ruler, are reading is but their dregs and sediments!"

[6] *Zhang Zi*, the second famous founder of Daoism, is the legendary author of the text that bears his name. We have provided our own translations, but references to English translations of this work, as well as to the Chinese originals that we consulted appear in the References.

Discussions like this emphasise the embodied and tacit nature of knowledge and skills. In the Chinese philosophical tradition there is a presumption that such knowledge and skills, which cannot be completely verbalized or fully codified, play an important role in practice. However, this point is pushed to a deeper level in the ancient Chinese tradition. The Daoist philosophers emphasise that knower and known are parts of the same dynamical process. Thus there is self-similarity such that self-understanding and other-understanding becomes a spectrum with only differences in degree rather than kinds.

The dynamic understanding of the cosmos and human body reveals the following features in knowing and acting:

- Knowing is to be attuned to the dynamic rhythm of the process to be known; in other words, the knower has to cultivate his body to be sensitive in perceiving and responding to the subtle signs of dynamic changes in the processes of interest.
- Knowing requires all the sensory faculties that one has to become aware of subtle signs of changes or patterns of change, applying them to discern changes and patterns of changes; in other words, the knower has to train her vision, hearing, smelling, tasting, and touching faculties so that she would be able to detect subtle signs before changes are fully manifest.
- Knowing is the ability to respond to changes in step with the natural rhythms of the process one is trying to understand and manipulate, whether to activate it or deactivate it depending on the purpose of intervention.
- Knowledge consist of the capacity to perceive rhythm patterns of dynamic processes, and skill is the ability to act in accordance with the rhythm at the right time and place just as an orchestral member participates or withdraws from a symphony performance.
- The capacity of such perception and response has to be cultivated through practice over a long time period.

This conception of knowledge and skill are most clearly demonstrated in Chinese medicine. We next turn to applying a complex systems dynamics perspective to interpreting the Chinese medical tradition in order to illuminate its unique perspective in scientific terms.

2 ARTICULATING, TRANSLATING, INTERPRETING CHINESE MEDICINE

Many aspects of Chinese medicine remain opaque to the Western scientific mind. Douglas Allchin [1996] has argued that Chinese medicine works under a different paradigm than Western science and hence the two systems are incommensurable. There is no doubt that Chinese medicine emerges from a perspective that from the point of view of Western science appears exotic. However, neither Chinese nor Western medicine are static formal systems whereby the question of commensurability is decidable *a priori.* Instead these are living traditions. This presents the possibility that translations can be constructed through a dialogue between the two approaches. A big impediment to such an exercise is that Western science during the modern period has been dominated by reductionism: a project of explaining processes by breaking them up into smaller sub-processes. There are two hopes in this strategy. The first is that the sub-processes will be simpler and hence easier to explain than the whole process. And second, that there will be a simple way to combine explanations of sub-processes into explanations of processes. Reductionism works when the dynamics of the system to be explained is fully decomposable, most clearly when a closed form analytic representation of the dynamics is possible, with linear dynamics the degenerate case. This enterprise was so successful — it enabled us to put men on the moon — that many concluded that it was universally applicable. However, in the second half of the Twentieth Century the interest in complex dynamic systems started unravelling the "universality" of the decomposable dynamic paradigm. The discovery of chaotic dynamics was one such development. When the parts of a system are related nonlinearly, understanding the motion of the parts individually is not sufficient to understanding the motion of the ensemble. From our perspective, the major contribution that the complex systems dynamics approach has made is an articulation of holism that makes sense to one subscribing to reductionism.

Reductionism never took hold in China. Ancient Chinese science and Chinese medicine in particular, sought to capture the overall patterns of flow in the details of their complexity. Of course, just like in the Western tradition, cataloguing symptoms and correlating them to treatment strategies was important to the development of the art of medicine. However, the obsession with change in Chinese philosophy led to a medicine with adynamic conception at its core. Thus Chinese medicine models the *processes* of living systems. Health is seen in terms of the harmonious flow of living processes, diseases are disruptions to harmonious flow. Recognising the patterns of flow and reasoning in terms of organic dynamics is at the heart of Chinese medical diagnosis and treatment. The dynamic tendencies of Chinese medicine fail to sit comfortably within Western medicine where the reflex is to seek atemporal explanation if at all possible. However, complex systems dynamics is focused, from a Western scientific perspective, on the characterisation of natural processes and this provides a unique bridge between the two approaches. The key to successfully exploiting this opportunity is the construction of a pidgin to facilitate communication across the paradigms of Chinese and Western medicines. Adopting complex systems dynamics vocabulary to articulate Chinese medicine provides the core contribution to the project of making sense of Chinese medicine in scientific terms.

2.1 Living systems as dissipative structures

In Chinese medicine the human body is animated by *qi*, or flow of energy. Even as early as the compilation of, the *Su Wen* (ca. 200 BCE) we see this characterisation:

> *Qi Bo* said... the external manifestations of the roots of life may be called the establishment of *qi*, and when *qi* stops, transformation will cease. Therefore, each kind of *qi* has its own regulator (*zhi*[7])... (*Su Wen* Ch. 70 (553)).
>
> The Yellow Emperor said: When *qi* starts flowing, there will be life and transformation; when *qi* differentiates into different forms (supporting various organ systems), *xing*[8] emerges... This applies to everything (*Su Wen* Ch.70 (554)).

Thus energy flow is essential to the emergence of life. In fact, the ancient Chinese viewed the whole world as energetic processes in complex interaction. The transformations in living systems are the result of *qi* (flow). It is in the differentiation of *qi* that structure (*xing*) emerges. We even see a glimpse of a notion of self-organisation: "Everything arises through intrinsic energetic processes (*hua*), and the final outcome (*ji*) is due to extrinsic energetic processes (*bian*); intrinsic and extrinsic energetic processes interact with each other, giving rise to success or failure in life" (*Su Wen* Ch. 68 (504)).[9]

Furthermore organic processes possess a dynamic character:

> As soon as expiration and inspiration stop, all organic processes and transformation cease; and as soon as upward and downward movements cease, all established *qi* are at risk of collapse. Therefore, birth, growth, maturity, aging, and dying are possible only with the presence of expiration and inspiration; and birth, growth, transformation, harvesting, and storage are possible only with the presence of upward and downward movements (*Su Wen* Chapter 68 (505)).

Upward and downward movement hence constitute what we would call the ecosystem. The movements are related in some way to the flow of energy (*qi*). The ancient text does not make the nature of this relationship explicit: however, this is characteristic of complex systems where the relationship may be complexly dynamic rather than linearly causal.

The organism in Chinese medicine can be understood as a dissipative structure. In complex dynamics systems theory such structure is observed in the Bénard system, consisting of a shallow vessel of fluid; heating it from the bottom introduces

[7] *Zhi* is difficult to translate. It refers to form and regularity as well as the characteristics of a process or a system.

[8] *Xing* is the structure exhibited in living creatures. It emerges when constant energy and material flow is established and stabilised.

[9] On *bian* and *hua* see Hsu [1999, p. 113].

an energy gradient.[10] When the temperature gradient is low, heat is transferred upward through *conduction*. Molecules near the bottom heat up which means they start vibrating more rapidly than before. Since the bottom is solid there is a slight upward component to this increased vibration. However, the viscosity of the liquid dampens upward movement. Increased vibration at the bottom is transferred up through the fluid; layer by layer the molecules vibrate faster throughout the fluid. Eventually, the air over the fluid vibrates a little faster, as well. Thus the heat is conducted through the system.

As the temperature gradient is increased, conduction cannot transfer the heat efficiently enough, and viscous force is overcome. Conduction becomes unstable as the more energetic molecules rise faster than their kinetic energy is dissipated through viscous drag, giving them a net upward velocity. Cooler molecules above must get out of the way by sinking. Randomly sinking and rising is not an efficient way of transferring the heat upward, so above a critical threshold of energy gradient the flow of molecules spontaneously becomes coordinated. This organized rising and sinking is called *convection*.

A common configuration of the coordination in the Bénard system is in the form of fairly uniform hexagonal cells resembling a honeycomb when viewed from above in which the fluid rises in the centre of the cells and descends at the edges. The self-organised cell structure of the system forms a set of constraints intrinsic to it, while their exact cellular form also depends on initial and boundary conditions. Since the convection cells, driven far from equilibrium by energy throughput, form a stable self-organised pattern, they are an example of a *dissipative structure*.

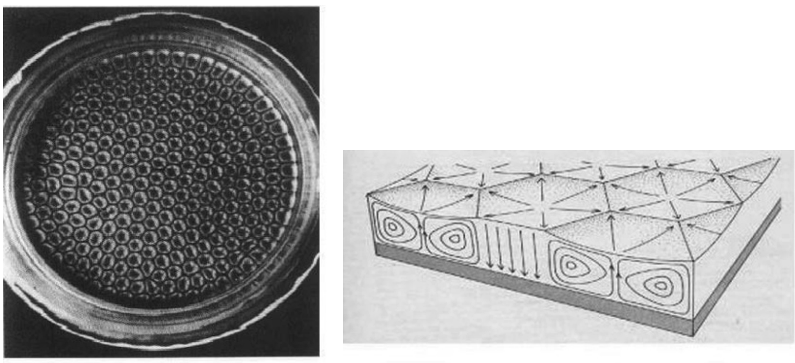


Figure 1. Bénard convection. Left: Bird's eye view showing cell structure. Freprinted by permission from Gregoire Nicolis, "Physics of far-from-equilibrium systems and self-organisation", in *The New Physics*, ed. Paul Davies (cambridge). Right: Schemaic view showing pattern of flow. Reprinted by permission of Alan D. Iselin.

[10]Bénard first studied this system experimentally, with Rayleigh providing the first theoretical treatment. See Nicolis [1989, pp. 316-20] for a general discussion in terms of far-from-equilibrium thermodynamics.

The appearance of the cells is deterministic; i.e. they will occur at the precise threshold each time a critical temperature gradient is crossed in repeated experiments with particular boundary conditions. The critical value can be predicted analytically from dynamical laws first proposed by Rayleigh.[11] Details of what happens after the onset of instability, e.g. where deviations from the generally expected hexagonal pattern occur (see Figure 1), aren't generally predictable and are held to be a function of chance microscopic fluctuation [Velarde and Normande, 1980].

An important point that must be considered concerns the ontology of *qi*. This issue has perplexed Western critics of Chinese medicine for decades. Classically *qi* is characterised as an energy-material flow in the body essential for life. *Qi* is contained in breath and in food. This raises several questions: Is this concept of *qi* the same as the scientific notion of energy? Does *qi* refer to a particular entity or is it a vague notion with no actual referent? Is there any evidence for the existence of such flow in the body? These ontological questions are important from a philosophical perspective, but for scientific understanding they are less important. What is important is to develop a coherent model of *qi*, establish a set of observable phenomena from the model, and correlate the theoretical phenomena with empirical observation. From this perspective the ontological issues are less important. Physics can model electricity as if it is a flow, even if it is not literally fluid. The model is a useful tool to characterise electricity. So too with *qi*. For the purposes of dynamic modelling, knowing that *qi flows* is more important than knowing what *qi is*. Classical Chinese medicine provides a model of *qi* that has been employed empirically for millennia. From a modern scientific perspective, more important than solving the ontological issue, the work that needs to be done is articulating the model mathematically, so that the dynamics of *qi* can be articulated and its empirical implications scrutinised.

In terms of complex systems dynamics, a dissipative structure is an emergent system stabilised by energy throughput. Viewed from this perspective, for Chinese medicine the entire world is constituted by dissipative structures in complex dynamic interaction. In particular, all organisms are dissipative structures. Their forms emerge via a process of development driven by energy throughput. Organisms are open systems. Thus matter and energy flows into and out of the system. More importantly, the energy is "degraded" in the process; the heat that is produced as the waste product of organic life has higher entropy than the food that nourishes it. It is this input of low entropy energy and the output of high entropy energy (energy throughput) which drives development and maintenance of the organism's structure. Another way of putting it is that developing and maintaining the organisation of an organism requires work. This work is performed through the degradation of the energy flowing through the organism. Without this energy flow the organism would never come into existence in the first place. Once the energy source is exhausted the process of decay rapidly puts the organism in equi-

[11] Chandrasekhar [1961] explores the standard model for the onset of Bénard convection in detail.

librium with its environment. The energy differential between which organisms exist makes them far-from-equilibrium systems. The fact that they are far from equilibrium enables the emergence of their dissipative structure.[12]

2.2 Patterns of flow

According to the traditional theories of Chinese medicine, the complex dynamics of the human body consist of material-energetic flows, including that of *qi* (energy), *xue* (blood), *jin* and *ye* (bodily fluids), *jing* (essence) and *shen* (mind). These substances flow between various "organ systems" (*zang-fu*) which constitute the subsystems of the body. It is important to note that the organ systems are not equivalent to the anatomic organs in Western biology. The *zang-fu* are characterised not structurally but functionally. Hence we talk of organ *systems* in Chinese medicine, whereas in western medicine *organs* are anatomical structures.[13] For instance, the functions of the lung *system* include being master of *qi*, governing respiration, descending and dispersing *qi* in the body, regulating water passages, governing skin and body hair and housing the corporeal soul (*po*).[14] Thus whilst some of the functions show some similarity with the anatomical structure of the lungs, e.g. governing respiration, others are more diverse with some functions spatially distributed throughout the body.

Misunderstanding the *zang-fu* is one source of difficulty in reconciling Chinese medicine with biomedicine. Because the names are familiar, e.g. heart, stomach, one not familiar with the Chinese medical tradition might assume that *zang-fu* are anatomical organs. This might lead one to search for the elusive structure, *san jiao*. Translated literally *san jiao* is "triple heater". Such a structure has never been found. But the *san jiao* is not an organ in the biological sense at all, but a functional representation of the spatial distribution of energy in the body. This sort of representation is common in the physical sciences. In particular, regulatory function can take place even though no particular element of the system is identifiable as a regulator. Garfinkel [1987] describes this situation for aggregation in populations of *D. discoideum*. Hooker *et al.* [1992] describe a case study in electrical engineering whereby regulation by means of a "virtual governor," distributed throughout the process, keeps voltage constant in an electricity grid fed by multiple generators. The *san jiao* is not unique in this regard; all organ systems in Chinese medicine are functional representations that the Chinese doctor discerns through relationships between inputs (e.g. acupuncture and herbs) and

[12]Note that everything we said about organisms applies equally to such structures as flames and tornadoes. Energy throughput maintains the structure of the flame and the tornado; once the energy gradient is exhausted both cease to exist. Hence while all living systems are dissipative structures, being a dissipative structure is not a sufficient condition for being alive.

[13]*Zang,* encompassing the liver, heart, spleen, lungs and kidneys, is translated as the "five *yin* organ systems." *Fu,* i.e. the stomach, large and small intestines, gall bladder, bladder and the triple burner are translated the "six *yang* organ systems." See discussion in [Porkert, 1974; Sivin, 1987]. Whilst we basically agree with Porkert's analysis of *zang* and *fu*, we find his rendering as "orbs" somewhat unnatural in English.

[14]`http://www.chinesemedicinesampler.com/theoryorgans.html`

outputs (e.g. signs and symptoms). For example, during acupuncture treatment, information about the state of *zang-fu* is discerned through pulse taking. The pulse is monitored throughout the treatment session to assess the progress of the treatment. The treatment is successfully completed when the pulses are balanced, indicating that the organ systems are functioning in harmony.

The organ systems are, like all entities in nature, interconnected energetic processes temporarily stabilised by energy throughput. Thus we can view bodies in Chinese medicine as dissipative structures that themselves are composed of dissipative structures. In order to characterise the dynamic interaction of the *zang-fu* Chinese medicine adopts the *wu xing* model, which in fact portrays the organ systems as a system of coupled oscillators.

2.3 Coupled oscillators

The *zang-fu* are modelled using the five phases (*wu xing*). The five *xing* are earth, metal, water, wood and fire. This has led to the confusing translation of *xing* as "element", akin to the traditional Greek notion. However the *xing* are not elements; rather they are five characteristic dynamic patterns found in nature. For example if a dynamic system is strong but flexible it has the characteristic dynamic of wood. More important than individual dynamic characteristics, however, is the interaction observed when the *xing* are coupled. This is construed in ancient Chinese cosmology in terms of *sheng* and *ke* relationships. Translated into complex systems terminology these terms mean driving and damping, respectively. From a complex dynamic perspective the *wu xing* system describes the interactions observed in the oscillations when the *xing* are coupled in a system.

The *Su Wen* (Chapter 70 (554)) discusses of the role of *qi* in establishing and sustaining life, where *Qi Bo* states, "Therefore, each kind of *qi* has its own regulator…each [system] has its subordinate (*shèng*), each [system] has its driver (*sheng*), and each [system] has a system that completes (*cheng*) it." This passage is a discussion of *zang-fu*. In order to model their dynamic relationships, in the "*wu xing* tradition," the various organ systems are mapped onto the five phases (*wu xing*), as shown in Figure 2.[15]

In the traditional schematic representation of the *wu xing*, the phases are explicitly construed in terms of damping and driving. In terms of dynamics, *sheng* is akin to a driving force and *ke* is akin to damping force. The *sheng* relationships are driving; for instance wood drives fire. The *ke* relationships are damping; for instance water damps fire, in this case literally. *Sheng* and *ke* lead to periodic cycling through the five phases. Hence, the oscillations of the *wu xing* are described in terms of the interaction of both types of force. Chapter 9 of the *Su Wen* describes the interaction between the seasons, *qi* and the human body in terms of the dynamic structure of the *wu xing*. It states, "…the five seasons *(wu*

[15]We follow John S. Major [1984] and Porkert [1974] in rendering *wuxing* as "five phases." See also Graham [1989, pp. 325-330]. For a good discussion of the evolution of the meaning of the term through early Chinese history see Major [1991].

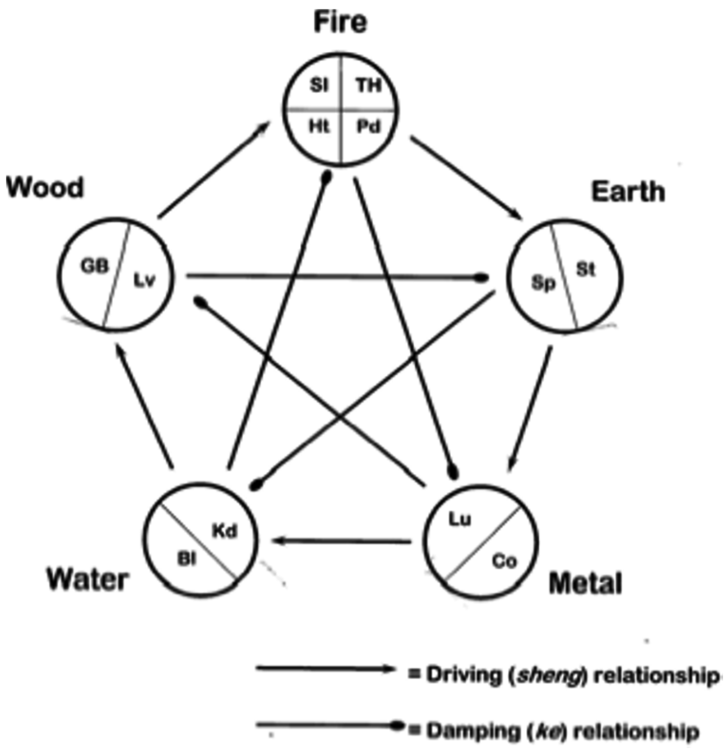


Figure 2. *Zangfu* modeled as wuxing. Ht = heart, SI = Small Intestine, TH = Three heater, Pd = Pericardium, Sp = Spleen, St = Stomach, Lu = Lung, Co = Colon, Kd = Kidney, Bl = Bladder, Lv = Liver, GB = Gall Bladder.

yun[16]) begin; they are like a ring with neither beginning nor end... The five *qi* are established in turn, and each has another that it subordinates so that it is natural for them to oscillate (*bian*) between abundance (*shèng*) and scarcity (*xu*)...." (*Su Wen* Ch. 9 (77)). Five phases cosmology describes dynamic transitions in natural phenomena.[17] In terms of nonlinear dynamics, each organ system has its own intrinsic oscillation pattern. The organ systems are connected together as a system of coupled oscillators.

The *Su Wen* (Ch. 9 (78)) continues by describing an example of the interaction between the seasons and the human body:

> The Yellow Emperor asked, How does this subordinating process work?
>
> *Qi Bo* replied: We can calculate the arrival of [the spring seasonal *qi*]

[16]Here the discussion is specific to the "five seasons;" however the *wu yun* map directly onto the *wu xing*.
[17]This is why the translation of *xing* as "element" is hopelessly misleading. Cf. [Graham, 1989, pp. 325-6].

> from the date of the spring equinox. If it arrives too early, then it [*qi*] is called "excessive." This results in it overpowering (*bo*) what ought not to be subordinate to it [in this circumstance] as well as overriding (*cheng*) that which is under its control.[18] This is called *qi* excess. ... If it arrives too late, then it [*qi*] is called "deficient." This results in what ought to be subordinate behaving inappropriately (*wang xing*), weakening that which it ought to drive, as well as overpowering that which it ought to subordinate. This is called *qi* deficiency. ... The seasonal *qi* ought to arrive as expected, but when they are not on schedule, the five seasons are not distinct. When this happens, harmful *qi* becomes internalised, and it cannot be stopped even by a skilful physician.

This passage describes, in terms of (over)damping and (over)driving of the five phase oscillation, how environmental factors can have a deleterious effect on the human body. If seasonal change does not occur in normal fashion, excess or deficiency of energy can cause disharmony in the organism's oscillation patterns. When the systems of the body oscillate harmoniously with the environment, and the internal oscillations within the body are harmonious, the organism is healthy. When there is disharmony, acupuncture, herbs, moxibustion, etc. can be employed to restore health. However, if the environmental disruption is so extreme that the five phases become indistinct (perhaps this is a reference to an oscillation that would be said to be "dynamically chaotic" in contemporary scientific terms), then no medical intervention can predictably restore harmony.

2.4 Dynamic conception of health and disease

Perhaps the most striking contrast between orthodox biomedicine and Chinese medicine is in their conceptions of health and disease [Herfel, *et al.*, 2007]. Whilst the contemporary medical model of disease is firmly rooted in Nineteenth Century science, which coincidentally was the age of equilibrium thermodynamics, the Chinese medical conception derives from ancient Chinese cosmology. Perhaps ironically, viewed from the perspective of complex systems dynamics, the Chinese model looks more progressive than biomedical orthodoxy.

On the one hand, health and disease, as presented in the *Su Wen*, are characterized in terms of dynamics: a subject is healthy when certain patterns of flow (e.g. *qi*) are in harmony. 'Harmony' in this context means that the patterns observed

[18]That is, the wood phase is overriding the earth phase. The five phases form a close-looped sequence with a specific order and specific relationships. The two fundamental relationships that hold between any two of them is the driving (*sheng*) relationship and the damping (*ke*) relationship, that is a) wood drives fire, fire drives earth, earth drives metal, metal drives water, water drives wood; b) wood damps earth, earth damps water, water damps fire, fire damps metal, metal damps wood. When the driving force and damping forces are in harmonious proportion to each other, the five phases follow each other with right order and pace and each phase arrives at its due time. When two forces are not in the right relationship with each other, oscillations can become out of phase.

reveal that the elements of the system stand in the right dynamic interrelationships. Of course these relationships are themselves dynamic: what is the right relationship itself evolves over time.

On the other hand, a contemporary anatomy and physiology textbook [Tortora, Grabowski, 1996, p. 10] discusses health and disease in terms of homeostasis:

> As long as the various body processes remain within normal physiological limits, body cells function efficiently, homeostasis is maintained, and the organism is healthy. When one or more components of the body lose their ability to contribute to homeostasis, however, body processes do not function efficiently. If the homeostatic imbalance is moderate, disease may result; if it is severe, death may result.

It defines disease in terms of "a pathological process with a definite set of characteristics in which part or all of the body is not carrying on its normal functions" [Tortora and Grabowski, p. 10]. This reflects a fundamental trait of the western tradition: health is characterized by stasis; departure from stasis is pathological. There are several physiological parameters that can be measured. Two paradigm examples that are relatively stable in healthy individuals are body temperature and resting heart rate (measured in beats per minute). Although these two parameters vary over time, in healthy individuals they remain regulated within a fairly narrow range. Of course, the processes that maintain stasis in these parameters are highly dynamic, and hence the values measured represent a dynamic equilibrium.[19] However, when discussing health and disease in Western biomedicine the focus is on the *number* not the process. This is a far cry from the Chinese medical notion of health as defined in terms of harmony. Harmony is not a 'number', but instead is a characteristic of a *process*. Consider the contrasting methods of pulse taking in the two traditions: in Western medicine the number of beats per minute is counted; in Chinese medicine the subtle patterns of blood flow are discerned. It is only by attending to the process as a whole that one can determine whether or not it is in harmony. Think of this in terms of a musical performance. Whether a particular voice is in harmony or not depends upon what the other voices are doing. And which particular note will be harmonious will change over time as the performance proceeds. Chinese medicine focuses on the dynamics of processes in order to restore and maintain health. Change is expected over time. Health is achieved, not when the numbers are stable, but when the processes are functioning properly.

From ancient times the Chinese tradition has been obsessed with *change*. So it is not surprising that the Chinese medical definition of health is in terms of *dynamics*. Furthermore, Chinese medical treatment does not focus on *restoring stasis* but *achieving harmony* in its dynamics. Without over-generalising, we can identify two aspects of this harmony. First, there must be internal coherence amongst

[19]Glass and Mackey [1988, pp. 4-6] discuss the role of homeostasis, pointing out that whilst homeostasis exists in physiological reality, it always does so within a larger physiological context that is essentially dynamic.

the flows within the body. Second, the processes that make up flows within the body must be in resonance with the flows that make up the environment. This is an emphatic statement of what it is to be healthy. Disease is defined negatively in contrast to health: it occurs when the flows are disrupted in particular ways. From this perspective, very different from the Western tradition, arose a unique approach to medicine.

The wisdom of the classic Chinese approach is made clear when viewed in light of contemporary developments in biomedical understanding of the nature of "dynamical disease" that, whilst not directly influenced by Chinese medicine, effectively follow in its footsteps. Recently, physiological and biomedical scientists have developed new models, theoretical and empirical, for understanding the dynamic character of illness. Building on the work of Reimann [1963], the first to document dynamic disease in western biomedicine, Leon Glass and colleagues did pioneering research much of which is summarized in their textbook [Glass and Mackey, 1988]. Arthur Winfree [1987a; 1987b] is another pioneer of this approach who applied methods he developed in his study of circadian rhythms to heart arrhythmia. Ary Goldberger [1997] and his colleagues have developed a number of tools for mathematically analysing heartbeat dynamics. Although the dynamic approach has been applied in other areas relevant to human health, those studying heart dynamics have been most prolific. In 1996, the journal *Cardiovascular Research* dedicated a whole issue to the topic of chaos in cardiology [Wagner, *et al.*, 1996]. Garfinkel [1983] reviews the dynamic approach for a wide range of physiological systems, including the effect of Parkinson's disease on the nervous system. Rapp [1986] is another pioneer of this approach for neurological systems. A good summary of how such research methods are applicable to behavioural psychology is provided by Heath's [2000] textbook. Anderson and May [1991] apply the dynamic paradigm to epidemiology. Daniel Kaplan and Leon Glass [1995] review mathematical techniques as they apply to dynamic disease. Pool [1989] provided an early news review of dynamic disease research in his article "Is it Healthy to be Chaotic?" The possibilities seem endless. Recently the dynamic approach has even been applied to dental caries [Featherstone, 2004].

Most interesting for our purposes though is the conception of disease emerging from the dynamic approach. In their textbook, *From Clocks to Chaos*, Canadian physiologists Glass and Mackey [1988, p. 172] define disease as follows:

> The normal individual displays a complex mosaic of rhythms in the various body systems. These rhythms rarely display absolute periodicity. ... Whether or not one interprets normal dynamics as chaos or some other type of dynamical behavior, it is clear that many pathologies are readily identifiable by abnormal rhythmicities.

If we replace "many pathologies" with "nearly all pathology" in the above passage we would have an elegant statement of the conception of disease operant in the *Su Wen* articulated in the terms of contemporary science.

3 MEDICAL TREATMENT AND TREATMENT VALIDATION

Once a complex dynamical perspective is adopted to interpret Chinese medicine, the contrast with the Western medical orthodoxy becomes transparent. This contrast affects Chinese medicine running from practice to the validation of practice. Not only does Chinese medicine have a different approach to diagnosis and treatment of disease, but also its novel approach calls into question whether orthodox biomedical research methods are most appropriate for the evaluation of Chinese medical intervention. The differences between biomedicine and Chinese medicine can be observed in the explanatory frameworks underwriting each. Whilst orthodox biomedicine follows a logico-causal model, Chinese medicine explains disease in terms of its detailed dynamics. In the following two subsections we will explore these contrasting explanatory frameworks and implications for research methodology. This exploration will be relevant not only for a better understanding of how research in Chinese medicine ought to proceed, but also will be relevant to dynamic disease explanations in biomedicine. Reflection on the relationship between the dynamic perspective and the aims and nature of explanation and methodology will also shed light on medical methods in general.

3.1 Explanation in orthodox and Chinese medicines

The dynamic approach to disease provides an alternative explanatory framework for medicine than the standard one offered by Paul Thagard in his *How Scientists Explain Disease*. In his book, Thagard [1999, p. 350] presents a hierarchy of disease categories (see Figure 3). On his account explaining disease is achieved by classification, that is, placing it into the appropriate category. For instance, influenza is a species of viral infection, which is one type of infectious disease.

'Dynamic disease' does not appear as one of the categories of Thagard's hierarchy. This is understandable as dynamical disease is not a separate category of disease. Disease dynamics transcends the categories in Thagard's hierarchy: diseases in all categories have their dynamics. For instance, influenza is quite a complex process. The spread of virus within the infected body follows certain patterns. The immune response follows certain patterns. The various manifestations of the disease, in terms of signs and symptoms, will have patterns at both the macroscopic (clinical) and microscopic (physiological) levels. In principle, these processes could all be studied *dynamically.* Nevertheless, on Thagard's account, explanation in orthodox biomedicine ignores dynamics of disease altogether. This omission impedes understanding of the Chinese medical approach to disease in western biomedical terms.

For Thagard, medical reasoning is formally logical. The diagnosis follows a conditional pattern: **if** *Helicobacter pylori* **then** stomach ulcer. Like all logical explanations, this sort of reasoning is atemproral. It is this atemporality that most significantly separates the approaches of biomedicine and Chinese medicine. Explanation in Chinese medicine is thoroughly dynamic. Hence the temporal

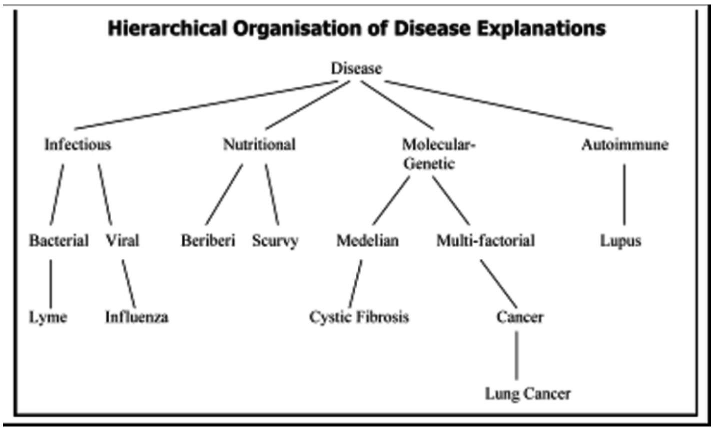


Figure 3. "How scientists explain disease". Reproduced by permission of Princeton University Press from Paul Thagard, *How Scientists Explain Disease* (Princeton: University Press, 1999), 35.

dimension is crucial to identifying, explaining, diagnosing, and treating disease. Thus reasoning in Chinese medicine is not strictly logical in the formal sense. Instead it involves complex pattern recognition and matching, as well as deriving conclusions from incomplete data. Of course the patterns involved change over time; thus formal logic is only tangentially relevant. In Chinese medicine it is not the cause of disease that is important; rather it is the dynamical pattern that needs harmonising that is the focus of the explanation. Thus in Chinese medicine it is typical that diagnosis is fluid and treatment varies over the course of the disease. Diagnosis and treatment are not standardised and they depend heavily on practitioner judgment. This reliance on judgment is a *strength* of Chinese medicine: it allows for rapid adaptation to what is understood to be a highly dynamic situation. But, at least until the advent of a theory of dynamical diseases and control theory for controlling complex dynamical processes, this variability and judgement seemed to stand at odds with the Western conception of a science-based medicine where the scientific paradigm requires exact reproduction of experimental and clinical demonstrations.

Of course, Western medical practitioners recognise that things change. It is just that their explanatory framework does not naturally accommodate reasoning about change. In some respects reasoning patterns dictate practice. Modelling disease as a uni-factorial, binary (either...or) condition facilitates straightforward remedies mainly in the form of single compound pharmaceuticals delivered in standardised dosages over pre-ordained courses of treatment. By contrast, in its attempt to capture the dynamic details of the situation, in order to exact more effective therapy, Chinese medicine models disease as multi-factorial interaction. Of course the difference between the two traditions is in degree. But, despite the work on dynamic disease cited above, in the mainstream biomedical tradition dynamic patterns do not play a significant role in diagnosing and treating disease. There is a reciprocal relationship between explanation and research in Western medicine. The explanatory framework serves the research methods employed in biomedicine. The atemporal explanatory framework underwrites standardisation which in turn justifies randomisation and orthodox statistical data analysis. We turn to these methodological issues in detail in the following section.

3.2 Dynamics and Chinese medical research methodology

In their paper on Chinese medicine Thagard and Zhu [2003] acknowledge that Chinese and Western medicines operate within different paradigms. However, in their view, whilst the incommensurability between them makes direct comparison of the two systems as a whole impossible, researchers can still use scientific methods to evaluate particular treatments based on Chinese medicine. It is always possible to conduct clinical trials assessing the effectiveness of treatment which need not invoke theoretical commitments from the either Chinese or Western approaches. In effect the diagnostic and treatment strategies are black-boxed: the strategies can be framed in any conceptual scheme. So long as outcome parameters are measurable, clinical trials can be designed to assess the *effect* of treatment independent of beliefs, practices and paradigms of those performing the treatments. If the effects of treatment are observable, from the point of view of a controlled clinical trial, it does not matter if the mechanism is thought to be chemical, mental or spiritual.

This is fine in principle. However, peculiarities of Chinese medical practice make designing clinical trials problematic. The methodological difficulties posed by Chinese medicine are well documented [Birch, 1997b; Hammerschlag 1998; Cardini, *et al.*, 2006]. For a given patient there is often lack of consensus between expert practitioners at either the level of diagnosis or treatment or both. Treatments are tailor-made for each individual. Treatment often varies from session to session. The same condition may be treated differently in different cases due to external factors. Health outcomes are assessed differently in Chinese medicine than they are in traditional Western biomedicine. To some extent the methodological issues have been superficially addressed in trial design, but seeing Chinese medicine from a complex systems perspective enables us to see that these difficulties are not mere superficial issues to be resolved through the refinement of clinical trial design. Instead these are deep issues that will only be properly addressed when we come to terms with health and disease in terms of the complex dynamic interactions within the organism as well as between the organism and environment. We will discuss each of these issues in turn, analysing them in terms of complex systems dynamics. This discussion provides a rationale in complex dynamical terms for taking the Chinese medical approach seriously, illustrating deep differences between the Chinese medical and traditional Western biomedical approaches to disease.

Several studies have documented inconsistency in diagnosis and treatment between Chinese medicine practitioners. For example, Hogeboom *et al.* [2001] found that when six acupuncture practitioners evaluated six patients the result was twenty different diagnoses yielding treatments involving 65 different acupoints. Other studies have reported similar results [Birch, *et al.*, 1999; Kalauokalani, *et al.*, 2001; Sherman, *et al.*, 2001; Zhang, *et al.*, 2004; Zhang, *et al.*, 2005]. One might conclude that these results reflect a problem with Chinese medical practice. However, if we consider Chinese medical practice from the perspective of complex systems dynamics we can see why such variation might be expected. We have already discussed the role that the *zang-fu* organ systems play in Chinese medical diagnosis and treatment. Using the *wu xing* as a coupled oscillator model of the interactions of the organ systems, each individual practitioner reasons dynamically to integrate observed signs and symptoms into a diagnosis. The *wu xing* model is thus a framework for constructing such disease explanations in terms of dynamic patterns of disease. As such it does not determine a particular interpretation of signs and symptoms. Nor does it dictate a particular course of action. A variety of equally valid approaches are available within the Chinese medical paradigm. And each may be equally effective. There is no more reason to believe that there is a particular strategy for curing a disease than there is for believing there is a particular strategy for winning a hand of bridge. Multiple pathways yield multiple strategies.

Another idiosyncrasy of Chinese medicine is the fact that treatment of the same patient with ostensibly the same complaint often varies from session to session. Viewed with a prejudice for standardised procedures, this may look like arbitrary inconsistency. However, a Chinese medical practitioner has perfectly rational motives based on beliefs about disease dynamics for eschewing standardisation. Since from the perspective of Chinese medicine the human body is a complex dynamic process its dynamics will be in constant flux. There will be several factors influencing this dynamics. A few examples are natural circadian fluctuation, environmental factors and response to treatment. Performing acupuncture, for instance, at the same points as prior treatments will only make sense if the patient is in the same state as they were before.The practitioner, alert to subtle changes, always has to be prepared to alter the course of treatment, even within a treatment session. The state of the patient is constantly monitored and assessed. An expert practitioner would never deliver a treatment preordained by standardised procedures.

Finally, Chinese medicine possesses a unique system of categorisation of disease. Biomedicine categorise diseases in terms of their causes [Thagard, 1999]. In Chinese medicine disease *is* dynamic disharmony. This provides a continuous, multi-dimensional, multi-parameterised characterisation of disease conditions, a characterisation for which there need be no simple distinguishable causes; rather, preceding conditions of the whole system play a generative role, as typically happens in dynamical systems. Thus the conditions of both the health and disease evolve over time. For the uninitiated, the shorthand for characterising syndromes, for instance, 'stomach *qi* deficiency' or 'liver *yin* deficiency' may belie the dynamic character of Chinese medicine, but it is important to realise that syndrome names describe, not fixed conditions, but patterns of flow within the body. The issue is deeper than there being no simple correspondence between Chinese and biomedical disease characterisations. The aim of restoring harmonic balance to biological rhythms makes Chinese medical practice unique.

We can thus see that Chinese medicine is extreme in its insistence on taking the details of highly dynamic situations into account when treating disease. This may make one accustomed to reasoning via linear deterministic causal pathways uncomfortable. However the human body is a complex system and there may well be multiple valid approaches to a particular problem. In fact, we face such situations commonly in everyday life, and even when the situation is not terribly complex. Consider the task of navigating by car across a busy city. There may well be dozens of equally efficient routes. Which one is best, at a particular time, will depend on many variables. Depending upon many details of the situation, traffic density, time of day, weather conditions, etc., one route may turn out to be better than another. However we rarely have enough information to make the most efficient decision. Instead we incorporate imperfect information, relying on past experience and following hunches to reach our destination in a reasonable time. The situation is the same for a Chinese medical practitioner. Practitioners may disagree, and still have valid approaches to a particular case. There may be several equally valid routes to curing a particular disease. In fact, given the complex dynamic nature of the organic world one would be inclined to view a belief that in a medical situation there is a particular correct diagnosis and optimal treatment upon which all (competent) practitioners ought to agree as a chimera.

Both Western biomedicine and Chinese medicine start from the assumption that health and disease are complex phenomena. Disease is always multi-factorial. From a dynamic perspective various processes in the human body are interconnected in complex ways; a problem in one part of the body cascades through the system creating a dynamically complex situation. However, the response to this complexity is different in the two traditions. In Western medicine focus is on the major cause of disease. Of course, there will be minor factors that affect how a disease is manifest in a particular situation, but standard treatment focuses on major causes in the belief that in most cases the effect of mitigating factors will be small enough not to effect the outcome of treatment targeting the major cause. This belief is reflected in research methodology. The purpose of randomisation in clinical trials is to increase the likelihood that confounding variables affecting treatment will be evenly distributed between treatment and control groups. Reflecting the core (and frequently unspoken) beliefs of Western medicine, clinical trial methodology focuses away from dynamical complexities. In a clinical trial nuances are randomised away. Chinese medicine takes these nuances into account. Tailor-made treatments are the result of the Chinese medical tradition taking small details into account. Every attempt is made to tailor treatment details to the patient's dynamical situation. So, two patients with a liver *qi* deficiency may be treated differently depending upon the respective states of their kidney or heart systems. But additionally treatment may take into account the time of day, season or weather, as well as the emotional state of the individual. Of course these factors often informally influence a Western medical doctor's treatment. In Chinese medicine such factors are systematically taken into account. Hence the rejection of standardised treatments.

In sum, viewing Chinese medicine from the perspective of complex systems dynamics allows us to articulate a rationale for its peculiarities. Empirically, it is necessary is to embed clinical evaluation in a fully dynamic framework which can provide adequate methods for Chinese medical research. What are inconveniences from the perspective of designing clinical trials of Chinese medicine may well be instrumental to the success of the practice. Thus Chinese medicine research calls many orthodox methodological assumptions into question. Complex systems dynamics allows us to illuminate where these assumptions fail. In particular, trials of Chinese medicine need to be designed to take into account the dynamic approach. Tailored treatments belie standardised diagnosis and treatment, making randomisation problematic. This in turn undermines the applicability of standard statistical evaluation. Furthermore, Chinese medicine aims at restoring dynamic harmony to its patients. Thus the outcome parameters for Chinese medical research will differ from those proposed in a biomedical framework.

4 THE ROLES OF PRACTITIONER JUDGMENT AND INDIVIDUALISED MEDICINE

Thus far we have emphasised the roles that individualised and flexible treatment strategies play in Chinese medical practice. We have argued that from a complex dynamic perspective, it makes perfect sense to adopt such strategies vis à vis promoting health in such a complex system as is the human body. We have also emphasised that these typical practises of Chinese medicine have serious implications for research methodology. However, such strategies obviate the standardisation that forms a significant facet of the scientific basis of biomedicine. In the absence of the possibility of standardisation, practitioner judgment becomes paramount to the success of Chinese medicine. In this section we explore the foundations of practitioner judgment and individualised medicine in terms of situated cognition and embodied knowledge.

4.1 Situated cognition and path-dependence in Chinese medicine

Chinese medical practitioners see versatility of response to a dynamical situation not as a deficiency, but as the key strength of the discipline. We will explore this issue in more detail by considering how situated knowledge evolves as a complex path-dependent process. In coming to understand any complex system, path dependence will play a crucial role, and as such the cognitive approach to such systems will be context-sensitive. Economist Bart Nooteboom has applied the principles of complex systems dynamics to the process of knowledge acquisition. His thesis is that knowledge acquisition, and hence the knowledge individuals possess, is path-dependent:

> Higher level, complex concepts, including categories of perception, interpretation and evaluation, are formed from modular elements that in turn develop from action, which is interaction with the physical and social environment. Hence our ability to perceive and to learn, in the sense of making new cognitive functions by reconnecting parts from existing ones, depends on past development. Hence, people will perceive, interpret and evaluate differently to the extent that they have developed in different contexts, along different paths. [Nooteboom, 1997, p. 68]

Furthermore, in Nooteboom's [1997, pp. 68ff.] view this path dependence applies equally to cognition in social organisations. Nooteboom's work focuses on human action. In particular, he sees an intimate relationship between action and cognition. In providing a cognitive framework for his model of knowledge acquisition he aligns himself with a "school" of cognitive science he calls "situated action" (SA). In Nooteboom's [1997, p. 61] words,

> According to SA the formation of patterns of neural activity is so much a matter of interaction with action that the logical sequence is in doubt. Every action to some extent modifies the patterns of neural activity, to such an extent that one could just as well say that action precedes neural patterning. It is meaningless to conceptualise the latter as being independent of context... Learning and action go together in the construction of neural patterns of activity.

Nooteboom provides no references to members of the SA school. Nevertheless, Edwin Hutchins [1995] has applied the principles of what he terms "situated cognition" to the human activity of navigation in his *Cognition in the Wild*. Although Hutchins does not speak of "path dependence" his example provides a fruitful case study of this concept.

Traces of the path dependence of the knowledge within navigation are embedded in nomenclature. The device for measuring the speed of a vessel is the "log." This name derives from the practice of throwing a piece of wood overboard and timing how long it takes the vessel to move away from it. "Starboard" and "port" derive from a time when vessels were steered with an oar ("steering board") attached to the right (looking forward) side, which had implications for how the vessels were docked in port. Hutchins points out that there are always several ways to solve the same problem in navigation. For instance, dead reckoning[20] computations of speed need to be carried out continually. Knowing the algebraic expression $D{=}RT$ allows one to compute speed (R) from the knowledge of the distance covered (D) since the last computation (T) [Hutchins, 1995, pp. 145ff.]. However, the computation takes many steps, each prone to error. Whilst algebra provides a general abstract tool that can solve a wide-range of problems, in practice, such a device as a nomogram dedicated to solving the problem at hand is preferable.

[20]By the way "dead reckoning" has nothing to do with death. It is shorthand for "deductive reckoning" = "ded. reckoning" which when spoken sounds like "dead reckoning"!

An even more direct way of calculating speed is the three-minute rule. It is based on the relationships between time, speed and distance and the nautical mile and the yard. The nautical mile is interesting enough in itself. A nautical mile equals one sixtieth of a degree latitude. The Mercator projection used in constructing nautical charts is designed to minimise distortion of these distances when the globe is projected onto a flat surface. Hence the latitude scale running along the side of the chart serves as a distance scale as well. It turns out that one nautical mile is equal to 2000 yards. We all know that there is 60 minutes in an hour. So if we measure the distance in yards a boat has travelled for three minutes it is very easy to see the average speed in knots (1knot = 1 nautical mile per hour) during the interval. Three minutes is 1/20 of an hour so the speed in knots equals the distance in yards divided by 100. So the navigator just moves the decimal point left two places to calculate speed from distance covered. For instance, if the boat has travelled 523 yards in three minutes then the speed is 5.23 knots.

This is crucial for dead reckoning. One needs to know the position, speed and direction the vessel is travelling in order to predict where it will be in the future. Position fixes are continually plotted by the navigator. The direction can be anticipated by the compass, the state of the helm, the current, etc. Speed must be calculated from the difference of the prior two position fixes. When near land,[21] the convention is to take position fixes at three-minute intervals. Hence the "three-minute rule" makes the determination of speed practically automatic. These practical considerations have resulted in the 'lock-in' of the three minute rule, the yard and the nautical mile, that is, their entrenchment in nautical methods of navigation so that there is resistance to introducing methods that depart from using them.

What a fuller treatment of Hutchin's subject than we have room for here would reveal is a rich history of the path dependent development of navigation. For instance, Hutchins [1995, ch.2] provides an elegant account using the principles of situated cognition to make sense of the Micronesian solution to the problem of navigation. The point here is not to recount that history, but to show how the concepts of path-dependence and lock-in illuminate the cognitive dimensions of human action within the context of history. The navigational practices we have today are the result of path dependent processes. Each procedure has a rationale that makes sense within its historical context. The definition of the nautical mile is arbitrary insofar as it could have been defined, with equal convenience, otherwise. The distance that makes up a yard was developed in a different context, and adopted in the marine context. What look like, out of context, an arbitrary set of conventions make sense as a whole. Practices evolve as a series of progressive constraint, each historical moment in the process limiting certain possibilities whilst enabling others.[22] Once evolved the whole system hangs together. A change in

[21] Land is the main source of danger for a boat. No vessel has ever run aground in the open sea!

[22] The first time we encountered this account of practice was in Dyke [1988]. Chapter 6 provides an account of bridge conventions in these terms without reference to "path dependency".

one element has implications throughout.

Nooteboom [1997, pp. 63-4] advocates modelling our practices on our minds and language:

> Scripts can be represented as lines, indicating the action sequence, with nodes, which function as variables that allow for the substitution by diverse items (such as tools, materials, and the like). This seems to be the appropriate manner to conceptualise reference: a fitting into scripts.... This implies allowance for dubious substitutions, which may, when successful, turn out to be innovations.... On the other hand, for a minimum of stability in practice, some obstacle to unorthodoxy must be present. This brings us back to an argument from Kuhn: some conservatism in maintaining incumbent practice is rational.... But too much conservatism inhibits adaptation to changed conditions. Survival requires a balance of, on the one hand, routine, habit, conservatism, continuity, and on the other hand, adaptability, innovation, shift. What the right balance is depends on the volatility of the environment. Our minds and language are masters at this.

In Chinese medicine dynamics is central to cognitive practices. Decision making is situated in a highly dynamic process requiring cognitive skills that are fully temporal in nature. Viewed from an atemporal perspective decision making in Chinese medicine may look *ad hoc* and arbitrary. But from within the context of a full understanding of cognition situated in the fully dynamic framework of Chinese medical practice such fluidity is a reasonable approach to health and disease. Just as reasoning in navigation must, in order to be rational, remain fluid and sensitive to context, so too Chinese medicine takes the temporal dimension into account. It is the complex interplay of body, environment and disease that dictates diagnosis and treatment in Chinese medicine.

4.2 Embodied knowledge in Chinese medicine

In Chinese medicine knowledge embodied in both doctor and patient is taken for granted as reliable source for diagnosis and therapy. The doctor uses all his senses to collect information from the patient, as Bray [1995, p. 238, *emphasis* original] describes:

> There were differences in the most favoured diagnostic technique according to school or period, but basically diagnosis comprised four essential methods involving all the senses. The Chinese doctor looked at the patient's complexion, *smelled* (and tasted) his breath and bodily odours, *listened* to his account of his disorder, and *touched* his wrist during the complex procedure of taking the pulse.

Furthermore, what Chinese physicians were seeing, smelling, listening and feeling is much more sophisticated than the observations of a layperson. When a Chinese physician examines the complexion of a patient, she sees the expression, the colour and the hue of the patient's face "in much the same way that a gardener judges the condition of plants, obvious signs of a plant in faltering heath include limpness, shrivelling, and desiccation" [Kuriyama, 1995, p. 228]. The colour, the hue, especially the lustre "offers the subtlest and most revealing index of vitality" [Kuriyama, 1995, p. 229]. Kuriyama argues that this botanic analogy of human body goes much deeper than a simple tool to illustrate some superficial similarity. A Chinese physician conceives human facial colour and hue botanically. For example, the *Su Wen* (Ch. 17; cf. [Kuriyama, 1995, p. 228]) states: "The five colours, is the flower (or flowering) of Qi." And "The heart gathers together the essences of the five organs ... The flowering visage is the bloom [of this essence]".

Kuriyama [1995, p. 228] further argues,

> The analogy to plants underlies, for example, the relationship between the various organs and the parts of the body that each governs: when the spleen ceases to nourish, the flesh becomes soft and the tongue wilts (*wei*); when the kidneys cease to nourish, bones become desiccated (*ku*). ... According to *Neijing*, the source of vital breath (*shengqi*) serves as the body's stem and roots; when the roots are severed, the branches and leaves wither.

Like the gardener who is able to detect the state of a plant by inspecting the colour and lustre of its leaves, the Chinese doctor is able to detect a patient's health state by examining the expression and facial colour of the patients, listening to the patient's speech, enquiring on the patients behaviour in eating, sleeping, excreting and palpating the *mo* (pulse) and other part of the body. An experienced gardener's ability to detect the state of a plant takes years to develop and is often impossible to articulate. The same is true with an experienced Chinese physician's ability to detect a disorder in a patient by a certain colour of the face, a particular quality in the pulse, a slight change in the colour and texture of the tongue, etc. Such extraordinary perceptions take years to develop even for the most talented person.

Such tacit/embodied knowledge and skills are illustrated by a diagnostic technique called "*quemo*", ie, palpitating the pulse at both wrists to detect the dynamic state of various organ systems in the body. What a Chinese physician's fingers are trying to feel at a patient's wrists is whether the pulse feels slippery or rough, floating or sunken, large or small.

> The slippery *mo* 'come and goes in slippery flow, rolling rapidly, continuously forward'.... The rough mo is the opposite: it is 'thin and slow, its movement is difficult and dispersed, and sometimes it pauses, momentarily, before arriving'; one has the impression of flow made rough by resistance, struggling forward laboriously, instead of in a smooth, easy glide. 'like sawing bamboo,' says the *Mojue*. [Kuriyama, 1999, pp. 50-1]

Such description of *mo* is clearly not a representation of what the *mo* is *like*, but how the *mo feels* at the finger tips of the Chinese physician. Although such description can be helpful for a novice to grasp how a *mo* should feel, it is certainly not enough. There are just too many perceptive and cognitive gaps that separate a description of a sensation from the sensation itself. For example, would a slippery *mo* feel the same for an artist who plays a violin and a bricklayer whose handles the brick all day? Would it feel the same for a person who never heard of Chinese medicine and is told to feel the slippery *mo* for the first time and a person who has studied Chinese medicine theory for a year then is told to feel the *mo* for the first time? It is possible that each experiences feels quite different. It is clear that even with a detailed description of how a slippery *mo* would feel, a novice would still be at lost without an experienced physician at her side when she feels a slippery *mo* for the first time telling her that she is feeling a slippery *mo*. A novice will need to feel many slippery *mo* to train her fingers how a slippery *mo* feels in order to grasp it, and many more slippery *mo* and *mo* of other qualities to distinguish a slippery *mo* from a rough *mo*, a floating slippery *mo* from a small, rough *mo*, etc.

Discerning the subtle shade and hue of facial colour and facial expression, the texture and colour of the tongue, detecting the slight changes in the pulse are not unique to Chinese medicine; biomedical practitioners conduct similar observations, although such activities play a more marginal role in diagnostic and therapeutic reasoning. We shall discuss some research in tacit/embodied knowledge and skill in the West in the following section. Unique in Chinese medicine is that tacit/embodied knowledge and skills are based on philosophic assumptions about the human body. Both physician and patient are parts of a cosmic dynamic process; tacit/embodied knowledge and skills are cultivated to discern the subtle signs of dynamic changes in the patient's body so that the doctor would be able to restore physiological dynamics out of sync with cosmic processes, the most obvious of which are circadian and annual cycles.

4.3 Research on tacit/embodied knowledge in the West

Literature on tacit and embodied knowledge has increased during the last three decades in the West. Previously, embodied/tacit knowledge and skills were not a major topic of philosophical inquiry. However, since 1980s, sociologists of science and technology realized the critical roles played by tacit knowledge and skills in scientific and technological research, teaching, and knowledge production with significant results. Insight into the various forms of embodied/tacit knowledge and skill in biomedical practice such as prenatal care, occupational therapy, nursing practice, surgical operation, has been provided. Research into knowledge management and research policy has also contributed to the understanding of the critical role of tacit knowledge and skill in technological design and knowledge as well as skill transfer and transmission.

Sociologists of science and technology have adopted Polanyi's concept of tacit knowledge. They define tacit knowledge as:

> Knowledge that is "imperfectly accessible to conscious thought" [Nelson and Winter, 1982, p. 79] yet will show its presence in the success

> of our performance using that knowledge. Using Polanyi's own famous example, he argues that a skilled swimmer are often unaware of the fact that keep one's lung full of air will enhance one's buoyancy while swimming yet will automatically do this while swimming.
>
> Knowledge that is difficult or impossible to communicate to others with symbolic forms such as spoken/written language, or picture to successfully execute an action, even when full self-awareness is achieved. That is, one has to perform the knowledge or skill in order to convey it to others. This is the knowledge that is shared by the master with his apprentices or the "learning by doing" kind of knowledge where observation, imitation, correction and repetition are the essential modes of learning [Polanyi, 1966]. In other words, tacit knowledge could only be acquired and communicated through experience. [Gertler, 2003, pp. 77-8]

For sociologists of science and technology, the tacitness of knowledge has significant implications for our understanding of process in scientific and technology research, training, and learning, and most importantly on the decision making process whereby tacit knowledge functions implicitly. The most distinctive characteristic of such tacit/embodied knowledge and skill is its implicit nature as convincingly argued by Polanyi [1966, p. 126]:

> The analysis of a skilful feat in terms of its constituent motions remains always incomplete. There are notorious cases, like the distinctive 'touch' of a pianist, in which the analysis of a skill has long been debated inconclusively; and common experience shows that no skill can be acquired by learning its constituent motions separately. Moreover, here too isolation modifies the particulars: their dynamic quality is lost. Indeed, the identification of the constituent motions of a skill tends to paralyse its performance. Only by turning our attention away from the particulars and towards their joint purpose, can we restore to the isolated motions the qualities required for achieving their purpose.

Research on embodied knowledge in medical practice is more specialized. Like sociologists of science and technology, medical researchers highlight the inexplicability of tacit knowledge. Schell and Schell [2008, p. 71] tell an illuminating story of gaining embodied knowledge as occupational therapists:

> I remember being told that spasticity was a function of abnormal muscle tone, and could be detected by rapidly stretching the affected muscles. If the person had spasticity you would detect a "clasp-knife type of catch and release." I confess that I had no idea what that meant. It wasn't until I started working with individuals who were recovering from brain injuries that I actually felt spasticity (i.e., used my own sense of touch and proprioception to gauge the muscle tension and release of that tension). Then I understood what was meant. Over time

> I learned to feel the difference between spasticity, rigidity, and other muscle problems, as opposed to tightness in joints capsule. This in turn led me to think about activity possibilities and limitations in different ways as I worked with the person to increase his or her functional abilities.

Such embodied knowledge involves all sensory faculties.

> For instance, therapists working with individuals with persistent mental illness will notice if a client has body odour or smells as if he is not keeping up with his personal hygiene. This becomes a clue to explore what is going on in terms of his bathing and laundry routines. Therapists working with young children with sensory processing problems and attention deficit syndromes will see the child move his body, hear the change in the child's rate and quality of speech, and interpret from these cues that the child is becoming agitated by sensory overload. This prompts the therapists to introduce calming strategies or to teach parents how to manage the child's environment to avoid sensory overload. [Schell and Schell, 2008, p. 71]

They also emphasise the embodiment of the tacit knowledge, i.e., the existential aspect of knowledge, that is, how the body/mind experiences it when acquiring such knowledge. For example, for nursing practice, the embodiment of the knowledge manifests in patients unique understanding of their bodily function. Wild [2003, p. 172] noticed that patients with a urine catcher linked recurring tract infection with disrupted urine flow and pay close attention to their urine flow and drinking more water to insure urine flow.

Medical anthropology also explicitly reflects on the significance of embodied knowledge in medical practice. Gordon [1988, p. 269] points out that "much of clinical and practical knowledge is 'embodied' knowledge — knowledge sensed through and with the body. This includes senses of sight, sound, touch, smell". Cassell [1998, p. 32] argues that being a surgeon is fundamentally "a physical proficiency.... [L]ike sports and the performing arts, surgery is based on body learning, body knowledge. ... One masters these skills by doing, not talking".

However, although the embodied knowledge is acknowledged by such researchers of Western medical practice, the significance of such knowledge is marginalized and ignored in accepted clinical decision-making processes. Instead, clinical decision-making is believed by practitioners to be exclusively based on evidence acquired through instruments, i.e., data from biochemical tests. Randomised-controlled trials and truth-preserving logical reasoning justify such techniques. Ellingson [2006, p. 301] argues that the mind-body dichotomy means that knowledge has to be disembodied, that is, the body of the researcher should not be involved in any way in the process of generating knowledge, let alone should it become the main source of information for clinical decision making:

> Research reports typically are written following strictly social scientific or medical conventions, in which the author's agency is obscured via

> passive voice (e.g., "the data were collected') or represented through a sanitized "I," who reports having taken actions without describing any details of the body through which the actions were taken.

Information acquired via a human body's sensory faculty is rejected as evidence because it is regarded as subjective and unreliable. Hunter [1991, p. 52] explains the underlying rationale thus:

> The aim of medical discourse is always to eliminate or control the purely personal and subjective, whether its source be patient or physician, so that the physical anomalies that characterized illness can receive the attention their successful treatment requires. Illness is a subjective experience, and the examining physician faces the task of translating it, locating the malady in the medical universe and conveying its characteristics and their meaning to others who know the medical language well ... All case presentations seek to turn an individual physician's interpretation of the patient's subjective and private experience of illness into an objective, scientific — or, from another viewpoint, a reliably intersubjective and medically recognizable — account of disease.

In other words, information that supports a diagnostic and therapeutic decision ought to be scientifically reliable and objectively measurable. As shown by sociologists and medical anthropologists, embodied knowledge and skills play an essential role in clinical practice. However such knowledge and skill are not sufficient to *justify* diagnostic or therapeutic decisions. The medical profession must rely on test results that do not contain 'purely personal and subjective' information.

4.4 Dynamic characteristics of tacit/embodied knowledge

A significant difference between the conceptions of Chinese and Western medicine is that in the Chinese tradition there is a dynamic perspective which makes tacit/embodied knowledge and skill indispensable for its practice. As discussed in above, Chinese medicine conceives that the processes of the body have to be in sync with cosmic process in order to be healthy. As stated in *Ling Shu*,[23] subtle changes in a person's facial expression and colour precedes other signs of physical pathology:

> The Yellow Emperor asked: when a pathogen attacks a person, how is the disorder manifested? *Qi Bo* replied: when a pathogen attacks (a patient suffering deficiency), the patient shivers with chill; when seasonal changes injure a patient, the disorder is mild. It appears in the patient's facial colour and expression first. There is no physical

[23]The *Ling Shu* is the classical text on acupuncture, forming, along with the *Nei Jing* the major part of the canon of Chinese medicine. We have provided our own translations, following the original Chinese, Ma [1998].

> change (known to the patient). It seems there now and not there then; it seems gone now and returns then; it seems discernable now and undiscernible then. It is hard to know the details of the disorder. [Ma, 1998, p. 27]

The role of the Chinese physician is to detect such subtle signs of change that signal disharmony between the processes in the human body and the cosmos. The embodiment of knowledge and skills in Chinese medicine are not only embodied in the sense that they have to be learned by the body, but also they are embodied in the sense that Chinese medicine practitioners' bodies share the same characteristic dynamics with those of their patients. Practitioners' bodies are instruments to measure the patients' bodies, and they are cultivated to develop particular sensory sensitivities to pick up subtle signals revealing the dynamic state of a patient's condition. Kuriyama's [1999, p. 51] analysis of the differences between the perceptions of pulse by ancient Greek medicine and *mo* by classic Chinese medicine demonstrates clearly the dynamic perception of the *mo*:

> Instead of the vertical rise and fall of the arteries toward and away from the body surface (as perceived by Greek doctors), Chinese doctors sought to feel the horizontal streaming of the blood and breath parallel to the skin. The *Suwen* thus glosses slippery and rough in terms of the opposition of "following" (*cong*) and "resisting" (*ni*), and the *Lingshu* relates both pairs — slipperiness and roughness, and *cong* and *ni* — to the lessons of hydraulic engineering. To *cong* was to be in the flow, or to go with the flow; to *ni* was to go against it. The eagerness to ascertain slipperiness or roughness mirrored the belief that life flowed.

Such dynamic perception of human bodily processes and conditions make it necessary to integrate different, but complementary, diagnostic and therapeutic strategies in order to bring out of phase bodily processes into sync with other processes inside and outside the body. For example, facial colour and pulse reading are often used together to check whether different organ systems are in harmony or how serious a condition is. The *Ling Shu* [Ma, 1998, p. 28] states:

> The resonation among the facial expression/colour, the pulse, and the skin colour at the wrist is like that of the drum, drum stick and the sound of drum; [the drum stick beat the drum followed by the sound], one could not exist without the other. ... [all three of] facial expression/colour, pulse, and the skin of wrist has to be in harmony [to be healthy].

In a current Chinese medicine diagnosis textbook, there are several different streams of diagnostic framework, including those based on eight principles, four sectors, z*ang-fu*, etc. These different streams do not exclude each other in terms of differentiating patterns of disorder, neither do they recommend different treatment strategies; instead they offer interpretations of dynamic bodily processes from various perspectives which are often complementary. For example, eight principle diagnosis focuses on abnormal changes in terms of excess and deficiency (*shi* and *xu*) in physiology, body fluid retention or deprivation (*shi* and *zao*), etc., while the four sectors focus on energy, material and bodily fluid circulation (*qi, xue, jin* and *ye*). The findings from these different diagnostic strategies present a multi-faceted picture of the condition of a patient. Approaching disorder from multiple perspectives often sheds light on signs that might be obscure from one perspective but apparent from another.

The dynamic perception of the human body implies individualised and context sensitive treatment strategies. Like all complex adaptive systems, the human body is capable of exhibiting path-dependence and initial conditions sensitivity. In clinical context, this implies that patients suffering from similar symptoms, but having different medical histories, may respond to treatment differently because their past trajectories have put their bodies on different pathological paths.

5 CONCLUSION: COMPLEXITY, EVIDENCE AND CHINESE MEDICINE

Recasting Chinese medicine in terms of complex systems dynamics allows us to scientifically articulate its foundations. When viewed this way the fundamental contrast between mainstream biomedicine and Chinese medicine becomes clear. Whilst it easy to rationally reconstruct biomedicine within a logico-causal model of explanation, Chinese medicine is thoroughly dynamic in its orientation. Recognising this fact explains what, from the point of view of orthodox medical practice, theory and research, appear to be idiosyncrasies of Chinese medicine: its eschewal of standardisation, its converse insistence on individualised treatments, its attention to minor details, its frequent changes in treatment strategies, its focus on diagnostic indicators and treatment practices foreign to Western medicine.

Insofar as they both view health and disease in terms of normal and abnormal rhythmicities, Chinese medicine shares a common basis with the dynamic disease paradigm emerging in Western biomedicine. Of course the two approaches are not identical. A chief difference, and a key strength, is Chinese medicine's 5,000 year tradition of experimenting with a range of techniques and substances not known in Western medicine. This tradition has resulted in a wealth of practical empirical knowledge that has been largely untapped by Western scientists. A marriage of these two traditions promises deep new insights into the workings of human bodies and their relationships with their environments. The dynamic disease approach brings to the relationship powerful techniques for quantifying subtle characteristics of complex data sets. These techniques will provide a methodological approach appropriate for the design of research for evaluating Chinese medicine compatible with its complex dynamic orientation.

The power of orthodox statistically-based methods derives from observing the relationships between a few variables for a large number of cases. Randomising subjects into treatment and control groups strives to render confounding variables irrelevant to the study. In effect, from a Chinese medical perspective 'confounding variables' are *dynamic factors* crucial to understanding the illness. Traditional research methods 'randomise away' an essential part of Chinese medical diagnosis and treatment. The power of complex dynamics methods derives from following individual cases in great detail over time. Several variables may be tracked at once, and more importantly, patterns in data are discerned over extended time periods. Neither method is superior to the other. In fact they are complementary. Orthodox medical methods are useful for generalising about multiple cases; complex dynamic methods are useful for discerning complex patterns in particular cases. Since it operates by concentrating on the dynamic details of individual cases, it is bound to be useful to adopt complex dynamic approaches in Chinese medicine research.

Biomedical research methods no doubt play an important role in understanding and evaluating Chinese medicine. This raises a basic question about the relationship between research and practice. The issue has been considered for orthodox medicine by advocates of 'Evidence-based Medicine' (EBM). From an EBM perspective, the proper relationship between research and practice is that clinical evidence (mainly derived from systematic reviews of randomised-controlled trials) ought to form the *basis* for clinical decision making. Rhetorically this position is unassailable (if perhaps vapid). No one would want to advocate *not* basing treatment on evidence [Rees, 2000]. Theoretically, as well, EBM appears straightforward: simply harness the result of decades of medical research, providing this information to doctors so that they can base their clinical decisions on tested treatments. One difficulty is the sheer volume of medical research literature. The main thrust of the EBM effort has been to summarise the research in systematic reviews making them easily accessible by publishing them in databases like that of the *Cochrane Collaboration*. Insofar as it has remained true to Archie Cochrane's vision, EBM has provided such information on efficacy through databases and evidence-based clinical guidelines. At its heart EBM is about disseminating research findings not about influencing research practice.

From a practical perspective, however, delivering information in a usable form is not the only issue. Despite its rhetorical and theoretical attractiveness, EBM has received extensive criticism. The main criticism concerns the evaluation and provision of evidence for use by practitioners. The traditional vehicle for this is systematic review of randomised-controlled trials (RCT's). In theory, a systematic review ought to provide an *objective* evaluation of all clinical evidence relevant to a particular illness that a doctor can access and apply in clinical practice. The idea is that a doctor could readily find information on an illness when a patient presents with it in clinic by accessing an electronic database. The database would contain a summary of the best evidence on treatments in the form of systematic review. There are several problems with this model:

1. *Systematic reviews are not the objective vehicle for evaluation that they may appear on the surface.* Although systematic, systematic reviews are not mechanical. Individual reviewers must make judgments on the quality of a given trial. In addition, any meta-analysis provided is also reliant on reviewer judgment [Bailar, 1995]. Various statistical techniques can provide a range of results depending on the statistical model chosen. Furthermore, since

many doctors may not be entirely *au fait* with meta-analytical techniques, results of meta-analysis must be *interpreted* by the reviewer. Clinicians rely on reviewer judgment for decisions about when a result is large, small, statistically or clinically significant.

2. *Overreliance on RCT's can yield a misleading picture of the evidence.* Although in principle the tenets of EBM do not exclude evidence obtained from other research designs, in practice reviews are often systematically, yet arbitrarily, restricted to RCT's. This is done mainly for expediency. A reviewer might reason like this: there is a vast literature concerning the question under review. Many of the studies are RCT's. RCT's provide "gold standard" evidence. Therefore a review restricted to RCT's will yield an accurate appraisal of the evidence. This reasoning is flawed. In many cases RCT's do not provide better evidence than other types of studies [Grossman and Mackenzie, 2005]. Particularly if the RCT's are of low quality and other high quality studies exist, restricting a review to RCT's can give a misleading result.

3. *Meta-analysis is an imperfect tool for amalgamating and assessing data.* Bailar [1995] points out several serious flaws common in meta-analyses of medical research. These flaws include understating bias and overstating statistical confidence, formulaic ("job shop") meta-analysis and lack of review of original articles (due to excessive work involved in producing such reviews). Notably, Bailar emphasises that meta-analysis cannot be routinised; in fact, human judgment is necessary to carry out accurate review of evidence.

4. *EBM principles may inappropriately influence research design and policy.* Archie Cochrane's concern was that often the latest evidence was not properly integrated into clinical practice. True to principles advocated by Cochrane, EBM is concerned with summarising, evaluating and disseminating empirical evidence. The aim is to improve clinical decision making. However as Grossman and Mackenzie [2005] argue, the EBM hierarchy often intrudes into research decision making. Researchers are often reluctant to propose and funding bodies are often reluctant to fund studies employing designs other than RCT's. Often a well-designed case-controlled or retrospective cohort analysis is superior to a poorly designed or underpowered RCT. Nevertheless RCT's are often favoured due to uncritical adherence to EBM principles.

5. *The relevance of epidemiological evidence for clinical practice is often limited.* RCT's are most suited to testing small delayed effects in pharmaceutical interventions. Larger scale immediate effects are easily observed, and surgery is difficult if not impossible to blind [Fairley, 2007]. Furthermore, there is a difference between statistical significance and clinical significance. A large study may reveal a small improvement on average in a treatment cohort; however it is not always apparent from average results whether the treatment is clinically effective. Translating average results to optimise treatment in individual cases is not a straightforward process [Schneider, 2005].

6. *There is actually little evidence that the patients of doctors adopting an "evidence-based" approach fare better than patients of doctors who do not.* This is the reflexive critique of EBM. From an EBM perspective, clinical practice is to be improved by incorporating the best evidence. But this claim itself is accepted by EBM without evidence [Dearlove, *et al.*, 1995]. In fact there is only limited evidence from empirical studies of the effectiveness of adopting an evidence-based approach [Fairley, 2007].

7. *Most GP's do not have sufficient time to perform the research necessary to integrate latest research findings into practice.* "In primary care, GPs have less than five hours a week for reading, educational courses, and teaching" [Jacobson, *et al.*, 1997]. Couple this with the fact that busy GP's may easily see over 240 patients per week and a little arithmetic reveals the problem: seeking the latest evidence on each condition seen in a week would require researching as many as 48 conditions per hour (assuming each case is unique) which is hardly feasible. (This workload pattern is itself a consequence of Western conceptions of standardised diagnosis and treatment.) Searching a database for evidence to apply in each individual case is not a practical model for clinical practice.

8. *Whilst EBM emphasises putting more science into medical practice it overlooks the arts of diagnosis and treatment.* No one would argue against basing medical treatment on evidence. However, like any complex human endeavour medical practice requires individuals to make judgments. In particular, diagnosis requires the assimilation and interpretation of complex and subtle information. It is acknowledged that diagnosis requires complex judgment; medical training takes this into account. However, EBM may give the impression that treatment can be routinised. Many treatments, e.g. surgery, require special skills. It is difficult to gather more than crude overall success rates and like evidence for such skill-based treatments, and it is impossible to routinise such procedures. There is art and technique to medicine, and these cannot be *derived* from evidence [Goodman, 2003; Grahame-Smith, 1995].

9. *There is a risk that EBM will erode confidence in practitioner judgment, replacing it with the meta-clinical decisions made by managers, government agencies and insurance company actuaries.* The intention of EBM is merely to make evidence available to practitioners in order to improve clinical decision making [Sackett, *et al.*, 1996; Rosenberg and Donald, 1995b]. However, to those not familiar with the complexities of clinical practice, some sound clinical decisions may appear arbitrary or contrary to best evidence. Even evidence-informed practice requires human judgment [Rosenberg and Donald, 1995b]. However, evidence-based principles could be co-opted by regulators in order to override clinical judgment [Grahame-Smith, 1995].

These criticisms do not undermine evidence-based medicine. Rather, the debate they have engendered has brought into relief a fundamental duplicity at the heart of evidence-based medicine. There have always been two aspects of EBM: one programmatic the other rhetorical. As a program EBM has made epidemiological evidence, collected in such databases as that of the Cochrane Collaboration, more easily accessible to medical practitioners. As it stands, what use is actually made of such information is entirely a matter of practitioner judgment. However, the rhetorical aspect is revealed even in the name 'evidence-*based* medicine'. The name seems to imply more than simply making evidence available. If medical practice was truly to be *based* on evidence, clinical decisions ought to be somehow derivable from the information in the systematic reviews. Even its leading advocates do not argue such a strong line. Guidelines for EBM stress the important role of practitioner judgment [Sackett, *et al.*, 1996; Rosenberg and Donald, 1995b).

The programmatic aspect of EBM seems beyond reproach. To our knowledge no one has argued that the Cochrane Collaboration is a bad idea. Rather, the debate has focussed on three issues: 1. the reliability of epidemiological evidence, 2. the limits of meta-analytical information and 3. how such evidence should be incorporated into practice. In fact the first two issues bear directly on the third. If epidemiological methods, including meta-analysis, are limited and unreliable, we must use care, and good judgment, when applying their results as evidence in practice. Thus practice cannot be evidence-*based* in a strong sense. It will be a matter of human judgment, rather than logical derivation, how evidence is incorporated into clinical decision making. In this sense the debate has been cautionary: it is a debate about the extent to which evidence-*based* rhetoric ought to be part of the program of evidence-based practice.

The critique of EBM bears directly on the issue of what role ought scientific evidence play in the future development of *Chinese* medical theory, practice and research. We have argued that conventional medical research methods fail to do justice to the complex dynamic commitments of Chinese medicine. Reliance purely on evidence gained from conventional methods would lead to a diminished Chinese medicine. All of the issues that have been raised in the context of orthodox medicine of course apply to Chinese medicine. From the perspective of this analysis the most important issue raised thus far concerns the role of practitioner judgment and the relationship between art and science in medical practice.

We have shown how Chinese medicine's dynamic approach yields a particular reliance on practitioner judgment. A sketch of a rationale for how such judgment can be made sensible was provided in terms of embodied knowledge and situated cognition. Although scientific evidence plays some role in Chinese medical practice, much knowledge that underwrites diagnosis and treatment is embodied in individual practitioners. This knowledge is transmitted through apprenticeship and experience rather than through studying textbooks and systematic reviews. The reasoning patterns in Chinese medicine incorporate its dynamic orientation. A key aspect of situated cognition is that it incorporates temporal factors into the reasoning process. Thus situated cognition is a better model for reasoning in Chinese medicine than is formal logic, which is atemporal in character. Although some themes implicit to situated cognition and embodied knowledge have been discussed (by e.g. [Farquhar, 1995; Kuriyama, 1995]) there is no sustained account of Chinese medicine explicitly from a situated cognition or embodied knowledge perspective published in the literature. Despite this, such a project would make a significant contribution to the understanding, legitimation and transmission of Chinese medicine. In particular, it would shed light on the rationality of practitioner judgment and make explicit the art of Chinese medical practice.

We should note that the issues of balancing art and science, as well as those of the legitimacy of practitioner judgment arise in Western biomedicine as well. And in fact many, perhaps a majority of, clinical decisions in medicine are not based primarily on scientific evidence from clinical trials [Morreim, 2003]. The main reason for this is that specific clinical trial results are difficult to generalise to real-life situations with diversity and co-morbidity playing confounding roles [Clark-Grill, 2007]. Thus Western biomedicine relies on practitioner judgment as well. And of course, this does not undermine the legitimacy of the practice. In fact one might argue just the opposite. (Would anyone seriously hold the view that medical care would be improved by replacing GP's with computerised expert systems and data bases?) However a careful articulation of the role of practitioner judgment, in terms of situated cognition and embodied knowledge would not only aid in *legitimising* the role of the individual practitioner in biomedicine, but could also contribute to the *improvement* of practice as well.

On this score the major difference between the Western and Chinese medical traditions is the rhetorical tradition of legitimating medicine by tying it to science; dating at least back to the Nineteenth Century in the West, this tradition has only recently crept into Traditional Chinese Medicine.[24] This rhetoric is reflected in an article appearing in a recent cognitive science book where Thagard and Zhu [2003] associate Chinese medicine with "homeopathic magic" and "pre-scientific and pseudoscientific thinking" whilst Western medicine is associated with "modern science" and "evidence-based medicine". In 1998 the *Journal of the American Medical Association* published an issue featuring several clinical trials of complementary and alternative medicine (CAM). The editorial prefacing this research stated, "There cannot be two kinds of medicine–conventional and alternative. There is only medicine that has been adequately tested and medicine that has not, medicine that works and medicine that may or may not work" [Fontanarosa and Lundberg, 1998, p. 841]. The editors of the *New England Journal of Medicine* express the same sentiment: "There is no alternative medicine. There is only scientifically proven, evidence-based medicine supported by solid data or unproven medicine, for which evidence is lacking" [Angell and Kassirer, 1998, p. 1618].

This rhetoric emerged with the scientific revolution over five hundred years ago,

[24]'Traditional Chinese Medicine' is an explicitly scientised version of Chinese medicine revived by Chairman Mao and developed in the mid 20^{th} Century after Chinese medicine was outlawed in China at turn of the century in favour of scientifically-based Western medicine (see [Scheid, 2002]).

and it has been deployed as legitimation discourse following a strategy which begins by drawing a sharp distinction between *superstition* and *truth*. The details vary depending upon context, but various dichotomies are established, for example between, magic and science, supernatural and natural, religion and science, pseudoscience and science, nonsense and sense. In recent medical discourse the contrast is drawn between "evidence-based" and "eminence-based" medicine highlighting the ideological nature of the issue. The examples quoted above are deficient in slightly different ways. On the one hand, the cognitive science chapter at least acknowledges that the two paradigms cannot be compared holistically head to head, whilst the medical journals, although they do not explicitly state the position, give the impression that the canon of Western medicine is monolithic and is by in large scientifically proven. (In fact from the perspective of clinical trials there is no "medicine", in general, only "medical treatments", in particular). On the other hand, at least the medical journals' contrast of "proven" with "unproven" medicine reflects some progress over Thagard and Zhu's contrast of "modern science" with "homeopathic magic". Lacking empirical evidence for a belief implies neither that it is false nor that it is based on superstition; it simply indicates ignorance. In any case, talk couched in such dichotomies is simplistic and unhelpful for furthering the cause of understanding health and medicine.

We would prefer to base our understanding on the premise that in all medical traditions practitioner judgment plays a crucial role; such judgment is informed by evidence provided by tradition, training, theory, experience and scientific research. It is only by providing an account whereby these influences are integrated that we will gain understanding of the complex process of the rationality of medical decision making and an articulation of the intelligence of medical practice. Such an approach to understanding medicine, including such unorthodox alternatives as Chinese medicine, would transcend such legitimation strategies providing a deep understanding of the intricacies of medical practice.

For Chinese medicine this understanding would have complex dynamics at its heart, acknowledging not only its conceptual role in diagnosis and treatment but also the interactive character of patient-environment-practitioner relationships. Such a program could be *broadly* scientific, moving beyond clinical trial methodology, including randomisation, control, placebo, blinding, probability and statistics, that is taught as though it comprises THE scientific method in medical and health sciences faculties. The research would acknowledge the legitimacy of Chinese medicine as a paradigm, and it would be premised on achieving outcome parameters as defined dynamically by Chinese medicine. By adopting an approach based on complex systems dynamics, a new research program will emerge, shedding light not only on Chinese medicine, but on health and disease in general.

BIBLIOGRAPHY

[Allchin, 1996] D. Allchin. Points East and West: Acupuncture and Comparative Philosophy of Science. Proceedings, PSA96. *Philosophy of Science*, 63, S107-S115, 1996.

[Ames and Hall, 2003] R. T. Ames and D. L. Hall. *Dao De Jing: A Philosophical Translation.* Ballantine Books, New York, 2003.

[Anderson and May, 1991] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control.* Oxford University Press, Oxford, 1991.

[Angell and Kassirer, 1998] M. Angell and J. P. Kassirer. Alternative medicine –The risks of untested and unregulated remedies. *The New England Journal of Medicine*, 339, 839–841, 1998.

[Bailar III, 1995] J. C. Bailar III. The practice of meta-analysis. *Journal of Clinical Epidemiology*, 48(1), 149-157, 1995.

[Bailar III, 1997] J. C. Bailar III. The promise and problems of meta-analysis. *New England Journal of Medicine*, 337, 559–61, 1997.

[Birch, 1997a] S. Birch. Testing the claims of traditionally based acupuncture. *Complementary Therapies in Medicine*, 5, 147-151, 1997.

[Birch, 1997b] S. Birch. Issues to consider in determining an adequate treatment in a clinical trial of acupuncture. *Complementary Therapies in Medicine*, 5, 8-I2, 1997

[Birch *et al.*, 1999] S. Birch, K. Sherman, and Y. Zhong. Acupuncture and low-back pain: Traditional Chinese medical acupuncture differential diagnoses and treatments for chronic lumbar pain. *Alternative and Complementary Medicine*, 5, 415-425, 1999.

[Bray, 1995] F. Bray. A Deathly Disorder: Understanding Woman's Health in Late Imperial China. In D. Bates (ed.), *Knowledge and the Scholarly Medical Traditions.* Cambridge University Press, Cambridge, 1995.

[Cardini *et al.*, 1999] F. Cardini, C. Wade, A. L. Regalia, S. Gui, W. Li, R. Raschetti, and F. Kronenberg. Clinical research in traditional medicine: Priorities and methods. *Complementary Therapies in Medicine*, 14(4), 282-287, 2006.

[Cassell, 1998] J. Cassell. *The Woman in the Surgeon's Body.* Harvard University Press, Cambridge MA, 1998.

[Chandrasekhar, 1961] S. Chandrasekhar. *Hydrodynamic and Hydromagnetic Stability.* Clarendon Press, Oxford, 1961.

[Charlton, 1996] B. G. Charlton. The scope and nature of epidemiology. *Journal of Clinical Epidemiology*, 49, 623– 6, 1996.

[Clark-Grill, 2007] M. Clark-Grill. Questionable Gate-keeping: Scientific Evidence for Complementary and Alternative Medicines (CAM): Response to Malcolm Parker. *Journal of Bioethical Inquiry*, 4, 21-28, 2007.

[Coetaux *et al.*, 2006] R. R. Coetaux, W. Chen, C. E. Lindemuth, Y. Tan, and A. C. Reilly. Variability in the diagnosis and point selection for persons with frequent headache by traditional Chinese medicine acupuncturists. *Alternative and Complementary Medicine*, 12, 863-872, 2006.

[Couto, 1998] J. S. Couto. Evidence-based medicine: a Kuhnian perspective of a transvestite non-theory. *Journal of Evaluation in Clinical Practice*, 4(4), 267-275, 1998.

[Davidoff *et al.*, 1995] F. Davidoff, B. Haynes, D. Sackett, and R. Smith. Evidence based medicine (Editorial). *British Medical Journal*, 310, 1085-1086, 1995.

[Dearlove *et al.*, 1995] O. Dearlove, A. Sharples, K. O'Brien, and C. Dunkley. Many questions cannot be answered by evidence based medicine. *British Medical Journal*, 311, 257-258, 1995.

[Dyke, 1988] C. Dyke. *The evolutionary dynamics of complex systems: a study in biosocial complexity.* Oxford University Press, New York, 1988.

[Ellingson, 2006] L. L. Ellingson. Embodied Knowledge: Writing Researchers' Bodies into Qualitative Heath Research. *Qualitative Health Research*, 16,298-310, 2006.

[EBMWG, 1992] Evidence Based Medicine Working Group. Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA*, 268, 2420-2425, 1992.

[Fairley, 2007a] J. W. Fairley. Evidence based medicine - historical perspective and critique, 2007. `http://www.entkent.com/ebm-historical.html` (accessed 7 April 2009).

[Fairley, 2007b] J. W. Fairley. Philosophical, scientific and statistical background to evidence based medicine, 2007. `http://www.entkent.com/ebm-philosbkd.html` (accessed 8 April 2009).

[Farquhar, 1995] J. Farquhar. Re-writing Traditional Medicine in Post-Maoist China. In D. Bates (ed.), *Knowledge and the Scholarly Medical Traditions.* University Press, Cambridge, 1995.

[Featherstone, 2004] J. D. B. Featherstone. The Continuum of Dental Caries — Evidence for a Dynamic Disease Process. *Journal of Dental Research*, 83(Special Issue C), C39-C42, 2004.

[Fontanarosa and Lundberg, 1998] P. B. Fontanarosa, G. D. Lundberg. Alternative medicine meets science. *Journal of the American Medical Association*, 280, 1618–1619, 1998.

[Gao *et al.*, 2008] Y. Gao, W.E. Herfel, and D. Rodrigues. Toward a New Research Paradigm: Traditional Chinese Medicine from a Nonlinear Dynamical Perspective. *Journal of Medicine & Philosophy* 29(April), 70-71, 2008.

[Garfinkel, 1983] A. Garfinkel. A mathematics for physiology. *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology*, 245, 455-466, 1983.

[Garfinkel, 1987] A. Garfinkel. The slime mold Dictyostelium as a model of self-organization in social systems. In F. E. Yates (ed.), *Self Organizing Systems: The Emergence of Order*. Plenum, New York, 181-212, 1987.

[Gertler, 2003] M. S. Gertler. Tacit Knowledge and the Economic Geography of Context, or the Undefinable Tacitness of Being (there). *Journal of Economic Geography*, 3, 75-99, 2003.

[Glass and Mackey, 1988] L. Glass and M. Mackey. *From Clocks to Chaos.* Princeton University Press, Princeton, 1988.

[Goldberger, 1997] A. L. Goldberger. Fractal Variability versus Pathologic Periodicity: Complexity Loss and Stereotypy in Disease. *Perspectives in Biology and Medicine*, 40(4), 543, 1997.

[Goodman, 2002] K. W. Goodman. *Ethics and Evidence-Based Medicine: Fallibility and Responsibility in Clinical Science*. Cambridge, University Press, Cambridge, 2002.

[Goodman, 2002] N. W. Goodman. Evidence-Based Medicine Needs Proper Critical Review. *Anesthesiology and Analgesia*, 95, 1817, 2002.

[Goodman, 2003] N. W. Goodman. Review of Goodman 2002. *Journal of the Royal Society of Medicine*, 96, 251, 2002.

[Gordon, 1988] D. R. Gordon. Clinical Science and Clinical Expertise: Changing Boundaries between Art and Science in Medicine. In M. Lock, D. R. Gordon (eds.), *Biomedicine Examined*. Kluwer, Dordrecht, 1988.

[Graham, 1989] A. C. Graham. *Disputers of the Tao: Philosophical Argument in Ancient China*. Open Court, La Salle, Ill., 1989.

[Grahame-Smith, 1995] D. Grahame-Smith. Evidence based medicine: Socratic dissent. *British Medical Journal*, 310(6987), 1126-1127, 1995.

[Greenhalgh, 1999] T. Greenhalgh. Narrative based medicine in an evidence based world. *British Medical Journal*, 318, 323-325, 1999.

[Grossman and Mackenzie, 2005] J. Grossman and F. J. Mackenzie. The Randomized Controlled Trial: gold standard, or merely standard? *Perspectives in Biology and Medicine*, 48(4), 516–34, 2005.

[Grube, 1980] G. M. A. Grube. *Plato's Thought*. Hackett, Indianapolis, 1980.

[Hammerschlag, 1998] R. Hammerschlag. Methodological and ethical issues in clinical trials of acupuncture. *Journal of Alternative and Complementary Medicine*, 4(2), 159–171, 1998.

[Heath, 2000] R. A. Heath. *Nonlinear Dynamics: Techniques and Applications in Psychology.* Erlbaum, Mahwah NJ, 2000.

[Herfel *et al.*, 2007] W. E. Herfel, D. Rodrigues, and Y.Gao. Chinese Medicine and the Dynamic Conceptions of Health and Disease. *Journal of Chinese Philosophy*, 34(s1), 57-79, 2007.

[Hochsmann and Yang, 2007] H. Hochsmann, G. Yang. *Zhuangzi*, Pearson/Longman, New York, 2007.

[Hogeboom *et al.*, 2001] C. J. Hogeboom, K. J. Sherman, and D. C. Cherkin. Variation in diagnosis and treatment of chronic low back pain by traditional Chinese medicine acupuncturists. *Complementary Therapies in Medicine*, 9, 154-166, 2001.

[Howells, 1996] J. Howells. Tacit Knowledge, Innovation and Technology Transfer. *Technology Analysis & Strategic* Managemen, 8(2), 91-106, 1996.

[Hooker *et al.*, 1992] C. A. Hooker, H. B. Penfold, and R. J. Evans. Towards a Theory of Cognition Under a New Control Paradigm. *Topoi*, 11, 71-88, 1992.

[Hsu, 1999] E. Hsu. *TheTransmission of Chinese medicine.* Cambridge University Press, New York, 1999.

[Hunter, 1991] K. M. Hunter. *Doctor's Stories.* Princeton University Press, Princeton, NY, 1991.

[Hutchins, 1995] E. Hutchins. *Cognition in the Wild.* MIT Press, Cambridge MA, 1995.

[I Ching, 1968] *The I Ching, or, Book of Changes: The Richard Wilhelm Translation*. Tr. C. F. Baynes. 1968. London, Routledge & Kegan Paul, 1968.

[Issacs and Fitzgerald, 1999] D. Issacs and D. Fitzgerald. Seven alternatives to evidence based medicine. *British Medical Journal*, 319, 1618, 1999.

[Jacobson *et al.*, 1997] L. D. Jacobson, A. G. K. Edwards, S. K. Granier, and C. C. Butler. Evidence-based medicine and general practice. *British Journal of General Practice*, 47**,** 449-452, 1997.

[Kalauokalani *et al.*, 2001] D. Kalauokalani, K. J. Sherman, and D. C. Cherkin. Acupuncture for chronic low back pain: Diagnosis and treatment patterns among acupuncturists evaluating the same patient. *Southern Medical Journal*, 94, 486-492, 2001.

[Kaplan and Glass, 1995] D. Kaplan and L. Glass. *Understanding Nonlinear Dynamics.* Springer-Verlag, New York, 1995.

[Kuriyama, 1995] S. Kuriyama. Visual Knowledge in Classical Chinese Medicine. In D. Bates (ed.), *Knowledge and the Scholarly Medical Traditions*, Cambridge University Press, Cambridge, 1995.

[Kuriyama, 1999] S. Kuriyama. *The expressiveness of the body and the divergence of Greek and Chinese medicine*. Zone Books, New York, 1999.

[Lao Tzu, 1996] Lao Tzu. *The Tao Te Ching*. M. Kwok, M. Palmer, J. Ramsay. Tr. Element, Rockport, MA, 1996.

[Lu, 1978] H. C. Lu. Trans. A *Complete Translation of the Yellow Emperor's Classic of Internal Medicine and the Difficult Classic*. The Academy of Oriental Heritage, Vancouver, 1978.

[Lippman, 1999] A. Lippman. Embodied Knowledge and Making Sense of Prenatal Diagnosis. *Journal of Genetic Counseling*, 8(5), 255-74, 1999.

[Ma, 1986] S. Ma. *Huang Di Nei Jing Suwen Zhu Zheng Fa Wei.* Academic Publishing House, Beijing, 1986.

[Ma, 1998] S. Ma. *Ling Shu*. Science and Technology Publishing House, Beijing, 1998.

[MacKenzie and Spinardi, 1995] D. MacKenzie and G. Spinardi. Tacit Knowledge, Weapons Design, and the Uninvention of Nuclear Weapons. *The American Journal of Sociology*, 101(1), 44-99, 1995.

[Major, 1984] J. S. Major. The Five Phases, Magic Squares and Semantic Cosmography. In H. Rosemont Jr. (ed.), *Explorations in Early Chinese Cosmology.* American Academy of Religion Studies, L(2), Scholars Press, Chico, 133-166, 1984.

[Major, 1991] J. S. Major. Substance, Process, Phase: *Wuxing* in the Huainanzi. In H. Rosemont Jr. (ed.), *Chinese Texts and Philosophical Contexts: Essays dedicated to Angus Graham.* Open Court, LaSalle Ill, 67-78, 1991.

[Mariotto, 2000] A. Mariotto. Propaganda based medicine is an alternative . . . *British Medical Journal*, 321, 239, 2000.

[Morreim, 2003] E. H. Morreim. A Dose of Our Own Medicine: Alternative Medicine, Conventional Medicine, and the Standards of Science. *Journal of Law, Medicine & Ethics*, 31, 222-235, 2003.

[NCMCCG, 1997] Nanjing Chinese Medicine College Classics Group (eds.). *Huangdi Neijing Suwen Yi Shi*. (*The Yellow Emperor's Inner Canon Plain Questions: Modern Chinese Translation and Interpretation.*) Shanghai Science and Technology Publishing House, Shanghai, 1997.

[Nelson and Winter, 1982] R, R. Nelson and S. G. Winter. *An Evolutional Theory of Economic Change.* Harvard University Press, Cambridge MA, 1982.

[Nooteboom, 1997] B. Nooteboom. Path Dependence of Knowledge: Implications for the Theory of the Firm. In L. Magnusson, J. Ottosson (eds.), *Evolutionary Economics and Path Dependence*. Edward Elgar, Cheltenham, 1997.

[Parker, 2005] M. Parker. False dichotomies: EBM, clinical freedom, and the art of Medicine. *Medical Humanities*, 31, 23-30, 2005.

[Parker, 2007] M. Parker. Two into one won't go: Conceptual, clinical, ethical and legal impedimenta to the convergence of CAM and orthodox medicine. *Journal of Bioethical Inquiry*, 4(1), 7-19, 2007.

[Plato, 1961] Plato. *The Collected Dialogues*. E. Hamilton, H. Cairns (eds.). Princeton University Press, Princeton, 1961.

[Polanyi, 1966] M. Polanyi. *The Tacit Dimension*. Rinehart, New York, 1966.

[Pool, 1989] R. Pool. Is it Healthy to be Chaotic? *Science*, 243, 604-7, 1989.

[Porkert, 1974] M. Porkert. *The theoretical foundations of Chinese medicine*. MIT Press, Cambridge MA, 1974.

[Rapp, 1986] P. E. Rapp. Oscillations and Chaos in Cellular Metabolism and Physiological Systems. In A. V. Holden (ed.), *Chaos*. Princeton University Press, Princeton, 179-208, 1986.

[Rees, 2000] J. Rees. Evidence-based medicine: the epistemiology that isn't. *Journal of the American Academy of Dermatolology*, 43, 727–9, 2000.

[Reimann, 1963] H. A. Reimann. *Periodic Diseases.* F.A. Davis, Philadelphia, 1963.

[Rodrigues *et al.*, 2008] D. Rodrigues, Y. Gao, and W. E. Herfel. Analysing a Chinese Medical Case Study in terms of Self-Organisation. *Journal of Medicine & Philosophy* 29(March), 56-7, 2008.

[Rosenberg and Donald, 1995a] W. Rosenberg and A. Donald. Evidence based medicine: an approach to clinical problem-solving. *British Medical Journal*, 310, 1122-1126, 1995.

[Rosenberg and Donald, 1995b] W. Rosenberg and A. Donald. Authors' Reply. *British Medical Journal*, 311, 259, 1995.

[Sackett *et al.*, 1996] D. L. Sackett, W. M. Rosenberg, J. A. M. Gray, R. B. Haynes, and W. S. Richardson. Evidence based medicine: what it is and what it isn't. *British Medical Journal*, 312, 71-72, 1996.

[Saviotti, 1998] P. P. Saviotti. On the Dynamics of Appropriability, of Tacit and of Codified Knowledge. *Research Policy*, 26, 843-56, 1998.

[Scheid, 1993] V. Scheid. Orientalism revisited: reflections on scholarship, research & professionalization. *The European Journal of Oriental Medicine*, 1(2), 22-33, 1993.

[Scheid, 2002] V. Scheid. *Chinese medicine in contemporary China: plurality and synthesis.* Duke University Press, Durham, 2002.

[Scheid, 2007] V. Scheid. Traditional Chinese medicine — What are we investigating? The case of menopause. *Complementary Therapies in Medicine*, 15(1), 54-68, 2007.

[Schneider and Lane, 2005] S. H. Schneider and J. Lane. *The Patient from Hell: How I worked with my doctors to get the best of modern medicine and how you can too.* Da Capo, Cambridge, 2005.

[Sherman *et al.*, 2001] K. J. Sherman, D. C. Cherkin, and C. J. Hogeboom. The diagnosis and treatment of patients with chronic low-back pain by traditional Chinese medical acupuncturists. *Alternative and Complementary Medicine*, 7**,** 641-650, 2001.

[Sivin, 1987] N. Sivin. *Traditional Medicine in Contemporary China.* University of Michigan Press, Ann Arbor, 1987.

[Slavin, 1995] R. E. Slavin. Best evidence synthesis: an intelligent alternative to meta-analysis. *Journal of Clinical Epidemiology*, 48(1), 9-18, 1995.

[Song *et al.*, 2009] F. Song, Y. K. Loke, T. Walsh, A.-M. Glenny, A. J. Eastwood, and D. G. Altmann. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *British Medical Journal*, 338, 1147, 2009.

[Thagard, 1999] P. Thagard. *How Scientists Explain Disease.* Princeton University Press, Princeton, 1999.

[Thagard and Zhu, 2003] P. Thagard and J. Zhu. Acupuncture, Incommensurability, and Conceptual Change. In G. M. Sinatra, P. R. Pintrich (eds.), *Intentional Conceptual Change.* Erlbaum, Mahwah NJ, 2003.

[Tortora and Grabowski, 1996] G. J. Tortora and S. R. Grabowski. *Principles of Anatomy and Physiology.* 8^{th} edn. HarperCollins, New York, 1996.

[Valarde, 1980] M. Valarde, C. Normande. Convection. *Scientific American.* 243(1), 78-93, 1980.

[Wagner *et al.*, 1996] C. Wagner, D. B. Nafz, and P. B. Persson. Chaos in blood pressure control. *Cardiovascular Research*, 31, 380-387, 1996.

[Wang, 1998] R. Wang. *Zhuangzi with modern translation and notes (in Chinese).* Qi-lu Books, People's Republic of China, 1998.

[Wang, 1999] R. Wang. *Zhuangzi (inEnglish)*, Hunan People's Publishing House, People's Republic of China , 1999.

[Wilde, 2003] M. H. Wilde. Embodied Knowledge in Chronic Illness and Injury. *Nursing Inquiry*, 10(3), 170-76, 2003.

[Winfree, 1987a] A. Winfree. *The Timing of Biological Clocks.* Scientific American Library. Freeman, New York, 1987.

[Winfree, 1987b] A. Winfree. *When Time Breaks Down.* Princeton, University Press, Princeton, 1987.

[Zhang *et al.*, 2004] G. G. Zhang, W. L. Lee, L. Lao, B. Bausell, B. Berman, and B. Handwerger. The variability of TCM pattern diagnosis and herbal prescription on rheumatoid arthritis patients. *Alternative Therapies in Health and Medicine*, 10, 58-63, 2004.

[Zhang *et al.*, 2005] G. G. Zhang, W. Lee, B. Bausell, L. Lao, B. Handwerger, and B Berman. Variability in the traditional Chinese medicine (TCM) diagnoses and herbal prescriptions provided by three TCM practitioners for 40 patients with rheumatoid arthritis. *Alternative Therapies in Health and Medicine*, 11, 415-421, 2005.

[Zhang, 2006] X. Zhang. *Returning to Chinese Medicine: Rediscovering the fundamental theory of Chinese medicine*. Qingdao Publishing House, Qingdao, 2006.

Part X

Military Science

MILITARY APPLICATIONS OF COMPLEX SYSTEMS

Alex J. Ryan

1 INTRODUCTION

Military theory is replete with assertions that contemporary conflict is becoming more complex. In *Network Centric Warfare*, one of the most highly cited books on military theory of the last two decades, Alberts *et al.* claim that "it is clear that our missions have become far more complex and our challenges and adversaries less predictable" [Alberts *et al.*, 1999, p. 60]. The National Security Strategy of the United States of America adds that the new adversaries of the West "make today's security environment more complex and dangerous"[Bush, 2002, p. 13]. Undoubtedly, globalisation and the spread of information and communications technology have led to greater interconnectivity and increased interdependence over much shorter time scales between previously weakly coupled events and actors. Modern forces are larger, more specialised, more networked, and decisions are more distributed compared to conventional militaries of the past. Further, the competitive dynamic that pervades both war and preparations against the threat of war generates continual pressure for military complexification, to the extent that complexity yields new potential for military advantage.

Yet this neither implies that complexity is a new phenomenon in war, nor that future missions will necessarily be more complex. Thucydides' [1996] detailed history of the Peloponnesian War between 431–411BC reminds us that military operations have always been a dynamic and violent nexus of political, diplomatic, social, cultural, moral, ideological, religious, economic, scientific, technological, informational, geographical and environmental interactions that have decisively shaped world history. Military operations are a perpetual source of novelty and complexity that challenge our ability to comprehend, explain and influence them, and carry the most severe consequences for those who fail to appreciate their complexity. Hence, it is hardly surprising that militaries are eager consumers — and indeed often producers — of new ways of coping with and even exploiting complexity. This chapter will show how an appreciation of complex systems is reflected in military theory and practice. The military applications of science are reviewed, with a particular focus on the systems sciences. Within this context, the contributions of complex systems to military problems is assessed. Perhaps more significantly, this chapter critically examines military theory from the perspective

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

of complex systems, in order to identify further applications with the potential to significantly transform military systems and operations.

2 SCIENCE IN WARFARE

Clausewitz's unfinished magnum opus *On War*, first published posthumously in 1832, argued that war is neither an art nor a science, but rather is a part of man's social existence — specifically the part involving the resolution of political ends via the means of organised violence [Clausewitz, 1984, p. 149]. Clausewitz emphasised this point to warn of the dangers of "a mass of incorrect analogies" between war and the arts and sciences, and also as an antidote to those theorists of Napoleonic warfare who wanted to make military strategy more scientific by supplanting creative thought and critical judgement in strategy with prescriptive geometrical forms and patterns. Both points are legitimate concerns, but they by no means preclude the use of science *within* the human activity of war. Przemieniecki [1994, p. 2] traces the recorded use of scientific methods to enhance the effectiveness of military operations to Archimedes' war machines, which helped to break the Roman naval siege of Syracuse in about 214 B.C. The desire to achieve an edge in military materiel has long been a driver for investment in new technology, and to the extent that scientific understanding underpins technological innovation, the evolution of military equipment from catapults to the atom bomb documents a tangible record of the application of science to war. However, our interests here will focus on how science has influenced the way that warfare is organised and understood, rather than the equipment with which war is waged.[1] Thus, the research and development that French scientists (including Fourier) conducted in support of Napoleon's Egyptian expedition of 1798 is a more appropriate precedent for this use of science in war.

During World War I, two individual contributions to a scientific understanding of war stand out. In 1916, Lanchester developed a simple pair of coupled differential equations as a model of air combat. The Lanchester equations later became the fundamental equations upon which most modern theories of combat attrition are based [Przemieniecki, 1994; Ilachinski, 2000]. The original form of the Lanchester equations for directed fire warfare [Lanchester, 1916] is:

$$\frac{dB}{dt} = -\alpha_R R(t) \tag{1}$$

$$\frac{dR}{dt} = -\alpha_B B(t), \tag{2}$$

where t denotes time, B, R are the number of force units in the Blue and Red sides, their derivatives represent the attrition rate, and α_B, α_R are the effective firing rate of a single force unit for the Blue and Red forces respectively. The

[1] For a discussion of the co-constitutive nature of science and technology as it applies to warfare, see [Bousquet, 2009]

Lanchester equation has a closed form solution known as the Lanchester square law. Intuitively, the Lanchester square law shows that increases in the initial size of the Blue force $B(0)$ produces quadratic improvements in battlefield superiority compared to improvements in the effective firing rate α_B. Lanchester's square law provides theoretical support for concentrating effort at a focal point. This tactic, first used to great effect by Epaminondas against the Spartans in 371 B.C.E, was later called *schwerpunkt* by the Germans [Gabriel, 2001, p. 91], and was described by Jomini [1992, p. 70] as one of the four maxims of the fundamental principle of war, "To maneuver to engage fractions of the hostile army with the bulk of one's forces".

World War I also motivated Richardson, an English physicist who served in the Friends' Ambulance Unit, to amass the largest data set of casualties from human quarrels dating back to 1820. Richardson deliberately did not differentiate between wars, riots and murders, arguing that they shared a common psychological basis, and in each case the casualties were the result of deliberate human action. Instead, he classified quarrels by magnitude of the quarrel dead in bins of unit range of the base 10 logarithm. The two world wars (Richardson maintained his data set following World War II) were the only conflicts with a magnitude of 7 ± 0.5, while at the other end of the scale an estimated six million murders of magnitude less than 0.5 (3 or fewer deaths) occurred over the 126 years between 1820 and 1945. By graphing the world total for the number of quarrels over eight orders of magnitude of quarrel dead on a log-log plot, Richardson raised the possibility that these apparently disparate kinds of conflicts could be described by a single distribution, and showed that the region of the war data (magnitude greater than 2.5) was described by a power law $P(x) \propto x^{-\alpha}$, where α was estimated to be 1.50 [Richardson, 1948, p. 540-541].

In addition to demonstrating the scale invariance of conflict casualties, Richardson showed that both the start and end dates of war were consistent with a Poisson distribution, which suggests that the outbreak of war and peace is a random, memoryless process comprised of independent events [Richardson, 1960, p. 128]. He produced an important model of arms race dynamics and detailed statistical analyses of the causes of war, which considered political, economic, religious, linguistic and geographical factors in conflict. Interestingly, in trying to assess the correlation between the length of a shared boundary between two countries and the likelihood of war, Richardson noted the dependence of the length of the boundary on the scale of the unit of measurement, and the paper reporting this result was one of the inspirations for Mandelbrot's theory of fractals [Hayes, 2002, p. 13].

2.1 Operations Research

The above examples demonstrate isolated applications of science to understanding warfare. It was during World War II that a more systematic approach developed. In the Bawdsey Research Station in the UK, the problem of coordinating the operational employment of the new radar equipment as an integrated system was addressed by a team of "operational researchers". The success of this approach during the Battle of Britain in 1940 contributed to the use of teams of civilian operational analysts in the RAF, Navy and Army headquarters as well as at the highest levels of government planning [Matthews, 2004, p. 228]. Notable applications during the war included optimising depth charge settings, convoy sizes and servicing schedules for long range aircraft in the Battle of the Atlantic, and targeting enemy logistics and analysing fighter losses in Operation Overlord. Radar also stimulated the growth of Operations Research (OR) in the U.S. Army Air Forces.

Following the war, OR was institutionalised in both countries. The Operational Research Society in Britain, formed in 1948, and the Operations Research Society of America, formed in 1952, created interdisciplinary communities of practice and journals for OR. This helped to migrate OR into industrial applications, although OR continued to play an important role in promoting the military application of scientific methods. The broader scope of OR was encompassed by Morse and Kimball's [Morse and Kimball, 1951] definition of OR as "a scientific method of providing executive departments with a quantitative basis for decisions regarding the operations under their control".

2.2 Systems Analysis

A number of other interdisciplinary fields were forming in the wake of World War II. In the newly established RAND (Research ANd Development) Corporation, the mathematician Paxson (whose work has not been declassified) was more interested in decisions affecting the next generation of military equipment, than configuring a fixed set of platforms already in operational service. Early applications of OR to what Paxson called "systems analysis" were criticised because they did not adequately consider costs [Digby, 1989]. Consequently, systems analysis became a collaborative venture between mathematicians/engineers and economists, differing most significantly from OR in the larger scope and less bounded nature of the problems it addressed.

After a decade where RAND led the formalisation and application of the systems analysis methodology, in 1961 the Kennedy Administration brought RAND staff and systems analysis techniques into the Office of the Secretary of Defense to provide a quantitative basis for broad decision-making problems [Digby, 1989]. Secretary of Defense McNamara installed Hitch as Comptroller and Enthoven as Assistant Secretary of Systems Analysis, and implemented a systems analysis approach still used today known as the Planning, Programming, Budgeting System (PPBS) to ensure that all important funding decisions were based on quantitative evidence. When the U.S. escalated its involvement in Vietnam, McNamara oversaw the collection and analysis of data on an unprecedented scale. Most notorious was the use of body count statistics to measure progress of the war. When Army Generals were surveyed about the utility of the Measurement of Progress system, 2% of respondents felt the system was valid, only 4% believed that the kill ratio based upon body counts was a valuable and necessary indicator, and 61%

thought the body count was often inflated [Kinnard, 2007]. One General made this comment about body counts:

> The bane of my existence and just about got me fired as a division commander. They were grossly exaggerated by many units primarily because of the incredible interest shown by people like McNamara and [General] Westmoreland. I shudder to think how many of our soldiers were killed on a body-counting mission—what a waste.

The obsession with measurement encouraged the counterproductive yet career-enhancing practices of incurring needless civilian and military casualties in pursuit of a higher body count, as well as widespread inflation of reported body counts. However, the Vietnam strategy was more deeply flawed, because it misframed a messy political problem situation as the narrow but more easily quantified problem of maximising Viet Cong attrition. As Summers [1992, pp. 44-51] noted, the Vietnam strategy demonstrated the inappropriateness of PPBS, a management tool for the preparation of war, for the purposes of war proper, a distinction emphasised by Clausewitz. Although Enthoven [2005, p. 270] disingenuously attempted to downplay the role of the Systems Analysis Office in the Vietnam war, most theorists were unconvinced (see for example [Summers, 1992; Hoos, 1969; Gray, 1971; Cohen, 1980; Ilachinski, 2004]). Gray [1971, p. 124] characterised the RAND influence on both Vietnam and the costly nuclear arms race as "a style and a content of defense management that has sought "science" where it was not to be found". This contains a familiar echo of Clausewitz's warnings on the misapplication of science to war. The systems analysis experience is notable both for the scale of impact[2] of an early systems approach, as well as for the lessons to be learned about the dangers of applying overly simplistic quantitative techniques to social and strategic problems.

2.3 Systems Engineering

Another interdisciplinary field motivated by military problems was forming towards the end of the second World War, this time in engineering. The need for systems engineering arose from problems in the design and implementation of solutions to large scale engineering challenges spanning multiple engineering disciplines. A multidisciplinary team of engineers required a lead engineer whose focus was not the design of individual components, but how they integrated. Consequently, management concerns were as significant as technical challenges for a systems engineer. The Bell Telephone Laboratories and Western Electric Company's design and manufacture of the Nike Missile Air Defense System, commenced in 1945, is widely cited as one of the first systems engineering projects.

The surface to air missile defense program integrated ground-based tracking radars, computers and radio controlled anti-aircraft rockets, in order to protect

[2]As well as its military impact, President Johnson directed all federal departments to adopt the DoD's systems analysis approach.

large areas from high altitude bombers. It was novel because unlike conventional anti-aircraft artillery, Nike allowed continuous missile guidance: the radars and computers enabled feedback and control. Bell Labs were the prime contractor for the project, while much of the detailed engineering was undertaken by the major subcontractor, Douglas Aircraft Company. The 1945 Bell Labs report *A Study of an Antiaircraft Guided Missile System* was considered a classic in applied interdisciplinary research due to its depth of insight, scope, and influential role in the systems engineering of Nike [Fagen, 1978].

Following the success of individual systems engineering projects, Bell Labs structured itself around the new systems engineering approach. Bell Labs was organised into three areas: basic research, systems engineering and manufacturing projects [Kelly, 1950]. The systems engineering area provided the interface between advances in communications theories and the manufacture of commercial systems. Because of the "whole system" perspective within the systems engineering area, it was responsible for prioritising the activation of projects with the greatest potential user benefit, within the technical feasibility of communications theory. The responsibility of the systems engineer was the choice of the "technical path" between theory and application in order to create new services; improve the quality of existing services; or lower their cost. Because of its emphasis on user benefit, standards were seen to play a vital role in systems engineering. Standards were used to measure quality of service, which enabled cost benefit analysis of different technical paths.

Outside Bell Labs, Project Apollo was one of the highest profile early successes of the systems engineering approach, which quickly spread from its origins in defence to also become the standard approach to large scale civilian projects. A thumbnail sketch of the traditional systems engineering process is as follows: 1) Customer needs are captured in precise, quantified requirements specifications; 2) System requirements are decomposed into requirements for subsystems, until the requirements for each subsystem are sufficiently simple; 3) Design synthesis integrates subsystems; and 4) Test and evaluation identifies unintended interactions between subsystems, which may generate additional requirements for some subsystems. If there are unintended consequences, the process returns to stage 2, and repeats until the system meets the requirements. Even today, all major defence acquisitions essentially follow this basic schema.

2.4 Cybernetics

Wiener's research on self-correcting missile control systems during World War II was a significant motivation for early cybernetics research, as was research in communications theory stimulated by the war. According to Wiener [1948], who suggested the name for the field, "cybernetics attempts to find the common elements in the functioning of automatic machines and of the human nervous system, and to develop a theory which will cover the entire field of control and communication in machines and in living organisms." Although this ambitious research agenda was never fully realised, cybernetics did inspire novel conceptualisations of war.

The most influential of these was Boyd's OODA loop, which modeled military decision-making as a cybernetic feedback loop. Colonel Boyd, a top Vietnam fighter pilot who subsequently became a critic and reformer of military strategy within the Pentagon, drew on a diverse range of military and scientific theories to develop a comprehensive theory of war. Although Boyd never published his work, the influence of his thinking on U.S. Marine Corps and Army doctrine has since been documented (see for example [Hammond, 2001; Coram, 2002; Osinga, 2005]). Osinga [2005, p. 120], who traces the influence of Boyd's education and reading on his theory of military strategy, identifies the concepts of entropy, feedback, interdependency, adaptation, and the role of experience and genetically inherited traits in shaping mental models as the key influences from cybernetics and systems theory.

The OODA loop, shown in Figure 2.4, contains four steps: observe, orient, decide, and act. A frequent interpretation of the OODA loop is that it demonstrates the importance of rapid decision-making in war, which enables the warfighter to operate "inside the decision cycle" of their adversary. However, this is a shallow and potentially dangerous interpretation of Boyd's work. A broader reading of Boyd's theory places the OODA loop in the context of understanding how information flows and feedback in military decision-making relates to military strategy. By viewing the protagonists in conflict as open systems, Boyd argued that the essence of military strategy involves mentally, morally, and physically[3] isolating the enemy, while mentally, morally, and physically interacting with the society in which the conflict takes place. The OODA loop describes Boyd's mechanism for achieving strategic interaction, and shows the importance of disorienting the adversary by exploiting ambiguity and unpredictability as part of the strategy of isolation.

While Boyd's scientific influences merit a cybernetic classification, Bousquet [2009, p. 187-196] convincingly argues that Boyd emphasised the inevitable uncertainty of war and focused "on the conditions of emergence and transformation of systems through information rather than merely the manner in which information is processed by a fixed organisational schema". Whereas cybernetics focused on negative feedback and sought to eliminate instability, Boyd challenged these assumptions and pioneered an approach that turns out to be much more closely aligned with the insights of chaos and complexity.

2.5 General System Theory

General System Theory (GST) and cybernetics were highly interwoven interdisciplinary approaches to science. GST had biological roots in the research of Berta-

[3]Boyd's emphasis on the mental and moral dimensions of warfare were a direct reaction to what he viewed as the overly quantitative and physical focus of the Pentagon during the Vietnam war.

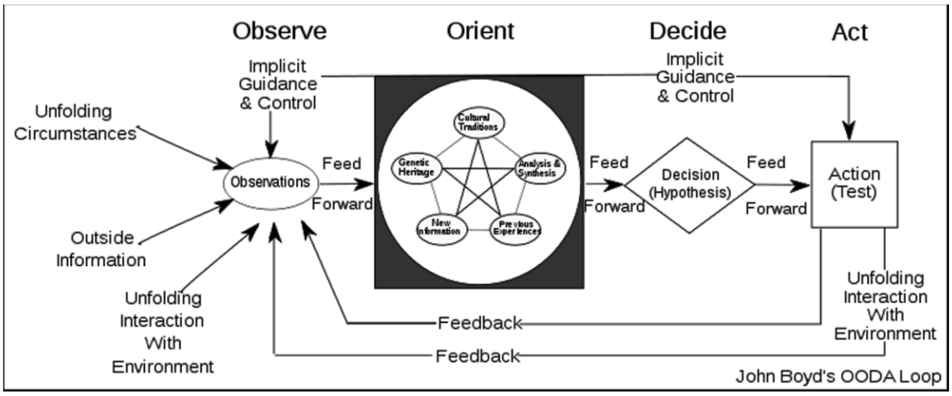


Figure 1. John Boyd's OODA loop.

lanffy, who first published an agenda for a general theory of systems [von Bertalanffy, 1950] in *Science* in 1950. Bertalanffy had even more ambitious aspirations than the cyberneticists, and argued that GST could establish a unified foundation for science [von Bertalanffy, 1956]. Relative to cybernetics, GST was more weakly associated with military applications. Some of the founding members of the Society for General Systems Research were vocal critics of the military-industrial complex, such as Boulding and Rapoport, who established the Center for Peace Research and Conflict Resolution at the University of Michigan in 1956. Bertalanffy himself wrote "I, personally, am not enthusiastic about the applications of systems in industry and politics but they are a fact" [Hammond, 2003].

In Russia, Bogdanov's [1925 to 1929] Tectology, a remarkably similar but independently developed theory, attempted to understand the universal science of organisation. Russian society in general, and the Red Army in particular, was influenced by Bogdanov's approach to systems thinking and the science of organisation. As leader of the Proletkult movement and President of the Academy of Social Sciences, Bogdanov's influence spread through the Scientific Organisation of Labour (*Nachnaia Organizatsiia Truda*) movement to the Red Army, and helped to shape the development of the Russian military's novel theoretical construct of "operational art".

It was not until the 1990s that a significant, explicit military application of GST was developed by Brigadier General (Retired) Naveh, co-founder of Israel's Operational Theory Research Institute. Naveh's PhD thesis, subsequently published as a book [Naveh, 1997], used GST as a heuristic framework for interpreting the history of operational art, from *Blitzkreig* and Soviet Deep Operation theory in the 1920s and 1930s through to its American expression in Airland Battle doctrine in the 1980s. Naveh [Naveh, 1997, p. 9] claimed that the operational level of war — the link between strategy and tactics — is the implementation of Bertalanffy's concept of systems as goal-directed open complex wholes in the military sphere. Following his historical review of the evolution of operational art, Naveh developed a novel systems approach to operations called Systemic Operational Design (SOD). Although Naveh has not yet published his methodology, articles by American colleagues [Wass de Czege, 2009] and students [Davison, 2008] describe how SOD is used to frame understanding of a complex operational environment and design an intervention to transform the system.

3 COMPLEX SYSTEMS IN WARFARE

The founding of the Santa Fe Institute in 1984 led to the recognition of complex systems as a distinct systems approach, which inherited heavily[4] if somewhat implicitly from GST and cybernetics. Complex systems was initially closely associated with chaos theory, both being nonlinear sciences, and sometimes jointly referred to rather exuberantly as the "new sciences". Roughly, the way in which complex systems differentiated from its contemporary, chaos theory, is that where chaos theory studied low dimensional deterministic systems with fixed transformation rules using continuous mathematics, complex systems was more interested with high dimensional non-deterministic systems with discrete, discontinuous and adaptive rules. These systems were more often explored using computer models based on discrete mathematics, such as cellular automata, genetic algorithms, agent based models and boolean networks.

Los Alamos National Laboratory was the first security establishment to embrace complex systems. Its historically significant contribution to nonlinear dynamics (including the work of von Neumann, Ulam and Feigenbaum), its Center for Nonlinear Studies formed in 1980, and its geographical proximity to the Santa Fe Institute positioned it to have a significant impact on the science of complex systems. In 1988, Farmer founded the Complex Systems Group in the theoretical division of Los Alamos. The main research areas were complex networks, nonequilibrium statistical physics, science based prediction for complex systems, and quantum information, with applications to weapons physics research, materials modelling, and nuclear stockpile stewardship [Los Alamos National Laboratory, 2007].

Another early application of complex systems was the development of a cellular automata battlefield simulation by Woodcock, Cobb and Dockery [1988], who showed that the extremely simple and computationally efficient interaction rules of cellular automata could be used to reproduce some of the complex phenomena of combat. They argued that the potential of their approach was not to replicate in detail the dynamics of combat, but rather to provide a new understanding of the maximum amount of behavioural complexity generated from the least complicated set of rules. Their initial study investigated the impact of topographical barriers and rudimentary command and control (with context-sensitive rule sets) on be-

[4]See [Ryan, 2007a] for a comparison between the research agenda of GST and cybernetics with complex systems.

haviour, both of which were difficult problems to solve analytically with formal mathematical models of combat.

In 1992, Beyerchen [1992] published perhaps the first application of the non-linear sciences to the interpretation of the theory of war. Beyerchen suggested that the notorious difficulties of interpreting *On War* were at least in part due to the predominance of a linear approach to analysis, when Clausewitz "perceived and articulated the nature of war as an energy-consuming phenomenon involving competing and interactive factors, attention to which reveals a messy mix of order and unpredictability". Beyerchen traced deep connections between Clausewitz's discussions of interaction, friction and chance with key concepts from nonlinear science, including positive feedback, instability, entropy and chaos. Beyerchen's two major conclusions were that an understanding of complex systems (or at least a non-linear intuition) may be a prerequisite for fully understanding Clausewitz; and that the non-linear sciences may help to establish fundamental limits to predictability in war.

3.1 Setting the Research Agenda

Ilachinski, a scientist at the U.S. Center for Naval Analyses (CNA), provided the first comprehensive assessment of the Santa Fe approach to complex systems as it applied to warfare. Following a review in 1994 of new discoveries capable of improving the profession of arms by the U.S. Marine Corps, in 1996 the commanding General, Marine Corps Combat Development Command (MCCDC), Lieutenant General Van Riper, asked the CNA to assess the general applicability of the new sciences — nonlinear dynamics and complex systems — to land warfare. The result was a report by Ilachinski on Land Warfare and Complexity in two parts. The first part [Ilachinski, 1996a] summarised the state of the art, providing the mathematical foundation for chaos theory and complex systems. Part two [Ilachinski, 1996b] assessed the applicability of complexity to land warfare, which contained the principal novel contributions of the study. Notably, this was the first serious attempt to bring together two previously separate fields — OR and complex systems — with a view to identifying lessons for OR.[5] Ilachinski identified eight tiers of applicability to land warfare, of roughly increasing payoff and risk:

- Tier I: General metaphors for complexity in war
- Tier II: Policy and general guidelines for strategy
- Tier III: Conventional warfare models and approaches
- Tier IV: Description of the complexity of combat
- Tier V: Combat technology enhancement

[5]CNA was the first OR group in the U.S. — it was originally the Antisubmarine Warfare Operations Research Group formed by Morse, with Kimball as deputy director, during World War II.

- Tier VI: Combat aids for the battlefield
- Tier VII: Synthetic combat environments
- Tier VIII: Original conceptualisations of combat

Although Ilachinski cautioned against the use of metaphors alone, he acknowledged their potential to reframe thinking. By replacing some vestigial Newtonian metaphors used by Clausewitz, such as "centre of gravity" and "friction", Ilachinski argued that complex systems metaphors may help to develop "non-linear intuition". The second tier involved changes to organisational and command and control structures to accommodate complex systems insights, for instance by identifying global patterns, exploiting decentralised control, and learning how to adapt better. Tier III would apply the tools of complex systems to develop nonlinear extensions to the Lanchester attrition equations and use genetic algorithms within existing simulation and wargaming models. Tier IV suggested empirical studies of combat data to identify chaotic dynamics, measure the complexity of combat and reconstruct attractors from real world combat data for the purpose of short term prediction. Tier V would improve the technology used in warfare, such as fractal data compression, chaotic encryption techniques and genetic algorithms in manufacturing. In Tier VI, near real-time decision support systems for evolving tactics and adaptive information processing, as well as autonomous ground vehicles could employ complex systems techniques on the battlefield. Tier VII considered the applicability of cellular automata and agent based models to provide a new paradigm for modelling land warfare as a system. Ilachinski's personal research largely continued within this tier, leading the development of the Irreducible Semi-Autonomous Adaptive Combat (ISAAC) and Enhanced ISAAC Neural Simulation Toolkit (EINSTein) cellular automata modelling toolkits for land combat [Ilachinski, 2004]. The last tier was reserved for fundamentally new conceptualisations of combat, which may uncover universal features of combat. Examples given included evolving low-level rule sets that explain observed high level combat behaviours, neural networks capable of inducing un-noticed battlefield patterns, exploiting chaos to drive combat into preferred regions, and using swarms of simple micro-bots to generate collective intelligence. In 2004, Ilachinski published the book *Artificial War* [Ilachinski, 2004], which summarised the CNA research program on complex systems between 1996-2003. The book contains additional detail, but maintains the same eight tiers of applicability of nonlinearity and complexity to warfare.

In his landmark report, Ilachinski [1996b, p. 123] also posed nine open research questions:

1. Are there measures of combat "complexity"?
2. Can patterns of observed chaotic data be exploited?
3. What is an appropriate phase space description for combat?

4. Can the chaos of combat be "controlled" or "tamed"?

5. What are the "optimal" strategies of adaptation on the battlefield?

6. What role does the psychology of the individual combatant play in shaping the combat process?

7. How "complex" must a combat system be in order for it to be amenable to the tools of complex systems theory?

8. How can one quantify the true value and nature of "information" on a battlefield?

9. Does the presence of fractals in combat point to something fundamental?

Progress on the first question has been predictably slow, given the controversy over measures of complexity in the broader field of complex systems. This controversy stems from the proliferation of proposed measures,[6] many of which are unmeasurable in practice, and none of which have gained broad acceptance. This is hardly surprising, since it is unrealistic to expect the reduction of a high dimensional non-determinate dynamical system to a single number to provide any general insight into the nature of the system. However, the narrower but related questions of quantifying the intermittency of attrition data, the use of complexity-inspired measures to predict combat outcomes, and identifying critical points and phase transitions within crowd dynamics, have produced interesting results that advance our understanding of combat complexity [Lauren, 1999; Dobias, 2008].

By the second question, Ilachinski meant that if an aspect of land combat could be shown to be deterministically chaotic, then this knowledge could be exploited for short term prediction. While many authors — most notably James [1996] — have explored this avenue, Hughes' assessment that "the vineyards of chaos theory have been pretty sterile of fruit insofar as military applications are concerned" [Hughes Jr., 1994] holds true to date. One quantitative study by Ingber [1993] investigated whether combat data from Janus wargames and National Training Center combat exercises were chaotic, or merely stochastic. Ingber concluded that "the impulse to cry "chaos!" in combat appears to have been shown to be premature. It is not supported by the facts, tentative as they are because of sparse data."

The third question on developing a phase space description for combat is intriguing, but challenging, due to the high dimensionality of any real combat situation. A number of techniques have since been proposed for exploring combat phase spaces, mostly to visualise the results of agent based models of combat in two or three dimensions. Dobias [2009] has identified sub-critical and critical phases in combat casualty data, which is discussed below.

[6] Even by 1997, Edmonds [1997] had collected 386 references offering distinct measures of complexity.

Question four was inspired by the results of Ott *et al.* [Ott *et al.*, 1990] and Romeiras *et al.* [1992], who showed that small controlling perturbations could stabilise chaotic dynamics and even "direct" chaotic trajectories towards a desired end state. While theoretical developments on the exploitation of chaos have continued, as with question two, these results have not yet found military application.

Question five asked whether complex systems theory could provide guidance on the best way to adapt to a changing environment. Although the question as stated here is framed with an unmistakable OR bent, the research topic of fostering better force adaptation is a rich and central theme within the military applications of complex systems, and will be discussed in detail below.

Question six may not initially appear to be a question of complex systems. However, this question was motivated by the tendency of OR combat models to include detailed physical models of projectiles, platforms and vegetation, whilst almost entirely ignoring intangible yet causally significant factors of human psychology, such as morale, stress and fear, as well as how the environment shapes the combatant's behaviour. The emphasis within complex systems on modelling interactions between a system and its environment provides the impetus for better understanding the role of individual psychology within combat. Project Albert, discussed below, made significant advances in the inclusion of intangibles in agent based models of combat.

The seventh question is actually one of scale rather than complexity. Ilachinski intended to determine the minimal temporal and spatial scales at which complex systems techniques became relevant. The domain of applicability of systems techniques has always been rather fuzzy, but as Weaver [1948] noted, "organised complexity" lies somewhere between the realm of mechanics (several components) and statistical mechanics (several billion components). This implies there is also an upper limit to the scale of systems where complex systems techniques apply, beyond which individual components do not disproportionately influence collective dynamics, and the assumptions of classical statistics (such as the law of large numbers) hold. It is not possible to give precise upper and lower limits in general, because the domain of complex behaviour depends on the nature, strength and topology of interactions between components in addition to the number of components. However, quantitative studies by Lauren [1999, p. 16] using ISAAC show that for given parameter settings, groups of homogeneous automata larger than 15 tend to display mob behaviour. The model then has fewer degrees of freedom and behaves more simply, because the automata constrain themselves to fighting as a single group. Lauren also showed that adding more automata types increases complexity, allowing the force to break up more readily into groups that then behave similarly to individual automata. These results confirm the importance of heterogeneity for complex behaviour, and show that even small numbers of agents are capable of complex behaviour.

Ilachinski's eighth question was also motivated by the exclusively physical nature of OR combat models, which simulated combat as the movement of matter (platforms such as tanks and helicopters) and the exchange of energy (kinetic and high explosive weapons). Within complex systems science, flows of information are at least as important as matter and energy, so the challenge is to understand combat "not just as a set of local fire-storms, each consisting of physical skirmishes among individual combatants, but as a complicated interleaving network of physical action resulting from local interpretations of local information" [Ilachinski, 1996b, p. 126]. Over the last decade, agent based models of warfare have been designed that incorporate local interpretations of local information, and network theory has been used to better capture information flows in distributed networked forces, even though a theoretical expression of the value and nature of information in combat has not been achieved.

The last question has seen the most theoretical progress over the last decade. Following Richardson, the significance of fractals (scale invariant statistical distributions) in combat casualties have been further explored. Power law statistics have been shown to naturally arise in agent based models of combat [Lauren, 2001], a fractal attrition equation has been proposed as a refinement of Lanchester's model of combat [Lauren *et al.*, 2007], and mechanisms to explain the power laws for casualties in wars and terrorist attacks have been proposed [Clauset *et al.*, 2007; Johnson *et al.*, 2006; Dobias, 2009].

The U.S. Marine Corps and CNA collaboration on complex systems was exposed to the broader defence community in the U.S. when the Military Operations Research Society organised a Mini-symposium and Workshop on Warfare Analysis and Complexity in 1997 [Palmore, 1997]. The Mini-symposium included presentations from the distinguished complexity theorist Kauffman, as well as Ilachinski and Lieutenant General Van Riper, that spanned from basic science to the military challenges of complexity. During the workshop, six working groups were formed to explore defence applications of complexity. Each group identified many applications for complex systems tools, techniques and approaches during brainstorming sessions, although in the workshop report, the excitement surrounding the "new sciences" was frequently tempered by cautions that traditional OR techniques would still be needed, and that the greatest challenge was "integrating "new" and "old" ideas without much hyperbole or insult" [Palmore, 1997, p. 3]. The workshop found that:

> The most important single implication for warfare analysis that emerges from the mini-symposium is the *centrality* of planning under uncertainty by emphasizing robust capabilities and adaptations. This will require fundamental changes in the paradigms under which the Department of Defense, the military departments, and the analysts who serve them have typically operated for decades.

The central finding was thus a broadening of Ilachinski's fifth open question, from adaptation on the battlefield, to considering how the organisations supporting the warfighters could help to deliver a more robust and adaptive capability. Davis, a Senior RAND Analyst who helped organise the Mini-symposium, had already published on this topic as early as 1993 [Davis and Finch, 1993], and would later apply the same ideas towards more robust and adaptive modelling and simulation of complex warfare [Davis and Henninger, 2007].

3.2 Applications to Military Problems

The growing interest in complex systems within the Marine Corps, exemplified by Ilachinski's report on the applicability of complexity to land warfare, was evident in the 1997 update of the Marine Corps' primary manual *Warfighting* [Headquarters U.S. Marine Corps, 1997a]. In the foreword to the manual, General Krulak, commandant of the Marine Corps, listed the need "to emphasize war's complexity and unpredictability" as the first motivation for the update of the manual. *Warfighting* described the nature of warfare as fundamentally complex.

> [W]ar is not governed by the actions or decisions of a single individual in any one place but emerges from the collective behavior of all the individual parts in the system interacting locally in response to local conditions and incomplete information. A military action is not the monolithic execution of a single decision by a single entity but necessarily involves near-countless independent but interrelated decisions and actions being taken simultaneously throughout the organization. Efforts to fully centralize military operations and to exert complete control by a single decisionmaker are inconsistent with the intrinsically complex and distributed nature of war [Headquarters U.S. Marine Corps, 1997a, p. 12].

The doctrinal publications for planning [Headquarters U.S. Marine Corps, 1997b] and command and control (C2) [Headquarters U.S. Marine Corps, 1996] explicitly included sections on complexity theory. Complexity justified the need for planning, while at the same time recognising the limits on prediction and precise, positive control. For C2, complexity was used to explain the nonlinear phenomena of war, the complex web of interacting causes and effects, and described "the military organization as an open system, interacting with its surroundings (especially the enemy), rather than as a closed system focused on internal efficiency" [Headquarters U.S. Marine Corps, 1996, p. 46]. The C2 doctrine described command and control as a continuous feedback cycle of reciprocal influence, and used a simplified version of Boyd's OODA loop as the conceptual model for C2.

In 1998, the congressionally funded U.S. Marine Corps initiative Project Albert began a nine year international program on complex systems modelling. Project Albert held 12 workshops, and published a series of proceedings named *Maneuver Warfare Science* [Hoffman and Horne, 1998] between 1998 and 2003.[7] Led by the U.S., Project Albert had strong international participation that included

[7]Following the end of Project Albert funding in 2006, the Simulation Experiments & Efficient Designs (SEED) Center for Data Farming at the Naval Postgraduate School continued the Project Albert research agenda, organising bi-annual International Data Farming Workshops, with the proceedings published in *Scythe* [Meyer and Horne, 2007].

modelling teams from Australia, Canada, Germany, Mexico, Netherlands, New Zealand, Norway, Portugal, Singapore, South Korea, Sweden and the U.K. The Project Albert approach was question-based. It involved bringing together teams from the different nations, each with a real world question that traditional OR techniques could not adequately address, and where there was potential for complex systems to provide some insight.

Three significant military subjects not easily modelled with traditional OR combat models, but amenable to complex systems analysis, are manoeuvre warfare, network centric warfare, and irregular warfare. Manoeuvre warfare, championed by Boyd and contrasted with attrition warfare, is a theory of how agile movement in the battlespace can generate shock and surprise to achieve mission objectives, rather than relying on "grinding" through the enemy's source of power (the kind of attrition strategy that led to the horror of the Battle of the Somme). Manoeuvre is thought to be more important on the increasingly lethal modern battlefield, because the advantages of mass offensives can more readily be outweighed by the risks of providing the opponent with a dense target vulnerable to long range weapons. Because manoeuvre warfare eschews attempts at centralised control and increased predictability in favour of decentralised execution and exploiting uncertainty, it is a natural fit with complex systems.

Network centric warfare, in simple terms, is the application of information and communications technologies to reduce the stovepiping of military capabilities, such that command and control arrangements can dynamically connect sensors, decision-makers and engagement assets to make better use of latent resources. Network centric warfare is best viewed as an integrating concept, concerned with how to coevolve a network of capabilities to deliver the best operational effect. Similar to traditional military command and control, traditional OR combat models typically contain fixed hierarchical C2 architectures, and as such are unsuitable for measuring the value of networking without fundamental modifications.

The proponents of network centric warfare explicitly draw on complexity theory. For example, Cebrowski and Garstka [1998] explain their central concept of self-synchronisation as "the ability of a well-informed force to organize and synchronize complex warfare activities from the bottom-up. The organizing principles are unity of effort, clearly articulated commander's intent, and carefully crafted rules of engagement". However, as Bousquet [2009, p. 229] notes, self-synchronisation is actually a misinterpretation of the principles of complexity, since self-organisation is driven by local information, whereas network centric warfare seeks to achieve self-synchronisation with a centralised, global "common operating picture". Bousquet [2009, p. 7] argues that "despite nods to chaos theory and complexity science, [network centric warfare] is found to be still largely in thrall to the principles of cybernetic warfare", by which he means the objective is predictability and scientific control over war. While the validity of network centric warfare's interpretation of complexity theory can and should be debated, the importance of complexity theory for implementing network centric concepts is clear.

Irregular warfare is characterised by asymmetry, where a powerful conventional force is challenged by an uprising that is incapable of directly confronting the conventional force, but seeks to outlast it through destabilisation and popular revolt. Reflecting on leading a successful revolt in Arabia, Lawrence observed that “irregular war is far more intellectual than a bayonet charge” [Lawrence, 1990], which vividly contrasts the subtle exploitation of asymmetries with the brute force of attrition warfare. One of Lawrence’s tactical principles for insurgency was to achieve what he called maximum articulation. “The Arab war was simple and individual. Every enrolled man served in the line of battle, and was self-contained. We had no lines of communication or labour troops. The efficiency of each man was his personal efficiency. We thought that in our condition of warfare the sum yielded by single men would be at least equal to the product of a compound system, and it was certainly easier to adjust to tribal life and manners, given elasticity and understanding on the part of the commanding officers.” The principle of maximum articulation, discovered by Lawrence through practice, is supported in complex systems theory by Bar-Yam’s law of multiscale variety [Bar-Yam, 2004]. Because there is a tradeoff between scale and complexity of an organisation, insurgents, who cannot compete with conventional forces on the basis of scale alone, seek to increase the fine scale complexity of the operational environment by operating in small autonomous groups [Bar-Yam, 2005; Ryan, 2006]. Because the dynamics of irregular warfare are attracted towards increasing fine scale complexity, complex systems provides a more appropriate paradigm than traditional OR, which due to its origins in large scale conflict in World War II, emphasises the importance of coordination to optimise large scale effects.

Horne, the executive director of Project Albert, described the two unifying themes within Project Albert as Operational Synthesis and Data Farming. Operational Synthesis means integrating agent based models, aggregate equations, high fidelity simulations, and wargames/exercises to “answer questions involving one or more of the phenomena of nonlinearities, intangibles, or coevolving landscapes” [Horne, 2001]. Data Farming, first proposed by Horne [1997] prior to Project Albert, means using agent based models; supercomputing to explore vast spaces of possibilities; and analytical techniques capable of manipulating, organising and visualising large amounts of data. Data Farming, a playful twist on the metaphor of data mining, involves ‘fertilising’ the minds of military professionals with new ideas for modelling conflict, ‘cultivating’ ideas on the important factors for the question of interest, ‘planting’ these ideas within the model to grow data across the parameter landscape, and ‘harvesting’ the output data for analysis [Brandstein and Horne, 1998]. The Maui High Performance Computing Center was a critical enabler for Data Farming, because it allowed modelling teams to grow new data landscapes every night during Project Albert workshops, which provided rapid feedback to shape subsequent inquiry.

Ilachinski’s early work on the military applications of complex systems was clearly a significant influence within Project Albert. Applications within all eight of Ilachinski’s tiers of applicability were published in *Maneuver Warfare Science*, even if like Ilachinski, Project Albert participants tended to focus on Tier VII — the applicability of the new agent based modelling paradigm. The ISAAC and EINSTein toolkits developed by Ilachinski formed the basis for much of the Project Albert modelling up until about 2002. Ilachinski's open research questions, particularly three, six and nine received repeated attention. By moving away from the traditional OR approach of detailed modelling of point scenarios towards exploring high dimensional landscapes, Data Farming began to develop phase space descriptions for combat. The emphasis on combat intangibles — morale, leadership, timing, intuition, and adaptability [Brandstein and Horne, 1998, p. 96] — can be seen as a response to question six on the role of individual psychology in combat. New Zealand provided some of the most important contributions to Project Albert, including work that would provide one answer to Ilachinski's ninth question on the significance of fractals in combat.

In 1998, U.S. Marine Corps Chief Scientist Brandstein and Project Albert director Horne visited New Zealand's Defence Technology Agency (DTA[8]). During the visit, DTA operations analyst Lauren obtained a copy of ISAAC and was invited to attend the inaugural Project Albert workshop in Maui in 1999. Following the visit, Lauren quickly recognised the significance of simple cellular automata models of combat because they could replicate emergent features of combat that were entirely absent from conventional detailed physical combat models. More importantly, Lauren was able to statistically differentiate between what he characterised as complex adaptive and conventional combat models. Lauren's [Lauren, 1999] initial report consisted of a number of small experiments that were at this early stage suggestive rather than comprehensive.[9] Even so, the report already contained many of the insights that would shape the way complex systems models could be used to represent combat. Lauren's critique of the conventional combat models developed for military OR consisted of two claims: conventional combat models are little more than a sequence of Lanchester engagements, and they do not capture the non-linearities of war.

There are many problems with the Lanchester equations as a model of combat,[10] but the shortcomings of concern to Lauren were that manoeuvre could never impact upon the attrition rate, and the differential equations resulted in continuous attrition, when historical conflict showed "clumpiness", where casu-

[8] At the time DTA was named the Defence Operational Technology Support Establishment.

[9] For more extensive excursions and analysis see [Lauren, 2001; Lauren, 2002a; Lauren, 2002b; Lauren and Stephen, 2002a; McIntosh and Lauren, 2002; Lauren, 2005; Lauren *et al.*, 2007; McIntosh and Lauren, 2007].

[10] Notable complex systems modeller Epstein [1988] summarised the four most significant ways the Lanchester equations fundamentally misrepresent combat: defensive withdrawal confers no benefit; the most fundamental tactic in military history, the trading of space for time, is precluded; the squaring of mass but not firing efficiency is a baseless implicit assumption; and there is no empirical support for the Lanchester equations. Further limitations of the Lanchester equations are that the original form of the Lanchester equations assume that every combatant is within detection and engagement range, forces are homogeneous, fire is uniformly distributed over surviving units, no force to space constraints ever set in, and factors such as manoeuvre, logistics, command and control and morale have no effect on casualties.

alties would come in bursts. Even though conventional combat models allowed manoeuvre, this would typically serve only to move opposing forces into combat range, where a Lanchester-style battle would occur until some arbitrary level of attrition was reached, at which point the forces would manoeuvre to a new position for the next Lanchester clash. Therefore, for all the physical detail of the movement of military1 platforms, conventional combat models reduce to a sequence of attrition-driven fights approximated by the Lanchester equations. Further, Lauren argued that many non-linear features of combat are neglected in conventional combat models, including self-organisation, co-evolution, adaptation, attractors, phase transitions, extreme events, and combat intangibles.

In contrast, Lauren's initial experiments with ISAAC naturally generated bursts of casualties, even without explicit programming of tactics. Lauren described two qualitatively different phases: the linear battlefield, where forces would form a stable linear battle front, and a turbulent battlefield, which was characterised by fractal statistics. For the turbulent battlefield, probability density functions estimating the number of time steps required to reach 50% casualties exhibited fat tails, and in some cases the variance diverged as more replications of the simulation were run, suggesting the extreme values belong to separate statistical populations. The turbulent battlefield exhibited a fractal geometry, which led Lauren to propose a deeper connection between combat and fluid dynamics. An analogy between combat casualties on the turbulent battlefield and Kolmogorov's proposed statistical structure function for turbulence led to the relation:

$$\left\langle |n(T_0 + t) - n(t_0)|^2 \right\rangle \propto t^{2/3} \tag{3}$$

where n is the number of remaining automata, T_0 is a reference time, t is time, and the exponent of 2/3 follows Kolmogorov. Because the structure function depends on the fractal dimension, it accounts for spatial organisation in a way that the Lanchester equation cannot. Further, it does not assume either continuity or smoothness, meaning it is consistent with clumpy attrition data. Lauren speculated that this structure function might be the "new Lanchester equation" of the science of complexity. This was an application Ilachinski had foreseen — non-linear extensions to the Lanchester equation — in Tier III of the applicability of complexity to land combat.

Lauren's initial experiments with ISAAC showed that seemingly simplistic cellular automata models of combat were in fact deceptively rich and genuinely novel. However, limitations of the available tools led Lauren, with Stephen and other DTA colleagues, to develop a new cellular automata agent based modelling platform called Map Aware Non-uniform Automata (MANA) [Lauren and Stephen, 2002b]. MANA, inspired by ISAAC, quickly became the most popular combat modelling environment due to its ease of use, engaging graphics, and novel features. MANA agents have access to their squad's map of the world, which can differ from the ground truth, and communication between squads allows the sharing of contact memories. This means that MANA can model different levels of situational awareness, and answer questions about the value of information on the battlefield (Ilachinski's eighth question). Event-driven personality changes can occur at individual and collective levels, allowing dynamic behavioural patterns in response to events such as reaching a goal, getting injured, or contact with the enemy. This enables many combat intangibles to be represented via the agent personality weightings and trigger states.

MANA was applied to the focus topics identified by Project Albert above: manoeuvre warfare [Lauren *et al.*, 2002], network centric warfare [Galligan, 2004], and irregular warfare [Lauren, 2002b]. Other applications included military operations in urban terrain, convoy protection, humanitarian operations, peace support operations, use of non-lethal weapons, effects of chemical and biological weapons, the global war on terrorism, homeland defense, and historical events [Engleback *et al.*, 2003]. However, MANA was just as important to the science of military complexity research as its application, by providing a laboratory for *in silico* experiments on the presence of fractals in combat. Subsequent collaborations involving Lauren generalised equation 3 by considering non-integer exponents for t given by the fractal dimension D, and showed how this corresponds to the Lanchester equations when the agents are evenly distributed on the battlefield [Lauren and Stephen, 2002a; Lauren *et al.*, 2007]. For agents on a two dimensional grid,

$$E\left(\frac{\Delta B}{\Delta t}\right) \propto \alpha_R^{D/2} \Delta t^{(D/2-1)} R(t). \tag{4}$$

It is clear that by substituting $D = 2$, this reduces to a form consistent with equation 1. However, the ability to also predict the attrition for fractional values of D means Lauren's metamodel quantifies the value of increased organisation for a force in terms of a decreased casualty rate. A high fractal dimension implies coordination across multiple scales, which requires effective C2, timely information flow, and resilience against disruptive enemy action. Maintenance of a high fractal dimension allows a force to concentrate its firepower, which explains why it is an important determinate of the attrition rate.

Back in the U.S., the progress of the U.S. Marine Corps had been noticed by the National Defense University (NDU), RAND and the Department of Defense Command and Control Research Program (DoD CCRP[11]), who organised a symposium and published a collection of papers on Complexity, Global Politics, and National Security [Alberts and Czerwinski, 1997]. The purpose of the symposium, hosted at the National Defense University in 1996, was to promote the application of complexity in the policy and strategic domains of the national security arena. Gell-Mann presented a paper that drew heavily on his book *The Quark and the Jaguar* [Gell-Mann, 1994], which called on the need for "a crude look at the whole". The purpose of the paper was to apply the notion of sustainability — quality that is not purchased mainly at the expense of the future — to defence and security. Gell-Mann called for sustainable peace, sustainable preparedness for possible conflict, sustainable global security arrangements, sustainable democracy

[11]The DoD CCRP is the lead U.S. agency for Network Centric Warfare.

and human rights, and sustainable communities and institutions, emphasising that these challenges were all closely interlinked. In addition to the symposium proceedings, two books on the military applications of complexity theory would be published by the DoD CCRP and NDU. In 1998, Czerwinski's [Czerwinski, 2008] book developed a neo-Clauswitzian view of war built around Beyerchen's article, which was reproduced in full. The second book, written by British scientist Moffat [2003], examined the applications of complexity theory to network centric warfare, particularly focused on modelling and OR. Moffat reviewed Lauren's work in detail, and provided an alternative derivation of equation 4 by assuming that warfare behaves as a scale-free system.

The New England Complex Systems Institute (NECSI), best known for hosting the International Conference on Complex Systems, became involved in military problems through several of its educational initiatives. Interestingly, the military applications stimulated by NECSI's courses in complex systems were the first not to fit within Ilachinski's tiers of applicability, nor were they addressed by his open questions. The interpretation of complex systems provided by NECSI's president, Bar-Yam, injected a different perspective and started new research threads that had not previously been pursued by the military OR community. At the Strategic Studies Group of the Naval War College, Bar-Yam taught complex systems and analysed military networking initiatives and future operating concepts from a complex systems perspective. The results were summarised in Bar-Yam's book *Making Things Work*, and also helped inspire Cares' book *Distributed Networked Operations*. Bar-Yam's principal contributions were to show how the tradeoff between scale and complexity in organisations changes the way forces should be organised for irregular warfare, and to investigate the importance of boundary formation in conflict resolution [Bar-Yam, 2005]. The latter was elaborated in a 2007 *Science* article, which analysed the process of global pattern formation that caused cultural differentiation, to investigate the relationship between group size and ethnic violence [Lim *et al.*, 2007]. Cares developed a novel model for networks that fight other networks, and developed an alternative foundation for network centric warfare more consistent with the principles of complex systems [Cares, 1996].

NECSI was exposed to a different class of military problems when Bar-Yam provided complex systems courses for executives and systems engineers from the MITRE Corp. These interactions, also discussed in *Making Things Work*, led Bar-Yam to develop an approach he called 'enlightened evolutionary engineering'. Bar-Yam challenged the existing engineering paradigm of designing complex systems from scratch, and proposed an alternative approach that utilised a nuanced understanding of biological evolution to grow solutions in context that featured overlapping generations of new and legacy subsystems. Norman and Kuras, MITRE engineers, subsequently used complex systems theory to explain how they successfully designed the Air and Space Operations Center for the U.S. Air Force, despite the fact that they did not centrally control or manage the funding or development of the 81 major subsystems, nor was the specific desired outcome for the project ever specified in advance. Because the systems engineering methodology outlined in Section 2.3 does not apply under these conditions, Norman and Kuras developed an alternative ‘regimen’ based on complex systems insights. Their regimen proposed creating a development environment for accelerated evolution, where developmental precepts force the rules for interaction, and continual selection pressure rewards results achieved in context.

In the Australian Army, the recognition of the central importance of complexity was marked by the publication of the Future Land Operating Concept Complex Warfighting [Australian Army Headquarters, 2004] in 2004. Within the Future Land Warfare Branch of Australian Army Headquarters, Brigadier Kelly and Lieutenant Colonel Kilcullen wrote Complex Warfighting, which became the endorsed Army assessment of the future operating environment, and framed an integrated land force response to the demands of complex war. As the capstone warfighting concept for Army, the FLOC shapes all modernisation efforts, including capability development, doctrine updates, future Army Objective Force design, the raise, train and sustain functions of the current force, and experimentation.

Complex Warfighting assessed complex multidimensional security environments, including the physical, political, cultural, economic and informational dimensions. Two environmental drivers, globalisation and U.S. military dominance, were identified as significant to land warfare. Globalisation was seen as a source of real and perceived inequalities, increasing the contact and interactions between competing ideologies, and extending the interests of nation states well beyond their geographical boundaries. U.S. military dominance provided a rationale for asymmetric avoidance behaviour by likely adversaries, who would avoid massing conventional forces in favour of irregular methods of war.

From these drivers, Complex Warfighting concluded that the future environment would be characterised as complex, diverse, diffuse and lethal. The source of complexity derived from a dynamic, interdependent, multilateral environment, with adversaries retreating into complex physical, human and informational “terrain”. The diversity was due to the sheer number of actors. Rather than a homogeneous “Red versus Blue” bipolar engagement (the sort of conflict assumed by the Lanchester equations), the future environment would involve heterogeneous alliances, non-military Governmental instruments of power, a plethora of non-Government organisations and non-state actors, contractors, and a set of adversaries ranging from conventional elements to militia, insurgents, terrorists, criminals and gangs within the conflict ecosystem. Diffusion refers to the increasing irrelevance of traditional distinctions between peace and war, strategic, operational and tactical levels of war, state and non-state actors, and conventional versus special operations. The implication of increasing lethality is that non-state actors and even individuals have access to weapons capable of producing mass casualties with strategic implications. The broad response to these challenges called for small, highly autonomous, modular combined arms teams, capable of swarming and reconfiguring in different ways.

The Australian Defence Science and Technology Organisation (DSTO) had been engaged in complex systems research since the early Project Albert meetings.

So had the Australian Defence Force Academy at the University of New South Wales, particularly the Artificial Life and Adaptive Robotics Laboratory headed by Abbass, who supervised among other projects the development of the first fully networked agent based model ABM, WISDOM-II [Yang *et al.*, 2005] and evolutionary computational methods for red teaming[12] [Yang *et al.*, 2006]. However, a recognised research community did not form until 2004, when the inaugural *Complex Adaptive Systems for Defence* workshop attracted over 100 participants. The workshop series was an initiative of a long range research task in complex adaptive systems for defence, led by Grisogono and managed by Ryan.

In 2006, The Technical Cooperation Panel (TTCP) — an inter-governmental collaboration on defence science between the U.S., U.K., Canada, Australia and New Zealand — charged Grisogono with organising an international Symposium on Complex Adaptive Systems for Defence with broad participation across technological, systems, operational, and human science domains within TTCP, in order to assess how they should address this emerging area of defence science. The result was the formation of the international Joint Systems Analysis Action Group on Complex Adaptive Systems for Defence. Grisogono chaired the Action Group, which initiated both basic and applied cooperative research activities in complex systems science. Many of the early contributors to military complex systems research — Lauren, Horne, Davis, Moffat, Cares, and Norman — participated in the DSTO and TTCP workshops. The four active nations (Australia, U.K., Canada and the U.S.) in the Action Group took turns hosting twice-yearly workshops and sought to develop collaborative connections not only between the defence scientists involved, but also with defence stakeholders and complex systems researchers in other government organisations, academia and industry in each country. The Action Group's research thrusts included further development of conceptual frameworks encompassing processes and phenomena such as adaptation and causal and influence networks, in order to identify generic concepts that could be applied in the two main categories of defence issues: systems and operations.

Adaptation and emergence were central themes of this effort and led to the development of a very generalised and rich view of adaptation [Grisogono, 2005; Grisogono, 2006a] and a novel framework for understanding the central systems concept of emergence [Ryan, 2007b]. This provided a theoretical foundation for a series of insights into implications for command and control [Grisogono, 2004a; Grisogono, 2006b], applications to operational concepts [Smith Jr. and Grisogono, 2006; Ryan, 2006; Grisogono and Ryan, 2007; Smith Jr. *et al.*, 2008], system design [Ryan and Grisogono, 2004; Ryan *et al.*, 2005] and organisational and force design and structure [Grisogono, 2004b; Ryan and Norman, 2006; Unewisse and Grisogono, 2007; Grisogono and Spaans, 2008; Grisogono and Unewisse, 2009; Spaans *et al.*, 2009].

Within the Australian Army, the research of the TTCP Action Group helped shape doctrine, operations, and force structure. Grisogono and Ryan were asked

[12]Red teaming is a formal devil's advocate role played from an adversary's perspective intended to challenge the assumptions of military planners.

to assist Future Land Warfare in devising the Army's response to the challenge laid out in Complex Warfighting. The strongest theme of the TTCP Action Group — that adaptation is the best way to cope with complexity — provided a scientific foundation [Grisogono and Ryan, 2007] for Future Land Warfare's new capstone doctrinal concept, Adaptive Campaigning [Australian Army Headquarters, 2006]. Adaptive Campaigning introduced the concept of operating across five simultaneous, interdependent and mutually reinforcing lines of operation: joint land combat, population support, indigenous capacity building, population protection, and public information. The purpose was to recognise that traditional military combat actions alone could not be decisive in complex warfighting: to cause an attractor change to a new stable society, multiple concurrent actions are required. Adaptive Campaigning also proposed a modified form of Boyd's OODA loop, comprising the sequence Act — Sense — Decide — Adapt. Act is first in the cycle to emphasise the need to act on incomplete information. Sense follows action to ensure that the response of the system to stimulus is collected. The decide function evaluates the implications of the interaction, while adapt explicitly recognises the need to implement decisions by changing the structure or behaviour of one's own system.

At the same time Adaptive Campaigning was being written, the Australian Army was deploying the first Reconstruction Task Force (RTF) to Afghanistan. As an engineer-led combined arms team, the RTF was a novel force package, and the commander, Lieutenant Colonel Ryan, placed a science and technology request with DSTO to "help win the adaptation battle". Lt Col. Ryan described how complex systems informed his operations: "Possessing this 'human map', continually updated using complex adaptive systems theory, equipped the commanders and staff with insights into the dynamics of the province and the ability to assess the impact of RTF projects on the various local actors" [Ryan, 2007c, p. 129]. The RTF's adaptability was enhanced through ensuring the results of after action reviews informed planning for future operations, formal weekly training on lessons learned, the development of web pages to capture lessons from the experiences of others, as well as reachback support to DSTO on adaptive approaches to military operations [Ryan, 2007c, p. 139]. A significant innovation of the 1st RTF was to explicitly plan counter-adaptation operations, designed to inhibit the Taliban's ability to adapt. Lt Col. Ryan's performance was recognised in 2008 when he was awarded Member of the Order of Australia for his leadership, planning and execution of the 1st RTF deployment.

In 2008, the Australian Army extended its adoption of an adaptive approach from the way it conducts operations to the way the entire army is structured. The Chief of Army, Lieutenant General Gillespie, released a public information paper[13] on the Adaptive Army initiative, which outlined the biggest restructuring of the Army since 1973 [Australian Army Headquarters, 2008]. The aims of the restructure included improving "the quality and timeliness of information flows throughout Army in order to enhance Army's adaptation mechanisms at all levels." This was achieved by organising the Army into two Commands responsible

[13]Lt Col. Ryan was a significant contributor to the public information paper.

for adapting at different time scales. Headquarters 1st Division was given responsibility for the shortest adaptation cycle of feeding lessons learned from current deployments back into preparations for the next rotation. Forces Command was charged with the longer adaptation cycle of raising, training and sustaining forces for future threats. This organisational structure facilitates the balance between being adapted to the current fight, and adaptability for future operations. Support from the TTCP Action Group informed the restructuring and experimented with improving individual adaptability, drawing on the theory and microworlds developed by Dörner [1996].

In Canada, the TTCP Action Group influenced the development of the Integrated Capstone Concept for the future Canadian Force, which identified the need for integrated and adaptive capabilities for a comprehensive, networked approach to the complex operational environment. Defence Research and Development Canada (DRDC) recognised complex systems science as a new field of research, and initiated new complex systems projects within operations research, complex systems engineering, and complex cognition. In the U.K., complexity research primarily focused on countering the threat of improvised explosive devices, using complex systems models to better target insurgent networks.

In the U.S. in 2008, a collaboration between MITRE and Australia's DSTO resulted in a research agenda for complex systems engineering. The study identified the six most relevant elements of complexity science for systems engineering as complexity, self-organisation, adaptation, autonomous agents, phase changes, and emergence, which framed a list of open research questions for complex systems engineering [DeRosa *et al.*, 2008]. Also in 2008, the U.S. Office of the Secretary of Defense (OSD) ordered a review of adaptability training, which in addition to U.S. efforts, identified Israel's Systemic Operational Design (SOD) and Australia's Adaptive Campaigning as promising approaches to adaptability [Burns and Freeman, 2008]. At the School of Advanced Military Studies (SAMS) in the U.S. Army Command and General Staff College, SOD had already been taught by Naveh for several years, and explicit links were being drawn between SOD and Adaptive Campaigning [Wass de Czege, 2009]. The TTCP Action Group on Complex Adaptive Systems for Defence engaged with both SAMS and OSD, which resulted in collaboration on adaptability training, and an attachment for Ryan at SAMS. Ryan taught complex systems[14] as a theoretical foundation for design, introduced microworld experiments into the SAMS curriculum to improve adaptability, and participated in the writing of U.S. Army doctrine on design. As discussed above, Israel's SOD was based in GST, and began by questioning what had emerged within the operational environment to invalidate the previous strategy. As the U.S. Army developed its own methodology for design, the concepts of complexity, self-organisation and adaptation were added to include insights from complex systems theory. Design provided a methodology for framing and refram-

[14]Founding SAMS professor Schneider was quick to realise the significance of complex systems for military theory, which he introduced into the SAMS curriculum in 1996. Schneider also published on the importance of chaos and complexity for information warfare [Schneider, 1997].

ing problems to continuously adapt the theoretical framework used to understand the operational environment and plan extended operations and campaigns. Collaborative research between DSTO, DRDC, OSD and SAMS under the umbrella of the TTCP Action Group continues to explore the utility of microworlds for improving individual and team adaptability, as well as the application of complex systems to operational design.

Recently, Richardson's pioneering research on the scale invariant relationship between the frequency and magnitude of conflict has received renewed attention due to the availability of detailed databases of casualties and injuries from terrorist attacks. A study by Santa Fe Institute and University of New Mexico researcher Clauset *et al.* showed that the relationship $P(x) \propto x^{-\alpha}$ held for worldwide terrorist attacks since 1968, with $\alpha = 2.0\hat{2} \pm 0.09$ for OECD nations, $\alpha = 2.5\hat{1} \pm 0.07$ for non-OECD nations [Clauset *et al.*, 2007], and $\alpha = 2.\hat{0}4 \pm 0.2$ overall [Clauset *et al.*, 2009]. Their statistical analysis of the data showed that severe terrorist events, such as the 1998 Nairobi car bombing and the 2001 New York City attack are not outliers, but can be explained from the same statistical distribution that explains the many smaller and more common terrorist attacks. One implication is that current models for predicting the incidence of terrorist attacks do not account for the fat tail of terrorist attack severity, therefore they dramatically underestimate future casualties from terrorism. To explain the power law, Clauset *et al.* suggested an exponential sampling mechanism where states and terrorists compete to decide which planned events are realised.

An alternative explanation was provided by Johnson *et al.* [2005; 2006; 2008], who developed an agent based model with self-organised critical dynamics [Bak *et al.*, 1987] that replicated Clauset's non-OECD power law exponent $\alpha = 2.5$. Johnson *et al.* also showed that while the power law coefficient for Iraq was initially lower than 2.5, and the coefficient for the Colombian conflict was initially higher than 2.5, both were converging towards 2.5. Clauset and Wiegel [2009] subsequently generalised Johnson's model, showing analytically that power law distributions are a universal feature for a family of fission-fusion models of the formation and disintegration of terrorist groups. Further, they showed that providing the number of "terrorism inclined" individuals is significantly large, the scaling coefficient α is independent of the causes of terrorist cell disintegration. Dobias, a Canadian scientist at DRDC and a member of the TTCP Action Group on Complex Adaptive Systems, also found evidence of self-organised criticality in asymmetric warfare data from Iraq and Afghanistan [Dobias, 2009]. Dobias found that for data up to 2007, while the Afghanistan conflict exhibited sub-critical dynamics, Iraq was already in a critical state. While both critical and sub-critical states produce a power law relation between the incidence and magnitude of casualties, fatalities in critical states are correlated in time, while for sub-critical states they are anti-correlated. This insight provides both a way of determining the phase of a conflict, and a guide to decision-makers on what impact a casualty event has on the likelihood of future casualties.

3.3 Summary of Military Applications

Together, the research efforts documented in this section demonstrate widespread and sustained acknowledgement of the inherent complexity of war. This in itself is nothing new, but it has paved the way for complex systems science to provide a foundation for revised doctrine in the U.S. Marine Corps, the U.S. Army, and the Australian Army, both as a description of the operational environment and a guide for action. Recognition of the complexity of the operational environment emphasises limits to predictability due to nonlinear instabilities and the emergence of novelty. Casualty data from conflicts as well as agent based models of conflict show that war is not simply stochastic, it exhibits fractal and scale invariant statistics. This implies that it is not meaningful to discuss, for example, the average size of a terrorist attack, which has significant implications for risk assessment and policy. Limits to predictability provide the rationale for policy that favours robust and adaptive strategies over those reliant on prediction, optimisation and central control.

If war is fundamentally unpredictable, then good strategy should exploit complexity and unpredictability, rather than fight to eliminate it. This was the approach taken in the field during Reconstruction Task Force operations in Afghanistan, by enhancing RTF adaptability and conducting counter-adaptation operations against the Taliban. As important as the operational force is, it constitutes only a small fraction of the total defence effort.[15] Countless decisions are made prior to deployment that impact upon the force's ability to cope with complexity, such as the equipment that is acquired and individual and collective training of soldiers. Australia's Adaptive Army initiative provides an example of applying complex systems insights across the broader institution, even though this too is still limited to the military employees of one Service within defence.

To date, complex systems has provided new insight into how to understand and describe complex operational environments; how to analyse and model conflict; how to organise military forces; how to train for adaptability; how to design new defence systems; and how to design operations and campaigns.

[15]For example, in 2005 the U.S. had approximately 250,000 troops deployed globally out of a total of 3 million active, reserve, guard and civilian employees within the Department of Defense [Global Security, 2005].

4 FUTURE MILITARY APPLICATIONS OF COMPLEX SYSTEMS

Ilachinski's tiers of applicability provided a useful structure for the initial research agenda. Yet after more than a decade of military complex systems research, it is time to reassess the tiers of applicability in the light of the development of both complex systems science and its military applications. The purpose of this section is to reframe the research agenda for the military applications of complex systems.

A useful starting point is to examine where Ilachinski's framework falls short of our aims. Ilachinski's scope was limited to land warfare. Although land forces have largely taken the lead in applying complex systems, applications exist across the whole defence sector, including air, maritime, space and cyberspace environments, as well as the supporting functions of defence, such as raise, train and sustain. Rather than being organised around either a top down analysis of existing military problems, or a bottom-up list of complex systems techniques, the eight tiers ranked applications by their level of risk and reward. Ilachinski wanted to avoid the trap of assessing complexity only against conventional problems and questions, but equally to avoid viewing all military problems as nails from the perspective of a complexity hammer. However, the risk and return profile for different classes of applications necessarily changes as progress is made against research objectives. For instance, the use of agent based models as synthetic combat environments was assessed to be the second highest tier in terms of risk, whereas today they are a widely accepted alternative for combat modelling. Another significant change that dates Ilachinski's tiers of applicability is the rise of network science as a significant complement to complexity theory, and a corresponding decline in the perceived relevance of chaos to military affairs.

The tiers of applicability mix together problems in defence science with problems in defence operations, which are separate concerns. Ilachinski's stated intent was to investigate the relevance of complexity and nonlinearity to OR, but defence science is much broader than OR. In particular, applications to systems engineering and the human sciences were not considered within the tiers of applicability. When these other domains are included, not only are new applications uncovered within the domains, but a new possibility also emerges: the use of complex systems within each domain provides a common framework for communicating between domains, opening up new interdisciplinary research opportunities that depend on interactions across human, operational, technical and systems domains. Finally, the possibility of applications at a meta-level were not considered. That is, defence as a whole can be viewed as a system, so that complex systems analysis can be used to improve the way the different programs of defence coordinate to deliver and employ capability. The same applies to defence science. On a more practical level, much work remains to be done to fully translate the insights of complex systems into military organisations, operations and culture.

In order to reframe and expand the scope of the military complex systems research agenda, we will make a distinction between defence science and defence problems. The latter uses complex systems directly to solve a real world problem, whereas the former applies complex systems to an existing domain of defence science, which can then be applied to a whole class of military problems. Within each category, complex systems applications will be divided into three interrelated themes: pattern formation, multiscale analysis and adaptation. These themes provide a framework to assess promising directions for future applications of complex systems to defence science and defence problems. We then identify grand challenges for complex systems by considering the converse of this relationship: how will the particular demands of defence drive future complex systems research?

4.1 Future Applications to Defence Science

Complex systems can be used to explore, understand, represent, predict, influence, manage, or create systems, as is the case for any other scientific framework. But complex systems can also be used to create the context within which the desired system can self-organise [Bar-Yam, 2003, p. 3], and can shed light on the limits of knowledge [Cilliers, 2007]. Depending on the purpose and constraints, techniques with different depths of analysis may be appropriate. Roughly in order of increasing depth are metaphor, systems thinking, conceptual analysis, agent based modelling, and data-driven modelling. Metaphor can act as a useful inspiration for inquiry, a method for the transfer of ideas between civilian and military disciplines, as well as a memorable and highly compressible summary of careful study. However, analysis that starts and ends with complexity as a metaphor may be a dangerous source of unintended consequences, and would not usually be considered appropriate within defence science. This is not to say that there is no role for metaphor in responding to defence problems — it is essential that military decision-makers are aware of and capable of updating the metaphors that orient their understanding of the operational environment — but rather that analysis must be more than metaphorical to be considered scientific.

Systems thinking can be used to consider a system of interest "as a complex adaptive system", by describing system boundaries, flows, feedbacks, patterns, and attractors, as well as how they evolve. At this depth, there is little distinction between complex systems and other systems approaches, except that perhaps complex systems places more emphasis on decentralised bottom-up sources of order than earlier systems approaches. This can be made more precise through conceptual analysis, which draws on mathematical theorems of complex systems without building an explicit model of the system [Bar-Yam, 2005, Ryan, 2007a, Ch. 4]. Such an approach may be useful when a detailed quantitative model is inappropriate because the problem is soft and fuzzy, or data, time and other resource constraints preclude quantitative model building. To some extent, the same criticisms of complexity as a metaphor apply to systems thinking and conceptual analysis. However, if problems are sufficiently complex that they do not have *a priori* analytic solutions, but require interaction and adaptation in context to resolve, then the purpose of analysis is limited to guiding the initial strategy for intervention. In this case, analysis of the system only needs get the initial strategy within "adaptive range" of a satisfactory resolution to the problem, which requires less depth and detail of analysis than if the exact solution needs to be developed up front.

For more structured problems, agent based models can provide a dynamic representation of a heterogeneous and adaptive system. Data-driven modelling incorporates empirical measurements into a complex systems model to correct differences between the model and the real world. This approach is exemplified by weather forecasting models, although it is rare that our understanding of a socio-technical military system's dynamics is sufficient to provide predictions based on a tight coupling between the model of the system and the real world data stream. Where such problems do exist, they tend to be tactical and local in nature, because the unencapsulated nature of the phenomenon of war means that strategic problems are inevitably contingent on political ambiguities and unbounded interdependencies within the geopolitical context. Consequently, the legitimacy of scientific methods within defence is not governed solely by depth. Applications to defence science that span the range of purpose and depth can be grouped within the three themes of pattern formation, multiscale analysis and adaptation.

Pattern formation

Pattern formation studies the dynamics of structural and behavioural patterns in complex systems. Self-organisation — the spontaneous increase in organisation of open systems far from equilibrium — is a key concept within pattern formation. Self-organising systems are open to flows of energy, matter and information from their environment, but the flow of information must be insufficient to account for the increase in organisation in relation to the internal state of the system. The dynamics of self-organising systems are driven by a combination of positive and negative feedback that transforms energy into information.

Pattern formation can also be analysed from a topological perspective. Topological approaches are currently very popular because they reveal certain invariant features of the whole system. The majority of topological approaches use graph theoretic models and their associated metrics, which are also commonly referred to as network science. Any complex system can be represented as a network, and most network metrics are computable and straightforward to apply to empirical data. However, current network theory is mostly limited to static network properties. Future advances in the dynamics of networks (growth and evolution) and dynamics on networks (flows and synchronisation), as well as how topology and flow coevolve, will drive future applications to defence science.

Pattern formation within phase space draws on techniques from nonlinear dynamics to identify attractors, bifurcations, phase changes and tipping points. When they apply, these are powerful tools for understanding the qualitative range of patterns accessible to a system, as well as the stability of and transition dynamics between attractors. The main challenges for applying nonlinear dynamical techniques to real world military situations are the lack of data, the extremely high dimensionality of phase space, and the mixture of continuous and discrete variables including social and human factors.

An alternative approach to pattern formation is to study the causal and influence networks that generate the system's dynamics. Understanding these networks may unlock an enormous potential to change their topology and flow and thereby indirectly control both intended and unintended effects. A series of workshops hosted by the TTCP Action Group on complex adaptive systems for defence aims to advance beyond techniques from Bayesian analysis, system dynamics, network theory, and nonlinear dynamics for modelling systemic causation [Abbott, 2009].

Cellular automata and agent based models provide useful environments for exploring global patterns produced by local rule sets. New tools that combine agents and networks could combine the analytical strengths of graph theory with the ability of agent based models to represent computation, decision-making and adaptation.

Within defence, the utility of fundamental research into pattern formation is obvious. The object of any war is a stable and enforceable peace. Countries go to war when their policy objectives cannot be achieved without military action. Conflict produces temporary instability while protagonists compete to transform the system towards incompatible attractors. Without understanding the dynamics of pattern formation within the operational environment, the risks of failing to achieve the military objective and producing unintended consequences increase dramatically. Pattern formation has further applications to understanding the potential, strengths and weaknesses of one's own organisation and of the adversary. From attempting cultural change to disrupting the adversary's system, understanding pattern formation provides an opportunity to use nonlinearities within the system as levers for transformation. In a complex conflict environment with many different kinds of actors, the need to understand pattern formation is even more pronounced.

Research in causal and influence networks may also lead to a novel approach to defence experimentation. The international standard for science-based defence experimentation is articulated within the GUIDEx manual produced by TTCP's Joint Systems Analysis Group. While the GUIDEx is valuable for many types of defence experiments, it has only limited applicability to complex systems experimentation. This is because its experimental method relies on isolating linear causal chains. According to GUIDEx, "Any national or coalition capability problem may be stated as: **Does A cause B?**" [TTCP, 2006, p. 8]. One of four GUIDEx criteria for a valid experiment is the ability to isolate the reason for change in the effect **B**. For a complex open system, knowledge of isolated, individual hypotheses provides little insight into the dynamics of the broader causal and influence network. If these networks are to be understood, traditional approaches to experimentation need to be augmented with methods that do not depend on closing the system and controlling for variation.

Complex systems experiments require tighter coupling between theory and experiment. The experimental framework needs to augment its metrics hierarchy, responsible for measuring the effects of change at multiple levels within the system, with a systems model of the causal and influence network. Within a campaign of experimentation, the systems model coevolves with the experimental design, so that each experiment refines understanding of the causal and influence network, while the causal and influence network enables synthesis and interpretation of the experimental results. Causality in social systems is not just networked, but the network itself is dynamic. As individuals and organisations anticipate, learn, and adapt, their behaviour alters the causal and influence network in novel and unpredictable ways. To cope with this, complex systems experimentation must be interactive and adaptive. Instead of large, discrete, infrequent experiments, more continual interaction in context would allow complex systems experiments to track changes in the causal network. An adaptive approach to coevolving the experimental design with the systems model would emulate the "virtuous bootstrap" between learning in action and learning how to learn, described by Hooker [Hooker, 2009] as a general approach to resolving open, ill-defined problems.

Many of the most significant capability investment decisions depend more on the ripple effects of new acquisitions on the network of legacy systems still in service, than on measuring the outcome of a single cause-effect chain. In operations too, Clausewitz [1984, p. 158] advised that "in war, as in life generally, all parts of a whole are interconnected and thus the effects produced, however small their cause, must influence all subsequent military operations and modify their final outcome to some degree, however slight". What Clausewitz did not mention is that in such interconnected systems, the accumulation of indirect effects may actually exceed the effects of direct action. Even if it is not possible to ever fully understand the relevant dynamic causal and influence networks, a tentative model of the whole system would be more valuable to decision-makers than more certain but disconnected knowledge of detailed cause-effect chains.

Multiscale analysis

The theme of multiscale analysis is motivated by the inability to separate the different scales within a complex system. Due to cross-scale interactions, a model of a complex system at any single scale will be incomplete. As a consequence, analysing a system across multiple scales can produce novel insights. Emergence is a key concept within complex systems science because an emergent property relates different scales within a system. Specifically, a novel emergent property is a difference between relative microscale and macroscale descriptions of a system: the emergent property is present in a macrostate even though it is not present in a microstate [Ryan, 2007b]. The description of a complex system cannot be reduced to just a description of the fine scale details, because this ignores the emergent properties at larger scales.

Richardson, Lauren, Clauset and Johnson have shown that conflict is fundamentally a multiscale phenomena. However, the use of multiscale analysis within defence science is not widespread. The majority of modelling is still focused at a single scale of analysis. The danger of interventions based on models at a single scale is they may produce unintended consequences at other scales that ultimately undermine improvements at the focal scale. For example, the fusion of intelligence from low-level tracking data up to high-level inferences and situation assessments is only just beginning to be explored from a multiscale perspective [Lingard, 2009]. However, this analytical approach has enormous potential value for defence science.

One approach to analysing multiscale systems with emergent properties is the complexity profile, which analyses the amount of information required to describe a system at every scale [Bar-Yam, 2004]. The complexity profile of an organisation can reveal how well it matches the complexity of its environment, and identify whether either increasing fine scale variety or enhancing large scale coordination is likely to improve the organisation's fitness.

Scale invariance, fractal statistics, the fractal dimension and measures of self-similarity also provide insight into the relationship between scales within a system. For example, these techniques may reveal limits to the utility of averages, the dependence of a measure on the scale of measurement, and the mutual information between scales of a system.

Renormalisation groups [Wilson, 1979] provide useful insights for multiscale analysis. Renormalisation operates by abstracting interdependencies between fine scale variables into interdependencies between coarse scale variables, and can be applied recursively. This allows an analyst to identify which details and relationships in the fine scale representation of a system have large scale implications, and which details disappear at coarser scales. Renormalisation may not apply directly to many defence systems, since it only applies when the system is close to a critical point, and in addition it abstracts away potentially important local details, losing context. However, used in conjunction with other tools, renormalisation can contribute to understanding multiscale systems.

Hierarchy theory is an older approach to multiscale analysis that arose within GST. A hierarchy is a partially ordered set that may be nested, where the members of one level include those at the level below, or non-nested. Simon [Simon, 1962] argued that because nested hierarchies are composed of loosely coupled building blocks, or stable intermediate forms, they could can be assembled orders of magnitude faster than a non-hierarchical assembly process. Hierarchy theory does assume that higher and lower levels are dynamically screened off from each other due to order of magnitude differences between levels, and that information exchange must therefore be interpreted at the boundaries between levels [Salthe, 2002]. Living systems exhibit massive cross-scale communication and energetic feedback, and increasingly so do engineered systems, such as adaptive robotics and adaptive organisations. Hierarchy theory only applies to those multiscale systems where the time scales of different levels can be separated, and cross-scale interactions are insignificant.

Adaptation

Adaptation is a theme that centres on the ability of systems to improve the fit with their environment over time. Adaptation works by combining variation and interactive selection in a feedback loop with retention (VSR), which biases changes that increase fitness. Evolution and learning are both special cases of adaptation. In addition to natural evolution, the immune system of vertebrates, prions, Artificial Life programs, neuronal group selection, memes and artifacts also exhibit evolutionary dynamics. Dawkins coined the term Universal Darwinism for this more general class of evolutionary processes [Dawkins, 1983].

Learning includes both individual and organisational learning. A simple model developed in machine learning, called reinforcement learning, implements VSR for a single agent with a fixed set of available actions, where the environmental feedback is a single real valued reward plus the observed state at each time step, and favourable moves are retained by updating an internal model, in order to estimate the future value of each state, which influences the likelihood that the agent will choose that state in future [Sutton and Barto, 1998]. However, reinforcement learning does not allow anticipation, and represents a primitive form of learning. A more sophisticated model for learning developed by Christensen and Hooker [1998], called self-directed anticipative learning (SDAL), describes learning as an increasingly developed regulation and internalisation of the VSR process. The SDAL cycle requires the capacity to anticipate the interaction process between the learning system and its environment, evaluate the interaction against normative signals, and modify the interaction as a result of anticipation and evaluation [Hooker, 2009].

An area of intersection between the themes of multiscale analysis and adaptation is the growing field of multilevel selection. Group selection, rejected by most evolutionary biologists and sociobiologists in the 1960s, is now supported by significant empirical evidence [Wilson and Sober, 1994]. One insight from this research is that weak global influences can dominate direct strong local selection pressures over the longer term, which reinforces the need to understand causal and influence networks and indirect approaches to system transformation. Niche selection, the evolution of evolvability, symbiosis, and coevolution are some of the more promising concepts within evolutionary biology to offer insights for defence systems. Understanding learning and evolution within the broader framework of adaptation provides a basis for designing new adaptive mechanisms, to ensure that the system is adaptive over the most relevant temporal and spatial scales.

There are many degrees of freedom for designing adaptive mechanisms, many of which are sensitive to the context of application. In particular, an adaptive system must be able to discriminate between those environmental features that trigger adaptive change and those that do not, which is context-dependent, and must be reflected in the adaptive system's internal informational state. A deep understanding of both adaptation and the context is required for the successful design of adaptive systems.

Defence applications for adaptation are ubiquitous. Whenever problems are sufficiently complex that *a priori* analytic solutions are not possible, some form of adaptation is required. Within defence science, the adaptability of individuals, groups and enterprises can be improved. Engineered systems can move beyond mimicking biological traits to incorporating genuinely adaptive mechanisms that allow for novel patterns of behaviour to emerge in response to unforseen circumstances.

Experiments in the human factors of adaptability are in their infancy. Measuring the utility of training, educating, mentoring and coaching adaptability could help to improve individual and team adaptivity. Second-order adaptation — applying adaptation to the adaptive mechanism itself — can accelerate the rate of adaptation. Counter-adaptation can deliberately steer and limit the adaptation of adversary systems. For example, deliberately targeting an adversary's sources of variation, using deception to provide misleading feedback, and disrupting feedback loops that allow the adversary to learn from success and failure all constitute counter-adaptation operations. Even when these tactics have been used in the past, they have rarely been explicitly designed and combined as a system. Multilevel selection can provide insight into evolving capabilities that meet strategic, operational and tactical ends. Conflict models and wargames can simulate an adaptive adversary and explore coevolutionary dynamics in order to identify intransitivities and avoid costly arms races.

Synthesis

The most powerful applications of complex systems to defence science will draw on all three of these themes. Ultimately, complex systems should aim to provide a framework for the design and management of organisations capable of solving complex multiscale global problems. For organisations that are too complex to design directly, complex systems must provide guidance for the design of environments within which these organisations can evolve, to foster the emergence of desired properties as well as the suppression of undesired properties. Because context is so important, this should integrate experimentation and evolution with the real-world operational environment. Rather than a clear distinction between experiment and intervention, all actions should serve a dual purpose of partly acting to learn and partly acting to improve. This ensures that the organisation is capable of both exploration and exploitation, and can continually adapt. An understanding of complex systems leads to a significantly different approach to interacting with, learning about, influencing, managing and designing defence systems. It also provides an awareness of the limits to predictability and control, which should reorient the relationship between defence science and military decision-makers. Defence science needs to be utilised to exploit uncertainty and as a source of potentially disruptive novelty, not as a way of eliminating uncertainty and risk. This is potentially the most profound implication of complex systems for defence science.

As the domain of applicability of complex systems is not limited by either the nature of system or the kind of intervention, it overlaps with the established branches of defence science. This presents both a challenge and an opportunity. Challenges arise when complex systems methods compete with classical approaches. For example, a culture of optimisation, quantitative prediction, control and efficiency may be threatened by methods that instead seek adaptation, qualitative pattern recognition, influence and effectiveness. Opportunities are presented by the creative tensions that arise from looking at the same problem situation from multiple perspectives. Because complex systems is inherently interdisciplinary, it is well structured for communicating between disciplines. This presents the opportunity for complex systems research to build bridges between the branches of defence science, enhancing communication and promoting systemic solutions to messy real world problems. Intellectually, complex systems is highly complementary with existing branches, since each branch includes many domain-specific results that can furnish the more abstract models of complex systems. In the short term, low risk (but also low payoff) applications of complex systems fall within existing paradigms, and indeed this is the pattern of early adoption portrayed in the historical narrative above, as complex systems infiltrated OR, systems engineering, and finally human factors research. Complex systems will be increasingly relevant for systems that process large amounts of information and where multiple fields intersect, such as the Nano-Bio-Info-Cogno convergence [Bar-Yam, 2003]. Complex systems will increasingly be recognised as a distinct branch of defence science, which asks new questions that are not studied in other branches. This includes questions about multiscale correlations, coevolutionary dynamics and self-organisation. However, it is essential that as complex systems differentiates itself from other branches, it maintains porous boundaries, to continue to foster interdisciplinary communication and collaboration.

Complex systems has meta-level applications to defence science organisations. The relationship between different branches of defence science, such as OR, systems engineering, human factors, and the information sciences, can be hostile, because each branch competes for resources and influence. While competition is an important driver of innovation, cooperation is equally important to tackling multidisciplinary real-world problems. Complex systems can provide insights into the appropriate balance of competition and cooperation between branches of defence science.

4.2 Future Applications to Defence Problems

Complex systems applies to social problems at all levels within defence, from individuals, teams, units, and enterprises to societies. Complex systems equally applies to technical systems, from subcomponents, components, and systems to systems of systems. More importantly, it applies to socio-technical systems where social interactions are mediated by hardware and software, and can examine the relationships between levels and with the environment. Within this extremely broad scope, several illustrative future applications to defence problems will be outlined along the same three themes of pattern formation, multiscale analysis and adaptation. The following comments will be focused on Western militaries. It is possible to comment on the West in general, because the commitment of Western nations to fighting in multinational coalitions and to achieving interoperability — especially with the militarily dominant U.S. — has resulted in large commonalities in doctrine, equipment and training.

Pattern formation

Culturally, the military is predisposed towards focusing on the commander, on official authority, and formal organisational structures. Research in self-organisation and pattern formation shows that informal mechanisms can be even more important in the dynamics of complex systems. Understanding the informal flow of power and influence within a complex system requires a different approach to collecting, interpreting, organising and disseminating intelligence. During operations, many patterns are observed but not reported, reported but not subject to information management, collated in databases but never analysed, not studied over the right time scale, or not shared with the right people at the right time. Typically, information management procedures are unit-specific, which means every rotation starts data collection afresh. Therefore, patterns longer than one rotation cannot be detected. The mantra of U.S. officers is that they have not been in Korea for over fifty years, they have been there for a year over fifty times. Better information management is an essential prerequisite for an improved understanding of pattern formation in conflict. In addition, if the self-organising mechanisms within an operational environment are recognised, the strategy for intervention will be substantially different. Those problems that cannot be solved with a direct frontal assault require a consideration of the network of pathways for influencing the system. This applies equally to cultural change within DoD as to an operational context. It applies especially within an inter-agency context, where military force alone cannot solve the problem, and the military commander does not have other agencies under his or her chain of command.

Network theory is vital for understanding and countering terrorist networks. However, the current models from network science do not directly apply to clandestine networks, because the form, function and logic of clandestine networks means they are not scale-free or decentralised, and moreover they develop distinct motifs within the overt, covert, and supporting components of the network [Jones, 2009]. More realistic models of clandestine networks can lead to more effective targeting strategies that disrupt entire networks, rather than removing highly connected but easily replaceable nodes. Improvised Explosive Devices (IEDs) are the number one cause of coalition casualties in Iraq. Initial efforts to counter the IED threat focused excessively on defeating the IED device itself. However, efforts to counter Radio Frequency (RF) activated devices at great expense were largely nullified when insurgents switched to pressure plate and other non-RF IEDs. Counter-IED experts soon realised that rather than accelerating an asymmetric arms race by targeting the symptom and steering the insurgents towards devices that have no effective counter, a more effective strategy is to proactively target the network that supports the emplacement of IEDs. However, this strategy is also predicated on an understanding of clandestine networks.

Network theory also has applications to planning. Current planning processes produce monolithic plans, which are time-consuming to develop, deconflict and synchronise. Uncertainty is mitigated by developing detailed branches and sequels. Subordinate units develop further detail only once their higher headquarters produce their plan. However, a plan can also be viewed as a directed network, where the vertices are planning elements and the edges are relationships between elements. By developing plan fragments rather than monolithic plans, many more possible plans can be produced by rearranging the relationships between elements. By allowing the connections to self-organise based on local information, planning networks that span multiple echelons can retain coherence even as new fragments are added to the network and the operational environment changes. There are obvious risks and challenges for such a radical approach to planning, but the rewards could include increased responsiveness and the ability to maintain a faster operational tempo.

Multiscale analysis

Emergence will be increasingly important to the design of both military systems and operations. The competitive dynamic of warfare drives the demand for systems and operations of greater complexity, which continually pushes the limits of designers to predict and control the consequences of design decisions. An understanding of emergence in complex systems may open up new methodologies for designing beyond the thresholds of predictability and control. A simple example is the emergence of robust systems constructed from sloppy parts. In this context, robustness is defined as "the maintenance of some desired system characteristics despite fluctuations in the behavior of its component parts or its environment [Carlson and Doyle, 2002]. Carlson and Doyle [2002] argue that "advanced technologies and organisms, at their best, use complicated architectures with sloppy parts to create systems so robust as to create the illusion of very simple, reliable, and consistent behavior apparently unperturbed by the environment." Within biology, the design principle of tensegrity has been used to explain how robust mechanical cytoskeletal networks emerge from sloppy parts (flexible molecular filaments) [Ingber, 2003a; Ingber, 2003b]. Circadian rhythms with a period close to 24 hours, observed in nearly all living organisms, are hypothesised to emerge when neurons with cell-autonomous oscillations sychronise via regulatory networks [Welsh *et al.*, 1995]. Circadian rhythms are uniform and stable, despite the fact that individual neurons have periods ranging from 20 to 28 hours [Bernard *et al.*, 2007].

There are many military applications for robust systems composed of sloppy parts. Networked swarms of cheap, disposable and individually unreliable unmanned aerial vehicles could provide robust and reliable surveillance of the battlespace. The cost of many procurement projects could be dramatically reduced if the reliability of the system did not depend as heavily on the reliability of its parts, many of which are orders of magnitude more expensive because they are required to meet military specifications. The operational level of war consists of designing campaigns to coordinate large numbers of engagements, each with highly uncertain outcomes, to reliably achieve strategic objectives. Intentionally designing systems to have emergent properties will open up many currently untapped opportunities.

An understanding of emergence may also help the military to avoid unintended emergent properties. Just as the widespread use of antibiotics may lead to increasing antimicrobial resistance, well-intentioned anti-terrorism operations could conceivably place a selection pressure on terrorist organisations. This could lead to the emergence of more virulent strains of terrorism, unless strategies are developed to limit such an effect. An ability to detect and measure emergent properties would enable the evolution of a network of relationships between micro actions and macro effects that can exploit desired emergent properties and inhibit unintended consequences.

Bar-Yam notes that "the military is typically far beyond other organizations in recognizing the implications of complex systems knowledge from a practical perspective" [Bar-Yam, 2005, p. 96]. This recognition is embodied in force structures that match fine scale Special Forces with complex terrain, and large scale aircraft carriers with simpler large scale ocean environments. Bar-Yam also points to the dramatic differences between operations in Afghanistan and the 1991 Gulf War as evidence of awareness of the difference between complex and large scale operations. However, not all aspects of military organisation reflect this understanding equally. Vertical integration of planning in operational units is a continual challenge, and the current approach to synchronisation lacks a theoretical foundation. Even the very objective of synchronisation is questionable, because it encourages mechanistic coordination of action according to the lowest common denominator. Instead of synchronisation, a complex systems approach would seek coherence, which may be achieved through central coordination, but could equally be realised with bottom-up self-organisation. A robust force would be capable of both centralised and decentralised planning and execution, so that it is able to adjust its organisational structure according to the demands of its environment.

Scale invariance in quarrels is theoretically well established, yet it has had surprisingly little practical impact. There are significant implications ranging across and connecting global security, international relations, national policy, military strategy, military tactics, homeland defence, emergency response, and law enforcement. The distribution that connects intentional acts of violence over all magnitudes repeats itself within the spectrum of terrorist violence, and repeats itself again within individual conflicts. The self-organised critical explanations of this power law should at least raise the question of what cross-scale effects interventions at one scale of violence have on the frequency of violence at other scales. It is possible that measures to decrease the incidence of frequent, low magnitude violence increases the probability of extreme violent events. Fat tailed distributions dramatically alter assessments of the risk of large wars, large terrorist attacks, and large troop casualties from singular events, because extreme events that would be vanishingly improbable within a Gaussian distribution will still occur with finite probability. The use of averages can be highly misleading, and common modelling assumptions are invalidated by fat tailed distributions. Incorporating the implications of scale invariance into risk assessments and decision support models will lead to more informed decision-making within the defence and security arena.

Because of the difficulties of obtaining empirical evidence in war, and providing realistic training in peace, the military places a high value on the study of military history. However, most military history is not very complex. In their classic work *Military Misfortunes*, Cohen and Gooch [1990, p. 29-43] dissect the reasons why complex events like the American failure at Pearl Harbor are often misunderstood in military history. Those reasons include the proclivity for utilitarian military history, which focuses undue attention on the dueling commanders, and falls prey to the fallacy of homogeneity;[16] the interests of Governments in producing official military histories for particular purposes; and the tendency to write military history in horizontal layers, as battle history, campaign history, or the history of strategy making. Following Clausewitz's method for *kritik* [Clausewitz, 1984, p. 156], Cohen and Gooch propose a multilevel approach for analysing military misfortune. They construct a matrix where the horizontal axis specifies the echelons of command and the vertical axis identifies the key events relevant to the failure in question. This enables the analyst to trace "pathways to misfortune" through the military organisation, to uncover systemic sources of failure. Cohen and Gooch's methodology presents a systems approach to history, revealing organisational roots of failures that have traditionally been blamed on incompetent commanders, military culture, institutional failure, or even national culture [Cohen and Gooch, 1990, p. 5-16]. Further, they are able to discriminate between the failure to anticipate, learn and adapt, and conclude that the failure to adapt is by far the most serious, because with adaptation, a force can overcome unanticipated surprises and a failure to learn lessons from the past.

Cohen and Gooch demonstrate the power of a multilevel systems approach to historical analysis. However, they still fall short of fully embracing a complex systems approach to history. In all six of their case studies, the primary pathways to misfortune flow from the top down, yet an understanding of complex systems would suggest that bottom-up flows are equally important to understand. Further, the flows always pass between adjacent levels, which ignores cross-level effects. Each of the matrices is one-sided, and does not represent the interaction between protagonists in the conflict. The case studies reveal organisational deficiencies resulting from a lack of coordination, but do not show how organisations coordinated for large scale effects can fail to cope with fine scale complexity. Allowing for these modifications, their approach provides a framework for a complex systems analysis of military history.

Adaptation

The operational force is often confronted by the need to "adapt or die". The selective pressure is strongest at the small unit level over short time scales. Opportunities for improved adaptation include better learning across rotations, across

[16] Cohen and Gooch [1990, p. 37] define the fallacy of homogeneity as "the habit of speaking of a large organization as a unitary whole rather than as a collection of suborganizations with definable subcultures, routines, and modes of operation."

organisational stovepipes,[17] and balancing the tradeoff between being adapted for the current fight and adaptable to future conflicts. Also, there are many ways to improve the adaptability of the defence bureaucracy and in particular the way new equipment is acquired. Because the acquisition of major defence materiel can take more than a decade, it is essential to have a flexible capability development process in the face of a changing strategic context and evolving technology base. A theoretical understanding of adaptation can lead to many more opportunities to improve adaptivity at the scales that are most important for military success.

The value of capturing and disseminating lessons learned is already appreciated within the military culture. Dedicated centres for lessons learned and the use of after-action reviews to continually improve processes contribute to organisational learning. However, the enormous responsibility that accompanies the capacity for lethal force can promote a culture that is intolerant of failure. If failure is not tolerated, mistakes and near misses will not be reported, which means further opportunities for organisational learning are missed. A system for blame-free error reporting with the power to implement changes as organisational flaws are identified could dramatically improve organisational learning within the military.[18]

There are many well-intentioned mechanisms within the military that promote standardisation. The intrinsic merit of standardisation is so deeply embedded into military culture that a synonym for military personnel is 'uniform'. Dixon's [1976] classic work on the psychology of military incompetence identifies uniformity, the love of regularity, and regimentation as organisational sources of incompetence in the military. Standardisation improves predictability, simplifies control, and can allow forces to produce large scale effects. However, this often comes at the expense of individual effectiveness, and necessarily decreases variety, which limits the potential of the force to adapt. The emphasis on standardised and approved doctrine over rigorous peer-reviewed discourse, and training what to think over educating how to think, exemplify excessive standardisation. An adaptive approach requires a re-examination of the role of doctrine, and a shift in emphasis from training towards education.

Synthesis

Complex systems can also help improve the way the military resolves cross departmental issues. The typical military response to a new crisis is to assign individual responsibility for solving or managing it. However, as Bar-Yam notes, for complex systems, we need to learn how to improve the system without putting someone in

[17] Organisational stovepipes are structures that restrict the horizontal flow of information and cooperation across an organisation. They tend to occur between departments and geographically separated sites in large, hierarchical organisations. Organisational stovepipes are reinforced by performance measures keyed to individual achievement against departmental objectives, because cooperation that betters the interests of the overall organisation has an opportunity cost in terms of individual achievement [Tobin, 2006, p. 82]. Clearly, stovepipes inhibit feedback, learning and adaptation.

[18] This opportunity for organisational learning in the military was first identified by Grisogono.

charge [Bar-Yam, 2005, p. 14]. If the system is too complex for one individual to understand, authority, responsibility and good intentions alone will not prevent failure [Dörner, 1996]. In principle, it would seem that there exist ways of organising to achieve collective intelligence exceeding individual capacities, since the human brain is organised to perform tasks greatly exceeding the complexity of individual neurons. Research on the wisdom of crowds has identified a relatively narrow range of conditions where large numbers of independently deciding individuals can outperform individual experts [Surowiecki, 2005]. Additionally, effective strategies for the governance of commons point to the need for mutual coercion, mutually agreed upon [Hardin, 1968], and complex, redundant and layered institutions in order to facilitate experimentation, learning and change [Dietz *et al.*, 2003]. However, little is known about the principles underlying collective intelligence. Achieving collective intelligence in human organisations in faster than evolutionary time scales will require advances in understanding the mathematics of collectives (see for example [Tumer and Wolpert, 2004]), as well as advances in social psychology.

There are also many problems within defence that persist because no single organisation or authority has responsibility. An example is the whole of life integration across capabilities that are funded, designed, deployed, maintained, and replaced by program managers that are only focused on achieving cost, schedule and content targets for their individual projects. Complex systems can provide insights into how to resolve systemic issues — not only by creating an integration authority, but by analysing the dynamics of pattern formation across multiple scales, and on this basis designing environments where the programs of defence have incentives to improve the fitness and adaptability of defence as a whole.

4.3 Future Challenges for Complex Systems

Just as the development of the science and technology of flight, radar, rocketry, computing, and atomic energy were driven by military problems, new developments in complex systems may be stimulated by military drivers. This section identifies nine grand challenges for complex systems motivated by military needs. Each open question includes a description of the potential contribution of complex systems towards addressing it.

1. What are the fundamental limits to the predictability of war, and what opportunities do they present?

   Earlier, we noted that the first part of this question was raised by Beyerchen. However, to be useful, establishing limits to predictability must provide a tangible benefit. Crutchfield [2002] notes that far from being a setback, the discovery of fundamental limits in mathematics, deterministic chaos, psychology and philosophy during the 20^{th} century fostered an unprecedented increase in our knowledge of nature: "Somewhat ironically, the realization of each limit, rather than being only a disappointment, often showed nature to be much richer than before, when seen through the dark sunglasses of simplifying assumptions." One way of viewing complex systems science is from the perspective of the traditional simplifying assumptions it challenges, such as the validity of superpositionality, closed system boundaries, smoothness, averaging, spatial mixing, separation of scales, centralised control, and linear causality [DeRosa *et al.*, 2008]. If complex systems science could help to determine fundamental limits to the predictability of war, decision-makers might reinvest effort away from anticipating the unknowable towards increasing resilience and adaptive options. Understanding the sources of fundamental unpredictability could be a source of shock and surprise by leveraging uncertainty to disorient the adversary. Understanding how predictability decays over time and space could shape operational tempo, information dissemination policies, and inform the degree of confidence assigned to intelligence assessments. By challenging the simplifying assumptions of current approaches to anticipating the future operational environment, complex systems could stimulate the growth of a richer understanding of the phenomena of war.

2. How should a complex adaptive system dynamically balance exploration and exploitation?

   According to Clausewitz [1984], more than in any other human activity, war is the province of chance. In the face of fundamental uncertainty, the military commander must allocate resources between exploiting the current understanding and exploring alternative explanations. This balance should be sensitive to the degree of uncertainty and the type of variables containing uncertainty. The greater the uncertainty, the more resources should be allocated to exploration, since the effectiveness of strategy that exploits the current understanding will quickly erode. It is important to estimate the uncertainty for different types of variables, because some may be much more stable and certain than others. In biological evolution, organisms have evolved the ability to modulate the rate, location and extent of genetic variation [Caporale, 2002]. This allows evolution to occur much faster than if all variation were purely random. By analogy, if military organisations could adaptively set differing exploration rates for different types of variables, this could increase their rate of adaptation. Core and safety-critical functions would have very low rates of variation, while functions dependent upon rapidly evolving technologies would encourage wide exploration.

   Achieving the right balance between exploration and exploitation is not a trivial task. In a complex system, every action is also a learning opportunity. Actions to intervene stimulate a response, which contains information about the causal and influence network within the system. Through a series of interactions, information from the system's response can be organised, compressed and represented as a model, which informs future interventions, leading to further refinements of the model. The model may be an explicit mathematical model but it could equally be an implicit mental model. Small scale probing actions may be purely intended to gather information — to explore — while large scale decisive action may primarily exploit understanding from within the current model to transform the system. However, most actions will contain a mix of exploration and exploitation. This presents two challenges: how are the lessons from all interactions with the system captured, and how is the model of the system coevolved with the strategy for intervention?

The variance of battlefield metrics is an important determinant for the rate at which lessons captured can be incorporated into the model. If the metric has a low signal to noise ratio and a correspondingly high variance, then changing too quickly will only react to random noise, rather than the true signal. This can be exacerbated when the system's dynamics includes long time lags. A real world military example is the surge strategy in Iraq, initiated and led by General Petraeus in 2007. Initially, the surge increased U.S. casualties as troops moved out of their secure "super Forward Operating Bases" in order to protect the people. It took time for the new strategy to gain the trust of the local population and reap the rewards, which included better intelligence and local militias switching to the American side [Ricks, 2009]. There are risks from adapting too slowly, but there are also risks from changing strategy too quickly when metrics contain large variance and time lags. The rate at which learning from action affects models that inform future interventions must be sensitive to the variance and lags in metrics that measure operational success. This presents a challenge for complex systems to provide a theoretical framework and a methodology for dynamically balancing exploration and exploitation.

3. How do complex systems balance robustness, resilience and adaptability to maintain and improve their organisation?

   The purpose of collective training prior to operations, and command and control during military operations, is to maintain coherent organisation in the face of disruptions caused by the adversary and the environment, while simultaneously acting to disrupt the adversary. Three basic properties that help to maintain organisation are robustness, resilience and adaptability. A robust system is able to maintain organisation because it is insensitive to perturbation. Castles and tanks are good examples of military systems designed to be robust to adversary attacks up to some energy injection threshold. A resilient system is capable of self-repair when it is damaged by external shocks. Resilience is achieved by the availability of excess capacity in combination with a sense and respond network capable of deploying the capacity when required. This is not difficult to design, but often the pressures of short-term efficiency and performance demands reduce resilience by eliminating perceived redundancies, overheads and infrastructure that provide real options to respond to shocks. Commanders have long recognised the value of holding combat forces in reserve for increasing resilience to the inevitable shocks of battle. Terrorist networks too are highly resilient to the removal of individual nodes because they have effective systems for re-establishing connections and roles within the organisation. Adaptable systems are able to proactively improve the fit with their environment over time. Adaptation will cause certain changes in organisation, but this may be necessary for the long term survival of the system.

The challenge for complex systems is understanding how mechanisms for robustness, resilience and adaptability work together to ensure the survival of the system over both short and long time scales. There clearly exist tradeoffs between these desirable properties. For example, robustness tends to reduce adaptive options and is insensitive to environmental perturbation, whereas adaptability requires flexibility and sensitivity to environmental signals. New theory is needed to understand the trade space between robustness, resilience and adaptability. Methodological guidance is also required to inform the design of systems capable of maintaining and improving their organisation during operations. Robustness, resilience and adaptability all come at a cost, but the cost benefit tradeoff of these traits at the whole system level is never explicitly assessed. Consideration of this tradeoff is essential to improving overall system effectiveness in the face of fundamental uncertainty and an interactive adversary.

4. How can the duality between systems and operations be exploited?

There is a fundamental duality between systems and operations: systems must be developed in the context of current and future operational requirements, while operations are both constrained and enabled by the systems that have been developed. In spite of this duality, the way defence is currently organised encourages separation between operational and systems concerns. This artificial division has significant consequences, resulting in both ineffective and inefficient capability decisions from the perspective of the whole system. While it seems that a tighter coupling between systems and operations would improve the design of both, exactly how this should occur relies on a deeper understanding of the relationship between them. If done poorly, a closer link between systems and operations could constrain rather than enable new possibilities, and lead to disruptive conflict between these two perspectives.

If the systems/operations duality was better understood, defence could be organised to reflect and exploit it. The applications of complex systems science to complex systems engineering and operational design reviewed above reveals that the challenges of design are deeply analogous within each domain. Complex systems could further help to build a bridge between systems and operations, to enable a more systemic approach to coevolving system and operational designs.

5. Is decentralised adaptation possible?

   The current understanding of adaptation is useful for identifying many ways to improve the adaptability of a system when there is central oversight. This is a crucial and foundational assumption, because there is no guarantee that adaptation based on local selection will not be maladaptive for the system as a whole. Therefore, a central, global map of the system is currently needed to understand and improve adaptive mechanisms at the spatial and temporal scales that are most important to the system. Whether or not adaptation can be implemented without centralised variation, selection and retention is currently an open question. Variation arises naturally in a decentralised system, however selection and retention are much more difficult to implement across a network. Advances in network theory and distributed control theory are needed before organisations can identify and amplify useful novelty without centralised oversight.

6. How can the interaction between multilevel organisations and multiscale problems be represented?

   The current generation of Agent Based Models (ABMs) are designed to explore phenomena at a single scale. Most analytical tools then measure change over a single time scale. Military organisations consist of many levels and concurrently address many interdependent problems over multiple temporal and spatial scales. New tools are needed that are capable of representing multilevel goal-directed organisations operating in a complex multiscale environment. A monolithic model may not be appropriate for this purpose. A multiscale framework capable of managing context-specific interactions between models of different aspects of a complex system would provide a flexible, reusable approach to multiscale modelling.[19] Such a framework could integrate legacy component models at different scales, and provide logging, plotting, and analytical tools for multiscale analysis. ABMs could be plugged into this framework to be synthesised with alternative modelling approaches, to better realise the Project Albert objective of Operational Synthesis. This could still exploit advances in multiscale modelling within ABM tool kits. ABMs capable of simultaneously representing military operations at the strategic, operational and tactical levels could help to address a fundamental operational question: how should tactical actions be orchestrated to achieve the desired strategic effects?

   Perhaps the biggest challenge for both multiscale modelling and agent based modelling is validation and verification. A network of models activated within a context-dependent simulation framework becomes difficult to verify, even if all of the component models have been verified. Validation has always been an issue for agent based models, because ABMs tend to be massively

[19] The nearest precedents to this approach are the dynamic information architecture system (DIAS) [Campbell and Hummel, 1998] and Multiresolution, Multiperspective Modeling (MRMPM) [Bigelow and Davis, 2003].

overparameterised and consequently at risk of overfitting. These issues are amplified within a multiscale framework. Part of the answer is that complex systems models require a different approach to validation. Because ABMs provide a natural description of a system in terms of the behaviour of its constituent entities, they are able to be validated and calibrated through expert judgment applied to individual behaviours, rather than by fitting aggregate system-level transition rates [Bonabeau, 2002]. The complex systems community must address the validation and verification challenges associated with the next generation of modelling techniques before they can be relied upon to support decision-makers.

7. How can data-intensive techniques be employed for near real-time decision support?

   The central theme of Boyd's work was developing a time-based theory of war [Coram, 2002, p. 328]. The purpose was not to promote speed at any cost, but to acknowledge the pervasive importance of the temporal dimension in military decision-making. Because complex systems models explore fitness landscapes rather than point scenarios, and because they model individual agents rather than aggregate transition rates, they are extremely data-intensive. The cost of data collection in lethal combat environments can be very high in terms of blood, treasure and time. Unreliable sources, deception, information operations, and classification restrictions together provide a formidable challenge for treating the data before it can be used for modelling, which takes further time. Military decision-makers are always confronted with the need to balance depth of analysis with operational tempo. This provides a strong driver for complex systems to develop techniques for fusing, manipulating, and visualising large amounts of data to provide support to decision-makers in a timely fashion.

   A closely related issue is the meta-decision of when decisions are made. The distribution of when decisions are made over time will affect the balance between the choices taken in the preparation for conflict and those left open for online adaptation during conflict. Making decisions early limits the quantity and currency of information informing the decision, but is essential for decisions that have long time lags to implementation, and for decisions that have many follow-on decisions depending on a resolution. Some decisions will rely on inherently uncertain variables, such as enemy intentions, but other decisions can be made on the basis of factors it is possible to anticipate, such as regional capabilities. Extending techniques for analysing temporal interdependencies (such as the Program Evaluation Review Technique developed by the U.S. Navy) to apply to interdependencies between decisions within a dynamic complex system could lead to better allocation of decision support resources, and more effective decision-making.

8. How can swarms of unmanned vehicles be controlled to robustly perform a wide range of dynamic activities?

Unmanned Aerial Vehicles (UAVs) were first developed in the 1930s by the British and the U.S. as target drones for anti-aircraft gunnery training. By the 1960s the U.S. were also using UAVs for reconnaissance over Vietnam, China and North Korea, and battlefield UAVs were used by Israel during the 1982 operations in Lebanon [Goebel, 2009]. However, it was during the so-called Global War on Terror that the UAV proved its operational significance. The silent Predator endurance UAVs, invisible to the naked eye at cruise altitudes and equipped with supersonic Hellfire missiles, conducted a number of high profile attacks in Afghanistan, Yemen and Iraq by pilots stationed in the U.S. The U.S. Air Force fleet of Predators has grown from the three tested in Bosnia in 1995 to 195 by 2009 [Drew, 2009]. Unmanned Ground Vehicles (UGVs) and Unmanned Undersea Vehicles (UUVs) have been slower to reach maturity, because their environments are far more complex. Unmanned vehicles are especially suited to dirty, dangerous and dull missions. Likely missions for UGVs include decoy, surveillance, reconnaissance, explosive ordinance disposal, minefield clearance, logistics, and weapon delivery. The size of these platforms will vary from unmanned variants of armoured fighting vehicles to insect sized mobile sensors.

As the production of unmanned robotics platforms continues to rise and the costs fall, the prospect of large swarms of unmanned vehicles becomes a possibility. However, current UAVs are still manpower intensive. The Predator system requires 55 personnel to operate four UAVs. The primary impediment to swarming is the lack of a robust, scalable, real-time algorithm for command and control of the swarm. The challenge for complex systems is to develop and implement a theory for the distributed control of collectives that allows a swarm of communicating unmanned vehicles to be controlled by a single operator. The swarm would need to be able to dynamically switch between multiple kinds of missions, degrade gracefully when individual vehicles are lost, and communicate despite enemy jamming, deception and spoofing. The benefits of a swarm of cheap, disposable unmanned vehicles compared to today's expensive individual platforms potentially include lower vulnerability, greater robustness, persistent presence, multiple simultaneous effects, new mission possibilities, as well as greater psychological impact on the adversary. Research on probability collectives [Wolpert, 2006] and autonomic computing [Kephart and Chess, 2003] provide a starting point for this challenge.

9. Where are the complexity catastrophes?

   The competitive dynamic of military arms races provided inspiration in evolutionary biology for the theory that a biological system requires continual development just in order to maintain its fitness relative to the systems it co-evolves and competes with [van Valen, 1973]. This theory, called the Red Queen Hypothesis,[20] applies not just to arms races, but more gener-

[20] In Carroll's *Through the Looking Glass*, the Red Queen said "It takes all the running you

ally to entire militaries. Continual improvement is required just to maintain fitness. This in turn tends to promote complexification. Ashby's law of requisite variety explains why a more complex environment requires a more complicated regulatory system to counter environmental perturbations [Ashby, 1956]. Competition acts like a ratchet to increase the complexity of the environment, which creates a need for adaptation, which creates greater complexity for other actors, which stimulates their co-adaptation, which further increases the complexity of the environment for the original actor.

Complexity introduces new, often characteristic risks. The greater the complexity of one's own system, the greater the internal friction between components. Complexity is a common source of system incoherence, inter-scale conflict and cascading failure. Kauffman [Kauffman, 1993, p. 54] showed that complexity catastrophes due to conflicting constraints are a general property of complex systems. As variables within the system become more interdependent, the local fitness peaks available to adaptive processes decrease in height. According to Tainter, within a social context, in order to solve problems, societies respond by increasing their complexity, which creates new problems that require further increases in complexity [Tainter, 1988]. However, every increase in complexity requires more resources. Tainter argues that eventually diminishing returns on investment in complexity set in, and are the ultimate cause of societal collapse.

The challenge for complex systems is to identify potential complexity catastrophes. Whether military competition is accelerating towards catastrophe, or broader societal pressures result in regional collapse, complexity catastrophes will always have military implications. If complexity catastrophes can be identified in advance, strategies can be developed for avoiding complexification to the brink of collapse.

5 CONCLUSION

In this chapter we have shown that the science of complex systems has already had a significant impact on the way Western militaries understand and operate within exceedingly complex and dangerous environments. With the development of new methodologies, tools and techniques within complex systems, current military practices have been re-examined, and many have been found to be consistent with complex systems theory. In this case, complex systems provides theoretical support to explain why best practice works. Complex systems theory may also provide clues as to what changes in the context would cause current best practice to fail. In other instances, complex systems challenges the status quo. Agent based models of combat help reveal the limitations of mechanistic, detailed, physical wargames as representations of battlefield dynamics. Adaptive Campaigning

can do, to keep in the same place."

challenges the military transformation advocates' concept of information dominance, by arguing that armies will have to continue to fight for, rather than with, information. According to Adaptive Campaigning, acting on the basis of incomplete information will always be necessary to stimulate the system and assess its response. Whether or not it aligns with current practice, complex systems provides a valuable contribution to military theory.

In spite of the progress to date, the application of complex systems insights is not uniform. Real world applications often lag advances in theory. For example, the implications of Richardson's 1948 result that all intentional human violence, including all war, can be described by a single distribution, have not been operationalised in any international or national policy. The previous section presented a research agenda for complex systems applications to defence science and to defence problems. Using a framework comprised of pattern formation, multiscale analysis, and adaptation, applications were identified where complex systems has not yet had a significant impact. New tools can be adopted rapidly, but military culture is much more difficult to change. Some of the most promising defence applications of complex systems depend upon a more permissive cultural context. The power of adaptive, self-organising networks cannot be fully realised until old patterns of thinking change.

Perhaps the ultimate complex systems capability would be the general capacity to design adaptive organisations for solving global multiscale problems. While this is still largely aspirational, complex systems already offers unique insights that improve the design of systems and operations within defence. Nine further challenges for complex systems science with substantial military payoffs are identified above. Complex systems will continue to have increasing relevance to defence. Its applicability does not rely on increasing complexity, nor is it limited to the current operational challenges of stability, counter-insurgency, and counter-terrorism operations. War has always been complex, and new ways of comprehending its complexity will always be in demand.

ACKNOWLEDGEMENTS

Suggestions and critical review by Cliff Hooker and Anne-Marie Grisogono are greatly appreciated.

BIBLIOGRAPHY

[Abbott, 2009] R. Abbott. Causality in complex systems. `http://cs.calstatela.edu/wiki/index.php/Causality_in_complex_systems`, accessed 10/8/2009 2009.

[Alberts and Czerwinski, 1997] D. S. Alberts and T. J. Czerwinski, editors. *Complexity, Global Politics, and National Security.* CCRP Publication Series, 1997.

[Alberts *et al.*, 1999] D. S. Alberts, J. J. Garstka, and F. P. Stein. *Network Centric Warfare: Developing and Leveraging Information Superiority.* CCRP Publication Series, 2nd edition, 1999.

[Ashby, 1956] W. R. Ashby. *An Introduction to Cybernetics*. Chapman & Hall, London, UK, 1956.

[Australian Army Headquarters, 2004] Australian Army Headquarters. *Complex Warfighting*. Commonwealth of Australia, 2004.

[Australian Army Headquarters, 2006] Australian Army Headquarters. *Adaptive Campaigning: The Land Force Response to Complex Warfighting*. Commonwealth of Australia, 2006.

[Australian Army Headquarters, 2008] Australian Army Headquarters. *Adaptive Army Public Information Paper*. Commonwealth of Australia, 2008.

[Bak *et al.*, 1987] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality: An explanation of 1/f noise. *Physical Review Letters*, 59:381–385, 1987.

[Bar-Yam, 2003] Y. Bar-Yam. Unifying Principles in Complex Systems. In M. C. Roco and W. S. Bainbridge, editors, *Converging Technology (NBIC) for Improving Human Performance*. Kluwer, Dordrecht, The Netherlands, 2003.

[Bar-Yam, 2004] Y. Bar-Yam. Multiscale Variety in Complex Systems. *Complexity*, 9(4):37–45, 2004.

[Bar-Yam, 2005] Y. Bar-Yam. *Making Things Work: Solving Complex Problems in a Complex World*. NECSI Knowledge Press, Cambridge, 2005.

[Bernard *et al.*, 2007] S. Bernard, D. Gonze, B. Čajavec, H. Herzel, and A. Kramer. Synchronization-Induced Rhythmicity of Circadian Oscillators in the Suprachiasmatic Nucleus. *PLoS Computational Biology*, 3(4):0667–0679, 2007.

[Beyerchen, 1992] A. D. Beyerchen. Clausewitz, Nonlinearity and the Unpredictability of War. *International Security*, 17(3), 1992.

[Bigelow and Davis, 2003] J. H. Bigelow and P. K. Davis. Implications for Model Validation of Multiresolution, Multiperspective Modeling (MRMPM) and Exploratory Analysis. Technical Report MR-1750-AF, The RAND Corporation, 2003.

[Bogdanov, 1925 to 1929] A. A. Bogdanov. *Vseobshchaia organizatsionnaia nauka (tektologiia) (A Universal Organizational Science (Tectology), Vols. I-III*. Ekonomika, Moscow/Leningrad, 1925 to 1929.

[Bonabeau, 2002] E. Bonabeau. Agent-based modeling: methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99 Suppl 3:7280–7287, 2002.

[Bousquet, 2009] A. Bousquet. *The Scientific Way of Warfare: Order and Chaos on the Battlefields of Modernity*. Columbia University Press, New York, USA, 2009.

[Brandstein and Horne, 1998] A. G. Brandstein and G. E. Horne. Data Farming: A Metatechnique for Research in the 21st Century. In F. G. Hoffman and G. E. Horne, editors, *Maneuver Warfare Science 1998*. Marine Corps Combat Development Command, Quantico, VA, USA, 1998.

[Burns and Freeman, 2008] W. L. Burns and MAJ GEN (R) W. D. Freeman. Developing an Adaptability Training Strategy and Policy for the DoD Interim Report. Technical Report P-4358, Institute for Defense Analyses, 2008.

[Bush, 2002] G. W. Bush. *The National Security Strategy of the United States of America*. The White House, Washington, DC, USA, 2002.

[Campbell and Hummel, 1998] A. P. Campbell and J. R. Hummel. The dynamic information architecture system : an advanced simulation framework for military and civilian applications. In *Proceedings of the Advanced Simulation Technologies Conference*, Boston, USA, 1998.

[Caporale, 2002] L. H. Caporale. *Darwin In the Genome: Molecular Strategies in Biological Evolution*. McGraw-Hill, Maidenhead, UK, 2002.

[Cares, 1996] J. Cares. *Distributed Networked Operations: The Foundations of Network Centric Warfare*. Alidade Press, Rhode Island, USA, 1996.

[Carlson and Doyle, 2002] J. M. Carlson and J. Doyle. Complexity and Robustness. *PNAS*, 99(Suppl 1):2538–2545, 2002.

[Cebrowski and Garstka, 1998] VADM (R) A. K. Cebrowski and J. J. Garstka. Network-Centric Warfare: Its Origin and Future. In *Proceedings of the U.S. Naval Institute*, volume 124, pages 28–35, 1998.

[Christensen and Hooker, 1998] W. D. Christensen and C. A. Hooker. *Our Cultural Heritage*, chapter From cell to scientist: Toward an organisational theory of life and mind, pages 275–326. Australian Academy of Humanities, University House, Canberra, Australia, 1998.

[Cilliers, 2007] P. Cilliers. Knowing Complex Systems: The limits of understanding. In *International Conference on Complex Systems*, Boston MA, USA, 2007.

[Clauset and Wiegel, 2009] A. Clauset and F. W. Wiegel. A generalized fission-fusion model for the frequency of severe terrorist attacks. *Submitted to Journal of Conflict Resolution*, 2009.

[Clauset *et al.*, 2007] A. Clauset, M. Young, and K. S. Gledistch. On the Frequency of Severe Terrorist Events. *Journal of Conflict Resolution*, 51(1):58–88, 2007.

[Clauset *et al.*, 2009] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *Accepted by SIAM Review*, 2009.

[Clausewitz, 1984] C. von Clausewitz. *On War.* transl. M. Howard and P. Paret, Princeton University Press, Princeton, USA, 1984.

[Cohen and Gooch, 1990] E. A. Cohen and J. Gooch. *Military Misfortunes: The Anatomy of Failure in War.* The Free Press, London, UK, 1990.

[Cohen, 1980] E. Cohen. Systems Paralysis. *American Spectator*, 13, 1980.

[Coram, 2002] R. Coram. *Boyd: The Fighter Pilot Who Changed the Art of War.* Little, Brown and Company, New York, USA, 2002.

[Crutchfield, 2002] J. P. Crutchfield. What Lies Between Order and Chaos? In J. Casti, editor, *Art and Complexity.* Oxford University Press, New York, USA, 2002.

[Czerwinski, 2008] T. J. Czerwinski. *Coping with the Bounds: A Neo-Clausewitzean Primer.* CCRP Publication Series, revised edition edition, 2008.

[Davis and Finch, 1993] P. K. Davis and L. Finch. Defense Planning for the Post-Cold War Era: Giving Meaning to Flexibility, Adaptiveness, and Robustness of Capability. Technical Report MR-322-JS, The RAND Corporation, 1993.

[Davis and Henninger, 2007] P. K. Davis and A. Henninger. Analysis, Analysis Practices, and Implications for Modeling and Simulation. Technical Report OP-176-OSD, The RAND Corporation, 2007.

[Davison, 2008] MAJ K. Davison. From Tactical Planning to Operational Design. *Military Review*, September-October:33–39, 2008.

[Dawkins, 1983] R. Dawkins. Universal Darwinism. In D.S. Bendall, editor, *Evolution from Molecules to Men.* Cambridge University Press, 1983.

[DeRosa *et al.*, 2008] J. K. DeRosa, A-M. Grisogono, A. J. Ryan, and D. O. Norman. A Research Agenda for the Engineering of Complex Systems. *Proceedings of the 2nd Annual IEEE Systems Conference*, September-October, 7-10 April 2008.

[Dietz *et al.*, 2003] T. Dietz, E. Ostrom, and P. C. Stern. The struggle to govern the commons. *Science*, 302(5652):1907–1912, 2003.

[Digby, 1989] J. Digby. Operations Research and Systems Analysis at RAND, 1948-1967. Technical Report N-2936-RC, The RAND Corporation, 1989.

[Dixon, 1976] N. F. Dixon. *On the Psychology of Military Incompetence.* Basic Books, New York, USA, 1976.

[Dobias, 2008] P. Dobias. Complexity-Based Assessment in Crowd Confrontation Modelling. *Journal of Battlefield Technology*, 11(2):1–8, 2008.

[Dobias, 2009] P. Dobias. Self-Organized Criticality in Asymmetric Warfare. *Fractals*, 17:1–7, 2009.

[Dörner, 1996] D. Dörner. *The Logic of Failure: Recognizing and Avoiding Error in Complex Situations.* Metropolitan Books, New York, USA, 1996.

[Drew, 2009] C. Drew. Drones are weapons of choice in fighting qaeda. `http://www.nytimes.com/2009/03/17/business/17uav.html?_r=1&hp` accessed 28/6/09, March 2009.

[Edmonds, 1997] B. Edmonds. Hypertext Bibliography of Measures of Complexity. `http://www.bruce.edmonds.name/combib`, accessed 11/7/2008 1997.

[Engleback *et al.*, 2003] MAJ N. Engleback, M. A. Anderson, and LTC (Ret.) J. Wasser. Some Uses of Agent-Based Models in the New Zealand Defence Force. In G. E. Horne and S. Johnson, editors, *Maneuver Warfare Science 2003*, pages 155–170. Marine Corps Combat Development Command, Quantico, VA, USA, 2003.

[Enthoven and Smith, 2005] A. C. Enthoven and K. W. Smith. *Methods of Operations Research.* RAND Corporation, Santa Monica, CA, USA, 2005.

[Epstein, 1988] J. M. Epstein. Dynamic Analysis and the Conventional Balance in Europe. *International Security*, 12(4):154–165, 1988.

[Fagen, 1978] M. D. Fagen, editor. *A History of Engineering and Science in the Bell System: National Service in War and Peace (1925-1975).* Bell Telephone Laboratories, Inc., 1978.

[Gabriel, 2001] R. A. Gabriel. *Great Captains of Antiquity.* Greenwood Press, Westport, CT, USA, 2001.

[Galligan, 2004] D. P. Galligan. Modelling Shared Siturational Awareness Using the MANA Model. *Journal of Battlefield Technology*, 7(3):35–40, 2004.

[Gell-Mann, 1994] M. Gell-Mann. *The Quark and the Jaguar: Adventures in the Simple and the Complex*. Henry Holt and Company, New York, USA, 1994.

[Global Security, 2005] Global Security. Where are the Legions? Global Deployments of US Forces. `http://www.globalsecurity.org/military/ops/global-deployments.htm` accessed 4/6/2009, 2005.

[Goebel, 2009] G. Goebel. Unmanned Aerial Vehicles. `http://www.vectorsite.net/twuav.html` accessed 28/6/2009, 2009.

[Gray, 1971] C. S. Gray. What RAND Hath Wrought. *Foreign Policy*, (4):111–129, 1971.

[Grisogono and Ryan, 2007] A-M. Grisogono and A. J. Ryan. Operationalising Adaptive Campaigning. In *Proceedings of the CCRTS*, Washington, D.C., USA, 2007.

[Grisogono and Spaans, 2008] A-M. Grisogono and M. Spaans. Adaptive Use of Networks to Generate an Adaptive Task Force. In *Proceedings of the 13th ICCRTS*, Seattle, USA, 2008.

[Grisogono and Unewisse, 2009] A-M. Grisogono and M. Unewisse. Engendering Flexibility in Defence Forces. In *Proceedings of the 14th ICCRTS*, Washington, USA, 2009.

[Grisogono, 2004a] A-M. Grisogono. A Generic Framework for Generating and Exploring C2 Concepts. In *Proceedings of the 9th ICCRTS*, Copenhagen, Denmark, 2004.

[Grisogono, 2004b] A-M. Grisogono. What Do Natural Complex Adaptive Systems Teach Us About Creating A Robustly Adaptive Force? In *Proceedings of the 9th ICCRTS*, Copenhagen, Denmark, 2004.

[Grisogono, 2005] A-M. Grisogono. Co-Adaptation. In *SPIE Symposium on Microelectronics, MEMS and Nanotechnology*, volume Paper 6039-1, Brisbane, Australia, 2005.

[Grisogono, 2006a] A-M. Grisogono. Success and Failure in Adaptation. *Interjournal*, (1901), 2006.

[Grisogono, 2006b] A-M. Grisogono. The Implications of Complex Adaptive Systems Theory for Command and Control. In *Proceedings of the CCRTS*, San Diego, USA, 2006.

[Hammond, 2001] G. T. Hammond. *The Mind of War: John Boyd and American Security*. Smithsonian Institution Press, Washington, D.C, 2001.

[Hammond, 2003] D. Hammond. *The Science of Synthesis: Exploring the Social Implications of General Systems Theory*. University Press of Colorado, 2003.

[Hardin, 1968] G. Hardin. The Tragedy of the Commons. *Science*, 162(3859):1243–1248, 1968.

[Hayes, 2002] B. Hayes. Statistics of Deadly Quarrels. *American Scientist*, 90(1):10–15, 2002.

[Headquarters U.S. Marine Corps, 1996] Headquarters U.S. Marine Corps. *Marine Corps Doctrinal Publication 6, Command and Control*. U.S. Government Printing Office, Washington, DC, USA, 1996.

[Headquarters U.S. Marine Corps, 1997a] Headquarters U.S. Marine Corps. *Marine Corps Doctrinal Publication 1, Warfighting*. U.S. Government Printing Office, Washington, DC, USA, 1997.

[Headquarters U.S. Marine Corps, 1997b] Headquarters U.S. Marine Corps. *Marine Corps Doctrinal Publication 5, Planning*. U.S. Government Printing Office, Washington, DC, USA, 1997.

[Hoffman and Horne, 1998] F. G. Hoffman and G. E. Horne. *Maneuver Warfare Science 1998*. Marine Corps Combat Development Command, Quantico, VA, USA, 1998.

[Hooker, 2009] C. A. Hooker. Interaction and bio-cognitive order. *Synthese*, 166(3):513–546, 2009.

[Hoos, 1969] I. R. Hoos. *Systems analysis in social policy: a critical review*. Institute of Economic Affairs, London, UK, 1969.

[Horne, 1997] G. E. Horne. Data Farming: a Meta-technique for Research on 21st Century Questions. Presented to Naval Strategic Studies Group, Newport RI, USA, 1997.

[Horne, 2001] G. E. Horne. *Maneuver Warfare Science 2001*, chapter Beyond Point Estimates: Operational Synthesis and Data Farming. Marine Corps Combat Development Command, Quantico, VA, USA, 2001.

[Hughes Jr., 1994] W. P. Hughes Jr. Review of Complexity: The Emerging Science at the Edge of Order and Chaos, by M. Mitchell Waldrop. *PHALANX*, 27(3), 1994.

[Ilachinski, 1996a] A. Ilachinski. Land Warfare and Complexity, Part I: Mathematical Background and Technical Sourcebook. Technical report, Center for Naval Analyses, 1996.

[Ilachinski, 1996b] A. Ilachinski. Land Warfare and Complexity, Part II: An Assessment of the Applicability of Nonlinear Dynamics and Complex Systems Theory to the Study of Land Warfare. Technical report, Center for Naval Analyses, 1996.

[Ilachinski, 2000] A. Ilachinski. Irreducible Semi-Autonomous Adaptive Combat (ISAAC): An Artificial-Life Approach to Land Combat. *Military Operations Research*, 5(3):29–46, 2000.

[Ilachinski, 2004] A. Ilachinski. *Artificial War: Multiagent-based Simulation of Combat.* World Scientific Publishing, Hackensack, NJ, USA, 2004.

[Ingber, 1993] L. Ingber. Statistical Mechanics of Combat and Extensions. In C. Jones, editor, *Toward a Science of Command, Control, and Communications*, pages 117–149. American Institute of Aeronautics and Astronautics, Washington, D.C., USA, 1993.

[Ingber, 2003a] D. E. Ingber. Tensegrity I. Cell structure and hierarchical systems biology. *Journal of Cell Science*, 116:1157–1173, 2003.

[Ingber, 2003b] D. E. Ingber. Tensegrity II. How structural networks influence cellular information processing networks. *Journal of Cell Science*, 116:1397–1408, 2003.

[James, 1996] G. E. James. Chaos Theory: The Essentials for Military Applications. Technical Report Tenth in the Series, The Newport Papers, 1996.

[Johnson *et al.*, 2005] N. F. Johnson, M. Spagat, J. A. Restrepo, J. C. Bohorquez, N. Suarez, E. M. Restrepo, and R. Zarama. From old wars to new wars and global terrorism. *arXiv.org:physics/0506213*, 2005.

[Johnson *et al.*, 2006] N. F. Johnson, M. Spagat, J. A. Restrepo, O. Becerra, J. C. Bohorquez, N. Suarez, E. M. Restrepo, and R. Zarama. Universal patterns underlying ongoing wars and terrorism. *arXiv.org:physics/0605035*, 2006.

[Johnson, 2008] N. F. Johnson. Complexity in Human Conflict. In D. Helbing, editor, *Managing Complexity: Insights, Concepts, Applications*, pages 303–320. Springer, Berlin, Germany, 2008.

[Jomini, 1992] A. H. de Jomini. *The Art of War.* Greenhill Books, London, UK, 1992.

[Jones, 2009] MAJ D. Jones. Understanding the Form, Function, and Logic of Clandestine Cellular Networks: The First Step in Effective Counternetwork Operations. Master's thesis, School of Advanced Military Studies, 2009.

[Kauffman, 1993] S. A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution.* Oxford University Press, Oxford, 1993.

[Kelly, 1950] M. J. Kelly. The Bell Telephone Laboratories - An Example of an Institute of Creative Technology. *Proceedings of the Royal Society of London A*, 203(1074):287–301, 1950.

[Kephart and Chess, 2003] J. O. Kephart and D. M. Chess. The vision of autonomic computing. *Computer*, 36(1):41–50, 2003.

[Kinnard, 2007] D. Kinnard. *The War Managers.* Naval Institute Press, 2007.

[Lanchester, 1916] F. W. Lanchester. *Aircraft in Warfare: The dawn of the Fourth Arm.* Constable and Co., London, UK, 1916.

[Lauren and Stephen, 2002a] M. K. Lauren and R. T. Stephen. Fractals and combat modeling: Using MANA to explore the role of entropy in complexity science. *Fractals*, 10(4):481–489, 2002.

[Lauren and Stephen, 2002b] M. K. Lauren and R. T. Stephen. Map-Aware Non-uniform Automata (MANA)A New Zealand Approach to Scenario Modelling. *Journal of Battlefield Technology*, 5(1):27–31, 2002.

[Lauren *et al.*, 2002] M. K. Lauren, MAJ N. Engleback, R. T. Stephen, and M. A. Anderson. Modeling Precision Maneuver Using Mana. In G. E. Horne and S. Johnson, editors, *Maneuver Warfare Science 2002*, pages 109–130. Marine Corps Combat Development Command, Quantico, VA, USA, 2002.

[Lauren *et al.*, 2007] M. K. Lauren, G. C. McIntosh, N. Perry, and J. Moffat. Art of war hidden in Kolmogorov's equations. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 17(1), 2007.

[Lauren, 1999] M. K. Lauren. Characterising the Difference Between Complex Adaptive and Conventional Combat Models. Technical Report 169, DOTSE, 1999.

[Lauren, 2001] M. K. Lauren. Fractal Methods Applied to Describe Cellular Automaton Combat Models. *Fractals*, 9(2):177–185, 2001.

[Lauren, 2002a] M. K. Lauren. A Fractal Based Approach to Equations of Attrition. *Military Operations Research*, 7(3):17–30, 2002.

[Lauren, 2002b] M. K. Lauren. How Agent Models Can Address Asymmetric Warfare: An ANZUS Collaboration. In *SimTecT 2002 Conference Proceedings*, Melbourne, Australia, 2002.

[Lauren, 2005] M. K. Lauren. A Metamodel for Describing the Outcomes of the Mana Cellular Automaton Combat Model Based on Lauren's Attrition Equation. Technical Report 205, DTA, 2005.

[Lawrence, 1990] T. E. Lawrence. The Evolution of a Revolt. Technical report, Combat Studies Institute Reprint, 1990.

[Lim *et al.*, 2007] M. Lim, R. Metzler, and Y. Bar-Yam. Global Pattern Formation and Ethnic/Cultural Violence. *Science*, 317(5844):1540–1544, 2007.

[Lingard, 2009] D. M. Lingard. Perturbation Avalanches in Random Inference Networks. *Submitted to Complexity*, 2009.

[Los Alamos National Laboratory, 2007] Los Alamos National Laboratory. Complex systems group. Technical report, T-13 Theoretical Division, 2007.

[Matthews, 2004] D. Matthews. *Rethinking Systems Thinking: Towards a Postmodern Understanding of the Nature of Systemic Inquiry*. PhD thesis, The University of Adelaide, 2004.

[McIntosh and Lauren, 2002] G. C. McIntosh and M. K. Lauren. Firepower Concentration in Cellular Automata ModelsAn Alternative to Lanchester. *Journal of the Operational Research Society*, 53:672–679, 2002.

[McIntosh and Lauren, 2007] G. C. McIntosh and M. K. Lauren. Incorporating fractal concepts into equations of attrition for military conflicts. *Journal of the Operational Research Society*, 59:703–713, 2007.

[Meyer and Horne, 2007] T. Meyer and G. E. Horne, editors. *Scythe*, volume 1 of *Proceedings and Bulletin of the International Data Farming Community*. Naval Postgraduate School, 2007.

[Moffat, 2003] J. Moffat. *Complexity Theory and Network Centric Warfare*. CCRP Publication Series, 2003.

[Morse and Kimball, 1951] P. M. Morse and G. E. Kimball. *Methods of Operations Research*. John Wiley & Sons, New York, USA, 1951.

[Naveh, 1997] BG(R) S. Naveh. *In Pursuit of Military Excellence: The Evolution of Operational Theory*. Taylor & Francis, London, UK, 1997.

[Osinga, 2005] F. Osinga. *Science, Strategy and War: The Strategic Theory of John Boyd*. Eburon Academic Publishers, Delft, Netherlands, 2005.

[Ott *et al.*, 1990] E. Ott, C. Grebogi, and J. A. Yorke. Controlling chaos. *Physical Review Letters*, 64:1196, 1990.

[Palmore, 1997] J. Palmore, editor. *Warfare Analysis and Complexity*. Military Operations Research Society, 1997.

[Przemieniecki, 1994] J. S. Przemieniecki. *Mathematical Methods in Defense Analyses*. American Institute of Aeronautics and Astronautics, Inc., Washington, DC, USA, 2nd edition, 1994.

[Richardson, 1948] L. F. Richardson. Variation of the frequencey of fatal quarrels with magnitude. *Journal of the American Statistical Association*, 43(244):523–546, 1948.

[Richardson, 1960] L. F. Richardson. *Statistics of Deadly Quarrels*. Boxwood Press, Pittsburgh, USA, 1960.

[Ricks, 2009] T. E. Ricks. *The Gamble: General David Petraeus and the American Military Adventure in Iraq, 2006-2008*. Penguin Press HC, The, 2009.

[Romeiras *et al.*, 1992] F. Romeiras, C. Grebogi, E. Ott, and P. Dayawansa. Controlling chaotic dynamical systems. *Physica D*, 58:165, 1992.

[Ryan and Grisogono, 2004] A. J. Ryan and A-M. Grisogono. Hybrid Complex Adaptive Engineered Systems: A Case Study in Defence. *Interjournal*, 2004.

[Ryan and Norman, 2006] A. J. Ryan and D. O. Norman. Agile and Flexible Capability Development. In *Proceedings of the Land Warfare Conference*, Brisbane, Australia, 2006.

[Ryan *et al.*, 2005] A. J. Ryan, S. Wheeler, and T. Turner. Impact of Correlations in Logistics Supply Chains. In L. Caccetta and V. Rehbock, editors, *Proceedings ASOR Conference*, pages 146–153, Curtin University of Technology, Perth, Australia, 2005.

[Ryan, 2006] A. J. Ryan. About the bears and the bees: Adaptive approaches to asymmetric warfare. *Interjournal*, 2006.

[Ryan, 2007a] A. J. Ryan. *A Multidisciplinary Approach to Complex Systems Design*. PhD thesis, The University of Adelaide, 2007.

[Ryan, 2007b] A. J. Ryan. Emergence is coupled to scope, not level. *Complexity*, 13(2):67–77, 2007.

[Ryan, 2007c] LTC M. Ryan. The Other Side of the COIN: Reconstruction Operations in Southern Afghanistan. *Australian Army Journal*, IV(2):125–144, 2007.

[Salthe, 2002] S. N. Salthe. Summary of the Principles of Hierarchy Theory, 2002. `http://www.harmeny.com/twiki/pub/Main/SaltheResearchOnline/HT_principles.pdf`, accessed 10/6/2006.

[Schneider, 1997] J. J. Schneider. Black Lights: Chaos, Complexity and the Promise of Information Warfare. *Joint Forces Quarterly*, Spring:21–28, 1997.

[Simon, 1962] H. A. Simon. The Architecture of Complexity. *Proceedings of the American Philosphical Society*, 106(6), 1962.

[Smith Jr. and Grisogono, 2006] E. A. Smith Jr. and A-M. Grisogono. Towards a Synthesis of Complex Adaptive Systems Theory and Effects Based Approaches to Operations: Warfighters to Coalition - A Case Study of Multi-level Adaptation in Effects-Based Operations. In *Proceedings of the 11th ICCRTS*, Cambridge, UK, 2006.

[Smith Jr. *et al.*, 2008] E. A. Smith Jr., A-M. Grisogono, and M. N. Clements. Cajole and Coordinate? C2 in Whole-of: -Government, -Nation, and -Coalition Action. In *Proceedings of the 13th ICCRTS*, Seattle, USA, 2008.

[Spaans *et al.*, 2009] M. Spaans, M. Spoelstra, E. Douze, R. Pieneman, and A-M. Grisogono. Learning to be Adaptive. In *Proceedings of the 14th ICCRTS*, Washington, USA, 2009.

[Summers, 1992] H. G. Summers. *On Strategy: A Critical Analysis of the Vietnam War*. Presidio Press, London, UK, 1992.

[Surowiecki, 2005] James Surowiecki. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. Anchor, New York, USA, 2005.

[Sutton and Barto, 1998] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. A Bradford Book, The MIT Press, Cambridge, 1998.

[Tainter, 1988] J. Tainter. *The Collapse of Complex Societies*. Princeton University Press, 1988.

[Thucydides, 1996] Thucydides. *The Landmark Thucydides: A comprehensive guide to the Peloponnesian war*. Simon & Schuster, New York, NY, USA, 1996.

[Tobin, 2006] D. R. Tobin. *Re-educating the corporation: Foundation for the Learning Organization*. John Wiley, New York, USA, 2006.

[TTCP, 2006] TTCP. *Guide for Understanding and Implementing Defense Experimentation (GUIDEx)*. The Technical Cooperation Program, Ottawa, Canada, 2006.

[Tumer and Wolpert, 2004] K. Tumer and D. H. Wolpert. *Collectives and the Design of Complex Systems*. Springer, New York, USA, 2004.

[Unewisse and Grisogono, 2007] M. H. Unewisse and A-M. Grisogono. Adaptivity Led Networked Force Capability. In *Proceedings of 12th ICCRTS*, Newport, Rhode Island, USA, 2007.

[van Valen, 1973] L. van Valen. A new evolutionary law. *Evolutionary Theory*, 1:1–30, 1973.

[von Bertalanffy, 1950] L. von Bertalanffy. The Theory of Open Systems in Physics and Biology. *Science*, 111(2872):23–29, 1950.

[von Bertalanffy, 1956] L. von Bertalanffy. General System Theory. *General Systems*, 1:1–10, 1956.

[Wass de Czege, 2009] BG(R) H. Wass de Czege. Systemic Operational Design: Learning and Adapting in Complex Missions. *Military Review*, January-February:2–12, 2009.

[Weaver, 1948] W. Weaver. Science and Complexity. *American Scientist*, 36(536), 1948.

[Welsh *et al.*, 1995] D. K. Welsh, D. E. Logothetis, M. Meister, and S. M. Reppert. Individual neurons dissociated from rat suprachiasmatic nucleus express independently phased circadian firing rhythms. *Neuron*, 14(4):697–706, 1995.

[Wiener, 1948] N. Wiener. Cybernetics. *Scientific American*, pages 14–19, 1948.

[Wilson and Sober, 1994] D. S. Wilson and E. Sober. Reintroducing group selection to the human behavioral sciences. *Behavioral and Brain Sciences*, 17(4):585–654, 1994.

[Wilson, 1979] K. G. Wilson. Problems in Physics with Many Scales of Lengths. *Scientific American*, 241(2):158–179, 1979.

[Wolpert, 2006] D. H. Wolpert. *Complex Engineered Systems: Science Meets Technology*, chapter Information Theory — The Bridge Connecting Bounded Rational Game Theory and Statistical Physics . Understanding Complex Systems. Springer, New York, USA, 2006.

[Woodcock *et al.*, 1988] A. E. R. Woodcock, L. Cobb, and J. T. Dockery. Cellular Automata: A New Method for Battlefield Simulation. *Signal*, 42(5):39–50, 1988.

[Yang *et al.*, 2005] A. Yang, H. A. Abbass, and R. Sarker. *Knowledge-Based Intelligent Information and Engineering Systems*, chapter WISDOM-II: A Network Centric Model for Warfare. Lecture Notes in Computer Science. Springer, New York, USA, 2005.

[Yang *et al.*, 2006] A. Yang, H. A. Abbass, and R. Sarker. Characterizing warfare in red teaming. *IEEE Transactions on Systems, Man, Cybernetics, Part B*, 36(2):268–285, 2006.

Part XI

Public Policy / Management

COMPLEXITY AND MANAGEMENT

Peter M Allen

1 INTRODUCTION

The scientific study of open systems has led to the science of complexity — that is the science of evolutionary change and self-transformation. It applies to systems that can undergo spontaneous, symmetry breaking transformations corresponding to qualitative change with new emergent features, capabilities and processes and do not simply grow or decline within a fixed set of dimensions. Until the study of open systems scientific understanding was sought by considering the results of "repeatable" experiments. In this way a "fact" could be stated which, by repeating the experiment, could be verified by others. Clearly, this applied to closed systems, which were independent of the particular environment. Whilst this incredibly simple, but powerful, idea was able to give us much of our science and technology, it clearly excluded open systems — systems whose existence arose from an evolutionary process and was sustained through their interactions with the environment.

One key difference between the science of closed mechanical systems of classical physics and social systems is that discovering and publishing a "law" in physics does not cause the behaviour involved to change, while in social science it often does. This is because it changes peoples' perceptions and understanding of their situation and therefore leads them to change behaviour, which can in turn invalidate the original findings. This establishes a fundamental difference between the fixed backcloth of physics and the "laws of nature", and the shifting, evolving and changing nature of any observed patterns and regularities of social science. The co-evolution of interacting elements, with emergent collective capabilities, is the characteristic feature of social systems, and this open-ended evolution is what is difficult to capture in a mathematical model, since usually such models consist of a fixed set of equations, only able to represent the behaviour resulting from a fixed set of mechanisms characteristic of a closed system. Although such closure can be useful in considering operational matters in the very short term, individuals, groups, firms, markets, towns and cities are always strongly connected to their changing environment and therefore a representation of these as closed or fixed will be misleading and, the long run, will lead to failure.

This evolutionary view agrees with that of Schumpeter's "creative destruction" where the most important factors in markets are the things that enter and leave the system over time — innovations and failures — rather than the particular

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

firms or organizations present at any particular time [Schumpeter, 1975]. In such a view learning — at any rate a capacity to adapt and change — is therefore more important over the longer term than being optimal in the immediate. There is a trade-off between being optimal at a given time and being able to learn, adapt and change, because the former focuses organisation and resources of perfecting currently winning technique, anything else considered wasteful, while the latter requires devoting organisation and resources to possible changes, most of which will never be exercised. Thus we see that the strategies and organizational properties required for long term survival are not those of a calculable, rational efficiency.

This has important implications for policy and management for such complex systems. For simple predictable systems whose behavioural laws are independent of management intervention, and for systems that can for practical purposes be treated in this way, policies can aim to optimise some (independently specified) valuable condition or product output, and the managerial means for doing so can be deduced from the system laws, so that management can be focused on the optimally efficient means for achieving the policy goal, which is entirely an exercise in technically calculable rational efficiency. Call this the minimal, classic account of policy and management formation, employing classical instrumental rationality. But complex open systems deeply challenge this approach: (a) epistemologically, the system cannot be known in advance independently of managerial behaviour, so there are insufficient, accessible and predictable interventions to support the application of any simple optimisation and (b) rationally, there is thus insufficient basis for applying the classical instrumental optimising conception of rationality as the basis of a managerial approach. Instead, the requirement of optimal efficiency needs to be replaced by one of "sufficient efficiency" ('satisficing' — [Simon, 1957; 1978; 1983]) combined with the capacities to learn and adapt.

The irreducible uncertainty of the open-ended co-evolution of things means that a messy, micro-diversity is the only insurance against an unknown future, and that social evolution will proceed through successive periods of drift and diversification separated by shorter spans of selective elimination. Evolution and co-evolution only demonstrate what is not viable at a particular time, and do not imply that what remains is an optimal structure that achieves anything in particular. Models can be built that capture the behaviour of multiple interacting agents and the way in which their beliefs are confounded, reinforced or updated as they struggle to make sense of their changing circumstances. Complex systems models can therefore help us explore the consequences of different possible practices, values and beliefs, perhaps indicating some basic features that will underlie any functioning society.

2 ORGANISATION, COMPLEXITY AND MANAGEMENT

Policies and management practices are not in general simply imposed from outside, as the classic account might suggest, but are normally generated from within a firm or government and must be carried out by that same system. Thus the organisational nature of the system matters to policy and management, especially when it is itself impacted by complexity.

The context of an organization is of course its physical, economic and technological environment, as well as other organizations, individuals and the cultural and social realities of the moment. As Tsoukas and Chia [2002] say: "Firstly, organization is the attempt to order the intrinsic flux of human action, to channel it towards certain ends by generalizing and institutionalizing particular cognitive representations. Secondly, organization is a pattern that is constituted, shaped, and emerging from change." The co-evolution of an organization with its context is therefore about the continual to and fro of modification as the "inside" and the "outside" of the organization restructure over time, blurring the separation and indeed, sometimes radically re-defining the boundary. So, organizations may separate into different specialist arms, or outsource work that was previously done in-house. A supply chain may become the relevant competitive unit rather than the firm, and indeed, we may see that evolution is governed by an ecology of interacting entities, none of which control the changes that are occurring.

In an recent review of complexity and organizational change Burnes [2005] points out that there is in fact a very large consensus that organizations are facing unprecedented levels of change and that their ability to manage change is therefore of great importance.[1] However, despite this, organizational change has proved to be very difficult to achieve successfully, with up to 80% failure rate.[2] This suggests that the traditional way of looking at and planning organizational change is somewhat flawed, and that perhaps complexity can offer us some help in improving this performance.

Despite the ubiquitous nature of evolutionary processes, we still tend to understand and make sense of what is occurring by looking at successive "snapshots" of the organization at different moments. Understanding these changes becomes related to seeing these as successive "stable" (temporarily) regimes of operation. These arise and persist (temporarily) when the interactions between their heterogeneous elements are such as to lead to a flow of resources from the environment, necessary to maintain the elements and their coordination. This implies a continual co-evolution between the identity of an individual or organization and its environment. Clearly, any structure or organization can either persist by finding an environment that will maintain it, or it must adapt itself to draw sustenance from where it is. Ultimately then, it is the second law of thermodynamics that dictates that the persistence of any coherent structure requires supporting flows of energy and matter, and to a necessary co-evolution of structure and flows.

The importance of the openness of an organization to its environment points to the absolutely fundamental significance of the idea of the "fit" between an

[1]See, e.g., [Brown and Eisenhardt, 1997; Cooper and Jackson, 1997; Dawson, 2003; Dunphy, *et al.*, 2003; Greenwald, 1996; Johnson and Scholes, 2002; Kanter, *et al.*, 1997; Kotter, 1996; Peters, 1997; Romanelli and Tushman, 1994].

[2]See e.g. [Beer and Nohria, 2000; Brodbeck, 2002; Bryant, 1998; Burnes, 2004b; Clarke, 1999; Harung, *et al.*, 1999; Huczynski and Buchanan, 2001; Stickland, 1998; Styhre, 2002; Whyte and Witcher, 1992; Witcher, 1993; Zairi, *et al.*, 1994].

organization and its environment. This immediately tells us that the "fitness" of any organization or structure is a measure of its ability to elicit, capture or merit resources from its environment. In order to maintain "fitness" in a changing environment then, it will be necessary for the organization to be capable of actively transforming itself over time. This has to be the central concern of management and the principles on which it is based. This creates a second level of explanation of organizational behaviour and of strategic management. The first level is that it must be such as to obtain resources from the current environment. The second is that it must also be capable of changing itself in such a way as to respond to the changing environment. This environment may well consist among other things of other organizations with similar objectives, and so we can directly see that there will be two directions of change: 1) the ability to out compete similar organizations, or 2) the ability to discover other "niches" which can still command resources, but which escape the competition.

Instead of management being based on the idea of creating a set of maximally efficient operations that will produce some good or service for a particular market, complexity recognizes the reality of openness, change in both the supply and demand situations, and the continual appearance of new ideas, technologies and competitors. It therefore replaces the requirement of "maximal efficiency" to one "sufficient efficiency" combined with the capacity to adapt, change and learn.

Instead of management being based on the idea of creating a set of maximally efficient operations that will produce some good or service for a particular market and organisation being designed to support that process, complexity recognizes the reality of openness, not only in change in both the supply and demand situations and the continual appearance of new ideas, technologies and competitors, but also in organisation itself in response to these as demands for learning and adapting arise and their focus continually shifts to track change. Call the organisational capacities for learning and adaptation, adaptability. This discussion of organisational dynamics reinforces the replacement of maximal efficiency with "sufficient efficiency" combined with sufficient adaptability, but emphasises the significance of self-organised, organisational change in underwriting this process.

3 EVOLUTIONARY DRIVE AND ORGANISATIONAL ADAPTIVENESS

"Evolutionary Drive" was put forward some years ago [Allen and McGlade, 1987a] as the underlying mechanism that describes the change and transformation of complex systems. In this view evolution is driven by the interplay over time of processes that create micro-diversity at the elemental level of the system and the selection operated by the collective dynamic that results from their interaction together with that of the system with its environment.

This co-evolution is seen as a continuous, on-going process and not one that has already "run its course", as in the case of "evolutionary stable strategies" [Maynard-Smith, 1979]. Because of the ignorance of individuals as to the pay-offs that will occur over time for a given behaviour, there are always new aspects of micro-diversity that can occur, so that co-evolution never reaches an "optimal" outcome, as in the Game Theory approach. Instead, we see this multi-level exploration and retention process as an on-going process that is occurring in real time, enabling the system to respond to currently undefined changes in the environment. History is still running. Each behavioural type is in interaction with others, and therefore evolutionary improvements may lead to greater synergy or conflict between behaviours, and in turn to lead to a chain of responses without any obvious end. And if there is no end, then the most that can be said of the behaviour of any particular individual or population is that its continued existence proves *only that it has been "good enough" — but not that it is* optimal.

And this brings "evolutionary drive" very close to the ideas of creative destruction and evolutionary economics expressed initially by Schumpeter [1975], Foster and Kaplan [2001] and Metcalfe [1998; 1999], as well as to the views of Tsoukas and Chia [2002] who see organization as an emergent and passing outcome of an on-going evolution. We shall therefore establish the basis of the complex evolutionary processes that give rise to the emergence and development of organisations and the behaviours and identities of the participating elements and individuals.

In this review of the ideas that link complexity to the management of organizations the guiding premise is that successful organizations require underlying mechanisms that continuously create internal micro-diversity of ideas, practices, schemata and routines — not that they will all be taken up, but so that they may be discussed, possibly tried out and the either retained or rejected. It is this that will drive an evolving, emergent system that is characterised by qualitative, structural change. These mechanisms also explain exaptation [Gould and Vrba, 1982; Gould, 2002] since the micro-diversity pre-exists the "uses" and "niches" that they later embody.

It firmly anchors success in the future on the tolerance of, and ultimately organisational understanding of, seemingly unnecessary perspectives, views, and ideas, since it is through the future implementation of some of these that survival will be achieved. If we define the diversity within an organization as the different functional types that are present — the number of different boxes in a systems diagram — then this System Diversity is in fact driven by the existence of micro-diversity within the functional types — at the level below. In other words the organizational behaviour and the functional types that comprise it **now**, have been created from the competitive and/or cooperative interactions of the micro-diversity that occurred within them in the **past**.

4 IGNORANCE AND LEARNING

Let us now consider a number of interacting elements that are of the same type, but differ from each other in detail. Nobody may know whether these differences will make a difference as they are just random variations around a reasonable average. But, consider that there is in fact a "fitness" landscape which actually reflects better and worse "fit" of the diverse individuals to the environment. In biological evolution, the variation is caused mainly by genetic variation which leads to phenotypic heterogeneity for which the fitness landscape will provide differential survival and reproduction rates, thus defining and amplifying the fitter, and suppressing the less fit. In this way, the existence of mechanisms that provoke genetic and phenotypic variation will automatically produce the exploration of the fitness landscape.

In a competitive market it will be true that differential performance will also serve to automatically amplify the fitter firms, and suppress the less fit. But of course, now "fitter" will be defined by customers and investors through the choices they make. Within organizations, however, evolutionary change will require the differences in performance of the different individuals, ideas, practices or routines, to be noticed and then for what works well to be deliberately reinforced, and what works less well to be discouraged. It is necessary to have multiple possible ideas, routines, practices and behaviours inside organizations in order to provide a spread of behaviours upon which differential dynamics can act. This differential dynamics is really the "selection" term of Darwin, operated by the environment, the customers and investors of the market place, or by the beliefs of the upper echelons of the organization. The organization chooses between different possible practices, routines etc. and the market chooses between different possible products and services made by firms — and the only thing that is certain is that if there is no diversity to choose between, then nothing can happen. Evolution can only occur if there is something that will generate heterogeneity spontaneously. And this is in fact the current ignorance about future outcomes and a lack of commonly agreed norms of how things should be done. It is this ignorance and freedom that allows creative individual diverse beliefs about what may work. Individual freedom and non-conformity is the secret power behind learning.

We can devise a simple computer program to demonstrate this idea, by considering a population that initially sits at a low point of a fitness landscape, and then has random variation of individual fitness. These are equally distributed left and right, but those that are of higher fitness are amplified slightly compared to those of lower fitness, and so gradually the population will "climb" the local fitness landscape. This tells us that the population increases in fitness — climbs the hill — because of processes of "exploration" in character space. Ignorance and consequent randomness are very robust sources of such exploration. Clearly random changes in teh design of any complicated entity will mean that most experiments are simply non-viable, and only half (randomly) of those that remain are "better". This effectively tells us that ther eis an "opportunity cost" to behavioural exploration. Such simple ideas can be extended to examine "how much" diversity creation (randomness) wins in the discovery of better performance. Now we consider two populations simultaneously at the foot of the fitness hill, and see how successfully they climb it. One population has a higher rate of randomness in character space than the other and therefore generates more heterogeneity and behavioural exploration. However, there is also a considerable "opportunity" cost in exploring rather than making perfect copies as most novel individuals will in fact be non-viable. Initially, the "explorer" population wins, because, despite its cost in non-viable individuals, diffusing and innovating faster is rewarded by the fitness slopes it discovers. Later, however, when the landscape has been explored and climbed, faster diffusion is no longer rewarded and the more conservative population with less exploration eventually dominates.

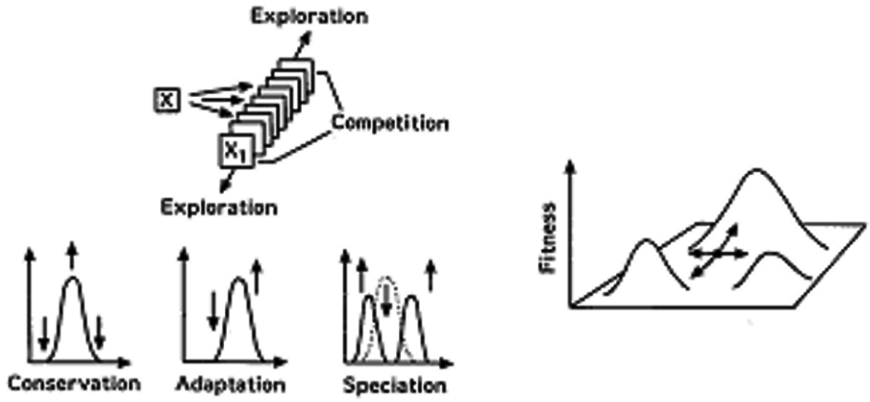


Figure 1. The coupled effects of mechanisms producing individual heterogeneity and differential performance in the collective dynamics leads to Evolutionary Drive.

When a new domain has been revealed and there is much to learn — then high rates of exploration pay off. However, as the system matures and much terrain has already been thoroughly explored the return on further exploration falls off and investments switch from this sector to another, newer one. We will see an overall pattern in which new sectors are discovered, and after an initial flowering of innovation and diversity the successful behaviour switches to that of fixed routines of exploitation.

The production of individual heterogeneity leads to differential performance of the collective dynamics, leading to the amplification of some individual types and the suppression of others. This in turn leads to changing performance at the level of the organization as a whole, and to a changing role in the larger environment. The performance of the organization within the wider environment changes as a result and so this will tend to amplify organizations in which internal learning is working well, and to suppress those where it is not. This also means that where there is selection at the lower level in favour of behaviours that do not improve the organizational performance as a whole then these will tend to be eliminated as a result of competition in the environment. Evolutionary drive will automatically lead to the selection of organizations in which there is an alignment between the selection processes at the individual level and at the organizational level.

In human systems, the typical development of economic sectors and markets, as shown in the work of Hirooka [2003], both expresses and results from the fact that initially search is rewarded, and therefore the bundling of components attracts investments. However, as the sector becomes well-established the pay-offs to new types of products falls, and so investment falls. It then switches to some other area that may seem to offer better opportunities. This is exactly what our simple theoretical model above predicts.

The presence of firms with different levels of exploration and exploitation (error-making and accuracy) will automatically lead to evolution selecting whichever is most appropriate. So, evolution will be driven by the amount of diversity generation to which it leads.

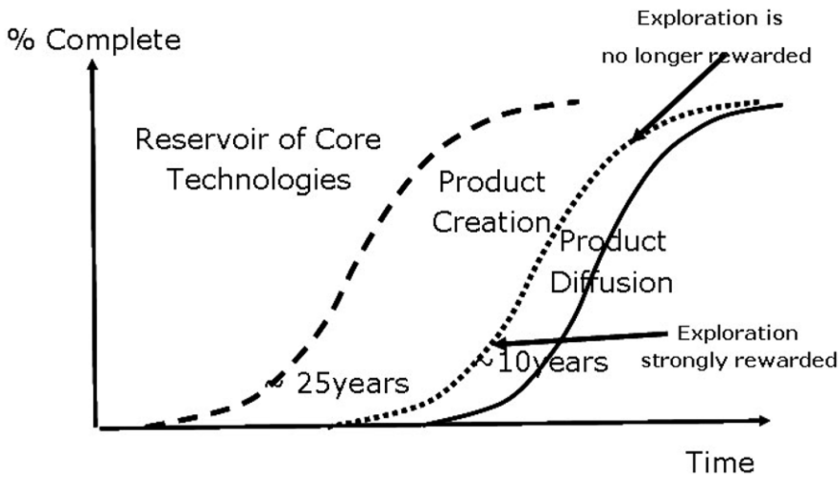


Figure 2. From [Hirooka, 2003]. The evolution of any industrial sector (here consumer electronics) may be explained by a similar dynamics, where investment gradually switches from exploration to exploitation.

Evolution selects for an appropriate capacity to evolve, and this will be governed by the balance between the costs of experimental "failures" (the non-viable individuals created) and the improved performance capabilities discovered by the exploration. This is what holds "total diversity generation" in check.

5 COMPLEXITY, EVOLUTIONARY DRIVE AND SIMPLIFIED REPRESENTATION

Evolutionary Drive tells us that evolution is driven by the noise to which it leads. Providing that microscopic diversity (noise) is produced in systems of interacting populations, the dynamics resulting from the interactions will lead to the retention and amplification of some, and the suppression of others. Not only that, the selection process operated by the "level above" will automatically select for individuals and groups that retain the ability to create micro-diversity. This process will select for the "ability to evolve" as well as for particular types of micro-diversity at a given time. This situation reinforces the earlier epistemic theme of our limited knowledge of our own systems. There will never be a completely clear understanding of any evolving system at a given time, because it will always contain micro-diverse elements that may or may not turn out to be successful. The understanding that we can have of reality is obtained by creating a "system" of interacting entities that are sufficiently correct to describe the current situation, but inadequate to predict the future structural evolution that may occur.

We understand situations by making creative, but simplifying assumptions. We define the domain in question (the boundary) and by establishing some rules of classification (a dictionary) that allow us to say what things were present when. This means that we describe things strategically in terms of words that stand for classes of objects. The "evolutionary tree" is an abstraction concerning types of thing rather than things themselves. In order to go further in our thinking, and get more information about an actual situation, we then consider only the present, and ask, what is this system made of NOW, and how is it operating NOW? This is "operational" not strategic. It therefore assumes structural stability and takes us away from open, evolutionary change, to the effects of running a fixed set of processes. If the events considered are discreet, then the running is according to a probabilistic dynamics, and we have what is called stochastic non-linear dynamics, where different regimes of operation are possible, but the underlying elements never change nor learn, nor tire of their behaviours. If we assume that we can use average rates instead of probabilities for the events, then we arrive at deterministic, system dynamics. This is in general, non-linear dynamics, and may be cycles or chaos or at equilibrium, but what happens is certain, simple and easy to understand.

In figure (3) we show how successive assumptions are made in order to "understand" the real situation. On the left hand side we have the "cloud" of reality and practice. Here, we are in the realm of non-science, in which people try to sum their experiences informally, and come up with heuristic rules and folklore of various kinds to deal with the problems of the real world.

Science begins by deciding on a boundary within which explanation will be attempted, in the context of the environment outside. The second assumption is that of classification. The elements present within the boundary are classified into types, so that potentially, previously established behaviour and responses of similar types can be used to predict behaviour. In examining any system of interest over some long time, however, it will be found that the components and elements present in the system have in fact changed over time. Qualitative evolution has occurred in which some types of component have disappeared, others have changed and transformed, and others still have appeared in the system initially as innovations and novelties.

This possibility of transformation and evolution is the domain of the complex system that co-evolves with its environment.

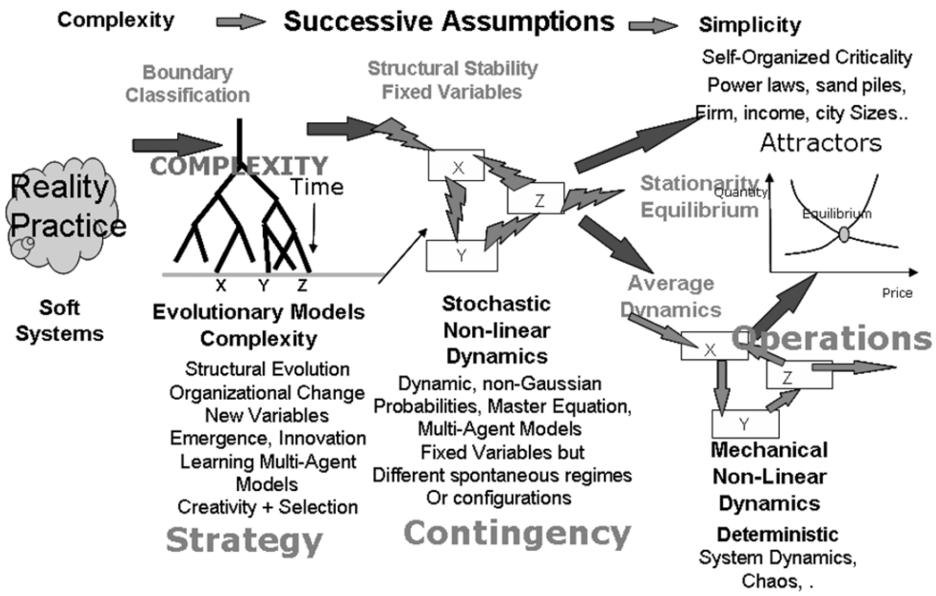


Figure 3. The choice of successive assumptions that lead to "scientific" understanding of a situation.

If we are interested in understanding the behaviour of the existing system then we can simply take the inventory and description now, and consider the "working" of the components. This will lead to probabilistic, coupled equations describing their interactions. These will be probabilistic in the absence of detailed knowledge of the precise, mechanical operation of the parts. This probabilistic dynamics will give rise to a dynamics of average values, of variances and higher moments, as a result of the coupled processes. Because there may be different possible attractor basins for the dynamics, the statistical distribution will reflect this, being multimodal and not a simple Gaussian or normal distribution. The dynamical system will itself generate the distribution and all its moments, including the variance, and so there will be no simplification into an "average" dynamic with a given random variance. Instead, all the moments will really play a role in defining the evolution of the distribution. These kinds of probabilistic dynamics are described by what are called the Master or Kolmogorov equations governing the change in the probability distributions as a result of the non-linear interactions.

There are now two different routes to simplification:

- Consider the stationary solution of the probabilistic non-linear dynamics — the Master or Kolmogorov equation
- Consider the dynamics of the average, mean or first moment of the probability distribution, and assume that this can be uncoupled from the higher

Table 1. The general complexity framework

Number	**Assumption Made**	**Resulting Model**
1	Boundary assumed	Some local sense-making possible — no structure supposed.
2	Classification assumed	Strategic, Open-ended Evolutionary — structural change occurs. Statistical Distributions part of evolutionary process, can be multi-modal
3	Average Types	Operational, Probabilistic, Non-Linear Equations, Master Equations, Kolmogorov Equations — assumed structurally stable. Statistical distributions can be multi-modal or power laws
	First Pathway	
4	Statistical Attractors	Self-Organized Criticality, Power law Distributions. Power law distributions
	Second Pathway	
4	Average events, dynamics of average agents	Deterministic Mechanical Equations, System Dynamics — assumed structurally stable. No statistical distribution.
5	Attractors of Non-Linear Dynamics	Study of attractors, Catastrophe Theory. Non-Linear dynamics with point, cyclic or chaotic/strange attractors.

moments. This leads to deterministic system dynamics — a mechanical representation of the system. We can then study the attractors of this simplified system, and find either point, cyclic or chaotic attractor dynamics as the long term outcome.

This succession of models arises from making successive, simplifying assumptions, and therefore models on the right are increasingly easy to understand and picture, but increasingly far from reality. The operation of a mechanical system may be easy to understand but that simplicity has assumed away the more complex sources of its adaptability. A mechanical model is more like a "description" of the system at a particular moment, but does not contain the magic ingredient of micro-diversity whose constant replenishment generates Evolutionary Drive. The capacity to evolve is generated by the behaviours that are averaged by assumptions 3 and 4 — average types and average events — and therefore organisations or individuals that can adapt and transform themselves, do so as a result of the generation of micro-diversity and the interactions with micro-contextualities. This tells us the difference between a reality that is "becoming" and our simplified understanding of this that is merely "being" [Prigogine, 1981].

6 MANAGING IN HUMAN SYSTEMS

Behaviours, practices, routines and technologies are invented, learned and transmitted over time between successive actors and firms, and we shall discuss how the principles of Evolutionary Drive can be used to understand them as well in human systems as in the biological.

6.1 A fisheries example

A detailed model was developed of Canadian Atlantic fisheries [Allen and McGlade, 1987b; Allen, 1994; 2000]. This consists of a spatial model with 40 zones over which up to eight fleets of fishermen roam, attempting to make a living by finding and catching different fish populations. The fishermen's movements are informed by their "mental maps" of potential profit, and this is based on the information they have from other fishermen. Clearly, members of their own fleet will exchange information, but this may not be true for other fleets. The model therefore describes fishing as a result of the acquisition and exploitation of knowledge about where profits are being made by other fishermen. This means that there is a spatial positive feedback mechanism that structures fishing patterns. Of course, in deciding which zone to go to, fishermen take into account the distances involved to go there, and to return to port and the cost of fuel.

In addition to these effects, however, our equation takes another very important factor into account. This factor R expresses how "rationally", how "homogeneously" or with what probability a particular skipper will respond to the information he is receiving. For example, if R is small, then whatever the "real" attraction of a zone i, the probability of going to any zone is roughly the same. In other words, "information" is largely disregarded, and movement is "random". We have called this type of skipper a *stochast*. Alternatively if R is large, then it means that even the smallest difference in the attraction of several zones will result in every skipper of that fleet going, with probability 1, to the most attractive zone. In other words, such deciders put complete faith in the information they have, and do not "risk" moving outside of what they *know*. These "ultra rationalists" we have called *Cartesians*. The movement of the boats around the system is generated by the difference at a given time between the number of boats that would like to be in the each zone, compared to the number that actually are there. As the boats congregate in particular locations of high catch, so they fish out the fish population that originally attracted them. They must then move on the next zone that attracts them, and in this way there is a continuing dynamic evolution of fish populations and of the pattern of fishing effort.

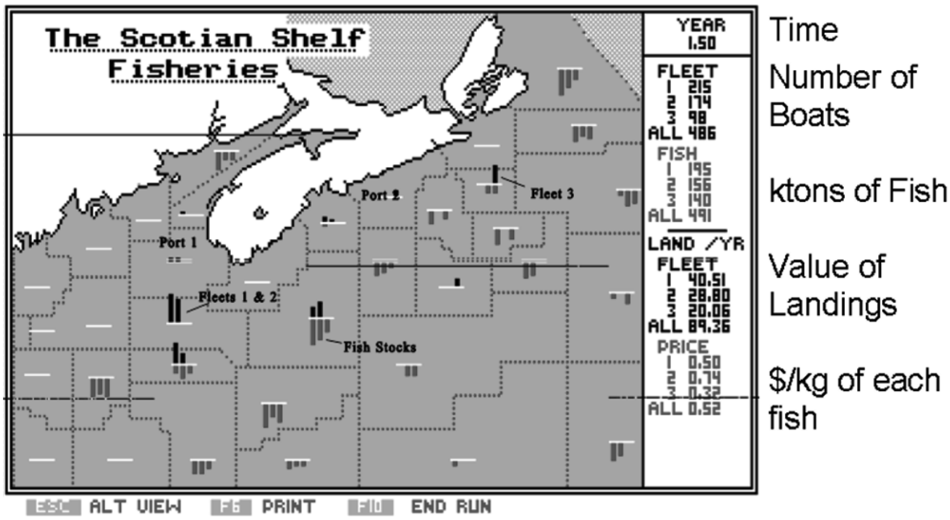


Figure 4. The dynamic, spatial model of learning fishing fleets, based on real data, showing that economically optimising fishermen are unsuccessful.

Running these fishery simulation models shows us that it is not true that fleets seeking profit with the highest possible economic rationality win. Indeed, the models show that it is important NOT to seek profit too ardently. Over time, higher profits are actually generated by behaviour that does not seek profit maximally at any moment!

This is because in order to fish for any length of time it will be necessary not only to exploit current information about fish stocks, but to discover new, currently unknown fish stocks, and exploit them. So exploitation alone is inadequate and some mix of exploration and exploitation is required. This tells us that the ability to "learn" is more important over time than the ability to "exploit" efficiently. Essentially it means that there has to be exploration and then a selection using the new information generated by the exploration. Successful management requires the generation of micro-diversity in the knowledge of the operatives in the "organization" – in this case a fishing fleet. It further requires decisions based on the positive pieces of information, while being careful to continue with exploratory behaviour. This model also shows us that there is no such thing as an Optimal Strategy. As soon as any particular strategy becomes dominant in the system, then it will always be vulnerable to the invasion of some other strategy. More importantly, we see that the knowledge generation of fleets arise from their ability and willingness to explore. So, instead of this corresponding to ultra efficiency and instrumental rationality, it actually arises from the opposite — a lower level of instrumental rationality, and a freedom to take creative action.

Another key result was that where fleets were able to evolve strategies concerning "leadership" — that is which fleets to follow, then instead of finding that the fleet with the most effective strategy was followed by a number of imitators, the results were completely inconclusive. Some fleets followed others with less effective strategies than themselves, and indeed, often the chain of followers looped round on itself, wiping out the distinction between followers and leaders. This is an important result which points to the idea that "leadership" is a process of continual construction and deconstruction, affected by chance as well as real qualities. As such it cannot serve to resurrect instrumental rationality management but will only form one factor among several in strategy formation.

6.2 *Emergent market structure*

The ideas developed in the sections above can also show us how important dynamic capabilities are for firms in the market place. Here we see how these dynamic capabilities are what is responsible for structuring of economic markets, as competition creates ecologies of firms, creating and filling interacting market niches. The fundamental process can be explored initially using a simple model in which we consider the possible growth/decline of several firms that are attempting to produce and sell goods on the same market. The potential customers of course will see the different products according to their particular desires and needs, and in the simple case examined here, we shall simply consider that customers are differentiated by their revenue, and therefore have different sensitivities to price.

This model has been discussed in Allen and Varga [2006], and so we shall not give much of the detail here. Inputs and labour are necessary for production and the cost of these added to the fixed and start-up costs, produce goods that are sold by sales staff who must "interact" with potential customers in order to turn them into actual customers. The potential market for a product is related to its qualities and price, and although in this simple case we have assumed that customers all like the same qualities, they have a different response to the price charged. The price charged is made up of the cost of production (variable cost) to which is added a mark-up. The mark-up needs to be such that it will turn out to cover the fixed and start-up costs as well as the sales staff wages. Depending on the quality and price, therefore, there are different sized potential markets coming from the different customer segments.

When customers buy a product, they cease to be potential customers for a time that is related to the lifetime of the product. For high quality goods this may be longer than for low quality, but of course, many goods are bought in order to follow fashion and style rather than through absolute necessity. Indeed, different strategies would be required depending on whether or not this is the case, and so this is one of the many explorations that can be made with the model.

The model calculates the relative attractivity of a product (of given quality and price) for a customer of a given type (poor, medium or rich) which calculates the "potential market" for each firm at each moment. Sales staff must interact with these potential customers in order to turn them into customers. When a sale is made, then the potential customer becomes a customer and disappears from the market for a time that depends on the product lifetime. The revenue from the sales of a firm is used to pay the fixed and variable costs of production, and any profit can be used either to increase production or to decrease the bank debt if there is any. In this way, the firm tries to finance its growth and to avoid going near its credit limit. If a firm exceeds its credit limit then it is declared bankrupt and is closed down.

A very important issue that arises in the modelling concerns the rationality of the managers involved in the new market. In traditional economic theories, firms might be expected to use the "profit" experienced as the driving force of their strategies. But the problem is that in launching a new product the initial situation will be one of investment, not profit. So, if the model regulates production on the basis of the (negative) profit currently experienced, then it will shut down. Clearly this would be a foolish model, and so a more realistic model would link production to the expected profit over some future time.But this would also be an oversimplification for two reasons. Firstly, the manager of firm A cannot calculate expected profits without knowing the pricing strategy of competing firms B,C D etc., and they in their turn cannot decide what their expected profits will be without knowing the strategy of firm A. Secondly, whatever firms A, B, C etc. do choose as a strategy, and what expectations they have, many of them clearly get it wrong because a large fraction in fact make losses and go bankrupt. This tells us that firms do not calculate expected profits instrumentally rationally at all — since as we see *this is impossible*. Instead, they simply believe in a strategy, and then discover heuristically whether or not it works as they thought. The market is therefore the arena for learning, in which beliefs are either reinforced or destroyed. A wise manager will be ready to modify his beliefs and strategy as rapidly as possible in order to avoid bankruptcy. Our model shows that it is the economies and diseconomies of production and distribution that will determine the number, size and scale of the niches that may be discovered.

We can use our model to explore the effect of different learning strategies of firms. The strategy space in view here is that of what % profit to charge and what quality of product to make. Obviously, a lower mark-up will provide larger market share, and lead to economies of scale, but these may not come soon enough to avoid bankruptcy. Alternatively, a firm can have a very high mark-up and make much profit on each unit it sells. However, its sales may be too small to allow survival.

Our model can allow us to study the effects of different learning strategies of managers. We distinguish four different types:

- **Darwinian Learning:** In this case we launch firms with random strategies, and if a firm goes bankrupt, we replace it with a new firm, with a random strategy. In this way, in theory the performance of the market will improve as unsuccessful strategies are eliminated, and only successful ones remain after a long time.

- **All Imitate:** Here, firms are launched initially with random strategies, but firms move to imitate whichever strategy is winning. In this way, in theory, the resulting market should evolve to a collection of firms all using a very successful strategy.

- **All Learn:** In this case, after being launched with random strategies, firms each explore the effects on profits of changing quality and mark up. They then move up the profit slope — if possible. In this way, they demonstrate the effect of a learning strategy.

- **Mixed Strategies:** Here, after a random launch, we have two Darwinists, two imitators and two learners. This leads to an evolution of the market structure gradually exploring and changing as profit and loss select for the winners and losers.

The model demonstrates the degree to which luck plays an important role, as for all these four different learning strategies, results depend considerably on the particular random strategies initially chosen and on the subsequent failures and random launchings that occurred. For different random sequences, very different market evolutions occur. If we use the model to calculate the total profits of all firms present, and allow for the cost of bankruptcies, then we can plot the performance of the emerging markets over time. The model shows us that for the Darwinian learning strategy, average market value achieved by the process, including the cost of bankruptcies is actually negative. In other words, Darwinism, applied to market evolution, using bankruptcy as the selection process is so costly on average that the industry as whole makes a loss. There is in fact enormous variance in the performance of the market as a whole, with some simulations showing very effective organization by chance, and others with very negative performance following a different sequence of choices.

The next block of simulations looks at the performance of the market when the players all imitate any winning strategy. This does perform better than Darwinian learning on average, with an average final value of over 800,000, compared to 114,000 for Darwinists. But the strategy seems to provide the most unstable trajectories, with some examples of market crashes and severe set-backs in the general pattern of improvement.

Figure (5). Simulation of markets with differing marketing strategies. Darwinian learning is a very ineffective strategy, the best overall performance being achieved by firms with learning mechanisms.

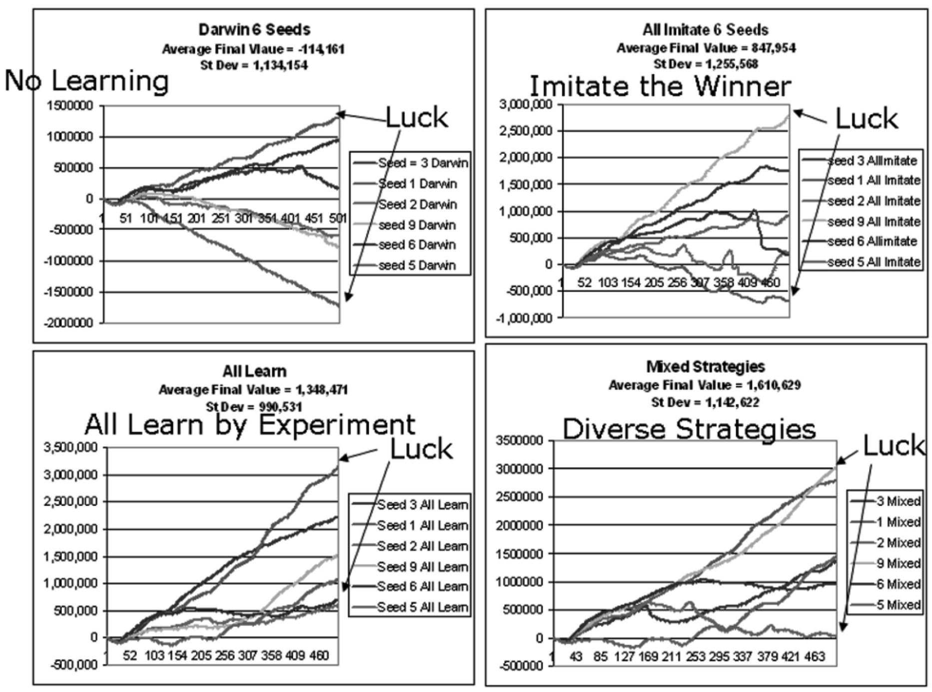
The most consistently successful performance corresponds to the “learning” strategy, where firms adopt an “experimental” strategy of probing their “profit” landscape and moving in the direction of increase. The fourth set of mixed strategies produces the largest average value of market performance, but has very wide variance

These simulations show us that organizational behaviours are linked through the market interactions in which they participate, and that the most behaviour of a firm is in fact that of "learning" what price and quality are successful in the market. In turn this requires that the firm can adapt its product and produce it competitively, and that it also conducts experiments with quality and mark-up in order to find out how to improve its performance within the market.

6.3 Evolution of manufacturing organisations

The study of organizational change and strategy can be looked at by reflecting on the organizations in terms of their constituent practices and techniques. The changing patterns of practices and routines that are observed in the evolution of firms and organisations can be studied using the ideas of Evolutionary Drive. We would see a "cladistic diagram" (a diagram showing evolutionary history) showing the history of successive new practices and innovative ideas in an economic sector. It would generate an evolutionary history of both artefacts and the organisational forms that underlie their production [McKelvey, 1982; 1994; McCarthy, 1995; McCarthy, *et al.*, 1997]. Let us consider manufacturing organisations in the automobile sector.

Table 2. Characteristics of manufacturing Organisations



With these characteristics (Table 2) as our "dictionary" we can also identify 16 distinct organisational forms:

- Ancient craft system
- Standardised craft system
- Modern craft system
- Neocraft system
- Flexible manufacturing
- Toyota production
- Lean producers
- Agile producers
- Just in time
- Intensive mass producers
- European mass producers
- Modern mass producers
- Pseudo lean producers
- Fordist mass producers
- Large scale producers
- Skilled large scale producers

Cladistic theory calculates backwards the most probable evolutionary sequence of events Figure (5). Again, in agreement with the ideas of Evolutionary Drive, we shall look at this as being the result of micro-explorations, and then a differential amplification of systems with emergent capabilities. We have studied the evolution of the automobile production industry by conducting a survey of manufacturers, and obtaining their estimates of the pair-wise interactions between each pair of practices. In this approach, the microscopic explorations consist in the attempts to connect in new practices to an existing system, with the object of improving performance and creating positive emergent capabilities. As has been reported before, we can understand and make retrospective sense of the evolution of the automobile industry.

We have then been able to develop an evolutionary simulation model, in which a manufacturing firm attempts to incorporate successive new practices at some characteristic rate. The “receptivity” of the existing complex determines which new practice will in fact be amplified or suppressed if it tries to “invade”.

Figure (6) shows us one possible history of a firm over the entire period of the development of automobile production. The particular choices of practices introduced and their timing allows us to assess how their performance evolved

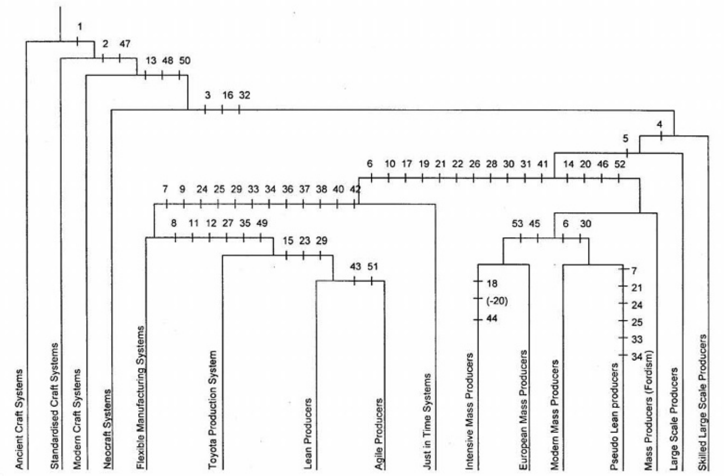


Figure 5. The cladistic diagram for automobile manufacturing organisational forms.

over time, and also assess whether they would have been eliminated by other firms. As a result of the different firms experimenting over time, there is an incredible range of possible structures that can emerge, depending simply on the order in which practices are tried. But, each time a new practice is adopted within an organisation it changes the "invadability" or "receptivity" of the organisation for any new innovations in the future. This illustrates the "path dependent evolution" that characterises organisational change, already visible in the luck-dependent strategies of the preceding market strategy simulations. Here successful evolution is about the "discovery" or "creation" of highly synergetic structures of interacting practices and their subtle organisational requirements mean that their emergence is highly sensitive to early events in their development.

The model starts off from a craft structure. New practices are launched with an "experimental" value of 5. Sometimes the behaviour declines and disappears, and sometimes it grows and becomes part of the "formal" structure that then changes which innovative behaviour can invade next. The model shows how the 16 different organizational forms have increasingly high synergy as they change in the direction of lean and agile Japanese practices. Overall performance is a function of the synergy of the practices that are tried successfully. The particular emergent attributes and capabilities of the organisation are a function of the particular combination of practices that constitute it. Different simulations lead to different structures, and there are a very large number of possible "histories". This demonstrates a key idea in complex systems thinking. The explorations/innovations that are tried out at a given time cannot be logically or rationally deduced because their overall effects cannot be known ahead of time. Therefore, the impossibility of prediction gives the system "choice".

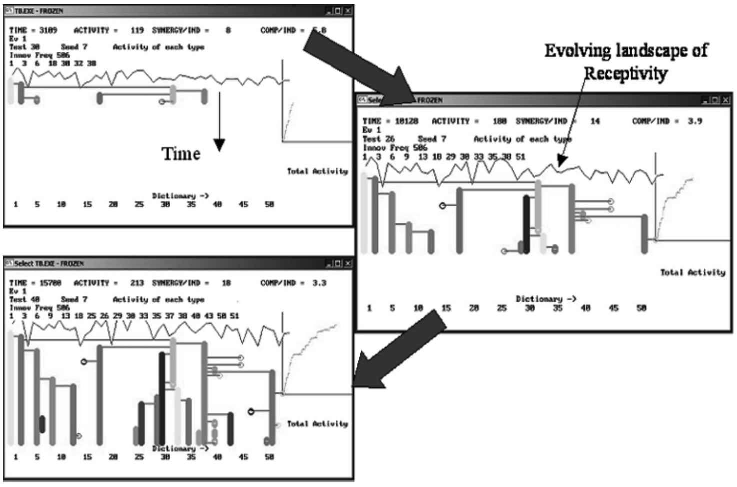


Figure 6. Successive moments (t=3000, 10000 and 15000) in the evolution of a particular firm. The evolutionary tree of the organisation emerges over time.

The competition between different firms' exploratory pathways through time means that those who for one reason or another fail to find synergetic combinations of practice, will be eliminated. Once again we find the principle of Evolutionary Drive, where the micro-explorations involving the testing of new practices leads to microscopic diversity among firms, and in turn these are either amplified or suppressed by the economic competition.

7 CONCLUSIONS

As with many concepts that change our view of what is 'normal', work leading to the realisation that determinism (things being causal, predictable and going to plan) is not the norm came from a number of directions. This started in the late nineteenth century with some American philosophers (James [1995]; Buchler [1955] on Peirce), a mathematician [Poincaré, 1890] and several physicists, through the advent of quantum physics. Cybernetics [Ashby, 1956], General Systems Theory [von Bertalanffy, 1968], Lorenz's [1963] mathematical exploration of weather patterns, Haken's [1977] work on synergetics, and Prigogine's non-equilibrium thermodynamics [Nicolis, Prigogine, 1989] built on this theme of uncertainty. They, variously, were able to develop the ideas further through the advantage of increased computer power which allowed the exploration of situations of interest through mathematical modelling.

Prigogine's work [1984], for which he was awarded the Nobel prize in chemistry in 1977, is key. He was intrigued by the mystery of evolution. How can it be, he asked, that evolution takes species and social systems into new and generally more sophisticated forms, when theories of physics seem to indicate something quite to the contrary? The answer was to be revealed by complexity science.

Complexity theory offers a 'scientific' theoretical lens through which to view many other perspectives, derived more empirically (see [Wheatley, 1992]). The ideas encompass the notion of learning and adapting, with strategy viewed as an emergent process developed by learning through experience. This so-called learning approach to strategy development encompasses a sense of incremental change however (see Quinn [1980]) and logical incrementalism). The underlying dynamic from the perspective of complexity theory may be much more sudden and dramatic if the situation in the environment is at a tipping point. So, the adaptation implied in such circumstance may be more radical and extreme than the notion of learning may imply. However the idea that implementation is in fact the only way of testing "strategy" speaks directly to a "pragmatist" point of view. In the real world almost everything in our social and economic systems are the result of pragmatic evolution — not of theoretical principles. The "creative destruction" of Schumpeter is indeed just a recognition of the fact that economic systems actually evolve in this pragmatic way. What complexity brings to management is the realization that successful evolution requires the facilitation of exploration and hence of internal heterogeneity and openness. In essence, if an organization is to persist over a longer time scale, success and specialization should not be allowed to fossilize its identity.

In a complex systems perspective there is perhaps more focus on the interplay between organisation and environment than in much of the literature on core competences [Hamel and Prahalad, 1989] although the notion of dynamic capabilities [Teece, *et al.*, 1997; Eisenhardt and Martin, 2000] do indeed recognise that the mechanisms through which firms accumulate and dissipate new skills and capabilities is the source of competitive advantage. Hamel and Prahalad [1989] also emphasise the concept of strategic intent. The need for this proactive orientation in the market place is embraced within a complexity perspective and we are at pains to emphasise this point.

From the discussions and models presented above we can derive some key points about organizational behaviour and its evolution. Organizational behaviour must be such as to allow organizational evolution, or the organization will fail. The rules that allow organizational evolution are:

- The presence of mechanisms that produce internal heterogeneity, which will involve freedom, ignorance and underlying error-making, exploratory processes

- Differential performance needs to be detected and evaluated with respect to their alignment with higher level goals. This will then provide the selection process that will amplify or suppress different elements of individual behaviour.

- The relative performance of the organization within the wider environment needs to be constantly reviewed and evaluated to see how the selection criteria are changing and how this may affect the capabilities and competences that the organization needs.

- For successful organizations, as Chia and Tsoukas [2002] point out, aggregate descriptions will always be short term emergent properties of an evolving system.

Successful management must behave as evolution does and make sure that mechanisms of exploration and experiment are present in the organization. Though they will not be profitable in the short term they are the only guarantee of survival into the longer term. In reality, the organizations that we observe and describe formally at any given moment are "structural attractors" [Allen, *et al.*, 2005; Strathern and Baldwin, 2005], which, if they persist over time, will change qualitatively as successive organizational forms emerge.

In hard science a new theory must be capable of being falsified, and therefore must produce testable predictions, which if not falsified lead to genuine cumulative knowledge. In complex systems however, clean predictions are no longer possible and the knowledge evolves under less severe selection criteria. In ecological and human systems emergent structural attractors can occur simply because their particular emergent capabilities/behaviour succeeds in getting resources from the environment. Fashion, lifestyles, art, artefacts and communities of practice can emerge and survive providing that there is a clientele for them [Arthur, 1994]. They are not about being "true or false" but simply about whether there is a "market" for them. Living systems create a world of connected, co-evolved, multi-level structures which may be temporally self-consistent, but will evolve and change over time. In our fishing example, the organizational behaviour of a successful fleet will be heterogeneous, having explorers and exploiters, who share information and allow the fleet to discover new fish concentrations, and therefore to have a future beyond current information.

In a market place, we find that our firms need to experiment with their strategy in order to find out how to improve profits in the moving constellation of other firms. Luck plays a role, but learning will do better than just hoping. Similarly in our study of automobile manufacturing, we can actually break down organizational behaviour into its "atomic" components of working practices, skills and techniques. Complexity tells us that as well as the "organization you see", the set of practices, there also has to be additional agents that decide to try out new practices, and which ones these should be. This is really the role of management, although hopefully it uses information from throughout the organization. However, the main point is that although a particular bundle of practices is what is observable at any given time, a successful evolution will require the additional presence of agents that suggest new practices, discuss how to bring them in and implement them, and then how to evaluate their performance.

Fundamentally, the development of complexity theory supports a pragmatic view of the world, since it points to the real limitations of theoretical understanding and prediction. Indeed it suggests that our own internal representations of the world are themselves under constant reconstruction in the light of our actions and experiences and as a result the world is also under constant evolution as a result of the internal evolution of its interacting inhabitants. It is the natural result of evolutionary drive — a multi-layered co-evolution of the different levels of description and interaction that emerge in open, non-linear systems.

This has important implications for policy and management for such complex systems. For simple predictable systems whose behavioural laws are independent of management intervention, and for systems that can for practical purposes be treated in this way, policies can aim to optimise some (independently specified) valuable condition or product output, and the managerial means for doing so can be deduced from the system laws, so that management can be focused on the optimally efficient means for achieving the policy goal, which is entirely an exercise in technically calculable rational efficiency. Call this the minimal, classic account of policy and management formation, employing classical instrumental rationality. But complex open systems deeply challenge this approach: (a) epistemologically, the system cannot be known in advance independently of managerial behaviour, so there are insufficient, accessible and predictable interventions to support the application of any simple optimisation and (b) rationally, there is thus insufficient basis for applying the classical instrumental optimising conception of rationality as the basis of a managerial approach. Instead, the requirement of optimal efficiency needs to be replaced by one of "sufficient efficiency" ('satisficing' — [Simon, 1957; 1978; 1983]) combined with the capacities to learn and adapt.

This process will select for the "ability to evolve" as well as for particular types of micro-diversity at a given time. This situation reinforces the earlier epistemic theme of our limited knowledge of our own systems. + Sea strategy (incl. leadership) and inst. rat. This tells us that firms do not calculate expected profits instrumentally rationally at all — since as we see this is impossible.

Instead of management being based on the idea of creating a set of maximally efficient operations that will produce some good or service for a particular market and organisation being designed to support that process, complexity recognizes the reality of openness, not only in change in both the supply and demand situations and the continual appearance of new ideas, technologies and competitors, but also in organisation itself in response to these as demands for learning and adapting arise and their focus continually shifts to track change. Call the organisational capacities for learning and adaptation, adaptability. This discussion of organisational dynamics reinforces the replacement of maximal efficiency with "sufficient efficiency" combined with the adaptability, but emphasises the significance of self-organised, organisational change in underwriting this process.

BIBLIOGRAPHY

[Allen, 1994] P. M. Allen. Evolutionary Complex Systems: Models of Technology Change. In L. Leydesdorff and P. van den Besselaar (eds.), *Chaos and Economic Theory*. Pinter, London, 1994.

[Allen, 2000] P. M. Allen. Knowledge, Ignorance and Learning. *Emergence*, 2(4), 78-103, 2000.

[Allen, 2001a] P. M. Allen. Knowledge, Ignorance and the Evolution of Complex Systems. In J. Foster and S. Metcalfe (eds.). *Frontiers of Evolutionary Economics: Competition, Self-Organisation and Innovation Policy*. Edward Elgar, Cheltenham, 2001.

[Allen, 2001b] P. M. Allen. A Complex Systems Approach to Learning, Adaptive Networks. *International Journal of Innovation Management*, 5(2), 149-180, 2001.

[Allen and McGlade, 1987a] P. M. Allen and J. M. McGlade. Evolutionary Drive: The Effect of Microscopic Diversity, Error Making & Noise. *Foundations of Physics*, 17(7), 723-728, 1987.

[Allen and McGlade, 1987b] P. M. Allen and J. M. McGlade. Modelling Complex Human Systems: A Fisheries Example. *European Journal of Operational Research*, 30, 147-167, 1987.

[Allen, *et al.*, 2005] P. M. Allen, M. Strathern, and J. S. Baldwin. Models of Self-organisation and Evolution in Socio-economic Systems. *European Journal of Economic and Social Systems.* "Self-organization in Economics" Guest Ed Bernard Paulré, 18(2), 171-199, 2005.

[Allen and Varga, 2006] P. M. Allen and L. Varga. Complexity: The Co-evolution of Epistemology, Axiology and Ontology. *Nonlinear Dynamics, Psychology, and Life Sciences*, 11(1), 19-50, 2006.

[Arthur, 1994] B. Arthur. *Increasing Returns and Path Dependence in the Economy.* University of Michigan Press, Ann Arbor, 1994.

[Ashby, 1956] W. R. Ashby. *An Introduction to Cybernetics.* Chapman and Hall, London, 1956.

[Beer and Nohria, 2000] M. Beer and N. Nohria (eds.). *Breaking the Code of Change.* Harvard Business School Press, Boston, MA, 2000.

[Brodbeck, 2002] P. W. Brodbeck. Implications for Organization Design: Teams as Pockets of Excellence. *Team Performance Management: an International Journal*, 8(1/2), 21–38, 2002.

[Brown and Eisenhardt, 1997] S. L. Brown and K. M. Eisenhardt. The Art of Continuous Change: Linking Complexity Theory and Timepaced Evolution in Relentlessly Shifting Organizations. *Administrative Science Quarterly*, 42, 1–34, 1997.

[Bryant, 1998] A. Bryant. Beyond BPR: Confronting the Organizational Legacy. *Management Decision*, 36(1), 25–30, 1998.

[Buchler, 1955] J. Buchler. *Philosophical Writings of Peirce.* Dover, New York, 1955.

[Burnes, 2004] B. Burnes. *Managing Change*, 4^{th} edn. FT/Prentice Hall, Harlow, 2004.

[Burnes, 2005] B. Burnes. Complexity Theories and Organizational Change. *International Journal of Management Reviews*, 7(2), 73-90, 2005.

[Clarke, 1999] M. Clarke. Management Development: a New Role in Social Change. *Management Decision*, 37(10), 767–777, 1999.

[Cooper and Jackson, 1997] C. L. Cooper and S. E. Jackson, eds. *Creating Tomorrow's Organizations Today: a Handbook for Future Research in Organizational Behavior.* Wiley, Chichester, 1997.

[Dawson, 2003] P. Dawson. *Reshaping Change: a Processual Perspective.* Routledge, London, 2003.

[Dunphy, *et al.*, 2003] D. Dunphy, A. Griffiths, and S. Benn. *Organizational Change for Corporate Sustainability.* Routledge, London, 2003.

[Eisenhardt and Martin, 2000] K. M. Eisenhardt and J. A. Martin. Dynamic Capabilities: What are They? *Strategic Management Journal*, 21(10/11), 1105-1121, 2000.

[Foster and Kaplan, 2001] R. Foster and S. Kaplan. *Creative Destruction.* Doubleday, New York, 2001.

[Gould and Vrba, 1982] S. J. Gould and E. S. Vrba. Exaptation — a Missing Term in the Science of Form. *Paleobiology*, 8(1), 4-15, 1982.

[Gould, 2002] S. J. Gould. *The Structure of Evolutionary Theory.* Belknap, Cambridge, MA, 2002.

[Greenwald, 1996] J. Greenwald. *Reinventing Sears.* Time, 23 December, 53–55, 1996.

[Haken, 1977] H. Haken. *Synergetics, An Introduction: Non-equilibrium Phase Transitions and Self-Organization in Physics, Chemistry and Biology.* Springer-Verlag, Berlin, 1977.

[Hamel and Prahalad, 1989] G. Hamel and C. K. Prahalad. Strategic Intent. *Harvard Business Review*, (May-June), 1989.

[Harung, *et al.*, 1999] H. S. Harung, D. P. Heaton, and C. N. Alexander. Evolution of Organizations in the New Millennium. *Leadership and Organization Development Journal*, 20(4), 198–207, 1999.

[Hirooka, 2003] M. Hirooka. Non-Linear Dynamism of Innovation and Knowledge Transfer. Paper presented at the ASEAT & I&IR Conference, Manchester, 2003.

[Huczynski and Buchanan, 2001] A. Huczynski and D. Buchanan. *Organizational Behaviour*, 4^{th} edn. FT/Prentice Hall, Harlow, 2001.

[James, 1995] W. James. *Selected Writings*. Everyman, London, 1995.

[Johnson and Scholes, 2002] G. Johnson and K. Scholes. *Exploring Corporate Strategy*, 6^{th} edn. FT/Prentice Hall, Harlow, 2002.

[Kanter, *et al.*, 1997] R. M. Kanter, J. Kao, and F. Wiersema, eds. *Innovation: Breakthrough Thinking at 3M DuPont, GE. Pfizer, and Rubbermaid*. Harper Business, New York, 1997.

[Kotter, 1996] J. P. Kotter. *Leading Change*. Harvard Business School Press, Boston, MA, 1996.

[Maguire, *et al.*, 2006] S. Maguire, B. McKelvey, L. Mirabeau, and N. Oztas. Complexity Science and Organization Studies. In S. Clegg, C. Hardy, W. Nord, T. Lawrence (eds.). *Handbook of Organization Studies*, 2^{nd} edn. London: Sage, 165-214, 2006.

[Maynard-Smith, 1979] J. Maynard-Smith. Game Theory and the Evolution of Behaviour. *Proceedings of the Royal Society*, London B, 205, 475-488, 1979.

[McCarthy, 1995] I. McCarthy. Manufacturing Classifications: Lessons from Organisational Systematics and Biological Taxonomy. *Journal of Manufacturing and Technology Management — Integrated Manufacturing systems*, 6(6), 37-49, 1995.

[McCarthy, *et al.*, 1997] I. McCarthy, M. Leseure, K. Ridgeway, and N. Fieller. Building a Manufacturing Cladogram. *International Journal of Technology Management*, 13(3), 2269-2296, 1997.

[McKelvey, 1982] B. McKelvey. Organizational Systematics. University of California Press, 1982.

[McKelvey, 1994] B. McKelvey. Evolution and Organizational Science. In J. Baum, J. Singh, *Evolutionary Dynamics of Organizations*, Oxford University press, Oxford, 314-326, 1994.

[Metcalfe, 1998] J. S. Metcalfe. Evolutionary Economics and Creative Destruction. Routledge, London, 1998.

[Metcalfe, 1999] J. S. Metcalfe. Restless Capitalism, Returns and Growth in Enterprise Economics. CRIC, Univ. of Manchester, Manchester, 1999.

[Peters, 1997] T. Peters. *The Circle of Innovation: You Can't Shrink Your Way to Greatness*. Knopf, New York, 1997.

[Poincaré, 1890] H. Poincaré. Sur le Problème des Trois Corps et les Équations de la Dynamique. *Acta Mathematica*, 13, 1-270, 1890.

[Prigogine and Stengers, 1984] I. Prigogine and I. Stengers. *Order out of Chaos: Man's new Dialogue with Nature*. Bantam, New York, 1984.

[Quinn, 1980] J. B. Quinn. *Strategies for Change: Logical Incrementalism*. Irwin, Homewood, IL, 1980.

[Romanelli and Tushman, 1994] E. Romanelli and M. L. Tushman. Organizational Transformation as Punctuated Equilibrium: an Empirical Test. *Academy of Management Journal*, 37(5), 1141–1166, 1994.

[Schumpeter, 1975] J. A. Schumpeter. *Capitalism, Socialism and Democracy*. Harper, New York, 1975. (Orig. pub. 1942.)

[Simon, 1957] H. A. Simon. *Models of Man: Social and Rational*. Wiley, New York, 1957.

[Simon, 1978] H. A. Simon. Rationality as a Process and Product of Thought. *American Economic Review*, 68, 1-16, 1978.

[Simon, 1983] H. A. Simon. *Reason in Human Affairs*. Stanford University Press, Stanford, 1983.

[Stickland, 1998] F. Stickland. *The Dynamics of Change: Insights into Organisational Transition from the Natural World*. Routledge, London, 1998.

[Styhre, 2002] A. Styhre. Non-linear Change in Organizations: Organization Change Management Informed by Complexity Theory. *Leadership and Organization Development Journal*, 23(6), 343–351, 2002.

[Teece, 1997] D. J. Teece, G. Pisano, A. Shuen. Dynamic Capabilities and Strategic Management. *Strategic Management Journal*,18(7), 509-533, 1997.

[Tsoukas, 002] H. Tsoukas, R. Chia. On Organizational Becoming: Rethinking Organizational Change. *Organization Science*, 13(5), 567-582, 2002.
[Von Bertalanffy, 1968] L. Von Bertalanffy. *General Systems Theory*. Braziller, New York, 1968.
[Wheatley, 1992] M. J. Wheatley. *Leadership and the New Science*. Berrett-Koehler, San Francisco, 1992.
[Zairi, *et al.*, 1994] M. Zairi, S. Letza, and J. Oakland. Does TQM Impact on Bottom Line Results? *TQM Magazine*, 6(1), 38–43, 1994.

COMPLEX SYSTEMS DYNAMICS AND SUSTAINABILITY: CONCEPTION, METHOD AND POLICY

Thomas S. Brinsmead and Cliff Hooker

What complex systems dynamics makes necessary, and possible, is a profound change in the structure of policy making. This revolution is clearly revealed in sustainability policy, where it is still coming to fruition. This chapter explores the implications of a complex dynamic systems perspective for the adequate conception of sustainability and satisfactory sustainability policy. The essence of sustainability policy is the appropriate management of human interaction with the natural world, with a particular emphasis on interactions that involve risks to valuable, especially life-supporting, environmental processes. Early practical attempts, before the development of the scientific discipline of ecology, to deal with issues that we now refer to as sustainability exploited only relatively crude models of the interaction between humans and their natural environment, including linear correlational, static nonlinear, linear dynamical, and nonlinear stable equilibrium and periodic models. More recently however, the rigorous representation and precise analysis of a much richer class of dynamical systems than previously — in particular nonlinear aperiodic dynamical models, has had a significant influence on the conception of adequate sustainability policy, though not yet on its formal management in practice.

In broad terms, the general aim of sustainability policy is to ensure that valued processes and conditions are preserved and/or enhanced over time. Both human and environmental processes and conditions are included, they may be valued on ecological or more pragmatic grounds, and the preservation/enhancement is to occur in the presence of ongoing human-environment interaction. Because of multiple causal interactions among human and environmental subsystems, sustainability policy (and its antecedents) has always required organisational mechanisms for collectively co-ordinated action. However earlier formal approaches to "environmental management" were based on cause-and-effect models that in retrospect can be recognised as a narrowly limited class, though nevertheless previously successful in supporting significant advancements in ecological protection and some components of human wellbeing. This chapter briefly reviews limitations of using dynamical models belonging to these earlier limited classes — including those that are restricted to being linear proportional, unidirectional causation, close to stable equilibria, homogenously constituted, and/or completely deterministic — in the light of the appreciation of the ubiquity of other, more complex, dynamical behaviours in natural and human systems.

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

Important enhancements to sustainability conception and policy arise from a greater appreciation of the significance of dynamical behavioural complexity (that cannot be obviously simplified by, for example, statistical treatment) due to networked non-linear interactions among many heterogeneous dynamical components. The hitherto dominant conceptions of sustainability — i) remove negative environmental impacts and ii) maximise human plus natural capital — are challenged by a radically different approach: iii) sustainability as maintenance or enhancement of adaptive resilience, the capacity to robustly preserve continued functioning through short term perturbations and long term change. This third concept is arguably more fundamental than the others for several reasons. (a) It prescribes sustaining what is essential to continued system existence in a highly dynamical, changing world of evolution-development. (b) It (i) leads to better outcomes than the impact approach because it explicitly recognizes, within its models, long term dynamical interaction between human and environmental processes and thereby offers synergies between sustaining ecology and human society and (ii) is more practical than the capital notion because it does not require trying to predict the long term future. In fact (c) it is their natural successor in policy strategy because it alone incorporates fundamental uncertainty into the dynamical models and decision strategies.

Many dynamically behaviourally complex systems possess several equilibria and so path dependence is ubiquitous, for example because often important global dynamical constraints have their origin in the amplification of relatively small past fluctuations. This and other aspects of dynamical behavioural complexity result in significant limitations on the practical feasibility of precise prediction over time (even where precise, local causal explanation could be provided by retrospective dynamical analysis).

Prediction limitation has several significant consequences for sustainability policy. Firstly, it implies that management for sustainability should be an adaptive interactive process, contingent on feedback, rather than an unconditional commitment to specific action. It also implies that rather than being able to be treated as a deterministic optimisation, sustainability policy becomes a problem of risk management under uncertainty, and thus must be directed to the less ambitious goal of enhancing the resilient achievement of some acceptable condition in the face of ignorance. More subtly, limitations on state prediction precision suggest that sustainability policy acceptability conditions should be framed in terms of broad constraints on dynamically relevant system organisation, rather than precise specific targets for subsystem states. Furthermore, path dependence and prediction limitations point towards backcasting as a promising framework for sustainability policy.

The limitations on predictability also suggest that responsible sustainability policy will be a co-evolutionary learning process, not only at tactical scales among intention, prediction and action, but also at strategic scales between theory and observation, and between values, which direct policy goals, and descriptive understanding. For example, dry Australian grasslands have organisational thresholds for species diversity and productivity that have to be discovered, as do behavioural sensitivities near thresholds; further, thresholds themselves can shift as a consequence of unusual event sequences, so that the formulation of resiliency policy must develop along with emerging knowledge of system thresholds and their dynamics.[1] Inherent limitations on dynamical modelling, prediction and normative development imply that a portfolio of precautionary do-a-little-to-learn activities is as important a component of sustainability strategy as is activity directed at maintaining current substantively valued conditions and processes. The dynamics of these learning processes are also characterised by strong positive feedback loops and path-dependent sensitivity to initial conditions.

Furthermore, and finally, achieving the adequate management of risk in behaviourally complex systems implies a requirement for a robust appreciation of the significance of values implicit in both learning and other methodological norms, in addition to the traditionally recognised significance of values in the determination of substantive policy goals. The explicit appreciation of relevant values assists in constructing appropriately simplified representations of the relevant dynamics of a given problem situation, which is a significant sub-problem within sustainability policy, and backcasting in particular.

1 INTRODUCTION: SUSTAINABILITY

The main theme of this book is the nature and importance of the impact of the increasing understanding of complex dynamical behaviours on pure and applied science. From a cultural and historical perspective, our increasing ability to represent and analyse the dynamics of increasingly wider classes of causal dynamical systems has led to profound changes in our appreciation of the types of dynamical behaviours that are empirically possible. This has led in turn to similarly profound changes in the conception of frameworks considered appropriate for individual and social human interaction with the wider world. In particular, within the class of dynamical systems models described by deterministic flows, the focus here is on the transition in understanding from that of the classes described by linear statistical correlation, linear dynamical, and nonlinear periodic systems, to that of nonlinear aperiodic systems and systems with a large range of qualitative dynamical states. This chapter describes how this general theme applies in the specific case of sustainable *collective*, ecological human material and cultural development. It focuses on interaction with the wider environment, as humanity's spheres of control, influence, and potential influence have expanded to cover wider and more complex, previously natural — now partially and increasingly artefactual — environmental subsystems.

Survival and maintenance of human society, and support for sufficient human wellbeing to maintain social coherence, has always been at least an implicit goal of almost all successfully surviving human cultures. Historically, a large proportion

[1] For further discussion see e.g. [Walker, *et al.*, 2004; Gunderson and Holling, 2002].

of the population has been required to interact directly with the natural (primarily ecological, including agricultural) environment (that is, in addition to interaction with other members in society) in order to meet their individual needs and wants. However, individual human interaction with the environment invariably also has indirect consequences that are of relevance to general social survival and individual wellbeing. Examples include simple competition for finite resource stocks (such as fresh water), and otherwise desirable activities that result in harm to the productivity of a resource renewal process (such as toxic pollutants from mineral extraction that reduce the fertility of game fish). Each culture has developed social (including ethical and legal) norms that constrain what are considered to be appropriate environmental interactions. These norms maintain processes in society's immediate environment including both those that support human progress, or at least the maintenance of wellbeing, and those environmental processes that are culturally valued for their own sake.

In contemporary parlance, improvement in the general condition of humankind is called "development", and the term "sustainability" is used to mean the condition of maintaining indefinitely into the future both human development and valued environmental processes, especially those that support human development. But, lest this formulation smack too strongly of anthropocentrism, we shall see that managing for sustainable resilience provides for sufficient autonomy to ecological dynamics as to provide some inherent, but well informed, ethical obligation to ecological care.

1.1 Sustainability: Relevance and Realisation

Human wellbeing and development has been historically maintained in large part by the interaction with, including manipulation of, natural sub-systems to serve human needs and desires. Examples include hunting and foraging, and later agriculture, to serve nutritional requirements, exploitation of forests for energy, construction materials and paper, and the construction of modern vast systems for water harvesting, treatment and reticulation for urban settlements. A significant proportion of such interaction must be co-ordinated for satisfactory or preferred outcomes to result. These include the avoidance of overexploitation and cross-activity interference, achieving communal agreement on appropriate trade-offs between environmental protection standards and individual human convenience, and the implementation of large scale projects. Such co-ordination requires the existence of organisational mechanisms, which are invariably realised via social interaction. The capacity to organise human activity through social interaction to meet many human needs is one function of culture (including knowledge of both social and natural worlds), whose maintenance is served by societal education.

The simplest cultural mechanisms for regulating human interaction with the natural environment will take the form of culturally stabilised social norms prescribing or proscribing particular types of activity (such as a sake-making ritual involving the requirement to sing particular songs, or the forbidding of beginning the salmon fishing season until the first salmon caught has been ritually dried). They will not necessarily be associated with a detailed articulation of the accompanying aims (ensuring the correct brewing process is followed, or maintenance of a viable salmon population). Nor may they be associated with an understanding that, and if so how in detail, following the given social norms will promote the achievement of relevant aims (the songs regulate the timing for various events in the brewing process; the period of several days required for the salmon to dry allows the fastest and fittest salmon to swim upstream to spawn, thereby ensuring a high quality stock for the following year).

Nor will the simplest mechanisms be capable of sophisticated contingency alternatives within the normative prescription, including an understanding of the contingent conditions under which they ought not be applied (broadcasting the sake-brewing songs from a compact disk recording is appropriate; the use of the latest technology such as a microwave oven, which speeds up the salmon drying process, is not appropriate). Such relatively simple (i.e. inflexibly prescriptive) regulatory (social governance) structures may be contrasted with, for example, more sophisticated modern engineering projects where the rationale for undertaking various construction and operations tasks are informed by a detailed empirical understanding. Such understanding might include site geology, the physics, chemistry and biology of relevant engineered processes, and some understanding of ecological and other environmental consequences, in addition to economic cost, profit, human health and safety and other human factors considerations. Our main concern in this chapter is how, as human (scientific) understanding of these relevant material processes has progressed in sophistication, the conception of what satisfactory sustainability policy entails has also become increasingly sophisticated.

The sustainability of the potential for human development, and environmental processes that support that potential, is not always required to be paid special attention. For some natural subsystems that contribute to human wellbeing, their human relevant conditions are continually renewed by an external energetic throughput (evapo-transpiration and subsequent precipitation (rainfall) for water purification, e.g., is driven by solar energy flows). For such subsystems, sustainability is realised automatically without requiring further human management, either deliberate or unconscious. For other natural subsystems, the human relevant aspects of their natural dynamics can be sustained with fairly simple maintenance interactions (e.g., composting for nutrient recycling or mulching for water storage in gardens). Thus sustainability of these conditions is not particularly complicated and may be maintained by the existence of relatively simple (e.g. prescriptive) management rules (although it may require significant effort).

However, there are also many interactions between humans via multiple indirect interactions through environmental sub-systems that affect their opportunities for improving their individual wellbeing. For example, small garden plots can be damaged through trampling or foraging by cattle or other livestock wandering in from grazing in nearby unfenced areas and sources of clean water can easily become affected by pollution in the catchment area.[2] This gives rise to a requirement for social coordination (norms governing the responsibility for livestock and fencing, constraints on pollution emissions in water catchments and/or the construction of large scale water treatment plants) in order to prevent significant harm to the interests of human society at large via the undesirable effects on the environmental subsystems.

On the other hand, human interaction with natural subsystems also results in some of them being irreversibly altered in relevant respects so that historically rewarding types of human interaction become no longer productive (excessive overgrazing of pasture by livestock can permit soil erosion and permanent loss of nutrients). Furthermore, some such changes in the dynamics of natural subsystems may also occur irrespective of human influence (for instance, changing climactic conditions such as the "Little Ice Age" after medieval times). In either case, whether or not changes in the dynamics of natural subsystems have been anthropogenic, the maintenance of human wellbeing and development has historically also ultimately been achieved by the development of new technological capabilities that allow humans to manipulate new subsystems or new niches (e.g. the introduction and spread of new crops suited to colder climates, such as the potato; the development of the internal combustion engine allowing the exploitation of the chemical energy in oil). It is worth observing that in many cases the manipulation of the new niche does not necessarily yield human value directly through a simple binary interaction (as in the ability to liberate nuclear energy, or to use radio waves to dramatically improve communication range and speed), but often (eventually) does so through the modulation or adaptation of traditional existing interactions (as in breeding techniques that develop new herbicide-resistant strains of canola, resulting in the modulation of agricultural practices, or the ubiquity of mobile phones that has significantly altered the way that social arrangements are planned).

Thus, as human technical innovation occurs (or indeed as natural environmental dynamics autonomously change in relevant respects, that is as natural dynamics innovate) there is always the risk that it may become increasingly difficult, or even impossible, continue to maintain and develop human wellbeing, and/or valued environmental processes (for instance, consider the threat of natural resource shortages, or the potential for gene modification technology to disrupt agricultural productivity). This risk increases with the abilities of technologically equipped humans to create larger environmental disturbances via the manipulation of increasing quantities of energy and via intervention in previously inaccessible subsystems (the threat of climate change is testament to human capacity to alter the global climate). Because of this risk it appears to have become increasingly necessary to explicitly develop sustainability strategies in order to manage the risk of the undesirable consequences of technological exploitation, arguably including the risk to the very survival of the human species (not to mention that of various diverse ecosystem processes). It appears to have become correspondingly increasingly possible to manage such risk because experience with technological innovation has

[2]See Chapter 2 in particular, but also Chapter 8, of [Leahy, Undated].

resulted in the development of increasingly sophisticated risk management techniques. Risk management is now a standard component of professional training and there exist widely recognised, explicitly documented risk management standards (e.g. Australian Standards 4360).

1.2 Explicit Sustainability Management: Early conceptions

We trace, in broad outline, the development in sophistication of the regulation of human-environment interaction, the antecedents of modern "sustainability policy". We note first the simplicity of the earliest heuristic environmental interaction practices that take the form of ritual and taboos and that are justified explicitly primarily for their own sake as norms of human social or cultural (including religious) significance, for instance, fertile crescent religious restrictions on eating pork.[3] Slightly more sophisticated are those social policies that are explicitly grounded in the anticipation of environmental sustainability, whether or not there is an associated anticipation that this will contribute to human welfare sustainability; for example, preserving a small stream out of deference to the associated water spirit may not be intended to serve any direct human interest, in contrast to doing so in order to preserve a source of water for livestock. The increased sophistication is a consequence of the necessity of an associated, at least implicit, assumed causal relationship between the policies and some particular desirable environmental conditions.

At its most basic, such an assumed relationship may be no more than a tendency towards correlation, an assumption perhaps based on past observations of similar correlations (e.g. of Australian aboriginals noticing bushfires followed by the return of breeding marsupials, or the observation that smelly, swampy areas are associated with a higher incidence of fevers). However, as human understanding of material dynamics becomes increasingly sophisticated the middle-order conceptions (sub-paradigms) of sustainability are transformed as the dynamical framework sub-paradigms informing sustainability policy are themselves transformed. Australian aboriginals come to understand that burning produces new growth which is preferred food for breeding marsupials and thereafter burn off locally in systematic patterns to maximise marsupial diversity and productivity ('fire stick farming')[4]. More explicitly still, noticing a swamp/fever correlation and identifying the swamp disease as malaria might lead to investigation of a causal link, leading to the discovery of the connection to the mosquito as parasite vector, leading in turn to a shift from eliminating malodorous air to eliminating mosquitoes as an environmental standard.

This chapter is a story of the sub-paradigmatic transformation in the conception of sustainability that arises as a consequence of the appreciation of the importance of dynamical complexity in both environmental (especially ecological) and human

[3] Or New Guinea religious rituals that maintain socially useful communities of biota, see [Rappaport, 1967].

[4] See e.g. [Flannery, 2002].

(especially social) systems. We trace this transformation quasi-historically.

While understanding complex dynamics has improved very rapidly over the past 50 years, it yet remains primitive compared to what is clearly required for rigorous sustainability policy, and this deficiency may yet prove inherent and thus irreducible. Indeed, as is evident on consideration of the other chapters in this book, we do not yet have a precise and widely agreed upon definition of complex dynamic systems, nor a useful taxonomical classification. However, for the purposes of discussion in this chapter, we distinguish simple models that are effectively static from those that explicitly represent the time dimension in a dynamical model, and we distinguish simple dynamical models which can be sufficiently accurately represented and controlled as deterministic from those which cannot.

2 STAGE I: REDUCE ENVIRONMENTAL BURDEN — LINEAR STATIC MODELS

The contemporary use of the concept of sustainability first received widespread international recognition as part of the phrase "sustainable development" in the well known Brundtland report to the United Nations on the environment and development.[5] At the time, an arguably dominant conception of sustainability was one based on a broad goal of environmental or ecological sustainability via eliminating negative impacts of human activities — most obviously toxic wastes — upon the wider environment. The central idea is that this would leave the environment pristine and so automatically preserved, thus achieving sustainability.

At their least sophisticated, the causal models underlying the corresponding strategies are unidirectional and usually implicitly non-interactive, although there may be a number of links in the causal chain. A typical causal chain is from some human activity to an environmental burden and from the environmental burden to some undesirable environmental or ecological (negative) impact. Consider for example, the human activity of mosquito elimination by the use of insecticides, resulting in a DDT load as the environmental burden, and increased infant mortality of wedge-tailed eagles as an undesirable impact. Such a causal chain model admits corresponding alternative strategies involving interventions in one or more alternative locations along the chain (eliminating the mosquito eradication programme, using alternative pesticides, implementing a wedge-tail eagle breeding programme). However, since the causal relationship between the burden and impact is often mediated by complexly interrelated non-human (e.g. climactic, ecological, biological) processes, it is generally less well understood than the relationship between the human activity and burden, a process which is often more accessible to human investigation and understanding, particularly because it is sometimes due to the implementation of a human technological design. It follows that effective intervention further down the causal chain (increasing wedge-tail eagle fertility rather than using alternative pesticides) is typically less achievable.

[5] [Brundtland, 1987]

This leads to the meta-strategic principle of preferring, if possible, the elimination or modification of the process that results in the burden — for instance, by capturing and processing human sewage for end-products usable within the human system – rather than end-of-pipe solutions that attempt to merely mitigate the negative impact by the treating the burden itself — notoriously, by diluting it in water, as was done with human sewage waste until relatively recently, and often still is the preferred, because simplest and cheapest, method. This type of unidirectional causal model is arguably what underpins and justifies the least sophisticated of the culturally mediated process for regulating human-environment interaction, the proscription of certain types of environmentally harmful activities (the release of pollutants) and the prescription of environmentally beneficial ones (planting trees and regenerating indigenous vegetation).

An only slightly more sophisticated human-environment regulatory process might involve some attempt to quantify the magnitude of the burden and significance of the impact. This enables a more sophisticated decision regarding the extent to which modifying the burden-producing human activity (which is valued, perhaps on some socio-economic and/or socio-cultural grounds) is worth the corresponding reduction in negative impact. On occasion it may be decided that the (human) costs of curbing the burden-producing activity make it not worthwhile overall to do so, and/or that it is acceptable to continue the activity provided the resulting burden is maintained below a specified sufficiently "safe" quantitative threshold. Water disposal of human sewage well illustrates both of these aspects, where thresholds concern disease vectors such as hepatitis, eutrophication rates and the like, although recently the economic value of processed by-products has risen to the point where industrial processing is the preferred option.

We claim that a complete framework for sustainability, if it is to be intellectually coherent and practically manageable, must be associated, at least implicitly, with the following common abstract elements: a normative ideal, a theory of the relationships between alternative courses of action and the normative evaluations of their corresponding consequences, and a corresponding collection of strategic advisory prescriptions regarding appropriate human behaviour in order to promote the realisation of that normative ideal.

Using this framework we can summarise this "remove negative impacts" conception of sustainability as follows. The implicit normative ideal is that of a natural ecology and other natural physical processes unspoiled by human interference. While the ultimate rationale for concern with sustainability is a desire to eliminate environmental negative impacts, the measurement focus is typically on environmental burden. The analysis tools used, which represent the understanding of the relevant causal relationships, such as life cycle analysis and resource demand assessment (including footprint analysis), tend to be concerned more with the representation of the human technological processes of burden production (such as tonnes of carbon dioxide production from coal combustion for electricity generation) rather than the dynamics of the natural, non-human, external environment (the global carbon cycle, atmospheric chemistry and thermodynamics). This is because the presumption is that most valuable (less complicated, more reliable) strategies will involve modification of the human burden production process to remove significant impacts.

2.1 Deficiencies of Stage I — from static to dynamical

One of the deficiencies of the conception of sustainability as the removal of negative impacts is that it does not have an explicit positive conception of an ideal. In particular, it has no specific normative vision for human development — nor does it have an especially detailed specific normative vision for the natural environment, only that it ought to be in its "natural" condition.[6] In one sense, this conception makes ecological norms primary — in the sense that the development of human wellbeing is to be tolerated only provided it unfolds in accordance with limited interference with, and limited negative impact upon, natural ecologies. However, since the emphasis is on the protection of the natural environment from the primarily destructive influence of human society, there is no clear conception of what might be understood by the concept of a flourishing ecology, nor of how ecological flourishing might be promoted. Moreover, as soon as it is realised that human impacts on the environment cannot be eliminated — natural species do not achieve that either, only ecological adaptation to, and incorporation of, their impacts — and moreover that many of the largest contributions to human welfare (and sometimes also to environmental welfare) stem from environmental interventions, this approach falls into either incoherence or requires a severe pre-historic utopian naturalism that few humans would accept.

However from the standpoint of this book's theme, the most significant aspect of this conception is not its normative character. Rather it is that, in principle, it does not require a particularly sophisticated dynamical cause and effect model. This is because the typical recommended strategy is for humans to avoid or modify particular activities in order to prevent placing a burden upon the wider environment. The most useful knowledge is the knowledge that particular types of human activities result in a burden and that a burden results in negative impact. A detailed understanding of the causal mechanism is unnecessary.

In fact, within a sufficiently narrow scope of analysis (for example, human engineered individual processes only, rather than human-environmental interaction) unidirectional causation models are sufficiently adequate. (This is even more likely to be validly accurate when there is in fact no significant environmental impacts that might result in feedback, such as the disposal of treated sewage from an ocean outfall.) At sufficiently large scale (economy rather than individual engineering

[6] Although not logically required, the natural purity ideal generalises to humans themselves being in an ideal "natural" condition, that is, unspoiled by the excesses of technology, and living in harmony with their natural ecological environment. Movements/communities with some such goal regularly recur. While this is a possible human ideal, we do not consider it further here because it is essentially an abrogation of any development policy and hence of sustainable development, our subject here, and ultimately because it is both impracticable and arguably unethical.

process wide) sufficient accuracy can be achieved using models that are linear proportional in burden to production relation, e.g. Life Cycle Analysis and Input-Output resource models. A complex dynamical model is not necessary because the strategic response to a perceived impact is relatively crude (stop this type of production — over-fishing — or substitute a lower impact producing activity — lower greenhouse emissions wind turbines rather than gas turbines for electricity generation) rather than a more subtle intervention in human socio-technical or engineering process dynamics (introduction of fishing licences or regulation of fishing methods, designing high efficiency gas combustion methods) or in the dynamics of wider environmental processes (for example fertilisation of the Southern ocean with the limiting nutrient of iron in order to stimulate carbon sequestration by phytoplankton[7], or the deliberate introduction of sulphate aerosols into the atmosphere in order to increase albedo and mitigate the enhanced greenhouse effect). Conversely, because only such a simple dynamical framework is used to represent the situation, a more sophisticated strategy is not possible.

3 STAGE II: OPTIMAL ENVIRONMENTAL EXPLOITATION TRAJECTORIES — FROM STATIC TO DETERMINISTIC DYNAMICAL MODELS

This brings us to consider those sustainability strategies based on the optimisation of a trajectory within a dynamical model, including many of the ecological and environmental economics conceptions of appropriate environmental management. These are distinct from the simple "remove negative impact" in that both the normative calculus and the relevant causal models employed are somewhat more sophisticated than those associated with the environmental burden reduction strategies described above.

The normative underpinning of such approaches is often based on human wellbeing as primary, and the value of preserving particular environmental and ecological conditions is derived from their contribution to human welfare. This approach makes environmental norms secondary — they are justified only provided they are derived from their contribution to satisfying human norms and development. Thus the natural environment is seen, not as valuable for its own sake — inherently, intrinsically — but primarily as a provider of ecological services (water filtration, carbon dioxide recycling and oxygen production, timber, fisheries, photosynthesis, mineral resources). In fact under this conception, desirable natural conditions are seen as a type of "capital": natural capital, a category in addition to traditional "built" or "manufactured" capital.[8]

[7] [Smetacek and Naqvi, 2008]

[8] Yet other categories of capital that have been added are "human" capital — skills and education, and "social" capital — the relationship ties that enable social networks to operate efficiently, including to cooperate. In this setting the nearest we can come to an inherent value being allocated to environmental conditions are cases where human wellbeing includes appreciation of the existence of natural ecosystems for their own sake. But this in no wise guarantees human interests

Because the natural environment is regarded as providing services and to be exploited, rather than protected and not interfered with, this implies a requirement to understand its dynamics. Under this conception of sustainability, in order to sustain human wellbeing the appropriate management of "natural" capital requires the careful management of trade-offs between the over-exploitation and the over-protection of the environment. Over-exploitation would result in its premature degradation, and its over-protection would result in rates of extraction that were too slight to best serve human needs, either of which results in less than optimum outcomes for human wellbeing. That this can require sophisticated dynamical models of human-environment interaction extending over long periods was beautifully demonstrated early on in the work of Richard Levins and others on North American fresh water fisheries. There a simple policy of restoring game fish populations annually by fingerling seeding — a policy that the 'remove impacts' approach might have favoured — was shown in fact to promote fishery collapse (through over-grazing at times of natural plant population lows in the marine fish-plant cycle) and that maximizing sustainable yield was instead a multi-oscillation policy that required accepting to some degree the natural fish-plant oscillations and intervening only at the right phases of the process. Similar dynamical outcomes now apply for many of the ecosystem management problems since investigated.[9]

In its standard formulation, the capital approach deploys the same economic decision making apparatus that is used for manufactured capital, namely to consider a suite of appropriate alternative decision strategies and determine the optimal one. Each appropriate strategy would stretch over a sufficiently long time to capture the major system dynamics involved (e.g. natural fish-plant population cycles) and apply a collection of intervention decisions that are either fully pre-determined or are contingent in pre-determined ways on the current state of the (human society and) environment, taken in the light of the preceding history and the dynamics at play. In the case of the North American fresh water fisheries, a strategy might provide specific pre-determined decision criteria for determining the number of fishing licences issued, the size of allowable catches and the number of fingerlings released, in any given year, as a function of location in the natural fish-plant population cycles at work. For a given strategy, at each time interval its present value is determined by discounting its anticipated delivered value at that time (e.g. fish caught) back to the present; the Net Present Value [NPV] of the strategy is then the sum over time of these discounted interval values. The optimal strategy is then that with the greatest NPV.

In terms of the framework for sustainability identified above, the normative ideal implicit in the "Natural Capital" conception of sustainability is that of max-

---

that coincide with the preservation of environmental process resilience at any time scale — communal, ecological or evolutionary — and in practice economic capital considerations have always dominated these approaches. See the broad review of positions at `http://www.gnudung.com/` and e.g. [Dasgupta and Serageldin, 2000; Adler and Kwon, 2002].

[9] See e.g. Allen herein; [Levins, 1968; Holling, *et al.*,1995; Gunderson, Holling, 2002, note 1; Walker, *et al.*, 2002] . More generally see `http://www.beijer.kva.se/`, and `http://rs.resalliance.org/tag/resilience-2008/`.

imising human wellbeing across time. While this provides the ultimate rationale for concern with sustainability, the measurement focus includes those states of the natural environmental that maintain its capacity for the provision of ongoing services useful to humans. Thus, the theoretical models used such as those representing the dynamics of ecosystem population, resource extraction, and pollution sinks tend to represent the natural environment, while the human component is represented in much less detail — in the utility (i.e. human value) of the production yield (e.g. fish caught) and in one or more choice variables (e.g. fingerling seeding times). More detailed understanding of the human system is typically not represented, because it is typically assumed that the human decision makers are "rational" and so will always choose the utility (e.g. NPV) maximising strategy. Nevertheless, the normative evaluation is slightly more sophisticated than that corresponding to the quantification of burden reduction in that it is required to be commensurate across service type (characterised by marginal rates of substitution) and time (typically characterised by discount factor).

The efficient management of natural capital is more general than the elimination of negative environmental impact, and reduces to it in particular circumstances. These circumstances may include where the condition that is being prospectively impacted upon negatively is crucial for human survival (so its values dominates), or where non-destructive enjoyment of the ecology is more highly valued than its destructive exploitation, or for temporary periods where it is preferable to invest in the development of natural capital in the present in order to enjoy improved ecosystem services in the future. In this sense, natural capital is a natural generalisation of the impact approach as the class of dynamical models widens.

In comparison to the "Reduce environmental burden" conception of sustainability, the "Natural Capital" conception provides both a richer and more flexible (in principle) normative structure and a wider systems perspective which tends to require a more sophisticated modelling apparatus. The normative structure is, in principle, more flexible because it explicitly permits comparison to be made between (and among) direct human and ecological (indirect human) values rather than maintaining an ideal that promotes hard constraints against compromise of ecological values (though in practice the conception tends to give greater weight to direct human values). The modelling structure is wider in scope because the natural environment is to be explicitly modelled so that it can be optimally exploited. It is more sophisticated, requiring the explicit representation of the time dimension and of human environment interactions, because the human interventions permitted in the natural environment result in the potential for significant changes over time and for feedback interactions.

3.1 Deficiencies of Stage II: from stable equilibria to complex dynamical models

Unfortunately however, experience involving practical applications of this approach to natural resource management (for example, "maximum sustainable yield" strategies), the impracticalities of implementing Hotelling's rule for resource extraction, and other examples such as those involving non-equilibrium dynamics, reveal that pursuing an "optimal" strategy under assumptions of perfect knowledge is nevertheless not necessarily robust to uncertainties.[10] Such under-modelling includes, in increasing significance, uncontrollable disturbance to inputs, errors in modelled dynamical parameters, or large-scale changes in the relevant behavioural dynamics (e.g. through self-organisation). We have noted above that arguably the main advantage in moving to the capital formulation model from the relatively simpler aim and dynamical models of environmental burden reduction lies in the normative deficiencies of the latter for formulating an adequate notion of sustainability where, in its extreme form, it provided only limited capacity to incorporate human benefit within the decision making process. In contrast, a significant deficiency of the "natural capital" approach (in addition to the difficulties of incorporating ecological norms for their own sake within the evaluation process) is the descriptive inadequacy of its dynamical framework. To explain.

The precision afforded by a nonlinear dynamical modelling apparatus of the natural capital approach to sustainability reveals that the accuracy of, in various cases, the model parameter estimates, the system initial conditions, state estimates, actual or predicted perturbations, is important in order to be able to effectively realise a reasonable approximation to an optimal deterministic temporal trajectory. In some cases, for example where the relevant trajectories are close to a simple equilibrium (one with a point or limit cycle attractor), these errors can be neglected or effectively handled statistically via the applicability of central limit principles. However, these issues can become particularly significant for dynamical systems that are either close to dynamical thresholds (even if at equilibrium), e.g. dry Australian grasslands near an organisational threshold for species diversity and productivity (see note 1 and text) or not close to a simple equilibrium, of which chaotic systems are an important subset (and include the weather).

The behavioural complexity explicitly recognised by descriptions of complex dynamic systems models emphasises the practical limitations of precise predictability. For example, locally divergent deterministic dynamics, such as that realised by positive feedback, chaos, bifurcations or any other differential amplification process, result in sensitive dependence of the dynamical trajectory to initial conditions. This implies that, due to uncertainty arising from irresolvable inaccuracies in the measurement of the system state (including through noise), accurate prediction is a practical impossibility, since unmeasured deviations from the actual state will be

[10]The problem with maximum sustainable yield strategies is that in seeking to maintain the optimal equilibrium condition, they presume that dynamical equilibrium is possible and desirable. Even if equilibrium can be maintained, small errors in the estimate of the parameters of the resource dynamics can result in recommendations that are far from optimal [Ludwig and Hilborn, 1983]. Hotelling's rule is that the percentage change in net-price per unit of time should equal the discount rate in order to maximise the present value of the resource capital over the extraction period; one problem with this is that it is not the market price that will prevail under competitive conditions, so that a monopoly market structure or appropriately calibrated, internationally co-ordinated, resource extraction royalties are necessary; another problem is that the appropriate discount rate is never obvious. See e.g. chapter 18 of [Pearce and Turner, 1990].

amplified over time. Due to such sensitivity, the range of possible future pathways that are compatible with a given (error-bound) estimate of the initial state very quickly becomes too large for any practically relevant prediction to be possible.[11] An epistemically equivalent source of state uncertainty derives from within the computational modelling of complex system dynamics that is nearly universally necessary, namely from the unavoidable finite rounding errors involved and the practical resource limitations on computational finesse.

As a direct consequence of the limitations of prediction accuracy with respect to the management of sustainability, deterministic optimisation is not a useful method for determining management choices. Instead it becomes necessary to *explicitly* consider a range of possible substantive consequences of any particular management strategy, that is, to consider contingent strategies. However, the uncertainties involved in managing complex systems are radical in the sense of also precluding optimisation over simple deterministic contingent strategies (concatenations of 'if state A, then do B with outcome C') taken either individually or in sets (e.g. corresponding to ranges of conditions, actions and outcomes) or over probability-weighted versions of these. The use of sets of simple contingencies is ineffective, not only because uncertainties can soon spread to encompass much or all of the state space and hence of the total action and consequence ranges as well, but because the dynamics may shift within the uncertain range so that no specific kind of action can be fixed upon and no specific kind of outcome can be calculated. Similarly, probabilities of subsequent states given initial states and of consequences can also become unspecifiable, limiting use of probabilistic versions as well.

In sum, for a conception of sustainability based on trajectory optimisation, behavioural complexity places significant informational demands on modelling, and on the management practices that rely on modelling, whenever a model requires a high degree of predictive and/or measurement accuracy in order to be validly applicable. In practice full information, especially across sufficient time, is never available (and often inaccessible in principle); often only much cruder information is known than could practically be made available. This places severe limits on the capital methodology — severe to the point of rendering most longer-term analysis quite illusory.

Put conversely, the capital approach tacitly assumes either that it is dealing only with dynamical models sufficiently simple that complete prediction and control over long time periods is feasible or that the discount rate is sufficiently large that

[11] Nevertheless, a dynamical system subject to uncertainty is not necessarily associated with significant prediction limitations. If the relevant trajectories are close to stable equilibrium ones, for example, sufficiently small uncertainties will result in essentially arbitrarily accurate predictability so that attempting to follow a unique optimal trajectory derived from a deterministic model represents a reasonable management strategy. This is because there exists a single deterministic trajectory that is capable of representing the range of possible trajectories to a sufficient degree of accuracy. But this is an unlikely situation for most of the complex systems with which sustainability management is concerned and becomes increasingly unlikely as the management time span increases.

all but short term differences are irrelevant to strategy evaluation. But the former is fundamentally wrong (the world is overwhelmingly dynamically complex) and the latter is an abdication of serious policy (the point is policy for long term outcomes), and each in its own way avoids confronting the challenges of making sustainability policy for a complex dynamical world.

4 STAGE III: SUSTAINED ADAPTIVE RESILIENCE — FROM SIMPLE TO COMPLEX DYNAMICAL MODELS

The lesson from the complexity of natural (and human) dynamics is clear: if we cannot predict and control the future then we should not try to aim for optimality (should implies can, so cannot implies should not).[12] We should instead try to ensure the satisfactory robustness of satisfactorily valuable system behaviour under uncertainty. The point of a building sprinkler system is not to predict the occurrence of fires, but precisely not to have to by ensuring the robustness of the building to the unpredictable outbreak of fire (while suffering only interim disruption). Moreover, the point of fire insurance is to ensure the same, this time via restoration of the building, in the unpredictable event of sprinkler failure (or where the costs of sprinkler use are deemed too great). This is a general insight, applicable in all management and policy domains: where prediction and control is infeasible or too costly, the focus should instead be on achievable robust outcomes, thereby relaxing the requirement for prediction. It is necessary instead to consider strategies that increase the ability of systems to sustain desirable conditions, and/or increase the ability of humans to influence the system towards doing so. This requires adopting a more sophisticated, higher-order, risk management policy framework: a qualitatively distinct conception of sustainability than that which results from a deterministic, full information, perspective.

With respect to risk, the preceding forms of fire insurance are not good specific models for responding to the unpredictability of ecological and eco-human dynamics, since it is often impossible or too costly, and often also too risky, to try to prevent all robustness-violating behaviours. So we need to reconsider responses to risk. Whenever there is the possibility of an untoward event (a threat) we can either try to prevent its occurrence (e.g., avoid war through peace-making), or try to prevent its damaging us, once occurred (e.g. have battle-proof defences, such as personal armour), or try to have a response to its impact that limits disruption while ultimately restoring important functionality (e.g. an effective tactical response force). All three options have in common that the requirement to pre-

[12]Technically, this does not follow if the only threats are those where all the probabilities of adverse events occurring are known, along with their impacts should they occur (that is, technically, cases of risk rather than uncertainty), for here we can optimise risk-weighted expected pay-off. If there are outcomes we cannot tolerate, then the next best class of policies are those that optimise conditionally on avoiding all the intolerable outcomes, or minimise their probabilities if avoidance is impossible. But this degree of knowledge is, as we shall see, very unlikely at any time and essentially impossible for any substantial time period into the future.

dict disturbance is relaxed (though prediction may contribute additional value if cheaply available). Where prediction and control is infeasible or too costly, only this last remains as response focus. It is the acquisition of relevant adaptive resilience.

The natural world uses a combination of aiming to prevent threats — principally through niche specialisation — or remove threats — principally through armour — but most often by acquiring adaptive resilience. Niche specialisation and armour illustrate the limitations of their respective strategies: they quickly limit other options (freedom of movement, etc.) so have a narrow success spectrum, and are often costly to maintain. Adaptive resilience, by contrast, encourages further synergistic options (like learning while creating new behaviour) and can have a wide a success spectrum that spans several categories of change; it is the overwhelmingly most important form of response and its efficacy underwrites the evolution of nervous system intelligence.

Within a policy context, design for dynamical resilience stands in contrast to attempting to tightly control the relevant variables directly via a small subset of the system states (often by manipulating faster acting decision variables) which is difficult in the absence of full information, especially if the corresponding basin of attraction is small. In fact, whilst ever the world includes within it partially autonomous subsystems, there will inevitably be wider global dynamics that are not under direct human control (and given that individual humans have some degree of autonomy, there will as well be global human dynamics that are not under the direct control of anyone). Instead design for dynamical resilience may be effected by modifying the parameters governing dynamical behaviour so that the basin of attraction of acceptable behaviour is enlarged. (Dynamical parameters are often more slowly changing and can be considered as belonging to a distinct subset of the system state.) Further, in some cases where there is local sensitivity to initial conditions, the trajectory followed may depend strongly on the particular realisation of even an energetically small fluctuation, the consequences of which are then effectively (energetically) amplified by the existing dynamics and result in the setting up of new global dynamical constraints. In this case the initial fluctuation has effectively resulted in a permanent and wide ranging consequence (the 'fixation of historical constraints'). The existence of such path dependencies, particularly those resulting from positive feedback, not only pose a problem for any general approach to policy formulation but, conversely point to the ability to have a significant influence in shaping, though not fine control in determining, future conditions. All this opens up the possibility of a wider range of behavioural design possibilities, indirectly modulating the behaviour of the variables of interest in order to widen their ranges of acceptable values, rather than attempting their direct control to within a narrower range.

The desirability of generalised resilience of satisfactory conditions, in the face of uncontrollable perturbations and pervasive uncertainty generated by unpredictable perturbations, gives rise to a requirement for appropriate adaptiveness. Adaptiveness is the capacity to appropriately respond to change contingently on the particular realisation of an actual perturbation. A contingent strategy whose subsequent decisions depend indirectly on the particular realisation of the uncertainty via direct dependence on feedback measurements made during the course of policy implementation represents a basic type of adaptability (i.e. an ability to adapt to the conditions signified by the feedback.) Even simple homeostatic feedback control, which is directed towards maintaining some given variable along a particular (non-contingent) trajectory, requires at least some type of adaptability since, although the target trajectory is independent of the particular uncertainty realisation, the actions required to approximately follow that trajectory are not. More complex feedback strategies, such as those that cannot be described as directed at maintaining a particular contingency-independent trajectory (such as those where the target trajectory itself is contingent on, that is adaptively responds to, the realised perturbations) require still more adaptability. The capacity for *anticipative* adaptiveness — to alter some condition in anticipation of improvement, often brings with it not only the ability to survive through significant environmental change, but also to flourish.

The development of appropriate anticipative adaptiveness (the development and appropriate exercise of alternatives) is more general than the maximisation of capital (natural or otherwise), although it reduces to it where the future is known with certainty and there is perfect control of outcomes based on knowledge of the relationship between decision alternatives and resulting consequences — but only because among the policy actions to secure maximal capital will be those conditioned on responding to the now-known perturbations. Whence the exercise of options, and with it the still-requisite resilience thus manifested, remains implicit in the assumption of policy effectiveness and unexamined. This makes the adaptiveness approach a natural generalization of the capital approach, just as that approach was itself a natural generalization of the impact approach. However, it also typically requires more analysis effort to evaluate the desirability of a given design (sequence of contingent adaptations) and consequently to derive specific desirable designs. Nevertheless, an overall increase in the scope and range of adaptive options will, all else equal, always improve the potential to sustain valued conditions.

The recognition of the fact of unpredictability of the precise trajectory that will result from any particular given (contingent) strategy requires a further development in evaluation method. The environmental impact approach to sustainability policy requires, in principle, evaluative comparisons to be made among impact categories (in those cases where it is impossible to completely avoid all impacts). The natural capital approach requires, in addition, evaluative comparisons to be made between natural and human value categories but also, novelly, comparisons to be made among valued processes realised at alternative points across time. The adaptiveness approach, due to the explicit recognition of unpredictability, requires (novelly) comparisons to be made among sets of possible futures, with or without associated probabilities. Thus the necessity to develop a risk-modulated evaluation methodology is yet another policy implication of the recognition of complex dynamics.

This third conception of sustainability is more fundamental than either of the other two because it, unlike the others, is explicitly able to cope with a fundamental feature of the world: ongoing and essentially unpredictable change, including those changes resulting from generalised evolution, development and learning. It shares the advantage of the natural capital approach in that it is normatively general enough to recognise the relevance of the human subsystem (in its capacity to enhance adaptiveness via both technological development and cultural development that leads to more capable means of social co-ordination), in addition to the relevance of natural ecosystems (in their capacity to enhance adaptiveness via species evolution, symbiosis and succession). However, it is explicitly concerned with the capacity to develop innovative capabilities, which provide the means by which to ensure that productivity can be maintained in the face of uncertainty and deep change for both systems.

Both an advantage and a disadvantage is that the adaptiveness approach to sustainability management potentially incorporates understanding of a much wider scope of concern and much longer time scales than its predecessors. The impact reduction approach focuses on the details of a narrow human technical subsystem and is primarily concerned with instantaneous flows. The natural capital approach considers a slighter wider natural dynamical subsystem into which some human technical process is inserted and extends concern to system states including levels of stocks over much longer time scales. (But because it needs to directly analyse and evaluate actual future productivity over time, it requires a reasonable degree of precise predictability and is in practice typically valid only over relatively shorter term time scales.) The adaptiveness approach is potentially concerned with whole-of-system adaptiveness, a capacity that often results from the synergistic capabilities of the combination of more than one human technology (long distance communications networks plus precision robotic surgical equipment enables telemedicine), ecosystem processes (bacteria with oxygen respiration capability plus prokaryote-provided protection of the plasma and membrane of the extended cell enables eukaryotic cells with nuclei) or human and natural processes (naturally occurring analgesic salicylic acid plus industrial manufacturing enables the mass production of aspirin). While these adaptiveness capabilities are ultimately materially realised as physical system states, these states are not only outside the scope of those identified within the other two approaches, but their own dynamics are themselves subject to significant predictive uncertainty. The precise consequences of the exercise of a new adaptiveness often cannot be accurately known, nor is it often known whether, and how, a new adaptiveness can be acquired.

In terms of the framework for sustainability identified above, the normative ideal implicit in the "Adaptiveness" conception of sustainability is that of maximising the scope and range of valuable options open to human society and the ecology jointly. Both human and ecological survival, development, and flourishing are valued inherently. While the ultimate rationale for concern with sustainability is a desire to maximise the potential for survival across a broad range of uncertainties, a reasonable measurement focus would include the capacity to change, and the ability to exercise that capacity appropriately. A more complete adaptiveness sustainability framework would entail a theory of the multi-timescale development of both biological organisms and ecologies (evolution, succession, community adaptation ...), and of human development (technological and cultural learning and innovation ...), together with an understanding of how best to promote the development of adaptive options in these domains.

This represents a radical revision and generalization of standard conceptions of sustainability and the profoundest change in that field brought about by the recognition of complex dynamics underlying public policy making. The uncertainty consequences of complex dynamics have yet to work their way through sustainability policy making. For instance, government and industry policy makers alike still persist in the view that the only rational energy technology policy is to estimate either the most promising future energy technologies or at least a guaranteed profitable technology, or refrain from investment — that is, not to invest unless prediction is available — when there is strong evidence that the domain is and will remain unpredictable for some considerable time and share market investment already offers a rational alternative: invest in a technology portfolio that spreads the risk and ensures a capacity to respond as opportunities arise, that is, achieve energy policy resilience through building technical adaptability. In most other areas of policy making these considerations have yet to make an impact; they have a recognised foothold only in ecological policy and military policy, where in each case they have deep consequences.[13] The essay concludes by exploring the complex systems aspects of these issues a little further.

5 COMPLEX DYNAMICS: CONSEQUENCES FOR SUSTAINABILITY POLICY

The recognition of the behavioural complexity of nonlinear dynamic systems has three further implications for sustainability policy, indeed for policy more generally, worth noting. One implication is the applicability of robust backcasting as a generalised policy strategy. Another is the general issue of solving open problems, and in particular the recognition that directed learning capacity is important, especially a capacity to learn while doing. A third implication is the intimate and explicit roles of values, not only in policy formation and validation, but also in the supporting learning processes that result in the formation of relevant understanding.

Robust Backcasting. Despite the limitations on predictability in complex systems, it is often possible to learn a system's dynamical possibility structure, that is, the structure of its possible trajectories. (Often this is done by exploring various

[13]See [Walker and Salt, 2006; Allen, this volume; Bishop, this volume] on ecological policy. See also [Ryan, this volume] on military policy.

bundles of trajectories through computational simulation, using empirical investigation and limited prediction to estimate the simulation model.) Accepting that prediction limitations require dropping the policy aim of realising a unique optimal trajectory, the focus of policy is shifted to one of trying to realise any among a class of sufficiently valuable or satisfactory futures through approximate shaping. Then the possibility structure can be made the basis of effective policy making in the following way.[14]

In contrast to forecasting, backcasting begins from a future satisfactory end state and works backwards in time toward the present in order to construct possible sequences of preceding states that, begun from the present state and run forward in time, would foreseeably yield the selected end state (or a state sufficiently nearby to be equally satisfactory). The collection of trajectories that end in the same state forms its backcast scenario possibility cone and provides information about the key requirements for arriving there from here given various contingencies along the way.[15] Conversely, the only way to identify all the developmental issues involved in planning is to start with the required outcome and ask, at each step back in time toward the present, what would have to be in place to proceed towards the goal. Repeating this scenario construction, for each of a range of satisfactory end states that encompass the main alternatives for a domain, yields a backcast scenario space for the domain. Within this space select those scenarios whose performance is at least satisfactory.[16] While uncertainties undermine both forecasting and backcasting, the latter has a natural extension that substantially mitigates it in most cases: *adaptive backcasting*, whose ideal is to proceed in a way that holds as many as possible of the satisfactory scenarios open to pursue, for as long as each can possibly deliver value. In practice, limited knowledge and resources quite reasonably tightens ‘possible’ to ‘plausible’ or ‘sufficiently likely’, while over time developing technological learning may, e.g., reveal poorer than hoped performance for some technologies, leading to their elimination, and also point to new technologies that need to be included.[17]

---

[14]For those familiar with rationality theory, this is to take a *satisficing* [Simon, 1982] approach in contrast to forecasting’s optimisation or maximising approach. For more on backcasting as a tool for policy formation under uncertainty, see [Brinsmead and Hooker, 2008] and [Hooker, Crossley, 1987].

[15]Consider, for instance, a 2050 end state in which Australia meets its responsible greenhouse gas emissions targets by replacing coal-fired with nuclear electricity power generation; then by 2030 it likely would need to be commissioning a new gigawatt nuclear station per year or better; clearly this will require substantial industrial development before that date, along with government regulatory, skills training and planning developments a decade or more earlier still.

[16]In the case of energy planning, e.g., one might select scenarios that deliver sufficient energy security throughout, and cost less to implement than the value they ultimately deliver over the amortisation of their infrastructure. Note that to learn about the possibility structure of a system for satisfactory end states and trajectory performances will often require also investigating unsatisfactory end states and trajectories.

[17]Forecasting has a related, but less natural and powerful, extension in response to uncertainty wherein it is augmented so as to simulate some of the comparative information that adaptive backcasting method provides. Essentially, this is achieved by (i) inserting into the forecasting base a set of extra data corresponding to the implementation of various outside-trend scenario

conditions that are not normally accessible, (ii) adding Real Options Valuation [ROV] — evaluating the having of various real options to take up each scenario at various given times and also various conditional compounds of such options, and then (iii) further adding a robust adaptive strategy analysis to ameliorate risk. However, the strength and limitations of ROV arises from its use of expected net present value [ENPV] (and variants) as its evaluation methodology, the same methodology that is used by the capital approach to sustainability and defeated by uncertainty. The consequence here is that this augmentation presumes at least the backcasting analysis of the scenario space and its key backcast decision points and estimates of relevant uncertainty ranges for its options, and it must re-create adaptive backcasting strategy analysis in order to provide a principled risk amelioration analysis. It must also artificially track synergies and other interdependencies among scenarios in order to find optimal pathways, but optimisation itself remains a risky methodology in the face of structural uncertainty. In short, adaptive backcasting is the foundational requirement and the thus-augmented forecasting method appears as a somewhat contrived and limited extension whose decision structure is framed by it. Nonetheless, it can be practically useful in appropriate circumstances.

We append here a note on the role of Dynamic Programming [DP] in ROV versus adaptive backcasting, because DP is a well known analysis technique and there is a danger of supposing backcasting reduces to its use. Optimal ROV strategies are calculated using DP. The importance of DP is that it implicitly considers all possible explicitly identified pathways to a specified objective in order to evaluate the optimal one, doing this by working backwards in time from the end goal, just as backcasting does. However it does this simply to reduce the number of already identified paths that must actually be evaluated, as a calculational efficiency exercise, rather than as a necessary part of a process to explore (and hence uncover as yet unidentified) possibilities. This is the root difference between DP and BC.

The assumption underpinning DP limits its applicability when there are path-dependence effects, such as where investing in a technology lowers its subsequent capital or operating cost or where it increases the value of a product through fashion and other 'positional' effects. By contrast, BC is able to take these and other possibilities into account, though without the rigorously thorough search. To see this note that, in order to structure evaluation, DP adopts the following partition assumption DPP: if P is any point on the optimal path from start state X to end state Y then the optimal path from P to X is the P-X part of the overall optimal path. Given DPP, there is no point in evaluating all possible paths from X to Y in order to identify the optimal one, instead one chooses a later state P and evaluates the optimal path just from P to X then, re-applying DPP, steps back in time to an earlier point O and evaluates the optimal path just from O to P, and so on back to Y. Then, if there are several candidate O states, one repeats this latter pathway evaluation for those and selects that O yielding the highest path value and, working forward, if there are several candidate pre-final P states, one repeats the overall pathway evaluation for those and selects that P yielding the highest path value to thus identify the optimal X-Y path and its value.

This makes it clear that DPP excludes path-dependence, since the optimality of the sub-path P-Y is assumed determinable independently of all X-P paths. But if, e.g., technology learning were in play then a path from X to P involving investment in technology T1 would raise the value of a P-Y path using T1 relative to all P-Y paths using alternative technologies and so could change which P-Y path is optimal, and/or change which P yields a comparatively higher valued optimal path, contrary to DPP. (Technically, path dependence can be avoided by introducing additional parameters to keep track of path-dependent values along paths and add these variables to the node states like P and O, but while this formally restores DPP, it does so only at the expense of adding the entire array of distinct values of these variables to every path segment optimisation. If only one simple cumulative variable is involved, as it is for the simplest technology learning curves that depend only on total installed capacity, then the resulting calculation is manageable since many paths will show the same value, but in other cases the consequence is that all possible paths have to be considered and evaluated.)

Nonetheless, even though it is not searching for an optimal path, the BC exploration of possibilities is in fact analogous to the DP process and it achieves similar calculational efficiency improvements: searching backward in time across each time step constrains the possibilities that

A strategy that delivers adaptive coverage of options in this way is a *robust adaptive strategy*. It will specify a portfolio of real options — that is, actual technological, infrastructural, financial and other preparations — that retain the capacity to pursue the widest feasible class of the most satisfactory scenarios for the circumstances. Such a policy has as a motivating goal to achieve the realisation of a satisfactory end state, however its primary goal must be to create and sustain the *adaptive resilience* of the system planned for, because in conditions of uncertainty that goal provides the necessary as well as sufficient precondition of goal achievement. This applies to both natural and human systems. The adaptive resilience thus acquired confers the ability to meet many changes without having to predict them and, depending on response preparation, to respond adaptively to emerging circumstances with reduced lead times. Its development typically requires the development of generalisable adaptive capacities that may prove useful even in the event of some unforeseen contingencies. Adaptive resilience is the basic way to protect against or ameliorate the risks posed by uncertainty (and external uncontrollability); although it may be limited by the cost of carrying the adaptive capacity and the cost of exercising it effectively contingent on risk materialising, it is the primary form of insurance against uncertainty.

Both forecasting and backcasting have their strengths and weaknesses, with the former most useful for the short term, the latter most valuable over the long term and the more fundamental for exploring possibilities, and hence for setting the broader context for policy making. The short term usefulness of forecasting derives from predictability based on plentiful operational data and the limitations on our ability to intelligently intervene rapidly in our world; the longer term usefulness of backcasting derives from predictability based on the general structure of empirical possibilities and our greater ability to intelligently intervene slowly in our world. Backcasting has these capacities because a) many variables that are effectively fixed in the short term become open to decision in the longer term (because of the roles of constraints and initial conditions in dynamics) and b) it is not confined by current trends, while c) it can also allow for self-reinforcement along pathways and for synergistic effects among pathways.

Open problem solving. The recognition of complex dynamics has other consequences for policy than explaining the utility of Robust Backcasting. It also emphasises the importance of solving open problems. A problem is closed when it is fully, uniquely specified — when everything required for its solution is provided — and all that remains is to correctly carry out the solution method, obtain the solution and check that it satisfies the solution criteria. Textbook problems, indeed well rehearsed exercises of familiar expertise in any well understood domain, are of this kind. Abstracting, a closed problem has the following elements: a well-specified objective or goal state, including constraints, and well-specified criteria for determining whether or not the goal has been achieved or the constraints violated; a well-specified solution space, that is an account of what decision op-

---

continue to be explored to those that are relevant to reaching the segment end state, thus limiting the search analogously to that of DP.

tions are available to the problem solver; a well-defined observation space, that is an account of what kind of information is available and what sub-set suffices for determining a solution; and a well-defined causal constraint model, that is, an account of the relation between decision option and outcome (including the possibility of unpredictability, provided that the unpredictability itself is bounded and well-specified). Interaction with complex dynamical systems requires a capacity for solving open problems because of the failure of any or all of these conditions, especially the absence of an adequate causal constraint model.

Problems are open to the extent they are not closed. Openness is a matter of degree: in an open problem one or more of the required elements for a closed problem is under-specified, is left uncertain, or is to some degree missing. Acting within and on complex systems very often gives rise to an open problem since it is easy to miss subtle, long term dynamics, thresholds, the significance of combined but individually weak interactions, and so on, and this often means that the terms of the problem and its proper methods of investigation are not understood. Understanding North Atlantic fisheries policy [Allen, this volume] provides a nice case in point, as could any other chapter. Indeed, most problems central to life are profoundly open.[18] Every decrease in closure represents an increase in the depth of uncertainty characterising the problem situation. Deep uncertainty pertains to widespread uncertainty about the nature of the problem itself and its solutions. The character of what is missing and its significance may itself not be understood. And when it is provided it can 'turn the world (that is, the problem conception, its presuppositions and guiding criteria) upside down'. Its provision following experiment/inquiry may be deeply surprising.

The literature describing open problem solving is as widely distributed as its applications, marked by idiosyncrasy of ideas and terminology, and tentative. In engineering and industrial design, e.g., the term maieutics "the art of giving birth to an idea"[19] is used to describe the process of jointly exploring problem definitions and possible solution options with the end of sufficiently constraining the definition of the engineering problem to a degree that it becomes a closed problem so that optimisation methods can be applied. The creative aspects of problem framing are usually described as being more "art" than "science" and French (note 19) demonstrates this by presenting numerous case studies emphasising the creativity

[18]How to design an award-winning house, validly research a new domain, have stable economic markets, achieve personal maturity, construct an adequate sustainability policy, ... One example of an open problem is how to "throw a nice party" [Hatchuel, 2001]. The criteria for determining what defines a "nice party" are not well-specified, and may have to be discovered as part of the design or problem solving process. The option space is not well-defined either and could include, for example, a choice between a variety of fixed locations, a moving location, or a distributed or virtual location. The choice of location will constrain the options for entertainment which may include party games, dancing, conversation, music, theatre, dining, or sport. There are no clear boundaries on either class of options: for location or entertainment. Neither are the criteria of success well defined: is it to be judged a nice party if the guests stay long, or drift in relaxed and go again in high spirits? Should they behave with quiet absorption, or 'go wild'? Should they express appreciation for the memories, or be unable to remember anything? ...

[19][French, 1988]

aspects of the engineering design process. Hatchuel names the ability of humans to solve open ended design problems "expandable rationality", suggesting that this is a better metaphor for intelligent creative design ability than Simon's notion of "bounded rationality" because it emphasises possibility rather than limitation, and suggests that human collective action creates richer designs than could be achieved by any single individual.[20]

The concept of "double-loop learning" in management theory refers to a problem solving process suited to tackling open problems in which both the problem model is learned in conjunction with the techniques to solve the problem as contemporaneously represented[21]; the process is called 'adaptive control' in the process engineering literature. It applies throughout the handling of complex systems, e.g. to the development of accurate predictive capacity: learning how to interpret measurements as observations in the light of theoretical understanding, while simultaneously undertaking learning that improves theory in the light of those observations, including the theory of the observational processes. It similarly applies to the focus of attention on relevant dynamics: learning which aspects of environmental dynamics can be reasonably ignored in the light of given sustainability principles and goals, while simultaneously undertaking learning that refines those principles and goals in the light of the dynamics.

Christensen and Hooker provide a general characterisation of open problem solving processes in what they refer to as self-directed anticipative learning (SDAL).[22] Here anticipative action is behaviour shaped by the anticipation that, in the circumstances, it will result in desired (rewarding) feedback, including both further information and non-informational resources (for example, tools). Self-directedness is the capacity of an organism to evaluate the feedback received against life needs and to selectively shape (i.e. direct) its future behaviour so as to improve both the immediate value of the feedback achieved (e.g. better feeding) and, more importantly, its anticipative capacity to act so as to generate valuable immediate feedback more effectively (better identification of good feeding sites). An SDAL process is the establishment of a feedback loop in which directed interaction with the environment generates information that improves the system's anticipation and thereby modifies the system's interaction processes, generating yet more refined information, and so on. In this way the system uses interaction to modify itself and/or its environment in ways that simultaneously move it towards its goal and improve its capacity to move towards its goal. The solution, the specific method for achieving it, and the proper formulation of the goal itself, are all progressively acquired (or, equivalently, refined).

To illustrate its power, the SDAL model has been applied to the study of the scientific research process itself, in particular to the early history of ape language research[23] but a first intuitive picture of the process is provided by a detective

[20] [Hatchuel, 2001; Simon, 1982]
[21] [Sterman, 2000; Argyris, 1999]
[22] See e.g. [Christensen and Hooker, 2000a; 2000b].
[23] See [Farrell and Hooker, 2007a; 2007b; 2009; Hooker, 2009].

investigating a new crime.[24] It is the interplay between the discovery of clues, the construction of a suspect profile and subsequent modification of the anticipative investigation strategy that makes the process an example of self-directed learning of anticipative action. The consequent feedback produced at each iteration serves to evaluate the success of the anticipations, whilst the anticipations themselves help the system improve its recognition of relevant information and evaluate its performance more precisely. Here we see the interrelations between integration and localisation: successful detecting is, first and foremost, an integration problem — the perpetrator is caught only when a complex array of factors (clues, motives, options, arguments, methods and so on) are brought into coordination; however, learning to achieve this coordination requires the localisation of important factors within the complex (such as clues and valid methods to investigate them). As the system interacts in an SDAL process its improving anticipative models and interaction processes allow it to: a) improve its recognition of relevant information, b) perform more focused activity, c) evaluate its performance more precisely and d) learn about its problem domain more effectively. Indeed, in this setting error itself can be a rich source of context-sensitive information that can be used to further refine these four features (note 23). It is proposed that all methods for solving open problems are instances of SDAL.

The preceding discussion makes it clear that learning is central to SDAL processes. The richer the system's anticipative structure is the more selective or directed its learning can be, and the more potential there is that learning will improve the system's capacity to form successful anticipative models of interaction with the problem domain. When successful, SDAL results in a "pushme-pullyou" effect as learning is pushed forward by the construction of new anticipations and pulled forward by the environmental feedback generated, creating an unfolding self-directing learning sequence. Because of its characteristic self-improvement SDAL can begin with poor quality information, vague hypotheses, tentative methods and without specific success criteria, and conjointly refine these as the process proceeds. This makes SDAL powerful because it allows successful learning to arise in both rich and sparse cognitive conditions. If the anticipations prove unsuccessful the sys-

[24]A detective's investigation, say of a murder, is typically an SDAL process. The detective's role provides a general problem: bring to justice the perpetrator of any crime committed. But this problem is typically highly ill-defined. The detective is initially ignorant of the nature of the crime, whether it was from passion, 'business', psychopathology, ..., of what features of the crime scene constitute informative clues (the bed clothes?, bank book?, the half-consumed sandwich?, ...), of what investigatory methods will ultimately pay off (psychological?, financial, ...), of what specific criteria to apply to achieving a solution (proving sufficient jealousy?, or random victimhood? ...), and perhaps even whether a crime has been committed. She begins, then, with a purely mechanical collection of standard information concerning any disrupted order, from a broken dish to a stranger's presence. The aim is to build a profile of the suspect, however initially tentative and vague, and then use this profile to further refine the direction and methods of the investigation. The profile tells the detective what the murderer is like and what types of clues to look for. This in turn sets new intermediate goals that focus the investigation, such as search for organised crime links to the victim, and if the profile is at least partially accurate the modified investigation should uncover further evidence that in turn further refines the search process, ultimately (hopefully) culminating in capture of the murderer.

tem will hunt fruitlessly, but if they are even partially successful the system can progressively improve its ability, bootstrapping its way to a solution.

It is also clear that the dynamics of learning is itself characterised by positive feedback path dependencies, e.g. for double loop learning the more accurately understood are the underlying system dynamics, the easier it is to manipulate trajectories within those dynamics, which then provides the enhanced capacity to (safely) deliberately explore the dynamics in order to understand them better. The uncertainty resulting from this amplification process during learning reinforces the unpredictability that is generated anyway through both creativity and accident in innovation. Managing the process of learning, which is required to support the risk management of a complex dynamic sustainability problem, then becomes itself a risk management problem.

In these circumstances, faced with an open problem, how should one proceed? Other than wishing for 'global inspiration', suggestions in the literature come down to a combination of (i) do what is accessible and will pay off now, even if of modest value ('pick the low risk fruit'); (ii) using i and background knowledge, break the problem up into manageable sub-problems that fall under i; and (iii) integrating what comes of i + ii at each stage, successively iterate towards a global solution.

This abstract iterative process is the skeleton of that specified earlier by the SDAL model and, as there, the detective's investigation provides an intuitive illustration. Faced with a crime of unknown kind, the detective can immediately collect all accessible information from the scene (cf. (i)), then formulate the most prominent options suggested by that, e.g. a crime of passion, business or social pressure, and for each build a tentative, sketchy profile and determine the best few next investigatory steps to take (cf. (ii)), then take these steps for each case and assess the impact of all the resulting new data on each profile, develop or eliminate each profile (cf. (iii)) and repeat the iteration until just one, increasingly rich and consistent profile remains. However, much of the power of the SDAL process lies in the openness to radical change, conceptual, methodological and so on, embodied in step ii; e.g. a new clue may switch profiles from a crime of passion to a business crime and with that switch conceptions of motive and relevant evidence, switch methods from relationship construction to financial auditing, and so on; in the skeletal formulation all this is hidden behind 'using ... break ... up', its unavailability sometimes reinforced by the incorrect assumption that the break-up can be complete rather than partial. In the policy literature this general procedure is often referred to as a 'Do-a-Little-And-Learn' method, and has been argued to be superior to any less incremental approach in many open problem circumstances.[25]

Roles of values. Sustainability policy for complex systems involves being normatively directed by guiding values in a distinctively intimate manner. First, while sustainability is already a normative goal since what is to be sustained is a matter of decision about what is considered valuable (whether in itself or to humans), it

[25] See, e.g. [Lempert and Schlesinger, 2002]. Resilience-based adaptive management approaches to environmental management are typically of this sort (see note 1 and text), as is management's 'double-loop learning' process (see [Sterman, 2002]).

is possible to suppress the role of values in the characterisation of policy formation, e.g. under the factual-sounding policy goal of removing pollution. However the presence of complex systems forces the use of backcasting methodology for policy formation and backcasting makes the driving values explicit in the choice of satisfactory end states. Second and more subtly, while policy construction involves solving the problem of what is the effective empirical basis for constructing specific policy and problem solving is a normative process, its evaluative nature can be suppressed by adopting the factual-sounding goal of finding the closed-problem method that suffices to automatically produce the solution. However the presence of complex systems forces on policy open problems that require learning the problem and the solution together and this requires making explicit additional policy-relevant values like sensitivity, decisiveness, timeliness, feasibility, controllability and robustness, e.g. in system model construction of relevant dynamics.

Third, while policy construction involves learning the best specific form for policy and learning is a normative process directed by general epistemic values like consistency, accuracy, precision, and scope, it is possible to suppress the role of values in the characterisation of learning by adopting the factual-sounding goal of explanatory adequacy. However the presence of complex systems forces the epistemic values to be made explicit through the roles they play in the choice of learning decisions, e.g. in the management of the risks of learning reinforcement. Finally, all three sets of values are forced to be explicit through the tension between practical resource and timing constraints on policy development and the realisation of policy values resulting from the implementation of the resultant strategies.[26]

6 SUMMARY

A complex system dynamics perspective has implications for sustainability policy that — compared to a relatively less sophisticated mechanistic perspective — are subtle and profound in their insight regarding the limitations of human understanding and control. A complex systems dynamics perspective emphasises the sustainability policy relevance of dynamical parameters such as stability time "constants", basins of stable attraction and their size (resilience), and higher order dynamical capacities such as adaptability, flexibility, insurance, innovation, learning and regulatory capacity, all of which depend on wider system organisation, rather than more directly accessible measurements of physical system state. It emphasises the importance of values in directing the development of the understanding that supports sustainability policy, including inter-temporal and risk-modulated considerations with respect to substantive policy values, and methodological values that must guide the policy process.

The existence of powerful positive feedback dynamics explains the leverage that may exist for human choices to influence future outcomes in some contexts, yet the

[26] For all these reasons, Brinsmead [2005] finds that the most effective construction of integrated complex systems models for applied policy support always begin with, and have their construction explicitly driven by, these values.

profound impotence of human endeavour in others. The ubiquity of dynamics with extremely sensitive dependence on initial conditions emphasises the profound limitations on the possibilities of alleviating ignorance, and the necessity of developing contingent strategies for ameliorating the material effect of smaller uncertainties, and developing contingency strategies for mitigating the negative evaluative consequences (and enhancing the positive) of more significant uncertainties.

Compared to conclusions drawn from a mechanistic perspective, the wider environment is apparently much less controllable, implying that human action ought to be undertaken with as much humility and caution as innovation and optimism. Thus the perspective recommends an approach to sustainability policy based not merely on boldly backcasting from desired futures, but also with the prudence of contingency planning and insurance, supplemented by cautiously innovative doing-a-little and learning. This learning approach to innovation requires sensitivity to the contingencies of empirical context-dependence and, in practice, the application of heuristic learning and interaction strategies. Thus a deeper understanding of complex dynamical systems offers the potential to more precisely define the boundaries between the knowable and unknowable; and the controllable, the shapable and the independent (autonomous); and to narrow the gaps between them.

In sum, the entry of a complex systems perspective into sustainability management transforms its conception, methodology and policy formulation, in the process providing generalisable policy tools for all such complex dynamical systems situations. Such changes are now spreading throughout the socio-political domain and promise to equally transform our conception of intelligent governance generally.

BIBLIOGRAPHY

[Adler and Kwon, 2002] P. Adler and S.-K. Kwon. Social Capital: prospects for a new concept. *Academy of Management Review* **27**(1), 17-40, 2002.

[Argyris, 1999] C. Argyris. *On organizational learning.* Blackwell Business, Malden Mass., 1999.

[Brinsmead and Hooker, 2008] T. S. Brinsmead and C.A. Hooker. Adaptive backcasting: the method of possibility and design for planning the future. Discussion Paper VIII, CRC for Coal in Sustainable Development, `http://pandora.nla.gov.au/pan/64389/20080828-1328/www.ccsd.biz/publications/industry.html`, 2008.

[Brinsmead, 2005] T. S. Brinsmead. *Integrated Sustainability Assessment: Identifying Methodological Options.* Report to the National Academies Forum, `http://www.naf.org.au/naf-forum/brinsmead.pdf` , 2005.

[Brundtland, 1987] G. H. Brundtland. *Our Common Future: World Commission on Environment and Development.* Oxford University Press, Oxford, 1987.

[Christensen and Hooker, 2000a] W. D. Christensen and C. A. Hooker. Autonomy and the emergence of intelligence: organised interactive construction. *Communication and Cognition -AI* **17**(3-4):Special Edition, The Contribution of Artificial Life and the Sciences of Complexity to the Understanding of Autonomous Systems, 133-158, 2000.

[Christensen and Hooker, 2000b] W. D. Christensen and C. A. Hooker. An interactivist-constructivist approach to intelligence: self-directed anticipative learning. *Philosophical Psychology* **13**, 5-45, 2000.

[Dasgupta and Serageldin, 2000] P. Dasgupta and I. Serageldin. *Social Capital: a multi-faceted perspective.* The International Bank for Reconstruction and Development, Washington USA, 2000.

[Farrell and Hooker, 2007a] R. Farrell and C.A. Hooker. Applying self-directed anticipative learning to science. a: Agency and the interactive exploration of possibility space in ape language research. *Perspectives on Science* **15**(1), 86-123, 2007

[Farrell and Hooker, 2007b] R. Farrell and C. A. Hooker. Applying self-directed anticipative learning to science. Learning how to learn across revolutions. *Perspectives on Science* **15**(2), 220-253, 2007.

[Farrell and Hooker, 2009] R. Farrell and C. A. Hooker. Applying self-directed anticipative learning to science: identifying error and severe testing through interactive exploration of possibility". *Foundations of Science*, 14(4), 249-271, 2009.

[Flannery, 2002] T. Flannery. *The Future Eaters*. Grove Press, New York, 2002.

[French, 1988] M. J. French. *Invention and evolution: design in nature and engineering.* Cambridge University Press, Cambridge, 1988.

[Gunderson and Holling, 2002] L. H. Gunderson and C. S. Holling. eds. *Panarchy: Understanding Transformations in Human and Natural Systems.* Island Press, Washington, 2002.

[Hatchuel, 2001] A. Hatchuel. Towards design theory and expandable rationality: the unfinished programme of Herbert Simon. *Journal of Management and Governance* 5(3-4), 260-171, 2001.

[Hooker, 2009] C. A. Hooker. Interaction and bio-cognitive order. *Synthese* **166**(3), 513-546, special edition on interactivism, M. Bickhard ed., 2009.

[Hooker and Crossley, 1987] C. A. Hooker and D. J. Crossley. Retrospective scenario analysis and government energy planning: toward an effective synergy. *Energy Systems & Policy* **10**(2), 93-115. 1987.

[Holling, *et al*, 1995] C. S. Holling, D. W. Schindler, B. Walker, and J. Roughgarden. Biodiversity in the functioning of ecosystems: an ecological primer and synthesis. In C. Perrings, K.-G. Maler, C. Folke, B.-O. (eds.), Jansson, *Biodiversity Loss: Economic and Ecological Issues*, Cambridge University Press, Cambridge, 1995.

[Leahy, Undated] T. Leahy, ed. *Mending the Broken Land: land care and food security for the South African villages.* University of Newcastle 2308, Australia, undated.

[Lempert and Schlesinger, 2002] R. J. Lempert and M. E. Schlesinger. Adaptive strategies for climate change. In *Innovative Energy Strategies for CO_2 Stabilisation.* (ed. R. Watts), Cambridge University Press, 2002.

[Levins, 1968] R. Levins. *Evolution in Changing Environments, Some Theoretical Explorations.* Princeton, NJ: Princeton Univ. Press, 1968.

[Ludwig and Hilborn, 1983] D. Ludwig and R. Hilborn. Adaptive probing strategies for age-structured fish stocks. *Canadian Journal of fisheries and aquatic science* **40**, 559-569, 1983.

[Pearce and Turner, 1990] D. W. Pearce and R. K. Turner. *Economics of natural resources and the Environment.* John Hopkins University Press, Baltimore, 1990.

[Rappaport, 1967] R. A. Rappaport. Ritual regulation of environmental relations among New Guinea people. *Ethnology* **6**(1), 17-30, 1967.

[Simon, 1982] H. Simon. *Models of Bounded rationality (Volume III).* MIT Press, Cambridge Mass., 1982.

[Smetacek and Naqvi, 2008] V. Smetacek and S. W. A. Naqvi. The next generation of iron fertilization experiments in the Southern Ocean. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* **366**(1882), 3947-67, 2008.

[Sterman, 2000] J. D. Sterman. *Business Dynamics: systems thinking and dynamics for a complex world.* Irwin McGraw-Hill, Boston, 2000.

[Walker *et al.*, 2002] B. Walker, S. Carpenter, J. Anderies, N. Abel, G. Cumming, M. Janssen, L. Lebel, J. Norberg, G. D. Peterson, and R. Pritchard. Resilience Management in Social-ecological Systems: a working hypothesis for a participatory approach. *Conservation Ecology* **6**(1), 14ff, 2002.

[Walker *et al.*, 2004] B. Walker, C. S. Holling, S. Carpenter, and A. Kinzig. Resilience, Adaptability and Transformability in Social-Ecological Systems. *Ecology and Society* **9**(2), art5., 2004.

[Walker and Salt, 2006] B. Walker and D. Salt. *Resilience Thinking.* Island Press, Washington DC, USA, 2006.

Part XII

Philosophy of Science

INTRODUCTION TO PHILOSOPHY OF COMPLEX SYSTEMS

Cliff Hooker

Part B: Scientific paradigm + philosophy of science for complex systems: a first presentation c. 2009

Pursuit of every scientific framework — that is, of a paradigm and philosophy for science — is underwritten by a practical act of faith that its cognitive apparatus — including concepts, classes of models and underlying mathematics, and experimental instruments, techniques and interpretations — is adequate to understand the domain concerned [Hooker-a, this volume, section 2]. The focus of this essay is the consequences of the cognitive apparatus of complex systems for methodology, epistemology and metaphysics.

The world has turned. The old orthodox framework for science that sufficed for the study of simpler systems — physical systems of a few components and of many-component equilibrium systems — no longer suffices: science has discovered the phenomenal richness and explanatory power of complex systems. In every science to some degree and in many sciences completely, the Bénard cell, Belousov-Zhabotinsky reaction, logistic reproduction and *Dictyostelium* aggregation have replaced the slingshot, gas and crystal as model systems. With that there has emerged a general model of complex organisation of dynamic processes. Typically there are many processes running simultaneously on different groups of timescales and across different spatial scales under multiple constraints, sometimes modular and sometimes reproduced by other processes. The processes are otherwise all mutually interacting, their non-linearities establishing complicated behaviours with amplifications and bifurcation thresholds. Typically, the slower, more extended processes set conditions and constraints for faster, less extended processes, while products of these latter accumulate to alter the slower, more extended processes, but with the whole interwoven with self-organised assembly of near-to-chaos criticality and resolutions of it. These dynamics are in turn developed within one of two broad classes of mathematical models, either a set of coupled partial differential equations or a multi-component/multi-agent network, a third option being model-free machine learning representations [Hooker-a, this volume, note 104; section 6.2.4 below]). This development calls for the articulation of a new paradigm adequate for the de-facto deployment of complex systems models throughout the sciences and a correlative articulation of a new philosophy of science.

Handbook of the Philosophy of Science. Volume 10: Philosophy of Complex Systems.
Volume editor: Cliff Hooker. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

I cannot do this systematically in any deep sense. Although there are plenty of fragments of a mature framework coming into focus, as the essays herein uncover, it is impossible to present a finished product at the present time because of the immaturity of the field. At this stage in its development much of the substance of a new framework lies more or less implicit in domain-specific detail and a major goal of the essays herein has been to make some of that explicit. In what follows I try to briefly limn some of the more general and prominent elements of the emerging scientific paradigm and correlative philosophy of science for a complex systems world. In fact, the apparently disparate scientific domains unsurprisingly do turn out to reveal shared systems-related problems in experimental design, error treatment, and other methodological issues, and just this is what the essays herein begin to reveal. In short, complex systems methodology and philosophy promises to form a substantially interdisciplinary study, as the heirs of general systems theory and cybernetics might have hoped. And scientists work toward the same goal for the sciences.

5 ELEMENTS OF A NEW SCIENTIFIC PARADIGM FOR COMPLEX SYSTEMS

Although much progress has been made in the last 20 years in understanding and using complex systems, any current discussion of a scientific paradigm unavoidably takes place between, and responds to, two classes of limits on our knowledge. On the one side is the absence of general mathematical foundations. On the other side lie practical knowledge constraints on applying complex systems models. These will be briefly explained before going on to explore some aspects of the emerging paradigm.

5.1 Two classes of knowledge limits

(I) *L*imits on mathematical foundations. Most fundamentally, there is a relative absence of any developed general (i.e. encompassing) mathematical framework for complex systems features that would correspond to, and replace, the sophisticated framework of analytical dynamics that supported the preceding classical paradigm [Hooker-c, this volume]. In illustration, what follows is a quasi-systematic summary of those knowledge gaps, derived from the section 3 feature list. Note that we are concerned with current limits on general system-level foundations here; many individual systems have been studied using (typically partial) mathematical models. This is not surprising because the basic laws of interaction still apply among suitably fundamental (typically chemical) components. However, these cases lack a framework of general systems principles.

(I.1) There is no general account of the necessary properties that would be needed to characterise any complex dynamics, or which extensions of these are severally sufficient for complex dynamics. We currently posses only the rudimentary beginnings of even a qualitative general mathematical account of the kinds of dynamics non-linearity can induce and the conditions under which they occur. For instance, and most generally, though we know that non-linearity is necessary to complex behaviour, we also know that it is not sufficient (e.g. 1-body planetary gravitation). Though we know of classes of cases where it suffices, we have no general characterisation of the sufficiency boundary. Again, though non-holonomicity of constraints [Hooker-a, this volume, section 3] is clearly central to manifesting the complex dynamics that arise from circumstances where constraints are influenced by the dynamics, non-holonomicity is neither necessary (n-body planetary gravitation) nor sufficient (contained gases) for complex dynamics. We evidently currently lack any general mathematical account, even qualitative, of the kinds of non-holonomic constraints there are and the kinds of dynamics they can induce.

(I.2) Turning to more specific, but still broad, properties, though thermodynamic irreversibility is neither necessary nor sufficient for complex dynamics (respectively, compound pendula, expanding gases), it captures perhaps the largest domain of complex systems. Yet, though general irreversible thermodynamics has begun development, we currently lack any general mathematical dynamics for irreversible processes and for characterising their dynamical stabilities and instabilities in terms of producing complex behaviours.[1] Further, there is no simple criteria for the character and severity of unpredictability resulting from dynamical instability (e.g. [Bishop, this volume]), though we have increasingly useful tools for characterising complex trajectories in particular cases. Indeed, though we know the available dynamical symmetries and progress has been made characterising the kinds of bifurcations there are [Hooker-a, this volume, note 41], we generally, perhaps wholly, lack detailed dynamical models for the shifts across them. Given their characterisations, we thus also lack general understanding of self-organisation and of emergence, specifically of when a relatively macro constraint is formed, and to what the 'self' refers [Hooker-a, this volume, note 42]. Whence we also lack any general mathematical account of the kinds and conditions of path-dependence and historicity, and of constraint duality and super-system formation.

(I.3a) There is as yet no useful general theory of global coherence constraints or of the forms and conditions for their multiplexing and multitasking designs. We have to proceed empirically, case by case. Underlying that, global,

[1]Perhaps the most sustained discussion of the issues is that by Prigogine — see [Hooker-a, this volume, note 10]. At one point Prigogine [1980] even suggested modifying the Schrodinger equation in order to accommodate irreversible dissipation. In the light of [Hooker-c, this volume, 4.2] this is at least a legitimate exploratory response, if unlikely to address the fundamental underlying issues.

spatio-temporally extended constraints, like autonomy, are at present not representable in differential equation/ phase space formalism. These modelling resources, powerful though they are for modelling the energetics of processes, do not explicitly describe the physical organisation of the system — a metabolic cycle and a pendulum, for instance, may be modelled as equivalent dynamical oscillators. In a phase space only the global dynamical states and their time evolution is specified (through a Lagrangian formalism), not the organised processes that produce the dynamics. The reverse engineering problem of specifying organisation (as distinct from order) from dynamics is currently unsolvable (cf. note 72). In this sense current dynamical representation cannot capture either global coherence constraints or organisation, and hence also not the mechanisms that realise them. Further, while global coherence constraints, like any dynamical constraint, must in principle be representable as a limitation on system accessibility to dynamical states and hence as a hyper-surface in system state space, there is at present no modelling methodology for constructing those accessibility limits for inherently global, spatio-temporally extended constraints like autonomy and the mechanisms that realise it. Nor, in its present form, does network dynamics fare any better since current networks are defined by local none-node relations, leaving them as ill-equipped as phase space to represent organisation and global constraints.[2]

(I.3b) This in turn means that we lack a general or foundational understanding of biological organisation — whether cellular, organismic, communal or ecological — and of autonomy in particular. This poses a serious challenge to theoretical biology, since autonomy and organisation are arguably the most fundamental features distinctive of biology.[3] We similarly lack a general or foundational understanding of social organisation, e.g. cities, business organisation (in particular, firms), military organisation, and so on. Further, we currently lack any general account of the kinds of modularity and hierarchy that occur in complex systems, especially in relation to multiplexing and multitasking designs, and of the conditions under which they might arise. These knowledge gaps in turn mean that we have at present no general understanding of functionally organised spatial differentiation, nor therefore of the correlative subtle constraints that spatial

[2]There is, however, the possibility of defining more complex relations among nodes, e.g. imposing a condition on the nodes completing closed paths encircling some specified node. These kinds of constraints may or may not prove useful in modelling global constraints and functional organisation, but to my knowledge have yet to be explored.

[3]The recent development of systems biology and synthetic biology [Hooker-a, this volume, note 18] has begun to address these issues for sub-cellular dynamics through explicit, iterative cellular modelling. There is much promising progress, and aspects of some simple systems, e.g. commercial yeasts, seem well understood, and well modelled. But this is an indirect, 'reverse engineering' approach for which any theoretical principles have to be extracted in retrospect from the diversity of displayed model dynamics, a difficult, partial and risky process (see also note 72 and text, 6.2.5 below).

> process coherence places back on functional organisation — except in the simplest cases of industrial machine and process design. Nor therefore can we relate use of spatial organisation to modularity and hierarchy in relation to multiplexing and multitasking designs.[4] And we similarly lack any general account of the kinds and conditions of multi-scale and multi-order functional organisation.

This completes the survey of current limitations to knowledge of general mathematical foundations of complex systems dynamics. It bears repeating that these limits refer only to having general (i.e. encompassing) accounts. Many individual systems have been successfully studied, and it is reasonable to expect that these studies will be increasingly joined up over time, at least to within the limits of idealisation (see section 6.2.5 below). But this remains to be seen, and meanwhile science needs must operate substantially without general systems guidance.

Before enumerating limits on applying complex systems models in practice, consider as reference dealing with the dynamically simple systems that are the staple of 'textbook' study and practice, e.g. simple pendula and single-planet motion. There are evidently no significant limits on knowing them. In each case the correct choice of interaction interrelations and constraints is fully specified, hence the appropriate variables (degrees of freedom) and differential equation (d.e.) models are known, as are the order parameters (that characterise the general form of the system dynamics), and it is clear what measurement sequences are required to empirically fix parameter values and initial conditions. These are the conditions that give rise to simple, crisply formulated experimental paradigms for isolating the laws of the system. For instance, fix or calculate the initial (including parametric) and constraint conditions so as to isolate a primary contributing law and determine it through measurement, repeating for all primary laws as required. Then treat secondary laws similarly, as small perturbations of primary laws. Even if 2 planets are involved, e.g., then each may first be treated separately, revealing primary Kepler-law motion for each, and then their mutual interaction can be revealed as small perturbations of the primary-law orbits. In this case, there proves to be a single underlying law for all the interactions (the law of gravitation), so the analysis has a particularly neat, integrated resolution. The success of this kind of approach with simple systems encourages its extension to what has been called 'Galilean method' (e.g. [McMullin, 1985]): differentiate out the manifestation of basic laws, singly and in concert, from secondary laws, treating these latter as small perturbations of the former to form the basic dynamics; differentiate the basic dynamics from other complicating factors such as biases and residual

[4]In the domain of biological evolvability, e.g., a prominent theme is that trait modularity - that is, mutual trait independence - supports lineage evolvability since individual trait variations don't affect other traits (e.g. [Schlosser and Wagner, 2004]). However, this has to also be reconciled with the evidence that the major evolutionary transitions [Maynard-Smith and Szathmary, 1995] require complex coordination of multiple constraints. It should be noted that object modelling in computing is now in widespread use and this not only provides tools for exploring these organisational issues but also forces investigation of effective organisation of complex functional systems. (Thanks to David Green for this point.)

perturbations (including noise), dealing with these latter as corrections to the basic dynamics according as methodological opportunity and practical motivation demand. This is sensible methodology, highly successful in simple circumstances and an obvious first option when faced with understanding any new system.

But already Cartwright [1983; 1989] finds a fundamental problem here. The issue is that the more general the basic laws sought the further their operation is from the complicated 'messy' behaviour actually shown and the more extreme are the modelling assumptions required to differentiate their effects from all the other factors, e.g. assuming the idealisation of frictionless motion to remove dissipation and thermal noise. The result is that generality and so explanatory power is held to be in tension with phenomenological empirical adequacy, the more of one the less of the other. Ultimately basic laws are threatened with empirical irrelevance, since empirical adequacy no longer confirms them [Rueger and Sharp, 1996]. There is no doubt that complex systems compound the methodological situation here. But it is important to distinguish the respects in which this is so. Complex systems can complicate the issues; e.g. time series data for complex dynamics (especially chaos) can be difficult to distinguish from noise, requiring much longer data runs (often too long for practicality) and sophisticated statistical testing (cf. [Rickles, this volume]). And quite generally, complex systems force the issues of dealing with partial and idealised models, issues that also arise from Cartwright's general concerns. For systems like cities and organisms, there is no workable option but to investigate partial aspects of them one at a time, often of necessity using knowingly simplified models. For instance, processes at one level or scale or category might be modelled, treating all cross-level/scale/category impinging processes as non-basic factors. The best of holistic intentions cannot avoid these investigations resulting in partial knowledges that only imperfectly unify. But while they complicate it, of themselves complex systems do not alter the general methodological issue of dealing with non-basic factors, that is, of the limits, if any, of the Galilean method. These are problems, if problems they be, that must anyway be dealt with by any general scientific methodology and epistemology.[5] They will be set aside here as beyond this book's scope, except to the extent they impinge on other complex systems issues, e.g. concerning scientific idealisation and unification (see below and section 6.2.5).

Rather, of relevance here are cases where complex systems pose additional, distinctive limits to practical application. Before turning to these note that, perhaps contrary to expectations, the holism per se of many complex systems is not neces-

[5]In fact [Rueger and Sharp, 1996] argue that systems with complex behaviours, like the Belousov-Zhabotinsky reaction system, constitute a counter-example to Cartwright's general contentions because their basic dynamics can be directly confronted with empirical data (e.g. through dimensional reduction and statistical time series analysis), and in a model-independent way, and yet are also generalisable. Strictly, this response deals with dynamical complexity in behaviour while sidestepping the roles of bias, small perturbations and noise, and to that extent does not deal with Cartwright's challenge. But it does provide a riposte to any claim that it is impossible to deal directly with holistic basic dynamical laws for complex systems. And to the extent that all complicating effects are lumped together under the label 'behavioural messiness', it does illustrate the importance of separating out dynamical complexity from other sources.

sarily one of them. Systems like rolling (Bénard cell) boiling will globally coordinate all convection cells to avoid colliding convection streams, while autonomous systems globally coordinate all their metabolism repair processes to successfully re-constitute the system. In these systems any attempt to isolate component basic laws by confining or freezing a part of them will in general only produce distorted 'laws' (if any) that offer little or no insight into the real principles on which they operate. But if we had assumed otherwise, the advent of holistic dynamics producing complex behaviours that is nonetheless holistically testable (note 6) demonstrates that isolating component basic laws in this way is no part of Galilean method. On the other hand, there is a distinctive complexity limitation here. The underlying usefulness of learning the basic physics and chemistry of the smallest relevant system components treated as isolates, e.g. ions in biology, cannot be directly extended to more complex components where, as in intra-cellular dynamics, their behaviour (e.g. protein folding) is itself a function of the holistic systemic conditions into which they are inserted.

For this last sort of reason, there emerge a number of relatively 'hard' practical limitations on using complex systems models in science. What is meant by a 'hard' limitation? Roughly, one that cannot feasibly be avoided. Whether what lies behind the 'hardness' of the limit is a theoretical, operational or cost barrier, the cases can all be treated as one where cost-benefit considerations, epistemic and/or pragmatic, render it infeasible to pursue the knowledge without undermining the value that it is hoped to realise by doing so. Principled limits derive from finitude and are usually unobvious in dynamical modelling, but nonetheless numerous. Various NP-hard problems — e.g., the 'travelling salesman' problem: find the shortest route connecting N towns once only — furnish one class of examples. Another class of examples derives from systems of great logical depth, that is, those systems whose dynamics cannot be computationally simulated any faster than their literal material emergence took. Evolved living systems may be cases in point.[6] More pragmatic limits do derive from more happenstance constraints, such as research funding decision intervals, but where these constraints cannot be met over intervals shorter than most imaginable future human histories they remain

[6]For finitude examples see Cherniak [1986], who has emphasised the surprising restrictions they impose, while for logical depth see note 37. Checking the consistency of a set of beliefs using truth table methods provides a simple logical example: "Suppose", says Cherniak (p. 93), "that each line of the truth table ... could be checked in the time a light ray takes to traverse the diameter of a proton," roughly 3×10^{-23} seconds, a 'super-computer' indeed. Then for a universe of age 20 billion years, or roughly 6×10^{17} seconds, this super-computer can check roughly $6 \times 10^{17}/3 \times 10^{-23}$, or 2×10^{40} lines. But a truth table for an argument of just 137 premises contains 2^{137} lines, which is more than 10^{41} lines, i.e., more than even this super-computer could check running for the lifetime of the universe. Even where parallel computations permit it, still quite small belief sets would require a computer which also exhausted the known matter of the universe, e.g., belief sets of say 2×137 components for a universe of 10^{50} particles and parallel components each of 10^{10} particles. These are limitations in principle, impossibilities holding in any finite, classical universe, they are not simply products of happenstance human resource limitations. Contemporary complexity theory provides many other examples of similarly intractable problems. Cherniak argues that complexity theory also suggests that such results are quite resilient in the face of alternative computational algorithms.

relatively hard. There follows a brief review of complexity-originating issues.

(II) *Limits on practical application.* The polar complement to the absence of fundamental knowledge is the complex of practical knowledge constraints on applying complex systems models.[7]

(II.1) *Discriminating model structure.* While it need not, complex dynamics often requires many parameters and variables, subtly interrelated, to faithfully characterise it. This creates the problem of discriminating the required interaction interrelations, variables and parameters for validly modelling a particular dynamical system. This can be very difficult. First there are obvious issues such as obtaining sufficiently sensitive measurements to discriminate locally very close but ultimately diverging trajectories, or obtaining a sufficiently long run of measurements to discriminate long transient trajectories from slow periodic trajectories or transitions among attractor basins from transitions across thresholds. Carpenter and Lathrop [2008], e.g., show how difficult it is to predict a threshold, even with very good data, because of a subtle constraint: since state parameters characteristically change only slowly near thresholds [van Nes and Scheffer, 2007], "in situations where precise determination of thresholds is most crucial, parameter estimation is most difficult and threshold estimation is least precise."(p.611). It proves necessary to express threshold location values as probabilistically distributed, making explicit their uncertainty. What is 'hardest' here are the relations between measurement discrimination or time periods and knowledge obtained, but for given instrumentation possibilities and feasible data runs, these will translate into 'hard' limits on obtainable knowledge.

Another problem arises when crucial aspects of dynamics can be contributed by coordinated weak interactions, perhaps across long distances, making it very difficult to discriminate these from the stronger interactions that dominate local dynamics. There are 3000+ reactions proceeding in the living cell and their organisation is such that between them each reaction — inputs and constraints — is supported by the outputs of others; but we as yet know far too little about what to expect of the spatio-temporal organisation required to achieve coherent cellular function to know how to search more than very partially for meaningful parameters and variables governing the processes.[8] And the sheer numbers do matter, because as the numbers of parameters, variables and interaction interrelations increase, the number of combinations of their values that need to be explored for

[7]See, e.g., [Mitchell, 2003; 2004; Westerhoff and Kell, 2007; Wolkenhauer and Muir, this volume] for general discussions focussed on cell biology.

[8]See, for instance, the property of respiration, under 'Multi-level and multi-order functional organisation', [Hooker-a, this volume, section 3], and [Bechtel and Abrahamsen, this volume] on Krebs cycles, cf. [Wolkenhauer and Muir, this volume].

valid discrimination explodes exponentially and the likelihood correspondingly diminishes of exploring the relevant ones within time, computational and other constraints.

The converse problem is knowing when a simplified model is valid for the decision purposes at play, whether epistemic or more practical decisions. There is, e.g., no right rule for suitable modelling. A common modelling principle is the 3-level/scale model: a focal level/scale where the main phenomena are located + one level/scale above and below it. But modelling respiration, like modelling climate (see [Snyder, *et al.*, this volume]), involves recognising effects extending across 4 or more levels/scales. In these cases the 3-level/scale principle does not suffice. This 'principle' in fact offers a pragmatic heuristic guide. Ultimately the only valid principle is that the actual dynamics should dictate the modelling. But while we are still learning what those dynamics are, or whenever resources or human understanding dictate simplification, any k-level/scale model we choose will make erroneous approximations and cut-offs. With luck, however, the models will still support further learning. As Wimsatt [2007] emphasises, heuristics must play a central role in any account of workable methods for complex systems that are usable by finite, fallible agents and it is the way errors are handled, rather than the inevitable making of them, that matters. For instance we often construct models known to be false in order to efficiently investigate other aspects of them (cf. [Cartwright, 1983], see also section 5.2 below).

Levins, a great pioneer of developing complex human-ecological dynamical interaction models as a basis for policy formulation (cf. [Allen, this volume]), remarks that model building involves "juggling the partially opposing requirements of realism, generality and precision", so as to produce useful insight and advice [Levins, 1966]. Edmonds and Moss [2005] reverse the standard KISS principle (Keep It Simple, Stupid), presumed where reducing complexity is desirable, to adopt KIDS (Keep It Descriptive, Stupid). They do so on the grounds that in complexity research we do not want to prematurely throw out relationships and data that, initially insignificant, may prove crucial to understanding the dynamics; better to prune complexities later, when we know better how to do it. However, the more data we retain the more complex is the problem of learning the system generating it. There is in fact no right trade-off among Levins' values or KISS and KIDS in modelling, or other similar trade-offs. Rather, the trade-off chosen must intelligently depend on context, on what will likely yield most value, epistemic or practical. The necessity of trade-offs and their condition-dependence are the 'hardest' features, and again they will quickly generate 'hard' knowledge limits under a variety of investigative conditions. Because of the trade-offs, any resulting model is unlikely to be error-free if treated as a globally valid model; instead it must be treated as acceptable within some class of applications, for some purpose, e.g. research or policy formation. If chosen for research purposes, the values engaged will also emphasise intelligibility, empirical accessibility, salience, reliability and the like, while if chosen for practical purposes the values engaged will emphasise risk-amelioration, efficiency, manageability, and the like. Again all these values are potentially competing in context and various of them will have to be mutually traded-off.[9]

(II.2) *Discriminating initial and constraint conditions.* The specific dynamics of any system depends on its initial and constraint (including 'boundary') conditions. This provides model dimensions to investigate in addition to interactions and variables. But for many simpler systems the dynamical form is independent of initial conditions and fixing constraints is straightforward. The dynamics of many complex systems, however, is a sensitive function of initial and constraint conditions, making it important to carefully discriminate both. An important sub-case is where system dynamics exhibit condition-dependent thresholds (that is, bifurcations), a hallmark of complex dynamics. Unless initial and constraint conditions are chosen that place the system state sufficiently near its thresholds to be able to investigate its behaviour across them, they will be hidden from view. Human misuse of ecological systems is replete with examples.[10] It can often be impossible or very difficult to tell from a sampled range of conditions and dynamics whether there are ignored attractors and thresholds. Moreover, system responses to change, whether generated internally or externally, are often a function of one or more of the size of the change induced, the rate of change induced or the total change over time induced. Moreover, the relative importance of these factors can depend on the initial conditions obtaining at the time. All this complicates the identification of thresholds. And once again, as the number of relevant initial conditions and partial constraints increase, the number of combinations of their values that need to be explored for valid discrimination explodes exponentially and the likeli-

[9]In all these cases, once complicated decisions concerning permissable simplifications, errors and outcome values have been decided, as discussed, then various formal techniques may also be of use in deciding among models, e.g. MML (minimum message length) and ML (machine learning) procedures. See respectively `http://en.wikipedia.org/wiki/Minimum_message_length`, `http://alumni.media.mit.edu/~tpminka/statlearn/glossary/` and [Hooker-a, this volume, note 104] and references. The former trades-off overall goodness-of-fit against overall algorithmic model complexity while the latter search for significant patterns in the data. (MML can also be used to optimise within classes of ML devices for a given data set.) Each has significant limitations in this model selection role, essentially because they do not deal with choice among error kinds in erroneous models and between fitting to crucial and superficial data and are blind to organisation. ML also often outputs high-dimensional patterns that cannot be related to physical system variables and parameters (see section 6.2.5 below). This emphasises the importance of the preceding, currently informal, methodological considerations.

[10]A common example is lake eutrophication, the case studied by Carpenter and Lathrop [2008]. As they make clear, one of the real dangers of pursuing optimisation strategies based on far-from-threshold dynamics is that the optimal strategy will place the system near an ignored threshold so that it crosses it irreversibly due to natural fluctuations or allowed measurement uncertainties (errors).

hood correspondingly diminishes of exploring the relevant ones within time, computational and other constraints. Conversely, valid simplification again becomes an issue.

Condition-dependent outcomes are forms of dynamical path-dependence. Path dependencies play important roles in the dynamics of historical systems, such as geo-dynamical and evolutionary histories, ecological and societal successions and individual lifecycles. They are a major contributor to their displaying historically unique paths (dynamical state sequences). On the one hand, these processes manifest so many condition sensitivities and feedback-led amplifications that most changes of constraints and states will lead to distinctive departures from the unchanged trajectory, hence the condition-dependencies. On the other hand, each process is immersed in such a sea of others that it is constantly being perturbed by them in ways improbable as well as probable, creating non-repeating conditions. This complicates the successful scientific investigation of such systems, since science requires repetition of tests, hence of test conditions, while these systems are always in unique, unrepeated states. However, it is their global states that are unrepeated, not their sub-states and this does open a window to scientific testing. Thus while the evolutionary spread of species across the planet is unique, the occupation of islands or other localised ecologies can be more readily repeated and allows hypotheses about the dynamics of the larger process to be tested. Doing so relies on sufficient sequential state data about various partial system dynamics segments to allow in effect for repeated detailed 'runs' of the dynamics to be studied, agreeing on key organisational dynamics but having different initial conditions. This extends all the way 'down' to testing the chemistry of the basic component interactions, e.g. the physics and chemistry of magma flows. It is easy to see how the penetration of idiosyncrasy can limit this approach, ultimately to the 'hard' category, and this will likely prove accurate for much knowledge, e.g. of the detail of complex events like mountain range formation or weathering. On the other hand, the creation of simplified models of some processes, e.g. of segmentation in a rotating ball of cooling viscous fluid to model plate techtonics, and of computerised models of joint such processes, has proven quite successful at revealing the broader outlines of histories, e.g. of geological history.

A grand case of the historical uniqueness problem is the current systemic impact jointly of burgeoning human numbers and technological power on planetary ecosystems and resources. This historical contingency will shift many ecological (and human) dynamical systems across thresholds, changing their dynamical forms, not just their detail. So there are no direct precedents to study when trying to understand it. We are instead forced to study previous partial analogs (e.g. the island dynamics so poignantly portrayed in [Diamond, 1997; cf. 2005; Flannery, 1994]) and attempt to capture the whole in various simplified models extracted from these and other process studies, e.g. [Meadows, *et al.*, 1972; cf. Meadows, *et al.*, 2004] for resource consumption and [Snyder, *et al.*, this volume] for climate change. The uncertainties involved in so doing handicaps us in our two historical races against time to avoid planetary mess: our race to gain knowledge of ecosystems faster than our increasing capacity to perturb them, and our (more fearsome) race to develop more effective institutional self-regulation as we perturb them.

(II.3) *Discriminating data structures.* The preceding decision problems are typically 'information hungry', they require large amounts of sensitively chosen data to resolve. However, discriminating the required data structures within practical resources is often a difficult or insuperable problem. For instance, one kind of limitation is that the necessary data cannot be had with the available instruments because their use would unpredictably disturb the system (e.g. in much intra-cellular research). Another kind of limitation is that deriving from ineliminable measurement errors. For instance, where false or noisy signals overlap true signals there is no error-free detection threshold for a signal detector, only a choice of the relative frequencies of type 1 (false positive) and type 2 (false negative) errors. Such measurement errors may mask or distort sensitive dynamics. A third kind of limitation arises from timing constraints. Many complex systems involve dynamical processes running on different timescales, e.g. oscillations of many different frequencies. Whatever the variables of interest, discriminating these requires data sampling at appropriate rates for each, but available instruments and methods will often not support the requisite spreads of sampling rates, the consequence being the confirming of erroneous models.[11] To these must be added the usual sources of uncertainty found in finite measurement resolution and finite computational resources. All of these sources of discrimination uncertainty are then amplified over time by nonlinear amplification processes working off dynamical sensitivities to initial conditions. It should also be borne in mind that, while the point of data is to identify the system dynamics, knowing the type of underlying dynamics is often necessary to distinguish the status of data yielded by particular measurement processes, especially as between data concerning basic processes, data concerning initial/constraint conditions to which the system is sensitive, data representing merely dependent detail and data representing noise. This does not in itself make the resulting methodological circle vitiating, but it does demand a subtler multi-instrument, mutually iterative

[11] If, e.g., only a middling rate of sampling is used or available then fast processes will be averaged over and effectively disappear into their mean values while slow processes will only be sampled for part of their variation and so equally disappear into open trends (including constancy). Similar considerations apply to spatial sampling in case the different processes are spatially organised. If the faster processes are too fast for current instrumental technologies, and/or the slower processes are beyond the practical investigatory time horizon (e.g. longer than the lifetimes of related policy decisions, funding or research careers), common conditions, then these processes cannot be correctly detected and measured.

approach.

Two particular cases are worth mention for their distinctive timing constraints. Many systems manifest considerable logical depth (note 37). For these the time required to obtain the relevant state and process information to make a prediction is in principle large and may be much longer than the time horizon for relevant prediction and action. Whence one's interaction with them is always on the basis of some uncertainty. Slow, long running developmental processes in ecological dynamics or human personal development can produce surprising behaviours that defeat even decades of contrary data. A different version of the problem arises when an agent alters its own environment on too fast a time scale for it to know the consequences of its past actions before it acts again. Humans have always been in this predicament and continue notoriously to be so, as witness climate change, peak oil, nuclear proliferation, instability of financial markets and so on. These issues also complicate the identification of non-random distributions of variables, since variables may change their distributions faster than they can be investigated, e.g. in changing social fashion or preferences for housing (say, following hurricane Katrina in New Orleans). Combining these timing difficulties with dynamical path-dependence produces particularly complex demands for access to sufficient data. These demands, confronted with finite practical data accessibility, soon generate hard limits on learning.[12]

The dual problem to model validation generally is policy support. Every limit to our capacity to validate complex systems models becomes a corresponding uncertainty in how to formulate policy for these systems, from climate prediction and ecological management to medical intervention and social understanding. Both policy goals and policy means are affected: we are uncertain about what these systems will do if we intervene in them and, precisely for the same reason, we are equally uncertain about what about them to value so as to properly support them.[13] Conversely, there are in general many different practical interests in the same social or ecological

[12]For those who like a more fundamental perspective, and can accept the decoherence quantum field framework employed, Gell'man [1997] provides a classification of fundamental sources of uncertainty. These explicitly cover several, and perhaps cover all, those reviewed above, but do not identify them in their working methodological guises, as here.

[13]This latter is not intended to indicate the existence of a special moral or values theory for nature but to make a practical point. For instance suppose that, because we do not understand the ecosystem dynamics, we do not know whether frequent fires or seasonal fire variation represent a useful support for an ecosystem, or instead a constant threat to it (and perhaps both in different respects). Then we do not know whether, or when and how, to augment or suppress them to support ecosystem functioning. This problem recurs for autonomous systems (in the sense of sections 3, 4.1.1). Typically neither creatures themselves nor those that manage them know what their basic autonomy norms are. This is true even of ourselves, since it has required numerous human experiments to make clear, e.g., that underlying our blood-sugar mediated hunger-satiation norms is adequate nutrition, not the sweet and fat 'quick proxies' we have developed in more impoverished circumstances.

system, and so many different simplified models could on occasion be chosen to support policy formation for those pursuits, leaving a patchwork of models, often partially irreconcilable.

5.2 Methodological issues in response to knowledge limits

"... the more we strive for realism by incorporating [in our models] as many as possible of the different processes and parameters that we believe to be operating in the system, the more difficult it is for us to know if our tests are meaningful." [Oreskes, 2003]

The catalogue of epistemic limitations reviewed in section 5.1 might easily lead to scepticism about the usefulness of complex systems modelling to science and public policy. Oreskes, a thoughtful analyst, argues that such models may be confirmed but never validated and urges that they never be used to make long term predictions because these are invariably wrong and thus undermine respect for science and its role in public decision making.[14] Cartwright [1993; 1999], as noted in section 5.1, provides a supporting context, emphasising that all scientific models are approximations at some scale of precision, either because they excise some aspects of the system for modelling (e.g. focusing on one particular level only) or because they simplify through idealisation (e.g. frictionless motion).

However, any temptation to unbridled scepticism needs to be tempered by the knowledge that there has been enormous progress over the past 50 years in developing empirically useful complex systems models in various sciences, e.g. in control engineering, cellular biosynthetic pathways and climate science. Even allowing that there may always be surprises from running processes and that some, perhaps much, of the fine-grained data such models generate is false below some precision floor, there has been such powerful progress as to require explanation. For instance, at least for simple bacterial cells used in industrial production, systems and synthetic engineering can now regularly predict with considerable accuracy the system output changes induced by single gene mutations on the basis of complex models of biosynthetic pathways. Add to that a million other new products that have eventuated from applied scientific understanding of complex systems, from challenging new buildings and aircraft control systems to adaptive fisheries and grasslands management. In sum, it is clear that progress in understanding and application is being made.[15]

[14] See also [Oreskes and Belitz, 2001; Oreskes, *et al.*, 1994].

[15] The literature here already runs into the thousands of products and books and millions of papers and daily actions, literally too complex to review - our economy, and science itself, are becoming complex systems to cope with complex systems. As befits the subtlety of a complex system there are various 'gems' to be found on complex systems hidden away in unlikely websites; here is a very short list to (no more than) illustrate the point: `http://www.cscs.umich.edu/~crshalizi/notebooks/`; `http://www.beijer.kva.se/index.php`; `http://www.cs.sjsu.edu/faculty/rucker/`; `http://www.tjurunga.com/biography/roger.html`; `http://www.complexity.org.au/vlab/`.

The general problem then is to understand the means of progress in scientific understanding and application despite epistemic limitations, and where and why progress has proven difficult. The latter is largely due to the limitations noted in section 5.1. Here the focus is on the former, for which 5.1 also supplied some examples. Two large bodies of literature frame this study but lie beyond its scope, the general use of models in science and policy making on the on side and the relevant methodological detail in particular local scientific modelling endeavours on the other. These will not be reviewed here, the reader is directed instead to their substantial literatures[16] and reminded of the preceding discussion where many aspects were covered in passing. The construction of a methodological paradigm for complex systems knowledge sits between these poles, borrowing a general scaffolding of issues from general modelling theory and generalising lessons from particular studies where possible. In what follows I offer only the outlines of some of the major classes of strategies used in order to illustrate the paradigm now emerging piecemeal within the sciences.

Basic strategies (a) Instruments and thus their data structures can be improved in range, precision, and reliability. In cellular biology, e.g., high throughput technology for quantitative cellular measurements (the 'omics' — genomics, proteomics. . . — that now dominate cellular biology) has transformed the knowledge of biosynthetic pathway structure.[17] (b) Systems can be decomposed in various ways, e.g. into mechanisms, modules, or components, that will simplify their piece-wise investigation and its subsequent integration.[18] (c) Hard problems can be ameliorated in various ways, e.g. computational complexity problems can be solved through heuristic algorithms, ill-defined problems can be attacked through self-directed anticipative learning.[19] (d) Theoretical tools can be improved, e.g. the recent development of data mining tools for analysis of high dimensionality, low density data distributions.[20] (e) Ways to permit repeated measurements can be expanded. (This is helpful because systems whose dynamics can be repeatedly measured — because they are cyclic or because multiple copies of the system dynamics are available, e.g. lake eutrophication — can be better determined by empirical testing.) (f) Techniques for investigating and measuring various key

[16]For an introduction to the general use of models in science see, among many, [Hesse, 1961], and for the more detailed analysis of scientific models see, among many, [Cartwright, 1983; 1989; 1999; Kell and Knowles, 2006; Magnani and Nersession, 2002; Mitchell, 2003; 2004; Mitchell and Dietrich, 2006; Morgan and Morrisson, 1999; Morrisson, 2000; Nersession, 2008], the references offered by Oreskes (note 14) and the essays by [Snyder *et al.*, this volume; Wolkenhauer and Muir, this volume]. For a brief overview see [Frigg and Hartmann, 2006] and for an overview of formal model selection see [Shalizi, 2006].

[17]See e.g. [Fu, *et al.*, 2009; Wolkenhauer and Muir, this volume].

[18]See section 6.1 and e.g. [Bechtel and Richardson, 1993; Bechtel and Abrahamsen, this volume] on piece-wise mechanism isolation and testing in cellular systems.

[19]See e.g. [Simon, 1969; 1997; Wimsatt, 2007] on heuristics and [Christensen and Hooker, 2000; Hooker, 2009] on self-directed adaptive learning and [Farrell and Hooker, 2007a; 2007b; 2009; Hooker, 2009b] on its application to scientific method.

[20]On data mining and other methods, see e.g. [Shalizi, 2006; Rickles, this volume; Hooker-a, this volume, note 104] references.

dynamical features, e.g. thresholds, attractors and panarchy cycling, can be improved.[21] (g) The centrality of computational modelling provides the platform for making new complexes of modelling methods, e.g. the use of game theory, evolutionary theory and computers to form evolutionary games and various other multi-agent interaction models.[22]

(h) The use of model sensitivity analyses and robustness criteria more generally can be expanded and refined. Sensitivity analysis measures the sensitivity of model outcomes to small changes in model inputs and parameter values and sometimes architecture. Since these things will in general only be determinable to within error bands, results that rely sensitively on more finely discriminated values should be treated as more unreliable, and vice versa for insensitive results, unless there is independent reason to treat the model as accurate in those respects. Robustness measures the degree of invariance of some feature F across some class of entities E. Sensitivity analyses are forms of robustness measurement (F = model outcome, E = model inputs/parameters/structure variations). A form of robustness applied to model reliability, it is a form of methodological robustness.[23] Other relevant forms include robustness of data across variation of production/collection conditions, instruments, measurement methods, ...; robustness of outcome to model abstraction (e.g. to simplification by aggregating below some spatial resolution); robustness of external control — not always easy.[24]

[21] See e.g. [Gunderson and Holling, 2002] on panarchy and [Gunderson, *et al.*, 1997; Carpenter and Lathrop, 2008; Walker and Salt, 2006] on thresholds.

[22] See, e.g., [Skyrms, 2004; Harms, this volume] on evolutionary games, [Holland, 1995; Lansing, 2002, p. 285; 2005], Edmonds at `http://cfpm.org/~bruce/`, among many on new multi-agent interaction models, [Humphries, 2004] on computer-aided.procedures.

[23] This is as opposed to robustness mechanisms within the studied systems themselves, that is, forms of systems robustness. Only methodological robustness is discussed here. Systems robustness has its own huge literature, see e.g. the interesting review by [Kitano, 2007] and references. In the uncertainty terminology introduced below, Kitano is discussing robustness as system resilience, suppressive or adaptive.

[24] Well known cases of methodological robustness are the multiple agreeing determinations of the mass of the sun, Avogadro's atomic number and the charge on the electron. On robust control for ecologies see e.g. [Anderies, *et al.*, 2007]. They conclude that getting overall robustness of control of certain fisheries — that is, control that is maintained to some criteria across some range of inputs and characteristic system perturbations - is very difficult, and may not be possible. They suggest linking it to an ongoing learning process — cf. [Allen, this volume]. This latter raises a subtlety to robustness analysis when dealing with complex systems: because of the intricacies of complex dynamics the phenomenon of interest (e.g. crossing a threshold) may not occur robustly by any of the standards that can be applied, or applied at the time. This might happen if, e.g., it requires a very particular alignment of weak forces (as, say, disturbance of an orbit of an outer planet in our solar system would do). Under these conditions all but the exactly right, refined robustness wipes out the effect. Typically, the robustness can only be tested in retrospect when the system dynamics is sufficiently well known. The idea that meaningful tests can only be provided in retrospect is alarming to many methodologists (presumably because it blocks any straightforward 'mechanical' process for getting from an initial investigation to the final truth), but a commonplace in science. Farrell and Hooker [2009] provide a striking example where the effect is central to the development of ape language research. But it discomforts Mayo's otherwise insightful testing methodology [Mayo, 1996].

(i) The expanded use of reduction and unification as methodological tools also deserves elaboration. It covers the broad search for underlying dynamical processes, including mechanisms, guided by hypotheses about the ways in which higher order and higher level phenomena might be produced by these and so reduce to them (see section 6.1 below). One of the simplest examples was already provided by Nagel [1961]: the development of the kinetic theory of gases in physics was promoted by searching for gas process analogues of the terms in Boyle's law. Contemporary neuro-science provides an extended form of this investigation method, with its use of molecular neural modelling, single neuron recording, neural system modelling and neural scanning (fMRI, CAT, etc.). In both of these cases the result is the progressive modification of both accounts to expand the match-up mappings.[25] While the Boyle law case occurs within physics, neuro-science is already more inter-disciplinary and the outcomes of its mutual adjustment processes will bring more unity to a larger slice of science, a slice already extending through the psychology of decision making to economics and sociology, and the like.[26] As these examples show, pursuit of these methods can often guide research productively.[27]

These are fragments of a basic methodology for the investigation of complex systems. There are large methodological literatures underlying them in various forms within the sciences, e.g. on the mathematical characterisation of bifurcations, on system identification and control in control engineering and on data analysis in machine learning and statistics (see section 6.2.5 below). All these have yet to be fully integrated with scientific practice and are themselves still developing. Taken together, these fragments and others like them already present a powerful basic set of tools that explains the considerable contemporary progress in scientific knowledge of complex systems despite the epistemic limitations involved. And they will surely be increasingly extended, elaborated and integrated as methodological experience and understanding, and communication of it across the sciences (currently sparse), increases. Their reach as an integrated paradigm cannot yet be limned. Just as heuristics 'get around', albeit more riskily, the difficulties posed by computational complexity without denying that complexity, so it is not clear (to me, anyway) where the real limitations lie in the development of this methodological paradigm. Like the systems it studies, its future is open.[28]

---

[25]Though often distorted into various formalisms (e.g. the 'bridge law' form), this mapping match-up method (what is now called the 'interpretation' model of reduction) is in fact the original orientation of Nagel's model of reduction - see [Hooker, 2004, Part 5; 2005; Hooker-c, this volume]. For mutual adjustment to expand reduction in gas kinetic theory see [Hooker, 1981, Part I] and references.

[26]However this will not be any simple deductive unity of the sort to which Nagel was restricted by his time period, see section 6.2.6 below.

[27]Wimsatt, a pioneer of exploring this methodology, provides a rich discussion of these methods [Wimsatt, 2007] while Bechtel and Richardson [1993], e.g., apply them instructively to cell biology.

[28]A fitting part of the complex adaptive system that is science itself — see section 6 below. Likely only complex systems can represent other complex systems. More certainly, only a complex adaptive system process can allow communities of finite, fallible agents to investigate complex

However, the epistemic limitations earlier canvassed still do apply and working around them is seriously incomplete; there are many situations where full and detailed prediction and explanation are not, and may never be, possible. A useful strategy in these circumstances is to look for less epistemically demanding, but still methodologically practicable and epistemically or policy-wise useful, tasks. These are found in the pursuit of less dynamically specific properties.

Shifting to less dynamically specific methodological features A simple example is the consolation pronounced by Snyder *et al.* [this volume] that "Where state prediction fails, complexity prediction works" (because the complex features that lead to constraints on state prediction are themselves evidence of complex processes at work). 'Is dynamically complex in manner X' is a property abstracted from the detail of the complexity-producing processes involved, and that freeing from detail can make the prediction of its occurrence much less epistemically demanding. Inference from 'fat tailed' statistical distributions to the presence of complex dynamics provides an example.[29] Another strategy is to look for empirically accessible crucial tests decoupled from detailed prediction. Again Snyder *et al.* [this volume] provide a nice example. Climate models are complex (many scales, parameters and interaction coefficients) and much of their detail is methodologically inaccessible. Nonetheless, it is possible to isolate an empirically testable tell-tale indicator of the anthropogenetic origin of global warming in 'green house' gases emitted from human processes (namely, the vertical location of a temperature 'hot spot' in the atmospheric vertical temperature gradient). The presence per se of that hot spot is unspecific about the details of the dynamics that generated it. A similar strategy consists in locating dynamical bifurcations by matching parameter values across the threshold.[30]

A final option is to identify various dynamical possibilities for future system behaviour — in effect, by exploring the (geometrical) structure of its threshold delimited attractor landscapes. For instance, to determine that from a given class of initial conditions both a limit cycle attractor and a bifurcation are future possibilities, depending on the perturbations received, it sufficies to produce these different futures, measuring some sufficiently identifying feature of each and the relevant perturbations. In a paper on bumble bee colonies that pioneered modelling numbers of individual bees (today known as agent-based modelling), Hogeweg and Hester [1983] produced a colony model that shows emergent behaviour strikingly similar to real colonies, including differentiation of elite and common workers, switching to production of new queens and queen expulsion. The importance of this model lay, not in any detailed predictions — these were not possible because of emergent behaviour dependent on historical conditions, but in setting out the

systems.

[29] See, e.g., [Vallis, 2009; Rickles, this volume].

[30] The location of the threshold for forming rolling boiling (Bénard cells), e.g., is given by equating suitable parameters each side of the bifurcation, without having to specify any further detail to the fluid motions [Chandrasekhar, 1961].

structure of possibilities for colonies and their general interrelations. These possibilities, as noted, could be empirically observed. But the nature of their genesis, and therefore also the prediction of such general features and options for their control, could not be understood until their possibility space was revealed.

Similarly Meadows [2005], reflecting on the significance of the pioneering Limits to Growth model [Meadows, *et al.*, 1972], describes it as an exploration of possibilities highly disciplined by the demand for quantitative modelling (and therefore required to be precise about the forms of interaction, if not the numerical details) and compatibility with available data.[31] But of course it was precisely the prediction of the general conditions for selected classes of possible trajectories that made the model rightly famous, it is these and not more precise predictions that were then most useful for policy orientations. In addition, the models usefully guided the directions of further exploratory research and model refinement.

Shifting to a higher order methodological focus: resilience It is clear that there are a great many possibilities open to complex systems and that, for all the reasons already canvassed, there is in general considerable and unresolvable uncertainty about which of these any given actual system will in fact realise.[32] While the possibilities for responding to the epistemic challenges this poses have been canvassed above, there is available a further policy/ management response which has not yet been discussed: acquire relevant resilience. In fact this is also the key to responding to known but unavoidable risks and in what follows these are assumed included under 'uncertainty'. Essentially a functional notion, adaptive resilience is the capacity to survive functionally intact the perturbing realisation of uncertainty.[33] Set aside the option of inaction - thus gambling that nothing bad will happen - because it is irrational except as a last resort. Then there are just two kinds of ways to respond to uncertainties: (1) with suppressive resilience: prevent their occurrence or remove their effect if they occur, or (2) adaptive resilience: be anticipatively prepared to adapt to their occurrence. For example, in response

[31]Somewhere, Donella Meadows, reflecting on the development of the systems modelling work, remarks: "Their model was a research tool to extend their imagination — they could imagine a structure of global connections but needed a numerical model to calculate actual consequences of these imagined interactions. Such modelling is an explorative endeavour requiring both creativity and critical judgment, all constrained by the need to be consistent with real-world observation and computational feasibility. The act of committing ideas to code and running realisations provides a useful discipline, and indeed such quantification is the only way to study consequences of such interactions, such is the complexity and analytical intractability of the equations. Thus quantification is required to realise consequences of theory, but nor is it a complete exploration of the consequences of theory, so for scientists this art is more than a naïve handle-cranking exercise. The model was indeed a successful research tool, and has persisted in time as a valuable learning tool for global system dynamics modellers."

[32]See e.g. [McDaniel and Driebe, 2005].

[33]But there may be changes during, and sometimes after, the perturbation effect, so long as they don't affect essential functionality. Of course, for clear application we have to decide what the relevant functioning entity is, what functions have to be preserved, and what counts as sufficiently preserving them. Here these matters are presumed resolved - but note that they may become crucial in various practical contexts.

to the uncertain realisation of war, respectively, sign a peace accord (prevent), insert a barrier (remove), learn self-defence (adapt). The human immune system provides us adaptive resilience to the uncertain arrival of many disease pathogens: they make us sick (the perturbing realisation of uncertainty), but temporarily and we recover our healthy functioning. Adaptability is the capacity to adapt, it is what is needed to realise adaptive resilience.

The natural world uses a combination of gambling and resilience — it gambles wherever there has not been sufficient selection for resilience. And it uses a combination of forms of resilience, sometimes aiming to prevent risks — principally through niche specialisation and camouflage — or remove risk effects — principally through armour — but most often by acquiring adaptability, e.g. through adaptable individual behaviour (fight/flight) or through species diversity. Niche specialisation and armour illustrate the limitations of suppressive resilience: it quickly limits other options (freedom of movement, etc.) so has a narrow success spectrum and it is often fragile (breaks down rapidly and catastrophically) and costly to maintain. Adaptability, by contrast, encourages further synergistic options, like learning while creating new behaviour; such synergies have been a principle factor in the evolution of intelligence. The upshot is that in situations with significant uncertainty, acquiring adaptive resilience is typically the best system viability defence against uncertainty. Moreover, it has the advantage that realisations of uncertainty don't need to be predicted to achieve a resilient response - indeed, with a second-order capacity to develop (first-order) adaptive options, such as our immune systems provide, we don't even need to predict the specific kinds of uncertainty realisation to achieve a resilient response.[34] (This is valuable since uncertainty realisation is inherently unpredictable.)

However, adaptability is always realised as infrastructural redundancy,[35] so it competes with immediate increased efficiency. Spare parts, e.g., remain a pure cost until called upon for use. Human brains, that provide so much of our adaptiveness, take ∼30% of our energy budget to run. As uncertainty diminishes in importance the competition between investing in adaptability and in efficiency means that how much resilience, and of what kind, should be acquired, and when, is an uncertainty

[34]All this provided ramping up the actual adaptive response post-realisation is fast enough to prevent lasting damage to functionality; if not, then, the resilience system has to be supplemented with some capacity to anticipate (predict) uncertainty realisations and/or, like piles of spare car parts, the redundancies have to be already present and ready to operationalise.

[35]The redundancy is what provides the response options. The most concrete, specific forms of adaptability are functionally similar (or identical) spare parts, whether as ecological biodiversity (spare species), sector diversity (spare firms) or within-firm stock piles (spare parts). Its most abstract, generalised forms are capacities to replace what is damaged, whether a generalised economic capacity to do so (insurance of all kinds: spare cash) or a generalised ecological capacity to do so (soil with seed load: spare regeneration). In between lie varying forms of operational adaptabilities, from very specific, 'coal face' ones (e.g. capacity to switch among drill bits, or fuels) to higher order, more generalised ones (e.g., and successively, for a firm the capacity to switch mode of production, or markets, or decision organisation). The more abstract the redundancy the more reliance there is on the functional capacity of the environment, e.g. insurance assumes that in the economic environment money will suffice to procure the replacement of what was destroyed.

management issue. (It is normally called a risk management issue, but this label can obscure the importance of uncertainty.)

Thus promoting adaptive resilience yields viable systems compatibly with reducing prediction requirements. The need to predict their behaviour is reduced to predicting their relevant forms of infrastructural redundancy and its performance, with reduced or no need to predict uncertainty realisations. Fire insurance, a financial resilience product, is as cheap and simple as it is because it works without the need to predict particular fires and their details. On this basis an argument can be made that, in a self-transforming (self-organising) and uncertain world, the core of sustainability — which is concerned with sustaining value over long periods — must be resilience. (It follows that present common conceptions of sustainability, like negative impact reduction or natural capital maximisation, are inadequate or worse [Brinsmead and Hooker, this volume].)

Integrated design driven by stakeholder values A final way to reduce prediction requirements grows more directly out of applied experience. Where the management of a complex system (e.g. a water shed) by multiple interests (e.g. governmental departments, private industries, community interests) is involved, then the systems models used have to both capture the system complexities, typically requiring multi-disciplinary characterisation (e.g. ecological, industrial, economic, cultural) and the complex patterns of condition-dependent reinforcement and competition among the many values in play along with a variety of model performance features, e.g. intelligibility, salience, reliability and tolerable risk. Successful studies are required to integrate all these dimensions with a set of management policies and tactics and functional relations with the stakeholder groups. Experience with successful studies of this kind suggests that the following is best procedure.[36]

1. Gather stakeholders early on in the design process, elicit their value structures.
2. Initially employ broad system models and policy and technical goals.
3. Introduce broad strategies designed to accommodate the most important values involved (e.g. dominant, widely shared values), and set their performance criteria.
4. Develop the descriptive partial models used to reveal salient patterns and trace causes, and the partial strategies proposed to achieve set normative criteria.

[36]See [Brinsmead, 2005; Brinsmead, Hooker, 2005], based on a review of the most successful Australian integrated management projects. Of course, different values structures may lead to differing models and strategies. These can be regarded as differing management 'perspectives' available to interact with a complex system in a complex world. But these do not yet support any deep disunity; they do so only if there is added a normative assumption of the inherent disunity (incomparability) of values. See further section 6.2.6 below.

5. Iteratively narrow down to specific indicators, models and tactics, and relevant reporting and ongoing adaptation.

This process is superior from the perspective of the stakeholders who must ultimately support its management proposals. But the point to be made here is that it is also superior from the perspective of the modelling professionals involved because it reduces predictive requirements to just the collection of features and dynamical ('causal') pathways required to address management issues.[37]

5.3 Towards a new scientific paradigm

Explanation: from statics to process dynamics. The impact of complex systems ideas and principles on the sciences is most accessible where it has barely begun and mathematical models are less used. In what follows the response to the entrance of complex systems into archaeology reflected in recent writing, especially by van der Leeuw, provides an instructive illustration.[38] For Archaeology the change represents something of an intellectual shock. For instance, the thought that, rather than being deliberately invented, human institutions may emerge as the unintentional by-product of myriad disparate individual coping actions (cf. [Hayek, 1973]) introduces a very different conception of them, including that:

- actions come first, the institution to which they subsequently conform, later;
- institutional strength is typically the expression of many weak interactions rather than a few strong ones;
- diversity and myopia, not conformity and vision, are most often central to the process - in particular, any institutional competencies wider than the myopic individual actions from which they arose must find their roots in intra-communal diversity;
- ecological niches are communally created, not found.

Lansing, in the related discipline of Anthropology, has a similar view: "systems composed of heterogeneous agents with limited knowledge of their environments can exhibit emergent properties of order", correlated social-environmental organisation can emerge in this way and result in niche creation [Lansing, 2002, pp. 283-7; cf. 2003; 2005; this volume]. The eternal tension between the importance

[37]In this connection, note that the points of greatest policy leverage discussed by Meadows [2005] cover the key higher order features of complex systems discussed under 'less dynamically specific methodological foci' above. On the other hand complications ensue when the managed complex system contains organisms sufficiently intelligent to set their own goals and modify them as circumstances change (say, from mice 'up'), for then the setting of management values is not the only site of evaluation and goal setting and successful policy must integrate these additional agent processes as well.

[38]See [McGlade and van der Leeuw, 1997; Van der Leeuw, 2004], cf. [Kohler and Gumerman, 2000].

of individuals and that of the collective in understanding communal properties and behaviours remains, but here it is crucially reformulated: (i) the individual may be myopic rather than lead and still be essential to the emergence process and (ii) rather than any particular actions, the communal dynamics may be dominated by underlying diversity (e.g. in process rates and fluctuation statistics) even in those cases where leader roles emerge.[39]

It is not surprising then that van der Leeuw nominates a new *explanatory paradigm* as the major focus of the shift to complex systems modelling. For van der Leeuw, the traditional pre-complexity paradigms of ancient history are characterised by chronology, continuity and coherence (uniformity of settlement kind, etc.) and stasis, allowing construction of 'seamless narratives' of societal stage formation. or of the sequential change in societies over time, the ideal theoretical models.[40] This ideal societal model is expressed in the Perfect Saltation Staircase [PSS]: historical development comprises an ordered sequence of optimal stases (equilibria), forming the steps in the civilization staircase. A step does not require further explanation of its stability. The steps are interrupted by sharp, explanatorily disjoint changes — coups, natural disasters, etc. — that result in ascending from one step to the next. What explains the sharp changes and the progress they cause, and what is their social status, is left to one side as outside the focus on stasis and optimality. The explanatory focus is on stasis/equilibrium, uniformity and optimality. Disorder, discontinuity and transformation have no analytical place in this scheme. (Cf. [Schmidt, this volume] on the pre-complexity paradigm for physics.) There is a version of this in ecology and landscape science, focusing on climax ecological stases and 'punctuated equilibrium' stratigraphic histories of these (cf. [Green, *et al.*, 2005]. A different version of this explanatory ideal, emphasising dynamical process stasis, is Constant Expansion/ Differentiation [CED]: continual development, not requiring further explanation, leading to increasing size and differentiation. What generates increasing size and differentiation, especially differentiation, and what is its social status, is left to one side as outside the focus on continuity and coherence.

By contrast, the new paradigm is focussed around the dynamically emergent and adaptable society: societies are conceptualised as flexible, internally diverse, interactive, both internally and with their environment, and far-from-equilibrium, hence open to their environment. They survive by adapting, in particular through emergent behaviour and social organization. Potentially, both the unexplained sharp changes of the PSS step model and the unexplained differentiation of the

[39]This last can remain true compatibly with particular actions (not necessarily by leaders) leading to general outcomes, that is, raising the particular fluctuations that in fact become amplified and fixated in the community as historical constraints.

[40]There is a tantalizing suggestion that this orientation is motivated by the need to maximise the informativeness of sparse data — the explanatory constraints imposed by chronology, continuity and coherence are relatively severe and hence, relative to looser explanatory constraints, provide for fewer possible social trajectories from which sparse data is to isolate the actual trajectories.

CED model can be explained as self-organised transformations.[41] Lansing in Anthropology also emphasises that the dynamical re-conception of basic concepts, including adaptation as inherently a global (that is, communal-wide) self-organised emergent, provides the capacity to transcend static equilibrium theories while still providing social explanations [Lansing, 2002, pp. 283-7; 2003]. In short, this is a shift from socially characterised narratives to dynamically characterised processes. The ideal is to find a characterisation of societal dynamics — in a domain and period, say 1st and 2nd millennium BC Rhone valley settlement — whose associated interaction dynamics generates a dynamically coherent chronological model of its societal trajectory using only general complex systems modelling concepts, principles and methods.

Change is the normal condition and stasis now requires explanation. But equally requiring explanation are all the specific kinds of change that occur, whether recurrent (cyclic attractor), stasis (point attractor), (dynamically)stable macro-level constraint or bifurcation.[42] For van der Leeuw, the explanatory focus is now on, not stasis/(static)equilibrium, uniformity and optimality, but on (i) difference — disorder, variation ('aberration', including singular behaviours: "idiosyncrasy" and structured fluctuations: "stochastic risk taking behaviours"), spatial inhomogeneity, discontinuity and instability, and (ii) sub-optimality (includes roles of non-adapted and error-making behaviours). These features are seen as serving larger functional purposes, e.g. increasing societal resilience (through access to a range of presently sub-optimal responses to respond to environmental changes). Correlatively, Lansing accents the shift from stability to robustness in human-ecological accommodation [Lansing, 2002, pp. 287-8].

As van der Leeuw (and Lansing) recognises, this explanatory shift brings with it correlative shifts across the board. Here the main categories mentioned are briefly reviewed. *Salient data.* Pre-complexity method focussed on decreasing diversity of data by pooling data (e.g. across settlements) and processing for what is common, ignoring the character of the local environment, and so on. Post-complexity method emphasises acquiring sufficient data to be able to distinguish the specific dynamics that apply, e.g. to distinguish seasonal farming variations due to changing local soil properties, or distinguish disruptions caused by singular shocks — a flood, a local individual's rebellion — from disruptions caused by bifurcational re-organisation, e.g. as population growth causes crowding to increase and/or soil productivity to decrease. Lansing remarks that the emergent nature of

[41] This contrast is perhaps easiest to appreciate in the PSS model: although each step could be understood as a temporary dynamical stabilisation altered through self-organisation, the focus on the separate character of each step and the way it determines the step sequence gives an conception of step structure that is independent of dynamics - cf. the way in which Piaget ultimately came to characterise his developmental stages purely logically thus unhappily separating them from developmental dynamics and so undermining his greatest insights (cf. [Hooker, 1994c]).

[42] At various places van der Leeuw says that the shift is to explaining stasis instead of change, to the study of the change of change and to "the relation between the recurring and the unique". If these characterisations are read as referring to specific aspects of the dynamical modelling, rather than taking each as a universal claim, and are suitably interpreted in dynamical terms, then these characterisations are covered by the foregoing formulation.

institutions, and especially their consequent adaptive functions, may be invisible to standard anthropological data collecting and description theory, for instance (and notably) in the Bali water temple case [Lansing, 2003, pp. 198-9; cf. this volume]. *Systems dynamical methods*. Here belong the complex systems tool box of analytical methods for identifying and validating/ confirming dynamical patterns and processes discussed in section 5.2 above. *Research context*. This is now systematically expanded to include multi-disciplinary contributions on the basis of shared dynamical investigations, including learning and cognition (e.g. in niche construction).

Condition-structured knowledge Certainly there are new kinds of knowledge, specifically knowledge of the diverse dynamical processes and conditions generating social phenomena. Lansing speaks, e.g., of a new way of investigating the fundamental nature, types and conditions for functional integration in a society [Lansing, 2002, p. 287]. In addition, van der Leeuw has in mind a larger epistemic feature that he calls 'poly-ocularity'. I take the central metaphor to be the necessity of multiple foci for knowledge, and understand that it claims to be typical of complex systems like societies that they can only be investigated at some degree of system resolution, where a resolution includes at least space and time scales, functional order and scale, and some prevailing constraint conditions. Since the particular dynamical patterns and laws that characterise that dynamical condition are not in general transferable unaltered to other dynamical conditions, complex systems knowledge is structured by dynamical condition. This is a dynamical species of perspectivism arising from the nature of dynamical interaction — and no more. A dynamical perspective of a complex system CS is essentially a partial dynamical model of CS picked out by some specified class of dynamical interactions that other systems have with CS, e.g. a class of measurement technologies, or a specific interaction scale in space and/or time and/or energy.[43] Similarly, Rueger's [2005] perspectivism of singular perturbations is a dynamical perspectivism formed under dynamical asymptotics, a structured set of degener-

[43]A set of natural perspectives are those picked out by those natural technologies, the senses, and correspond indirectly to ecological niche. Another distinguished set of perspectives pick out a specific system level, treating interactions from other levels are constraints or perturbations. The constraints imposed by the class of interactions on measurement resolution, etc., are here the dynamical conditions constraining the partial models. (Note that this is only a special class of dynamical conditions, the general class being concerned with the dynamical constraints applying to any systems dynamics.) The roots of the discussion of perspectivism now to follow lie in [Wimsatt 1994]. There a perspective is characterised as follows: "Perspectives involve a set of variables which are used to characterize systems or to partition objects into parts, which together give a systematic account of a domain of phenomena, and which are peculiarly salient to an observer or a class of observers because of the ways those observers interact with the system or systems in question." The paper provides a dense fount of ideas for addressing the conceptual/methodological issues raised by complex systems, if it is itself somewhat of a 'conceptual thicket'. Here I try, as elsewhere, to have only dynamically characterised distinctions underwrite the analysis. Thus I do not include reference to 'parts' or to 'systematic accounts' or 'observers' per se. Wimsatt is more liberal, though his analyses are always aware of the underlying dynamical complexities.

ate idealisations unified by the single theory from which they derive (see further section 6.2.6 below).

However, the applicable epistemic limits are likewise a function of dynamical conditions, as therefore are valid methods also; so that all knowledge claims must be specified relative to their dynamical validity-making conditions. This yields dynamical condition-structured knowledge - but no more. Van der Leeuw speaks suggestively of preceding epistemologies (nominating positivist, Marxist, structuralist and post-structuralist versions) as "fixed", presumably to highlight their shared assumption that knowledge should form a condition-independent, universal set of laws. To be consistent with the complex system character of societies, and science, they need to be replaced by a poly-ocular knowledge structure, understood to refer to condition-dependence, as above.

It is possible, however, that van der Leeuw has in mind a more radical perspectivism. He also speaks of these epistemologies as suffering from the "Positivist versus Intentionalist dichotomy", however it is unclear how dynamically-read poly-ocularity would alleviate that problem. This suggests that poly-ocularity might instead be intended to somehow include both empirical and personal perspectives, each with its distinctive norms, or perhaps communal versions of these. Even here, if individual and community are construed dynamically (see section 4 above) and there is no added normative assumption of the inherent disunity (incomparability) of values, this will still fall within dynamical perspectivism. There are other forms of perspectivism — e.g. a (too) general epistemic perspectivism essentially asserting only that all knowledge claims are conditional on accepting some set of underlying theories and research history [Giere, 2007] and a radical normative perspectivism based on the added normative assumption of the inherent disunity (incomparability) of values — but these are not to be conflated with dynamical perspectivism, are not specifically relevant to complex systems issues and are not considered further here (but see also section 6.2.6 close below).

Lansing [2002, p.289] observes of these five connected shifts — in explanatory paradigm, salient data, systems methods, research context, condition-dependent knowledge — that, taken together, they begin to shape a new scientific paradigm for the social sciences — and not just for them but for all sciences.

6 ELEMENTS OF A NEW PHILOSOPHY OF SCIENCE FOR COMPLEX SYSTEMS

We now will also need a correlative philosophy of science. Pre-complex philosophy focussed on identification of universal a-temporal causal laws and universal prediction as the carrier of testability and explanation, plus reduction achieved through analysis followed by bottom-up synthesis to closed formed analytic solutions as the method for incorporating complexity into laws. Instead, we now need a philosophy of science focused on dealing with multi-scale partial interdependencies, validity of top-down as well as bottom-up analysis, partial unpredictability, failure of causality as the core naturalist explanatory tool, the entwinement of emergence and reduction, domain-bound (condition-dependent) dynamical laws that accept historically unique individuals (sometimes as the norm), the unavailability of closed-formed analytic solutions and the consequent model centredness and limited knowability and controllability that flows from all this. In short, a substantially revised philosophy of science is required.

6.1 Elements of a complex systems philosophy of science: metaphysics

The discussion of the ontology of complex systems [what kind of existents are they?] includes two main issues, the nature of their fundamental constituents or components and the nature of their internal features. These will be addressed in order. The discussion will be brief because this book is not the place for an extended examination of the former and most of the relevant discussion of the latter has already occurred.

6.1.1 Fundamental nature

(A) Basic components. The main general ontological issue concerning fundamental constituents or components is whether these should be considered things (dynamically: localised 'particles') or processes (dynamically: fields). These have very different ontological features: most fundamentally for dynamics, things move in space over time, while fields, being everywhere, don't move but have instead spatial concentrations that wax and wane over time.[44] Whichever of these is chosen, complex systems will be attributed the fundamental dynamical attributes of interaction, state, constraint, and time evolution, all developed in their terms. 'Things' is perhaps a natural choice because most complexity occurs well above ion formation. The stability of ions (radioactivity aside) grounds the condition-independence of the basic physico-chemical dynamical interaction rules, itself the chief simplification making complex systems scientifically accessible. For this reason it will certainly be the working assumption of most scientists. However, Bickhard [this volume] argues the case that 'processes' is the better choice, (i) because assumptions of natural metaphysical 'units' (with well-defined boundaries) can prove false and misleading at any level/scale (cf. lasing and superconductivity), (ii) because the currently fundamental conception of matter in physics is that of quantum fields and fields are comprised of processes, with things derivative dynamical constructs, (iii) because fields provide more naturally for fundamental features like bifurcation and emergence and (iv) because, as these reasons suggest, particle assumptions distract from exploring dynamic explanations of entities and boundaries when these do exist. And while the field account will have to do some non-trivial construction work before it gets to relevant things like ions (cf. [Hooker-a, this volume, note 92]), it might also have the advantage of gaining systematic explanations of all thing mutabilities along the way. In some fields, e.g. economics,

[44] For further discussion see e.g. [Hooker, 1973; 1974] and references.

the advent of complex system dynamics has made newly salient the issue of the nature of the fundamental units of analysis [Foster, this volume; Potts, 2000] and it may be that dynamical aspects of the process/thing difference become relevant to them (perhaps in ways analogous to that in which an earlier paradigm of complexity — the self-consistent field approximation — became relevant in economics, [Auyang, 1998]). In any event, these issues place complex systems at the cutting edge of the 2,500 year old thing/process debate.

(B) System identity. A further fundamental issue is what the identity criteria for complex systems should be. Identity criteria could be expected in turn to determine principled system boundaries. They are thus relevant for clarifying scientific practice. The problem looks simpler for thing metaphysics: a system is identical with the collection of its fundamental component things, whatever dynamical state they are in. By contrast, there is the immediate problem for fields that, since they are everywhere and have no parts, it is unclear how any spatially local system is to be distinguished. But a little further reflection suggests that each ultimately faces the same complications. What we would like to capture is the identity of a system *as internally complex*, not merely some demarcation that is neatly abstractly applicable but irrelevant to science.

From this perspective, the first problem with the things criterion is that it is arbitrary because there is no criterion provided for how to select the relevant collections of fundamental components. Within biochemistry, for instance, most possible choices of groups of molecules will not coincide with any relevant systems. The second problem with the things criterion is that it is both too wide, since systems may bifurcate and change their organisation and laws without changing fundamental components, and also too narrow, since every exchange of fundamental components with the environment counts as changing identity (e.g. breathing). These problems arise because the criterion does not select for the complexity-constituting relations.

Since these latter are dynamical, this suggests a dynamically grounded criterion of some kind instead. Strength (typically energy exchange) of dynamical interaction is the standard criterion: a system is identical with the collection of its fundamental component things that are sufficiently strongly dynamically interacting. But then "sufficiently" can never be simply specified, say as some energy exchange amount, because in many complex systems very weak interactions may nonetheless be crucial to determining the complexity dynamics and relatively strong interactions may nonetheless not disturb their form. Nor can any one particular dynamical behaviour, e.g. a particular dynamical trajectory flow pattern, be chosen as identifying criterion instead because complex systems can show a host of different such criteria.

For fields, the pressure to use a dynamical criterion is immediate since there are only arrays of field intensities and their changes over time under dynamical interaction. But then the same problems with using interaction strength and flow pattern as criteria recur here as for things. A third way is to consider same relational complexity, e.g. the same correlational interdependencies or same logical depth or perhaps the same architecture of mechanisms. This approach faces its own versions of the problems of narrowness and width: how much could these relations change without, on the one hand, disturbing the overall relational complexity and, on the other, while still preserving system identity? For example, how would self-organisation, or interaction between two such systems versus super-system formation, be adjudicated? Conversely, note that the same issues would apply if this approach were used with a thing ontology.

In sum, the interactions between systems (including environment) can be intimate and integrating in ways that trouble any boundary-dependent approach to system identity. As Nolfi [this volume] makes clear, system behaviour can emerge out of robotic system-environment interaction in such subtle and integrative ways as to make problematic any split between that due to the robot system and that due to its environment. This sort of situation undoubtedly applies also to organisms. Lewis, e.g., describes how the emotional regulation that will characterise us as adults emerges dynamically in very young children out of their interaction with their care givers [Lewis and Todd, 2007; Lewis and Cook, 2007]. Piaget [1971] and Vygotsky [1986] each emphasised similar regulatory dynamics for cognitive development. This dynamical integration of system and environment causes problems for understanding where the self generally, and mind in particular, ends and the environment begins.

However, the nature of organisms as autonomous (section 3 above) opens up a different approach to identity: same regulatory locus. Living systems exhibit significant self-regulatory capacity since their autonomy requires that they regulate their own behaviour (action) and internal processes (metabolism) to regenerate themselves, including their self-regulatory capacities. This self-regulatory capacity distinguishes them from their environment as a locus of regulation of behaviour, and thus affords them an individual identity (see further 4 above). This allows both components and processes to come and go — as they do in autonomous organisms — while preserving identity.

Even where internal regulation and behaviour emerges jointly from system and environment interacting, the system is self-regulating in a way that the environment is not. It remains an identifiable self-regulatory locus however much the additional character constructed on that basic self reveals the fixated history of interaction with its environment (or how little character remains, as with childhood abuse, drug addiction, trance states and some illnesses that prevent/ remove/ distort/ suppress much of healthy construction). And it provides a principled basis on which to distinguish super-systems that have individuals among their sub-systems from those that don't, e.g. a family from a robotic assembly line (however robustly controlled) and both from an organism with prostheses. However, this criterion is unavailable to all those systems that lack autonomy or some similar self-regulatory capacity.

In practice, identity is specified pragmatically via some mix of: same things, same behaviour patterns, same complexity structure and same regulatory locus.

This flexible arrangement suffices for practical research and is perhaps the best that can be achieved.[45]

6.1.2 *Nature of internal features*

(A) Reduction and emergence. The most immediate issue is the ontological nature of emergence, especially of self-organised emergence. Is the outcome a distinctive new existent, hence irreducible, or does the system remain reducible to its fundamental constituents? Here it seems to me science provides clear guidance. In the class of cases where a dynamical bifurcation occurs a changed dynamical form comes into being, doing mechanical work (transforming energy) in a changed way. This provides a clear dynamical basis for claiming that there is a new dynamical existent formed, whether or not it is more ordered or more organised that its predecessor. (But if it is, if a new relatively macro constraint is formed, then the case is particularly obvious.) Because of this [Hooker-c, this volume] argues that this dynamical demarcation provides the only well-defined dynamical basis on which to draw up concepts. If this is accepted, we have from [Hooker-a, this volume, section 3]: bifurcation = new dynamical form = changed constraints = wide emergence, bifurcation with relatively macro constraint formation (= top-down/downward causation formation) = narrow emergence, self-organisation = narrow emergence under internal system dynamics.[46] I prefer wide emergence on the grounds that it is the outcome, the changed dynamical form, that matters, not the route to it. Either conception of emergence leaves dynamical reduction and emergence delicately and unavoidably intertwined: emergence provides the constraints within which processes are well defined and reduction identifies these with their realising dynamical flows [Hooker, 2004, Part 5]. This places emergence exactly at the locus of the 'scandal' of mathematical dynamics underlying complex systems, see [Hooker-c, this volume].

(B) Mechanisms. The dynamical flows that realise functions are mechanisms, broadly construed [Hooker, 1981, Part III; Hooker-c, this volume]. For a variety of practical reasons most of the engineered mechanisms we are familiar with, like clockworks, are separately located and designed to repeatedly and reliably carry out a single function. That is, they are functionally modular (separate) and mono-functional (restricted to one function). It is important to appreciate that, for complex systems, the basic conception of a mechanism carries no such

[45] For further discussion of system identity see [Cumming and Collier, 2009].

[46] Many philosophers have had difficulty with this situation because they naturally gravitate to logical criteria but self-organised emergence presents a class of cases in which the new existent has the same fundamental components as the original entity and yet there is no logical combination of them that distinguishes between the old and new existents. At least one, and often both, are irreducible. Others, including [Hooker-c, this volume], opt to replace this use of logical criteria by dynamical criteria, so that a natural unity is preserved. Bickhard [this volume], for whom emergence is central, argues that the issue cannot even coherently arise in a process (dynamically: field) metaphysics.

implications. Bechtel and Abrahamsen [this volume, p.2] characterise the steps in a mechanist explanation as comprising "(1) the identification of the working parts of the mechanism, (2) the determination of the operations they perform, and (3) an account of how the parts and operations are organised so that, under specific contextual conditions, the mechanism realises the phenomenon of interest." Although Bechtel's language may sometimes suggest more permanency for components, in the foregoing conception it suffices that the working parts of the mechanism are those components whose stability is sufficiently long lived, in relation to the timescale of the whole mechanism operation, to consistently perform their component operation as often as appropriate (and at least once).

This is compatible with component lifetimes being only a fraction of the overall mechanism operation time as well as with their lasting many cycles, and with a mechanism being realised for a once-only performance, not to be re-created. It is compatible with component operations destroying themselves or other components — to be replaced by others from elsewhere in the same mechanism or from other mechanisms, or with their being transformed so as to enter some other mechanism or with their remaining outside salient mechanisms for a time. Moreover, it is compatible with the same component playing different roles in many mechanisms (multi-tasking) and the same overall function occasion to occasion being realised by different mechanisms using different components (multi-realisation/tasking) and therefore with mechanisms inter-weaving 'through' one another in very subtle patterns in space and time, yet still existing. Finally, it is compatible with a mechanism being self-organised from component mechanisms or basic components, e.g. with the mechanism serving the global constraint of respiration possibly emerging from the entrainment of cellular Krebs cycles plus cardio-vascular and other cycles, along lines suggested by Bechtel and Abrahamsen [this volume]. Outcomes of all these sorts are unavoidable for irreversible developmental systems carrying out the complexes of work-constraint cycle mechanisms required to satisfy global constraints like autonomy (see [Hooker-a, this volume, section 3]).[47]

(C) System decomposition. It would be nice if all complex systems were non-trivially decomposable under some decomposition principle that also grounded functional decomposition. Here several principles of decomposition are reviewed in pursuit of that goal. *Decomposition by flow.* The desired universality of decomposition can be achieved by construing mechanism broadly as any energy flow and construing function equally broadly as any many:one or one:one input/output map. These are the construals supporting the function-to-mechanism reduction schema of preceding paragraphs. They are useful for explicating the reduction relation between functions and dynamical processes, more structurally complex cases then following directly. But these construals fail the non-triviality requirement by evacuating the notion of mechanism of any distinctive content beyond dynamical description itself. *Decomposition by mechanism.* Another decomposition princi-

[47]For further analysis of biological, especially cellular, mechanisms see [Bechtel and Abrahamsen, 2005; 2007; this volume; Bechtel, 2007].

ple is decomposition into Bechtellian mechanisms as characterised above (hereafter mechanisms *simpliciter*). Bechtel has argued at length that they play a crucial role in the search for system decomposition [Bechtel and Richardson, 1993, note 47]. However, as noted above, mechanisms will not necessarily be amenable to simple decomposition into permanent 'parts' because of the possibilities of multi-plexing and multi-tasking, but especially because they may be dynamically emergent (although this does not need to, but may, alter their components).[48] Moreover, while they are important to understanding many systems, they will typically only provide a partial understanding of overall system functioning because there will also be many less organised flows constituting the processes of these systems.

Decomposition by pathway. An alternative but related decomposition is into pathways, that is, directed energy/material flows through complex systems, whether organised (in the sense of [Hooker-a, this volume, section 3]) or not. Pathways capture highly ordered energy/material flows that mechanisms exclude, e.g. simple linear transformation sequences without feedback/forward. It follows from the second law of thermodynamics that each irreversible system forms a pathway, since energy/material intake and waste ejection is unavoidable for them, but this is a trivial application. The notion of pathway has also proven useful in cell biology to capture sequences of chemical formation and transformation. While both mechanisms and pathways may each be decomposable in kind, also parts of mechanisms can constitute pathways and mechanisms can form parts of pathways. However the notion of pathway could be so broadly construed as to render it identical to that of energy flow and hence render pathway decomposition as trivial as flow decomposition. It is unclear if there is an interesting tightening of the concept that would render it a substantive alternative principle of decomposition to mechanism.
Decomposition by module. An obvious decompositional principle, often justified and (so) useful. It is clear that modular decomposition is nearly always partial because mechanisms and pathways can be supra-modular, correlating activity across many modules (e.g. respiration in biology), sub-modular (e.g. the sub-cellular Krebs cycle) and a-modular (e.g. turbulent heat convection).

Decomposition by function. This is only indirectly a dynamically-grounded principle of decomposition — grounded through the function-to-flow reduction of flow decomposition. Thus it must coincide, case by case, with flow/pathway decomposition. Nonetheless, it has an independent interest because for all living systems (and some others) their functional organisation is crucial to their material integrity [Coffman, this volume]. And as living systems increase in intelligence there is some suggestion (e.g. from the appearance of language) that their functional organisation is dominated by internal functional rationales. Functional organisation has ultimately to be instantiated in dynamical flows. And it would reproduce the design simplicity and transparency of current engineered devices if functional

[48] See e.g. [Bechtel, 2007, notes 4, 7, 10] and text. And any such parts may not coincide with stand-alone things because entering the organisation may change their characteristics either through self-organisation or through mutual co-evolution. Cf. [Levins, 1970].

organisation were directly realised in (homomorphically embedded in) a matching decomposition of flow organisation. But for multi-plexing, multi-tasking organisms there is no a priori reason why this has to be so, and many reasons, energetic and evolutionary, why it should not be expected. Flow/pathway realisation will be condition-dependent. Thus functional decomposition stands as an independent, partial decomposition principle.[49]

In sum, there are two clear dynamical principles of decomposition - by mechanism and module. These decompositions overlap and interweave, and severally and jointly remain partial. In addition, there is possibly a third dynamical decomposition principle - by pathway or flow - that might be universal and trivial, but in any case picks up all the system connections the other two omit. And there is a fourth, functional decomposition principle that complements the others, and is ultimately instantiated in, and condition-dependently reducible to, them.

(D) Condition-dependent laws. Another internal feature is the nature of the condition-dependent laws the system obeys. The basic notion of condition-dependency has already been discussed in [Hooker-a, this volume, section 3]. Dynamically, the interaction conditions characterising a system condition the expression of the dynamical interaction laws in that system. Thus while free electrons normally mutually repel and are attracted to ions, in a metallic crystal ionic lattice they freely drift together to form its Fermi conduction current. The issue here is to what ontological feature does lawlikeness refer (if any) and how do condition-dependent laws share in it (if they do). There is a philosophical debate about the nature of law-likeness — whether, e.g., it refers to some kind of necessary connection in reality or simply refers to some kind of constructed status, e.g. is an axiom or theorem of our best science. Here science has again provided a strong lead: over three centuries, through admittedly some elaboration and adjustment and still with hanging threads, there is a distinguished dynamical relation underlying every law-like regularity: energy transmission and transformation. It has constructed theoretical status and its referent (energy transmission and transformation) provides a natural nomic necessity to laws. Moreover it extends directly to condition-dependent laws since they are simply read as referring to the common dynamical-condition dependence of energy transmission and transformation.[50]

[49]Motif analysis is an obvious and rapidly spreading computational technique. (It amounts to looking for repeated patterns in a system, whether these be intra-cellular multi-gene DNA sequences or connection patterns in a network.) But this does not in general offer a dynamically well-defined principle of decomposition, although in some cases it will offer a principle of modular (network connections) or functional (DNA sequences) decomposition. In short, it is a sometimes valuable aid to existing decomposition principles. Thanks to David Green for drawing this to my attention.

[50]See [Hooker, 1998] for a review of the triad of available positions under each general alternative for the status of laws, but cf. [Carroll, 2006] for additional discussion and references. Energy transformation follows from transmission by the Second Law of thermodynamics. It is clear that ultimately a basis in a relativistic quantum field version of energy transformation should be provided; besides the general problem of unifying the two, among the more specific hanging threads are the status of the quantum mechanical Pauli exclusion principle and the Bohm-Aharanov

To see what this might involve, consider that all dynamics is specified as a constrained flow (across state space over time). The flow is driven by the energetics of the specified dynamical interactions and precisely characterises the behaviour of the system, no matter how complicated the interrelationships. The constraints specify the conditions by which the interaction dynamics is constrained. These are the other forces operating on the system, treated as external to its interaction dynamics, e.g. the container in which a gas is confined, together with any complementing limiting constraints (e.g. that interaction forces go to zero at infinite separations). Thus the system behaviour is explained as the joint product of interaction dynamics and constraints, together with the initial conditions that fix its particular trajectory from among the possible ones [Hooker-c, this volume]. Thus a condition-dependent law is the basic (unconstrained) dynamical flow law operating under the constraint conditions.

(E) Determinism. In the same spirit, dynamical analysis provides the appropriate framework for the consideration of the nature and status of determinism in complex systems. From a scientific perspective, since virtually all real complex systems occur above ion formation, and well below light speed and space-time warping mass, their dynamics can be adequately modelled as Newtonian and hence, presumably, as deterministic.[51] But specifying the nature of that determinism turns out to be complicated. For instance, determinism is often thought to exclude the presence of randomly distributed variables, but with the discovery of strange attractors and chaotic trajectories in deterministic systems we know that random features of dynamical trajectories are not indicators of indeterminism. Again, it is often assumed that complete prediction from complete system state information is the appropriate characterisation of deterministic dynamics, but we already know enough about the limits of prediction in complex systems to know that this will not provide a satisfactory criterion (and there are other complications). Finally, there is the more promising idea of unique evolution - that a dynamics is deterministic just in case if any two dynamical trajectories coincide at any one point they coincide at every point. But this is hard to apply, and can fail in surprising ways. For instance, it fails to apply fully to Newtonian mechanics, even when the basic component interactions are well defined. Essentially this is because when Newton's equations are applied to many complex systems they yield systems of differential equations whose solution trajectories are either globally inaccessible, or show surprising non-uniqueness features, or fail to yield analyseable dynamical trajectory representations (bifurcations). Thus the most appropriate ontological representation of determinism is left an open issue.[52]

---

effect.

[51] Quantum effects can mostly be treated as micro-generated noise and the cases of coherent macro quantum effects, such as lasing and superconductivity, given special treatment where necessary.

[52] For further, more precise discussion, see [Bishop, this volume, section 3.2; Earman, 1986; 2004; 2007].

(F) Causality. The dynamical basis of laws also provides an appropriate underpinning for an assessment of the place of causality in complex systems. In fact, it supports the dissolution of the traditional notion and its replacement by a more appropriate notion of dynamical interdependency. The traditional notion is essentially that cause is a unique relation between two entities, $A \to B$, such that the cause (A) causes the effect (B), typically expressed logically as A is necessary and sufficient for B. This conception is no doubt derived from everyday experiences of human agency in which a greatly simplified identification of initiator of action and the course of action mostly suffices for common practical matters. It is also reinforced by the symmetry/stability-equilibrium paradigm, since departures from these states can frequently be isolated. Nonetheless, this notion quickly runs into difficulties on its own terms. Never mind debates about what kind of entities A and B could be, the formulation quickly runs into complications trying to accommodate multiple component causes and effects, ending in the swamp of INUS conditions (Insufficient but Necessary factors of Unnecessary but Sufficient causal conditions).

However, in the presence of feedback loops, global constraints, emergence and the like, this formulation completely breaks down (cf. [Bishop, this volume, section 3.3; Coffman, this volume]). It is not just that the swamp of INUS conditions is intensified to the thicket of SI (stochastic interdependence) conditions, but that all of the subtle dynamical conditions that underlie them — the entrainments and nested correlations, entwinements of emergence and reduction, unique individuals, and so on — slip past this kind of analysis. The appeal to causality may once have been intended both as description of what was there and as explanation of what happened, but such uses are irredeemably lost, except as an unsystematic pragmatic convenience.[53] Rather, the sole general relationship capable of embracing these complex interrelations is that of dynamical interdependency. At the very least, tracing all the dynamical pathways in a complex system (trivial, but universal, pathway decomposition - see above) reveals all its specific dynamical interdependencies (including across levels). All its more complex interdependencies derive from these and exhaust its behavioural capacities. Thus this analysis wholly and adequately replaces causal analysis. Moreover, explanation is the over-arching concept under which causal explanation falls, and its function survives the demise of causality and its replacement by dynamical interdependency — see 6.2.3 below.

(G) Unification. Finally, there is the issue of the ontological nature and status of scientific unification, if any. According to section 5.3 knowledge is condition-structured. So then, how are these dynamical interaction conditions, or equiva-

[53] It is always possible to re-define cause in vaguer language so as to cover complex interrelationships — e.g. something like ‘causation specifies that events have substantial production conditions so are non-arbitrary and nomologically unavoidable’ or ‘causation is the principle of explanation of change, it is a relation, not between things, but between changes of states of things’ (suggested as Schopenhauer’s definition). But this only secures us a general label, it covers over the dynamical diversities rather than illuminating them.

lently the perspectives they define, interrelated? How do they 'hang together' to provide a coherent dynamical conception of reality? (This is in distinction to the issues of how that conception is epistemically constructed and justified.[54]) The answer is that the dynamical perspectives supported by dynamical interaction conditions are distinguishable but provide for a complex perspectival unity — reality as a dynamically structured unity of dynamical perspectives. For if a sufficiently complete model of the complex system involved is constructable then all of its dynamical perspectives can be reconstructed from it by modelling the corresponding classes of interactions with it. In this way the perspectives may be unified, in a manner analogous to the unification of 2-dimensional perspectives of an object in a 3-dimensional model. So long as this includes the dynamical independencies and multiple realisations characteristic of levels, we have an ultimate sense of unity. Even when practical dis-unity is expressed in working paradigms applicable at different levels and scales and concerning different properties, e.g. between molecular chemistry and cellular behaviour, the resulting science is ultimately unified by being 'matched up along their boundaries', e.g. by matching cellular respiration to molecular transformation, signalling to molecular transmission, including understanding levels as dynamical emergents that support functional reductions of just these kinds (but do not reduce the levels). But there is one group of perspectives, thus far omitted, that create a deeper sense of disunity because they cannot be unified in the foregoing sense — idealisations. These are examined at 6.2.6 close.[55]

The limits to the accesses we have to complex systems, e.g. modelling at one or a few scales at a time, is a natural consequence of the constitution of systems, including ourselves — it can itself be modelled in the world of complex systems. To try to make the individual accesses ontologically primary in either case is to perversely reverse cognitive achievement in both development and science. It reverses cognitive achievement in development because it reverses the achievement of transcending partial sensory perspectives by uniting them through embedding them in a perceiver-independent world. This is a major cognitive achievement because in that world the perceiver is itself located and hence the perceptual relation itself can be modelled and thus examined. It is not surprising that it proves superior as a vehicle for subsequent scientific investigation. Making the

[54]The literature is dominated by epistemically or values based perspectivisms, and there positions tend toward radical perspectival separatisms — reality as a collection of semantically, normatively or like strongly separated perspectives. Since these latter stem from apriori assumptions imposed on science from elsewhere, they are set aside here (see further 6.2.6 close). Note also that the dynamically-based perspectives discussed in the text will have been accepted under certain epistemic values (e.g. predictive, explanatory and unificatory powers, testability and fecundity). Although these values are often in tension with one another (= cannot all be met at once), they will nonetheless promote and be promoted by this complex unity as, e.g., 'matching up' across levels (see text below) increases explanatory reach and evidential support.

[55]In brief, many dynamically complex theories support different (but usually quite specific) idealisations of their dynamics under various limiting conditions (e.g. friction-free). but these are likewise unified by their being limiting cases of a single dynamics. If the idealisations are degenerate (collapse out structure) it will not be possible to logically re-construct that unifying dynamics from the idealisations and simple logical combinations of them will prove less explanatory, perhaps even mutually contradictory.

individual accesses ontologically primary reverses cognitive achievement in science because it is only through modelling ourselves as complex systems in interaction with other complex systems that we can adequately investigate the limitations of access to such systems, including for other species and systems besides ourselves, and seek ways to systematically improve that access. For this very reason, even making these accesses methodologically primary is unhelpfully restrictive (KIDS, not KISS — section 5.1 II.1), and far from reducing risk of error, greatly increases it. (For certain limited epistemic purposes it may make good sense, but that is a matter of epistemic strategy, not a fundamental ontological issue.) Thus, it seems to me, science provides no support for ontological perspectival separatism, although dynamics itself imposes a complexity, and practical incompleteness, on the unity of any science constructed by finite creatures (cf. also [Wimsatt, 2007]).

6.2 Elements of a complex systems philosophy of science: epistemology

A characteristic feature of our knowing complex systems is our use of models of them as intermediaries. This would be a banal observation if it referred only to the necessity of symbolically representing objects of knowledge as a means to investigate them (e.g. to mathematical models). It would remain banal if it referred only to the fact that complex systems, being complex, are typically most conveniently treated perspectively in terms of partial representations of them. These characterisations do not discriminate between complex and simple systems. The claim begins to bite when it is added that many complex systems are so complex that modelling them as whole is pragmatically 'hard' in the sense of section 5.1 above. A further dimension is added through the recognition that bifurcation and self-organisation add internal principled bases for partial modelling. The force of the claim is completed by adding that the mathematical dynamical representations of many complex systems, if they could be set down, would have no accessible solution and so they necessarily have to be explored and known through computational model simulation.

The standard modelling literature covers quite necessary topics such as (i) what models are, their various forms (formal, material, metaphorical, etc.) and various purposes (exploration, prediction, explanation, control, etc.), (ii) the corresponding variety of methods for constructing them (decomposition versus global modelling, differential versus difference equations versus programming, etc.), (iii) the steps involved in using them (construction, internal investigation, targeted application) and (iv) the use of robustness, sensitivity and goodness-of-fit methods to improve reliability. These features are significantly and subtly more complex for complex systems. Good examples are provided by the faithful-model assumption and goodness-of-fit methods. Much of the serious methodology on these latter issues is actually found in the sciences themselves and in theoretical statistics, rather than in philosophical writings, but no less contribute to epistemology. Most of these topics are discussed or touched on in the essays herein (cf. e.g. [Bishop, this volume] on the faithful model assumption) and some are briefly discussed in section 5.1. Here they will be largely assumed in order to focus just on those aspects of particular interest for complex systems. Similarly, general philosophical discussion of laws, explanation, confirmation and the like are set aside here in favour of issues of more particular relevance for complex systems.

6.2.1 Condition-dependent generalisation

In the traditional conception of science, an assumed world governed by universal laws supported the assumption of universal generalisations and explanations as the basic forms of these features in science. If generalisations and explanations had to be hedged around with conditions of applicability, it was not the universality of law that was at issue but the prospect that the particular laws invoked might not be the basic ones. The hedging was taken to signal the presence of ignorance of the underlying universal principles that would allow the hedged accounts to be replaced by appeal to the 'real' underlying laws plus special initial and/or constraint conditions that remained independent of them. In one sense this basic picture has not changed, since the basic component-component interactions still obey universal laws (at least up to cosmological condition-dependence). If this basic universality condition were not present, if condition-dependence were the only universal condition, then the situation would be significantly altered - and science that much harder, as it might yet turn out to be. And there are prominent voices extolling the prospects for high-order general laws of organisation that would apply to all complex systems.[56] How contentful any truly general laws will turn out to be remains unclear (cf. [Fox Keller, 2007]). The situation was probably summed up for all by Alm and Arkin [2003] speaking about biology: "... a recurring theme in biological modelling: the more we know, the more complex it is — even as some overarching principles become clearer." (p.193) Some overarching principles of organisation would be welcome; even so, it will remain true that once more than two relatively simple components are involved, condition-dependent laws become the rule. This means that condition-dependent generalisations and explanations also become the rule.

How to generalise then becomes an issue. Using functional similarity as a guide is fraught with risk, e.g. both algal and human cells require free energy to be made available to drive cellular processes, but this is achieved differently in the two cases. What is needed is some basis on which generalisation can proceed. In the light of the dynamical nature of condition-dependent laws, the most fundamental basis has to be that generalisation is valid across the same dynamics, at whatever degree of abstraction the dynamics is taken. For instance, a prominent strategy in biology is to appeal to evolutionary conservation of mechanisms for generalisation.[57] This relies on constancy of a complex combination of genome and

[56]See e.g. [Barabasi and Oltvai, 2002; Barabasi and Bonabeau, 2003; Kauffman, 2000; Westerhoff and Palsson, 2004]).

[57]Thanks to Bill Bechtel for reminding me of this here, cf. [Bechtel and Abrahamsen, this

epigenetic dynamics, presumably because developmental and selection constraints ensure that no alternative will prove viable. Abstractly, for a single system generalisation is valid across the same dynamical sub-domain, e.g. within an attractor or within a class of transient trajectories, and generalisation across systems is valid within the same attractor landscape structure (topological and projective). The latter preserves all higher order dynamical properties, all kinds of dynamical behaviour (attractors, bifurcation thresholds, etc.) and their interrelations, only smaller quantitative details may change. This is formally the requirement for transferring internal understanding of function and behaviour from one system to another. Although sameness of dynamics could in principle be brought about by different combinations of interactions and constraints it will in practice very likely be explained by the same underlying dynamical structure. But, as the evolutionary example illustrates, in practice generalisation assumptions will usually be developed in much more pragmatic ways as rules of thumb, and in ignorance of precise dynamics. The validity conditions of these generalisation rules work their way toward a better footing where they can be characterised by ranges of parameter values, and possibly ranges of inputs as well, e.g. where compound condition-dependent laws are active (see below).

6.2.2 Condition-dependent laws

Scientists in biology often regard their approach more as a model-building than as grand theory or even law centred. This is understandable in a domain where nearly every variation results in differing functional capacities and behavioural patterns. Compared to the grand universal, invariant laws of physics, these local idiosyncratic behavioural patterns don't count as laws. It is more useful to simply model each system, or even each system+context, and try to understand it on that basis. But, as noted, biologists do use laws in constructing their models, the laws of physics and (bio-)chemistry. Their constancy and universality grounds both the reliability of bio-chemical mechanisms, and of our reliably identifying them. Even so, complication arises from the fact that the operational invariance largely occurs at the ion-ion interaction level. How n-body, k-component ion systems operate is often a strong and sensitive function of the system initial and constraint conditions, especially organisational conditions and, vice versa, their initial and constraint conditions is often a strong and sensitive function of their dynamical trajectory (e.g. because of self-organisation).

Of particular interest is how to understand laws arising from system organisation. It is rare to encounter organised conditions in physics while it is normal in biology, sociology and like domains. In [Hooker-a, this volume, section 3], the discussion of condition-dependent laws concluded with assimilating the 'special' laws of biology and cognate domains to physics, seeing them as differing at most in the degree of intricacy in their characterising conditions. The problem posed by organisation is that we have no mathematical analytics for dealing with it.

---

volume].

Organisation has as its defining character the realisation of one or more global functions (that is, for the organised components as a whole). The root issue, I suggest, is the presence of these inherently global constraint conditions. We have no way to represent these since globalness per se has no dynamical representation (see section 5.1). Our present (and perhaps only) way out of the problem is to make use of whatever decompositions are available. The construction of electrical circuit laws from their components, e.g., is accomplished pair-wise using Kirchoff's laws, no matter how complicated the circuit or whether its topology includes closed loops that may realise organisation. Any global constraint that guided the design of topology is implicit, not explicit, in this analysis. Similar dynamical analyses using linear sub-functions that can be successively pair-wise interrelated plus generalised Kirchoff's laws can be applied to any network analogously to that for circuits.[58] Where this cannot be done, e.g. because component character is itself a function of its embedding organisation (cf. [Coffman, this volume]), we have no adequate formal tools for analysis of the required construction and have to proceed by trial and error supplemented by analysis of component behaviour. However constructed, it is clear that the law that governs the whole organisation is a condition-dependent compound of the interaction laws of the components, pair-wise or n-wise, in the same general way as it is for circuits. It is on that basis that the systematic functioning of organisation in complex systems is to be explained.

Consider, e.g., the Krebs cycle, a fundamental cellular process (cf. Bechtel and Abrahamsen herein). This is a cycle of sub-cycles of chemical interactions, the products of each step and sub-cycle providing input to the next respectively, as well as determining required external inputs and outputs. Each of the individual interactions is context-independently law-like, determined by the quantum mechanics of electron exchange interactions. But each such reaction occurs in the spatially organised context of all the others and only in that complex empirical constraint context do they constitute cyclical cellular dynamics. That oscillatory law derives then from a complex fusion of the local ion-ion interaction laws and empirical constraints pertaining to the conditions for each sub-cycle and the whole super-cycle and cannot be derived from the local ion-ion interaction laws alone. This is how mechanisms are lawlike. Change any component or organisation and the entire edifice becomes something different, with a different dynamics, exhibiting different laws. And where we do not know how to dynamically represent the constraints in local detail, as with most forms of autonomy still [Hooker-a, this volume, section 4.1.1], the form of the law will elude us even though the mechanisms remain (context-dependently) lawful. That is why no simple set of laws can be deduced in advance. And it explains why straight generalisation across organisms often doesn't work: the component material conditions vary because of differing developmental conditions.

Such considerations reasonably licence the extension of the notion of law to all these cases. For since constraint formation involves a new dynamical form,

[58] Cf. [Kell, 2006, p.880; Bechtel and Abrahamsen, this volume] for illustration with Krebs and other cycles.

it is reasonable to say that it obeys new dynamical laws characteristic of that form. Moreover, the idea that true laws have to be specified independently of any initial and constraint conditions is a conceit of physics, and perhaps ultimately not true there either considering that even fundamental laws evidently changed form as the big bang cosmos cooled. But once that independence requirement is dropped we are free to see biology as replete with real laws, it is just that they will be condition-dependent, or 'special' (as some philosophers say).[59] This applies even in commonplace physics; for instance, condition — a cooling mould of liquid iron in contact with a heat reservoir of lower temperature, emergent laws — rigid body (not fluid) dynamics, crystalline (not fluid) conduction of electricity, heat and sound. Put that way, condition-dependent laws are commonplace throughout even physics, and throughout all the other sciences. It is just that condition-dependent laws are often on that account hard to predict or use for prediction, but that is a different, epistemic issue.

Why not push condition-dependence further to include every instance of change in initial and/or constraint conditions? For instance, the specific force of gravity changes between the Sun-Saturn and Sun-Earth sub-systems, because of the changing masses involved. Why not claim all these as distinct, similarly condition-dependent, laws? In one undiscriminating sense they are, but in these cases it is the same general law that is involved, the diverse cases are unified by a single lawful interaction form and differ only in initial parametric (mass) and location (orbit) conditions. Thus it is more perspicuously identified as behavioural change without law change. This is not so for the iron bar and other cases involving constraint formation. However, surely the constraint formation cases are equally a consequence of the underlying universal dynamics, and simply produced under specific initial and/or constraint conditions; if it is just that at present we cannot analytically represent constraint formation then that should not stop us from allying them to the previous simpler cases. This is so, but there are two important differences marking off the constraint formation cases: (i) constraint formation heralds the presence of an irreducibly new dynamical existent, (ii) the dynamical form itself alters accordingly, so there is no common universal law form. Thus they represent a genuinely distinct set of conditions that justifies distinguishing the resultant laws.

The same considerations apply to dynamical models of the genome. There may be a wide variety of dynamically different forms that a genome can take up as various of its processes alter its own initial and/or constraint conditions so as to induce a dynamical bifurcation — e.g., by creating and inserting a new catalyst into the

[59]See [Polyani, 1968] for an early insight of this kind. Polanyi argued, in effect, that what was distinctive of living systems was that their governing laws were so strongly dependent on initial and/or constraint conditions. But we can now see that condition-dependence is not unique to life, e.g. it characterises at least all self-organising dynamics. Polanyi had in mind at least the way that information can alter the basis of behaviour in living systems. Since presumably the impact of an information-conveying signal on an organism on some occasion is dynamically equivalent to some dynamical change, from simple change of state to a radical bifurcation, Polanyi's living systems can after all be brought under the same dynamical paradigm.

protein dynamics — thus forming special laws for that condition. These effects can propagate developmentally and/or genetically. The emergence of a new constraint with new dynamics may lead to the subsequent dynamical formation of still further top-down constraints, and so new entities, which would not have been dynamically possible without that preceding formation event. This path-dependent cascade of dynamical consequences is marked by its initiating formation event and thus exhibits dynamical fixation of those historical constraints. Something like this must be the overall dynamical forms of both evolution and development and occurs throughout ecology (for example, somatic and niche symbioses).

Needless to say, biology will not splinter into an unprincipled disunity under these complex dynamics, and nor will any other science.[60] Once again, all these dynamical changes will still be dynamically determined by, and identified in terms of, their dynamical constituents and governed by laws which themselves are thus grounded in the underlying universal dynamics. For biology, this is precisely what the biochemistry of the dynamical network models is meant to show. It will also encompass the many changes that consist of less profound dynamical transitions falling under the same dynamical form. And the requirement to 'match-up' the dynamics of different spatio-temporal scales and domains provides a further important unifying component. For example, unifying molecular chemistry and cellular biology requires interpreting cellular processes in biochemical terms which immediately generates many penetrating tests because of the requirement to match up the two descriptions — for instance all of the function to process reductions.

All this provides a shared dynamical framework interconnecting emergent variety in intimate ways that make it possible to successfully model complex genome dynamics and even development, navigating through the complex world of emergent but interconnected intra-cellular and inter-cellular levels and laws. This gives a strong sense in which biological science remains unified even while acknowledging more strongly initial and/or constraint condition dependent laws than simple physics and chemistry was wont to consider.

The challenge to science is to recognise explicitly and better understand this plethora of law types and unifying inter-connections across its many domains, so as to make explicit their basis and their theoretical and methodological implications.

6.2.3 Condition-dependent explanation

In the traditional conception, laws are the basis for explanation: explanations consists in deducing the explained phenomenon from laws and initial conditions. But these 'covering law' explanations are intentionally condition-independent, the covering law is universal and picks out a universal pattern holding everywhere. The domains of organisms, by contrast, are deeply condition-dependent in their structure and function and demand equally deeply condition-dependent explanations of

[60]Cf., e.g., [Dupre, 1993] and for the complex dynamical unity response presented here see [Hooker, 2000; 2004, n.4, section V]. It should be added that the conception of laws as simple universal generalisations, commonly assumed by philosophers and scientists alike, is simplistic; science shows a far more complex and rich spectrum of kinds of laws — see [Hooker, 1998].

their workings. Fortunately, the development of the notion of condition-dependent laws allows the same general deductive formula for explanation to be retained, but with a restricted, though still valid, sense of universality that holds within the specified conditions. As observed in section 6.1.2, complex system behaviour is explained as the joint product of interaction dynamics (the true universal laws) and constraints (the operative conditions), together (as before) with initial conditions. A dynamical explanation then consists in specifying all three factors and deducing the behaviour from them, a direct generalisation of the original covering law formula. The fundamental aim of science continues to be provision of dynamical explanations of the behaviours of complex dynamical systems.

The dual role in explanation of interaction dynamics and constraints leaves room for their relative importance to vary from case to case. This is of some importance in correcting the widespread impression, e.g. left by most physics textbooks, that constraints are a minor consideration, merely background to the interaction dynamics, where all the explanatory work is done. In complex systems constraints play a major explanatory role. We have already seen that they define the foundational problem of providing a general analytical dynamics for complex systems [Hooker-a, this volume, section 4.1.1]. And they played key roles in structuring many of the central complex systems notions of [Hooker-a, this volume, section 3], from self-organisation to autonomy. But in addition it is also important to understand the particular dynamical, hence explanatory, roles they play in many processes.

Consider a diverse social group ordering from a restaurant menu. We normally assume that our order reflects our tastes, = our selection dominates the result. But suppose the menu to contain only 1 item. Then the menu constraints dominate, selection for taste has no force. If the menu has only 2 items, constraint still dominates selection. Even if it has a lot of items but they all belong to one category, whether of food (soups only) or cuisine (Korean only), there remain important constraints on the effectiveness of our selection to deliver optimal, or (depending on taste) even satisfactory, results. It follows that whatever the impact of the social group dynamics on the menu selections made, it must work within the menu constraints. In consequence, the explanation of the choices made must refer to both factors and across cases each can vary from having the dominant explanatory role to having negligible explanatory role. These sorts of consideration apply from physics to sociology. Consider a gas and a metal crystal with heat applied. The gas can conduct or convect and whichever it does it explores efficiently through random collision. But the metal has only crystal vibration and Fermi conduction to transport heat, its options are more constrained. No doubt what it does it does efficiently, but within the solid state crystalline constraints.

This provides the framework from within which to consider how far energetics might explain complex system behaviours. When Kleindon [2009] formulates a thermodynamic principle of maximum entropy production [MEP] for complex systems he is careful to qualify its application by "if there are sufficient degrees of freedom within a system ...". Only when the heat supply is made high enough to melt the crystal in the example above would MEP (perhaps) dominate the explanation of behaviour. Once feedback pathways are involved it is harder to see how MEP applies. An energetically small feedback signal can make an very large difference to energetic behaviour. Consider a flow monitor feeding back to a controller to raise/lower a dam sluice gate; the energy in the monitor-controller circuit will be miniscule compared to that in the water flow, and even that in the sluice motor will be tiny compared to that in the water flow. Nonetheless, because of the material constraints operating, the flow magnitude is dominated, not by its energy, but by these much smaller energy flows. Of course the basic laws of physics apply. But this doesn't help much because, as illustrated, by altering the constraints under which the law is applied its effect can often be significantly modified. As the discussion to follow illustrates, there is even stronger reason to suppose that the application of MEP will be constrained when we pass to complex systems comprised of organisms, particularly intelligent, socially organised ones whose behaviours are constrained by many layers of socially originating regulation.

In standard neo-Darwinian population genetics evolutionary dynamics is a function of variation (V - here genetic mutation), selection (S - here for genetic fitness), and reproduction (R - here reproduction of genomes). If selection is to be the driving force of evolution then it is important that effective mutation (mutations that are presented for environmental selection) provide an adequately varied 'menu' of mutations to select from, hence the standard assumption of the universality and randomness of mutation. To the extent effective mutation is constrained, as today we know it to be in various respects, to that extent selection is less potent in explaining evolutionary dynamics. The same applies if the selection force is otherwise constrained, e.g. by absence of resource competitors or predators, and/or retention is constrained, e.g. by group reproduction practices. To this add that new organism capacities can evolve through enabling coordination of constraints, including the capacities to modify niches and even genes (either through internal cellular gene editing or, in our case, external medical intervention). Then instead of a simple dominant selection dynamics, evolution now appears as a selective process operating under many layers of constraints, indeed where each of V, S and R are increasingly constrained (= regulated, [Hooker-a, this volume, section 4.1.3]). These constraints often play important, sometimes dominant, explanatory roles (cf. e.g. [Coffman, this volume]). This is also the larger explanatory setting within which to consider specific issues, such as Kauffman's [1993; 1995] self-organisational constraints. Kauffman [1995, p.112] remarks: "If biologists have ignored selforganization, it is not because selfordering is not pervasive and profound. It is because we biologists have yet to understand how to think about systems governed simultaneously by two sources of order ...if ever we are to attain a final theory in biology, we will surely, surely have to understand the commingling of selforganization and selection".

Explanation is the over-arching concept under which causal explanation falls and its function survives the demise of causality (section 6.1.2F). It is instructive at this point to consider reviving a notion of causal explanation by referring to constraints as the new form of causality, as does Juarrero [1999]. As we have just seen, Juarrero's general insight is right that constraints play important explanatory roles in complex systems. Indeed, we could generalise this position beyond what Juarrero considers by noticing that in canonical cases the dynamical flow itself can be considered as satisfying an internal constraint to minimum action. And we might then be tempted to add that the initial condition of the system provides the final constraint to a particular version of the flow form on a particular occasion. For then it could be claimed that it is constraints of one kind or another that do all the work of specifying what is specific to the dynamics of a complex system. And then 'cause' = 'constraint' returns causality to a universal explanatory role for natural systems.

But this last move is not warranted. First, there is still the character of the dynamical interactions involved that play a crucial role in explaining the dynamical behaviour.[61] Even ignoring this point, second, such a move would provide the label 'cause' at the expense of any distinctive content by deflating cause to any formally necessary condition of dynamical explanation.[62] It is true that in non-canonical systems this neat separation of roles can blur. Change of constraints or initial condition can result in qualitative changes in the character of the ensuing behaviour, just as can changing the interaction (cf. [Wilson, 1990; Bishop, this volume]). But an important feature of many non-canonical systems is that the notion of a dynamical action constraint is not available [Hooker-c, this volume]. Thus while interaction continues to be the locus of production of behaviour, it cannot be characterised as even the formal effect of a constraint. In sum, it is more conducive to clarity about the dynamics of complex systems to drop the notion of cause, retaining that of dynamical explanation. (Of course, this does not affect Juarrero's substantive claims, e.g. that intention is better thought of as a

[61]They do this by specifying the basic dynamical variables and determining their inherent (constraint-free) interrelations, thus determining the particular dynamical degrees of freedom the system possesses. While the action constraint is used as a basis for constructing a dynamical framework within which the motion behaviour may be deduced [Hooker-c, this volume], its dynamical effect is only to introduce a further relationship among the variables, reducing the degrees of freedom by one, but not determining the behaviour.

[62]Cf. [Hooker-a, this volume, note 49]. In canonical systems the three factors in dynamical behaviour — interaction, constraint, initial condition — are typically of very different character: interaction provides what is physically productive of outcome, constraint provides a shaping form and an initial condition simply selects among the differing possibilities of the same form provided by interaction and constraint. For the marble rolling in the bowl, e.g., gravitational force drives the motion, the bowl shapes it by constraining the motion to its surface, and the marble's initial beginning place and velocity selects a particular trajectory on the bowl's surface from among all the possible ones. Only the first of these inherits the distinctive sense of producing the outcome that characterises physical causality. To label all three factors causal is to treat causality as a purely formal notion, so that 'cause' = 'formally necessary condition of dynamical explanation' and 'causal explanation' = 'dynamical explanation by appeal to its formal factors'. It is analogous to asserting that all properties are functional on the grounds that dynamical flow itself may be formally considered the function of carrying an initial system state into a final system state under dynamical action; the formal analogy is granted, but the physical need to distinguish dynamical from other functions remains.

dynamical constraint shaping extended interactions — cf. in support [Christensen and Hooker, 2002; Skewes and Hooker, 2009; Van Orden, this volume], and claims like them.)

Condition-dependent explanation provides the appropriate framework for addressing mechanistic explanation, that is, the explanation of functional effects in terms of their realising mechanisms. Mechanistic explanation requires locating and specifying the operation of the components that make up the mechanism and specifying how their organised interrelations jointly produce, that is, realise, the function to be explained [Bechtel and Abrahamsen, 2005]. This is different in immediate form from the standard covering law formula above. However, once the notion of condition-dependent laws applied to organised networks is available, it can be seen to rely on the same elements. An organisational law is compounded from component universal laws and network constraint conditions (section 6.2.2 above) and this, together with inputs as its initial conditions, is used to explain (by reducing) the function. Cases where the mechanism organisation is self-organised can be included in this formulation so long as the compounding of the organisation law is done after the mechanism organisation has emerged. The mechanism organisation law will then also have emerged, e.g. as the laws of solid iron do on cooling. With or without self-organisation, the mechanism function will still reduce to the mechanism process, being identified with the realising dynamical map provided by the compound law. It is then also possible to explain in detail how and why generalisation of mechanism across organisms succeeds and fails.

However, this leaves two respects in which mechanistic explanation differs from standard covering law explanation. The first respect is that changes in conditions can change the explaining law. Generalisation may fail (commonly enough) because the component material conditions entering the compound law vary sufficiently to change the compound law. For instance in biology differing developmental conditions (genetic or environmental) may suffice to alter the specifics of some mechanisms that have multiple alternative pathways for their realisation. More rarely generalisation may fail because one or more components have changed their character (e.g. through different protein folding) and so changed their basic interaction laws. The ongoing need to understand the particularity of mechanisms in organisms is grounded in such effects and limits the search for easy extraction of general organised network laws (cf. [Bechtel and Abrahamsen, this volume]). The second respect of difference is that the notion of organisation as globally constrained functioning is unique to mechanisms and left implicit because we have no analytic tools for explicitly addressing it. This means that if a model is adequate we often cannot, or cannot easily, tell analytically whether it is redundant (cf. CLOCK, [Bechtel, in press]) and if it is not adequate we cannot systematically analyse where and why.

6.2.4 Condition-dependent confirmation

Another aspect that deserves attention is the enhanced condition-dependence of confirmation. To illustrate the general idea, consider the following extension of Russell's story about the fate of the unlucky rational chicken [Russell, 1959, p.35]. Consider a rational chick born in January and fed each day until December 22 of that year. This feeding evidence has >300 data points and no exceptions and thus, when combined with Reichenbach's straight rule of induction, yields for the chicken the conclusion that it will be fed on the morning of December 25. But instead Christmas dinner beckoned. Russell properly did not take the failure of the induction to indicate chicken irrationality - the chicken was unlucky, not irrational, it did the best it rationally could with the information available to it. Rather, Russell used the failure to note the limits of animal (and human) pattern extraction.[63] For if the very same evidence is inserted into a wider theory of human cultural practices than the chicken had available, we may deduce, with at least as high a probability as the chicken was entitled to, that on the morning of December 25 this bird will be slaughtered for the dinner table. So much for animal (and thus human) finitude. This is where Russell left the story. But extend the example a step further and we find that something else of importance is at issue: the very evidence that one may rationally take in one information context to support continued feeding (the better the feeding, the greater the evidence the chicken will be looked after) can in another information context be equally rationally taken as evidence that feeding will not continue (the better the feeding, the fatter the chicken, the more attractive it is as dinner). Rationally chosen method then is a function not only of logic but of our representation of the substantive context in which the method is to be applied.[64]

The example can be transformed into a dynamical setting by noting that it is equivalent to the distinction between modelling only processes on a timescale of a day or less (feeding) and modelling processes on timescales from days to years. The very same diurnal data may lead to very different outcomes between being taken separately and being coupled into a longer running annual process. Complex systems intensify this problem many fold. Slow running processes can fail to noticeably influence behaviour for long periods, only to then bring about large changes (e.g. aging processes, the 40+ latency period of the cancer mesothelioma, the centuries-long decline of civilisations due to desertification). Weak interactions, each remaining below measurement resolution thresholds, can massively influence systems at bifurcation points and can combine to produce measurable effects at unanticipatable times (e.g. gossip at elections, consequences of slow ecological

[63]He also wanted to query the extent of lawful pattern in reality, but this does not follow, and is anyway not at issue here.

[64]The phenomenon itself is commonplace enough, though it is sharper than just the idea that more information can change a conclusion, for this is the *same* evidence being used. What makes the case distinctive and of particular methodological significance is the way that the new information does not simply appear as an additional premise but leads to a change in the form of the inference. Setting E = feeding evidence, C = Christmas ritual, F = fed on Christmas day, the shift is from $E \rightarrow F$ to $E.C \rightarrow\sim F$.

species losses appearing under rare climate stressors like extended drought). These characteristic complex systems features and their like all accentuate the shifts in reasoning that would be made between remaining ignorant of them and becoming aware of them, in just the way that coming into possession of knowledge of annual cultural cycles as well as its daily cycles would do for the chicken. There is no way to know prior to studying it, what features a complex dynamics might present. Subsequent to study, we often enough still cannot be sure that we have studied a complex system in the right manner (long enough, under enough sets of initial conditions, with fine enough resolution, ...) to avoid being faced with a Russell chicken dilemma, and sometimes are pretty sure that we have not (cf. [Bishop, this volume, section 3.1]).

These are not just academic niceties, the problem infects the contemporary climate change debate [Snyder, *et al.*, this volume], sharply complicating its risk structure, and most other complex system policy issues (e.g. military policy - [Ryan, this issue]). This underlines the value of anticipatory response methodologies, from finding tell-tale tests that are relatively independent of having to predict long term system behaviour (e.g. testing for climate change anthropogenesis, [Snyder, *et al.*, this volume]), to developing resilient responses to possible future outcomes (e.g. mitigating climate change through development of new agricultural and silvicultural varieties, developing increased coastal dune protection, and so on), to reducing risk profiles (e.g. through carbon emission reduction, adopting precautionary principles for high risk emissions, etc.). Without those, method is continuously balanced between increasing confidence and the possibility of serious failure. This is no reason to reject either rationality or empirical science — and we have been wondrously lucky in science thus far[65] — but it is to appreciate a distinctive feature of complex systems epistemology.

6.2.5 Epistemic value, ontology and the limits of model-centred confirmation

The study of complex systems has given new twists to several versions of an old problem: where and why should science be data-driven and model-driven? By the former is meant, at the extreme, relying wholly on a methodology of generating data experimentally and searching for patterns in it, and relying on these for prediction and control of future system behaviour, i.e. future data. Models are by-passed. In contemporary *machine learning*/statistical inference parlance, the method is model-free or non-parametric, it makes no assumptions about the internal nature and dynamics of the system under study. By the latter is meant the more traditional *model learning* method where a dynamical model or class of alternative models of a system or system-perspective are proposed, typically specified in terms of differential equations, measurable consequences are deduced and, by testing these experimentally, the model is (models are) either

[65]See [Hooker, 1994b] on the lucky assumptions that lie behind Newton's brilliant success and make it possible in the circumstances. In the light of successive discoveries that workable fundamental theories were in fact lucky degenerate idealisations [Hooker, 2004], we may appreciate that the risk structure of method is likely always subtler and more important than supposed.

confirmed and the testing extended or the model(s) modified and re-tested. In the same contemporary parlance this is parametric modelling, because a quantitative mathematical model (e.g. a set of differential equations) is typically specified in terms of a set of parameters whose values determine the form of the equations and hence the character of the resulting dynamics.[66]

There has always been a tension between these two emphases in science; mostly it has been considered merely pragmatic, that ultimately model-free methods are simply complementary supports of model-driven science. But not always; for instance, last century it came to reflect the methodological tension between the impulse to direct testing driven by theoretical conjecture through Popperian hypothetico-deductive procedure and the opposing impulse to restrict theory to what appears in the data through empiricist inductive procedure, two approaches that were represented as mutually excluding.

Perhaps the best known early scientific version of this tension erupting into cleavage was the disagreement between Wundtian intuitionist psychology and the nascent behaviourist psychology of Watson and Skinner in the early twentieth century. Traditional psychology built a model of the inner person from personal or experimentally-based experience. Wundt's emphasis on intuition simply emphasises the chief point that, however accomplished, traditional psychologies agree on assuming that this inner-model method was the only one to adopt and that the inner model would be one that was cast in intuitively recognisable agency terms. Behaviourism rejected both assumptions, insisting that a pyschological science should be constructed only from external, physically (not psychologically) specified behavioural data — stimulus and response, i.e. input and output, data

[66]In model learning, a class of mathematical models — specified by output being some currently unknown function of input variables and internal parameters — is chosen as a presumed model of the underlying reality generating the empirical input/output relations expressed in the data. The parameters are interpreted in terms of the entities and potential dynamical processes thought to constitute the underlying reality. The problem for statistical methodology is then to use the input/output data in an unbiased way to estimate the parameter values and so fix the particular model involved. This model can subsequently be tested by its prediction of new data and the parameter values re-estimated as required. In machine learning, by contrast, no specific model is specified, rather a general model form is assumed and the data are used to refine it using a machine learning algorithm (some one of a large class of convergent mathematical adaptation or self-correction processes, for example neural nets, run on computers). The refined form is then used to predict further data and is tested against the outcomes and re-worked as necessary. Although it is described as a model-free method, there are assumptions about the adequacy, accuracy and reliability of the data to the system reality implicit in the method (cf. [Lewontin and Levins, 2000], and the 'no free lunch' theorems of Wolpert [1996; 2001] show that these methods are not entirely free of assumptions about the nature of the system generating the data since some general form must be assumed in order to begin the refinement process. However, it is typically much more general than a class of differential equations, as is made clear by the output refined form often having no obvious understanding in physical model terms; indeed its state dimensions, hence parameters, emerge from the refining process and may be very large. Nonetheless, in a variety of situations it provides predictive performance superior to that of parametric models similarly refined against the data. Moreover, with modelling goodness-of-fit tests often too weak to select among a variety of models, it emphasises prediction as a self-sufficient goal of science.

— and should construct psychology in terms of patterns in that data, eschewing all internal models. Since a person is a very complex system, we can today understand each side as in effect pushing one of the two methodologies above for this class of systems, and it would be interesting to re-read that history with this in mind. From a contemporary perspective a particular problem with the behaviourist approach was that (ironically) it's method was often made out to be an a priori normative matter of what science ought to be (cf. empiricism above), and its adherence (ironically) became tribal, sometimes semi-religious (cf. Skinner's utopian works), thus distracting from the main epistemic/ methodological issues. Methodologically, data analysis became a matter of using simple correlations and intervening variables (= constructed as input/output functions) to re-do associationist psychology externally. This proved a disappointingly stringent method in face of the complexity of persons and when internal computational models became available, with all their rich variety yet disciplined structure, they were rushed, returning the dominant method to parametric modelling in directly psychologically interpretable terms. Unfortunately, this equally buried the underlying epistemic issues.

Meanwhile, more pragmatic versions recurred throughout engineering — most closely to psychology, in control engineering. A closed-loop direct-feedback controller, where it is possible, is often much simpler and more reliable than an open loop controller computing inputs from the current model state fed to a model of the target system. Consider the problem of maintaining room temperature. One way to do this is to equip the room with a heater and a thermostat with a set-point feeding back to switch the heater on and off. No model of the room is involved, the feedback is entirely driven by the output temperature data. Alternatively, suppose the heater run by developing a thermodynamic model of the room, including air circulation patterns, the conductive heat loss through the walls, floor, ceiling, and windows, etc., and using its current state to determine future heat loss and thus an electrical input to the heater. The closed-loop controller is much simpler than the open-loop alternative, and more reliable across a broad range of common conditions. Under those conditions it is not necessary to know the internal nature of the room in order to control its temperature.

By contrast, the model-driven open-loop controller, precisely because it is more complex, is vulnerable to modelling errors — the more room parameters the more opportunities for error — and may also suffer from control delays due to computational processing time. Moreover, these problems are increased if the room's character alters during the control period (e.g. a door is opened), especially if the timescale of these changes is less than or comparable to the control signal computation time. The use of such a model in a closed-loop controller can improve performance because the temperature output is now available as feedback. However, neither of the issues of errors and computational complexity is removed. And if the errors problem is reduced by using an adaptive model-updating (learning) process then this is likely to exacerbate the computational burden and may fatally slow feedback. On the positive side, the model-driven control design provides the *possible* behaviours of the system. This allows discrimination of which controller actions, and which preventative non-controller actions (such as insulating the floor and ceiling), will be most *effective* in particular classes of circumstances; and can tell when the model's *validity* domain has been violated. In each of these respects it provides much greater insight than does direct model-free control. In these differences we already see the ever-present tensions among basic epistemic values driving science, here predictiveness, reliability and control versus validity, explanatoriness and effectiveness, with scopes split competingly between them.

To this the advent of new pattern extraction methodologies, driven by the need to cope with complex systems data, adds a new factor. Although it was not obvious to most at the time, this story also has an early example in psychology. The dominance of computational psychology was first challenged by the revival of neural net learning models, see [Hooker-a, this volume, section 4.2.5]. There the point was made that the secret of their success was sub-categorial processing: neural nets process information and store it in their internal ('hidden') layers in terms of non-logical (typically mathematical) functions of the input and output components, whence the output classification is not arrived at from the input by any process of purely logical analysis. Here the lesson to be drawn is that the terms in which the underlying dynamics is represented are sub-personal or sub-psychological (cf. [Smolensky, 1988]), using them involves rejecting the long and widely held assumption of adhering to agency categories when constructing psychological theory.[67]

This opens up a further issue. Neural nets seemed like a natural model of neural cognitive processes, i.e. like the central part of a model-driven method of modelling cognitive capacities. That is, e.g., how Churchland [1989; 1995] and others viewed them (including myself at the time). But they are only weakly neuron-like (namely, solely in their network interconnectedness), so there is limited reason to privilege them on natural modelling grounds. And as they are currently commonly used, they simply form one class among many pattern extraction learning processes (called 'learning machines' because their processes can be automated; see [Hooker-a, this volume, section 4.2.5, note 104] and text). So there is limited or no reason to privilege them when it comes to developing a general theory of learning. Finally, as noted above, neural nets, in common with other learning machines, may have intervening states that have no obvious understanding even in physical model terms, e.g. with emergent state dimensions, hence parameters, much larger than any cognitive process could explicitly use. However, there have been recent developments in control using neural nets that do provide them a potentially distinctive role (e.g. [Wang and Hill, 2006]), though whether either the functional distinctiveness or any special relation to nervous system function are born out remains to be seen. All told, from this perspective neural nets, which seemed a natural neural- network model, are ironically perhaps better seen as part

[67]There is the unpalatable alternative of creating an arbitrary implementation/modelling divide by stipulating that the neural foundations of psychology lie outside of the psychological domain. This option is not pursued here.

of a powerful class of data-driven, model-free learning methods in which attention shifts from physical interpretation and explanation as a goal to instead achieving predictive accuracy in the data domain.

Complex systems, allied with powerful contemporary measurement technologies, present us with data sets of huge size and (this is crucial) spanning many dimensions (i.e. their quantitative components characterised by many linear independencies). Further, often the data is sparsely distributed (i.e. data points expressing specific combinations of values are rare) and expresses many subtle interrelations across data dimensions, space and time. Such circumstances raise the salience of model-free methods because it is impossible to intuit patterns within the data sets and difficult to apply either traditional inductive inference or successful system model learning (cf. e.g. [Lorenz, 1989]). However machine learning models often work effectively off these data sets to reveal important patterns, such as multi-variate interdependencies, including attractors, strange or otherwise. And it often finds them efficiently, without being weighed down by ambiguous internal models when model goodness-of-fit discrimination is poor. This starts a turn to model-free methods that is reinforced by the high numbers of parameters and delicately balanced non-linear dynamics that many of these systems display. The turn is reinforced because the consequent large potential for modelling errors (cf. section 5.2 above) and the often radical difference of the underlying processes to visible surface behaviour (cf. neuron interaction with self-perceived agency action) makes effective modelling seems a distant prospect.

The situation was recently summed up in a paper by [Breiman, 2001]. There he opposes two approaches to data which he refers to as using stochastic data models and algorithmic models, i.e. model learning and machine learning, and argues for giving the latter recognition as a separate approach to scientific methodology.[68] Thus the larger methodological tension between explanation/understanding and prediction/control has been re-invigorated by the rise of complex systems, where it threatens to pull apart prediction from ontological-dynamical understanding as epistemic goals of science. There is currently a lively version of the debate in systems biology.[69]

Breiman is right about the rising importance of model-free pattern extraction methods. The extending reach of scientific investigation into this planet's complex systems will ensure the methodology expands its application range and this will also inevitably force its deepening. But will, or must, this be to the exclusion of model-driven learning? This seems unlikely and the engineering example above shows why, each approach has both strengths and weaknesses, with the strengths of one complementing the weaknesses of the other. For instance, small mistakes in model structure can become amplified by data to increasingly wrong system

[68]Parts of the following discussion of Breiman are adapted with modifications from [Fu and Hooker, 2009, section B.6], where they first appeared.

[69]See, e.g., [Kell, 2006; Kell and Knowles, 2006; Westerhoff and Kell, 2007]. Philosophers may fear that the machine learning approach may bolster either a new philosophical empiricism (cf. Behaviourism) or a post-modern anti-realism (machine learning solipsism); however these extremes do not follow and will not be pursued here.

models; but, equally, for machine learning small mistakes in data structure can become amplified by data to increasingly wrong predictions. The chief strength of the model-free approach is precisely that it is able to extract patterns without having to rely on model-specific information. This allows it to function when circumstances are as outlined in the preceding paragraph. Correlatively, it is just this reliance (on model specifics) that becomes the chief weakness of the model-driven approach under those circumstances. However, in other circumstances permitting some reliable grip on model specifics, e.g. as organic chemistry confers on intracellular biosynthetic pathways, the strengths of the model-driven method comes more into play.

For instance, the representation of possibility includes distinguishing law-like relations (as energy transform processes) from mere correlations or noisier relations, and also identify the sources of noise and bias, including in the interaction between system and data-gathering instruments. And model matching across scales and domains widens these discriminations. The outcomes structure future testing and assessment regimes, including the filtering and correction of data itself. Conversely, and because it cannot distinguish possibility from actuality and (relative) necessity, the machine learning approach still relies on the choice of data categories, experimental set-up and appropriate instruments and probes to generate its data. But all such choices make presumptions about the character and salience of features in the underlying reality. For instance, instruments have systematic errors and limitations and unless we have sound insight into how the instruments work we cannot know what their defects are and hence how to process data. Thus instruments themselves are understood through empirically validated theoretical models.[70] In short, the strength of model-driven research rests on the way a confirmed dynamical model can direct research much more effectively than simply trying to collect more data per se. In all these cases machine learning can seemingly only combine the data pools without direction. But the same data that would convince, say, of instrumental bias is available to machine learning and while the methodology is as yet in its infancy, it is not yet clear that data lists admit no methods for simulating dynamical discrimination and unification, systematic data errors and data limitations.[71] Certainly, all this leads to increased systems size and hence to increased, perhaps insatiable, demands for sufficient data. For example, it might ultimately extend to encompass all science and the entire universe as a block, not a good theory of epistemic (learning) strategy for finite agents beginning in ignorance. So whether this will prove to be a self-undermining development remains to be seen.

[70]Often enough using the very theories they are used to test; but this is largely acceptable, see [Hooker, 1975]. A striking demonstration of this comes from the quantum theory proof that even core classical measuring devices have inherent error rates, of which science had been entirely unsuspecting. On data choices see, for example, [Lewontin and Levins, 2000].

[71]But the difficulties are clear. Even identifying random data errors may cause problems here, since these have somehow to be distinguished from inherent dynamical fluctuations in the system (e.g. chaos), the latter behaving as noise except near regions of strong nonlinear amplification, including bifurcations, where their form may be critical to understanding the system dynamics.

There remains the question of the in-principle relationship between the two methods. Are they, e.g., equivalent given complete data, so that in this sense they are equally fundamental, or is model-driven method more powerful, as the engineering example suggests, even if machine learning method is more widely applicable and faster? Suppose it to be true that External Sufficiency [ES]: every internal difference must ultimately make an input-output difference. Then we might suppose that, given sufficient input-output data, a complete internal model can be inferred (in engineering parlance, reverse engineered). But this would be too fast. First, it is not obvious that ES always holds. Compare here Nelson's earlier attack on behaviourism by arguing that there were internal states inaccessible in principle to characterisation by intervening variables, and hence that ES is false (see [Nelson, 1969; 1976]). Second, ES in itself does not guarantee that the external differences will be distinctive. If the external effects of two or more internal conditions coincided then worthwhile reverse engineering would be defeated. Thus ES is necessary but not sufficient to show machine learning equivalent to model learning. It is also necessary to show Unique Reverse Engineering [URE]: it is always possible to reverse engineer a unique underlying model from a complete set of intervening variables. URE is immediately dubious on the grounds that organisation cannot be directly expressed, only as collections of interdependencies, so it seems impossible to single some interdependencies out as those originating from organisation.[72] Finally, though necessary, URE is also not sufficient for the purpose. To guarantee the constructibility of the unique underlying model it is further necessary to show that Machine Learning Equivalence [MLE]: complete input-output data forces all machine learning processes to converge on the same intervening variables. Given the wide diversity of representations across machine learning devices (e.g. hidden layers and tree structures), MLE looks unprovable and in fact unlikely.

These considerations, taken overall, reinforce the view that it is a false dichotomy to oppose prediction to understanding, since each is necessary to the other: understanding without prediction is uninformed and uncritical, prediction without understanding is weak (*qua* undirected) and fragmented. This suggests that in scientific practice a pragmatic methodological mixed strategy is called for, reinforced by the many approaches in use that combine parametric and non-parametric methods. If you know nothing about the domain but have enough data (data rich, hypothesis poor), then machine learning may be the best approach, while if you know a lot about the domain then, especially if only a small range of data is available (hypothesis rich, data poor), model learning is the best method. And in between, knowledge-wise and data-wise, the features of the best mixed model will no doubt vary complexly with context. In short, where our

[72] Current RE programmes presume simple system structures that do not include global constraints and organisation, cf. [Bongard and Lipson, 2007; Tegnér *et al.*, 2003] This is the counterpart to the severe constraints on programmes for the extraction of causal order in systems, that presume that not even simple feedback loops are present (cf. [Shalizi, 2006, 2.2.1]). However work continues on increasing the discriminatory capacities of RE – for an interesting development see [Schmidt and Lipson, 2009] — and it is too soon to pronounce on its limits.

dynamical intuition works modestly reliably, model-learning offers faster rates of learning (model improvement) on new data, but where it doesn't machine learning is faster. Machine learning works best where data categories plus completeness are clearest, model learning where dynamical intuitions are most effective.

6.2.6 Condition-dependence, idealisation and epistemic context dependence

Finally, we re-visit the more general issue of context dependence, briefly discussed under ontological issues (6.1.2G). The discussion of van der Leeuw in section 5.3 made the case that condition-dependence is a *dynamical* species of perspectivism arising from the multi-scaling and levels of complex systems and our finite interactions with them. This may already be enough dis-unity for some; the point is that nothing of a more normatively dictated or otherwise ontologically fragmented nature need be involved. In tractable cases these perspective models can be unified as so many special cases of a more complete system model. Although there are many different sorts of traffic jams [Hooker-a, this volume, p. 2], e.g., they have in common delays in driver responses to changed traffic conditions, and a general model may be approached in these terms. The section 5.3 discussion also concluded that the dynamical character of perspectives was not broken even though the applicable epistemic limits are likewise a function of dynamical interaction conditions, as therefore are valid methods also. This much perspectivism, dynamical-ontological and epistemic, already accommodates much of the detailed work (as opposed to the 'philosophy') of locally situated models of science of the sort often favoured by historians and anthropologists of science (Gallison, Latour, etc.), the related analysis of the roles of models in science (Cartwright, Morrisson, etc.) and distributed, agent-oriented models of science (Giere, Shi).[73] However, these researches have raised two kinds of challenge to the degree of unity — and so philosophical realism — that dynamical perspectivism indicates, and to the methodology and epistemology based on it. Since unification, even if complex, constitutes a goal that is fundamental to scientific method and knowledge, and complex systems are implicated in the frustration of its achievement, these challenges are now examined.

The *first* of these challenges to unification arises from the analysis of the roles of models in science. Cartwright [1983; 1999] presents case studies illustrating that good models are often false and that theories are only distantly and complexly related to empirical data via 'nomological machines', = artificially simplified models of particular processes aimed at revealing underlying fundamental laws (see [Cartwright, 1999]). Morrisson [2000] points out that scientists often successfully use several incompatible models of one and the same system, and that attempting to combine even just their 'good points' would lead to weakened explanatory and predictive power. Moreover, scientists often have little interest in reconciling

[73]See respectively history: [Shapin and Schaffer, 1985], anthropology: [Latour, 1987; Gallison, 1997], distributed, agent-oriented approaches: [Giere, 2006; Shi 2001], and for modelling: in science see note 16, cf. the discussion of dynamical models of science in [Hooker-a, this volume, section 4.1.3].

them or elaborating larger models that would encompass them. These kinds of situations encourage the conclusions that scientific knowledge is more seriously dis-unified than dynamical perspectivism suggests, and perhaps ultimately driven by other goals than truth. All this can also be taken as encouragement to drop the supposition of a complex but single real world as the referent of scientific knowledge. These are serious and complex challenges, especially for complex systems science. It is impossible to do more than comment briefly here. The upshot will be to the effect that these challenges indicate inevitable epistemic complications in the process of creating knowledge of complex systems but do not destroy the general realist ontological conception offered by dynamical perspectivism or the methodology and epistemology based on it.

The underlying issue here is not the phenomenon of model falsity per se. The idea that models are often not true, and as often known by their users not to be true, is a commonplace in science (cf. [Wolkenhauer, this volume; Westerhoff and Kell, 2007]) and often remarked on in philosophy of science. Setting aside preparatory learning (fiddling with a model to get to know it), the thrust of their scientific use is captured in the title of a paper by Wimsatt: 'False models as means to truer theories' [Wimsatt, 1987], in which Wimsatt offers a wide range of methodological procedures used in science to this end. Scientists understand that every time a model is artificially simplified it can, and likely will, produce false results in real situations. This causes no problems as long as the nature and size of the errors is understood and manageable (cf. using Newtonian mechanics for space science). But to achieve this understanding it is typically necessary to have the kind of insightful model of the system that initially using artificially simplified models is designed to reveal. The implications of all this for learning with false models needs to be understood.

Nonetheless, the underlying issue of relevance here is that models must reduce complexity to be manageable and instructive, and it is the kinds of reduction of complexity in models that create the falsities. The same issue lies behind the use of incompatible models, for these arise from the fact that it is often possible to reduce complexity in many different, and sometimes incompatible, ways. It was observed at the outset of section 6.2 that many systems may be sufficiently complex that it is most practical to use selected partial models of them, tailored to purpose. However convenience turns increasingly into necessity as complexity overwhelms any value to more holistic modelling. This shift is reinforced by bifurcation and self-organisation, which add internal principled bases for partial modelling, and by absence of holistic analytic representations, so that exploration through computational model simulation is unavoidable. Each of these conditions adds to the pressure to reduce complexity in order to gain methodological and thus epistemic grip on these systems. It is necessary then to investigate model simplification in relation to the challenges raised above to methodology and unity.

Simplifications, that is, complexity reductions, can be divided up into substitutions and idealisations, and idealisation in turn broken up into non-degenerate and degenerate kinds. *Substitutions* reduce complexity by excising parts of a model, typically substituting for them some aggregated signal, while all the remaining parts are unchanged. Examples are (i) an automobile system (engine, drive chain, etc.) with the electronics replaced by just its input and output signals, and a board of directors of a firm with departmental activities similarly replaced, and (ii) Fermi-band conduction electrons with their metallic crystal lattice replaced by general constraints and an office with higher management similarly replaced. These substitutions are reversible, sub-, respectively supra-, systems substituted with 'black box' signal or constraint generators can be undone by adding more detailed models of them to the substituted system model. Constraining a system's dynamics to just part of its state space, e.g. to just an attractor basin, to better reveal the dynamics there, could also count as a substitution (of original constraints by stronger constraints) if it is carefully done so that the process is indeed reversible, e.g. so that the removal of constraints does not alter the dynamical landscape.

While the usefulness of substitution is obvious, it is not possible for sufficiently integrated systems, e.g. those with global constraints, and this includes many complex systems. *Idealisations* instead strip all the affected processes in a system down in complexity, e.g. by assuming frictionless movement, point masses, instantaneous signalling, omniscience, perfect rationality, etc. All idealised conditions can be approximated in practice, but typically only under severely constrained conditions. Frictionless movement, e.g., can be approximated by hovercraft-type air support and perfect rationality by decision making in situations reduced to a binary choice among simple alternatives assessed on single outcome values, but these are rare conditions to encounter. Idealisations are of two very different kinds. *Non-degenerate* idealisations are reversible by re-adding to the idealised model what the idealisation stripped out. Thus to the equation for sliding on a frictionless slope (the idealised model) one may add a frictional force and retrieve the frictional sliding model, to L-system models of organic growth may be added reaction-diffusion processes to set the timescale for the L-system growth rates, retrieving a unified growth model [Green and Leishman, this volume]. *Degenerate* idealisations, by contrast, are not thus reversible. While Newtonian mechanics is the result of idealising relativistic mechanics by setting $1/c = 0$, c = speed of light, there is no set of assumptions that can be added to Newtonian mechanics that will then deliver relativistic mechanics again. Similarly for idealising quantum mechanics by setting h = o, h = Planck's constant, there is no set of assumptions that can be added to Newtonian mechanics that will then deliver quantum mechanics again.[74]

[74]Hooker [1994a, section 4] coined the term 'degenerate idealisation' to refer to the irreversible idealisations because degeneracy in the mathematician's sense refers to structure that is collapsed together. As these last examples show, the major inter-theory relations in physics are related as degenerate idealisations, structured as singular asymptotic relations — see [Batterman, 2002; Hooker, 2004]. In an insightful discussion of models, Humphries [2004] discusses the 'correction set' of a model, = the set of ways in which, by undoing, a model can be made more realistic. in that discussion (pp. 78-9) he uses a different terminology to specify a non-equivalent partition of the same range of cases: idealisation = degenerate or non-degenerate idealisation;

For present purposes, the key feature is the relationship between reversibility and compatibility and the contra-positive relationship between irreversibility and incompatibility. Because they are reversible, substitutions and non-degenerate idealisations (of the same system model) are essentially compatible, any group of them can be reversed to yield the full model.[75] For the same reason the resulting substituted/idealised models will improve in accuracy as these complexity reductions are reversed. This is essentially Laymon's [1985] response to Cartwright. This leads to a simple Galilean method and epistemology (section 5.1 II) for their investigation: start with simplified reversible models of convenient sub-systems and slowly build a more complete and accurate model by conjoining them (cf. [Simon, 1961]). The need to study complex systems piecemeal is reduced to a pragmatic methodological problem. However, even here if sub-system models are not systemically interrelated when they are re-integrated, but simply logically conjoined, then they too might appear to show the defects Morrisson decries.

However, as noted, many complex systems are not amenable to these ways of reducing complexity and, as the examples above show, the resulting degenerately idealised models are not mutually compatible. Even though they share Newtonian mechanics as a common idealisation, relativistic and quantum mechanics make incompatible predictions. Indeed, there is still no successful unification of them in some more encompassing model structure (presumably of which they would

abstraction = substitution (subtracting variables) or non-degenerate idealisation (subtracting friction); (restricting) constraints = substitution (reduces case range); (coarsening) approximation = substitution (aggregation) or non-degenerate idealisation (replacing detail by a variable). This division is perhaps more recognisable to scientific practice, but it misses identifying the basic kinds of changes that can be made to models. (It also encourages his mistaking, at p. 147, an increase in power for an increase in complexity wrought by some degenerate idealisations, e.g. an ideally rational agent, and thus to needlessly reject use of simplification terminology.)

[75]Technically, any set of mutually independent substitutions and non-degenerate idealisations (of the same system model) are compatible in the sense that they can be reversed in any order and will yield the unique full model, while for each set of substitutions and non-degenerate idealisations that contains at least one mutually dependent pair there is at least one order of reversals that will yield that full model. To illustrate consider the movement of a rotating spherical object O with a small protruding bump at one location falling vertically through a viscous fluid F under two cases (A) O is rigid, (B) O is deformable, and for the application of two idealisations (I1) O is spherical, (I2) F has zero density. Further assume that the shape of O differs as between crumpling under pressure the O with bump (shape 1) and inserting a bump into a crumpled spherical O (shape 2), due to non-linear interaction between crumpling and bump-making processes. In both cases, A and B, the idealised motion under I1 + I2 is a straight line while the motion in the full model is both retarded (due to viscous drag) and curved (due to asymmetrical drag). In case A first re-introducing the bump makes no difference (0 drag still) and then raising the density yields the full motion since it acts on both sphere (retardation) and bump (drag), while first raising the density produces retardation then re-introducing the bump adds in the asymmetrical drag. The reason for this interchangeability between I1 and I2 is that the case A rigidity condition ensures that I1 and I2 are mutually independent in their effects. But in case B the non-linear interaction between crumpling and bumping removes the independence. First re-introducing the bump again makes no difference (0 drag still) and then raising the density yields shape 1 motion, while first raising the density produces both retardation and spherical crumpling, leading to a retarded and already curved motion, and then re-introducing the bump yields shape 2 motion. No matter which of shape1 or shape2 motion is chosen as the full motion, only one order of reversal returns us there.

themselves in turn be different idealisations). Conversely, the same system may have several incompatible degenerate idealisations, e.g. the fluid dynamics case presented by Rueger [2005], where motion near a wall is very different to that well out in the stream even though both are idealisations of the full flow equations. Such differences should not be surprising. While substitutions and non-degenerate idealisations remove structure from a model, they leave the form of its processes untouched. In the mechanical cases, e.g., the kinematics (possibilities of motion) remain untouched. But degenerate idealisations systematically alter the forms of the processes. In the cases of relativistic and quantum mechanics, e.g., the idealisations strip out two very different higher order structures. So if there is more than one way to do this we cannot expect to interrelate the diverse pairs in any simple way, and we cannot expect any simple conjoining of the idealisations, or even of their 'good points', to be any improvement, indeed we should mostly expect it to be incoherent.[76, 77] This is consistent with Morrisson's complaints [2000] and Rueger's response [2005].

None of this complexity is reason to relinquish the natural realism and unity of dynamical perspectivism, since it follows from the dynamics itself. And recall, there is a single general dynamics underlying and generating all this diversity. To the contrary, it has drawn attention to an underlying current of unification that tends to be neglected in the focus on partial models, namely the discovery of increasingly encompassing theories that unify the foundations of particular disciplinary domains, both by undoing substitutions and non-degenerate idealisations and (most impressively) by transcending and unifying degenerate idealisations, local and domain-wide.[78] This provides its own sense of explanatory power and objectivity. The real situation in science is that there are two currents operating, the more encompassing one pressing the use of partial models of complex systems for dynamical and epistemic reasons, with dynamical perspectivism its outcome and, compatibly within that, the other pressing the expansion of direct dynamical

[76]This should remain so whether idealisations of diverse models result in a common model, as in Newtonian mechanics, or the pair results from diverse idealisations of the same model, as its proponents might hope to happen from a String Theory.

[77]Nor, given their epistemic interests and working constraints, should scientist disinterest in undoing idealisations be surprising. Quite reasonably, scientists only go to the trouble of undoing reduced complexity when the result is expected to illuminate some new feature. Since they can do so relatively straightforwardly for substitutions and non-degenerate idealisations, it detracts from the discovery value (but not from the effort) of doing so. On the other hand, there being no expectation of doing so at all for degenerate idealisations, there is little incentive to try.

[78]The spectacular case is physics, with Newton unifying the dynamics of gravity in heaven and earth, Maxwell unifying the dynamics of optics, electricity and magnetism and Einstein unifying Newton's and Maxwell's theories in Relativity, with the prospect of unifying all the forces under relativistic quantum fields now in play. But there is also the unification of life forms under cellular biology and of biology with bio-chemistry now well advanced, the unification of inorganic chemistry under quantum bonding and the unification of organic chemistry now in play, the spreading unification of functional cellular biology and control engineering under synthetic biology, and so on. It is this sense of strong, if complex, unity that ultimately defeats those who try to appeal to isolated cases to discredit this or that unfavoured scientific theory or finding (e.g. anti-evolutionists, creation scientists, racists, etc.)

unification of foundations. Both share robustness-structured notions of objectivity, evidential reliability, and the like. Rather than seeing these two currents as inconsistent or ignoring foundational unification, as there is currently a tendency in philosophy to do, this complex situation is reason to reflect on what the learning process might be that generates both.[79]

The *second*, and in my view not so serious, challenge to a dynamical perspectivism comes from more general claims about the inherently personal (subjective) and/or values (normatively) based nature of knowledge. The more radical historians and anthropologists, together with kindred postmodernists and model analysts, supported more radical contextualisms than dynamical dependence. In this respect, there is an inviting slide from a dynamical to a larger epistemic reading: conditions are a species of contexts and assertions of context-dependent knowledge or perspectivism have a recent history in which they refer to everything from personal subjectivities (normative perspectival personalism) and communally constructed knowledge (normative perspectival communalism) to philosophical relativism (normative perspectival solipsism). The key to understanding this shift is that the larger positions address the normative status of cross-context relations, while condition-dependence addresses only their dynamically-based interrelations. (This essay tries to confine 'condition' to dynamical features and 'context' to non-dynamical features.) In particular, the radical versions all join in restricting or denying subjective and/or normative comparison across contexts.

Condition-dependence does not in itself yield any of these. It could be transposed into any of the radical context-dependencies by adding a claim denying in principle cross-context subjective and/or normative comparison. But it could equally also be equipped with a notion of emergent global normative evaluation (e.g. a Peircean pragmatism) plus an ontological unification relation to yield a plausible perspectival realism.[80] Again, from a scientific perspective the radical alternatives are unmotivated except by apriori non-naturalist decisions. Values, including those epistemic values driving science, are clearly variable and amenable to being modified by interaction. Given that, it is theoretically more productive as well as better empirically supported to treat social dynamics, including science processes, as normatively variegated and evolving/developing through negotiation dynamics (see [Hooker-a, this volume, section 4.1.3; Harms, this volume]) than

[79]On foundational unification see [Hooker, 1991] and on unificatory objectivity see [Hooker, 2000] and references. In fairness it needs to be noted that [Morrisson, 2000] provides a careful historical scrutiny and sceptical assessment of this unificatory current, arguing that it neither provides increased (causal) explanatory power or support for realist ontological claims or even provides any single conception of what unification is. It can be agreed immediately that formal dynamical unification by itself carries none of these things (cf. the same wave equation for fluid and heat flow), but this is scarcely the end of the matter. However, evaluating all the cases is beyond this essay.

[80]Orange [1992; 1995], concerned with personal contexts, may be close to the latter pragmatist position while many of her compatriots in the psycho-analytic tradition tend toward more radical personalist contextualisms. This is not just an arcane matter; it will take clarifying such issues before the impact of complex systems on clinical psychology, currently entering it through this route and family dynamics, among others, can be settled.

to treat it as normatively partitioned by decree (that is, by a decision made on non-natural normative grounds located in apriori assumptions, philosophical or ideological). And to similarly treat values as emerging from biology (e.g. as per [Hooker-a, this volume, section 4.1.1]) and constructed communally analogously to the way the epistemic regulation of science is (as per [Hooker-a, this volume, section 4.1.3]).

Beyond this debate about perspectival unification, any further discussion of epistemic unification is really about developing a methodological model of how unity is constructed (see e.g. [Bechtel and Hamilton, 2007] for history). Everyone agrees that the old model of unification through deduction from universal fundamental laws of physics was at most only part of the story. But for finite, error-prone and initially ignorant agents like us the complexity of methods required to investigate the world is so large, and new to our thought, that no settled unification model has emerged. Mitchell [2003; 2004; Mitchell and Dietrich, 2006]), e.g., emphasises the prima facie large variety of integrations that may occur where two scientific domains overlap, from part-whole and cause-effect relations to mutual re-configuring to reduction.[81] She draws them together (unifies?) them under three kinds of condition-dependent model integration — mechanical rules, local theoretical unification and explanatory concrete integration[82] — to form an 'integrative pluralism'. Hooker [2000] earlier suggested a similar ecological model of unity, to emphasise the interrelatedness of diverse perspectives and the methodological rule that partial models should be 'matched up along the edges' - that is, in any sub-domain where two models overlap, for instance the bond formation and quantum electron exchange models of chemical bonding, or atmospheric and ocean flow models, the models should agree. But these are not the only principles at work (see above and [Hooker-a, this volume, section 4.1]).[83] These first efforts at characterising the complex unity through plurality of science are likely to prove both too weak and too simple.

Unity is forged from a maze of mutually correcting interconnections among scientific investigations. We discover in practice sub-disciplinary, inter-'level' and high order abstract cross-disciplinary models (Farmer's Rosetta stone for adaptive processes, small worlds network dynamics) and more, utilising a myriad fixated historical conditions and condition-dependent laws, useful approximations and ide-

[81]She refers especially to the work of Dardin [Dardin and Maull, 1977; Dardin, 1986] and Bechtel [Bechtel, 1986; Bechtel and Richardson, 1993]. See also [Bechtel and Abrahamsen, 2007; Bechtel and Hamilton, 2007; Dardin, 2005; 2006].

[82]Roughly and briefly, *mechanical rules* are used to add together independent causes from the two domains (e.g. by vector force addition), *local theoretical unification* is developing a model that unifies all information for a specific class of conditions (only), e.g. the unification of top-down and bottom-up models of some phenomenon, and *explanatory concrete integration* occurs when several partial processes of a complex system are integrated into in a single dynamical model (more or less tightly).

[83]Hooker [2000] also discussed the profound inner unity required by Newton's methodology and exhibited in his planetary theory, a further important dimension that has not been discussed so far and only partly captured in the unity of organisms of the ecological model. It is not further pursued here.

alisations, all creating idiosyncratic webs of demands to match up. If it is more tightly and idiosyncratically interconnected than an ecology, it has more underlying systematicity and connectivity than may appear in Mitchellian integrative or Cartwrightian dappling models.[84] Something like a neuronal web may prove more apt for scientific knowledge for, like a young brain having its networks sculpted as its learning develops, scientific learning is expressed in science's developing structure. The work on models and unification in science (cf. note 16) is much richer than just the discussion of simplification here and provides a rich domain for understanding the complexities introduced by complex systems models across the sciences, one that largely remains unexploited. This will also lead to a better understanding of the kind of unity complex systems science provides. Achieving maximal unificatory interconnection is uneven and far off nearly everywhere. Even so, it is a single real world, and it includes the complex systems that are learning about it.

7 CONCLUSION

The complex systems revolution is currently exploding through science, transforming its concepts, principles, methods and conclusions. It is also transforming its disciplinary structure, both creating new, distinctive 'complexity' disciplines, such as climate science, systems and synthetic biology and self-assembling/repairing and social robotics, and transforming older disciplinary relations, e.g. between developmental biology, psychology and sociology. This revolution creates a plethora of new problems and challenges for foundations and philosophy of science. These have a special intellectual appeal, because the foundations of the science of complex systems is itself still being invented. This dual revolution in science and philosophy is the most important large scale development in scientific cognition for a century. It invites the urgent attention of scientists and philosophers alike.

[84]Da Costa and French [2003] claim that introducing an extended logical formalism of partial models and partial truth can resolve many of the difficulties. (They deploy the technical apparatus of formal model theory, where a model of a set of propositions P is any set of formal objects and relations that satisfy P = for which P is true.) It may well be the case that their formalism clarifies a number of formal issues, it can certainly provide a correctly disciplined way to speak of partial models. But it cannot substitute for dynamical relations and the assumption that it can, that dynamical relations can be logically captured by some formal relation, however complicated, is likely to profoundly mislead. For instance, it proved demonstrably inadequate to capture reduction/emergence relations, where it cannot distinguish (i) aggregated or summed collective patterns where reduction succeeds from (ii) dynamical form transformation with top-down constraint formation where structural reduction fails but, for that reason, functional reduction succeeds [Hooker-c, this volume]. Thus dynamical determination, = there being only one dynamical possibility for the collective dynamical state/property, cannot be equated with logical determination, = the collective dynamical state/property is logically derivable from, can be expressed as a logical sum of, its constituent states/properties. More generally, it does not provide the appropriate form for scientific application, hence testing — see e.g. [Humphries, 2004, section 5.4]. Similarly, perspectival relations are dynamically determined and logical structures are unlikely to faithfully model them.

This book is designed to pull together the complex strands of this revolution and focus attention on its foundational/philosophical problems and challenges.

At the outset [Hooker-a, this volume, section 1] I spoke of this volume as a cross-border bastard, properly belonging to neither philosophy or science but an inheritor, and supporter, of both. I hope, gentle reader, that having read thus far you will find the revolutionary phenomenon it discusses to form a vigorous and promising — if still somewhat neotenous and dishevelled — bastard with a life worth living and, indeed, worth your contributing to.

ACKNOWLEDGMENTS

This essay and its opening companion essay have especially benefited from constructive comment on an earlier draft by Bill Bechtel, Mark Bickhard, Thomas Brinsmead, David Green, Jan Schmidt and Olaf Wolkenhauer. I am grateful for their substantial efforts, and to several other authors for particular helpful remarks. Above all I am grateful to my wife Jean whose practical generosity amid her own obligations made a long, time-consuming task so much easier and whose love enlightens everything.

BIBLIOGRAPHY

[Anderies *et al.*, 2007] J. M. Andries, A. A. Rodriguez, M. A. Janssen, O, Cifdaloz. Panaceas, uncertainty and the robust control framework in sustainability science. *Proceedings of the National Academy of Science*, 104(39), 15194-15199, 2007.

[Alm and Arkin, 2003] E. A. Alm and A. Arkin. Biological networks. *Current opinion in structural biology*, 13, 193-202, 2003.

[Auyang, 1998] S. Y. Auyang. *Foundations of complex systems theories in economics, biology and statistical physics.* Cambridge University Press, Cambridge, 1998.

[Barabási and Bonabeau, 2003] A.-L. Barabási and E. Bonabeau. Scale-free networks. *Scientific American*, May, 50-59, 2003.

[Barabasi and Oltvai, 2002] A.-L. Barabási and Z. Oltvai. Life's complexity pyramid. *Science,* 298, 763-4, 2002.

[Batterman, 2002] R. W. Batterman. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction and Emergence*, Boston: MIT Press, 2002.

[Bechtel, 1986] W. Bechtel (ed.). Integrating scientific disciplines. Martinus Nijhoff, Dordrecht, 1986.

[Bechtel, 2007] W. Bechtel. Biological mechanisms: organised to maintain autonomy. In [Boogerd, *et al.*, 2007], 269-302, 2007.

[Bechtel, in press] W. Bechtel. Computational modeling and mechanistic explanation: understanding circadian rhythms. *Studies in History and Philosophy of Science Part A*, in press.

[Bechtel and Abrahamsen, 2005] W. Bechtel, A. Abrahamsen. Explanation: a mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 421-441, 2005.

[Bechtel and Abrahamsen, 2007] W. Bechtel and A. Abrahamsen. In search of mitochondrial mechanisms: interfield excursions between cell biology and biochemistry. *Journal of the History of Biology*, 40, 133, 2007.

[Bechtel and Hamilton, 2007] n96 W. Bechtel, A. Hamilton. Reductionism, integration, and the unity of the sciences. In T. Kuipers (ed.), *Philosophy of Science: Focal Issues.* Handbook of the Philosophy of Science, vol.1. Elsevier, New York, 2007.

[Bechtel and Richardson, 1993] W. Bechtel and R. Richardson. *Discovering complexity.* Princeton University Press, Princeton, 1993.

[Bedau and Humphries, 2008] M. Bedau and P. Humphries (eds.), *Emergence: contemporary readings in philosophy and science*, Bradford/MIT Press, Cambridge, Mass., 2008.

[Bongard and Lipson, 2007] J. Bongard and H. Lipson. Automated reverse engineering of nonlinear dynamical systems. *Proceedings of the National Academy of Science*, 104(24), 9943–9948, 2007.

[Boogerd *et al.*, 2007] F. Boogerd, F. Bruggeman, J-H. Hofmeyr, H. V. Westerhoff (eds.), *Systems Biology: Philosophical Foundations*, Amsterdam, Elsevier, 2007.

[Breiman, 2001] L. Breiman. Statistical modelling: the two cultures. *Statistical Science*, 16, 199-215, 2001.

[Brinsmead, 2005] T. S. Brinsmead. *Integrated Sustainability Assessment: Identifying Methodological Options*, Report to the National Academies Forum, Australia, 2005. Available at `http://www.naf.org.au/naf-forum/brinsmead.pdf`.

[Brinsmead and Hooker, 2005] T. S. Brinsmead and C. A. Hooker. Integration in Sustainability Assessment. Brisbane: Cooperative Research Centre for Coal in Sustainable Development. Available at `http://www.newcastle.edu.au/centre/casrg/`.

[Carpenter and Lathrop, 2008] S. R. Carpenterand R. C. Lathrop. Probabilistic estimate of a threshold for eutrophication. *Ecosystems*, 11, 601-613, 2008.

[Carroll, 2006] J. W. Carroll. Laws of nature. *Stanford Encyclopedia of Philosophy*, `http://plato.stanford.edu/entries/laws{-}of{-}nature/`, 2006.

[Cartwright, 1983] N. Cartwright. *How the laws of physics lie*. Clarendon Press, Oxford, 1983.

[Cartwright, 1989] N. Cartwright. Nature's capacities and their measurement. Oxford University Press, Oxford, 1989.

[Cartwright, 1999] N. Cartwright. *The dappled world. A study of the boundaries of science*. Cambridge University Press, Cambridge, 1999.

[Chandrasekhar, 1961] S. Chandrasekhar. *Hydrodynamic and Hydromagnetic Stability*. Clarendon Press, Oxford, 1961.

[Cherniak, 1986] C. Cherniak. *Minimal Rationality*. MIT, Cambridge Mass., 1986.

[Christensen and Hooker, 2000] W. D. Christensen, C. A. Hooker. An Interactivist-Constructivist Approach to Intelligence: Self-directed Anticipative Learning. *Philosophical Psychology*, 13, 5-45, 2000.

[Christensen and Hooker, 2002] W. D. Christensen, C. A. Hooker. Self-directed agents. In J. MacIntosh (ed.), *Naturalism, Evolution and Intentionality*. Ottawa, *Canadian Journal of Philosophy*, Special Supplementary Volume 27, 19-52, 2002.

[Churchland, 1989] P. M. Churchland. *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Bradford/MIT Press, Cambridge, Mass., 1989.

[Churchland, 1995] P. M. Churchland. *The Engine of Reason, the Seat of the Soul*, MIT Press, Cambridge Mass.,1995.

[Cumming and Collier, 2009] G. Cumming and J. Collier. Change and identity in complex systems. In B. Brown, K. de Laplante, K. Peacock (eds.), *Philosophy of Ecology*. Handbook of the Philosophy of Science, vol.13. Elsevier, Amsterdam, 2009.

[Da Costa and French, 2003] N. Da Costa and S. French. *Science and partial truth*. Oxford University Press, Oxford, 2003.

[Dardin, 1986] L. Dardin. Relations among Fields in the Evolutionary Synthesis. In [Bechtel, 1986, 113-123], 1986.

[Dardin, 2005] L. Dardin. Relations Among Fields: Mendelian, Cytological and Molecular Mechanisms. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 357371, 2005.

[Dardin, 2006] L. Dardin. *Reasoning in biological discoveries: Essays on mechanisms, interfield relations, and anomaly resolution*. Cambridge University Press, Cambridge, 2006.

[Dardin and Maull, 1977] L. Darden and N. Maull. Interfield theories. *Philosophy of Science*, 44, 4364, 1977.

[Diamond, 1997] J. M. Diamond. *Guns, germs, and steel: the fates of human societies*. W.W. Norton & Co, New York, 1997.

[Diamond, 2005] J. M. Diamond. *Collapse: how societies choose to fail or succeed*. Viking Books, New York, 2005.

[Earman, 1986] J. Earman. *A Primer on Determinism*. University of Western Ontario Series in the Philosophy of Science 32, Reidel, Dordrecht, 1986.

[Earman, 1986] J. Earman. Laws, Symmetry, and Symmetry Breaking; Invariance, Conservation Principles, and Objectivity. *Philosophy of Science*, 71, 1227-1241, 2004.

[Earman, 2007] J. Earman. Aspects of Determinism in Modern Physics. In J. Butterfield, J. Earman (eds.), *Philosophy of Physics*. Handbook of the Philosophy of Science, vol.2. Elsevier, Amsterdam, 2007.

[Edmonds and Moss, 2005] B. Edmonds and S. Moss. From KISS to KIDS — an 'antisimplistic' modelling approach. In P. Davidsson et al. (eds.), *Multi Agent Based Simulation*. Springer, Lecture Notes in Artificial Intelligence 3415, Berlin, 130–144, 2004. Available at `http://cfpm.org/~bruce/`.

[Farrell and Hooker, 2007a] R. Farrell and C. A. Hooker. Applying self-directed anticipative learning to science: agency and the interactive exploration of possibility space in ape language research. *Perspectives on Science*, 15(1), 86-123, 2007.

[Farrell and Hooker, 2007b] R. Farrell and C. A. Hooker. Applying self-directed anticipative learning to science: learning how to learn across revolutions. *Perspectives on Science*, 15(2), 220-253, 2007.

[Farrell and Hooker, 2009] R. Farrell and C. A. Hooker. Applying self-directed anticipative learning to science: identifying error and severe testing through interactive exploration of possibility. *Foundations of Science*, 2009.

[Flannery, 1994] T. Flannery. *The future eaters: an ecological history of the Australasian lands and people*. Reed Books, Melbourne, 1994.

[Fox Keller, 2007] E. Fox Keller. The disappearance of function from 'self-organizing systems'. In: F. Boogerd, F. Bruggeman, J-H. Hofmeyr, H. V. Westerhoff (eds.), *Systems Biology: Philosophical Foundations*, Elsevier, Ámsterdam, 303-317, 2007.

[Frigg and Hartmann, 2006] R. Frigg, S. Hartmann. Models in Science. *Stanford Encyclopedia of Philosophy*, `http://plato.stanford.edu/entries/models-science/`, 2006.

[Fu and Hooker, 2009] P. Fu and C. A. Hooker. Outstanding Issues in systems and synthetic biology. In [Fu and Panke, 2009], 615-646, 2009.

[Fu and Panke, 2009] P. Fu and S. Panke (eds.). *Systems Biology and Synthetic Biology*. Wiley, Hoboken N.J., 2009.

[Gallison, 1997] P. Galison. *Image and logic*: *a material culture of microphysics*. Chicago: University of Chicago Press, Chicago, 1997.

[Gell-Mann, 1994] M. Gell-Mann. *The Quark and the Jaguar: Adventures in the Simple and the Complex*. Henry Holt, New York, 1994.

[Giere, 2006] R. Giere. *Scientific Perspectivism*. University of Chicago Press, Chicago, 2006.

[Green *et al.*, 2005] D. G. Green, N. Klomp, G. Rimmington, S. Sadedin. *Complexity in landscape ecology*. Springer, New York, 2005.

[Gunderson *et al.*, 1997] L. Gunderson, C. Holling, L. Pritchard, G. Peterson. *Resilience in Ecosystems, Institutions and Societies*. Beijer Discussion Paper Series, No. 95. Beijer International Institute of Ecological Economics, Stockholm, 1997.

[Gunderson and Holling, 2002] L. Gunderson and C. Holling (eds.). *Panarchy: Understanding Transformations in Human and Natural Systems*. Washington, Island Press, 2002.

[Hayek, 1973] F. Hayek. *Rules and order*, Vol. 1 of Law, legislation and liberty. University of Chicago Press, Chicago, 1973

[Hesse, 1963] M. Hesse. *Models and analogies in science*. Sheed & Ward, London, 1963.

[Holland, 1995] J. H. Holland. *Hidden order: how adaptation builds complexity*. Addison Wesley, Redwood City, Cal.; Perseus Books, Cambridge MA, 1995.

[Hooker, 1973] C. A. Hooker. Physics and metaphysics: a prolegomnon for the riddles of quantum theory. In C. A. Hooker (ed.), *Contemporary Research in the Foundations and Philosophy of Quantum Theory*. Reidel, Holland, 174-304, 1973.

[Hooker, 1974] C. A. Hooker. The Metaphysics of Science: Atoms versus Plena. *International Logic Review*, 9, 111-146, 1974.

[Hooker, 1975] C. A. Hooker. Global theories. *Philosophy of Science*, 42, 152-179, 1975. Reprinted in C. A. Hooker, *A Realistic Theory of Science*, State University of New York Press, Albany NY, 1987.

[Hooker, 1981] C. A. Hooker. Towards a general theory of reduction. *Dialogue*, XX, Part I, Historical framework, 38-59, Part II, Identity and reduction, 201-36, Part III, Cross-categorial reduction, 496-529, 1981.

[Hooker, 1991] C. A. Hooker. Between formalism and anarchism: a reasonable middle way. In G. Munevar (ed.), *Beyond reason: essays on the philosophy of Paul Feyerabend*. Kluwer, Boston, 41-107, 1991.

[Hooker, 1994a] C. A. Hooker. Idealisation, naturalism, and rationality: some lessons from minimal rationality. *Synthese*, 99, 181-231, 1994.

[Hooker, 1994b] C. A. Hooker. From phenomena to metaphysics. In D. Prawitz, D. Westerstahl (eds.), *Logic and Philosophy of Science in Uppsala*. Dordrecht, Kluwer, 159-184, 1994.

[Hooker, 1994c] C. A. Hooker. Regulatory constructivism: on the relation between evolutionary epistemology and Piaget's genetic epistemology. *Biology and Philosophy*, 9, 197-244, 1994.

[Hooker, 1998] C. A. Hooker. Laws, Natural. In E. Craig (ed.), *Routledge Encyclopaedia of Philosophy*, 10 vols. Routledge, London, 1998.

[Hooker, 2000] C. A. Hooker. Unity of Science. In W. H. Newton-Smith (ed.), *A Companion to the Philosophy of Science*. Blackwell, Oxford, 540-549, 2000.

[Hooker, 2004] C. A. Hooker. Asymptotics, reduction and emergence. *British Journal for the Philosophy of Science*, 55, 435-479, 2004.

[Hooker, 2006] C. A. Hooker. Reduction as Cognitive Strategy. In B. Keeley (ed.), *Paul Churchland*. Cambridge University Press, New York, 154-174, 2006.

[Hooker, 2009] C. A. Hooker. Interaction and bio-cognitive order. *Synthese*, 166(3), 513-546, 2009.

[Hooker-a, this volume] C. A. Hooker. Introduction to philosophy of complex systems. Part A: Towards a framework for complex systems. In this volume.

[Hooker-c, this volume] C. A. Hooker. Conceptualising reduction, emergence and self-organisation in complex dynamical systems. In this volume.

[Humphries, 2004] P. Humphries. *Extending ourselves: computational science, empiricism and scientific method*. Oxford University Press, Oxford, 2004.

[Juarrero, 2002] A. Juarrero. *Dynamics in action: intentional behavior as a complex system*. MIT Press, Cambridge, Mass., 2002.

[Kauffman, 1993] S. Kauffman. *Origins of order*. Oxford University Press, Oxford, 1993.

[Kauffman, 1995] S. Kauffman. *At Home in the Universe*. Oxford University Press, Oxford, 1995.

[Kauffman, 2000] S. Kauffman. *Investigations*. Oxford University Press, New York, 2000.

[Kell, 2006] D. B. Kell. Metabolomics, modelling and machine learning in systems biology: towards an understanding of the languages of cells. The 2005 Theodor Boucher lecture. *FEBS Journal*, 273, 873894, 2006.

[Kell and Knowles, 2006] D. B. Kell and J. Knowles. The role of modelling in systems biology. In Z. Szallasi, J. Sterling, V. Periwal (eds.), *System modelling in cellular biology: from concepts to nuts and bolts*. MIT Press, Cambridge Mass., 2006.

[Kitano, 2007] H. Kitano. Towards a theory of biological robustness. *Molecular Systems Biology*, 3, Article 137, 2007. Available at `http://www.nature.com/msb/index.html`.

[Kleidon, 2009] A. Kleidon. Nonequilibrium thermodynamics and maximum entropy production in the Earth system: Applications and implications. *Naturwissenschaften*, 96, 653-677, 2009.

[Kohler and Gumerman, 2000] T. Kohler and G. Gumerman, (eds.). *Dynamics in human and primitive societies: agent-based modeling of social and spatial processes*. Santa Fe Institute Studies in the Science of Complexity, Oxford University press, New York, 2000.

[Lansing snd Axtell, 2005] J. S. Lansing and R. L. Axtell. Afterword. In C. S. Beekman, W. W. Baden (eds.), *Nonlinear Models for Archaeology and Anthropology: Continuing the Revolution*. Ashgate Press, Farnham Surrey, 139-146, 2005.

[Lansing, 2003] J. S. Lansing. Complex Adaptive Systems. *Annual Review of Anthropology*, 32, 183-204, 2003.

[Lansing, 2002] J. S. Lansing. 'Artificial Societies' and the Social Sciences. *Santa Fe Institute Working Papers* and *Artificial Life* 8, 279-292, 2002.

[Latour, 1987] B. Latour. *Science in action*. Harvard University Press, Cambridge Mass., 1987.

[Laymon, 1985] R. Laymon. Idealizations and the testing of theories by experimentation. In P. Achinstein, O. Hannaway (eds.), *Observation Experiment and Hypothesis in Modern Physical Science*. MIT Press, Cambridge Mass., 147173, 1985.

[Levins, 1966] R. Levins. The strategy of model building in population biology. *American Scientist*, 54, 421-431, 1966.

[Levins, 1970] R. Levins. Complex systems. In C. H. Waddington (ed.), *Towards a theoretical biology*, vol.3. Aldine, Chicago, 73-88, 1970.

[Lewis and Todd, 2007,] M. D. Lewis and R. M. Todd. The selfregulating brain: cortical-subcortical feedback and the development of intelligent action. *Cognitive Development*, 22, 406430, 2007.

[Lewis and Cook, 2007] M. D. Lewis and M. L. Cook. Changing habits of emotion regulation at transition points in infancy: A dynamic systems analysis. *Journal of Developmental Processes*, 3, 6789, 2007.

[Lewontin and Levins, 2000] R. Lewontin and R. Levins. Let the numbers speak. *International Journal of Health Services* 30(4), 873-7, 2000.

[Lorenz, 1997] H.-W. Lorenz. *Nonlinear Dynamical Economics and Chaotic Motion.* (1^{st} edn 1989, 2^{nd} rev. & enl. edn, 1993.) Springer, Berlin, 1997.

[Mackie, 1974] J. Mackie. *The Cement of the Universe.* Oxford: Clarendon Press, 1974.

[Maynard-Smith and Szathmary, 1995] J. Maynard-Smith and E. Szathmary. *The major transitions in evolution.* New York, Freeman, 1995.

[Mayo, 1996] D. Mayo. *Error and the Growth of Experimental Knowledge.* University of Chicago Press, Chicago, 1996.

[Magnani and Nersessian, 2002] L. Magnani and N. Nersessian, eds. Model-based reasoning: science, technology, values. New York, Kluwer, 2002.

[McDaniel and Driebe, 2005] R. McDaniel, Jr. and D. Driebe (eds.). *Uncertainty and surprise in complex systems.* Springer, New York, 2005.

[McGlade and van der Leeuw, 1997] J. McGlade and S. E. van der Leeuw. Archeology and Nonlinear Dynamics: new approaches to long term change. In S. van der Leeuw, J. McGlade, (eds.), *Archaeology: Time, Process and Structural Transformations.* Routledge, London, 1-23, 1997.

[McMullin, 1985] E. McMullin. Galilean Idealization. *Studies in the History and Philosophy of Science*, 16, 24773, 1985.

[Meadows *et al.*, 1972] D. H. Meadows, J. Randers, D. L. Meadows, W. W. Behrens III. *Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind.* Universe Books, 1972. (New American Library, 1977.)

[Meadows *et al.*, 2004] D. H. Meadows, J. Randers, D. L. Meadows. *Limits to growth the 30 year update.* Chelsea Green, White River Jct. Vermont, 2004.

[Meadows, 2005] D. H. Meadows. Places to intervene in a system. `http://www.developerdotstar.com/mag/articles/places_intervene_system.html`, accessed April 2009.

[Mitchell, 2003] S. Mitchell. *Biological complexity and integrative pluralism.* Cambridge University Press, Cambridge, 2003.

[Mitchell, 2004] S. Mitchell. Why integrative pluralism? *E:CO* 6(1-2), 81-91, 2004.

[Mitchell and Dietrich, 2006] S. D. Mitchell and M. R. Dietrich. Integration without unification: an argument for pluralism in the biological sciences. *American Naturalist*, 168, S73–S79, 2006.

[Morrisson, 2000] M. Morrisson. *Unifying scientific theories.* Cambridge University Press, Cambridge, 2000.

[Morgan and Morrisson, 1999] M. Morgan and M. Morrisson. *Models as mediators.* Cambridge University Press, Cambridge, 1999.

[Nagel, 1961] E. Nagel. *The Structure of Science.* Routledge & Kegan Paul, London, 1961.

[Nelson, 1969] R. J. Nelson. Behaviorism is false. *Journal of Philosophy*, 66, 41752, 1969.

[Nelson, 1976] R. J. Nelson. Mechanism, functionalism and the identity theory. *Journal of Philosophy*, 73, 36486, 1976.

[Nersessian, 2008] N. Nersessian. *Creating scientific concepts.* MIT/Bradford Press, Cambridge, Mass., 2008.

[Orange, 1992] D. M. Orange. Perspectival realism and social constructivism: commentary on Irwin Hoffman's "Discussion: toward a socialconstructivist view of the psychoanalytic situation". *Psychoanalytic Dialogues*, 2, 561565, 1992.

[Orange, 1995] D. M. Orange *Emotional understanding: studies in psychoanalytic epistemology.* Guilford Press, New York, 1995.

[Oreskes, 2003] N. Oreskes. The role of quantitative models in science. In C. D. Canham, J. J. Cole, W. K. Lauenroth (eds.) *Models in Ecosystem Science.* Princeton, Princeton University Press, 1331, 2003.

[Oreskes and Belitz, 2001] N. Oreskes and K. Belitz. Philosophical Issues in Model Assessment. In M.G. Anderson and P.D. Bates (eds) *Model Validation: Perspectives in Hydrological Science.* London, John Wiley and Sons, 2341, 2001.

[Oreskes *et al.*, 1994] N. Oreskes, K. ShraderFrechette, K. Belitz. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263, 641646, 1994.

[Piaget, 1971] J. Piaget. *Biology and knowledge: an essay on the relations between organic regulations and cognitive processes* (trans. B. Walsh). University of Chicago Press, Chicago, 1971.

[Potts, 2000] J, Potts, *The New Evolutionary Microeconomics: Complexity, Competence and Adaptive Behaviour*. Edward Elgar, Cheltenham, 2000.

[Rueger, 2005] A. Rueger. Perspectival Models and Theory Unification. *British Journal for the Philosophy of Science*, 56(3), 579-594, 2005. n123 p95

[Rueger and Sharp, 1996] A. Rueger and W. D. Sharp. Simple theories of a messy world: theory and explanatory power in nonlinear dynamics. *British Journal for the Philosophy of Science*, 47, 93-112, 1996.

[Russell, 1959] B. Russell. *The problems of philosophy.* Oxford University Press, Oxford, 1959.

[Schlosser and Wagner, 2004] G. Schlosser and G. Wagner. *Modularity in Development and Evolution.* University of Chicago Press, Chicago, 2004.

[Schmidt and Lipson, 2009] M. Schmidt and H. Lipson. Distilling FreeForm Natural Laws from Experimental Data. *Science*, 324(5923), 81 85, 2009.

[Shalizi, 2006] C. Shalizi. Methods and techniques of complex systems science: an overview. In T. Deisboeck, J. Kresh (eds.), *Complex systems science in bio-medicine.* Springer, New York, 2006. Also available at `http://arXiv.org/abs/nlin/0307015`

[Shapin and Schaffer, 1985] S. Shapin and S. Schaffer. *Leviathan and the AirPump: Hobbes, Boyle, and the Experimental Life.* Princeton University Press, Princeton, 985.

[Shi, 2001] Y. Shi. *Economics of scientific knowledge: A rational choice neo-insitutionalist theory of science.* Edward Elgar, London, 2001.

[Simon, 1969] H. A. Simon. *The sciences of the artificial.* MIT Press, Cambridge, Mass., 1969.

[Simon, 1997] H. A. Simon. *Models of bounded rationality, vol.3.* MIT Press, Cambridge, Mass., 1997.

[Skewes and Hooker, 2009] J. Skewes, C. A. Hooker. Bio-agency and the problem of action. *Biology and Philosophy*, 24(3), 283-300, 2009.

[Skyrms, 2004] B. Skyrms. *The Stag Hunt and the Evolution of Social Structure.* Cambridge University Press, Cambridge, 2004.

[Smolensky, 1988] P. Smolensky. On the proper treatment of connectionism. *Behavioral and Brain Science*, 11, 1-74, 1988.

[Tegnér *et al.*, 2003] J. Tegnér, M. K. S. Yeung, J. Hasty, J. J. Collins. Reverse engineering gene networks: integrating genetic perturbations with dynamical modeling. *Proceedings of the National Academy of Science*, 100(10), 59445949, 2003.

[Vallis, 2009] G. K. Vallis. Mechanisms of climate variability from years to decades. In T. N. Palmer, P. Williams (eds.), *Stochastic physics and climate modelling.* Cambridge University Press, Cambridge, 2009.

[Van der Leeuw, 2004] S. Van der Leeuw. Nonlinear processes and archaeology. In A.C. Renfrew, P. Bahn (eds.), *Key concepts in archaeology.* Routledge, London, 182186, 2004.

[Van Ness and Scheffer, 2007] E. Van Nes, M. Scheffer. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *American Naturalist*, 169, 738-747, 2007.

[Vygotsky, 1986] L. S. Vygotsky. *Thought and language.* MIT Press, Cambridge Mass., 1986.

[Walker and Salt, 2006] B. Walker, D. Salt. *Resilience Thinking.* Island Press, Washington DC, USA, 2006.

[Wang and Hill, 2006] C. Wang, D. J. Hill. Learning from neural control. *IEEE Transactions on Neural Networks*, 17(1), 130-146, 2006.

[Westerhoff and Kell, 2007] H. Westerhoff, D. Kell. The methodologies of systems biology. In F. Boogerd, F. Bruggeman, J-H. Hofmeyr, J. Westerhoff (eds.), *Systems Biology: Philosophical Foundations.* Elsevier, Amsterdam, 23-70, 2007.

[Westerhoff and Paulsson, 2004] H. Westerhoff and B. Palsson. The evolution of molecular biology into systems biology. *Nature Biotechnology*, 22, 1249-1252, 2004.

[Wilson, 1990] M. Wilson. Law Along the Frontier: Differential Equations and Their Boundary Conditions. In A. Fine, M. Forbes, L. Wessels (eds.), *PSA 1990,* Vol. 2. PSA Press, East Lansing Mich., 565-575, 1990.

[Wimsatt, 1987] W. Wimsatt. False models as means to truer theories. In N. Nitecki, A. Hoffman (eds.), *Neutral models in biology.* Oxford University Press, Oxford, 1987. Reprinted as chapter 6 of [Wimsatt, 2007].

[Wimsatt, 1994] W. Wimsatt. The ontology of complex systems: levels, perspectives and causal thickets. In R. Ware, M. Mattern (eds.), *Canadian Journal of Philosophy*, Special Supplementary Volume 24, 207-274, 1994. Reprinted as chapter 7 of [Wimsatt, 2007].

[Wimsatt, 2007] W. Wimsatt. *Re-engineering philosophy for limited beings.* Harvard University Press, Cambridge, Mass., 2007.

[Wolpert, 1996] D. H. Wolpert. The lack of a priori distinctions between learning algorithms, the existence of a priori distinctions between learning algorithms. *Neural Computation*, 8, 1341-1390, 1391-1421, 1996.

[Wolpert, 2001] D. H. Wolpert. The supervised learning nofreelunch theorems. Invited contribution to World Conference on Soft Computing, 2001. Available at `http://ti.arc.nasa.gov/profile/dhw/statistical/`

[Wolpert, 2002] D. H. Wolpert. Supervised Learning Theory. In L. Nadel (ed.), *Encyclopedia of Cognitive Science*. Macmillan, London, 2002.

INDEX

This index is predominantly concerned with subject entries, almost wholly as chosen by the essay authors.