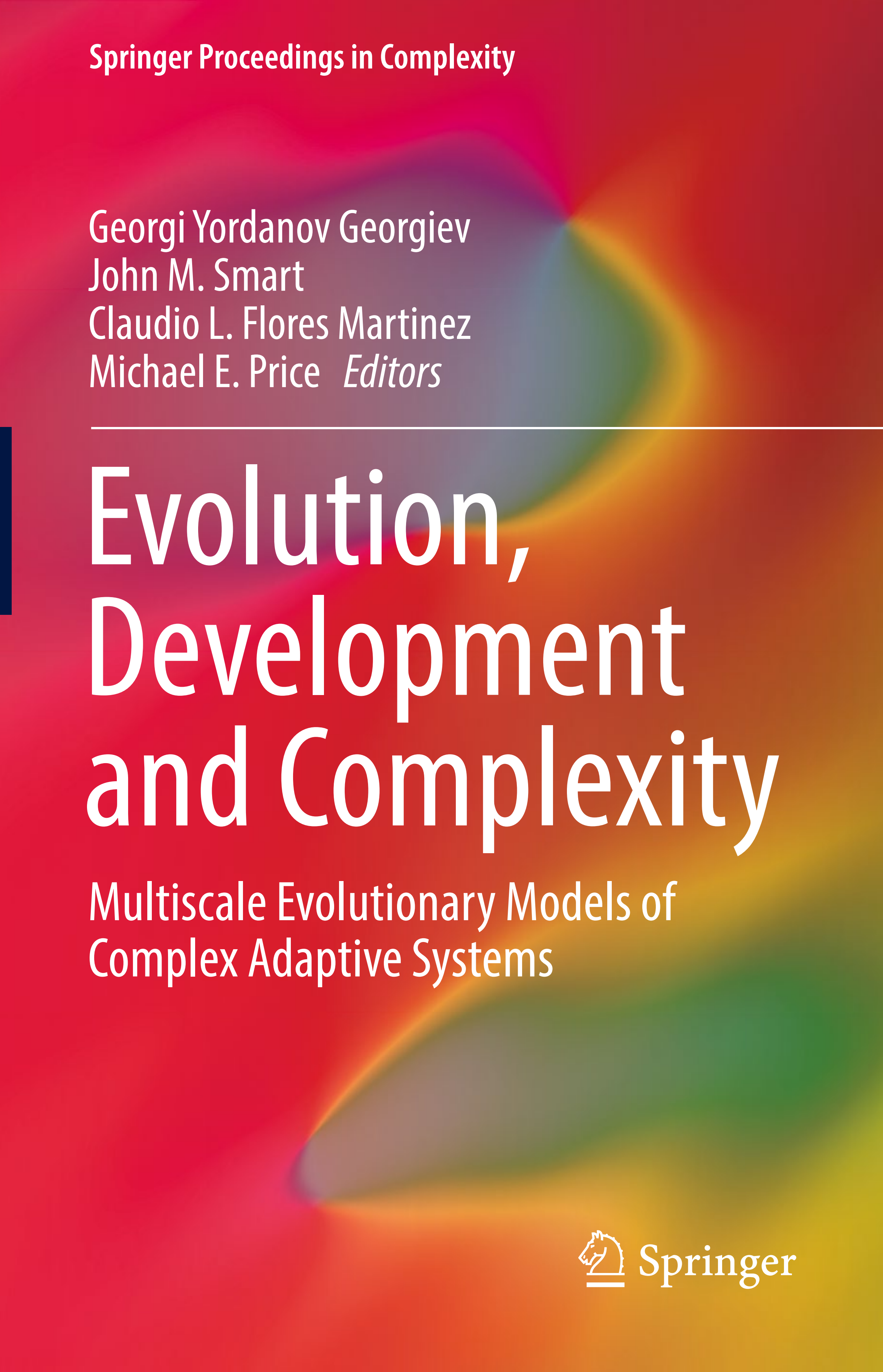


Springer Proceedings in Complexity

Georgi Yordanov Georgiev
John M. Smart
Claudio L. Flores Martinez
Michael E. Price *Editors*

Evolution, Development and Complexity

Multiscale Evolutionary Models of Complex Adaptive Systems

Springer

Springer Proceedings in Complexity

Springer Proceedings in Complexity publishes proceedings from scholarly meetings on all topics relating to the interdisciplinary studies of complex systems science. Springer welcomes book ideas from authors. The series is indexed in Scopus. Proposals must include the following: - name, place and date of the scientific meeting - a link to the committees (local organization, international advisors etc.) - scientific description of the meeting - list of invited/plenary speakers - an estimate of the planned proceedings book parameters (number of pages/articles, requested number of bulk copies, submission deadline). submit your proposals to: christopher.coughlin@springer.com

More information about this series at http://www.springer.com/series/11637

Georgi Yordanov Georgiev • John M. Smart
Claudio L. Flores Martinez • Michael E. Price
Editors

Evolution, Development and Complexity

Multiscale Evolutionary Models of Complex Adaptive Systems

Springer

Editors
Georgi Yordanov Georgiev
Department of Physics
Worcester Polytechnic Institute
Worcester, MA, USA

Department of Physics
Assumption College
Worcester, MA, USA

Department of Physics
Tufts University
Medford, MA, USA

Claudio L. Flores Martinez
Department of Biology
Institute of Zoology
Molecular Evolutionary Biology
University of Hamburg
Hamburg, Germany

John M. Smart
Acceleration Studies Foundation
Los Gatos, CA, USA

Naval Postgraduate School
Monterey, CA, USA

Michael E. Price
Department of Life Sciences
Centre for Culture and Evolution
Brunel University London
Uxbridge, UK

ISSN 2213-8684 ISSN 2213-8692 (electronic)
Springer Proceedings in Complexity
ISBN 978-3-030-00074-5 ISBN 978-3-030-00075-2 (eBook)
https://doi.org/10.1007/978-3-030-00075-2

© Springer Nature Switzerland AG 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

The Evo Devo Universe Community: Why This Volume Exists

The Evo Devo Universe (EDU) research and discussion community (http://EvoDevoUniverse.com) is a small international group of scholars with broad interdisciplinary interests in multiscale complexity. The EDU community explores Biologically-Inspired Complexity Science and Philosophy (BICS&P) as applied to the universe and its many complex subsystems. What unites our community is the recognition that all complex systems are in a continual tension between stochastic, creative, and unpredictable processes, which we may call *evolutionary*, and convergent, constraining, and predictable processes, which we can call *developmental*. We also recognize that all of our most complex systems engage in replication and inheritance and are subject to many forms of selection.

Within the biological sciences, one community that exemplifies this perspective is the field of evolutionary developmental (evo-devo) biology. Evo-devo biology seeks to understand how developmental processes evolved and how they both constrain and enable the generation of evolutionary novelty under replication and selection. As both evolutionary novelty and developmental constraint are found in all complex replicators, it is an open question, of potential interest to all scholars, whether an "evo-devo" perspective will prove particularly fundamental to understanding multiscale complexity and, perhaps, to the universe itself.

If we live in an evo-devo universe, one whose complexity has self-organized recursively, in a manner at least roughly analogous to organic systems, our universe may embody at least two fundamental *telos* (drives, purposes, goals): to evolve (create, vary, diverge, experiment) and to develop (conserve, converge, generate hierarchy, undergo a predictable life cycle). We can further imagine that both of these drives are subject to yet unclarified forms of selection and adaptation. If our universe indeed has this degree of functional similarity to living systems and replicates in some multiversal environment, we can predict that development at all system scales (organismic, ecological, biogeographic, cultural, technological, universal, etc.) will act as a constraint on the creation of evolutionary novelty at all system scales, just as it does in biology. Likewise, we can expect that evolutionary processes, via preferential replicative selection, will continually and slowly change future development, again, at all scales. By analogy with the evolutionary development of two genetically identical twins, many of our cosmology models now predict that two parametrically identical universes would each exhibit unpredictably separate and unique "evolutionary" variation over their lifespan and, at the same time, a broad set of predictable "developmental" milestones and shared structure and function between them.

We may ask: In what other ways does our universe appear to be an evolutionary developmental system? What other forms of evolutionary novelty and developmental constraint are missing from our current models? How do cyclic unpredictability and predictability interact in complex systems within our universe, from stars to chemistry to life, and what generic selection functions apply? What models suggest our universe itself may replicate and be selected upon in some extrauniversal environment?

The majority dynamical view in science, both in the modern evolutionary synthesis and in modern complexity science, is one of ever-increasing unpredictability, contingency, and diversity in our extant complex systems. But developmental thinking, and emerging evidence of convergent evolution, argues the opposite, that our universe is simultaneously winnowing toward a few implicit past and future forms and functions, at all scales, constrained to do so by initial conditions, universal laws, functional optima, and statistical determinism. Always looking for *both* creative and divergent, and constraining and convergent, physical and informational processes, and asking how they interact and simultaneously contribute to adaptation, is an important shift in perspective. We hope you find it valuable for your own work.

To summarize, we think taking an evo-devo approach can help us address a few of the "missing links" in complexity research today. If this view proves valid, there is much we will continue to learn from the biological and complexity sciences, and their philosophical bases, that will inform our models of complex replicators at all scales. A better understanding of organic development, for example, may inform our understanding of galactic development. A better understanding of convergent evolution in biology may inform our understanding of it in society and technology. A better informational and functional understanding of the origin of life may help us to understand the origin of other complex replicators, including our universe itself. Many other insights, unclear to us today, may be gained. We hope this volume conveys some of the great excitement and opportunity ahead.

A Brief History of EDU and How to Connect with Us

Our EDU community was co-founded in 2008 by systems theorist John M. Smart and philosopher Clément Vidal. We were soon joined by physicist Georgi Yordanov Georgiev and are now also led by biologists Claudio L. Flores Martinez and Michael E. Price. Four of these five directors are editors of this volume. The fifth, EDU co-founder Clément Vidal, bowed out from editorial tasks as a proud new father.

Our founding conference, The Evolution and Development of the Universe, occurred in 2008 at the École Normale Supérieure in Paris. It was attended by 35 scholars in physics, cosmology, chemistry, biology, neuroscience, philosophy, information, computation, social, and complexity science and was an early proof for us of the value of this perspective. In 2009, Clément Vidal selected via peer-review 12 papers for publication, most from our 2008 conference. This founding volume, also titled *The Evolution and Development of the Universe*, introduced a new way of thinking about cosmology and complexity. Vidal also organized open peer commentary for these papers. This volume and its commentaries are on arXiv at https://bit.ly/2snXyZ7.

Our current volume, with 19 papers, is an expansion in the scope and breadth of our last publication. Since 2008, our community has slowly but steadily grown and now comprises over 100 publishing scholars, representing more than 30 academic specializations. Our community is anchored by an active and insightful listserve, EDU-Talk, where we share, learn from, and respectfully critique each other's and other's published work. Our EDU community is still small relative to other interdisciplinary research communities, but it is of high professional quality, growing steadily, and has a bright future.

We welcome new scholars to join our listserve at http://evodevouniverse.com/wiki/EDU-Talk.

We are also happy to list relevant work at http://evodevouniverse.com/wiki/Bibliography. Finally, we hope to see you at our next conference, a satellite meeting at a leading complexity science conference like CCS. In the meantime, connect with us, and share your unique insights.

Together we can do much more, and see far further, than we can alone.

Your *EDC* Editors,
Georgi Yordanov Georgiev,
John M. Smart,
Claudio L. Flores Martinez,
Michael E. Price

Acknowledgments

Warm thanks to Carlos Gershenson at CCS for encouraging us to propose our satellite meeting and to Chris Coughlin at Springer for helping us produce this beautiful volume. Big thanks to all the scholars who answered our call for papers. Special thanks to the 14 presenters who came together in Cancún, to the hundreds or so scholars who attended our CCS satellite, and to the 37 authors who have here shared their important work.

To everyone reading this, may your personal and professional futures be as bright and laden with opportunity as a new and viable embryo, emerging into an astonishingly complex and nurturing environment.

Introduction: About This Volume

What you are reading is a product of a one-day satellite meeting, Evolution, Development, and Complexity, that took place on Sept 20, 2017, at the Conference on Complex Systems (CCS), Cancún, Mexico, Sept 17–22, 2017. CCS is organized by the Complex Systems Society, founded by Paul Bourgine of CNRS, Jeff Johnson of the Open University, Mina Teicher of Bar-Ilan University, and others in Torino, Italy, in 2004. Since its inception, CCS has grown from its European roots to become an international conference, and perhaps the most important annual meeting for the global complex systems research community.

The editors and contributors come from a variety of academic backgrounds, including astrophysics, physics, biology, psychology, complexity science, social sciences, robotics, philosophy, and systems theory. We all share a deep common interest in multiscale complexity and mutual respect for the value of interdisciplinary research.

Our EDC satellite was organized by the Evo Devo Universe (EDU) research community, founded in 2008. Our satellite's goal was to find and present work intended to advance conceptual and mathematical models, evidence, and empiricism in three potentially foundational research themes:

A. Evolution/diversification/phase space creation/unpredictability
B. Development/convergence/phase space reduction/predictability
C. Complexity/nonlinear dynamics/learning/intelligence/adaptation

We suggest that these three themes are particularly fundamental lenses for understanding complex systems. Fortunately for us, a growing number of scholars also seem to share this perspective. We had room for 14 presenters in Cancún, 10 of whom are included in this volume, and this volume includes 13 additional authors or coauthors who were unable to attend our satellite. Several others also answered our call for papers, and we apologize to those who were not selected. Those who joined us braved the formidable September Yucatán sun and enjoyed some unforgettable late-night dinners, conversations, and swims in the Caribbean Sea.

Scholars in this book ask such questions as:

- What can the interaction of divergent evolutionary stochasticity and convergent developmental determinism in biological systems teach us about these processes in other complex systems?
- How can bio-inspired complexity science and philosophy inform astrophysics? Geophysics? Astrobiology? Biological sciences? Cognitive sciences? Social sciences? Engineering and machine learning?
- Why are the laws, constants, and boundary conditions of the universe the way they are? Are they actually or only apparently fine-tuned? What apparent fine-tuning exists for the production of universal complexity? Life? Intelligence? Black holes (e.g., cosmological natural selection)?
- How may the apparent fine-tuning of universal parameters in cosmological development be analogous to the fine-tuning of genetic parameters in biological development? How can we assess fine-tuning relative to adaptiveness in different classes of replicator?
- What are our best network and collective models for adaptation? For learning and intelligence? For cooperation and competition? If our universe is a replicator, as in theories of cosmological natural selection, is there a fitness function (adaptive value) for evolved complexity or intelligence?
- How do cumulative information and network topologies and properties (relation/association) act to constrain physical dynamics in complex environments (modularity, hierarchy, top-down causation)?
- How do scaling laws (scale invariance, scale relativity) constrain evolutionary dynamics?
- Why do we see temporal deceleration of complexification in the early universe (e.g., by Chaisson’s energy rate density measure) and acceleration of complexity in the later universe? How does information influence acceleration?
- What are our best models of convergent evolution? When can we consider evolutionary convergence (e.g., eyes, intelligence) an example of universal development (an astrobiological universal)?
- Are any of Earth’s plausible convergences (e.g., organic chemistry, RNA, fermentation, photosynthesis, multicellularity, nervous systems) likely to be developmental bottlenecks (the only easy path to significantly faster and greater learning and adaptive capacity) at any given level of environmental complexity?
- Can we propose a universal ethics in relation to complex systems? Might such an ethics be more adaptive than our current anthropocentric approaches?
- Why hasn’t our current scientific world view, which posits a universe where life is a “lucky accident,” been updated in a Bayesian manner by new astrophysical discoveries since 1995, including the apparent ubiquity of Earth-like exoplanets? When will we shift from an “accidental” to a “biofelicitous” universe view?
- What is the proper relation between free energy and information in complex systems? How do better thermodynamic descriptions inform scaling limits in describing complex systems?

- How does free energy explain perception, action, and learning in biology and neuroscience? What is the fundamental relation between free energy and the evolutionary development of form and function in living systems? How does the free-energy functional illuminate biotic self-organization and autopoiesis?
- How do action and power efficiency (the least action principle) constrain cellular objectives and dynamics? What is the relation between least action and self-organization?
- In what ways are the interactions between evolutionary algorithms analogous to the interactions between evolutionary agents in swarms and in society? How can complex networks help us visualize and analyze these interactions?
- What is the fundamental relation between the second law of thermodynamics and the emergence of local order? How may we apply models of Darwinian selection not only to biology but to the cosmological, quantum, neural, and human cultural domains? In what sense do natural processes of selection create knowledge, and how may we model knowledge in Bayesian terms?
- What classes of complexity are applicable to morphogenetic systems (heterogeneous software-based agents that form nontrivial shapes and engage in dynamical behaviors)? How do simulations with these systems inform our general models of self-organization and evolution?
- What are the genomics of convergent evolution? Can they help us address general questions of convergence in complex systems?
- What is the hard problem of abiogenesis (the origin of life)? Why is functional genomics central to this problem, and what may this tell us about function in other complex systems?
- How do we account for the “progressive” nature of evolution in complex living systems over time, from the origin of life to “superorganisms” like humankind? What roles do synergy and synergistic selection play this process, and how does synergy offer us a bioeconomic theory of complexity?
- In what ways are immune systems effective and adaptive at their protective function, and why are they apparently scale invariant, working as well and fast in both large and small body sizes? How can a better understanding of immune systems improve the design of disease surveillance systems in society?
- In what ways are special operations forces increasingly used in war fighting and peacekeeping, like a global immune system? When should we use them versus our traditional, hierarchical, conventional, and military forces? How can we make our special operations forces more effective, efficient, and adaptive?
- Why does one small group perform much better than another, even if both groups seem to consist of members with equivalent skills and potential? How do internal structure and energy flows in groups, and their internal and external forces, explain differences in resilience and effectiveness?
- In what ways are religious belief and practice like complex adaptive systems? How do their individual and group-level effects generate communal order and social norms? How do we better understand and promote religious adaptiveness, for individuals and communities, in changing environments?

Contributions

We have organized this volume into four major sections and systems domains: Universe, Multiscale Complexity, Biology, and Society. Each domain is an important frontier for complexity science. Each has powerful insights to offer the others, and we hope you will find concepts, models, and applications in each section that will inform your own work. A brief discussion of each systems domain and its papers now follows.

I. Universe

Considering our universe as a replicative, intelligent system under selection, engaged in both evolutionary novelty production and a developmental hierarchy and life cycle, is the most speculative of potential evo-devo contributions to cosmology and complexity science. Yet, if it proves valid, it will also have the widest implications, as our universe is the largest system we know that may be subject to a single common destiny and set of selection pressures. The papers in this section offer a broad introduction to this idea and explore challenging questions of ethics and predictability of life transitions from a universal perspective.

We begin this section with "*Cosmological Natural Selection and the Function of Life*" by evolutionary psychologist Michael E. Price. Price presents a theory of cosmological natural selection with intelligence (CNSI), proposing that our universe evolved via a process of natural selection acting on a multiverse, with intelligent life emerging as a developmental subroutine of this process. This CNSI theory is essentially an adaptationist revision of Smolin's theory of cosmological natural selection (CNS). The key difference is that whereas CNS identifies black holes as the aspect of our universe's "phenotype" that is most likely to be an adaptation for universe reproduction, CNSI identifies life as this aspect. Price regards the CNSI conclusion as the more probable of the two, because it seems more consistent with a fundamental Darwinian rule: the more a phenotypic trait exhibits improbable complexity, the more likely it is to be an adaptation (as opposed to a nonadaptive trait). Because life appears to be the most improbable complex aspect of the universe, it therefore seems more likely than black holes (or anything else) to represent an adaptation for universe reproduction.

In "*Evolutionary Development: A Universal Perspective*" systems, theorist John M. Smart offers an autopoietic definition of the phrase evolutionary development (evo-devo) in which evolutionary process is broadly construed as any dynamics, in the system or its environment, that generates novelty, diversity, and contingency and developmental process is any dynamics that reduces options and diversity to converge on previously conserved form and function. He sketches a model of natural selection in autopoietic systems and relates this model to the genetics of evo-devo biology. He suggests evo-devo models may be necessary prerequisites for dynamical understanding of any autopoietic (self-reproducing) complex systems, including the universe as a system. John also explores the CNSI framework for universal autopoiesis and argues that intelligence is a particularly adaptive type of accumulated complexity, able to nonrandomly influence replication in a variety of selective environments. His paper explores some expected features in evo-devo models of cosmological fine-tuning, intelligence production, and convergent evolution, considers observation selection effects from an evo-devo perspective, and outlines a few theoretical and simulation advances that may be necessary to validate evo-devo models.

In "*Life, Intelligence, and the Selection of Universes*," astrophysicist and philosopher Rüdiger Vaas explores a variety of possible answers to the following question: Why are the laws, constants, and boundary conditions of nature the way they are? Vaas's subtitle does not refer to the theologically motivated theory of intelligent design (the universe was created by an intelligent entity), which he considers (as do all other authors in this volume) an outlandish claim, but rather to the possibility of cosmological artificial selection (another name for CNSI), a scheme of universal replication in which intelligence and its simulation capacity plays some useful selective role. Vaas critically explores some of the premises and implications of CAS/CNSI employing not only an evolutionary framework, as Price did, but also a cosmological one, considering it in relation to fundamental cosmological issues, such as the fine-tuning of fundamental constants, and to future cosmic doom scenarios.

In "*Universal Ethics: Organized Complexity as an Intrinsic Value*," mathematician Jean-Paul Delahaye and philosopher Clément Vidal stay with a universal perspective, applying it to the human social domain of ethics. While historical approaches to ethics have most often been anthropocentric in both assumptions and goals (e.g., trying to maximize human happiness), Delahaye and Vidal argue for another foundation: that we should preserve, augment, and recursively promote the creation of organized complexity. They use a definition of complexity based on the mathematical and computational concept of logical depth, leading them to three imperatives: the preservation, augmentation, and recursive promotion of organized complexity. They outline conditions under which destruction of complexity (e.g., redundant copies) may still favor these imperatives. Finally, they apply their proposal to a few traditional ethical issues and anticipate some ways this proposal may be misunderstood.

Our last paper in this section is "*Inductive Probabilities in Astrobiology and SETI: Have Sceptics Retreated?*" by astrophysicist Milan M. Ćirković. Ćirković challenges a long-standing and dogmatic assumption in the scientific community. The assumption is that the inductively derived probabilities for critical evolutionary transitions, such as the transition from chemistry to life, from prokaryotes to metazoans, and from simple metazoans to animals with intelligent culture, are astronomically small and thus best characterized as "lucky accidents" rather than processes of evolutionary development. He explores theories and astrophysical discoveries, since the detection of our first exoplanets in 1995, that should have already begun to raise these inductive probabilities in a Bayesian manner. He argues that such a shift has not yet occurred and suggests this lack of change is due to an implicit bias in current scientific culture toward accidental universe models. He also briefly outlines a research program for such a shift in coming years.

II. Multiscale Complexity

Complex systems exist at multiple scales. They employ both physical-energetic and informational-computational processes to create and manage their structure and function and adapt to their environments. They build order, knowledge, and intelligence, maintain boundaries, form networks, and exhibit evolutionary transitions to new regimes, under multiple forms of selection. The papers in this section propose foundational perspectives on such systems, potentially applicable at any scale. They address such universal processes as free-energy use, information production, self-organization, action and power efficiency, evolutionary, swarm and network dynamics, and the evolutionary origins of order.

We begin this section with "*The Equivalence of Free Energy and Information: Thermodynamic Descriptions as a Condition of Possibility of Objectivity*" by biochemist and physicist Joseph M. Brisendine. Brisendine reviews select advances in physics over the last few decades, and the modern concept of information in statistical physics, to illuminate the long-standing philosophical debate on "emergence vs. reductionism" and the ways we can use thermodynamic descriptions to understand the natural scales of emergence for complex systems and their dynamical and informational boundaries. Using examples from the evolution of biological macromolecules and the ordered phases of fluids, he illustrates the proper meaning of entropy in classical thermodynamics, and its relationship to Shannon information theory, and how these imply scaling limits in describing complex systems.

In "*A Multi-scale View of the Emergent Complexity of Life: A Free-Energy Proposal*," astrophysicist and neuroscientist Casper Hesp and coauthors Ramstead, Constant, Badcock, Kirchhoff, and Friston describe how free-energy minimization, and a process they call active inference, helps explain self-organization in living systems. Active inference has recently seen increased adoption in the field of deep machine learning. They introduce Markov blankets, a method for shielding sets of variables in machine learning to create conceptual boundaries and hierarchies, and propose a relation to free-energy minimization and active inference. They then generalize active inference across a number of spatiotemporal scales to propose a general framework for biological self-organization. They first examine active inference from subcellular to multicellular scales in a sketch of how active inference (i) emerges naturally from a primordial soup, (ii) channels dendritic self-organization of single neurons, and (iii) enables the collective organization of many cells into entire organs. They next turn to the organismic level, where we consider (i) the hierarchical brain, (ii) communication and dialogue through active inference, and (iii) cultural affordances and collective active inference. Finally, they consider the species scale, discussing how biological evolution can be viewed as a form of active inference over the order parameters of the lower levels. They end with a focus on niche construction and examine its role in both development and evolution to describe how species build their own eco-niche. In active inference, our brains, bodies, and niches help us to get better *either* at predicting in general, irrespective of our desires (a process we might call development), *or* at actualizing the particular and unique predictive visions we imagine (a process we might call evolution). How these two potentially conflicting neural goals are mediated is one of many questions for future research in this exciting and promising new field.

In "*Action and Power Efficiency in Self-Organization: The Case for Growth Efficiency as a Cellular Objective in* Escherichia coli," physicist Georgi Yordanov Georgiev and coauthors Aho, Kesseli, Yli-Harja, and Kauffman explore a potentially fundamental feature of self-organization in complex systems. They propose action and power efficiency as an objective in the process of self-organization and explore this model in a sample system of *E. coli*. The authors derive an expression for growth efficiency as a special case of action (power) efficiency and justify it through a first principle analysis in physics, using the principle of least action. Exploring when growth efficiency may be particularly adaptive, as in energy-restrictive environments, and when other cellular objectives may apply, is a potential area of future research. Action and power efficiency analysis may apply to multicellular systems and ecosystems as well.

In "*On Relation Between Swarm and Evolutionary Dynamics and Complex Networks*," computer scientists Ivan Zelinka and Roman Šenkeřík propose a novel, network visualization method for visualizing the dynamics of evolutionary algorithms. The paper integrates models in the fields of complex networks and evolutionary computation and proposes that interactions between evolutionary algorithms are in several ways analogous to interactions between evolutionary agents in swarm colonies and in society. The dynamics of differential evolution, particle swarm optimization, and firefly evolutionary algorithms are explored in this network visualization method. The analogy between network vertices and individual actors in a population and between network edges and communication among individuals, under any algorithm, is also discussed.

Our last paper in this section is "*Universal Darwinism and the Origins of Order*," by evolutionary scholar John O. Campbell and evolutionary psychologist Michael E. Price. Campbell and Price portray our universe as the outcome of two fundamental processes: one of ever-increasing disorder (due to the second law of thermodynamics) and one of ever-increasing local order (due to Darwinian selection not just in the biological domain but also in the cosmological, quantum, neural, and cultural domains). Campbell and Price flesh out how the substrate-neutral "Darwinian algorithm" (variation-inheritance-selection) can act so effectively as a universal source of order, despite the relentless destructive influence of the second law. Knowledge is key: each natural domain is characterized by a distinct "knowledge repository," that is, a cumulative store of information about existence requirements in that domain (e.g., a genome, wave function, or neural model). A knowledge repository is a probabilistic model which makes guesses about how to exist, which is tested for accuracy by an "embodied adapted system" (e.g., phenotype) that the repository encodes. The repository undergoes a Bayesian update based on test results, which results in it becoming less ignorant and less entropic.

III. Biology

Biological systems are the canonical example of complex systems. Though we know of only one example of their origin and history so far, life on Earth, that astonishing example spans roughly four billion years, nearly a third of the age of the universe itself. The better we understand life's evolution, development, intelligence, and selection, the better we may understand other complex adaptive systems. Papers in this section explore a range of life-critical phenomena including self-organization, convergent evolution, the origin of life, and the kinds of selection (competitive and synergistic) we find in intelligent systems.

We begin this section with "*Complexity, Development, and Evolution in Morphogenetic Collective Systems*" by complexity scientist and systems scientist Hiroki Sayama. Sayama reviews his recent empirical and theoretical work on morphogenetic collective systems, which are systems made of a large quantity of heterogeneous agents that interact with each other to form nontrivial shapes and dynamical behaviors. Sayama proposes four complexity classes applicable to morphogenetic collective systems based on properties of the individual agents. Using numerical simulations in his swarm chemistry model, he demonstrates that the heterogeneity of agents plays a crucial role in collective systems' self-organization and that dynamic redifferentiation of agents helps produce more diverse macroscopic behaviors. Sayama also studies evolutionary design of these morphogenetic collective systems, using both interactive and autonomous methods. The robustness and sensitivity of self-organization and evolution against two- and three-dimensional changes are also discussed.

In "*Comparative Genomics of Convergent Evolution*," molecular biologist Claudio L. Flores Martinez explores the phenomenon of convergent evolution in biological systems via the discipline of comparative genomics. As genome data availability has rapidly grown in recent years, studies of convergence in complex adaptive traits are now being published on newly sequenced genomes from all organizational levels of bio-complexity. Convergent evolution is even occurring at the molecular sequence level. Information and biological systems theory are offering us new models of convergence based on definitions of functional equivalence classes in adaptive traits, driven by top-down causation. The information-driven study of convergence is providing a new computational paradigm within systems biology for exploring sequence space and its relation to genomic functional equivalence classes across multiple levels and scales of bio-complexity. These information-based models may have even wider future applicability in science, as convergence appears to be a fundamental property of complex adaptive systems, found not only in biology but also in our social organization and technology.

In "*Why Functional Genomics Is the Central Concern of Biology and the Hard Problem of Abiogenesis*," developmental biologist James A. Coffman proposes that the defining characteristic of life is the nature of its genomic information. He argues, after Robert Rosen, that it is life's "self-entailing" system of metabolism, repair, and replication that is its unique feature, a system that also allows it to model and encode relations with its environment. In this view, genomic (and epigenomic) information is life's central concern, and it cannot be reduced to physics and chemistry, as it is physically indeterminate, contingent, and context dependent. The hard problem of the origin of life is the sequence of contingencies that produced the first genome. Paleontology has been an essential science for understanding aspects of life's evolutionary development to date, but as early genomes are not preserved in the fossil record, the hard problem of abiogenesis will likely remain unsolved until we learn how to create a viable, self-entailing organism in our empirical experiments.

Our last paper in this section is "*Synergistic Selection: A Bioeconomic Approach to Complexity in Evolution*" by biologist and complexity scholar Peter A. Corning. To account for the "progressive" evolution of complex living systems – from the origins of life itself to "superorganisms" like humankind – Corning proposes the synergism hypothesis. This economic (or bioeconomic) theory of complexity regards living systems as subject to basic economic criteria. Focusing on the costs and benefits of complexity and the unique creative power of functional synergy in the natural world, the theory proposes that the overall trajectory of evolution over the past 3.8 billion years or so has been shaped by synergies of various kinds. The synergies produced by cooperation among various elements, genes, parts, or individuals may create interdependent "units" of adaptation and evolutionary change that are favored in a dynamic that Maynard Smith termed synergistic selection. In this ambitious synthetic view, synergistic selection is an often overlooked subcategory of natural selection.

IV. Society

As biological collectives also employ extra-biological processes, societal systems are the most structurally, functionally, and computationally complex systems, per mass or volume, in our known universe. While our models of societal evolution, development, intelligence, and selection are preliminary, insights here are particularly valued, as we can potentially apply them to many formidable human problems. Our papers explore bioinspired and general applications of complexity research to societal phenomena, including an immunological approach to disease surveillance system construction, special operations forces as an element of an emerging societal immune system, cost and benefit analysis in species-ecosystem interactions, predicting performance in small groups, and religious belief networks as complex adaptive systems.

We begin this section with "*Movement and Spatial Specificity Support Scaling in Ant Colonies and Immune Systems: Application to National Biosurveillance*" by biologist and computer scientist Tatiana Flanagan and coauthors Beyeler, Levin, Finley, and Moses. Flanagan et al. propose a distributed, adaptive, and scalable framework for the use of public health biosurveillance information to minimize detection time and maximize accuracy in disease outbreaks. They consider the adaptive immune system and ant colonies, two living systems that efficiently sample, detect, and respond to diverse data sets in a wide range of environments. The apparent scale invariant nature of immune system response across body size, in which larger body size does not significantly slow adaptive response, is a curious and potentially critical feature of immune systems and one not found in many of our current large social organizations tasked with intelligence gathering, security, or defense. They identify measurable features, including decentralized search authority, automated distributed response, and the ways immune system agents communicate, move in space, construct hubs, and store memory, that may guide designers and managers to improve our existing disease surveillance systems and, perhaps, help other large organizations faced with intelligence, security, and defense challenges.

In "*Special Operations Forces as a Global Immune System*" by complexity scholar Joseph Norman and complexity scientist Yaneer Bar-Yam, the authors also explore the immune system metaphor in the context of special operations forces. Using an extension of Ashby's law of requisite variety, they identify a tradeoff between scale and complexity in control and between the use of hierarchical control and conventional military forces in response to large-scale challenges and the use of decentralized control and special operations forces in high-complexity, fine-scale challenges. Using the analogy of human civilization as a multicellular organism and our defense forces as immune systems, they consider special forces as an adaptive first response in preventing cascading disruption in social systems when they can rapidly respond to local crises while they remain small. They identify three ways, all found in healthy immune systems, that these forces might be particularly effective, including unmediated interaction with local cultures and peoples, persistent presence and embeddedness, and local autonomy and decision-making. We have seen dramatic increase in global use of special forces in war-fighting and peacekeeping efforts in recent decades. Understanding them from an immune systems lens may tell us how to make that use more effective, efficient, and adaptive.

In "*Applying Evolutionary Meta-Strategies to Human Problems*," complex systems analyst Valerie Gremillion proposes that complexity science can help us find general classes of evolutionary strategy and meta-strategies that we may apply to our increasingly intricate, entwined, and encompassing human problems. Gremillion uses a neural-derived network approach to modeling species-ecosystem interactions, with discrete balance equations and a methodology for analyzing higher-order network costs, benefits, and unintended consequences. General classes of ecocircuitry networks may be identified using this approach, as well as evolutionary meta-strategies used in these networks. Direct meta-strategies include alterations to the number and magnitude of beneficial and detrimental flows of energy, materials, services, and information and efforts to decrease the variance of all flows. Powerful indirect meta-strategies, which include externalizing costs, tool and infrastructure building, and novel strategy combinations, are also identified. Gremillion explores how human meta-strategies have led to unintended consequences, both beneficial and problematic, and considers how a more complex and deliberate ecocircuitry approach may help us find better solutions in an increasingly complex and rapidly changing world.

In "*Complex Dynamics in Small Groups*," evolutionary psychologists Holly Arrow and Kelly Bouas Henry use a dynamic systems perspective to tackle a fundamental puzzle in group dynamics: Why does one group often perform much better than another even if both groups seem to consist of members who possess equivalent skills and potential? Their analysis draws on the metaphor of group "thermodynamics" to illuminate how structure and energy in groups reciprocally influence each other as groups change over time. This framework attends to the interplay between the internal and external forces acting on a group and suggests that differences in developmental trajectories help explain why some groups become more resilient and effective than others. To illustrate their examination of the interplay between endogenous development and externally responsive adaptation, they use examples of groups from three contexts: the classroom, team sports, and military aviation. These case studies allow them to unpack their ideas about structure, dynamics, and development, using the concepts of heat and energy transfer as underlying drivers of performance.

Our last paper in this section is "*The Building Blocks of Religious Systems: Approaching Religion as a Complex Adaptive System*" by anthropologist Richard Sosis. Sosis proposes that religions consist of cognitive, neurological, affective, behavioral, and developmental traits that form a complex adaptive system. These systems function to foster cooperation and social coordination within communities, and they maintain eight primary interdependent building blocks: authority, meaning, moral obligation, myth, ritual, sacred, supernatural agents, and taboo. Religious systems are fuelled by ritual behaviors, which interact with other core building blocks to produce individual-level effects, including physiological and affective responses. Individual-level effects in turn yield group-level effects, such as shared cognitive schema, ethos, symbolic meaning, and identity. These group-level effects result in a sense of communal order and produce social norms, including – most importantly, from a group-level functional perspective – norms that encourage cooperative and coordinated behaviors among community members.

Contents

Part I Universe

Part II Multiscale Complexity

Part III Biology

Part IV Society

Editors and Contributors

About the Editors

Claudio L. Flores Martinez is a Director of the Evo Devo Universe research community and is a Ph.D. student at the University of Hamburg with a background in molecular biosciences and developmental biology (B.Sc. and M.Sc. from the University of Heidelberg). His expertise is covering areas in theoretical biology with a focus on major transitions across organizational levels of bio-complexity, the origin of life, genome evolution, convergence, top-down causation, philosophy of science, cognition, neuroscience, complex adaptive systems, technological innovation, and biomimetics. He is running his own innovation consultancy for biomimetic systems design and knowledge engineering.

Georgi Yordanov Georgiev is a Director of the Evo Devo Universe research community and is a Full Professor of Physics at Assumption College and Worcester Polytechnic Institute, Worcester, MA. His interests are in applying first principles from physics to explain the continuous self-organization and evolution in complex systems very far away from thermodynamic equilibrium. His approach is to use variational principles as drivers for progressive development of structure in those systems. He has proposed to use the efficiency of physical action as a measure for the degree of organization.

Michael E. Price is a Director of the Evo Devo Universe research community and is Senior Lecturer in Psychology at Brunel University London. He has a BA from Duke University and a Ph.D. from the UC Santa Barbara Center for Evolutionary Psychology. His research interests focused initially on biological evolutionary psychology (especially morality and human nature) but have expanded more recently to include universal Darwinism (the creative potential of Darwinian selection, across natural domains) and the evolution of religiosity.

John M. Smart is a Director and Co-founder of the Evo Devo Universe research community, is a Lecturer in Foresight at the Naval Postgraduate School (Monterey, CA), and President of the Acceleration Studies Foundation, a nonprofit promoting scientific understanding and technical management of accelerating change. He is Author of *Evo Devo Universe?* (2009), The transcension hypothesis (2012) and *The Foresight Guide* (2016).

Contributors

Holly Arrow is Professor of Psychology at the University of Oregon and former President of the Society for Chaos Theory in Psychology and Life Sciences (SCT-PLS). She studies the formation and development as groups as complex dynamic systems. Her book *Small Groups as Complex Systems: Formation, Coordination, Development, and Adaptation* (2000, with J. E. McGrath and J. L. Berdahl) has been widely cited. More recent publications include *Using Complexity to Promote Group Learning in Health Care* (2010, with K.B. Henry).

Yaneer Bar-Yam is a Professor and Founding President of the New England Complex Systems Institute. He has been a Founding Contributor to the field of complexity science since the 1980s, bringing mathematical rigor, real-world application, and educational programs to this field. He is Author of *Dynamics of Complex Systems* (1997) and *Making Things Work: Solving Complex Problems in a Complex World* (2004) and has published papers on a wide range of scientific and applied problems, ranging from cell biology to the global financial crisis.

Joseph M. Brisendine is a Researcher and Adjunct Professor in the Departments of Chemistry and Physics at the City College of New York. He is the Author of several research articles on de novo protein design and is currently coauthoring a book on the physics of living membranes in the Biochemistry Department of The Rockefeller University. His research uses the thermodynamics of computation to understand the evolution and function of biological molecular machines.

John O. Campbell is an independent Researcher based in Victoria, British Columbia, Canada. His research interests include universal Darwinism, Bayesian inference, and information theory. He is the Author of *Universal Darwinism: The Path of Knowledge*, *Darwin Does Physics*, and *Einstein's Enlightenment.*

Milan M. Ćirković is Professor of Physics at the Astronomical Observatory of Belgrade (Belgrade, Serbia) and Research Associate, Future of Humanity Institute, Oxford University (Oxford, UK). His research interests include astrobiology and SETI, the evolution of galaxies, baryonic dark matter, and the philosophy of science and cosmology. He is Author of *The Astrobiological Landscape: Philosophical Foundations of the Study of Cosmic Life* (2012) and *The Great Silence: Science and Philosophy of Fermi's Paradox* (2018) and has more than 150 journal publications.

James A. Coffman is Associate Professor at the MDI Biological Laboratory and is a Developmental Biologist and Experimental Embryologist interested in gene regulation and developmental plasticity, whose laboratory is currently focused on the question of how chronic stress during early development influences adult health and disease risk. He has a long-standing interest in the problem of biological change through development and evolution, particularly from the perspective of complex systems theory, and has published numerous papers and one book on this subject.

Peter A. Corning is the Director of the Institute for the Study of Complex Systems, a one-time Science Writer for *Newsweek* and Professor in Human Biology at Stanford University, and the Author of seven books and more than 200 professional and print media articles. For more information, see the ISCS website: www.complexsystems.org. Also, see the Amazon link and reviews of his newest book, *Synergistic Selection: How Cooperation Has Shaped Evolution and the Rise of Humankind* (2018): https://tinyurl.com/ybpxbu3j.

Jean-Paul Delahaye is Professor Emeritus of Computer Science at the University of Lille and Member of CRISTAL (Centre de Recherche en Informatique Signal et Automatique de Lille) du CNRS (Centre National de la Recherche Scientifique). His research includes Kolmogorov complexity and cryptocurrencies. He writes a blog, Complexites, at http://www.scilogs.fr/complexites/. His latest book is *Les Mathématiciens se Plient au Jeu*.

Tatiana Flanagan is a Postdoctoral Fellow at the University of New Mexico and a Technical Staff Member at Sandia National Laboratories. Her research interests include complex systems and ant colony behavior. She is the Author of "How Ants Turn Information into Food" (2015) http://digitalrepository.unm.edu/biol_etds/37/ and "Fast and Flexible: Argentine Ants Recruit from Nearby Trails" (2013) *PLoS ONE* 8(8): e70888.

Karl Friston is a Scientific Director and Principal Research Fellow at the Wellcome Trust Centre for Neuroimaging (UK) and a Professor at the Institute of Neurology, University College London. He is a theoretical neuroscientist, psychiatrist, and a global authority on brain imaging. He invented statistical parametric mapping (SPM), voxel-based morphometry (VBM), and dynamic causal modelling (DCM), new uses of probability theory to analyse neural imaging data. These contributions were motivated by schizophrenia research and theoretical studies of value learning. Mathematical contributions include variational Laplacian procedures and generalized filtering for hierarchical Bayesian model inversion. He currently works on models of functional integration in the human brain and the principles that underlie neuronal interactions. His main contribution to theoretical neurobiology is a free-energy principle for action and perception (active inference). His dynamic causal modelling is used to estimate how different cortical regions of the brain influence one another. His hypothesis that the minimization of surprise can explain many aspects of action and perception informs his ongoing empirical and theoretical work. He received a Golden Brain Award from the Minerva Foundation in 2003 and the Weldon Memorial Prize in 2013.

Valerie Gremillion initial Ph.D. in Neurosciences at UCSD led her to study complexity theory at Los Alamos National Lab and SFI. She expanded to cross-disciplinary private consulting work, applying systems thinking to problems in ecology, business, politics, healthcare, society, and brain optimization. Her latest public work will bridge knowledge of the brain from the academic ivory tower to the public, through neuroscience consulting, writing, and speaking. She is also a Research Assistant Professor at UNM in Biology.

Kelly Bouas Henry is a Professor of Psychology at Missouri Western State University in St. Joseph, Missouri. Her research interests focus on the role of group identification in team performance and cooperation and identification of dynamics associated with high-performing work groups.

Casper Hesp is a Researcher of Theoretical and Computational Neuroscience at the Wellcome Trust Centre for Neuroimaging (Karl Friston's group), the Amsterdam Institute for Brain and Cognition, and the Institute for Advanced Study of the University of Amsterdam. Besides neuroscience, he has a background and has published in computational astrophysics (M.Sc.). His research interests focus on the interface between theory and modelling and the cross-scale dynamics of neurocognitive evolution and development.

Stuart A. Kauffman is an American medical doctor, theoretical biologist, and complex systems researcher who studies the origin of life. He was a Professor at the University of Chicago, University of Pennsylvania, and University of Calgary. He is currently Emeritus Professor of biochemistry at the University of Pennsylvania and affiliate faculty at the Institute for Systems Biology. He has a number of awards including a MacArthur Fellowship and a Wiener Medal. He is best known for arguing that the complexity of biological systems and organisms might result as much from self-organization and far-from-equilibrium dynamics as from Darwinian natural selection as discussed in his book *Origins of Order* (1993). Using random Boolean networks to investigate self-organizing properties of gene regulatory networks, he proposed that cell types are dynamical attractors in such networks and that cell differentiation, dedifferentiation, and cancer can be modeled as transitions between attractors. He also proposed the self-organized emergence of autocatalytic sets of peptide polymers as the origin of adaptive molecular reproduction. These proposals remain active areas of research in complexity science.

Melanie Moses is a Professor in the Department of Computer Science at the University of New Mexico, and Director of the Moses Biological Computation Lab at UNM. She has a secondary appointment in the UNM Department of Biology and is External Faculty at the Santa Fe Institute. She studies complex biological and information systems, and the general rules governing the acquisition of energy and information in complex adaptive systems and networks. Her focus is on the efficiency of growth and information exchange in biological and computational networks, and how the size and topology of networks determine emergent system behavior. The Moses Lab develops computational and mathematical models of biological systems, with an emphasis on modeling the adaptive immune system, bio-inspired computation and swarm robotics, and understanding scaling in biology, immunology, and computation. She is also PI for two programs that work to diversify computer science education, NM CSforAll and the NASA Swarmathon swarm robotics challenge.

Joseph Norman is an Applied Complexity Scientist with a focus on self-organization and multiscale information theory for detecting and managing systemic risk in large-scale systems. His research interests include systemic risk and precaution, biological pattern formation, and multiscale variety in purposeful systems. He is an affiliate of the New England Complex Systems Institute and the Real World Risk Institute.

Hiroki Sayama is a Professor of Systems Science and Industrial Engineering and the Director of the Center for Collective Dynamics of Complex Systems at Binghamton University, State University of New York. His research interests include complex systems, dynamical networks, human and social dynamics, artificial life/chemistry, interactive systems, and other topics. He is the Author of *Introduction to the Modeling and Analysis of Complex Systems*, a free Open SUNY textbook that has been downloaded more than 44,000 times.

Roman Šenkeřík is Associate Professor of Computer Science at the Tomas Bata University in Zlin. His research interests include bioinspired computation and artificial intelligence. He is the Author of numerous journal papers, including *Performance comparison of differential evolution driving analytic programming for regression*.

Richard Sosis is James Barnett Professor of Humanistic Anthropology at the University of Connecticut. His work has focused on the evolution of religion and cooperation, with particular interests in ritual, magic, religious cognition, and the dynamics of religious systems. He is Co-founder and Coeditor of the journal *Religion, Brain & Behavior*, which publishes research on the bio-cultural study of religion.

Rüdiger Vaas is a Philosopher of Science, the Astronomy and Physics Editor of the monthly German science magazine *Bild der Wissenschaft*, Author of several books about cosmology and fundamental physics, Coeditor (together with Laura Mersini-Houghton) of the book *The Arrows of* Time, as well as an Editorial Board Member of Springer's book series The Frontiers Collection and Science and Fiction.

Clément Vidal is a Director and Co-founder of the Evo Devo Universe research community and is Assistant Professor of Philosophy at the Vrije Universiteit Brussel, Belgium. He is the Author of *The Beginning and the End: The Meaning of Life in a Cosmological Perspective* (2014). He is eager to tackle big philosophical questions in cosmology, physics, computer science, astrobiology, complexity science, and evolutionary theory.

Ivan Zelinka is a Professor of Computer Science at the VSB-Technical University of Ostrava. Ivan's research interests include unconventional computation and cybersecurity. She is the Author of the book *Evolutionary Algorithms, Swarm Dynamics and Complex Networks: Methodology, Perspectives and Implementation* and numerous journal articles, including Swarm viruses: next-generation virus and antivirus paradigm? He is a member of the British Computer Society and IEEE.

Part I
Universe

Cosmological Natural Selection and the Function of Life

Michael E. Price

1 Introduction

1.1 Universal Darwinism, Adaptation, and Cosmological Evolution

This chapter is essentially an adaptationist critique of Smolin's (1992, 1997) attempt to apply the Darwinian theory of natural selection at the cosmological level. I do not dispute, however, that Darwinian principles can be usefully applied at this level. On the contrary, I agree that the theoretical tool set provided by Darwinism is the most useful and powerful one we have in accounting for the natural origins of complex order at both the biological and cosmological levels. Indeed, the concept of universal Darwinism implies that Darwinian principles can potentially explain the origins of complexity not only at these levels but at many others as well. Universal Darwinism is founded on the observation that mechanisms of variation, selection, and heredity can operate, and seemingly 'design' complex structures, in a wide variety of domains besides the biological, including the cosmological, quantum, neural, cultural, and epistemological (D. Campbell 1974; J. Campbell 2011, 2015, 2016; Cziko 1997; Dennett 1995; Heylighen 1999; Smolin 1992, 1997; Zurek 2003, 2009). The Darwinian algorithm, in other words, is substrate-neutral (Dennett 1995).

The domain in which Darwinian mechanisms have been studied most intensively is, of course, the biological. In this domain, the complex structures produced by evolution are known as adaptations. The adaptation concept is central to achieving a

---

M. E. Price (✉)
Department of Life Sciences, Centre for Culture and Evolution, Brunel University London, Uxbridge, UK
e-mail: michael.price@brunel.ac.uk

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_1

clear and coherent understanding of how biological organisms are designed. This concept also has the potential to be highly illuminating when applied to non-biological domains. The central argument of this chapter will be that when Smolin uses Darwinian theory to understand the design of the universe, he underutilizes the adaptation concept. In particular, he does not apply the observation that improbable complexity is the hallmark of adaptation, and that therefore the phenotypic aspects most likely to be adaptations (as opposed to non-adaptive traits) are those displaying the highest degree of improbable complexity. As a result, he misidentifies black holes as the characteristic of the universe that is most likely to be an adaptation for universe reproduction. Intelligent life displays more improbable complexity than black holes, and so by adaptationist criteria, it is more likely than black holes to constitute such an adaptation.

Throughout this chapter, I will refer to Smolin's (1992, 1997) theory of cosmological natural selection, which proposes black holes to be the most likely mechanism of universe reproduction, as 'Smolin's CNS'. I will refer to the alternative adaptationist theory developed here and in Price (2017), which proposes intelligent life to be the most likely mechanism of universe reproduction, as 'cosmological natural selection with intelligence', or CNSI.

1.2 Chapter Structure: (1) The Scientific Case for and (2) Philosophical Implications of CNSI

Part 1 of this chapter proposes the scientific case that intelligent life functions ultimately as an adaptation for universe reproduction. The notion may seem outlandish, at least initially, and especially to those unfamiliar with the basic cosmological and Darwinian concepts on which it is based. I aim to establish, however, that regardless of how outlandish this idea may seem, it represents an adaptationist correction to a theory that, in other respects, seems basically reasonable – Smolin's theory of cosmological natural selection. As discussed below, Smolin's CNS is itself a logical extension of speculative but nevertheless quite conventional cosmological concepts, especially the ideas that we live in a multiverse, and that black holes generate new universes by leading to big bang-like events. The novel contribution of Smolin's CNS is the suggestion that this generation of new universes is governed by a Darwinian process. He reasonably suggests, in other words, that if some kinds of universes were especially good at reproducing, then they would become particularly well represented in the multiverse. When Smolin attempts to apply this Darwinian framework to the problem of cosmological evolution, however, he does not follow it through to its logical adaptationist conclusion. From an adaptationist perspective, this seems to have been an oversight on Smolin's part, and I will be aiming to correct it in this chapter.

Part 2 of this chapter considers CNSI's philosophical implications: if CNSI were true, what would the implications be for beliefs about life's fundamental 'purpose' and the nature of morality? For a naturalistic theory, CNSI seems uniquely able to accommodate a notion that is central to all major religious worldviews: that life itself has a ‘transcendent’ purpose, above and beyond the biologically evolved interests of organisms themselves. In other words, CNSI provides a naturalistic framework for thinking about life as if it were ‘designed’, by a ‘higher (or more fundamental) power’, to serve an ultimate ‘purpose’. These terms are all in quotes, because they are often used to describe what are perceived to be the actions of an intelligent designer, and I want to emphasize that CNSI is not a theory of intelligent design. CNSI proposes, rather, that life’s fundamental function (or ‘transcendent purpose’) was designed by a non-intelligent Darwinian process, operating at the cosmological level. CNSI’s Darwinian perspective is the reason it can propose that life has a transcendent purpose. This is true because Darwinian selection is the only known natural process that can endow an entity with function and, therefore, with apparent purpose.

2 Part 1: The Scientific Case for Cosmological Natural Selection with Intelligence

2.1 The Origins of Complexity

Perhaps the most basic law of physics is the second law of thermodynamics, which states that the total entropy (disorder) of an isolated system can only increase over time. In this law, science possesses a fundamental and universally acknowledged explanation for why such disorder tends to always be increasing in the universe. Science lacks, however, an ultimate explanation for complex order that is as universally acknowledged as the second law. Virtually everyone agrees that disorder is due to the second law, but what are the origins of order, including complex order? If complexity did not exist in the first place, nothing would decay into a less-ordered state, and the second law would not apply to begin with.

We could attempt to account for complexity by referring to the four forces of gravitation, electromagnetism, and the weak and strong interactions. But while these forces do impose order, referring to them would be a superficial explanation for complexity, because it would fail to account for the origins of the forces themselves. We seek a more fundamental process, one that could account for the origins of these forces and ultimately for all complexity in the universe, including life. This may sound like an ambitious goal, but a compelling attempt at it has already been made by Smolin (1992, 1997), in the form of cosmological natural selection theory (Smolin’s CNS). This theory utilizes a Darwinian selectionist framework to account for the origins of complexity, in recognition of natural selection’s unique power to generate complexity. To emphasize this power, Smolin (1997) refers to a quote from Dawkins (1986): ‘The theory of evolution by cumulative natural selection is the only theory we know of that is, in principle, capable of explaining the existence of organized complexity’.

2.2 A Summary of Smolin's Cosmological Natural Selection (CNS)

Smolin's CNS (1992, 1997; for a formal CNS model, see Gardner and Conlon 2013) proposes that our universe exists in a population (multiverse) of universes that reproduce slightly mutated versions of themselves, allowing for descent with modification. Universe designs achieve greater representation in this population if they reproduce relatively successfully. Smolin's CNS is thus highly analogous to Darwin's (1859) theory of biological natural selection, and can be expressed in the terminology of contemporary Darwinism. The laws, parameters, and other features exhibited by our universe represent its phenotypic design, and the universe is a vehicle of selection (sensu Dawkins 1986) that evolution designed for the ultimate goal of reproduction. The universe thus requires some mechanism that would enable it to reproduce, and Smolin proposes black holes to be this mechanism. The idea that a black hole could give birth to a baby universe seems intuitively plausible, if we regard a big bang as being similar to what we would expect to find on the 'other side' of a black hole. A black hole singularity is an infinitely dense concentration of space-time, matter, and energy, and a big bang is an explosive expansion of space-time, matter, and energy, which emerges from a singularity.

Smolin was not the first physicist to recognize that a black hole and big bang seem suspiciously like inverses of one another. John Wheeler, a leading figure of twentieth-century-physics, influenced him most directly on this point. Years before Smolin published his CNS theory, Wheeler (1977) had explored the idea that black holes give birth to new universes. Wheeler, however, suggested that a baby universe's laws and parameters would vary randomly from those of its parent, whereas Smolin proposed that the baby would inherit parental traits with a degree of fidelity that is comparable to biological inheritance.

The addition of this inheritance factor allowed Smolin to consider that universes may evolve by a process of natural selection, with the selective process favouring universe designs that, by maximizing black hole production, also maximized their reproductive success. In other words, according to Smolin's CNS, the laws and parameters that characterize our own universe were selected because they enable stars to develop and ultimately collapse into black holes. All order in our universe, therefore, including the four forces themselves, originates from selection for black hole production. And what is the status of life from this perspective? It exists not as a design feature of this evolutionary process, but as a by-product of selection for black hole production. As stars develop towards black holes, they generate elements such as carbon, nitrogen, and oxygen as by-products. Under a strictly narrow set of environmental parameters, these elements may end up interacting in a manner that accidently enables life to evolve. On at least one planet, this accident has actually occurred.

A limitation of Smolin's CNS is that it does not identify the precise mechanism by which a baby universe could inherit parental traits: if the universe's laws and parameters constitute its phenotype, what is the nature of its genotype? Where is the universe's 'DNA' located, and could it remain functionally intact after making the seemingly obliterative journey through a black hole? These are important questions, but our present inability to answer them conclusively should not preclude us from considering Smolin's theory as thoroughly as we presently can. As Smolin (1997, p. 316) notes, our ignorance here is analogous to Darwin's ignorance about the exact mechanism of biological inheritance. The structure of DNA would not be discovered until nearly a century after *On the Origin of Species* (1859) was published, and Darwin constructed a remarkably accurate and robust theory of biological selection while knowing nothing about genes.

2.3 Smolin's CNS Is a Logical Extension of Conventional Cosmology

Smolin's CNS may seem a sharp deviation from conventional cosmology, especially because it attempts to synthesize cosmology with a theoretical framework that is normally associated with biology. This direction may seem especially radical to those who are not already familiar with the cosmological concepts on which Smolin's CNS is based. It is important to emphasize, however, that these concepts are fairly conventional within cosmology and that Smolin's CNS extends them in a straightforward fashion. For example, Smolin's CNS depends on the notion that universes are constantly generating new universes and that our own universe is just one of many in the multiverse that results from this process. The idea that black holes represent the 'other side' of big bang-type events has been under consideration for several decades (Garriga et al. 2016; Wheeler 1977) and melds naturally with the multiverse idea. The existence of this multiverse is itself predicted by nearly all models of eternal inflation, which is an extension of Big Bang theory and currently the most widely accepted theory of universe origins (Guth 1981; Linde 1986, 2017).

The larger point I wish to make is that although Smolin's CNS is theoretically speculative, it is founded on cosmological ideas – especially, that we exist in a multiverse, which is the result of black holes leading to big bang-type events – that cannot be considered unconventional. These ideas converge on the general notion that our multiverse can be conceived as a population of universes that are constantly generating new universes. The contribution of Smolin's CNS was to extend this notion, by pointing out that if these universes exhibited variance in their 'designs' (i.e. laws and parameters) that affected their ability to generate new universes, and if there were any continuity ('inheritance') between universes and their offshoots, then some designs could become better represented than others in this multiverse. In other words, universes could evolve via Darwinian selection.

It is true that Smolin's CNS adds another layer of speculation to cosmological ideas that, despite being widely regarded as plausible within cosmology, were speculative already. On the other hand, the point that Smolin adds is quite simple: if a certain kind of entity is particularly good at reproducing, then it will become particularly prevalent. Given that Smolin is extending conventional cosmology in such a simple way, the theory that results from this extension is surprisingly powerful. It not only suggests that universes evolve according to the same Darwinian algorithm as organisms, but it provides an ultimate explanation for the origins of complexity in our universe. Because Smolin's CNS is so simple, powerful, and consilient with existing scientific ideas, it seems to merit serious consideration. My criticism of Smolin's CNS, therefore, relates not to the theory in general but only to one specific point: it seems to misapply the Darwinian principle of adaptationism.

2.4 The Adaptationist Perspective and Smolin's Underutilization of It

Contrary to Smolin's CNS, Darwinian adaptationist principles lead to the conclusion that intelligent life is more likely than black holes to be an adaptation for universe reproduction. To understand why this is the case, we must consider the adaptationist framework in evolutionary biology.

As noted above, Smolin chooses a Darwinian framework for explaining the origins of cosmological order because he recognizes selection's unique power to generate complexity. In his words: 'There is only one mode of explanation I know of, developed by science, to explain why a system has parameters that lead to much more complexity than typical values of those parameters. This is natural selection' (Smolin 2006). He also notes that due to its unrivalled ability to produce complexity, selection has created the least-entropic entities ever observed: living organisms. 'A living system… continually creates an enormous number of different kinds of molecules, each of which generally perform a unique function. The entropy of a living thing is consequently much lower, atom for atom, than anything else in the world' (Smolin 1997). Smolin's thoughts here closely align with those of Schrödinger (1944), who famously defined life in terms of its extraordinarily low entropy.

Smolin correctly identifies selection[1] as the designer of organized complexity and acknowledges organisms as the least-entropic (and most complexly ordered) known entities. His approach lacks, however, a rigorous adaptationist perspective. This perspective can illuminate both why organisms possess such low entropy and also how we can recognize which aspects of organisms are most likely to be adaptations (as opposed to non-adaptive traits).

The low entropy of organisms is due to the fact that they are bundles of genetically encoded, functionally complex, problem-solving biological devices which themselves have low entropy: adaptations. Adaptations have low entropy

---

[1] I'll use 'selection' as shorthand for all evolutionary selective processes that design adaptations, including natural selection (Darwin 1859), sexual selection (Darwin 1871), and kin selection (Hamilton 1964).

because they are functionally specialized for solving adaptive problems that are often highly nuanced, complex, and difficult. An adaptive problem is defined as any aspect of the environment that challenges the individual organism's ability to survive and reproduce successfully, in competition with local members of its own species. Organisms are generally equipped with numerous adaptations, each functionally specialized for solving a particular problem or set of problems (Williams 1966; Tooby and Cosmides 1992). For example, these problems may be physiological (pumping blood; digesting food), perceptual (processing visual information; avoiding painful tissue damage), behavioural (catching prey; avoiding drowning), or social (finding a mate; acquiring a language). Adaptations tend to increase in complexity over evolutionary time, because more complex adaptations can often solve problems more competently than simpler ones (Heylighen 1999). For instance, a more complex immune system can defend against a more diverse array of diseases, and a more complex eye design can utilize visual information under more diverse conditions.

That is not to say that a more complex adaptive design will always be favoured over a simpler one. On the contrary, there is often a trade-off between complexity and efficiency; more complex adaptations are costlier to the organism in terms of time, energy, and likelihood of malfunction (Heylighen 1999; Orr 2000). A simpler design may be favoured if it is significantly less costly to the organism, but does not perform significantly less competently, than a more complex design. Adaptations are expected to be no more complex than they need to be, in order to solve an adaptive problem at least as well as competing designs. Some adaptations can be relatively simple (as in the pigmentation example below). Nonetheless, many adaptive problems do require complex solutions, and selection is the sole process that can provide them.

2.5 *All Organismal Traits Are Adaptations, By-Products of Adaptations, or Noise*

The fundamental design principle of organismal tissue is to be organized into functionally specialized adaptations. This does not imply, however, that all organismal traits are adaptations. A trait may be (a) an adaptation, (b) the by-product of an adaptation, or (c) stochastic noise.[2] An adaptation is a trait that functions to solve some adaptive problem, a by-product is a trait that happens to be coupled with an

---

[2] Gould and Vrba (1982) propose 'exaptation' as another category of organismal trait. They define exaptation as a trait that has a function that it did not evolve to fulfil in the first place. This could be a by-product that is subsequently fashioned by selection to fulfil a function, but is more commonly conceived as an adaptation that evolved to fulfil one function and that then got co-opted by selection to fulfil a different function. For purposes of this chapter, however, so-called exaptations can be regarded as equivalent to adaptations. That is, the 'adaptation' category here includes all functional organismal traits and so subsumes the 'exaptation' category.

adaptation but that itself serves no functional purpose, and noise is residual variance in trait design that does not affect survival and reproduction (Buss et al. 1998; Tooby and Cosmides 1992). An illustration provided in Price (2017) relates to the umbilical cord, a functional trait of placental mammals, which enables the foetus to access nutritious blood from its mother's placenta. This cord falls off after birth, leaving behind a navel that varies in size and shape across individuals. Although the umbilical cord is itself a complex adaptation, the navel is just its by-product, and the navel's exact dimensions are just noise.

How can we be sure that the human umbilical cord is an adaptation? Because of its improbable complexity. This complexity can be partially assessed in terms of the cord's structure itself. For example, the cord contains one vein to bring oxygenated blood to the foetus and two arteries to carry deoxygenated blood away; its outer lining is composed of two layers (epithelial and mesenchymal) and beneath it is a layer of protective Wharton's jelly; and inside the foetus, its vessels branch out in order to reach all four heart chambers, and a special passage (the ductus arteriosus) leads blood to the aorta.

To fully appreciate the improbability of the cord's complexity, however, we need to consider this complexity in light of its function. In doing so, we look for evidence of 'special design' (Williams 1966): to what extent is the cord a well-tailored solution to a particular adaptive problem, like a key that has been precisely engineered to fit a specific lock? The cord obviously functions to solve the adaptive problem of enabling the foetus to access nutrients from the mother. What is the probability that a complex structure such as the umbilical cord would arise by chance, as the result of stochastic (as opposed to selective) evolutionary processes, and just happen to be characterized by the precise design that enabled it to solve this highly specific problem? It is difficult to assign an exact probability, but safe to say that it is virtually zero (Williams 1966; Dawkins 1986).

2.6 *Adaptations Exhibit Improbable Complexity; By-Products and Noise Do Not*

In sum, in order to determine how likely the umbilical cord is to be an adaptation, we must conduct a probability assessment: how likely would it have been to arise by chance, in the absence of a non-random process of selection? The less likely this outcome, the more likely the cord is to be an adaptation. Using this same adaptationist framework, we can also assess the probability that the navel is an adaptation. The navel does not have a particularly complex structure, nor is there any evidence that its structure enables it to fulfil any function related to survival or reproduction. In other words, it displays no evidence of Williams' (1966) special design. The navel exhibits vastly less improbable complexity than the umbilical cord, and the most parsimonious explanation for its existence is clearly that it is a by-product of this cord, and not an adaptation in its own right. And what of the variance in navel size and shape that is exhibited across individuals (e.g. 'innie' versus ‘outie’)? This is best characterized as noise. No two navels are exactly the same, but this residual variance has no effect on survival or reproduction, and so it is invisible to selection.

2.7 All Adaptations Exhibit Improbable Complexity, but Some More than Others

Just as we can utilize the adaptationist framework to distinguish adaptations from by-products and noise, we can also use it to distinguish a trait that is more likely to be an adaptation from one that is less likely to be an adaptation. We can do this even if both traits are probably adaptations. An illustration in Price (2017) compares two human adaptations: the mammalian eye and the amount of melanin produced by the skin. The eye is highly and famously complex, composed of diverse components (lens, cornea, iris, optic disk, optic nerve, and many more) that work together in a coordinated fashion. The improbability of this complexity is compounded by the fact that it represents not just structural but functional complexity: it solves the adaptive problem of converting electromagnetic radiation into information that the organism can use to survive and reproduce.

The skin’s degree of melanin content is probably also an adaptation, but a relatively simple one compared to the eye. I am not referring to melanin production itself, which is a complex process, achieved by a specialized class of cells (melanocytes) located in the skin’s lower layer and specialized organelles (melanosomes) which transport melanin along dendrites to the skin’s outer layer. I am referring, rather, to the pattern across populations whereby people have more melanin in their skin if their ancestors lived in sunnier climates. This pattern may at first seem random, until we consider that it probably represents a functional calibration (Jablonski and Chaplin 2010). Decreased melanin appears to increase the risk of skin cancer, but to improve the body’s ability to produce vitamin D. It is likely that skin cancer was a less severe adaptive problem than vitamin D production for more northerly populations and that the relative importance of these problems was reversed for more southerly populations. When these levels are observed across populations, the overall pattern presents a degree of improbable complexity. This is because adjustments to pigmentation level appear to be a well-designed trade-off solution to the problems of avoiding cancer and producing vitamin D, and it seems unlikely that evolution would have achieved this solution by chance. Nevertheless, this cross-population pattern seems much less improbably complex than the eye.

So both the eye and melanin levels appear to display improbable complexity. But if we are considering the human phenotype overall, and comparing the probability of the eye being an adaptation to that of melanin levels being an adaptation, we would need to conclude that this probability is much higher for the eye. Even if we can be reasonably confident that both traits are adaptations, the eye displays vastly more improbable complexity than melanin levels, and so is vastly more likely to constitute an adaptation.

2.8 The First Law of Darwinian Adaptation

The adaptationist considerations presented above can be distilled into a simple law: the more improbably complex a phenotypic aspect is, the more likely it is to be a design feature of selection, that is, an adaptation. I will refer to this as the ‘first law of Darwinian adaptation’. Why is it true? Because as outlined above, Darwinian selection is the strongest known antientropic process in the universe, and the only known process that can act as an ultimate designer of improbable complexity. Selection is the creator of adaptations, and its hallmark is improbable complexity. The more unambiguously a trait displays this hallmark, the more certain we can be that it is an adaptation.

2.9 An Adaptationist Revision of Smolin’s CNS: Cosmological Natural Selection with Intelligence (CNSI)

The first law of Darwinian adaptation makes it clear why Smolin is probably mistaken in his attempt to identify the phenotypic aspect of the universe that is most likely to be an adaptation. If we are serious about applying Darwinian reasoning in order to understand the organization of the universe, we should follow this reasoning to its logical conclusion. And this conclusion seems unambiguous: life is the most improbably complex known aspect of our universe’s phenotype, and so in line with the first law of Darwinian adaptation, it should be regarded as more likely than any other known aspect to be a mechanism of universe reproduction. From this perspective, the evolution of intelligence is a developmental subroutine of cosmological natural selection, and the function of intelligence is to develop the technological ability to reproduce its own cosmic habitat.

This general idea – that intelligent life is a mechanism of universe reproduction – has been in circulation almost as long as Smolin’s CNS (1992, 1997) itself; a helpful history of it is available on the Evo Devo Universe (2017) website. Crane (1994/2010) did the first widely acknowledged work on this idea, and Harrison (1995) published the first peer-reviewed article on it. Gardner has published several articles (2000, 2001, 2002, 2004, 2005) and two books (2003, 2007) on it and coined ‘selfish biocosm’ to describe the notion of a universe with an evolved interest in reproduction. Vidal (2014) also provides a detailed exploration of this idea, which he refers to as cosmological artificial selection. Smart (2009) explores universe reproduction from the framework of evo-devo biology and refers to it as cosmological natural selection with intelligence (CNSI). That is the term I am using here for reasons explained below. Whatever terminology is used, the general idea of CNSI is not particularly new. What *is* novel, however, is to approach this idea from an explicitly adaptationist point of view, as I do here and in Price (2017). This approach is unique in proposing that from the perspective of Darwinian adaptationism, CNSI is a more logically coherent model of universe organization than Smolin's CNS.

It is important to emphasize that from this adaptationist point of view, CNSI emerges as a more reasonable model than Smolin's CNS simply because that is where the logic more directly leads: because intelligence is the most improbably complex aspect of the universe's phenotype, it is also the aspect most likely to be an adaptation. This conclusion raises a host of secondary implications, the most important being that 'future intelligence' (i.e. intelligent life, in whatever form it may eventually evolve into) will probably someday acquire the ability to reproduce its natal universe. This implication, in turn, raises the question of how future intelligence would be able to do this. We cannot know the correct answer to this question, although we can speculate (as I do below). Nevertheless, we should keep in mind that the plausibility of CNSI does not hinge on our present ability to definitively answer this or any other question raised by the implications of CNSI. We arrived at CNSI not because any of these questions necessarily have clear answers at present, but because we followed the adaptationist approach to its logical conclusion.

2.10 If Universe Reproduction Is the Lock, Intelligence Is a Well-Crafted Key

We should also keep in mind that from the CNSI adaptationist perspective, intelligence is a solution, designed by cosmological evolution, to the problem of how the universe can reproduce itself. We should therefore expect for intelligence to seem well-designed to solve this problem, like a key that has been precisely engineered to fit a specific lock. Does intelligence fulfil this expectation? It plausibly does, for two reasons.

First, intelligence should possess the requisite *interests* in solving this problem. Like all forms of life, future intelligence should be fundamentally interested in pursuing its own evolved goals, related ultimately to survival and reproduction. To pursue these goals, it will always require a suitable cosmic habitat, that is, a universe characterized by the same laws and parameters as that to which it has already adapted. Future intelligence will at some point very likely find itself in need of, or simply advantaged by, the ability to create a new cosmic habitat. For instance, this scenario could arise in the very far future due to the inevitable heat death of our current universe (Vidal 2014), or because future intelligence had colonized its natal universe's habitable regions so comprehensively as to exhaust possibilities for further expansion. This scenario could also arise in the relatively near future, if some faction of future intelligence were motivated to create a universe that it could have to itself, perhaps to escape competition from another faction. In any of these scenarios, the aims of the universe and of future intelligence would be perfectly aligned: the universe would be 'interested' in reproducing (because, like all products of selection, this interest has been designed into it), and future intelligence would be motivated to reproduce the universe, in order to promote its own evolved interests.

The second aspect of intelligence that makes it a good solution to this problem is that it could plausibly acquire the requisite *ability* to solve it. We do not know exactly what the future process of universe reproduction will involve, but it will presumably require the creation of something like black holes (Crane 1994/2010). For reasons discussed below, however, we should expect for any black hole-like structures designed by future intelligence to be significantly better suited to the task of universe reproduction than existing, 'natural' black holes. Whatever this process will involve, it will require technology considerably more advanced than that which humans currently possess. Given the exponential pace of human technological progress, however, it is realistic to expect that future intelligence will succeed in developing this technology. It took about 1,600,000 years to develop the hafted bronze axe from the stone biface axe, another 3700 thousand years to invent the steam engine, and another 280 years to arrive at the Internet and smart phone. The pace of technological progress continues to accelerate (Kurzweil 2005), and provided that intelligent life can avoid extinction (especially by its own hand), this acceleration process has no recognizable upper limit. It is impossible to imagine how far human technology will have advanced even one hundred years from now, let alone one thousand or one billion. This is especially true given the real possibility that a 'technological singularity' – an explosion of technological advancement, sparked by the invention of artificial superintelligence – will occur in the near future (Kurzweil 2005).

2.11 If the Universe Could Already Reproduce Via Black Holes, Would It Need Intelligence Too?

Both adaptationist CNSI and Smolin's CNS suggest that black holes generate new universes. The main difference between them is that CNSI alone proposes that the purpose of intelligence is also to achieve this goal. CNSI may therefore seem to imply a redundant function for intelligence: if cosmological evolution endowed the universe with trillions of relatively simple black holes in order to produce new universes, why would it also prescribe something as improbably extravagant as intelligent life, in order to accomplish the exact same goal? This seems like an absurdly inefficient design strategy. Instead of embedding the evolution of intelligence as a developmental subroutine in the universe's DNA, could not cosmological selection have just specified some extra black holes?

These are good questions, and they elicit compelling responses. Intelligence probably does not, in fact, merely replicate the function of natural black holes, and this is because it will probably prove to be the superior mechanism of high-fidelity universe reproduction. We do not know how capable black holes actually are at faithfully reproducing the laws and parameters of their parent universe. Smolin's CNS requires for inheritance between parent and offspring universes to be high fidelity yet imperfect; there must be some room for random mutation, as in biological evolution, so selection has variance to act upon. Black holes may in fact be capable of such high-fidelity inheritance. On the other hand, they may not, even to the point where designs of offshoot universes vary completely randomly with regard to parental designs, as suggested by Wheeler (1977) and others (Susskind 2004).

2.12 A Sufficiently Knowledgeable Intentional Design Process Would Be Superior at Universe Reproduction Than an Unintentional CNS Process

Regardless of how competent natural black holes might be at high-fidelity universe reproduction, it is important to keep in mind that a sufficiently sophisticated intelligence would almost certainly prove superior at this task. This is because natural selection, at either the biological or cosmological level, can engage only in unintentional design. That is, natural selection unfolds as if directed by a 'blind watchmaker' (Dawkins 1986), who lacks foresight and makes progress only by randomly stumbling upon design modifications that happen to prove advantageous. That is not to imply, of course, that natural selection cannot be a truly impressive designer of functional complexity. Biological selection's masterfulness in this regard is the very reason why it has been able to project, for thousands of years and in so many people's minds, such a powerful illusion of intentional, intelligent design. The technical sophistication of many of its adaptations remains unmatched by current human technology. We still do not understand our own brains well enough, for instance, to create an artificial brain that comes anywhere close to faithfully executing all the functions of the original.

On the other hand, as skilful as biological selection can be, the limitations resulting from its lack of foresight are famously evident (Dawkins 1986). It creates organisms that are riddled with design flaws, such as organs that have far-from-optimal designs or that have been rendered completely vestigial. And with the ever-accelerating pace of technological progress noted above, it will not be long before our technical prowess as designers of functional complexity *does* outstrip that of biological selection (Kurzweil 2005). Given sufficiently precise knowledge about the design of biological selection's most impressively complex accomplishments, we will inevitably be able to exceed these accomplishments. This is because we alone can engage in intentional design, guided by our own intelligence and foresight, and this power will ultimately afford us vast advantages over unintentional selection. Furthermore, the advantages of intelligence-guided intentional design over chance-driven unintentional design will probably ultimately become evident not just at the biological level but at the cosmological level as well. It will take an enormous amount of new knowledge before future intelligence can exceed whatever competence black holes may possess as mechanisms of high-fidelity universe reproduction. But given the accelerating pace of technological progress, and the vast inherent advantages of intentional over unintentional design, this outcome is likely to occur.

2.13 CNSI Does Not Imply That Our Own Universe Was Intelligently Designed

CNSI is sometimes misunderstood to propose not only that intelligence is our own universe's primary mechanism of reproduction, but that our universe was itself created by intelligence. This 'past intelligent design' proposal actually is made explicitly by one CNSI model (Harrison 1995), but is not made by the CNSI model proposed here and in Price (2017), for two main reasons. First, it seems to lack parsimony. From observing biological natural selection, we already know that a blind process of mechanistic selection can serve as a highly skilled creator of functional complexity, even though it involves no intelligence, foresight, or intent. The proposition that intelligence is an adaptation for universe reproduction necessarily implies that life was designed by cosmological natural selection, but not that it was produced by intelligent design. This latter implication seems superfluous, because it goes beyond what we have observed to be true of natural selection processes in general. The second reason that 'past intelligent design' is not proposed by CNSI is that it seems infinitely recursive: if our universe were generated by an alien intelligence in a parental universe, then presumably this parental universe was itself created by an alien intelligence in a grandparental universe and so on ad infinitum. Past intelligent design does not seem able to explain, therefore, how cosmological evolution could produce intelligence in the first place.

The CNSI model proposed here suggests that future intelligence will enable our own universe to reproduce, by initiating a process of what could be termed 'cosmological artificial selection' (Vidal 2014). For reasons just noted, however, it does not suggest that *past* iterations of this cosmological evolutionary process have involved intentional artificial selection, rather than unintentional natural selection. The somewhat subtle reason I have chosen the term 'cosmological natural selection with intelligence' (CNSI; Smart 2009) to describe this process, instead of the more elegant 'cosmological artificial selection' (CAS), is because CNSI retains the key concepts of both natural selection and intelligence. CAS, on the other hand, might be misunderstood to imply that cosmological selection is expected to always be artificial, rather than natural. (It should also be emphasized, however, that this interpretation would indeed constitute a misunderstanding, as even Vidal [2014] himself does not actually regard the CAS term as implying that cosmological selection will always be artificial.)

2.14 CNSI Implies a Progressive Direction to Human Biocultural Evolution

CNSI may elicit scepticism, especially from those who would regard it as unacceptably ‘teleological’: it implies that we can understand human biocultural evolution as having the fundamental ‘purpose’ of enabling universe reproduction. Teleological thinking is often seen as a flawed way to reason about evolution, because organisms can be designed only by *past* selection pressures. Evolution, that is, can be guided only by what has happened and not by what will happen (Tooby and Cosmides 1990). Although it is true that adaptations can only be products of the past, it is also true that an adaptation can reasonably be said to have a function for which selection designed it and to therefore have a ‘purpose’. Adaptations can have functions because the environments of organisms have many stable aspects, and past selection pressures are therefore often reliable guides to future selection pressures. For example, if an organism evolves in an environment characterized by the stable presence of both gravity and a landscape of varying elevation, then it may evolve adaptations to protect it from falling (such as a fear of heights). Even though these adaptations are shaped entirely by the past, we can expect that in the future, they will fulfil the function (or purpose) of helping the organism to avoid falls.

By this same logic, if CNSI proposes that intelligent life is an adaptation for universe reproduction that evolved by cosmological selection, it implies that enabling this reproduction is the function, or purpose, of life. It also implies that we can reasonably expect for life to eventually succeed in fulfilling this purpose. That outcome is not, however, guaranteed; adaptations sometimes do fail to fulfil their functions, especially as the result of some atypical environmental change. Human lungs, for example, will start failing as soon as they find themselves in an environment lacking breathable oxygen. Nevertheless, the default expectation for any evolved adaptation is that it will actually develop and function successfully.

If we regard humanity as being currently engaged in the process of fulfilling its function, then this suggests a progressive direction to human biocultural evolution (Gardner 2000). Some of the most dominant trends in this evolutionary process, such as those towards greater intelligence, increasingly sophisticated science and technology, and higher levels of complex cooperation (Wright 1999; Pinker 2011), seem consistent with the notion that intelligence is indeed headed in the direction predicted by CNSI. These observations also have important implications for the evolution of human moral systems, as discussed below.

3 Part 2: Philosophical Implications of CNSI

Darwinian selection is the only known natural process that can generate function and, therefore, purpose. Biological natural selection is the reason why organisms are endowed with purposeful adaptations. It is also the reason why organisms can behave purposefully, with goals, motives, desires, interests, and so on. Just as selection is the only naturalistic theory that can explain why an individual organism would behave purposefully, it is also the only naturalistic theory that can explain how life itself could have a 'transcendent' purpose, above and beyond the bioculturally evolved interests of organisms themselves. Because CNSI is the only existing theory of how life could be a design feature of cosmological selection, it is unique in its ability to propose that life could have a purpose that is both natural and, in this sense, transcendent.

There are, of course, many non-naturalistic belief systems which propose a transcendent purpose for life. As these systems are only able to propose supernatural causation, however, they are generally regarded as being 'religious' and outside the realm of scientific inquiry. (For the sake of convenience, I will use 'religious' to refer to any belief system that proposes supernatural causation, even if its proponents might describe it not as religious but as, for instance, 'spiritual'.) These systems generally propose that reality is determined by some purposeful higher power or order and that life's purpose and meaning can be understood only with reference to this transcendent force. This force is often characterized as being instantiated in a god or gods, but may also be conceptualized more abstractly (e.g. Buddhism's karmic cycle).

I will refer to the kind of reasoning described above – the belief that reality is controlled by a purposeful transcendent force – as 'transcendent teleological thinking' (TTT). A typical and apparently universal feature of religious belief systems, including all major modern religions, is that they explicitly sanction and encourage TTT (Banerjee and Bloom 2014; Bering 2002; Heywood and Bering 2014; Norenzayan and Lee 2010). In contemporary industrialized societies, TTT may in fact be the single most fundamental aspect distinguishing whether a belief system is considered religious as opposed to naturalistic. This is true not just because all major religious systems explicitly sanction TTT, but also because naturalistic belief systems generally do not. CNSI represents an exception to this rule, however, because it is a naturalistic belief system that could accommodate TTT.

Another way CNSI is similar to religious philosophies is that in accommodating TTT, it also implies a certain moral order. In his classic work on the psychology of religion, James (1902) offers a compelling general definition of religiosity: 'the belief that there is an unseen order, and that our supreme good lies in harmoniously adjusting ourselves thereto'. In other words, religions provide not just TTT but also a basis for morality: adjusting one's self to an 'unseen order' means aligning one's values and moral beliefs to be in harmony with this order. The cosmic order implied by CNSI would seem to imply a kind of moral direction as well, although this direction seems identifiable only in fairly broad terms.

3.1 Moral Implications of CNSI

Broadly, then, there are two general categories of values that seem particularly compatible with the CNSI framework. The first category relates to the production of knowledge. CNSI predicts that future intelligence will ultimately succeed in creating new universes, a feat which will require scientific and technical capabilities that exceed, by a vast margin, those which we currently possess. CNSI thus strongly values continuous progress in the production of scientific knowledge. The second category relates to social cooperativeness, stability, and strength: although the production of scientific knowledge is propelled forward by the genius of individuals, it is ultimately a highly cooperative process. Progress depends on knowledge being widely shared, and on widespread support for large-scale scientific projects that are expensive and complexly cooperative. Such progress is more likely to occur in societies that are stable, educated, and economically productive.

The values implied by CNSI may seem vague, but this vagueness may be necessary. No matter how scientifically supported CNSI turns out to be, the moral framework it offers may always be sketchable in only quite general terms. These terms may be necessarily broad because, even if they did offer some instruction about how to align one's values with an 'unseen order' (i.e. the CNSI-implied direction of human progress), this is not the only evolutionary order, seen or unseen, of relevance to humans. Even if human minds and cultures, like all entities in the universe, could be seen as being ultimately caused by cosmological evolution, they are also the more proximate products of biocultural evolution. We cannot escape the fact that our minds and cultures were designed directly by biocultural evolutionary processes, even if these processes themselves could be regarded as developmental subroutines of CNSI. In short, a CNSI-guided moral system must be compatible not just with the implications of cosmological evolution, but also with the fact that humans were designed proximately by biocultural evolution.

Compatibility with biocultural evolution seems especially important when we consider the design of intelligence, in relation to its function. As noted above, intelligence seems like a well-engineered solution to the problem of universe reproduction, like a key designed for a specific lock. This is because intelligence possesses the requisite interests (in its own survival), and should eventually possess the requisite ability (technical expertise), to solve this problem. In order to enable intelligence to solve this problem, a CNSI-guided moral system would need to ensure that people remained maximally motivated both to maintain their own survival and to overcome formidable barriers to technological progress. This moral system must consider, therefore, the kinds of rewards that human nature is designed to find desirable and enjoyable. The more people perceive their lives as rewarding and worth living – the greater, that is, their sense of subjective well-being – the more they will value their own survival, and the more willing they will be to do the hard work that cultural evolution requires.

It is also worth asking why individuals would *want* to align themselves with a CNSI-guided moral system. Biological evolution has adapted the individual to prefer activities that further his or her own evolved interests (i.e. interests that would have promoted one's survival and reproduction in the ancestral past [Tooby and Cosmides 1992]). What would individuals gain from embracing a CNSI-guided morality? The answer is that they would gain the advantage of aligning themselves with the direction of cultural progress that is likely to prevail. If CNSI correctly indicated the direction of cultural evolution, it would provide people with an objective basis for defining cultural progress. By aligning their values with this direction, they would put themselves in the best position to be on the 'winning side' of cultural evolution. And individuals gain by being on teams that win: humans are adapted to pursue their individual interests, but often do so most effectively by joining successful groups (Alexander 1987).

4 Conclusion: As a Potentially Unifying Framework for Physics, Biology, and Moral Philosophy, CNSI Merits Further Consideration

CNSI is a speculative theory, but one that applies a simple Darwinian idea – that traits exhibiting higher improbable complexity are more likely to be adaptations – as parsimoniously as possible. The conclusion that life is more likely than black holes (or anything else) to be an adaptation for universe reproduction requires no contortion of adaptationist logic; it is simply the destination to which this logic leads. Smolin's CNS, in contrast, diverges from this adaptationist path.

The main goal of both this chapter and Price (2017) is simply to propose that if we wish to apply Darwinian adaptationist theory to cosmological evolution, we should do so as straightforwardly as possible. The efforts to do so here and in Price (2017) attempt to develop a CNSI theory that could ultimately and directly account for not just non-biological complexity in our universe but biological complexity as well. CNSI may facilitate the integration, within a single adaptationist framework, of biology, physics, and a rudimentary moral philosophy (i.e. beliefs about the 'purpose' of life and the general direction of moral progress). The potential explanatory power of CNSI, combined with its parsimony and plausibility, will hopefully encourage more researchers to develop it further.

Acknowledgements Thanks to John Smart and Clément Vidal for detailed comments on the manuscript, and thanks also to John Campbell, Georgi Yordanov Georgiev, Claudio Flores Martinez, and Jade Price.

References

Alexander, R. D. (1987). *The Biology of Moral Systems*. New York: Aldine de Gruyter.

Banerjee, K., & Bloom, P. (2014). Why did this happen to me? Religious believers' and non-believers' teleological reasoning about life events. *Cognition*, *133*(1), 277–303.

Bering, J. M. (2002). The existential theory of mind. *Review of General Psychology*, *6*(1), 3–24.

Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L., & Wakefield, J. C. (1998). Adaptations, exaptations, and spandrels. *American Psychologist*, *53*(5), 533.

Campbell, D. T. (1974). Evolutionary epistemology. In P.A. Schilpp (Ed.), *The Philosophy of Karl R. Popper*. The Library of Living Philosophers. Lasalle, IL: Open Court Publishing Company, Volume 14-1, 413–463.

Campbell, J. O. (2011). *Universal Darwinism: The Path of Knowledge*. s.l.: CreateSpace.

Campbell, J. O. (2015). *Darwin Does Physics*. s.l.: CreateSpace.

Campbell, J. O. (2016). Universal Darwinism as a process of Bayesian inference. *Frontiers in Systems Neuroscience*, *10*.

Crane, L. (1994/2010). Possible implications of the quantum theory of gravity: an introduction to the meduso-anthropic principle. arXiv:hep-th/9402104v1. Reprinted in Crane, L. (2010). Possible implications of the quantum theory of gravity: an introduction to the meduso-anthropic principle. *Foundations of Science*, vol. 15, pp. 369–373.

Cziko, G. (1997). *Without Miracles: Universal Selection Theory and the Second Darwinian Revolution*. MIT Press.

Darwin, C. R. (1859). *On the Origin of Species*. London: John Murray.

Darwin, C. R. (1871). *The Descent of Man, and Selection in Relation to Sex*. London: John Murray.

Dawkins, R. (1986). *The Blind Watchmaker*. New York: Norton.

Dennett, D. (1995). *Darwin's Dangerous Idea*. New York: Schuster.

Evo Devo Universe (2017). Cosmological natural selection (fecund universes). Viewed 12 Oct. 2017 at http://evodevouniverse.com/wiki/Cosmological_natural_selection_(fecund_universes).

Gardner, A. & Conlon, J. P. (2013). Cosmological natural selection and the purpose of the universe. *Complexity* 18: 48–56.

Gardner, J. N. (2000). The selfish biocosm. *Complexity* 5: 34–45.

Gardner, J. N. (2001). Assessing the robustness of the emergence of intelligence: Testing the selfish biocosm hypothesis. *Acta Astronautica* 48: 951–955.

Gardner, J. N. (2002). Assessing the computational potential of the eschaton – testing the selfish biocosm hypothesis. *Journal of the British Interplanetary Society* 55: 285–288.

Gardner, J. N. (2003). *Biocosm: The New Scientific Theory of Evolution: Intelligent Life is the Architect of the Universe*. Inner Ocean Publishing.

Gardner, J. N. (2004). The physical constants as biosignature: An anthropic retrodiction of the selfish biocosm hypothesis. *International Journal of Astrobiology* 3: 229–236.

Gardner, J. N. (2005). Coevolution of the cosmic past and future: The selfish biocosm as a closed timelike curve: A recipe for cosmic ontogeny and a blueprint for cosmic reproduction. *Complexity* 10: 14–21.

Gardner, J. N. (2007). *The Intelligent Universe: AI, ET, and the Emerging Mind of the Cosmos*. New Page Books.

Garriga, J., Vilenkin, A., & Zhang, J. (2016). Black holes and the multiverse. *Journal of Cosmology and Astroparticle Physics*, *2016*(02), 064.

Gould, S. J. & Vrba, E. S. (1982). Exaptation – a missing term in the science of form. *Paleobiology* 8 (1): 4–15.

Guth, A. H. (1981). Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, *23*(2), 347.

Hamilton, W. (1964). The genetical evolution of social behaviour I. *Journal of Theoretical Biology*, vol. 7, no. 1, pp. 1–16.

Harrison, E. R. (1995). The natural selection of universes containing intelligent life. *Quarterly Journal of the Royal Astronomical Society* 36: 193–203.

Heylighen F. (1999): "The Growth of Structural and Functional Complexity during Evolution", in: F. Heylighen, J. Bollen & A. Riegler (eds.) *The Evolution of Complexity* (Kluwer Academic, Dordrecht), p. 17–44.

Heywood, B. T., & Bering, J. M. (2014). "Meant to be": How religious beliefs and cultural religiosity affect the implicit bias to think teleologically. *Religion, Brain & Behavior*, *4*(3), 183–201.

Jablonski, N. G., & Chaplin, G. (2010). Human skin pigmentation as an adaptation to UV radiation. *Proceedings of the National Academy of Sciences*, *107*(Supplement 2), 8962–8968.

James, W. (1902). *The Varieties of Religious Experience: A Study in Human Nature.* New York: Longmans, Green & Co.

Kurzweil, R. (2005). *The Singularity is Near*. London: Penguin.

Linde, A. D. (1986). Eternally existing self-reproducing chaotic inflationary universe. *Physics Letters B*, 175(4), 395–400.

Linde, A. (2017). A brief history of the multiverse. *Reports on Progress in Physics*, *80*(2), 022001.

Norenzayan, A., & Lee, A. (2010). It was meant to happen: explaining cultural variations in fate attributions. *Journal of Personality and Social Psychology*, *98*(5), 702.

Orr, H. A. (2000). Adaptation and the cost of complexity. *Evolution*, *54*(1), 13–20.

Pinker, S. (2011). *The Better Angels of Our Nature: The Decline of Violence in History and its Causes*. London: Penguin.

Price, M. E. (2017). Entropy and selection: Life as an adaptation for universe replication. *Complexity*, vol. 2017, Article ID 4745379, 4 pages, 2017. doi:https://doi.org/10.1155/2017/4745379

Schrödinger, E. (1944). *What Is Life? The Physical Aspect of the Living Cell*. Cambridge University Press.

Smart, J. M. (2009). Evo devo universe? A framework for speculations on cosmic culture. In Cosmos and Culture: Cultural Evolution in a Cosmic Context, S. J. Dick and M. L. Lupisella, Eds., pp. 201–295, Government Printing Office, NASA SP-2009-4802, Washington, DC.

Smolin, L. (1992). Did the universe evolve? *Classical and Quantum Gravity*, vol. 9, no. 1, pp. 173–191.

Smolin, L. (1997). *The Life of the Cosmos*. New York: Oxford University Press.

Smolin, L. (2006). The case for background independence. In Dean Rickles, Steven French, and Juha Saatsi (Eds.), *The Structural Foundations of Quantum Gravity*, pp. 196–239. Oxford: Oxford University Press.

Susskind, L. (2004). Cosmic natural selection. arXiv:hep-th/0407266v1.

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology*, *11*(4–5), 375–424.

Tooby J. & Cosmides L. (1992). The psychological foundations of culture. In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, J. Barkow, L. Cosmides, and J. Tooby, Eds. Oxford University Press.

Vidal, C. (2014). *The Beginning and the End: The Meaning of Life in a Cosmological Perspective*. Springer.

Wheeler, J. A. (1977). Genesis and observership. In *Foundational Problems in the Special Sciences*. Butts, R. E. & Hintikka, J. (Eds.), pp. 3–33. Dordrecht: Reidel.

Williams, G. C. (1966). *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton University Press.

Wright, R. (1999). *Nonzero: The Logic of Human Destiny*. Pantheon.

Zurek, W. H. (2003). Quantum Darwinism and envariance. *arXiv:quant-ph/0308163.*

Zurek, W. H. (2009). Quantum Darwinism. *Nature Physics*, *5*(3), 181–188.

Evolutionary Development: A Universal Perspective

John M. Smart

1 Definition and Overview

"Evolutionary development," "evo devo" or "ED," is a term that can be used as a replacement for the more general term "evolution," whenever any scholar thinks that both experimental, creative, contingent, stochastic, and increasingly unpredictable or "evolutionary" processes and conservative, convergent, statistically deterministic (probabilistically predictable), or "developmental" processes, including replication, may be simultaneously contributing to selection and adaptation in any autopoetic (self-reproducing) complex system.

The hyphenated "evo-devo" is commonly used for living systems, most prominently in evo-devo genetics and epigenetics, and the unhyphenated "evo devo" can be used for the theory of any potentially replicating and adapting complex system (star, prebiotic system, gene, cell, organism, meme (concept), behavior, technology), whether living or nonliving. Occasionally, the hyphenated term "eco-evo-devo" is used to place evo-devo biology within ecological systems with their own evo devo dynamics (e.g., Pigliucci 2007; Gilbert et al. 2015). This is controversial, since the conventional neo-Darwinian Modern Synthesis does not recognize multi-level selection, and maintains that ecological dynamics are secondary to species competition. But the rise of theoretical and systems ecology and its models, including ecological energetics, panarchy, and ascendancy, can be viewed as supporting the

---

Support: No institutional support or grant was used for this work.

J. M. Smart (✉)
Acceleration Studies Foundation, Los Gatos, CA, USA

Naval Postgraduate School, Monterey, CA, USA
e-mail: ejsmart@nps.edu

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_2

idea that ecologies themselves both evolve and develop. Finally, future evo devo models may require what Lucia Jacobs refers to as “cog-evo-devo” (Jacobs 2012), the recognition that both information and cognition evolve and develop, are causal agents in the dynamics of complex replicators, and are increasingly important in determining their future as higher intelligence emerges. Inspired by the work of evo-devo biologists, evo devo systems theorists look for processes of evolutionary creativity and developmental constraint in any autopoetic complex systems, at any scale.

Evo devo systems theory thus redefines the much-used but increasingly multi-causal term “evolution,” to restrict evolutionary process to contingent, stochastic, information-creative, experimental, diversifying, and nonhierarchical processes of system change, when we are referring to evolution *within the boundaries of any proposed autopoetic system*. These processes are the dynamical and informational opposite of the predictable, information-conservative, convergent, unifying, and hierarchical processes of “development,” which work to replicate and maintain that system. Redefinitions of long-used words are never a popular choice, but this redefinition is potentially clarifying for autopoetic dynamics, from the perspective of information theory. If evolutionary processes necessarily generate new information, and developmental processes conserve and build upon old information, and we can determine “new” or “old” only in relation to the life cycle of the system under analysis, we may have a useful new perspective on both dynamics and their intrinsic predictability to observers within any self-reproducing system.

When we apply these definitions to the life cycle of an individual organism, such as a frog, we can observe evolutionary, information-creative processes in such events as stochastic gamete production, and in the stochastic cellular microarchitecture in any specific frog. Simultaneously, we can observe developmental, information-conservative processes in any replicative dynamics, informatics, and morphology that we empirically observe in all frogs of a specific species. Both evolutionary and developmental processes can thus be empirically differentiated in any living complex system via these definitions. Both processes are presumably fundamental to adaptation, and the ways each system encodes representations (models, intelligence) of itself and its environment.

Again, within any particular autopoetic system, evo devo models redefine the word “evolution” to refer specifically to variety-generating, experimental, divergent, and other “soon-unpredictable” processes that generate combinatorial explosions of contingent possibilities. They use the word “development” to refer to variety-reducing, conservative, convergent, and other “statistically predictable” processes that manage replication. These developmental processes are intrinsically predictable if you have the right models and sufficient computation capacity, the right perspective (often a collective, big picture, or long-term view) or if you have empirical experience, having seen a prior life cycle of the developing system in question (a cell, a tree, a human, a stellar system, a galaxy, a universe).

Independently, and via similar reasoning, some scholars occasionally use the term “evolutionary development” as a replacement for “evolution” as it juxtaposes two dynamically and informationally opposing concepts – “random,” unguided

Darwinian evolution and nonrandom, guided development – and thus is a more conservative and humble descriptive term to use when one is uncertain whether the change one is talking about is random or directional. For either reason, in addition to occasional early use by physicists (Turchin 1977) and origin of life scholars (Oparin 1968), a small but growing group of ecologists (Salthe 1993), biologists (Losos 2017), paleontologists (Conway-Morris 2004), theoretical biologists (Reid 2007), cosmologists (Munitz 1987), complexity theorists (Levin 1998), and systems theorists (Smart 2008) find it valuable to use the "evolutionary development" term.

The start of the journal *Evolution & Development* in 1999 signaled rise of evo-devo biology to a formal subdiscipline. Evo devo systems science and philosophy presently has no journal. If it did, *Evolution, Development, and Complexity*, the name we use for this volume, would be a reasonable title. This would be a journal within which the complexity science and systems theory of such topics as universal Darwinism, evo devo cosmology, evo-devo biology, eco-evo-devo, cog-evo-devo, living systems theory, technological evo devo, artificial and biological intelligence, hierarchy theory, accelerating change, and related topics might be modeled and critiqued. Biologically-Inspired Complexity Science and Philosophy (BICS&P) would be reasonable title for our emerging field itself. BICS&P is the self-description of our Evo Devo Universe research and discussion community (EvoDevoUniverse.com).

2 Two Polar Categories and Tensions

Table 1 (Smart 2008) introduces sets of two polar (equal and opposite) word pairs that can be associated with evolutionary and developmental processes in a range of complex systems. As you look them over, think of all the events, processes, and systems you have previously described with these words. These and similar words, and the concepts behind them, are often useful starts at categorizing social, economic, and technological events and processes into one of two camps.

Some systemic processes operate by chance, others by necessity, and some by both. Some processes are random, others predestined. Some events are indeterminate, others predetermined. Some processes are segregating, others integrating. Some act bottom-up, others are top-down. Some systems appear to be branching, others funneling. Some changes look reversible, others irreversible. Some are generating novelty, others conserving sameness. Some are exploring possibilities, others running into constraints. Some promote variability, others stability. Some degrade hierarchies, others create hierarchy. In the organization, good foresight and strategy requires a continual balance between divergent (innovative, experimental) and convergent (predictive, conservative) thinking. We can see these twin tensions, and their mixture, in all the ways humans use for knowing the world.

In the twentieth century, we learned that even our scientific laws fall neatly into these two categories. From our reference frame, not only have we discovered deterministic (developmental) types of laws that precisely describe the far future, like

Table 1 Common evo and devo terms

Evolution	Development
Unpredictability	**Predictability** (*statistical*)
Chance	Necessity
Indeterminacy	Determinism
Random	Destined
Divergent	Convergent
Reversible (long-term)	Irreversible (long-term)
Possibilities	Constraints
Variety/Many	Unity/Monism
Variability	Stability
Uniqueness	Sameness
Transformation	Transmission
Accidental	Self-organizing
Bottom-up	Top-down
Local	Global
Immaturity	Maturity
Individual	Collective
Instance	Average
Short-term	Long-term
Reductionism	Holism
Analysis (breaking)	Synthesis (joining)
Amorphous	Hierarchical/Directional
Innovative	Conservative
Creativity (of novelty)	Discovery (of constraint)
Period-doubling/Chaos	Period-halving/Order
Experimental	Optimal
Dispersion	Integration
Dedifferentiation	Differentiation
STEM recombination	STEM compression
Nonergodicity	Ergodicity
Innovation	**Sustainability**
Belief (unproven)	**Knowledge (verified)**

Source: ASF

the equations of classical mechanics and relativity, we've also found stochastic and statistical (evolutionary) physical laws, like quantum mechanics, thermodynamics, and nuclear physics.

We have learned we can view physical and informational systems as either deterministic or stochastic, depending on the analytical reference frame we adopt. Deterministic laws are highly conserved and predictable at the individual level (i.e., the laws of motion for individual objects), yet become unpredictable at the collective/emergent level (i.e., the N-body problem in physics). Stochastic laws are random, novel, and creative at the individual level (the quantum state or entropy of any particular system, the decay of any particular nucleus), and yet are probabilistically predictable at the collective level. We see a simple example in radioactive half-life, and more complex examples in non-equilibrium thermodynamics, self-organized criticality, and phase transition thermodynamics. Such factors as the reference frame of the observer with respect to the system, the scale at which they are observing the system, and the duration of observation relative to a (presumed) autopoetic cycle all seem to influence the ease and extent of predictability in nature.

Many social, economic, and political processes historically alternate between unpredictable and divergent (evolutionary) and predictable and convergent (developmental) phases (cf. Vermeij 2009). For every social issue, we can find processes simultaneously generating "evolutionary" variety and "developmental" convergence, in comparative analyses of different cities, counties, states, countries, or regions. For example, regarding economic inequality, we find great "evolutionary" variations, country by country, in the levels and quality of social services available to each citizen and in the cycles of increasing or decreasing inequality. Yet we also find a long-term "developmental" trend of predictably increasing total economic inequality (relative and absolute rich-poor divides) the greater the flow rates of capital, goods, and information in any societies we analyze (Bejan and Errera 2017). The two opposing perspectives and tensions of evolution and development (unpredictability or predictability) appear to be equally fundamentally useful ways to view the world.

Use of the paradox-containing evo devo term also communicates our humility and ignorance when we are asked whether evolutionary (divergent, contingent) or developmental (convergent, inevitable) processes are presently dominating in any system or environment. We usually don't know which processes are most in control of physical or informational dynamics, at first glance. Careful study, modeling, and data collection may be required to see where any complex system is presently headed, process by process.

3 Contingency Versus Inevitability: The Two Extremes of Scientific and Societal Bias

One might think that the existence of inevitable, developmental physical and informational processes of change is as obvious, in modern scientific practice and philosophy, as are unpredictable contingent, evolutionary processes of change. It has been more than 300 years since universally inevitable celestial mechanics was elucidated (Newton 1687), a century and a half since we discovered the second law of thermodynamics (Clausius 1851), and a century since Einstein reformulated mechanics into an even more general inevitabilist framework (Einstein 1915), allowing us a deeper understanding of both space-time and energy-matter and predicting such still incompletely understood emergent phenomena as black holes, and perhaps even dark energy, via the cosmological constant. These and many other physically well-characterized processes are developmental constraints within which all of life's stochastic evolutionary processes must occur. Surely such examples must lead us to realize that there are likely to be many other statistically predictable macrotrends and inevitable emergences in life, society, and technology, waiting to be discovered and measured empirically, and eventually more rigorously characterized by physical and informational theory and simulation.

Unfortunately, there is presently a strong practitioner and philosophical bias against inevitabilist thinking in most scientific, technical, economic, political, and cultural communities, particularly since the rise of chaos theory and nonlinear science in the 1970s and of subjectivist postmodernism in our academic institutions in the late twentieth century. Humanity is guilty of periods and domains of overapplied developmentalist thinking, as in the various clockwork universe models of the eighteenth and nineteenth centuries (most famously, Laplace 1812). More recently, many Western nations overapplied reductionist and logical positivist thinking in our think tanks, corporate strategy, and government plans in the mid-twentieth century. The above brief history illustrates that our dominant societal biases have tended to each of two extremes (contingency or inevitability in various human futures) in a chaotic and cyclic dynamic. A few decades or generations hence, perhaps after some particularly predictive scientific or technical advance, we may again swing to the opposite extreme, and adopt an overly developmentalist bias, at least in particular scientific or societal subcultures.

A more adaptive position, rather than swinging to extremes, might be to recognize that both predictable and unpredictable processes are always occurring within any complex system, and to try to better understand each. We can identify at least a few inevitable (developmental) processes, as well as many contingent (evolutionary) processes in any complex system we analyze, including the universe as a system. Chaos theory and sensitive dependence on initial conditions apply to some universal processes, but certainly not to all processes. Even our modern philosophy of science, while it acknowledges a "contingentist" and "inevitabilist" debate with respect to the results of scientific experiments (cf. Martin 2013) does not yet acknowledge that both positions are always true, in any complex system, from different perspectives, as we will discuss. More disturbingly, modern science and complexity theory also rarely ask how each apparently fundamental process interrelates, and how each must contribute to selection and adaptation.

Today, we are primarily contingentists, and so we are biased to under-recognize and under-seek statistically inevitable processes of change, and there are social blocks and professional costs to significant inevitablist thinking in the social and technical sciences. For one example of the costs of this bias, consider the following potential sociotechnical developmental process, one that I study and find particularly important. We can characterize a "general Moore's law" of exponentially growing computational capacity per dollar, observable since the 1890s at least (Kurzweil 1999), involving exponential growth in performance and resource efficiency the further our computing processes move into meso-, nano-, and quantum-scale realms (Smart 2000). This computational performance and efficiency acceleration via physical miniaturization appear to be a developmental macrotrend in human history, one likely to occur on all Earthlike planets, within some predictable stochastic envelope, if those planets harbor intelligent technology-using life, regardless of their political economies and cultures.

This example of one very socially relevant potential statistical inevitability, still poorly characterized in our physical and informational theory, has been largely ignored by academic and complexity science communities alike. Only futurists like myself, and a handful of philosophers and social theorists, seem interested in writing about it and asking about its causal dynamics. I have followed the literature in this area since the late 1990s, and I can assure you that the number of funded science or engineering researchers considering this process is minuscule, even today. Contingentist bias, in my view, is the simplest likely explanation for this state of affairs.

The Santa Fe Institute, a leading US complexity science research organization, tried at least three times (2009–2011) to get the NSF to fund a Performance Curve Database (PCDB) project (see http://pcdb.santafe.edu/), simply to collect better data on predictable exponential trends in technological performance efficiency, to aid in empirical and theoretical models of these fascinating and still accelerating processes. The requested modest funds were denied, and the SFI postdoc leading the grant applications, Bela Nagy, a personal friend, left his scientific career soon afterward, in part due to his disillusionment with the conservative funding priorities of Big Science. The PCDB remains unfunded today, and I know of no other similar project yet in any nation. Perhaps collecting data on technical exponentials, the fastest-changing and most economically and socially disruptive processes in human society today, wasn't considered a high enough priority for the grantors, due to our current scientific and societal bias toward a primarily contingentist view of social change. Perhaps also, NSF politicians didn't want the controversy of being seen as aiding the inevitablist perspective (see Kurzweil 1999, 2005) on scientific and technical acceleration. I do not know the details, but would be curious to see a causal study done.

Given the reality of contingentist bias, those who write about technological development or accelerating change from a macrohistorical perspective today are often pejoratively labeled as technophiles, utopians, or positivists, when all they are trying to do is establish that both unpredictable evolutionary paths, wherein we must exercise our free moral choice, and predictable yet causally opaque developmental processes, like technical acceleration, and destinations, like societal electrification, digital computers, or machine intelligence, appear to coexist in our complex universe. Our ability to see not only evolutionary change but also simultaneous processes of ecological, societal, technical, global, and universal development suffers greatly as a result.

4 The VCRIS Model of Natural Selection in Autopoetic Systems

If we wish to understand natural selection in autopoetic systems, both living and nonliving, we must better characterize dynamical change and develop better theories of information and intelligence. The VCRIS (“vee-kriss”) evo devo conceptual model (Smart 2017a) may be a useful, small step toward these challenges, especially when contrasted to the classic VIST model (variation, inheritance, selection, time/cumulative replication, Russell 2006) of dynamics offered by traditional evolutionary theory. The VCRIS model proposes that three sets of physical and informational dynamics must be modeled to understand and predict the outcomes of natural selection in autopoetic systems. The first two are fundamentally oppositional processes, and the third arises from their interaction. These are:

1. Variational or “evolutionary” processes that generate, maintain, and manage diversity, divergence, and experiment. When we observe them from within any autopoetic system, these processes grow increasingly unpredictable over time.
2. Convergent or “Developmental” processes that attract, constrain, and guide the system through hierarchical stages of form and function. When observed from within any autopoetic system, these processes grow increasingly predictable over time.
3. “Evo Devo” processes that are Replicative, with Inheritance of informational and physical parameters, under Selection for adaptation. Selection can favor either or both “evolutionary” (variational) or “developmental” (convergent) dynamics in the replicator, depending on context, and these two processes are particularly fundamental ways to understand selection. Adaptation, in turn, depends on the encoding of information (intelligence) in three places: Replication (organism, autocatalytic) processes, Inheritance (seed, gene, parametric) processes, and Selection (environment) processes (Fig. 1).

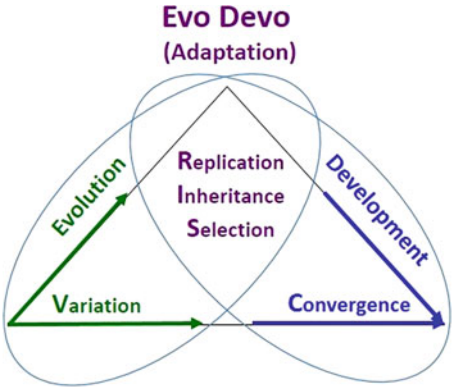


Fig. 1 The VCRIS conceptual model of natural selection in autopoetic systems. Variation (evolutionary process) and convergence (development), operating under replication and inheritance, can be viewed as fundamental physical and informational dynamics that mediate selection in self-reproducing systems. (Source: ASF)

In the VCRIS model, physical and informational processes that *change unpredictably in successive replication cycles*, to generate, maintain, and manage Variety, are in fundamental tension and opposition with physical and informational processes that *change predictably in successive replication cycles*, and thus generate, maintain, and manage Convergence. V and C are the first two terms in the VCRIS model, as these two oppositional processes are proposed as "root perspectives" in any model of physical and informational change in autopoetic systems, including our universe itself, if it is a replicating and adaptive system, as various theorists have proposed (Smolin 1992, 1997, 2004; Vaas 1998; Vidal 2010; Price 2017). Standard evolutionary theory offers no model of this fundamental opposition, of the inheritance and tension between two classes of informational-physical initiating parameters (evo and devo) at every scale at which replication occurs, including gene, epigene, organism, group, niche, environment, and universe.

If our universe is an autopoetic system, the VCRIS model offers us a new term to understand selection, a term that juxtaposes two fundamental binaries, those things that change and those that converge to stay the same, in any replication cycle. "Unpredictable predictable" is a term a physicist might favor, yet evolutionary development (evo devo) seems more precise, as it uses our model of replication (life cycle) as a way to define those things that predictably stay the same, in prior and parallel life cycles.

In toy cellular automata universe models, like Conway's Game of Life, the spatiotemporally repetitive structures and dynamics that we see in each successive game (replication cycle) can be defined as predictable, convergent, and developmental. Such reliably emergent structures and dynamics are robust to variation in most of the game's initial conditions (occupied configurations within the initiating matrix), yet they are also finely sensitive (finely tuned) to be critically dependent on a few of those conditions, such as the rules of the automata. The morphology and dynamics of other emergent structures in this game are essentially unpredictable and divergent and can be thought of as sources of evolutionary variety within the game. See Poundstone 1985 for an account of Conway's game from a universal perspective.

In real-world systems, such as individual living organisms, we can observe that the features of two genetically identical twins that look the same are (in theory) predictable, convergent, and developmental. The morphological, dynamical, and functional features that are stochastically different, which include their fingerprints, brain wiring, organ microarchitecture, and many (not all) of their ideas and behaviors, are unpredictable, variety-generating (within bounds), and "evolutionary," in an evo devo model. Most dynamical processes in two identical twins, when we observe them at the molecular scale, appear stochastic and evolutionary. It is only when we look at the twins from across the room (a great increase in observational space and time, from the molecular perspective) that we see a subset of developmental similarities. We will discuss this as the 95/5 rule in the next section and then consider how it may apply to the universe as an autopoetic system.

In living systems, Selection always appears to involve a majority of "tree-like" evolutionary processes (think of Darwin's "tree of life") driven by Variation, and a minority of "funnel-like" developmental processes (any cyclically stable attractors in phenospace) driven by Convergence. From the perspective of information theory, the first process generates new information, and the latter conserves old information, expressed in a prior cycle. These two informational and dynamical processes appear to work both cooperatively and competitively with each other, in service to adaptation. Consider how Replicating organisms are sometimes driven to variation, and sometimes to convergence in both their systems and subsystems. Inheritance units (seeds, genes) sometimes duplicate (think of gene duplication) and vary, and sometimes converge (with gene loss). Selection in the environment sometimes favors creation of phenotypic diversity, and sometimes favors convergence to a particular dominant phenotype. In the VCRIS model, evo and devo (variational and convergent) replication and inheritance under selection are the root source of adapted order.

Perhaps most promisingly, from my perspective, the RIS terms at the center of the model allow us to think of information, learning, and intelligence, all processes that may be central to the maintenance of autopoesis, from three separate systems perspectives, that of the replicator (organism, as an autocatalytic system), the inheritance system (informational parameters that guide variation and convergence), and the selective environment (environmental conditions). This seems particularly appropriate, and a clue toward a better autopoetic information theory, as all sufficiently complex organisms (such as any metazoans with culture) appear to store the fruits of their learning and intelligence in these three, partially decomposable systems. In other words, we can say that adapted intelligence (encoded information) in any evo devo system always appears opportunistically partitioned between three complex actors, Seed (inherited parameters), Organism (autocatalytic replicator), and selective Environment (SOE partitioning). Intelligence is never resident in only one of these actors. It always straddles all three (Smart 2008).

For a basic example of Environmental intelligence partitioning, genes use historically metastable features of the local environment to reliably guide the evolving and developing organism to its future destinations. Much information for embryo construction is not specified in the genome, but in the replication-stable features of the environment. For a more complex example, metazoans externalize their intelligence in “niche construction” of their local environment, to make it more co-adapted (Odling-Smee et al. 2003). This process is also called “stigmergy” by scholars (Heylighen 2008, 2016). Niche construction/stigmergy is a key informational process that appears to grow with the complexity of the replicator. It presumably exerts selective pressure toward certain forms of variation and of convergence, in ways not yet well characterized in evo devo theory.

Consider also that environments may also replicate, on some higher systems level, just as organisms and seeds replicate. This happens, for example, when we replicate an urban architecture or idea-complex (like capitalism or democracy) in global society, when stars replicate, when continents drift apart, or if our universe itself replicates. In this model, our selective environment is much more similar to an organism, one fated to produce a new seed or seeds in special high-complexity locales, than is commonly understood in complexity theory.

We may also use the VCRIS model to gain a new perspective on another long-used term in the complexity literature, self-organization. Self-organization is typically defined as the emergence of "spontaneous" order from a previously apparently disordered system. When a complexity theorist uses the term self-organization, they are calling attention to poorly understood, partly hidden ordering processes. In the VCRIS model, these ordering processes must be partly evolutionary (via inherited mechanisms of variation) but largely developmental (via inherited mechanisms of convergence). Both ordering processes interact to produce an autocatalytic life cycle (replication), and both appear to require inheritance factors that are selected upon. These five VCRIS processes, then, are the key ones we must strive to better understand in any autopoetic system.

To understand self-organization, we must find the hidden evolutionary (to some extent) and developmental (to a major extent) dynamics that have been tuned into the initial and boundary conditions of the replicating system, as a result of selection that occurred upon that system in previous autopoetic cycles. For example, when we randomly cut up viral DNA and proteins in a petri dish, and place those molecular fragments in another dish, many fragments will appear to "self-organize" (spontaneously form structure), at a rate much greater than chance. They do so because those molecules have become finely tuned, under prior selection, to use physically and informationally metastable features of the universal environment to produce both contingent evolutionary variety and deterministic developmental order (self-assembly), using processes of both bottom-up and top-down causation. In an evo devo universe, such classical self-organization discoveries as Rayleigh–Bénard convection and the Belousov–Zhabotinsky reaction can be called previously hidden, now understood forms of evolutionary developmental ordering. Once we have the appropriate model, such order is no longer spontaneous but becomes predictable, in a broad range of environmental conditions.

If our universe is an autopoetic system, it too must have many such hidden evolutionary and developmental ordering processes at work as well, most of which we do not yet model well. Complexity theorists who argue that self-organization under far-from equilibrium conditions is as much a source of biological complexity as genetic variation and natural selection can be classified as universal evolutionary developmentalists, though they may not self-describe with this term. See Jantsch (1980), Haken (1984), and Kauffman (1993) for three promising yet still early theoretical efforts exploring self-organization from a universal evo devo frame.

Self-organization is thus a helpful term to remind us that both evolutionary and developmental processes are occurring in any autopoetic system, and I will use it in that sense in this article. At the same time, it should be most helpful to use the full set of VCRIS terms as our models improve, as we should be able to model replication, inheritance, and selection in evolutionary and developmental terms. Again, in the VCRIS model of selection in autopoetic systems, adaptive processes are not called "evolutionary" but rather "evolutionary developmental" or evo devo, to remind us they are always a balance between diverging and converging dynamical processes. This small change in terms helps to correct a bias of standard models, which ignore or minimize convergence, particularly at the level of the universal environment. Even today, the topic of convergent evolution (apparent planetary, biogeographic, and ecosystem development) remains controversial and understudied in evolutionary (developmental) biology. This neglect is no longer acceptable, in my view.

Many biologists today would argue that macroevolutionary dynamics are overwhelmingly contingent, diversity generating, and unpredictable. So it is a small change in definition for us to restrict the term evolution to "only" such processes, within any autopoetic system. Many evolutionary biologists might not like that restriction, but from my perspective, evolutionary biology today offers a view of life and selection that is dangerously incomplete. It has long neglected the physical and informational roles of development in macroevolutionary change, and developmental processes in the selective environment. Fortunately, evo-devo biology is rehabilitating development as a process in living systems. We can hope this will lead us to better see development in the universe as well.

Finally, if autopoesis turns out to be the most efficient and effective way to generate advanced complexity that is both intelligent and stable to time and change, as I presently believe but cannot prove, then it seems most parsimonious to expect both that our future AI must be autopoetic (evo devo) in nature, and that our universe itself is an autopoetic system. Our reality may be, as Rod Swenson (1992) argues, autopoetic "turtles all the way down."

5 Evolutionary Development in Organisms: The 95/5 Rule

Since the mid-1990s, the interdisciplinary field of evolutionary developmental, or "evo-devo" biology has emerged to explore the relationship between evolutionary and developmental processes at the scale levels of single-celled and multicellular organisms (Steele 1981, 1998; Jablonka and Lamb 1995, 2005; Raff 1996; Sanderson and Hufford 1996; Arthur 2000; Wilkins 2001; Hall 2003; Müller and Newman 2003; Verhulst 2003; West-Eberhard 2003; Schlosser and Wagner 2004; Carroll 2005; Callebaut and Rasskin-Gutman 2005). Evo-devo biology includes such issues as:

- How developmental processes evolve
- The developmental basis for homology (similarity of form in species with a common ancestor)
- The process of homoplasy (convergent evolution of form and function in species with unique ancestors)
- The roles of modularity and path dependency in evolutionary and developmental process
- How the environment impacts evolutionary and developmental process.

Conceptual and technical advances in scientific disciplines including comparative phylogenetics, morphology and morphometrics, and statistics are allowing better insights into the evolutionary relationships among organisms, and inferences about how developmental processes influence those relationships. The best work in evo-devo recognizes that natural selection is a net subtractive process. Natural selection generates increasing physical diversity, as seen in ever-growing evolutionary "trees," but, at the same time, an even greater reduction in potential physical diversity (Johnson 2011).

The fundamental role of evolution can be hypothesized as cumulative mechanisms that generate experimental ("good bet") types of diversity, to improve the odds of survival under environmental selection. Evolutionary systems harness stochasticity in an increasingly information-driven and intelligent way as organic complexity grows, but evolutionary innovation itself is largely unpredictable (Shapiro 2011; Noble 2017). Living systems continually sense their internal states and environment, and they react to catastrophe and stress with bursts of such poorly predictable, information-driven innovation, a pattern some evolutionary biologists call punctuated equilibrium (Eldredge and Gould 1972).

The fundamental role of development can be hypothesized as cumulative mechanisms that conserve and execute a small subset of (in-principle) predictable processes that have worked in the past to guarantee replication, under a range of chaotic internal and external environmental conditions. Developmental systems encode future-predictive probabilistic models of themselves and their environment, models which we assume follow the rules of Bayesian probability in nervous systems and presumably even in single-celled organisms. Developmental prediction (a convergent form of "intelligence") is generated from special initial conditions (developmental genes), tuned via informational constancies that exist in genes, developing organisms, and the environment.

The theory of facilitated variation (Gerhart and Kirschner 2005, 2007), in which the genetic processes in living systems are assumed to sort into two groups, a conserved core, which regulate critical elements of development and physiology, and a set of changing genetic elements, whose variation is "facilitated" by the conserved core, presumably in ways that both reduce the lethality of experimental change and increase the utility of genetic variation ("experiments") subsequently retained by populations, is a model consistent with this view. In evo devo language, the conserved core are conserved developmental genetic, allelic, and epigenetic processes, and evolutionary genetic processes are those that facilitate genetic, allelic, and epigenetic variation within and across generations. Such processes presumably act in tension with and opposition to each other in very fundamental informational and dynamical ways.

In this model, natural selection can be argued to be a composite of two more fundamental kinds of selection. Evolutionary selection biases the system toward potentially useful, intelligence-guided innovation and disorder when needed, and developmental selection biases the system toward convergences and order that have historically allowed complexity conservation and replication. In this view, we must see both of these selective and often opposing processes, apparently at work at many scales in every system that replicates, to truly understand biological change. For example, we should be able to identify both structure- or function-divergent and structure- or function-convergent classes of gene flow operating between species in the terrestrial biosphere, via such processes as genetic drift and horizontal (or lateral) gene transfer. Such transfer is well documented in Prokarya, and has been greatly facilitated by viruses in Eukarya (Zimmer 2015).

One clarifying feature features of developmental selection is that it is always critically dependent on a small subset of control parameters (in biology, developmental genes and other regulatory molecules). While about half of metazoan genes are expressed in such processes as organ development, less than 20% of these (thus less than 10% of all genes) are substantially regulated during expression (Yi et al. 2010). A further subset of our genome, roughly 5% of DNA in human, mouse, and rat, is highly conserved across these and other metazoan species. This 5% of our genome typically cannot be changed without stopping, or causing major deleterious effects to, processes of development. The majority of this highly conserved DNA, 3.5% of our genome, is noncoding, yet presumably also constrains functional expression (Wagman and Stephens 2004). A subset of this conserved DNA is sometimes referred to as the developmental genetic toolkit (DGT), or less accurately, the evo-devo gene toolkit. These genes include the *Hox* genes which determine animal body plans, and they often involve initial symmetry breaking choices in spatial, dynamical, and informational form and function that commit the organism to a particular developmental path. Thus a subset of all metazoan genomes have become very finely tuned, over many past replications, for the production of complex, path-dependent modularity, hierarchy, and life cycle. Presumably, the other 95% of these genomes can change and generate diversity without such immediately deleterious effects.

Thus all genomes can be categorized into two groups, of conserved and non-conserved genes, and we can propose that all highly conserved genes which are also highly tuned (highly sensitive to change, with deleterious effect) are the core constraints on development itself. I call this observation the 95/5 rule, and have found early evidence for it in replicative systems at a wide variety of scales (Smart 2008). The rule proposes that some small subset of developmental parameters are always top-down causal, involving essentially one-way information flow (in this case, developmental genes to organism). They can no longer be easily changed, they can only be added to, as organisms get more complex. The remainder of the genome can be considered evolutionary, whether it controls evolutionary or developmental process, as all of those genes can be altered by two-way information flow with the environment, with feedback. But per the 95/5 rule, a small and highly tuned set of top-down constraints must always exist, in any evo devo system.

There are a variety of levels of biological hierarchy at which evo devo concepts can be applied, and evo-devo biologists believe developmental processes and genes must themselves act to constrain evolutionary processes, in ways not yet understood by traditional evolutionary theory (notably the neo-Darwinian Modern Synthesis), and that both evolutionary diversity and developmental constraint are important to understanding long range “macrobiological” change (Pigliucci 2007; Pigliucci and Müller 2010). Evo-devo genetics and epigenetics are rapidly improving fields, and they promise to improve our understanding of such complex yet central topics as biological constraint, adaptation, intelligence, and convergent evolution.

6 The Riddle of Development and the Challenge to Cosmology

There is nothing in science more magnificent and more mysterious than biological development, including genetic, embryonic, organismic, and psychological development. How is it that developing organisms can reliably converge on far-future form and function (from the molecular perspective), under chaotic and variable environmental conditions? How is this done with just a small percentage of highly conserved developmental genes? Development employs stochastic, contingent, and selectionist processes, presumably ranging from quantum to macroscopic scales, in service to statistically deterministic, modular, hierarchical and cyclic emergent change, from conception to organism, and from organism to reproduction, senescence, and death (recycling) (Salthe 2010). Our mathematical models of development are incomplete today, but they continue to make progress. Our models of evolution, of random genetic reassortment and selection in populations, are much more advanced.

Development also involves teleology, or the assumption of goal-driven, end-seeking behavior, including successful replication. For these and other reasons, most scientists have focused on the idea that our universe may be evolving, while ignoring the idea that it may also be developing. This oversight, more than any other, has motivated the creation of the EDU community. The great challenge to cosmology today is to change this state of affairs, to learn from biology to better understand universal change.

Biologically inspired hypotheses for cosmological change offer us a number of predictive models of the dynamics and large-scale properties of the universe. This is necessary for establishing the potential value of ED as an explanatory approach, but only falsifiable predictions can establish (or negate) its legitimacy. Unfortunately, falsifiability is not easy in our present level of cosmological understanding. Whenever these hypotheses appear (and we shall see some below) we may need to investigate many details before concluding that that the hypothesis is impossible or unfeasible. In such circumstances, it is best not to jump to negative conclusions on the basis of the greater familiarity that science has had to date with mechanistic, bottom-up reductionism.

Ever since Plato, scholars have occasionally compared our universe, in some ways, to a living organism. If evo devo models are correct, this organicist philosophy may be true in part, but we should also expect this analogy to be both overly and poorly applied at first. Fortunately, the rise of bio-inspired design, and the recent successes of bio-inspired approaches in deep machine learning, are showing the value of generalizing organic structure and function to other substrates. As our understanding of biological development grows, and we gain the ability to predict developmental outcomes in embryogenesis via partial dependence on top-down parameters like developmental genes, our understanding of causality will improve, and so too will our cosmology. Fortunately, a subset of scholars continue to call for more holistic and top-down approaches to understanding universal change (e.g., Vidal 2010; Ellis et al. 2012; Ellis 2015; Adams et al. 2016). The great challenge we have is in learning how to blend our best top-down and bottom-up perspectives.

Predictability, convergence, and constraint help explain our universe, but these concepts only take us so far. Consider symmetry. Discovering hidden symmetries underlying physical reality has been tremendously helpful in building our standard model of physics, allowing us to understand conservation laws, Maxwell's equations, the electroweak interaction, and predict fundamental particles like the charm quark. Exploration of the symmetries of very high dimensional shapes, like the Lie groups, may uncover a constraining relationship between our universe's forces and particles. But our attempts to use supersymmetry to arrive at a single "theory of everything" for our universe, and to make verifiable predictions in our particle accelerators, have stalled. I do not expect such a single theory exists, and would predict that supersymmetry, or any other fully top-down, constraint-based model, will never be enough to explain reality.

Our universe also seems to use unpredictability, divergence, and freedom just as fundamentally. Besides quantum theory, two other useful physical theories, eternal inflation and string theory, each offer a mathematics of diversity and unpredictability, in which our universe is just one in a multiverse of possible universes, and our fundamental parameters alone cannot fully specify all the features of this universe. Some scholars propose that these new multiverse models imply we live in an "accidental" universe (Lightman 2011), and that our ability to understand our universe in terms of fundamental principles, at a level below this essential randomness, is, like a fully deterministic (non-statistical) understanding of quantum theory, an objective we will never achieve.

I am sympathetic to both of these views, and expect each will continue to make progress, while each alone will remain incomplete. If we live in an evo devo universe, where universal dynamics and informatics have proceeded something like biology has, from simple to more complex, over many past cycles, then something like the 95/5 rule should apply. Our universe will increasingly be understood as both "accidental" and "purposeful." While the vast majority of our universe's mathematics will have this random, accidental, and evolutionary looking nature, we will also continue to discover a growing subset of top-down, constraining processes that guide our universes critical processes of development, including complexity and intelligence production, conservation, and replication. Some combination of environmental selection (for adaptation) and self-selection (for intelligence) should apply. We need to get smart enough to see both classes of process, and ask how they each relate to and support the growth of useful (intelligent, adapted) complexity.

Like living systems, our universe broadly exhibits both stochastic and deterministic components, in all historical epochs and at all levels of scale (Miller 1978). It has a definite birth and it is inevitably senescing toward heat death. The idea that we live in an "evo devo universe," one that has self-organized over past replications both to generate multilocal evolutionary variation (preselected diversity) and to convergently develop and pass to future generations selected aspects of its accumulated complexity ("intelligence"), is an obvious hypothesis. Living systems harness stochastic evolutionary processes to produce novel developments, especially under stress, in a variety of systems and scales (Noble 2017). If our universe is an adaptive replicator, it makes sense that it would do the same. Yet very few cosmologists or physicists, even in the community that theorizes universal replication and the multiverse, have entertained the idea that our universe may be both evolving and developing (engaging in both evolutionary innovation and goal-driven, teleological, directional change and a replicative life cycle).

There is a reasonable frequency of discussion, in the cosmology and astrophysics literature, of the idea of universal evolution. But none of it takes an evo devo approach. We find plenty of random, Monte Carlo models of change, applied to our universe's initial conditions (e.g., various chaotic inflationary multiverse models; Linde 1992), but no models in which adaptive complexity and modeling intelligence emerge via evolutionary development in replicating universes in the multiverse, just as it does in all living replicators, and in several nonliving ones, such as hierarchical prebiotic chemistries on the path to RNA and hierarchical populations of increasingly chemically complex stars. Even our best current models of universal replication, like Lee Smolin's cosmological natural selection, do not yet use the concept of universal development, or refer to development literature, or to any theories of intelligence. Yet intelligence and its causal implications are an emergent property of all organic replicators, and if our universe is a replicator it is reasonable to expect universal intelligence must be accounted for in our future cosmology, as we will describe.

Organisms are evolutionary, and most of their genes recombine and change to generate diversity, but they are also developmental, and a small subset of their DNA, on the order of 5% per the 95/5 rule, cannot be changed without disastrous effects on development. As previously mentioned, this DNA has become very finely tuned, over many past replications, for the production of complex, path-dependent modularity, hierarchy, and life cycle in all complex metazoans.

In the same fashion, a handful of our universe's fundamental parameters appear breathtakingly finely tuned, in their mathematical values, for producing stable, long-lived, complex universes (Barrow and Tipler 1986; Rees 1999; Smolin 2006, 2012). If our universe is an adaptive replicator, under some sort of selection (either self-selection or environmental selection), the most parsimonious explanation for our universe's incredible developmental fine-tuning would be past universal replication, with both optimization and path dependency of developmental parameters (conserved inheritance) aiding in universal complexity survival and adaptation. Virtually all known or proposed intrauniversal complex adaptive systems are replicators, with the exception of galaxies, which presumably replicate as dependents on their universes (Smart 2008), so it is conceptually the simplest to infer that the universe is also such a system, in my view.

In living systems, developed properties like intelligence, immunity, and morality strongly alter previously locally contingent environmental selection processes toward organism improvement and survival. See Corning 2018 for a nuanced argument that synergy (cooperative competition, interdependence) is central to adaptive selection in all intelligent systems. If our universe is a replicating system under selection, it is a reasonable hypothesis that aspects of its internal adaptive complexity, intelligence, immunity, and morality may not only be evolutionary (stochastic, unpredictable) but also developmental (fine-tuned, predictable) as well. Yet at present, the scientists exploring the fine-tuned universe problem presently do not consider explanations in terms of universal development. Instead, we find fine-tuning research disproportionately dominated by intelligent design creationists championing the idea of fine-tuning as "evidence for God," leading to much confusion in professional and lay circles.

Perhaps as a result, the field remains professionally controversial for orthodox science, and a minority of astrophysicists, seeking to debunk theists, argue that fine-tuning, beyond the weak anthropic principle (observer-selection effects) doesn't exist (Adams 2008; Stenger 2011; Carroll 2016). But since the anthropic principle was first clearly articulated in cosmology (Carter 1974), another community of scientists have offered their own reasonable evidence that such tuning appears baked into our standard model of physics and empirically observed cosmology. In recent decades, fine-tuning explanations are commonly done via appeal to the multiverse. Among multiverse models, the hypothesis of universal evolutionary development offers a naturalistic explanation for fine-tuning that is homologous to biological fine-tuning. It deserves elucidation and critique.

The primary bias that exists in our cosmological models today is not observer selection bias, which is real but overrated. The primary bias at present is our failure to consider the concept of universal development, the idea that our universe's special initial conditions and stunning internal complexity are likely self-organized, via evolutionary development, just as our initial conditions and complexity have self-organized in all living systems. If our universe is a replicator, then evo devo self-organization is the most parsimonious explanation for the surprising levels of fine-tuning, massive parallelism, and fitness for life we find in our universe, not randomness alone, and not "design." See the fine-tuned universe hypothesis – early evidence for universal evolutionary development below for further discussion.

In the meantime, our leading theories of universal change are presently missing the concept of evolutionary development. If our universe is an evo devo system, then cosmologists, astrophysicists, geochemists, planetary scientists, astrobiologists, information theorists, philosophers, Big Historians, anthropologists, sociologists, and scholars of long-range biological, social, and technological change will have to update their models of the future. For more on this perspective, see Smart (2015).

7 Do We Live in an Evo Devo Universe? The EDU Hypothesis

All replicating complex systems can be viewed from two fundamentally different perspectives. When we look at the system up close, whether it is a star, a prebiotic chemistry, a cell, or an organism, we see much that is locally unpredictable. Yet when we observe the same system either at a larger scale, or over a longer time

Fig. 2 Acorn and oak tree. (Source: Pixabay)

frame, long enough to see its replication cycle, we see much about it that is predictable – even when we don't yet know any of the math or causal forces behind its predictability.

Think of an acorn. Once you've seen one acorn grow into an oak tree (Fig. 2), you learn that the shape of the acorn seed tells you that it will make an oak tree, with its characteristic leaves, morphology, and behavioral proclivities. Once you've planted more than one acorn, you know, in advance, that most of the structural and molecular details of each oak tree will remain contingent, "random," and unpredictable. But you also know much about its future that is predictable. That predictability, in biology, is called development. The unpredictable, diversity-generating parts we can call evolution, in an evo devo model.

If our universe is a replicating system, it is very much like an oak tree, moving from a highly defined initial seed, to a very flexible, undifferentiated, and totipotent embryo, as we see in our universe's pregalactic era, then to an increasingly specified and constrained set of outcomes, like the increasingly terminally differentiated structure of the oak tree, or the terminally differentiated tissue types that emerge in a developing embryo. Some scholars have represented the latter as a "tree of differentiation" (Fig. 3), a developmental counterpart to the evolutionary "tree of

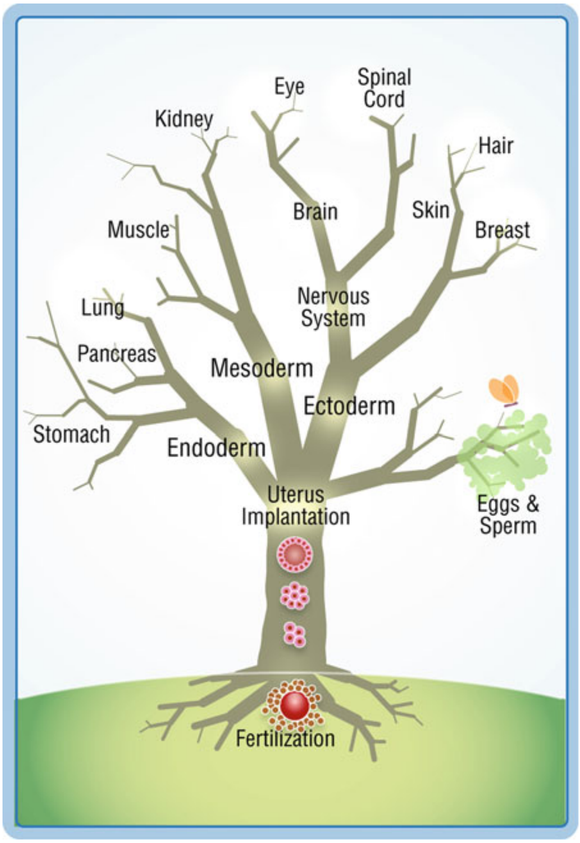


Fig. 3 A developmental "tree of differentiation." (Source: ASF)

life." The more we learn about the shape of the seed that created our universe, its nurturing environment (multiverse), and the "organism" itself, the more we'll know about both our evolutionary futures – what will stay unpredictable, and about our developmental future – what predictable and constraining "portals" and "terminal destinies" lie ahead, both for us and for all intelligent life.

The massive scale and isotropy (parallelism) of our particular universe, and its severe migration and communication constraints, can also be suspected, given the presumably sharply limited complexity of each local intelligence (constrained by physical law), to have been self-organized by the universe to maximize the local evolutionary variety of each intelligence prior to contact (Smart 2008). If universal intelligence plays a nonrandom role in universal replication, as it does in living systems, a bio-inspired case for the emergence of our kind of massively parallel yet apparently intelligence-compartmentalized universe can be made, as well as the prediction that a mechanism must exist for all end-of-universe intelligences to eventually be able to compare and contrast their computationally incomplete yet usefully locally unique models of reality. If future intelligences can survive a black hole transition, a number of arguments can be made that black holes themselves may uniquely offer such a merger and selection mechanism, in what I call the transcension hypothesis for universal intelligence (Crane 1994; Harrison 1995; Smart 2008, 2012; Vidal 2008, 2016).

Our intelligence can take these multiscale and macrotemporal views on our reality, even as we are physically stuck in one small corner of our universe. All the universe's most complex bits are curiously isolated, by astronomical distances, and thus each is constrained to follow its own unique evolutionary path toward common developmental destinations. When viewed from a cosmic perspective, we can also see that our computers are rapidly becoming the new leading local intelligence on our planet. They may soon (perhaps even this century) exceed us in their general adaptiveness, immunity, and intelligence. Such intelligences may be immune to environmental catastrophe, able to exist in near space, fully independent of our planet's nurturing environment. Using nuclear fusion technology, they would not even require our sun for energy. Using quantum computation, these intelligences might even function best in the cold environment of space. This accelerating transition to a new level of hierarchical complexity (and presumably, consciousness) may predictably occur on all planets that harbor intelligent biological life.

The evo devo universe (EDU) hypothesis proposes that our universe has two fundamental drives, to evolve (vary, diverge, create, experiment) and to develop (converge on a predictable, information-conservative hierarchy and life cycle). In the VCRIS model, the adaptive intelligence of any replicating complex system lives in, and is opportunistically partitioned between, at least three physical and informational actors: the initiating Seed, the Organism, and the selective Environment.

The cosmological natural selection (CNS) hypothesis (Smolin 1992, 1997, 2004), in which our universe replicates via black holes, with random reassortment of fundamental cosmological parameters at each replication, is one such evo devo model. In CNS, black holes can be considered the "seeds," the universe the "organism," and the multiverse the "environment." CNS remains controversial (see Vaas 1998 for one review of questions to be resolved). See Gardner and Conlon (2013) for an evolutionary biological approach to CNS using the Price equation to model selection for black hole replication. In my view, CNS as a model for an evo devo universe is an auspicious start, but has at least two shortcomings (Smart 2008).

First, CNS currently predicts that universes which replicate via black holes would select for a maximum of progeny (black holes), when real biological replicators always balance replication fecundity with other adaptive goals, including the resource cost to add somatic complexity to the organism (the universe, in this case). In a biologically analogous evo devo model, the qualities of the soma (universe), of the seeds (black holes) and of the environment itself (multiverse) can all be modified, both randomly (evo) and predictably (devo), and increasingly intelligently in more complex replicators, to make the system more adaptive. Some critics of CNS who state that our universe doesn't appear to maximize black hole production have assumed this insight makes the theory invalid, when in fact, any adaptive theory would rarely argue for black hole maximization.

Second, CNS has a very incomplete selection function, which does not yet account for intelligence (modeling ability) at any level. CNS assumes a random reassortment of our universe's fundamental parameters at the replication step, but this model is not appropriate even for the simplest biological replicators, as all living systems encode a kind of world modeling (intelligence) in both their evolutionary and developmental gene complexes. Genes reassort nonrandomly, based on developmental constraints. In higher complexity systems like human civilizations, consider the way ideas replicate in communities of brains. Idea replication is not random, but is increasingly selected by the intelligences responsible for modifying and passing them on. As internal intelligence grows in any replicator, it seems increasingly hard to neglect, in any good model of selection.

In my view, a theoretical framework we can call CNS with Intelligence (CNSI) (Smart 2008; Price 2017) will be necessary, if we are to use CNS to causally explain the roles of life and intelligence in our universe in coming years. Adding intelligence to our selection function allows us to consider to what extend the parameters of any seed have been *self-selected* (e.g., for greater capacity to simulate, and to engage in replication, independent of multiversal environment) by the growing evolutionary and developmental intelligence of the replicator itself. I will offer one such speculative model (Five goals of complex systems, Smart 2017b) for such self-selection later. At the same time, we must also consider if and how our universe's parameters have been environmentally selected, in some specific multiverse context, as we would do in a conventional Darwinian view of selection. For more on CNS and CNSI, see the EDU wiki page "cosmological natural selection (fecund universes)" (Wikipedia 2008, and Smart and Vidal 2008–2017).

If our universe has these general similarities to living systems, and is subject to selection, in some fashion, either self-selection or selection in the multiversal environment, we can predict that development at all system scales (organismic, ecological, biogeographic, cultural, technological, universal, etc.) will act as a constraint on evolution at all system scales. Likewise, we can expect that evolution, via preferential replicative selection, will continually and slowly change future development, again at all scales.

I expect a future information-centric theory of adaptation will find a number of evo devo processes (goals, values, drives, abilities) that are widely shared by complex systems. I can imagine, but not validate, one such speculative evo devo model, which we will see later. If we live in a noetic (information and intelligence-accumulating) universe, we may need such normative (goal and value-based) models to understand the way the growth of information and modeling abilities change complex environments. Modern hypotheses on how top-down causal information (Walker et al. 2017) and niche-constructing intelligence (Odling-Smee et al. 2003; Heylighen 2016; Noble and Noble 2017) constrain and direct innovation and selection in biology are an important start in this direction.

The better we understand the evo and devo roles for information-driven processes in bioadaptation (and today we often do not) the better we may understand their adaptive role for the universe as a replicator. When we discover and validate evolutionary process and structure, we can better describe innovation possibilities for complex systems in our universe. Likewise, when we find and model developmental process, we can predict or guess developmental constraints on those systems, and where they are striving to go. Most auspiciously for our moral and intellectual lives, we can better understand more of the evo and devo “purposes” or “telos” for ourselves, our societies, and the universe. We can recognize our natural drives to pursue both evolutionary goals (e.g., to create/innovate/experiment) and developmental goals (to conserve/sustain/discover), and seek to harness these two apparently fundamental processes to greater individual, organizational, and societal adaptiveness (Smart 2017a).

8 The Fine-Tuned Universe Hypothesis: Early Evidence for Universal ED

The fine-tuned universe hypothesis (Rees 1999, 2001) can be best understood as an important and early example of universal evolutionary development. In most organisms, you can change many genes and generate phenotypically different organisms, but they will still develop. We can call those “evolutionary” genes. But there is a subset of genes that are highly conserved in evolutionary history, and highly resistant to change. Nudge them just a bit, and you don’t get viable development.

In the same way, while our universe and multiverse simulation capacity are still emerging (Fig. 4), and our physical and informational theories are not yet complete, we know that among the known 26 or so fundamental parameters of our universe, most can be changed and simulations will still produce viable universes (Smolin 1997). We can call these the universe’s “evolutionary” parameters in its initiating “seed” or “genome” in an evo devo model. At the same time, there are a special subset of parameters that seem improbably precisely tuned (one, the cosmological constant, apparently even to 120 orders of magnitude), to work with the other finely tuned parameters to produce universes capable of rich internal complexity and longevity.

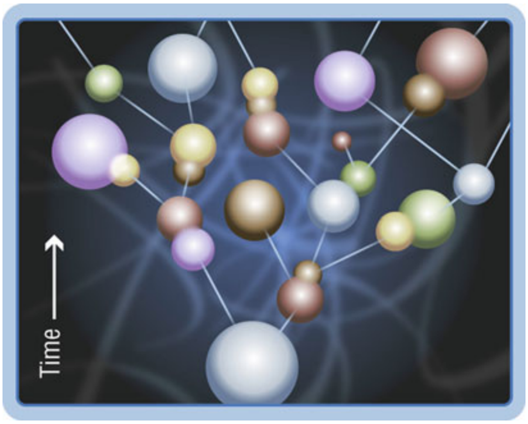


Fig. 4 Universe systematics must exhibit both evo and devo processes when viewed from a multiverse perspective, if our universe is an evo devo system. (Source: ASF)

When we nudge any of these precisely tuned parameters in our simulations, we don't get viable universes. We can call those "developmental" parameters, in an evo devo model. If our universe replicates, they seem homologous to the small subset of developmental genes in organisms. Edit any of those parameters and you never get viable organisms. They've been self-organized, over vast numbers of previous cycles, to work together to conserve the developmental forms, functions, hierarchies, and life cycle of the organism.

The proposition that our universe's laws are finely tuned for various evo devo outcomes can be made in a variety of ways. These claims can be considered various forms of the anthropic principle, and anthropic reasoning its field of inquiry. In its most useful variation, the anthropic principle is the idea that our universe's initial conditions and laws seem improbably biased toward the production of intelligent observers (Barrow and Tipler 1986). But there is a more fundamental bias that must be considered when evaluating fine-tuning models, the bias that results from observer-selection effects. Is the particular kind of physical and mathematical universe we live in logically necessary, if there are intelligent observers around to ask questions about it?

We will discuss anthropic selection effects in detail in a later section, but for now, let us make just one potentially useful observation. It does seem plausible that we must have a quantum universe in order to have a universe with observers. So that particular level of observer-selection bias may necessarily exist, at least. In perhaps the best-known "weirdness" of quantum physics in the standard ("Copenhagen") interpretation of the wave function, we observers alter physical reality (quantum states) by the manner in which we choose to observe them. But the apparent necessity for quantum physics in our observable universe in no way tells us that fine-tuning doesn't exist. If anything, it could point to necessity of some version of the "co-evolution" between universe and observers, perhaps as first sketched by John Archibald Wheeler (1977, 1988). The kind of quantum physics we have, and its relation to the rest of our physics, may be tuned for the necessary emergence of not just observers, but of intelligent observers, of "mind." Quantum physics doesn't presently integrate fully with other physical and mathematical features of our universe, such as the fundamental parameters, general relativity, symmetry, information, and meaning (whatever that is). So we don't know yet what our most fundamental universal theories are. As we don't yet have the ability to definitively answer such questions at present, the fine-tuning debates will continue, and continue to be productive.

Most physicists were strongly opposed to the teleological (purposeful, directional) idea of fine-tuning when the Barrow and Tipler book emerged in 1986. Today, many leading physicists, like Leonard Susskind (2006) and Steven Weinberg (2007), now argue that multiverse models offer us the simplest explanation (principle of parsimony) for the mathematically improbable levels of fine-tuning we find in several of our fundamental physical constants, and as a result, in the strengths and nature of the four forces in our Standard Model of Physics. So we have seen a shift of many leaders in the physics community toward multiverse explanations of fine-tuning, and thus an implicit recognition that our universe's apparent fine-tuning is a real problem that must be addressed.

In perhaps the most dramatic example presently known, the empirically observed value of our universe's cosmological constant appears to be tuned to one part in ten to the power of one hundred and twenty. In current models, any imperceptibly small change in that constant would lead to either near-immediate collapse or destructive inflation of the early universe. Likewise, Planck's constant, the gravitational constant, the neutron-proton mass difference, the strengths of electromagnetism, weak and strong nuclear forces, the masses of particles in early inflation after the Big Bang, and several other aspects of our physical universe appear fine-tuned for the production of long-lived universes that support high levels of emergent complexity.

Assuming other values of these constants are possible and would lead to alternative universes, there are many improbable transitions and architectures to explain, including the special subatomic physical resonance (quantum "fine-tuning") we call the triple-alpha process, famously predicted by astronomer Fred Hoyle (1954), which produces an abundance of carbon and oxygen in our particular universe. Recent calculations (Meissner 2013) continue to support the hypothesis of fine-tuning in the fundamental parameters of quantum chromodynamics and quantum electrodynamics for this fortuitous result to occur. Carbon, oxygen, and a handful of other elements (HNPS, and some metal cofactors) are developmental portals (unique gateways on the molecular phase space landscape) to redox organic chemistry, which is the structural and energetic foundation of life.

For another potential fine-tuning example, consider the way dark matter and a smattering of older Population II stars, often presenting as globular clusters, form an elliptical "halo" galactic superstructure, one that allows newer Population I stars to precipitate into elegant planar spiral and elliptical galaxies. The newer stars' rotation and metallicity gradients are apparently created and maintained by this halo dark matter distribution across vast ranges of space and time, giving the mature form of our most chemically complex galaxies the rough appearance of a complex biological development, like an ovarian follicle. Consider also the curiously scale-free and organic looking appearance of the large-scale structure of matter distribution in our universe. We also can suspect a variety of improbably life-generating conditions on the early cooling Earth, whose geochemistry may be catalytically optimized for life's emergence, including a predictable distribution of mineral cofactors able to catalyze the rTCA cycle on metal-rich planets around Population I stars, perhaps one of several critical preconditions for life (Smith and Morowitz 2016). Life appears to have arrived on Earth almost as soon as our crust sufficiently cooled.

Consider also the curiously biphasic nature of Earth's crust. It consists of a denser, continually recycling oceanic crust, which regulates CO_2 and other atmospheric parameters via plate tectonics, and a lighter, continually floating continental crust, which offers a stable nursery for the growth of land-based life. Stable continents, in turn, may be a developmental portal for the first emergence of complex social mimicry, language and tool use on any Earthlike, given the many physical advantages of air over water for such forms of social intelligence growth. We can also identify, with varying degrees of controversy, several life-stabilizing biogeohomoeostatic features (e.g., the Gaia hypothesis) in various atmospheric and ocean properties on our current Earth (Volk 2003). All these tunings and others may

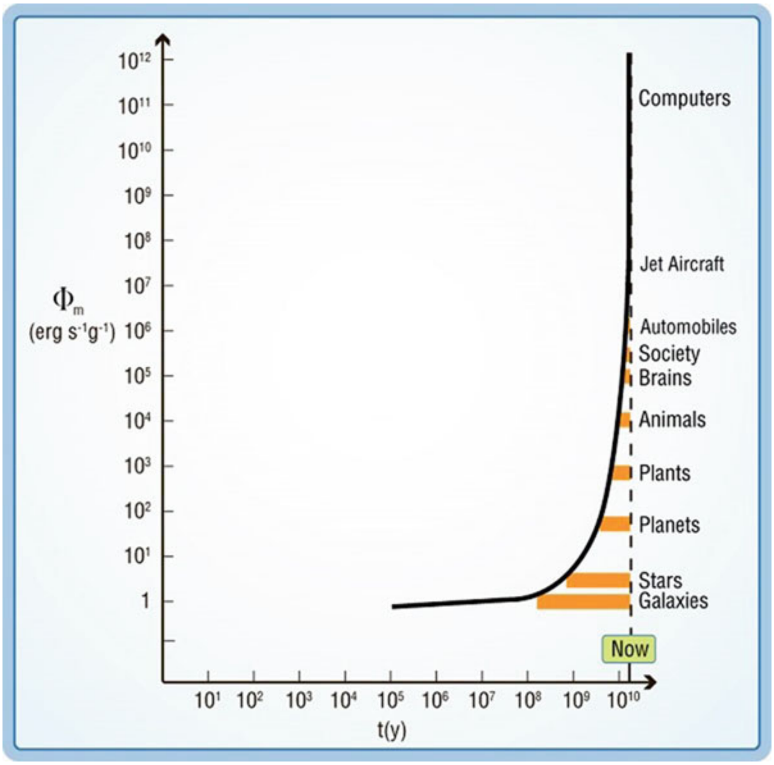


Fig. 5 Accelerating energy flow density control in our most complex (and rapidly learning) systems, in universal history. (Figure: Smart 2008. Data: Chaisson 2001)

be necessary for robust phase transitions to higher complexity, in special domains of space and time, allowing life to emerge, diversify, persist, and grow more complex and generally intelligent at its leading edge of evolutionary development.

But there's even more to explain, because not only does our universe support improbably high levels of emergent complexity and mind, it supports an even more improbable condition of continuously accelerating complexification in special environments, an acceleration that seems increasingly self-stabilizing under periodic, and often catalyzing, episodes of selective catastrophe (Smart 2000, 2008, 2012). Konrad Lorenz (1977) was an early advocate of the view that both energy transfer and information processing must work together to create the mode and tempo of biological change. If our universe is a replicator, we can expect both physical and informational causes are needed to explain its accelerative aspects as well. Chaisson (2001), Aunger (2007a, b), and others have proposed that it is the increasingly intelligent control of energy flow that drives structural-functional acceleration in our universe. Chaisson has estimated exponentially increasing energy flow density (free energy flow per gram or volume) in a special subset of complex adaptive systems over universal time (Fig. 5). Processes like galactic structure formation, stellar nucleosynthesis, and redox organic chemistry are themselves accelerative, in free energy flow measures, over previous complex systems, and each may be developmental portals (unique gateways) to further structural and functional complexification and intelligence growth. Life's accelerating complexification, in turn, has reliably produced a variety of social tool using species, and in humans, accelerating intelligence, immunity, and (though it is often debated) morality in recent millennia.

Curiously, our leading technology, digital computers, have a free energy density control rate that is now at least a millionfold faster than our biological neurons. This differential has grown exponentially over our "Moore's law" era of computing, and may grow by many additional orders of magnitude as we shift to future even more miniaturized, dense, and complex architectures and technologies including massive parallelism, single electron transistors and optical and quantum computing. I call this process of accelerating complexification "STEM compression" (Smart 2002, 2008, 2012), with "compression" referring to predictable growth in both physical and informational density and efficiency of critical spatial, temporal, energetic, and material (STEM) metabolic, effector, and thinking processes in our most dominantly adaptive systems over time. I consider energy flow density acceleration (e.g., Chaisson 2001) to be just one example of this apparent universal trend. Furthermore, now that our leading computers are using biologically inspired algorithms, and are developing increasingly general forms of intelligence, the adaptive goals they learn from their environment should be similarly accelerated, particularly if we can intelligently aid this apparently natural process. Later in this paper, I will propose five learnable goals (abilities, drives, ends, *telos*) that seem particularly universally adaptive and self-stabilizing for intelligent complex systems, if they are built from both evolutionary and developmental processes.

Empirically, this record of growing internal control and self-stabilization, and increasingly general adaptiveness of our most complex systems, which we can argue exists in the geophysical processes of Earth, in life, and in human civilization and our leading technology, seems unlikely in a randomly generated universe. Why hasn't our universe been far more disruptive to our general record of accelerating complexification? Some kind of developmental immunity may exist, tuned into every autopoetic system, including our universe's genes, soma, and environment, if our universe replicates with inherited characteristics, and if the acceleration of complexity and/or intelligence has had some past adaptive (selective) value.

In biological systems, we presumably need ever more sophisticated processes of immunity and (in more complex systems) of morality to stabilize growing individual intelligence. If we come to understand the inevitability of these processes in biological systems, we may come to understand them in replicating cultures and their technologies as well. Early developmentalist models of social immunity (stability) and morality (virtue) were championed by such 19th theorists as August Comte (1844) and Herbert Spencer (1864). In the twentieth century, the priest-paleontologist Pierre Teilhard de Chardin (1955) was perhaps their most famous advocate. In recent years, a few psychologists have offered us statistical arguments that both the average severity and average frequency of global social violence have substantially declined over human history, even as our potential for committing acts of violence at scale, via science and technology, has steadily grown (Pinker 2010). Causal models for this decline are still lacking, but Pinker is clearly arguing for both an evolutionary and developmental morality, and I believe our collective morality must grow predictably in nuance, force, and scale on all Earthlike planets if intelligence is to be stabilized as complexity accelerates.

We may also need to explain the impressive simplicity and comprehensibility of (most of) the mathematics that underlies nature. According to Leslie and Kuhn (2013), Gottfried Leibniz (1686) was perhaps the first to argue that while some mathematical equation could be found to fit any curve one might draw, the vast majority of the set of possible curves and equations would be exceedingly complex. Similarly to Leibniz, Vilenkin (2006) argued that one would expect "horrendously large and cumbersome" mathematics underlying a typical randomly derived universe in a multiverse ensemble. Yet the applied mathematics and physics that our minds can understand seems unreasonably effective for both scientific modeling and technological development (Wigner 1960).

We should also explain why our universe appears to use massive parallelism in its production of intelligent civilizations, and keeps them spatially separated for the majority of their evolutionary development. This is related to the Fermi paradox, namely, the observed absence of extraterrestrial intelligent beings and their artifacts in our past light cone (Lem 1971; Brin 1983; Webb 2015), even though we are likely to have emerged one to three billion years later than other Earthlike planets further in on our galactic habitable zone (Lineweaver et al. 2004). Although many hypotheses have been suggested for this curious absence, a large subset of these hypotheses require some form of fine-tuning (Ćirković 2009). It's as if the universe seeks to promote maximum evolutionary diversity in each civilization, while developmentally guiding each of them to a future in which their evolutionary learning might be instantly shared (Smart 2012). In short, there's a lot of apparent tuning that needs explaining.

Of course, there is also a lot of waste and danger and randomness in our universe as well. The vast majority of physical systems in our universe are simple and dead, not complex or adaptive, and there are catastrophes and danger everywhere. Intelligent life may be so hard to produce that our universe may need to evolve an entire galaxy of stars to develop one intelligent planet, on average. Evolution on Earth has seemed equally wasteful and violent, if we focus on all the species that have disappeared, rather than the intelligence, morality, and immunity that have survived and grown. In the far future, our Milky Way galaxy and Andromeda are destined to crash into each other, obliterating their beautiful spiral structures. Observing all this apparent waste, danger, and chaos has led many astrophysicists, including Neal DeGrasse Tyson (2006), to argue that fine-tuning doesn't exist.

Besides the desire to avoid the idea of a purposeful universe, and its historically theistic implications, scientists commonly seem to reject the idea of fine-tuning via two ways: misperception and mischaracterization. Let us consider each of these latter issues now, and propose an alternative description, the partially fine-tuned universe hypothesis, to try to reduce these problems. In my view, modeling fine-tuning in evo devo terms, in both physical and informational dimensions, and simulation testing it in both organisms and universes, is a core challenge science must address if it is to properly critique universal fine-tuning models.

9 The Partially Fine-Tuned Universe: Intelligence Is a Weak Selector, Not a Designer

Misperceptions can commonly cause us to reject fine-tuning, if we examine complex systems from inappropriate perspectives and scales. Consider a few examples:

- If it turns out to be true that it takes a galaxy of stellar "experiments" to produce just one (or a few) intelligent civilizations per galaxy, on average, that looks extremely wasteful and random (evolutionary) at the solar system scale, but simultaneously convergent and predictable (developmental) when we view the same process at the universal scale. A system that reliably produces hundreds of billions of something may very well be fine-tuned for that end.
- Intelligence in living systems looks very fragile and endangered (evolutionary) at the species scale, but very robust and accelerative (developmental) when viewed from the ecosystem or planetary scale (Heylighen 2008). For example, we can presume that very little of the "conserved core" of developmental genetic intelligence (Gerhart and Kirschner 2005) in Earth's species pool was eliminated by any of the major past catastrophes and extinctions that Earth's ecosystem has experienced. Instead, those catastrophes appear to have pruned back the evolutionary variety, created new exploration space, and catalyzed powerful new punctuations of evolutionary innovation (new phenotypic or sociotechnological morphology and function), while increasing immunity to further disruptions of the same type, shortly after each major catastrophe (K-T event, Permian extinction, Ice Ages, Toba event, many others). I call this process natural security, in general terms, or the catalytic catastrophe hypothesis, in relation to specific catastrophic events (Smart 2008, 2018). This hypothesis has been explored in biological systems by Gerhart and Kirschner 2005, Bhullar 2017 and many others. In social systems, economist Nick Taleb calls it antifragility (Taleb 2012). A general kind of immune learning appears to have operated throughout life's long history on our planet, as a central stabilizer of accelerating change.
- We can see chaos and randomness in galaxies colliding billions of years from now, as Tyson (2006) emphasizes. But if both low-intelligence universes (via CNS) and higher-intelligence civilizations (via the transcension hypothesis) use black holes (either to "randomly" produce new universes in CNS, or to do intelligence-guided replication in the transcension hypothesis), then future galaxy collisions long after many of the universe's black holes are created looks like normal aging and recycling of an evo devo system after it has aged past replicative maturity. All complex living systems are developmentally fated to senesce and be recycled. What looks fine-tuned, from that life cycle perspec-

tive, is that galaxies are stable for the billions of years necessary to produce complex life, and that mechanisms for universal replication and civilization communication (e.g., black holes, in the CNS and transcension hypotheses) have self-organized to be fecund in our universe.

- We can focus on how easy it is for planets to be outside a galactic or stellar habitable zone, and (in our solar system) become greenhouse hells like Venus, or lose their plate tectonics and atmospheres and dry up like Mars (taking an evolutionary perspective on planetary science), or we can consider the marvel of the apparently robust (developmental) existence of both habitable zones and Earthlike planets in our galaxy, and the unique features of Earthlikes as a cradle for life. Water-bearing Earthlike planets and yellow-white suns may be universally unique developmental portals (accelerative gateways) for life. Yellow-white stars like our Sun have their peak irradiance in the visible light range, optimal for water-based, photosynthetic life (Fig. 6). Hotter (blue) stars radiate much more in the dangerous, high energy range, and colder (red stars) radiate more in the infrared range, and with far lower specificity (their peak irradiance curve is much flatter). Our Sun's particular spectral type and our Earth's plentiful water vapor, water, and strong magnetic field efficiently shield life from radiation arriving outside our Sun's most useful range. Our plate tectonics, oceans and clouds, carbon and nitrogen cycles, and ecosystem itself stabilize many other features of our nurturing environment (Volk 2003). Why is our Sun-Earth energy transfer and geophysics so apparently co-adapted for the generation and buffering of life processes? Either of two kinds of observation selection would seem to be involved. Either the law of large numbers explains these local conditions (an evolutionary observer selection explanation), or some of the physical parameters in our universe have become biased toward the production and protection of life (an evolutionary developmental observer selection explanation). We must predict that the vast majority of planets would still be expected to be barren in an evo devo universe, as biological evolution always requires massive and "wasteful" stochastic variety to find new developmental optima. But we should also expect some improbable fine-tuning, in a small subset of parameters (the "5%"), for the robust emergence of a special class of life-supporting planets. Without taking this evo devo perspective, we might predict that self-aware life would typically emerge, on average, in both a much less efficient and less safe ecogeophysical environment.

Mischaracterizations, arising from incorrect models, can also cause us to reject fine-tuning. Perhaps the most common mischaracterization comes from focusing only on the evolutionary processes of adaptation, to the exclusion of the developmental processes. That can happen when we view a system from only one scale or perspective, as we have just described. But there is another mischaracterization that comes with the assumption, surprisingly common among fine-tuning critics, that fine-tuning must be extensive if intelligence is involved in universe replication. But as we'll argue now, if replicating universes are anything like replicating organisms, then extensive fine-tuning by any finite intelligence, whether internal or external to our universe, seems an unsupportable and non-naturalistic assumption.

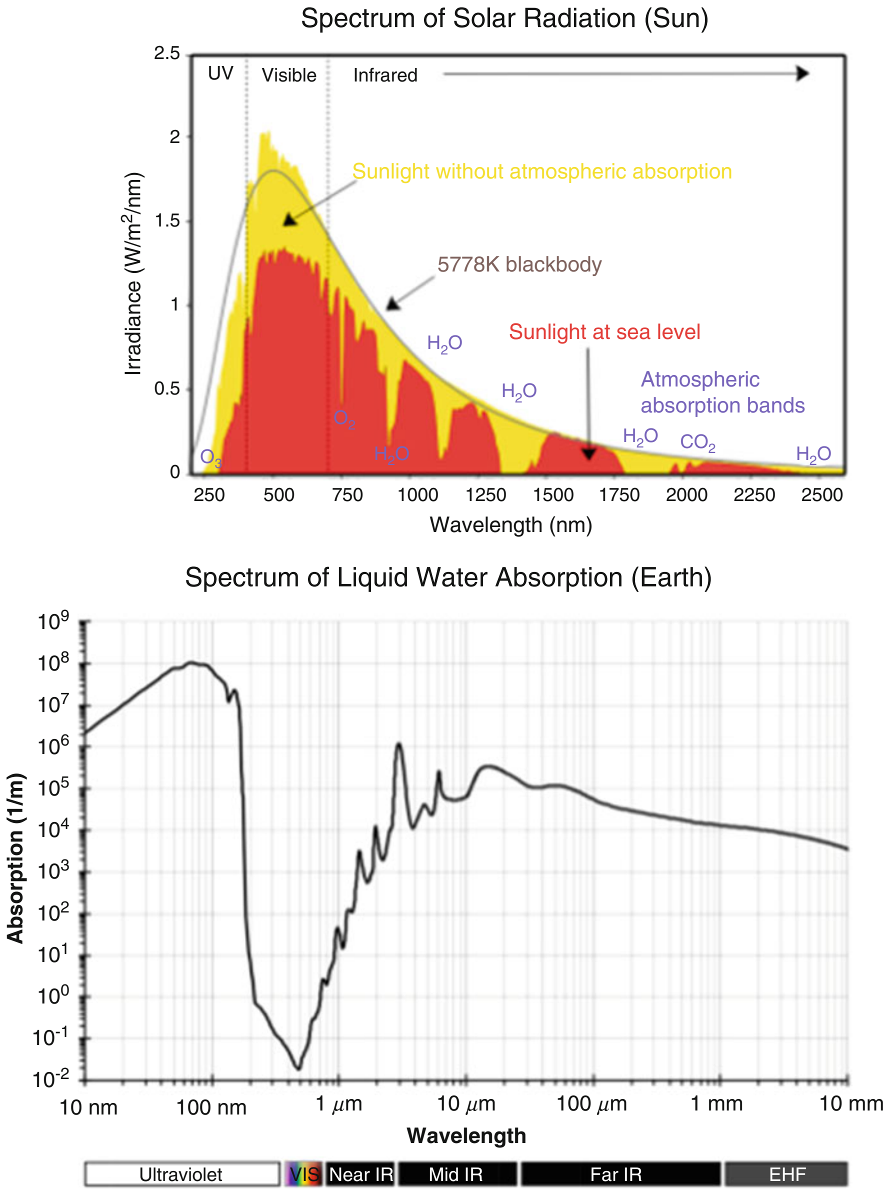


Fig. 6 Our highly energetically efficient Sun-Earth energy transmission and buffering system. The absorption spectrum for atmospheric water is very similar to liquid water, depicted. (Image sources: Wikipedia)

If we live in an evo devo universe, it can only be a partially fine-tuned universe, as the production of very limited and partial fine-tuning is how intelligence has always interacted with replicating, complex adaptive systems in biology, and with itself as a replicator. At best, intelligence always functions as an aid to selection in complex systems, and never as an omniscient or all-powerful designer.

What I call CNS with Intelligence (CNSI) (Smart 2008, 2012), or what my EDU colleague Clement Vidal calls Cosmological Artificial Selection (CAS) (Vidal 2008, 2010) is the hypothesis that intelligence, and its ability to simulate more and less adaptive futures, must play some useful role in the replication of universes. These hypotheses do not argue that intelligence can rationally design future universes, but rather that universes that self-organize intelligence are somehow more adaptive, in a nonrandom fashion, than universes that don't. In other words, some kind of multiversal selection occurs, in which intelligence, at both the very fundamental "genetic" level of self-replicating universal parameters, and at its higher levels, which includes conscious beings able to develop science and engineering, plays a nonrandomly beneficial role. That's the core hypothesis.

Even without the math it is easy to induce, in any biological replicator, that there is likely to be a nonrandom adaptive value to the emergence of general intelligence, of immunity (defensive intelligence), and of various forms of interdependence (collective intelligence, social morality, positive sum games), the latter starting with kin. We can imagine many circumstances when each of these computational systems that encode models of self-, others, and environment have adaptive value. These intelligence systems (general intelligence, immunity, and interdependence) may be present, in some fashion, in all complex adaptive systems. If we also consider evolutionary innovation and developmental sustainability as forms of intelligence, including the mix of stochastic and predictable genetic processes that generate our minds, we can imagine at least five potentially universal processes of intelligence, as we will discuss later (Smart 2017b). We can also identify multiple forms of intelligence (genetic, cellular, collective, neurological, societal, technological, etc.) on Earth, in living systems and their creations. Why should intelligence not also be a central property of the universe as a complex system, if it in fact is a self-replicator, existing in some larger environment (the multiverse)?

Yet we must also recognize that all simulations that any intelligence can do, either within our outside our universe, must be sharply finite and constrained, rapidly unable to predict most of the multivariate nonlinear dynamics and informatics of any complex adaptive system, the farther we extrapolate it to the future, or the more we include its evolutionary (vs. developmental) mechanisms. All such systems quickly become combinatorially explosive in their potential complexity, and all real intelligences are limited in their physical and computational complexity. With respect to the special case of logical-mathematical provability, this concept is as old as Gödel's Incompleteness Theorem (1931), and incompleteness seems intrinsic to the nature of informational complexity itself (Chaitin 1992; Calude and Jürgensen 2005).

The informational incompleteness of all intelligence, along with the inability to have perfect knowledge (simulation capacity) of initial conditions and all the relevant laws, may also explain why all complex actors with mind have "free will" (unpredictability to self), even under the most informationally ideal conditions. Philosophers since Lucas (1961) have tried to relate free will to Gödel's theorem, but such a relation starts by preassuming the physical universe conforms to Gödel's conditions for mathematics (it may not). At present, the full informational and physical nature of (our self-experienced) free will, and of conscious decision-making, remain a mystery to be solved.

Nevertheless, what we know and can guess so far about intelligence argues that any "design" that real intelligences can do, of their future selves and environments, will be highly limited. It does not seem defensible to imagine that end-of-universe or extrauniversal intelligences might be omniscient or Godlike, if they are physically real and we live in an evo devo universe, and they also seem unlikely to have the capacity to create "anything" out of "nothing." It is illuminating that "Why there is anything rather than nothing?" or, alternatively, "Why did anything begin?" is sometimes called "Question Zero" in physics. It may always remain a metaphysical question to real intelligences. Perhaps one of the most useful clues to its metaphysical nature is that the concept of nothing itself, just like the concept of infinity, while very useful in our mathematics, may be only an informational, not a physical concept (Aguirre 2016). Real intelligences may be forever stuck within the (evo devo) system (supporting universal environment) that they find themselves emerging within, a system that has its own physical and informational laws and constraints, only some of which are likely to be modifiable.

Discovering our universe's parameters and laws, and learning how to manipulate them to improve adaptation under selection, but always in finite and limited ways, seems to me to be the central benefit of intelligence in living systems. All living systems, while they possess some level of intelligence, still have many vestigial systems and errors and maladaptations in them which are beyond their control or even understanding. This should be true for our universe as well, if it is an evo devo system.

My current intuition for what an end-of-universe intelligence might be able to do with respect to "design" of future universes would be to alter some of the coupling constants which influence the developmental characteristics of the next universe, presumably to raise the probability that it will be complexity, life, and intelligence friendly. They might do this, for example, if intelligences in this universe can use their intelligence, and the laws of physics, to produce black holes which create other universes, and if the coupling constants can pass through the singularity of a black hole into another universe, as some physicists have postulated (Smolin 1992; Crane 1994). But notice that this kind of "universe engineering," though it is simulation-guided, is not the rationally engineered universes idea of James Gardner (2003, 2007). Gardner assumes that end of universe intelligences could change any of the constants, and might have extensive foreknowledge and control of the consequences of those changes. I would equate this view with intelligent design, which we shall discuss later. It sounds like non-naturalistic theology, not science.

Instead, tinkering with the values of our universe's coupling constants, in a way that might produce even more life- and intelligence-friendly universes, seems likely to be analogous to what human genetic “engineers” do today when we alter genes in “designer” organisms. What we are actually doing is making intelligence-guided engineering guesses at what will be more adaptive, and some of the most critical conserved genes are beyond our ability to tweak, without killing the organism. Our foreknowledge of these complex evolutionary systems must always be limited the further ahead we look. An honest assessment would be that we are not really “engineering” or “designing” new organisms, but are instead making our best experimental guesses, based on our finite simulation capacity and knowledge, working within the evo devo framework we have inherited, at what might be more adaptive.

Any hypothetical universe “design” would have to work the same way, in an evo devo universe. It would be a process of partly intelligence-guided selection, and partly unknowable experiment. It is not accurate to call such an undertaking by the word design. When we are talking about bottles and bridges, and other nonautonomous systems, it makes sense to use the word design. But the more complex adaptive and internally intelligent the system gets, the more the unpredictable evolutionary aspects of the system overwhelm the predictable developmental parts. Once we get to the “design” of things like living organisms, or new deep learning computers, or future universes, it makes more sense to call this process selection than it does design.

Gardening future universes using our own best science and intelligence would be directly analogous to the artificial selection we humans do on our domestic plants and animals, a process Darwin discussed at length in *Origin of Species* (1859). This is why Clement Vidal prefers the term Cosmological Artificial Selection to describe what this process might look like to any future intelligences that become competent enough to “engineer” intelligence-influenced black holes (if those are the seeds of new universes, per CNS), or to otherwise aid in the production of future universes.

In sum, there are at least five important levels of evo devo-related partial fine-tuning models that should be critiqued in future fine-tuning debates:

1. Level I. Our universe appears fine-tuned (self-organized) for the emergence of complex, long-lived universes and black holes (Smolin 1997; Rees 1999; Gardner and Conlon 2013).
2. Level II. Our universe appears fine-tuned for the fecund emergence of G-, K-, and M-class stars and biological life (Henderson 1913; Barrow et al. 2008; Lewis and Barnes 2016).
3. Level III. Our universe may be fine-tuned for the fecund and accelerating emergence of intelligent life (Piel 1972; Sagan 1977; Moravec 1979; Dick 1996; Kurzweil 1999, 2005).
4. Level IV. Our universe may be fine-tuned for the fecund emergence of intelligent life, which can then produce new universes (Crane 1994; Harrison 1995).
5. Level V. Our universe may be fine-tuned for the fecund and accelerating emergence of increasingly innovative, intelligent, immune, interdependent (defending evo devo values), and sustainable forms of complex life (Smart 2008, 2012, 2017b).

If this analogy between replicating organisms and universes holds up, models like Smolin's CNS, in some variation that also includes intelligence (CNSI), will continue to gain theoretical and empirical support. The better we understand and can simulate the operation of evolutionary and developmental parameters in living systems, the better we should be able to understand and simulate them in universes as well. Both look like finite and replicating systems, in an evo devo model.

10 The Riddle of Convergent Evolution: Limited Forms Most Beautiful

Convergent evolution is evidence or argument for physical attractors in the phase space of dynamical possibility which guide and constrain contingently adaptive evolutionary processes into statistically predictable future-specific structure or function, in a variety of physical and informational environments. When we look at evolutionary history, species morphology or function is often seen to converge to particular "archetypal forms and functions" in a variety of environments.

Such attractors have been called deep structure, guiding evolutionary process in predictable ways, regardless of local environmental differences. Organismic development depends on specific initial conditions (developmental genes in the "seed"), the emergence of hierarchies of modular structure and function in the unfolding organism, and persistent constancies (physical and chemical laws, stable biomes) in the environment. Likewise, some examples of convergent evolution may be best characterized as ecological, biogeographical, stellar-planetary, or universal evolutionary development (ED) if their emergence can be modeled, after adjusting for observer selection bias, to depend on specific universal initial conditions, emergent hierarchies, and environmental constancies.

A famous example of convergence is found in eyes, which appear to have evolved on Earth from different genetic lineages to work similarly (function as sensors for nervous systems) in all species possessing sight, and in the case of camera eyes, to also look very similar (form) in both vertebrate and invertebrate species, like humans and octopi (humans famously have a blind spot, however, as our eyes evolved via a different evolutionary developmental history than invertebrates. See Ogura et al. 2004). One can easily advance the argument that, in universes of our type, eyes, though first created by a process of evolutionary contingency, become a developmental archetype, an adaptive optimization for the great majority of multicellular species in Earthlike environments.

Presumably, the previously rapidly changing "evolutionary" gene groups that led to eye creation become part of an increasingly conserved "developmental" genetic toolkit for all eye-possessing species in environments where eyes are adaptive. Eventually, due to both path dependency and emergent hierarchies, some subset of these gene groups should be incapable of being changed without preventing development itself. Proving such genetic convergence arguments with evidence and theory is of course more difficult, yet it is a fertile area of investigation today.

Charles Darwin ended his foundational text on evolution, *On the Origin of Species* (Darwin 1859) with a well-known phrase, predicting "endless forms most beautiful" continuing to evolve. But as George McGhee describes in a well-titled book, *Convergent Evolution: Limited Forms Most Beautiful* (McGhee 2011), preexisting physical and informational optima in our particular universe mandate that only a very limited subset of forms and functions will ever emerge in biological evolution. Evolutionary development always grows morphological and functional diversity, and especially rapidly under stress, but developmental control and optimization makes it a net subtractive and constraining process, relative to its theoretical potential. Creative evolutionary process is continually reconverging to developmental optima, driven there by functional (environmental) and developmental (genetic) constraints. Better understanding and modeling convergence is one of the great and underappreciated challenges of modern evolutionary biology.

11 Less-Optimizing Convergence (LOC) Versus Optimizing Convergence (OC)

In our mostly chaotic, contingent, and deeply nonlinear universe, we can predict that the vast majority of examples of convergent evolution will not be driven by the evolving system's discovery of some hidden general optimization function in parameter space, like the eye archetype, but rather, the discovery of many less-valuable and less-permanent optima that do not lead to higher complexity, yet may still be required for the universe's evolutionary development. To understand convergence, we will need some kind of evo devo-guided general optimization theory. Let's consider two necessary features of that theory now.

1. We can predict that any optimization that occurs must be on a continuum, from highly optimizing convergence, which we will refer to simply as optimizing convergence (OC), conferring advantage in all the most competitive and complex environments, to a wide variety of other cases, which we can refer to collectively as less-optimizing convergence (LOC). LOC cases would include convergence that offers only some temporary or local adaptive advantage, to just a few specific species, or in some subset of specialized or less-complex environments, convergence that offers no advantage, or convergence that is deleterious but not fatal. Names for a few general classes of LOC cases have been offered by scholars, including passive convergence, parallel evolution, etc.
2. Optimizing convergence can occur via both physical and informational processes. Physically, we might see greater efficiency of employment of physical resources, as in Bejan's constructal law, or greater density of employment of physical resources for offense or defense, the escalation hypothesis (Vermeij 1987). Informationally, we might see efficiency or density gains via informational substitution for physical processes, what Fuller called ephemeralization, or greater general intelligence (modeling ability), greater immunity, or a more

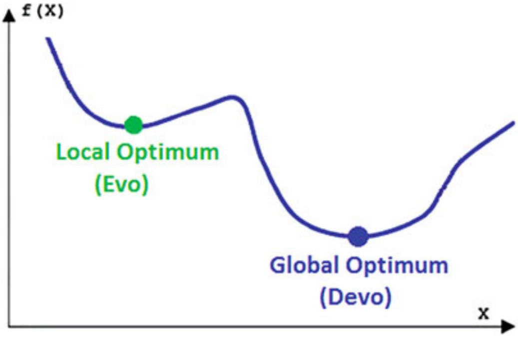


Fig. 7 Adaptive landscapes allow both local and global optima. (Source: ASF)

useful collective morality, offering more general and persistent adaptation to a wider range of environments than previous strategies. Intelligence also offers the ability to modify environments to suit the organism, what biologists call niche construction or stigmergy, as humans, social insects, and many other species do either in limited forms or extensively today. To understand OC, we will need a theory of optimization that tells us when a physical or informational advantage is likely to be more generally adaptive, particularly in the most complex, competitive, and rapidly changing environments (Fig. 7). We also need to know whether there are any other paths that can lead, in a competitive timeframe, toward a competitively superior new form of adaptiveness. If not, then we may have discovered a developmental portal, a global optimum that represents a bottleneck, a singular pathway toward greater adaptation at the leading edge of local complexity. Organic chemistry, RNA, photosynthesis, and oxidative phosphorylation are all potential examples of portals that all universal life must pass through first, on the way toward greater adaptive complexity. They may be the only global optima on their landscapes, at the relevant timeframes, that will allow the creation of vastly greater adaptive complexity.

Another complication of optimizing convergence at the leading edge of complexification is that over time, it must occur within an increasingly limited set of evolutionary possibilities, as increasing developmental genes and processes at the leading edge will progressively limit the evolutionary possibility space within any particular inheritance system. Processes like heterochrony, neoteny, and gene duplication (Wagner 2003) can temporarily reverse generally growing genetic constraint, but only the invention of a new class of inheritance system, in a metasystem inheritance transition (e.g., self-replicating genes in organisms inventing self-replicating ideas in brains, inventing self-replicating algorithms in technology), seems able to lead to large new regimes of evolutionary innovation (Turchin 1977).

If we live in a noetic (information accumulating and intelligence-centric) universe, nervous systems would surely qualify as OC. Based on neurotransmitter and genomic differences, Flores-Martinez (2017) argues that nervous systems were convergently invented three different times, by comb jellies, jellyfish, and bilaterians. But only in a small subset of prosocial, tool-using, land-based vertebrate bilaterians do we see a strong trend toward runaway brain size. OC is clearly multifactorial for developmental transitions to more rapid, more stable, and more complex evolutionary regimes (e.g., cultural evolution).

Consider eyes again. As with nervous systems, which are particularly helpful in complex environments, we can make a plausible case that eyes, at one point in time, became a necessary functional adaptation in the most complex environments. Andrew Parker's light switch theory (*In the Blink of an Eye* 2003) proposes that the development of vision in Precambrian animals directly caused the Cambrian explosion. Critics have observed problems with the timing, and that complex eyes may instead be a consequence of rapid body plan complexification, rather than a generator of new selection pressure for complexification. Either way, this is a fascinating theory, as it implies a necessary coevolution of intelligence and morphological and functional complexity. Once they emerged, it is easy to argue eyes were an evolutionary ratchet, and that all visible animals in the most complex environments would soon need them, or a handful of other uniquely effective defensive strategies, to survive.

Many other examples of OC can be proposed, in the most physically and informationally complex, and rapidly changing, environments on Earth, including the necessary emergence of eukaryotes, oxidative phosphorylation, multicellularity, bioluminescence, nervous systems, bilateral symmetry, jointed limbs, opposable thumbs, tool and language use on land (much faster-improving than aqueous environments), culture, and technology, including machine intelligence.

To make a few intelligence-related predictions in OC, I suspect that grasping limbs and tree niches on land are an early developmental portal (optimized convergence and phase transition in collective intelligence) leading to complex tool use on Earthlike planets, as tree swinging and grasping limbs offer an ideal training ground for complex, predictive brains, and as tool use and construction in air offer far greater mechanical advantage than in water. I am also a fan of Dale Russell's Dinosauroid hypothesis (Russell and Séguin 1982), which argues that the bilaterian tetrapodal humanoid form, which includes two locomotion and two prehensile (grasping) appendages, may be an optimizing convergence (minimum viable solution set) for becoming the most generally intelligent (and largest brain to body weight) land-dwelling bilaterian. I have also predicted that competitive-cooperative tool use on land, in the manner employed by early humans with Oldowan axes, is likely to be a universal developmental portal to runaway collective intelligence in bilaterians, as that environment seems to offer such strong selection pressures for generally adaptive defensive and offensive capacities, by contrast to animals that cannot collectively employ such "game-changing" early offensive and defensive tools as stone axes, clubs, and fire (Smart 2015).

Future science will need better theories of complexity, complexification, and optimization, to deeply understand such candidates for evolutionary convergence, and to distinguish the much greater variety of examples of less-optimized convergence from the most highly optimized forms.

12 Optimizing Convergence as Accelerating and Stabilizing ED on Many Scales

When convergence is viewed from the perspective not of the evolving species, but from some larger system scale (the biogeography, the planet, the universe) we can view optimizing convergent evolution as a process of not simply evolution, but of evolutionary development (ED), an ED that continually accelerates and stabilizes its complexity in special domains of space and time.

When we claim a convergence process is an example of ED, we are not only claiming that some kind of general optimization is occurring. We are also claiming that some kind of evolutionary developmental process, with both "random" and creative evolutionary search, and predictable convergence, directionality, hierarchy, modularity, life cycle, and perhaps other features found in biological development, is being followed, at some larger systems level. Consider embryogenesis. Viewed from the perspective of the individual actors (molecules), we see mostly stochastic, divergent, and contingent processes. As we zoom outward to larger and longer spatiotemporal scales, we can also see a few convergent, hierarchical, and life cycle processes. To view optimizing convergence as not simply evolution, but as evolutionary development, we often must take these wider scale views, as in the following examples:

- Galactic and universal change offers many potential examples of not only evolutionary but apparent developmental change, as we have discussed. Curiously, the evolutionary development of complexity seems strongly accelerative, with increasingly rapid complexity transitions in increasingly local spatial domains (Smart 2008). As a high school student contemplating this trend in 1972, I recognized the logical limit of that process was the black hole. The first of these puzzling objects, Cygnus X-1, had been discovered just the year previously, in 1971.
- Stellar-Planetary-Astrobiological change offers more examples. When we look down from early universal change to the stelliferous era, and the genesis of our life-permissive planet and its star, astrophysical theory tells us that the way stars have replicated, and chemically complexified, through three different populations over billions of years, has been not only evolutionary (a variety of randomly arrived at star and planet types and distributions) but evolutionary developmental, involving a progressive drive to complexification in a predictable subset of types. Many astrobiologists and planetologists argue that a subset of chaotic and nonlinear ("evolutionary") stellar-planetary change has reliably led, with high probability and massive parallelism, to G-(and perhaps some K- and M-) class stars and Earthlike planets ideal for the development of archaeal (geothermal vent) life and, from there, to prokaryotes and eukaryotes. See Nick Lane's *The Vital Question* (2016) and Smith and Morowitz's *The Origin and Nature of Life on Earth* (2016) for two such stories.

- Biogeography and Ecology offer more examples of not only evolutionary but apparent developmental change. In biogeography, we find scaling laws, like Copes rule, and biogeographic laws like Foster's rule and Bergmann's rule, with their predictable processes of optimizing convergent evolution, or evolutionary development. The famous convergence of form seen in placental and marsupial mammals, on separate continents, offers another example of not just evolution, but biogeographic ED. For many more examples, including intelligence traits, see Conway-Morris (2004, 2015), McGhee (2011), Losos (2017) and our list of examples of convergent evolution (Wikipedia 2012, and Smart and Chattergee 2012–2017) in species morphology and function. In ecologies, we can identify predictable patterns of ecological change, including ecological succession, ascendancy, and panarchy.
- Culture, Science, and Technology change offers yet more examples. When we look above individual cultures and do cross-cultural comparisons, we find many examples of developmental features at the leading edge of competitiveness, including parallel invention and/or convergent development of archetypal scientific and technological inventions like fire, language, stone tools, clubs, sticks, levers, written language, mathematics, hydraulic empires for our first great cities, wheels, electricity, computers, artificial neural networks, etc. In each of these cases, a high-order convergence has occurred. These and other specific examples of cultural change look not only evolutionary, but evolutionary developmental (ED). Once these archetypes and algorithms exist, there's no going back, for any culture seeking to stay on the leading edge of physical and informational complexification, and general adaptiveness. Some form of technologically-mediated global superorganism (see Heylighen 2007) is another planet-scale development that might be necessary to regulate future state, corporate, and individual technological misuse and rivalries. We also find many examples of developmental constraint laws that operate in social and economic systems, like scale laws (West 2017; Bejan and Zane 2013) and more generally, the least action principle (Georgiev et al. 2015).

In each of these rough hierarchies of complexity, our universe is not only generating local variation, creativity, and difference; it is also developing toward a small set (in our present understanding) of currently predictable destinations. While there is much about cosmogony that remains unclear, we know that dark energy is accelerating complex galactic groups away from each other, that total entropy increases, and that an increasing fraction of the mass-energy of our universe will end up in black holes. The better we understand the conservative and predictable features of our universe, and can distinguish them from creative and unpredictable ones, the better we may understand evolutionary and developmental processes at all scales.

There are two additional curious features of this developmental trajectory, two unexplained phenomena we can observe across all of these complex systems, which must now be mentioned. Understanding and modeling them are among the greatest challenges of modern science:

1. The first is the ever-faster complexification rates seen in the historical record of the most physically and informationally complex locations in our universe, since the emergence of G-class stars, Earthlike planets, and almost simultaneously, on our planet, life. This acceleration was famously summarized in Carl Sagan's metaphor of the Cosmic Calendar. Ever since August, on this calendar metaphor, leading-edge complexity environments have become exponentially faster, more complex, and more intelligent, on average, on Earth. Sagan said this phenomenon, which we can call acceleration studies, was an understudied area of science, in need of better understanding. See Sagan's *The Dragons of Eden* (1977) for his original account and Heylighen (2008) for more recent work. It is my hope that better models of early universe and astrophysical, chemical, biological, psychological, social, economic, technological, and other evolutionary development will help us understand our universe's emergence record of ever-faster and more physically and informationally complex local environments.
2. The second is the increasingly informationally stable (developmentally immune, antifragile) nature of complexity in ever more complex environments. In prehistory, species could easily be destroyed by environmental change. But once we began recording and simulating our world in nonbiological substrates, human-technological culture has gotten better every year at recording, simulating, and recreating both biological and cultural information (Malone 2012). As a result, such information has become far more resilient to catastrophe (Smart 2008; Taleb 2012; Dartnell 2014). There is something about mind, culture, science, and technology that makes the information it produces more stable to destruction via environmental fluctuations. Perhaps a growing intelligence typically provides increasingly useful sets of adaptive strategies for survival. Some kind of nonlinear input-to-state stability (a form of Lyapunov stability) may emerge as intelligence's potential to moderate environmental inputs grows. Perhaps the most intelligent collectives develop not only greater immunity but greater morality (both have been proposed as subtypes of intelligence). This latter view is controversial, given recent human history with advanced technology, but there are good early arguments for it as well (Pinker 2010). Perhaps it is simply that increasing intelligence allows progressively more durable (informationally immune) forms of cultural memory to be developed (Malone 2012). The best descriptor of local intelligence's ever-growing immunity may be niche construction (environmental engineering), of which memory is just one form. Niche construction has afforded humanity the ability to move our core complexity to increasingly time-stable architectural environments (books, villages, cities, computers), but these are nothing compared to what may soon come. Several scholars have argued that humanity appears just a few decades away from being able to port its essential complexity, in body and mind, into a technological substrate (substrate shift). Such postbiological entities seem likely to be vastly more stable to destruction via any imaginable universal process, and far more redundant, than today's bio-dependent culture, science, and technology. Such entities should be able to harness (and do intelligence-guided experiments

with) molecular nanotech, fusion energy, and perhaps even subatomic processes (femtotech), and should no longer require either planets or functioning stars to maintain their existence (Forward 1980; Smart 2008; Davies 2010; Rees 2015). Due to accelerating change, such stable new entities also seem likely to arrive much earlier, in cosmic time, than most of us would presently expect or predict.

Both of these features, our potentially developmentally guided acceleration and our progressive forms of informational stability, suggest that today's current models of existential risks are likely to be overstating the near-term risk to our species of many apparent species-threatening events (for a detailed review of such risks, see Bostrom and Ćirković (2008)). The time for which such risks actually threaten our informational complexity seems to be rapidly decaying. We appear to be on the edge of entering a far more stable substrate for life and intelligence, in a cosmologically insignificant fraction of future time.

13 "Tape of Life" (Identical Earths) Experiments: Simulating Ecogeophysical ED

If life emerges on two similar Earthlike planets, either in reality or in a sufficiently accurate simulation test, then by definition we can predict that the evolutionary aspects will almost always turn out differently in the two environments, and the developmental aspects will turn out the same. This is called the "Tape of Life" experiment, and it is commonly discussed in the philosophy of biology and by some of the more systems-oriented evolutionary (developmental) biologists.

Beginning in the 1970s, leading evolutionary theorist Stephen Jay Gould (1977, 2002) famously predicted that little of life's functions and morphologies on another similar Earth would turn out the same as those presently found on our Earth. He expected a few broad similarities, in kingdoms and some phyla, but most species would turn out very differently, in his view. Beginning in the 1990's, Simon Conway-Morris (1998, 2004, 2015) famously argued the opposite, that most functions and many morphologies would turn out the same, optimized for replication and adaptation in this particular Earth environment. We may aptly call this an evolutionary developmental perspective on Earth's history (Fig. 8). In the decades since, some biologists and most astrobiologists have migrated from Gould's to Conway-Morris's camp, though the dividing line between predictable and unpredictable processes of change remains a productive and contentious debate.

In recent years, there has been a surge of studies of evolutionary convergence, motivated by such wide-ranging questions as the structure of the protein space to experimental evolution to evolutionary genetics to ergodicity in biophysics to the attempted "neo-Gouldian" developmental account of homology versus homoplasy (Dryden et al. 2008; Turner 2011; Lobkovsky and Koonin 2012; Pearce 2012; Powell 2012; Orgogozo 2015; Powell and Mariscal 2015; McLeish 2015; O'Malley and Powell 2016; Louis 2016). Roughly speaking, most of these new results

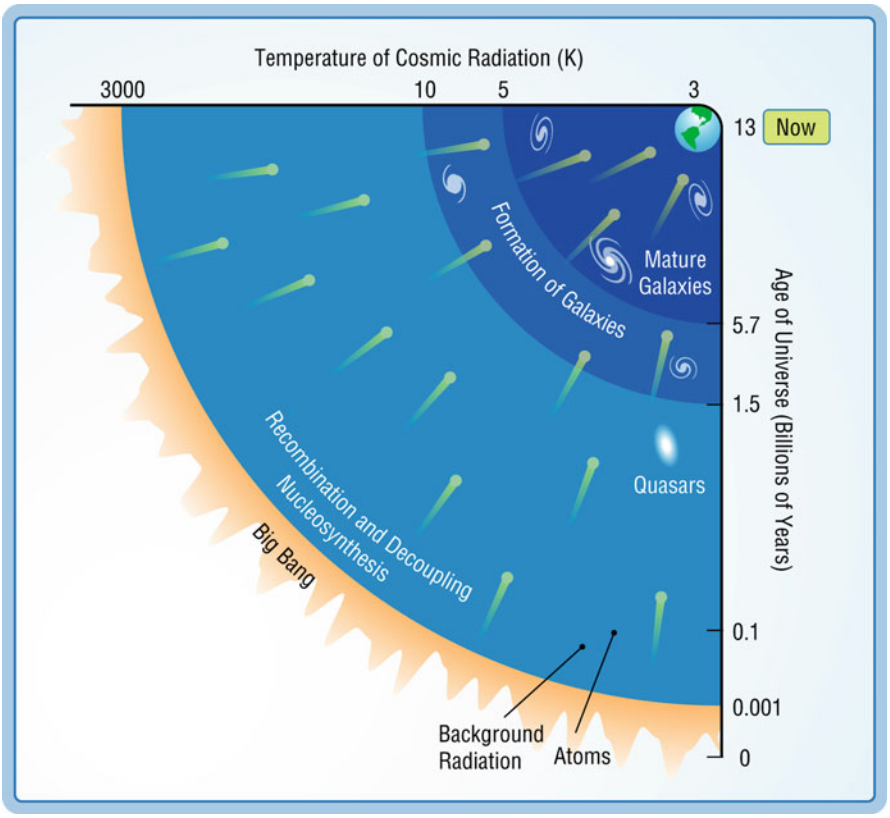


Fig. 8 Cartoon of a developmental Earth. Many Earthlikes, with stochastic differences (not depicted), would make an evo devo cartoon. (Source: ASF)

are strongly supportive of convergence – in more or less radical form – as the key feature of macroevolution. For example, Dryden et al. (2008) and McLeish (2015) argue that the accessible part of the genomic space is much smaller than conventional combinatorial wisdom suggests, and that evolution may have actually explored most of it by now. This is a powerful idea. Consider that a fully explored (statistically repetitive and no longer creative) phase space may be a necessary but not sufficient condition of all developmental portals (complexity transitions), to make such transitions appropriate guides (checks, funnels, gateways) to evolutionary exploration.

Convergent evolution, at all universal scales, can be productively modeled as a pull of attractors, and if those attractors are subject to replication and selection, as a process of evolutionary development. Such modeling should work, to varying degrees, whether we are describing physical, chemical, genetic, organismic, species, ecosystem, organizational, cultural, or technological evolutionary development. Perhaps the simplest phrase to encompass all these and other types is “universal evolutionary development." Applied to the universe, evo devo theory argues that both universal evolution (useful diversity) and universal development (useful similarity) must be aspects of any universal biology that some scientists and systems theorists (Flores-Martinez 2014; Mariscal 2016) are seeking. Though we seek simplicity in our models, discussing either unpredictable or predictable processes alone will lead to insufficient views of how adaptation actually occurs. We must learn how they blend, and when those blends are adaptive.

We must also recognize that just as in biological evo devo, our science and simulation skills will be insufficiently advanced to predict many of the developmental similarities ("convergent evolutionary developments") that emerge between two parametrically identical universes, two Earthlike planets, two similar but biogeographically separated continents, two highly similar cities or organizations, two genetically identical twins, or even two dividing cells.

Fortunately, the latter examples, and others, have happened many times on Earth. So we can look to these "natural experiments" to better understand processes of development, at all scales. As our science and simulation capacity gets better, we can also develop better and more predictive models of how our physical universe evolved and developed.

In a few of our more advanced biotechnological prosthetics (e.g., cochlear and vision implants, even hippocampal "chips"), our software and hardware models are good enough to substitute for the biological system without significant loss of function. We can hope that this intelligence substitution will also serve us as we learn to simulate universes in our future computers as well.

If so, we will increasingly be able to predict and validate ED hypotheses in at least two major ways. By discovering more natural experiments, at all scales, and by simulating the emergence of those experiments, at a level sufficient for the simulation to substitute for the physical process.

14 "Tape of the Cosmos" (Identical Universes) Experiments: Simulating Universal ED

Let's look now at convergent evolution on the largest scale we can presently imagine: our universe. In Carl Sagan's famous Cosmic Calendar metaphor of change (1977, 1980) (Fig. 9), we see that earlier stages of hierarchical evolutionary development, involving the emergence of large-scale structure, galaxies, and stelliferous and planetary change, are highly isomorphic and convergent, across the universe. Simply looking at the night sky shows us these amazing levels of convergence. In the last century, physicists have worked out many of the reasons this convergence is evolutionary developmental. It is written into the initial conditions and emergent laws of our particular universe.

Are the observable morphological, functional, and informational features of our universe that have clearly accelerated on Earth since the emergence of life, as depicted from August afterward in the Cosmic Calendar metaphor, also found

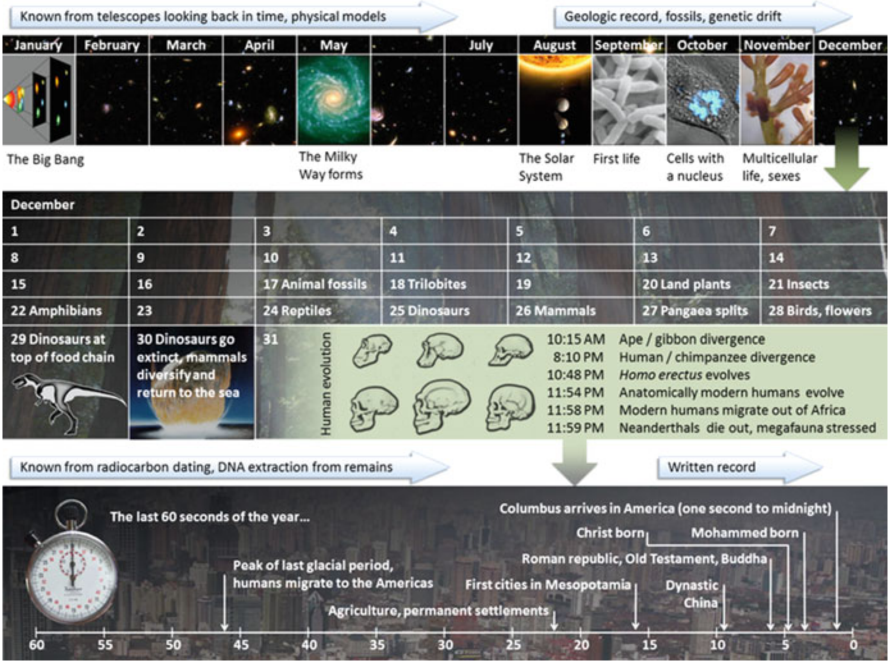


Fig. 9 The Cosmic Calendar: 13.7 billion years of universal history on a 12-month calendar. (Source: Eric Fisk, Wikipedia)

convergently throughout the universe? Is this convergence on multilocal complexity acceleration in our universe strong, happening with high frequency, as a developmental process, or is it random and happening weakly, as an evolutionary process? In other words, should we expect Earthlike acceleration in a multitude of special environments, such as those found on habitable planets around G-class stars? These are questions of universal ED. Astrophysicists and astrobiologists hope to answer such questions, by theory, simulation, and observation, in coming years.

Today we can conduct primitive simulation tests (“simulation experiments”) to explore the divergences and convergences we see in two model universes, but our science remains incomplete, and our cosmological simulations, both in their physical and informational dimensions, still do not capture all the reality they attempt to model. Fortunately, our experiments in simulating evo-devo phylogenetics in biology (Fig. 10) may lead the way to better simulations of all kinds of evo devo systems. If we live in an evo devo universe, our simulation experiments will get ever more predictive in their developmental components, and they’ll eventually convince even the most die-hard believers in contingency that we have a set of highly constrained futures ahead of us.

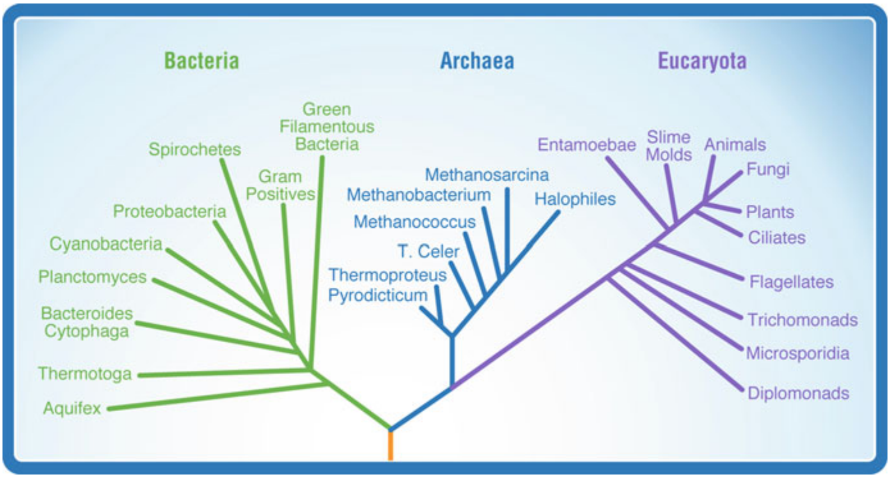


Fig. 10 Phylogenetic tree (simulation experiment). (Source: ASF)

Consider genetically identical twins. Most molecular and tissue-level aspects of two genetically identical twins look different when you view them up close (different fingerprints, organ microstructure, ideas, etc.). Those are "evolutionary" differences in an evo devo model. They are locally unique in myriad ways, either because the twins genetic systems aren't capable of ensuring perfect identicalness, or because there are adaptive (e.g., immunity) advantages to this local diversity. Genes are not a blueprint, but a recipe for building local complexity in a way that allows contingent local diversity, yet is also robust enough to local molecular chaos that each twin is reliably guided toward a set of useful far-future destinations in structure and function. All the aspects of the two genetically identical twins that turn out the same, we call "developmental."

Now consider that if our universe replicates, and its emergent features and intelligence undergo some form of self-selection or selection in the multiverse, this twin model helps us to define evo devo terms like universal evolution (variation between universes) and universal development (similarity between universes). Cosmology models typically assume that if our multiverse had two parametrically identical universes (universes with the same fundamental parameters and initial and boundary conditions), some aspects of those universes would turn out the same and some would turn out differently. Astrophysics guides our universe toward future-varying (evolutionary) and future-determined (developmental) form and function, at the same time. Both evo and devo processes, and a recognition of the adaptive value of both evolutionary variation and developmental conservation, would seem to be necessary to any accurate simulation.

15 Physical and Informational Adaptation: Autopoesis and Intelligence

Autopoesis is a term introduced by Chilean biologists Humberto Maturana and Francisco Varela (1973/1980) to describe the chemistry of living cells. It became popular with a few systems theorists in the late twentieth century to describe the capacity of some complex systems to self-reproduce and self-maintain. Autopoesis scholars seek to find general systems rules applicable to any stably self-reproducing complex systems, including not only living systems, but stars, the chemical origin of life, and ideas, behaviors, algorithms, organizational rulesets, and technologies in culture. Implicit to autopoetic models is the idea that a better information theory, including a theory of cumulative embodied and adaptive cognition (intelligence) in the replicator, its inheritance system, and its environment, will be necessary to understand dynamical change in complex systems. See Varela et al. 1974; Maturana and Varela 1987; Mingers 1995; Luhmann 2003/2013; Luisi 2003; Bourgine and Stewart 2004 for some autopoetic models. Mingers (1995) offers a particularly good introduction to rules, drivers, and research questions regarding autopoetic chemical, biological, social, and technological systems, though even this excellent work does not consider the universe as an autopoetic system.

While they have made little scientific progress to date, autopoetic models are focused on what we might call the right questions: the physical and informational sources of adaptation in self-producing, autocatalytic systems, and the ways adaptation changes over time in environments which are, in the most likely presumption, replicating, autocatalytic complex systems as well. At the least complex end of the spectrum, all stable replicating systems depend on the emergence of some set of predictable action-reaction couplings to their environment. Stars are autopoetic systems, dependent on physical action-reaction processes. Moving up the chain of information-production rate density (a form of complexity), a variety of self-reproducing prebiotic systems (clays, RNA, protein polymers) are dependent on not only physical but also chemical action-reaction processes afforded by the complexifying Earth environment. On the path to life, certain self-replicating chemical systems developed autocatalytic protometabolic networks (Kauffman 1993), and some developed sensory-motor cognitive Bayesian (predictive) chemical networks, including memory networks, observable in single-celled organisms like *Paramecium*, *Amoeba*, or *Stentor* (Bray 2011). At some point, gene-protein regulatory networks also emerged, and lipid cellularization. A subset of cells developed multicellularity, another subset developed specialized neural networks, and a subset of those, self-awareness. Understanding these increasingly complex set of adaptations, and the necessary emergence of life and mind as adaptations in certain environments, is one of the main challenges of modern science.

Scholars in such complementary fields as the origin of life (Smith and Morowitz 2016; Pross 2014), computational astrobiology (Pohorille 2012; Forgan et al. 2017), artificial life and information theory (Adami 2016), evolutionary escalation (Vermeij 1987), top-down causation (Walker et al. 2017), and evo-devo theory (see above) have all made progress in recent decades in understanding how successful cumulative replication changes the replicator, its seeds, and its local environment. Continued progress in such fields, especially in intelligence and information theory, will be critical to developing better models of adaptation.

Neo-Darwinian models of evolutionary adaptation, such as the adaptive landscape theory of Sewall Wright (1932) seek to model adaptation as phenotypic fitness to the environment, in some genotypic, morphological, or functional parameter space. Such models, while they have been tremendously useful, are also deeply incomplete, as they do not allow that the environment itself may change in predictable and highly nonrandom ways over time, as the growth of intelligence influences local environments, as described in the phenomenon of stigmergy/niche construction (Odling-Smee et al. 2003). They also do not consider that the selective environment itself may be both evolving and developing over time, changing the nature of selection and adaptation.

But if our universe itself is a replicator, as the evo devo universe hypothesis proposes, then it too is a selective environment that is not just evolving (experimenting, diversifying), it is also developing (complexifying, and engaged in a life cycle). From the perspective of biological evo-devo theory, much of this environmental complexification is both constraining, directional, progressive, and in-principle predictable, just as biological development is in-principle statistically predictable (though not always so in practice). If the universe is an evo devo replicator, at least some kinds of local environmental complexification will function to protect the replication and self-maintenance of the system (the universe). Any evolutionary (experimental, creative, contingent) activity that occurs within a developing organism must be increasingly constrained as that organism develops, in service to the organism's self-replication and self-maintenance. If our universe is an evolutionary developmental system, the local adaptive landscape must constantly be shifting toward certain developmental attractors, as evo devo complexity grows in certain local environments.

As with intrauniversal replicators, mechanisms that guide and protect universal replication may be very simple action-reaction and maximally energy-dissipative physical and informational processes, such as those that statistically guarantee stellar replication via star formation feedback in the nebular hypothesis (Krumholz and McKee 2005), or the reductive tricarboxylic acid (rTCA) cycle, a proposed universal intermediary metabolism (Smith and Morowitz 2004). The rTCA cycle generates the five fundamental precursors to all biosynthesis (acetate, pyruvate, oxaloacetate, succinate, and α-ketoglutarate), and may be a maximal free energy dissipator in high energy flow environments, like geothermal vents. The rTCA cycle can be catalyzed by inorganic mineral cofactors. When run in reverse, the rTCA cycle is the oxidative Kreb's cycle, central to all life. After the rTCA phase transition occurred in a local environment, presumably further phase transitions allowed combination with another cycle, oxidative phosphorylation, and an energy harvester, photosynthesis, to store energy for the first cells. It is possible energy storage wasn't part of the first metabolism, as some photosynthetic bacteria use the rTCA cycle (Kreb's in reverse) to produce carbon compounds (Smith and

Morowitz). Note that we've still left out DNA-guided protein synthesis, which is an information producing and environmental simulation system, and the transition that would merge it with metabolism, if we are going to describe the origin of life in evo devo terms. No wonder it is such a complex puzzle at present.

How advanced universal developmental processes may be, and how deeply they structure cascades of non-equilibrium phase transitions, may depend on the prior degree of universal replication, and the strength and nature of intra- and extrauniversal selection. At some point, the analogy with developmental genes in living organisms may apply, in which tuned parameters guide the emergence of planetary-scale social and technological processes that are functionally similar to biological intelligence, immunity, and morality. Such top-down causal informational mechanisms could be an integral part of our universe's self-maintaining processes. From a functional perspective, mind might inevitably emerge in a universal replicator, just as it has in biological replicators, if intrauniversal intelligence plays any usefully nonrandom role in universal replication and selection.

In summary, evo-devo biology may offer us the most complex and rigorous model for understanding not only convergent evolution in universal evolutionary development, but how adaptation itself must change in a universe that is itself a replicator. Once certain critical biological advantages, like eyes, emerge and are strongly adaptive in an environment, the majority of the most successful complex replicators in those environments may have to employ that advantage. Once certain critical technological advantages, like digital computers and machine intelligence, emerge and are adaptive in an environment, a subset of replicators (individuals, organizations, societies) must use those technologies if they wish to remain at the leading edge of adaptation. In this view only evo-devo biology, and its successive processes of molecular, genetic, physiological, and psychological evolutionary development, offer us a sufficiently complex analogy for understanding how adaptation may change in the universe, if it too is a self-reproducing, self-maintaining evo devo system.

16 Evo Devo Models Require Progress in a Variety of Theories, Especially of Intelligence

If universal evolutionary development is occurring, future science must show that each successive environment in the developmental hierarchy inherits certain initial conditions and physical constancies from the environment that preceded it, back to the birth of the universe, and that some of these initial conditions and constancies act to predictably constrain the future dynamics of each successive environment, to some degree. Such constraints have been called developmental attractors (or more commonly, just attractors) by a variety of scholars. If they are the *only* such attractors on the adaptive landscape at that level of complexity and timescale, in a universe where accelerating complexification is possible, and such acceleration results in local dominance of the most rapidly improving systems, then I think it is clarifying to call them developmental portals (gateways, checks, bottlenecks) as well (Smart 2016a). For specific examples, G-, K-, and M-class stars and organic chemistry may be necessary portals to planets capable of generating life. Fats, proteins, and nucleic acids may be necessary portals to cells. Eyes may be necessary portals to higher nervous systems. Tree niches (which support complex prediction), and animals with grasping appendages, language, and technology use may be necessary portals to human civilization acceleration, etc.

From an adaptive landscape (phase space) perspective, if ED is occurring, as the evolutionary "search" landscape gets more diverse and locally complex, certain portions must convert into developmental funnels, then portals. These portals must also work together to periodically produce a metasystem transition (a higher level of order or control), a new level of ED hierarchy. Both the landscape's tendency to produce funnels/portals as complexity emerges, and the number of portals (lower is generally better) are two obvious ways to maintain developmental control in any evolutionary (chaotic, creative, locally unpredictable) system.

It is widely agreed that physical complexification, and such riddles as the origin of life, must be described by non-equilibrium thermodynamics as a coupled cascade of phase transitions in energy degradation and information production. As Smith and Morowitz (2016) state, each emergence (phase transition) in the development of hierarchy creates new simplifying constraints and logic, and there is a floor and a ceiling of environmental complexity for which those constraints apply. Reductionism can be very successfully applied at each level to discover its internal constraints (laws of chemistry, biology, etc.). The holism problem comes at the transitions (portals) between hierarchies. Some combination of bottom-up (evo, atomistic) and top-down (devo, holistic) parameters are involved, but how this works in any transition still remains unclear.

Evo-devo genetic and epigenetic models, as they seek to differentiate developmental and "evolutionary" gene fitness landscapes, will have to incorporate phase space models and (wherever there is high dimensional reduction) landscape models as our theory, tools, and data advance. Unfortunately, there are many problems with current adaptive and fitness landscape models in depicting the hyperspace of structure and function, and as critics of the adaptive landscape metaphor point out (Kaplan 2008) few models incorporate any concept of probability of movement across the landscape. In useful landscape models, potential portals would have to emerge as persistent, and theoretically globally optimal peaks (or in a more thermodynamically useful depiction, valleys) on adaptive fitness landscapes. These models will eventually have to evolve into network-based models with search basins and portal paths, which include both "evolutionary" tangles of similar-fitness landscapes and portals (Crutchfield and van Nimwegen 2002) as well as regions that use portals to predictably transition to globally optimal, hierarchical and developmental forms, landscape locations offering the greatest resource (space, time, energy, matter) efficiency or density of adaptation.

Another field that will help evo devo models advance will be protein folding, which already use funnel (portal) landscapes to depict 2D to 3D transformation of protein structure, involving both energy minimization and information production or conservation, a key example of biomolecular evolutionary development. In evo devo models, alternative chemistries for life, periodically sought by astrobiologists (see Goodwin et al. 2014) if they continue to be undiscovered by observation or simulation, would be more evidence indicating a universe with a high level of ED (self-organizing) constraint on the life transition. Science fiction authors and origin of life theorists have been imagining them for decades, but so far we haven't found any particularly credible alternatives, in my view. Such constraint (only one physico-chemical portal for the life transition being accessible in reasonable astronomical time, see Koonin 2007), if it exists, might be due to strong or weak multiversal selection for life and intelligence with both evo and devo properties, over many past cyclings of our universe.

In addition to better simulation capacity, progress in any theory of evolutionary development will require better:

1. Complex systems theory – Seeing the appropriate system and scale at which ED is occurring, and any information-dependent processes (goals, drives) that may operate all in complex adaptive systems. I offer one such speculative model (Five goals of complex systems) in Smart 2017b.
2. Evo-devo theory – Better understanding organismic ED, modularity, reaction-diffusion systems, gene-protein regulatory networks, intelligence, immunity, morality, and other ED features of living systems, both individually and as collectives. This will require advances in evo-devo genetics and epigenetics, theoretical morphology, paleontology, evolutionary (developmental) biology and psychology, anthropology, sociology, economics, and many other fields.
3. Adaptation theory – Moving beyond the MES (modern evolutionary synthesis) to evo devo models, including self-selection (intelligence) and self-organization (development) as sources of adapted order.
4. Optimization theory – Reliably differentiating less-optimized convergence (LOC) and optimized convergence (OC) in the emerging study of convergent evolution, via better definitions, tools, data, models, and optimization functions.
5. Acceleration theory – Understanding accelerating change, in ED terms. When it happens as a physical process, acceleration always seems to involve both densification and miniaturization of critical adaptive processes in complex systems. Speculative proposals like the transcension hypothesis (Smart 2012) and the stellivore hypothesis (Vidal 2016) extrapolate accelerating densification trends in adaptive systems to their universal limit, a black hole. Acceleration also happens via informational or computational processes as well. For that we may need a better intelligence theory.
6. Intelligence theory (cog evo devo) – The Baldwin effect is the recognition, beginning with James M. Baldwin in 1896, that learned behavior affects an organism's reproductive success. It is a modest start in understanding learning and intelligence in evo devo systems, but we must go much farther. The better we understand the evo and devo roles for cognitive processes in adaptation (and today we often do not) the better we may understand the roles of intelligence

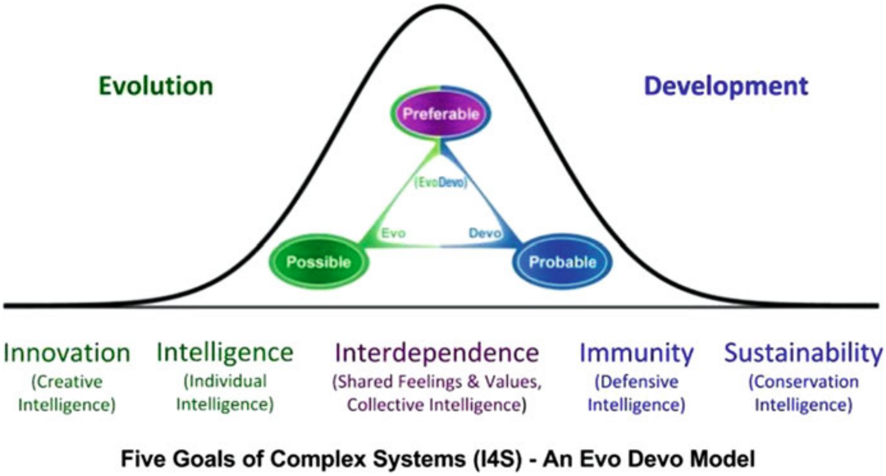


Fig. 11 Five goals (abilities, values, drives, telos) of complex systems, a possible evo devo self-selection function for the growth of adaptive intelligence. In this cartoon, interdependence (collectively shared feelings, values, preferences) is proposed as the central (most evo devo mixed) goal of complex systems (Smart 2017b)

for the universe as a replicator. I can imagine (Smart 2017b) at least five goals of complex systems, innovation, intelligence, interdependence, immunity, and sustainability, each of which may be considered a form of intelligence (Fig. 11). All of these goals may be self-selecting in evo devo systems, and their interaction a primary driver of adaptation, as follows:

6.1. Intelligence as innovation (exploratory intelligence) – Evolutionary process is the hallmark of this type of intelligence. As Shapiro 2011 and others propose, living systems harness stochasticity to generate selectable variety (experiments, possible futures), particularly under stress or after catastrophe. They seek to do this in increasingly clever (“good bet”) ways, the smarter they become. Evolutionary innovation is nonrandomly guided by intelligence, particularly in the “next adjacent” action and feedback cycle. At the same time, the complexity generated becomes rapidly unpredictable the farther ahead any intelligence looks.

6.2. Intelligence as intelligence (representation intelligence) – Most fundamentally, intelligence is a process of informational representation of environmental reality (Fischler and Firschein 1987). Informational representation (modeling) can be argued to be a dominantly divergent, evolutionary process. Our neural models conform to regularities of their environments, but they also generate astounding numbers of exploratory representations, only a fraction of which are universal (predictable) or adaptive. Think of imagination, fiction, or theoretical math, most of which has no known application. Being “intelligent” is also no guarantee of being adaptive.

Indeed, those with too much of this single ability may be maladaptive. The finite nature of all intelligence (its computational incompleteness) also strongly argues that massive parallelism is a fundamental adaptive evolutionary strategy. All models are wrong, but some are useful. Massive parallelism, and regular selection on that parallel variety, seems key to how genes, neural nets, populations, and civilizations get more adaptive.

6.3. Intelligence as interdependence (empathic-ethical intelligence) – Organisms engage in positive sum games, rules and algorithms (morality, ethics), involving not just self- and world-modeling but collective competition and cooperation, coordinated by other-modeling and other-feeling (empathy). Complex interdependent organisms develop culture, which evolves and develops independently from the individual, both faster and more resiliently, and allows them to view and optimize outcomes from either personal or group perspectives (which may conflict). A variety of universal evolutionary and developmental ethics (algorithms that protect collective adaptation and intelligence) may apply to all complex cultures. For more on how emergent synergies (interdependences) may have driven major transitions in evolutionary development, see Corning and Szathmáry 2015.

6.4. Intelligence as immunity (security intelligence) – Organisms use many strategies for differentiating self from other, and passively and actively countering degradation and predation. Chronic stress and stress avoidance both weaken immunity, but right-sized cyclic stress and catastrophes both build immune system capacity and accelerate evolutionary innovation. Taleb's concept of antifragility argues this for organizations, as does the catalytic catastrophe hypothesis. If there are universally discoverable security architectures and strategies (many ways to fail, only a few ways to survive), as I suspect, then immunity can be classed as a dominantly convergent and developmental process.

6.5. Intelligence as sustainability (predictive intelligence) – Developmental process itself is the hallmark of this type of intelligence. Organisms use their intelligence not just to explore possible (innovation, intelligence) and preferable (interdependent, immune) futures, but to build predictive, and presumably Bayesian, models of probable futures. A subset of these predictive models are encoded in an organisms developmental genes, in emergent properties of their soma, in their environment, and in more complex organisms, culture. The growth of knowledge, common sense, science, and all the processes of development that predict, but do not protect (immunity) can all be considered sustainability. These processes grow "truth" and understanding. This form of intelligence is in constant tension with innovation, which can rapidly cause both poorly understood and dangerous forms of complexity to emerge.

6.6. Intelligence substitution – Understanding precisely when information, or a computational process, can substitute for a physical process, and either retain or improve adaptiveness for the system under study. Some scholars call this dematerialization, or ephemeralization. Along with densification, dematerialization seems to be a central driver of accelerating complexification (Smart 2016b).

6.7. Intelligence partitioning – Adaptation and intelligence always exist in three interacting subsystems: seeds (with evo and devo initial conditions), organisms (which engage in a life cycle), and the selective Environment (some scholars call this ambient intelligence). Because of niche construction or stigmergy (intelligence always alters its local environment, in minor or major ways, changing adaptive pressures), environments essentially replicate along with seeds and organisms (think of the replication we see in city structure and function) and are a full partner with organisms and seeds in the ED of intelligence.

7. Hierarchy theory – Seeing the ED trajectory for the system as a whole. Stan Salthe's work on subsumptive hierarchies is an excellent example. Hierarchy theory (Salthe 1985, 1993, 2012) tells us how each new hierarchy is in some sense more free and in another more constrained than the latter. While we traditionally think of intelligence in an evolutionary role (increasing diversity and options), hierarchies tell us the ways that new "higher" systems are also more developmentally constrained than the ones from which they emerged. Using Volk's concept of "combogenesis" (Volk 2017), we can think of chemistry as both a set of new freedoms (to space- and time-efficiently create new adaptive structure and function) and new constraints on the local dynamics of physical laws. Biology locally enables and constrains chemistry, society locally enables and constrains biology, and so on. In a physical universe, such nesting and accelerating hierarchies must have a limit, a point at which further evolutionary development cannot proceed within this universe (Smart 2008, 2012).
8. Information theory – Convergent evolution in biology can be modeled as the result of networks made up by biomolecules or other agents that are organized and structured by information hierarchies emerging via top-down causation. The emergence of modularity and of functional equivalence classes in subroutines – both in biological and technological systems – can be explained via such information hierarchies. Top-down causation describes the process whereby higher levels of emergent informational organization in structural hierarchies constrain the dynamics of lower levels of organization. In a typical reductionist paradigm it is assumed that purely physical effects entirely determine the dynamics of lower levels of organization and, by extension, at higher levels as well. But an emerging school of investigators hypothesize that the transition from non-life to life, abiogenesis, requires a top-down transition in causation and information flow (Flack 2017; Walker et al. 2017).
9. Life cycle theory – Seeing the full replicative cycle of the developing system. If we can predict the remaining stages of the life cycle in any system, aided by comparisons with other evo devo systems, we can see its developmental futures, in broad outline at least. Its evolutionary futures, of course, remain intrinsically unpredictable at the same time. Both predictable and unpredictable processes are perennially found in complex systems, whether an organism, a culture, a star, a galaxy, or a universe.

Building better hypotheses and theory of evolutionary and developmental processes will be an immense amount of work. But this path may be the only viable way forward (a conceptual developmental portal) to fully understanding such scientific challenges as convergent evolution, galactic development, and the origin of life. If validated, the benefits we stand to gain, via better collective foresight, also seem comparatively immense.

17 Observation Selection Effects: The Challenge of Assessing Them in Evo Devo Terms

Any form of reasoning about traits or properties which constitute observers, or that are logically or physically necessary for the existence of observers, is subject to observation selection effects and biases. The importance of these selection effects and biases has only recently begun to be fully appreciated. For example, the number of small bodies' (asteroids and comets) impacts in Earth's history is constrained by our existence at the present time through the "anthropic shadow" effect (Ćirković et al. 2010).

Several detailed reviews of observation selection effects exist (for example, see Bostrom 2002). Observer selection arguments and models are often used to critique the fine-tuned universe hypothesis. Unfortunately, our cosmological models remain quite primitive, so it is easy to argue that either fine-tuning or observer selection bias is more important in such models. But even more fundamentally, as I have argued before (Smart 2008) all observer selection models in common use depend on a random observer self-sampling assumption, a random distribution of possible universes, or some other random, Monte Carlo-style mathematical framework in their evaluations. In other words, they assume we live in an essentially random, evolutionary universe, and this is a major limit to their utility.

Consider that if we actually live in an evo devo universe, such math must itself be incorrect. If our universe is not just randomly (contingently) evolving, but it also nonrandomly developing, then some subset of its critical probability distributions (informational and dynamical) will continually be skewed in the direction of the universe's developmental trajectory, given its special initial conditions and constraints. As complexity and hierarchy grow in local environments, those environments will further bias a special subset of locally constraining, nonrandom developmental processes to occur. Such developmental biases may be why accelerating complexification occurs in special environments (acceleration bias) and why spontaneous abortions (miscarriages) are so frequent early in gestational development, but so rare late in development, with miscarriage frequency in humans declining from 40% of pregnancies at conception to 0.1% of pregnancies at 42 weeks of gestation (Rosenstein et al. 2012). Presumably, the more developmentally complex both the fetus and the gestational environment become, the less often that any random perturbations of a standard size or duration are disruptive (stability bias).

Math that describes an increasing developmental bias toward both acceleration and toward growing informational stability during complexification in special environments is the kind of math we may need to properly model developmental processes, and to properly understand complex observers. Such observers are not random, they are privileged, in some proportion to their complexity. We surely do not live in an anthropic universe, if by that we mean one self-organized for the end purpose of producing biological humans. But we may well live in a noetic (intelligence-centric) universe, self-organized to produce accelerating and increasingly stable intelligent observers, as a central adaptive strategy for the universe itself. Biological humans may well be one of a long chain of developmental purposes, both an important and a transitory intelligence substrate (Hoyle 1983; Gardner 2007).

Even the frequency of evolutionary convergence, versus presumably much more commonplace evolutionary contingency in biological change, is a complex issue we don't yet agree upon. In the Signor-Lipps Effect, because rarer species are much less common in the fossil record, and the record itself is so sparsely sampled, rarer species will seem to disappear long before their actual time of extinction, simply due to chance. This makes the timing and speed of mass extinctions, and the ancestry of specific genera both much harder to determine from paleontology alone. As one consequence, quick and dramatic extinctions can go undetected, as they may look gradual due to selective sampling of a poor record, a classic observation selection effect (Signor and Lipps 1982).

Unfortunately, even the consequences of known biases like this are not clear. On one hand, if extreme ecological perturbations have overturned entire faunas via mass extinction events at much greater rates than we presently appreciate, and more rapidly switching Earth's macroevolutionary regimes (Jablonski 1986), then the contingency that life is subject to due to intrinsic evolutionary (diversity-generating) mechanisms may be less than is currently theorized. Our environment, not evolution itself, may be our greatest source of contingency. On the other hand, it has long been argued that environmental catastrophes themselves act as a major (and perhaps even the primary) catalyst for evolutionary innovation (Gerhart and Kirschner 2005; Bhullar 2017). We know that immune systems depend on catastrophe and hormesis to get stronger, and evolutionary gene complexes may also depend on catastrophe and chaotic stress to innovate. Without periodic catastrophe, some biologists have argued that evolutionary contingency must steadily reduce with time (Salthe 1993; Shapiro 2011). In stable niches, stabilizing genes presumably gain the upper hand over innovating genes. At other times, evolutionary innovation itself has provided the catastrophic environmental change, spurring further evolutionary innovation, as in the Great Oxygenation Event, causing massive dieoffs due to oxygen-excreting cyanobacteria, the End-Ediacaran extinction of large sessile organisms due to the emergence of mobile sighted animals during the Cambrian Explosion, and the Permian-Triassic extinction, perhaps precipitated by the emergence of *Methanosarcina*, a methane-synthesizing archaea (Ray 2017).

Geerat Vermeij has offered a particularly interesting and relevant argument supporting high frequency of convergence in evolution based on the logical asymmetry between singular and multiple events in the incomplete terrestrial fossil record (Vermeij 2006). If some specific character, call it Z, demonstrably evolved multiple times in different lineages in different ecological conditions, this is clearly an argument for the convergent nature of Z. However, if Z evolved only once to the best of our knowledge, it is still not an argument for uniqueness, since the incompleteness of our record may hide repeated instances of the independent evolution of Z. This logical asymmetry has intriguing consequences when a wider set of various important innovative characters scattered over all of the history of life are analyzed from a Bayesian point of view. Vermeij was able to show that the alleged singular innovations tend to be either more ancient or to appear in small clades. Small clades, in turn, are more invisible in the fossil record if located in the more distant past. In other words, purportedly unique innovations in small clades, or in the distant past, may be only the latest or the dominant instances of convergent evolutionary events, with most past convergences hidden from our view.

Vermeij concludes that few innovations are ever truly unique: "Purportedly unique innovations either arose from the union and integration of previously independent components or belong to classes of functionally similar innovations. Claims of singularity are therefore not well supported by the available evidence. Details of initial conditions, evolutionary pathways, phenotypes, and timing are contingent, but important ecological, functional, and directional aspects of the history of life are replicable and predictable" (Vermeij 2006, p. 1804). Insofar as key evolutionary innovations are largely determined by universal principles of physics and economics, they will lead to widely-to-universally useful designs. This is the classic view of environment-driven convergence.

18 Ambiguity of the Word "Evolution" and the Modern Evolutionary Synthesis

In the scientific literature, the term "evolution" is used to describe any process of growth or change that involves the accumulation of historical information, in either living or nonliving complex systems (Meyers 2009). When we restrict the term to refer to biology, and modern forms of Darwinian evolution, it is used to describe cumulative inherited change, via descent with modification from preexisting organisms. A classic conceptual model of Darwinian evolution, often taught in undergraduate classes, is the acronym VIST. Evolutionary change is proposed to happen via Variation, with Inheritance, and (Natural) Selection, over long amounts of Time (Russell 2006).

While it is a good start, there are three basic problems with the VIST model:

1. VIST does not explicitly consider the concept of development, and of developmental genes and processes, which act in opposition to processes of variation within the organism. Developmental genes and processes are those that keep the organism on a convergent, conservative life, and reproduction cycle. Their fundamental role is Convergence, funneling the organism toward a series of future-specific states. Variation, within the organism or within the environment, is the "enemy of development." It must be overcome by Convergence, if the organism is to develop in a predictable way. Unfortunately, both classical Darwinism and modern evolutionary theory deprioritize the influence of organismic development on macrobiological change.
2. VIST does not explicitly consider cumulative Replication, and its growing informational constraints on organizational change, in any substrate, over cumulative life cycles. Replication is implicitly considered as the factor of "Time" in the classic VIST model. But it is not Time that causes biological change. Organic change occurs via cumulative and increasingly ergodic cycles of Replication (of the organism), within any substrate, as guided by Inheritance factors (genes, brains, and other information carriers, or "seeds"), and Selection (in the environment). In all three of these interacting systems (organisms, seeds, and the environment) we find processes of Variation (evolutionary processes) and Convergence (developmental processes), working together in service to adaptation. Considered together, these five factors give us the VCRIS model of change. After variation and convergence themselves (what changes, and what doesn't), replication is the next most fundamental process we should acknowledge in any model of the self-organization of complex adaptive systems. Whether we are discussing replicating suns creating organic chemistry, replicating chemicals creating cells, replicating cells creating organisms, replicating organisms creating ideas, or replicating ideas creating self-replicating machines, we must recognize that the most complex forms of adaptation, learning, and intelligence always require replication, inheritance, and life cycle, in some kind of "organism" (system).
3. VIST doesn't recognize that the natural environment may itself be not only evolving (creating unpredictable experimental variety, by our definition above) but also developing, if our universe is itself a replicator. As a result, the Modern Evolutionary Synthesis, our current standard in biological investigations, is biased toward the idea of an "accidental" universe, and "random" experimentation and diversity as the primary (or in some views, exclusive) cause of macrobiological change.

Evo devo models, whether in biology or in other replicating systems, help us correct the biases of both the original Darwinian VIST view of evolution (Fig. 12, white oval), and of modern evolutionary theory (Fig. 12, light gray oval), both of which view diversification as the primary source of adaptiveness. Each of these views ignores or minimize the converging, conserving role of development, and the possibility of development on scales far larger than the organism. An evo devo-

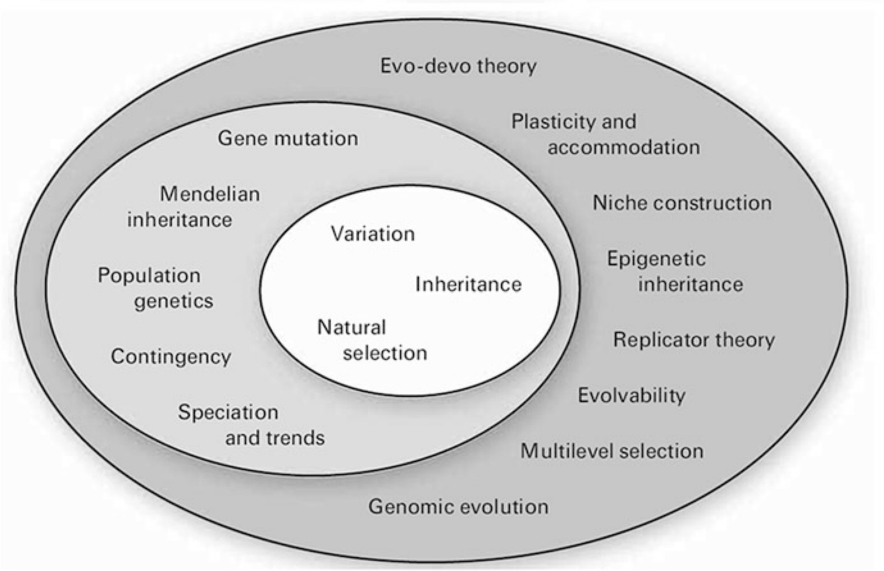


Fig. 12 Conceptual schematic of Darwinian evolution (1859, white oval), the modern evolutionary synthesis (post-1940, light gray oval) and current concepts in the coming extended evolutionary synthesis (post-2000, dark gray oval). (Source: Pigliucci and Müller 2010, with permission)

centric perspective (Fig. 12, dark gray oval, for the case of living systems) will allow us to see that complex adaptive systems must harness both unpredictable, divergent evolutionary stochasticity and predictable, convergent developmental destiny and life cycle in search of greater adaptiveness and that these two sets of mechanisms act in productive opposition to and tension with each other, at every scale. Evo devo models allow us to see evo devo self-organization as the natural source of adapted complexity and causal order in all successfully replicating systems, which we must come to understand from both physical and informational perspectives.

As we'll discuss shortly, understanding self-organization also shows us why challenges to Darwinism that have been launched by groups like the "intelligent design" community are more in line with supernatural belief, not science. They are typically motivated by belief in an "intelligent designer." But if the universe replicates, as several cosmologists propose, parsimony and evidence both argue that evo devo self-organization, via many past replications in a selective environment, not intelligent design, is the source of the intelligence we see.

After we have done our best to adjust for observer-selection effects, we still see many highly unreasonable examples of adaptedness for complexification, in the laws and processes of our universe as a system. The phenomenon of accelerating change, evolutionary constraint laws (like the constructal law and various scaling laws), terminal differentiation of morphological complexity, the fine-tuned universe hypothesis, the presumed fecundity of Earthlike planets, the collective morality of social animals, and the Gaia hypothesis (in a more rigorous form) all come to mind. To explain such unreasonable adaptedness for complexification in our universe we should think first of replicative self-organization under selection, not design. After all, such self-organization is our best model for the source of the intelligence that is reading this page, right now.

As we come to understand the complex phenomenon of convergent evolution, on myriad system levels (physical, chemical, genetic, morphological, functional, algorithmic, cognitive, technological, etc.), we will rectify the historical biases that the Modern Evolutionary Synthesis (MES) has perpetuated with respect to our presumably living in a "random," "directionless," and "purposeless" universe. To do this, we will need what Pigliucci and Müller (2010) and in a particularly comprehensive review, Laland et al. (2015) call an Extended Evolutionary Synthesis (EES). I expect this synthesis must include evo-devo and evo devo perspectives, a better theory of intelligence, better science and simulations, and more.

A large and well-funded group exploring an EES, led by evolutionary biologists Kevin Laland and Tobias Uller, can be found at ExtendedEvolutionarySynthesis.com. Another group working on an EES, led by biochemist and molecular biologist James Shapiro and physiologist Denis Noble, can be found at The Third Way of Evolution. This latter website is admirable, but not entirely error-free. As biologist Jerry Coyne points out in a post at the Richard Dawkins Foundation, the web editor of the Third Way website, Raju Pookottil, who does not have biology training, once argued that life "carries the hallmarks of design." That is a useful critique (see my section on the Fallacy of Intelligent Design below) but Coyne's post also ignores the scientific contributions of the many eminent meta-Darwinist scholars listed at the website. In exploring an EES, both poor evolutionary thinking and ultraorthodoxy with respect to the modern synthesis must be avoided.

"Ultra-Darwinists" like Coyne and Dawkins have attracted this label whenever they advocate, with a confident certainty, the position that contingent evolutionary selection (neo-Darwinism) can be the only force driving macrobiological change. Though Darwinism has deep evidence behind it, and its models appear to aptly describe the vast majority of (divergent and contingently convergent) organic change, they also seem insufficient to explain a small subset of phenomena that appear developmental and universally convergent. That subset includes the accelerating development of intelligence, and the increasingly nonrandom guidance of evolutionary innovation in intelligent systems (Smart 2008). In a similar way, the vast majority of change we can sample at the molecular scale in biology seems locally stochastic, randomly selectionist, and diversity generating ("evolutionary"), yet a small subset, as we have proposed, also seems deterministic, convergent, and predictively selectionist ("developmental"), particularly when viewed from larger or longer-range spatial, temporal, energetic, and material (STEM) scales. Better defining and empirically discovering that subset seems a reasonable next step in evo devo inquiries.

19 The Fallacy of Intelligent Design as an Explanation for Adapted Complexity

Religious belief is deeply valuable to many of us, and religious communities are globally important social institutions. Religion is humanity's first effort at universal moral philosophy, greatly predating Greek natural philosophy, and religion often ventures first into areas of moral prescription where science cannot yet easily tread. History shows that religious community has provided invaluable guidance and public benefit for millennia, and that all of our most socially successful religions are continuously reforming their beliefs and practices to be congruent with accelerating scientific knowledge.

Yet a key insight in the philosophy of science is that every intelligence is woefully finite and incomplete relative to both the current and future complexity of physical and informational reality. Thus we all must live with our own pragmatic sets of unproven beliefs, and many of us will seek communities that share those beliefs. It seems inevitable that self-aware artificial intelligences, if and when they eventually emerge, must also evolve and develop their own set of religious beliefs (read: philosophies of universal purpose, meaning and value), as there will remain many areas of reality about which they will know little. Fortunately, freedom of religious practice and freedom from religious discrimination are bedrocks of all modern democracies. Besides the traditional religions, atheism, agnosticism, possibilianism, universism (my own belief), and others are all belief systems that offer valuable, unproven beliefs about metaphysical reality.

As good practitioners of science and natural philosophy, all of us should attempt to make our unproven beliefs explicit and public, and seek to test them with evidence and experiment in whatever partial ways we can. When we feel we cannot separate our unproven beliefs from the practice of science, we should declare our influences. Unfortunately a number of scholars in the intelligent design community do not do this, and their religious belief has led some members of these communities to a variety of objectionable political acts, like seeking "equal treatment" for their evidence-poor hypotheses in the science classrooms of our public high schools. Given the intelligent design community's position on mixing religion and science, and not declaring their supernatural beliefs, we do not welcome scholars affiliated with the Discovery Institute or other intelligent design or creation science communities within the Evo Devo Universe (EDU) research community, and non-naturalistic discussion of religious belief is outside the scope of our community.

Given that there are fully naturalistic, evolutionary developmental explanations for the sources of "design" we see in living systems, and given the sharply limited value of all known physical intelligence, the concept of intelligent design, as it is generally proposed, appears fallacious. We must recognize that adapted intelligence has always had a useful but very minor influence on processes of selection in VCRIS systems. A crucial insight is that no physical intelligence ever becomes "Godlike" in its ability to predict either its own or its environmental future. We must acknowledge that all our present attempts to "rationally design" our own environment, including our genetically modified organisms, must be more accurately characterized as intelligence-guided guesses at more adaptive forms and functions. Human science and engineering are always evo devo process, like all other natural processes. They might be 95% bottom-up creativity/experiment/serendipity, and 5% top-down discovery/optimization/prediction, if we were to guess a very rough ratio. They are never a fully top-down or future-omniscient "design." The universe and all evo devo systems are far too chaotic and contingent to allow such omniscient foresight.

If we live in an evo devo universe, it is easy to argue that our future must continue to become rapidly computationally opaque to any finite and physical beings, the further ahead they look into their own futures. Combinatorial explosions of possibilities and contingencies, both in the universe itself and in our own mental processes, must always limit our foresight. No matter how advanced we become, any intelligences generated by this universe, or its ancestors, seem destined to remain evo devo "gardeners" as opposed to omniscient engineers, finite beings with "free will" (self-unpredictable evolutionary futures) not gods.

Supernaturalism takes many forms, some quite subtle. Even otherwise deeply insightful works, like EDU scholar and complexity theorist James Gardner's *Biocosm* (2003), which some critics read as an attempt to "split the difference" between a God-created and self-organized universe, run into trouble when they speculate that our universe may have been rationally constructed (read: intelligently designed) by "godlike" entities in a previous cycle. Such models simply don't fit with all materialist experience to date with respect to intelligence's role in replicating systems within our own universe.

Just as life's incredibly adapted complexity self-organized over many evo devo cycles, and just as everything that is complex and adaptive inside our universe is a replicating system, it is most parsimonious to assume that our universe is a replicating evo devo system as well. If it is, its evo devo intelligence will always remain a limited and incomplete aid to selection, not a "godlike" designer. We may think a highly adapted design offers evidence of a designer (Paley 1802) but this argument has been exhaustively refuted by the rise of evolutionary theory with respect to biological systems (Darwin 1859), and we should expect it to be defeated for an evo devo universe as well.

If our universe replicates, either in isolation or as part of a fractally replicating multiverse, as some cosmologists propose, evolutionary developmental self-organization under selection seems the simplest explanation for such curious universal features as our improbably fine-tuned initial conditions, the robust emergence of adaptive complexity and intelligence, our improbably self-correcting geophysical environment, our continually accelerating complexity on Earth (Sagan's "Cosmic calendar" metaphor), even under periodic catastrophe, and other puzzling aspects of our complexity emergence story so far. We have no need to invoke supernatural entities to explain such phenomena, and we have found no credible evidence, in our five hundred year epic of science advancement, for an intelligent designer.

20 Research Questions

- How do we best improve our physical and informational theories of unpredictable evolutionary and predictable developmental process?
- What improvements to complex systems theory, evo-devo theory, adaptation theory, optimization theory, acceleration theory, intelligence theory, hierarchy theory, life cycle theory, and other topics will help us better define, delineate, and compare evo and devo process in all replicating complex systems?
- How can we better define evolutionary and developmental process as sources of intelligence, in seeds (containing initiating evo and devo parameters), organisms, and environments?
- What evo and devo goals (purposes, telos) can we discover for intelligent complex systems?
- To what extent can we find modularity, reaction-diffusion systems, and other features of organismic ED in ecosystem ED? In biogeographical ED? In stellar-planetary ED? In galactic ED?
- What empirical and statistical tools and tests can help us to infer developmental processes in biology, based on past experience with other organisms, when we do not have the capacity to simulate development causally? Can we use those tools and tests to help us infer hierarchy and life cycle in the universe as well?
- How do we best improve our models, simulations, and tests, especially falsification tests, for the universe as an evo devo system?

Acknowledgments The author thanks Evo Devo Universe co-directors Clement Vidal, Georgi Yordanov Georgiev, Michael Price, and Claudio Flores-Martinez for helpful critiques. Special thanks go to EDU member Milan Ćirković who offered extensive constructive feedback on the earliest version of this paper. Thanks also to Anthony Aguirre, John Leslie, Denis Noble, Rüdiger Vaas and Tyler Volk for key insights, and to Carlos Gershenson and the CCS committee for approving our satellite on Evolution, Development, and Complexity at CCS2017, where these and other ideas were discussed.

References

Adami, Chris (2016) What is Information? *Phil. Trans. Royal Soc. A* 374(2063)20150230.

Adams, Fred C. (2008) Stars in Other Universes: Stellar structure with different fundamental constants, arXiv:0807.3697 [astro-ph].

Adams, A.M., Zenil, H., Davies, P.C.W., and Walker, S.I. (2016) Formal Definitions of Unbounded Evolution and Innovation Reveal Universal Mechanisms for Open-Ended Evolution in Dynamical Systems (preprint, ArXiv.org).

Aguirre, Anthony (2016) Why There is "Something" rather than "Nothing", Interview by Lawrence Kuhn for *Closer to Truth* (YouTube, 4 min).

Arthur, Wallace (2000) *The Origin of Animal Body Plans: A Study in Evolutionary Developmental Biology*, Cambridge U. Press.

Aunger, Robert (2007a) Major transitions in 'big' history, *Technological Forecasting and Social Change* 74(8):1137–1163.

——— (2007b) A rigorous periodization of ‘big’ history, *Technological Forecasting and Social Change* 74(8):1164–1178.

Barrow, John D. and Tipler, Frank (1986) *The Anthropic Cosmological Principle*, Oxford U. Press.

Barrow, John D. et al. (2008) *Fitness of the Cosmos for Life: Biochemistry and Fine-Tuning*, Cambridge U. Press.

Bejan, A. and Errera, M.R. (2017) Wealth inequality: The physics basis. *J. Applied Physics* 121(12):124903.

Bejan, A. and Zane, J.P. (2013) *Design in Nature: How the Constructal Law Governs Evolution in Biology, Physics, Technology, and Social Organizations*, Anchor.

Bhullar, B-A.S. (2017) Evolution: Catastrophe triggers diversification, *Nature* 542:304–305.

Bostrom, N. (2002) *Anthropic Bias: Observation Selection Effects in Science and Philosophy*, Routledge.

Bostrom, N. and Ćirković, M.M. (Eds.) (2008) *Global Catastrophic Risks, 1st Edition*, Oxford U. Press.

Bourgine, P. and Stewart, J. (2004) Autopoiesis and Cognition, *Artificial Life* 10: 327–345.

Bray, Dennis (2011) *Wetware: A Computer in Every Living Cell*, Yale U. Press.

Brin, David (1983) The ‘Great Silence’: the Controversy Concerning Extraterrestrial Intelligent Life, *Q.J.R. Astr. Soc.* 24:283–309.

Callebaut, Werner and Rasskin-Gutman, Diego (2005) *Modularity: Understanding the Development and Evolution of Natural Complex Systems*, MIT Press.

Calude, Cristian S. and Jürgensen, Helmut (2005) Is complexity a source of incompleteness? *Advances in Applied Math* 35(2005)1–15.

Carroll, Sean B. (2005) *Endless Forms Most Beautiful: The New Science of Evo Devo*, Norton.

Carroll, Sean M. (2016) *The Big Picture: On the Origins of Life, Meaning, and the Universe Itself*, Dutton.

Carter, Brandon (1974) Large Number Coincidences and the Anthropic Principle in Cosmology. *IAU Symposium 63*, Reidel. pp. 291–298.

Chaitin, Gregory J. (1992) *Information-Theoretic Incompleteness*, World Scientific.

Chaisson, Eric (2001) *Cosmic Evolution: The Rise of Complexity in Nature*, Harvard U. Press.

Ćirković, M.M. (2009) Fermi’s Paradox – The Last Challenge for Copernicanism? *Serbian Astronomical Journal* 178, 1–20.

Ćirković, M.M., Sandberg, A., and Bostrom, N. (2010) Anthropic Shadow: Observation Selection Effects and Human Extinction Risks, *Risk Analysis* 30:1495–1506.

Clausius, Rudolf (1851) On the Moving Force of Heat, and the Laws regarding the Nature of Heat itself which are deducible therefrom, *London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 4th. 2(VIII):1–21;102–119.

Comte, August (1844) *Discours sur l’Esprit Positif [A General View of Positivism]*, Paris.

Conway-Morris, Simon (1998) *The Crucible of Creation: The Burgess Shale and the Rise of Animals*, Oxford U. Press.

——— (2004) *Life’s Solution: Inevitable Humans in a Lonely Universe*, Cambridge U. Press.

——— (2015) *The Runes of Evolution*, Templeton Press.

Corning, Peter A. and Szathmáry, Eörs (2015) “Synergistic selection”: A Darwinian frame for the evolution of complexity. *J. Theoretical Biology* 371(2015)45–58.

Corning, Peter A. (2018) *Synergistic Selection: How Cooperation has Shaped Evolution and the Rise of Humankind*, World Scientific.

Crane, L. (1994) Possible Implications of the Quantum Theory of Gravity: An Introduction to the Meduso-Anthropic Principle (PDF). *Foundations of Science* Preprint 1994, no. Special Issue of the First Conference on the Evolution and Development of the Universe (EDU-2008).

Crutchfield, J.P. and van Nimwegen, E. (2002) The Evolutionary Unfolding of Complexity. In: *Evolution as Computation*, Landweber L.F., Winfree E. (Eds.), Natural Computing Series, Springer.

Dartnell, Lewis (2014) *The Knowledge: How to Rebuild Our World from Scratch*, Penguin Press.

Darwin, Charles (1859) *(On) The Origin of Species (by Means of Natural Selection)*, John Murray Press.

Davies, Paul (2010) *The Eerie Silence: Renewing Our Search for Alien Intelligence*, Houghton Mifflin Harcourt.

Dick, Steven J. (1996) *The Biological Universe: The Extraterrestrial Life Debate and the Limits of Science*, Cambridge U. Press.

Dryden, D.T.F., Thomson, A.R., and White, J.H. (2008) How much of protein sequence space has been explored by life on Earth?, *J. R. Soc. Interface* 5:953–956.

Einstein, Albert (1915) Die Feldgleichungen der Gravitation (The Field Equations of Gravitation), *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin*, 844–847.

Eldredge, N. and Gould, S.J. (1972) Punctuated equilibria: an alternative to phyletic gradualism, In: *Models in Paleobiology*, T.M. Schopf (ed.), Freeman & Cooper, pp. 82–115.

Ellis, G.F.R. (2015) Recognising Top-Down Causation, in A. Aguirre et al. (Eds.) *Questioning the Foundations of Physics*, Springer, pp. 17–44.

Ellis, G.F.R., Noble, D., and O'Connor, T. (2012) Top-down causation: an integrating theme within and across the sciences?, *Interface Focus* 2:1–3.

Fischler, Martin A. and Firschein, Oscar (1987) *Intelligence: The Eye, the Brain, and the Computer*, Addison-Wesley.

Flack, Jessica (2017) Life's Information Hierarchy. In: *From Matter to Life: Information and Causality*, Walker, Sara I. et al. (Eds.), Cambridge U. Press.

Flores-Martinez, Claudio L. (2014) SETI in the light of cosmic convergent evolution. *Acta Astronautica* 104(1):341–349.

——— (2017) Introducing Biomimomics: Combining Biomimetics and Comparative Genomics for Constraining Organismal and Technological Complexity. In: Mangan M. et al. (Eds.) *Biomimetic and Biohybrid Systems: Proceedings of Living Machines 2017*, Springer.

Forgan, D., Dayal, P., Cockell, C., and Libeskind, N. (2017) Evaluating galactic habitability using high-resolution cosmological simulations of galaxy formation, *International Journal of Astrobiology* 16:60–73.

Forward, Robert L. (1980) *Dragon's Egg*, Del Rey.

Gardner, Andy and Conlon, Joseph P. (2013) Cosmological natural selection and the purpose of the universe, *Complexity* 18:48–56.

Gardner, James (2003) *Biocosm: A New Scientific Theory of Evolution*, Inner Ocean Publishing.

——— (2007) *The Intelligent Universe: AI, ET, and the Emerging Mind of the Cosmos*, New Page.

Georgiev, Georgi Y. et al. (2015) Mechanism of organization increase in complex systems, *Complexity* 21(2)18–28.

Gerhart, John C. and Kirschner, Marc W. (2005) *The Plausibility of Life: Resolving Darwin's Dilemma*, Yale U. Press.

——— (2007) The theory of facilitated variation, *PNAS* 104 Suppl 1:8582–9.

Gilbert, S.F., Bosch, T.C.G., and Ledón-Rettig, C. (2015) Eco-Evo-Devo: Developmental symbiosis and developmental plasticity as evolutionary agents, *Nature Reviews Genetics* 16(10):611–622.

Gödel, Kurt (1931) On formally undecidable propositions of *Principia. Mathematica* and related systems (German), *Monatshefte für Mathematik und Physik*, 38(1)173–198.

Goodwin, Jay T. et al. (2014) *Alternative Chemistries of Life: Empirical Approaches*, NASA/NSF Workshop.

Gould, Stephen J. (1977) *Ontogeny and Phylogeny*, Harvard U. Press.

——— (2002) *The Structure of Evolutionary Theory*, Harvard U. Press.

Haken, Hermann (1984) *The Science of Structure: Synergetics*, Van Nostrand Reinhold.

Hall, Brian K. (Ed.) (2003) *Environment, Development, and Evolution*, MIT Press.

Harrison, Edward R. (1995) The Natural Selection of Universes Containing Intelligent Life (PDF). *Quarterly Journal of the Royal Astronomical Society* 36(3)193–203.

Henderson, Lawrence J. (1913) *The Fitness of the Environment: An Inquiry into the Biological Significance of the Properties of Matter*, Macmillan.

Heylighen, Francis (2007) The Global Superorganism: an evolutionary-cybernetic model of the emerging network society. Web published manuscript. Retrieved from: http://pespmc1.vub.ac.be/Papers/Superorganism.pdf

——— (2008) Accelerating socio-technological evolution: from ephemeralization and stigmergy to the Global Brain. In: *Globalization as Evolutionary Process: Modeling Global Change*, George Modelski, Tessaleno Devezas, and William R. Thompson (eds.), Routledge.

——— (2016) Stigmergy as a Universal Coordination Mechanism. *Cognitive Systems Research* 38:4–13.

Hoyle, Fred (1954) On nuclear reactions occurring in very hot stars: the synthesis of the elements from carbon to nickel, *Astrophysics Journal* Suppl. 1:121–146.

——— (1983) *The Intelligent Universe: A New View of Creation and Evolution*, Holt, Rinehart and Winston.

Jablonka, Eva and Lamb, Marion J. (1995) *Epigenetic Inheritance and Evolution: the Lamarckian Dimension*, Oxford U. Press.

——— (2005) *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*, MIT Press.

Jablonski, D. (1986) Background and Mass Extinctions: The Alternation of Macroevolutionary Regimes, *Science* 231:129–133.

Jacobs, Lucia (2012) From Chemotaxis to the Cognitive Map: The function of olfaction. *PNAS*, 109(1):10693–10700.

Jantsch, Erich (1980) *The Self-Organizing Universe: Scientific and Human Implications*, Pergamon.

Johnson, Norman L. (2011) What a Developmental View Can Do for You. In: *Thought Leader Forum*, Michael J. Mouboussin (Ed.), Credit Suisse/First Boston.

Kaplan, Jonathan (2008) The end of the adaptive landscape metaphor? *Biol Philos* 23:625–638.

Kauffman, Stuart (1993) *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford U. Press.

Koonin, Eugene V. (2007) The cosmological model of eternal inflation and the transition from chance to biological evolution in the history of life. *Biology Direct* 2:15–36.

Krumholz, M.R. and McKee, C.F. (2005) A general theory of turbulence-regulated star formation, from spirals to ultraluminous infrared galaxies. *The Astrophysical Journal* 630:250–268.

Kurzweil, Ray (1999) *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*, Penguin.

——— (2005) *The Singularity is Near*, Penguin.

Laland, Kevin et al. (2015) The extended evolutionary synthesis: Its structure, assumptions and predictions, *Proceedings of the Royal Society* B282:20151019.

Lane, Nick (2016) *The Vital Question*, W.W. Norton.

Laplace, Pierre-Simon (1812) *Theorie analytique des probabilites (A Philosophical Essay on Probabilities)*, Translated from the fifth French edition of 1825 with Notes by the Translator, Andrew I. Dale, Springer-Verlag, 1995.

Lem, Stanislaw (1971) The New Cosmogony, In: *A Perfect Vacuum* (trans. by M. Kandel), Northwestern U. Press, 1999, pp. 197–227.

Leibniz, Gottfried (1686) *The Discourse on Metaphysics (Fr: Discours de métaphysique).*

Leslie, John and Kuhn, Robert L. (2013) *The Mystery of Existence: Why Is There Anything At All?*, Wiley-Blackwell.

Levin, Simon A. (1998) Ecosystems and the Biosphere as Complex Adaptive Systems, *Ecosystems* 1:431–436.

Lewis, G.F. and Barnes, L.A. (2016) *A Fortunate Universe: Life in a Finely Tuned Cosmos*, Cambridge U. Press.

Lightman, Alan (2011) The Accidental Universe, *Harpers*, Dec 2011.

Lineweaver, Charles H., Fenner, Y., and Gibson, B.K. (2004) The Galactic Habitable Zone and the Age Distribution of Complex Life in the Milky Way, *Science* 303(5654):59–62.

Linde, Andrei D. (1992) Stochastic approach to tunneling and baby universe formation, *Nuclear Physics B* 372:421–442.

Lobkovsky, A.E. and Koonin, E.V. (2012) Replaying the tape of life: quantification of the predictability of evolution, *Frontiers in Genetics* 3:246–254.

Lorenz, Konrad (1977) *Behind the Mirror: A Search for a Natural History of Human Knowledge*, Harcourt, Brace, Jovanovich.

Losos, Johnathan B. (2017) *Improbable Destinies: Fate, Chance, and the Future of Evolution*, Riverhead Books.

Louis, A.A. (2016) Contingency, convergence and hyper-astronomical numbers in biological evolution, Studies in History and Philosophy of Biological and Biomedical Sciences 58:107–116.

Lucas, J.R. (1961) Minds, Machines and Gödel. *Philosophy* 36:112–137.

Luhmann, Niklas (2002/2013) *Introduction to Systems Theory*, Polity Press.

Luisi, Pier L. (2003) Autopoiesis: a review and a reappraisal. *Naturwissenschaften* 90 49–59.

Malone, Michael S. (2012) *The Guardian of All Things: The Epic Story of Human Memory*, St. Martin's Press.

Mariscal, Carlos (2016) Universal biology: assessing universality from a single example, In: *The Impact of Discovering Life beyond Earth*, Steven J. Dick (ed.), Cambridge U. Press.

Martin, Joseph D. (2013) Is the Contingentist/Inevitabilist Debate a Matter of Degrees? Philosophy of Science 80(5):919–930.

Maturana, H.R. and Varela, F.J. (1973/1980) *Autopoiesis and Cognition: The Realization of The Living*, D. Reidel.

——— (1987) *The Tree of Knowledge: The Biological Roots of Human Understanding*, Shambhala.

McGhee, George R. (2011) *Convergent Evolution: Limited Forms Most Beautiful*, MIT Press.

McLeish, T.C.B. (2015) Are there ergodic limits to evolution? Ergodic exploration of genome space and convergence, *Interface Focus* 5:41–53.

Meissner, Ulf-G. (2013) Life on earth – An accident? Chiral symmetry and the anthropic principle, *Int. J. Mod. Phys. E* 23.

Miller, James G. (1978) *Living Systems*, McGraw Hill.

Mingers, John (1995) *Self-Producing Systems*. Kluwer Academic/Plenum.

Moravec, Hans (1979) Today's Computers, Intelligent Machines, and Our Future. *Analog Science Fiction and Fact*, Feb 1979.

Müller, Gerd B. and Newman, Stuart A. (Eds.) (2003) *The Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology*, MIT Press.

Munitz, Milton K. (1987) *Cosmic Understanding: Philosophy and Science of the Universe*, Princeton U. Press.

Meyers, Robert A. (2009) *Encyclopedia of Complexity and Systems Science*, Springer.

Newton, Isaac (1687) *The Mathematical Principles of Natural Philosophy*, Daniel Adee, 1846.

Noble, Denis (2017) Evolution viewed from physics, physiology and medicine. *Interface Focus* 7(5):20160159.

Noble, Raymond and Noble, Denis (2017) Was the Watchmaker Blind? Or Was She One-Eyed? *Biology* 6(4):47 (Special Issue: Biology in the Early 21st Century: Evolution Beyond Selection).

Odling-Smee, John et al. (2003) *Niche Construction: The Neglected Process in Evolution*, Princeton U. Press.

Ogura A., Kazuho, I., and Gojobori, T. (2004) Comparative analysis of gene expression for convergent evolution of camera eye between octopus and human. *Genome Research* 14:1555–1561.

O'Malley, M.A. and Powell, R. (2016) Major problems in evolutionary transitions: how a metabolic perspective can enrich our understanding of macroevolution, *Biology and Philosophy* 31:159–189.

Oparin, Alexander I. (1968) *Genesis and Evolutionary Development of Life*, Academic Press.

Orgogozo, V. (2015) Replaying the tape of life in the twenty-first century, *Interface Focus* 5:57–68.

Paley, William (1802) *Natural Theology or Evidences of the Existence and Attributes of the Deity*, John Morgan Press.

Pearce, T. (2012) Convergence and Parallelism in Evolution: A Neo-Gouldian Account, *Brit. J. Phil. Sci.* 63:429–448.

Piel, Gerard (1972) *The Acceleration of History*, Knopf.

Pigliucci, M (2007) Do we need an extended evolutionary synthesis?, *Evolution* 61:2743–2749.

Pigliucci, M. and Müller, G.B. (Eds.) (2010) *Evolution: The Extended Synthesis*, MIT Press.

Pinker, Steven (2010) *The Better Angels of Our Nature: Why Violence Has Declined*, Penguin.

Parker, Andrew (2003) *In the Blink of an Eye*, Basic Books.

Powell, R. (2012) Convergent evolution and the limits of natural selection, *Euro. J. Phil. Sci.* 2:355–373.

Powell, R. and Mariscal, C. (2015) Convergent evolution as natural experiment: the tape of life reconsidered, *Interface Focus* 5:40–53.

Pohorille, Andrew (2012) Processes that Drove the Transition from Chemistry to Biology: Concepts and Evidence. *Origins of Life* 42(5)429–432.

Poundstone, William (1985) The Recursive Universe: Cosmic Complexity and the Limits of Scientific Knowledge, William Morrow & Co.

Price, Michael (2017) Entropy and selection: Life as an adaptation for universe replication (PDF). *Complexity*, Article ID 4745379, 4 pages, 2017. doi:https://doi.org/10.1155/2017/4745379.

Pross, Addy (2014) *What is Life? How Chemistry Becomes Biology*, Oxford U. Press.

Raff, Rudolf (1996) *The Shape of Life: Genes, Development, and the Evolution of Animal Form*, U. of Chicago Press.

Ray, Georgia (2017) Evolutionary Innovation as a Global Catastrophic Risk, Presentation at EA Global 2017, San Francisco, CA.

Rees, Martin (1999) *Just Six Numbers: The Deep Forces that Shape Our Universe*, Basic Books.

——— (2001) *Our Cosmic Habitat*, Princeton U. Press.

——— (2015) Post-human evolution on Earth and beyond. Presented at *Exploring Exoplanets: The Search for Extraterrestrial Life and Post Biological Intelligence*, John Templeton Foundation conference.

Reid, Robert G.B. (2007) *Biological Emergences: Evolution by Natural Experiment*, MIT Press.

Rosenstein, Melissa G. et al. (2012) Risk of Stillbirth and Infant Death Stratified by Gestational Age, *Obstet Gynecol* 120(1):76–82.

Russell, Cathy M. (2006) Retrieved from Epicofevolution.com/biological-evolution, 27 Mar 2017.

Russell, D. A. and Séguin, R. (1982) Reconstructions of the small Cretaceous theropod *Stenonychosaurus inequalis* and a hypothetical dinosauroid. Ottawa: National Museums of Canada, National Museum of Natural Sciences.

Sagan, Carl (1977) *The Dragons of Eden: Speculations on the Evolution of Human Intelligence*, Random House.

——— (1980) *Cosmos*, Random House.

Salthe, Stanley M. (1985) *Evolving Hierarchical Systems*, Columbia U. Press.

——— (1993) *Development and Evolution: Complexity and Change in Biology*, MIT Press.

——— (2010) Development (and evolution) of the universe. In: *The Evolution and Development of the Universe*, C. Vidal (ed.), *Foundations of Science* 15:357–367.

——— (2012) Hierarchical structures, *Axiomathes* 22:355–383.

Sanderson, M. and Hufford, L. (Eds.) (1996) *Homoplasy: The Recurrence of Similarity in Evolution*, Academic Press.

Schlosser, G. and Wagner, G.P. (Eds.) (2004) *Modularity in Development and Evolution*, U. of Chicago Press.

Shapiro, James A. (2011) *Evolution: A View From the 21st Century*, FT Press Science.

Signor, P.W. and Lipps, J.H. (1982) Sampling bias, gradual extinction patterns, and catastrophes in the fossil record, In: *Geological implications of impacts of large asteroids and comets on the Earth*, L.T. Silver and P.H. Schultz (Eds.), Geological Society of America Special Publication, 190:291–296.

Smart, John M. (2000) The developmental singularity hypothesis (DSH), AccelerationWatch.com. Retrieved from: Accelerationwatch.com/developmentalsinghypothesis.html, 29 Jan 2018.

——— (2002) Understanding STEM compression in universal change, AccelerationWatch.com. Retrieved from: Accelerationwatch.com/mest.html, 29 Jan 2018.

——— (2008) Evo Devo Universe? A Framework for Speculations on Cosmic Culture. In: *Cosmos and Culture: Cultural Evolution in a Cosmic Context*, Steven J. Dick, Mark L. Lupisella (Eds.), NASA Press.

——— (2012) The transcension hypothesis: Sufficiently advanced civilizations invariably leave our universe, and implications for METI and SETI, *Acta Astronautica* 78:55–68.

——— (2015) Humanity rising: Why evolutionary developmentalism will inherit the future, *World Future Review* 7(2–3):116–130.

——— (2016a) Portals (funnels, bottlenecks) and convergent evolution. In: *The Foresight Guide*. Retrieved from Foresightguide.com/portals-funnels-bottlenecks-and-convergent-evolution, 20 Jan 2018.

——— (2016b) The great race to inner space. In: *The Foresight Guide*. Retrieved from: Foresightguide.com/the-great-race-to-inner-space-our-surprising-future/, 28 Mar 2017.

——— (2017a) The VCRIS model of natural selection. In: *The Foresight Guide*. Retrieved from: Foresightguide.com/the-vcris-model-of-natural-selection-evolutionary-development-of-adaptive-complexity, 6 Jun 2018.

——— (2017b) Five goals of complex systems. In: *The Foresight Guide*. Retrieved from: Foresightguide.com/five-goals-of-complex-systems, 30 Dec 2017.

——— (2018) Catalytic catastrophes: Advancing the five goals. In: *The Foresight Guide*. Retrieved from: Foresightguide.com/catalytic-catastrophes-how-right-sized-catastrophes-advance-the-five-goals/, 20 Jan 2018.

Smith, Eric and Morowitz, Harold J. (2004) Universality in intermediary metabolism, *Proc. Natl. Acad. Sci.* 101:13168–13173.

——— (2016) *The Origin and Nature of Life on Earth*, Cambridge U. Press.

Smolin, Lee (1992) Did the Universe Evolve? *Classical and Quantum Gravity* 9:173–191.

——— (1997) *The Life of the Cosmos*, Oxford U. Press.

——— (2004) Cosmological natural selection as the explanation for the complexity of the universe. *Physica A* 340:705–713.

——— (2006) The status of cosmological natural selection. arXiv:hep-th/0612185v1.

——— (2012) Scientific Approaches to the Fine-Tuning Problem, *The Nature of Reality*, Nova.

Spencer, Herbert (1864) *Illustrations of Universal Progress: A Series of Discussions*, D. Appleton.

Steele, Edward J. (1981) *Somatic Selection and Adaptive Evolution: On the Inheritance of Acquired Characters, 2nd Ed.*, U. of Chicago Press.

——— (1998) *Lamarck's Signature: How retrogenes are changing the natural selection paradigm*, Perseus.

Stenger, Victor J. (2011) *The Fallacy of Fine-Tuning: Why the Universe is Not Designed for Us*, Prometheus.

Susskind, Leonard (2006) *The Cosmic Landscape*, Back Bay.

Swenson, Rod (1992) "Galileo, Babel, and Autopoiesis (It's Turtles All The Way Down)". *Int. J. General Systems* 21 (2): 267–269.

Taleb, Nicholas (2012) *Antifragile: Things That Gain from Disorder*, Random House.

Teilhard de Chardin, Pierre (1955) *The Phenomenon of Man*, Harper & Row.

Turchin, Valentin (1977) *The Phenomenon of Science*, Columbia U. Press.

Turner, D.D. (2011) Gould's replay revisited, *Biology and Philosophy* 26:65–79.

Tyson, Neal D. (2006) The Perimeter of Ignorance, Presentation at: *Beyond Belief Conference*, Nov 5–7, 2006, YouTube (41 mins).

Vaas, Rüdiger (1998) Is there a Darwinian evolution of the cosmos? In: *Proceedings of the MicroCosmos-MacroCosmos Conference*, Aachen.

Varela F.J., Maturana, H.R., and Uribe, R. (1974) Autopoiesis: the organization of living systems, its characterization and a model, *Biosystems* 5(4):187–196.

Verhulst, Jos (2003) *Discovering Evolutionary Principles through Comparative Morphology*, Adonis Press.

Vermeij, Geerat J. (1987) *Evolution and Escalation*, Princeton U. Press.

——— (2006) Historical contingency and the purported uniqueness of evolutionary innovations, *PNAS* 103:1804–1809.

——— (2009) *Nature: an economic history*, Princeton U. Press.

Vidal, Clement (2008a) The future of scientific simulations: from artificial life to artificial cosmogenesis (PDF). In: *Death and Anti-Death 6: Thirty Years After Kurt Gödel (1906–1978).*

——— (2010) Computational and biological analogies for understanding fine-tuned parameters in physics (PDF). *Foundations of Science* 15(4):375–393.

——— (2016) Stellivore Extraterrestrials? Binary Stars as Living Systems (PDF). *Acta Astronautica* 128:251–56.

Vilenkin, Alexander (2006) *Many Worlds In One: The Search for Other Universes*, Hill and Wang.

Volk, Tyler (2003) *Gaia's Body: Toward a Physiology of Earth*, MIT Press.

——— (2017) *Quarks to Culture: How We Came to Be*, Columbia U. Press.

Wagner, Gunther (2003) *Hox* cluster duplications and the opportunity for evolutionary novelties, *PNAS* 100(25):14603–14606.

Wagman and Stephens (2004) Surprising 'ultra-conserved' regions discovered in human genome. *UCSC Currents*.

Walker, Sara I. et al. (Eds.) (2017) *From Matter to Life: Information and Causality*, Cambridge U. Press.

Webb, Stephen (2015) *Where is Everybody? Seventy-Five Solutions to the Fermi Paradox and the Problem of Extraterrestrial Life*, Springer.

Weinberg, Steven (2007) "Living in the Multiverse", In: *Universe or Multiverse?* B. Carr (ed.), Cambridge U. Press.

West, Geoffrey (2017) *Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies*, Penguin.

West-Eberhard, Mary Jane (2003) *Developmental Plasticity and Evolution*, Oxford U. Press.

Wheeler, John A. (1977) Genesis and observership, In: *Foundational Problems in the Special Sciences*, R. E. Butts and J. Hintikka, (Eds.), D. Reidel, pp. 3–33.

——— (1988) World as system self-synthesized by quantum networking, *IBM Journal of Research and Development* 32:4–15.

Wigner, Eugene (1960) The unreasonable effectiveness of mathematics in the natural sciences. *Comm. on Pure and Applied Math.*, 13(1):1–14.

Wikipedia (2008b), and Smart, John M. and Vidal, Clement (2008–2017). Cosmological natural selection (fecund universes). EvoDevoUniverse.com wiki.

Wikipedia (2012), and Smart John M. and Chattergee, Atanu (2012–2017). List of examples of convergent evolution. EvoDevoUniverse.com wiki.

Wilkins, Adam S. (2001) *The Evolution of Developmental Pathways*, Sinauer Associates.

Wright, Sewall (1932) The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the Sixth International Congress on Genetics*. pp. 355–366.

Yi, Hong et al. (2010) Gene expression atlas for human embryogenesis. *FASEB Journal* 24(9):3341–3350.

Zimmer, Carl (2015) *A Planet of Viruses, 2nd Edition*, U. Chicago Press.

Life, Intelligence, and the Selection of Universes

Rüdiger Vaas

> *Many worlds might have been botched and bungled, throughout an eternity, 'ere this system was struck out. Much labour lost: Many fruitless trials made: And a slow, but continued improvement carried out during infinite ages in the art of world-making.*
>
> —David Hume (1779)

The far future of our universe and its beginning raise deep and difficult questions: How will it evolve and perhaps end, and why did the big bang occur in the first place? Furthermore, the possibility of life and intelligence is closely connected to these questions: Their final fate, if unchallenged, appears deadly dark, and their origin and continuation depend on specific boundary conditions as well as the laws and constants of nature, which seem to be special or extremely improbable.

The universe continually evolves and develops as life on Earth and human cultures do. The causes of these self-organized processes are different but depend on each other at least in one direction; and they also have some similarities, for example, the increase of complexity (see Chaisson 2001, 2011). Such an "evo devo" perspective (evolution and development) can offer new insights not only regarding underlying mechanisms of those processes but also in respect of the far future of the universe and intelligent beings – and perhaps even the nature of nature (see, e. g., Smart 2008 and this volume). This might be inevitable if there are giant feedback loops from technological cultures to galactic development at some point. Because if life and intelligence, including postbiological descendants (Dick 2003, 2008, 2009; Smart 2012; Sandberg et al. 2016; Vaas 2017a, b, 2018a), resist to be ultimately doomed on their planetary surfaces, radical new options have to be envisaged. Changing the future of the universe at large scales is an extreme possibility – if it is a possibility at all. But it is also a fascinating topic to speculate about, for it might additionally reveal something about the past and the fundamental properties of the universe as well as the boundary conditions of life.

R. Vaas (✉)
bild der wissenschaft, Leinfelden-Echterdingen, Germany

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_3

1 Is Our Universe Fine-Tuned?

Life as we know it depends crucially on the laws and constants of nature as well as the boundary conditions (e.g., Barrow and Tipler 1986; Leslie 1989; Vaas 2004a; Carr 2007; Barrow et al. 2008; Barnes 2012). Nevertheless, it is difficult to judge how fine-tuned it really is, both because it is unclear how modifications of many values together might compensate each other (see Aguirre 2001; Harnik et al. 2006; Jaffe et al. 2009; Stenger 2011; Fedrowa and Griest 2014 for interesting examples) and whether laws, constants, and initial conditions really could have been otherwise to begin with. It is also unclear how specific and improbable those values need to be for the development of information-processing structures – and, hence, intelligent observers. At least for life on Earth, as we know it, specific values of physical and cosmological parameters are necessary – this is an *empirical fact*, and it is not trivial but surprising in its far-reaching depth (and it was therefore also a matter of scientific discovery). If we accept, for the sake of argument, that at least some values are fine-tuned, we must ask how this can be explained.

Sometimes, fine-tuning is conflated with the anthropic principle (AP). This is confusing and misleading, because there are many AP versions – or, indeed, principles – and some are meant as explanations, some are as implications, some are mere consistency conditions, some are metaphysical postulates, and some are almost independent of fine-tuning issues (Barrow and Tipler 1986; Vaas 2004a). Thus, Nick Bostrom's lamentation (2002, p. 6) is justified: "A total of over thirty anthropic principles have been formulated and many of them have been defined several times over – in nonequivalent ways – by different authors, and sometimes even by the same authors on different occasions. Not surprisingly, the result has been some pretty wild confusion concerning what the whole thing is about."

And it is not just the issue of fine-tuning with its existential aspects, which demands explanation. Other features of physics and cosmology, which appear special or improbable, also enter the equation. One might leave the topic of anthropic reasoning aside altogether and can still ask, why the laws and constants of nature (as well as the boundary conditions) are the way they are. Therefore most of the explanatory possibilities and problems (summarized in Table 1) still apply. These are meaningful questions. Whether they can be partly answered in practice or in principle is an open issue.

It was argued (see, e.g., Callender 2004; Manson 2000; McGrew et al. 2001; Mosterín 2005) that searching for such explanation is either a fruitless waste of cognitive efforts and time or even illegitimate as long as the explanandum is not well-defined or because extravagant proposals belong to metaphysics, not serious science. Furthermore, it seems even inevitably that at some point, issues get too complex and difficult to be answered or understood (even if a supercomputer or an oracle would reveal the truth); perhaps we can never overcome some kinds of cognitive closure, which Colin McGinn (1989, p. 350) defined as follows: "A type of mind M is cognitively closed with respect to a property P (or a theory T), if and only if the concept-forming procedures at M's disposal cannot extend to a

Table 1 Digging deeper: Laws, constants, and boundary conditions are the basic constituents of cosmology and physics from a formal point of view (besides spacetime, energy, matter, fields, and forces or more fundamental entities like strings or spin-networks and their properties with regard to content). An ambitious goal and historically at least a successful heuristic attitude are reduction, derivation, and unification to achieve more fundamental, far-reaching, and simple descriptions and explanations. While uniqueness is much more economical and predictive, multiple realizations – presumably within a multiverse – have recently been proposed as an opposing (but not mutually exclusive) alternative. This table (adapted from Vaas 2015, p. 72) provides a summary of different approaches, possibilities, and problems; it is neither complete nor the only conceivable system

	Fundamental laws (L) of nature	Fundamental constants (C)	Boundary conditions (BC)
Uniqueness? (and just one universe?)	(1) *Irreducible* and disconnected? (2) Or *derived* from or unified within or reducible to one fundamental theory (“Theory of Everything,” TOE)? “Logically isolated”? Or even logically sufficient (self-consistent)? (bootstrap principle) Ultimately (logically?) deducible? (thus without empirical content if analytically true? Natural science as pure mathematics?) (3) Or *nonexistent* because emerging via self-organization from an underlying chaos (“law without law” approach)	(1) *Irreducibly* many? (2) Or *only a few or just one* (e.g., string length)? With a unique value? (just random or determined by L?) With (infinitely?) many possible different values? (according to a probability distribution determined by L?) (3) Or *ultimately none*? Because C is just conversion factors (e.g., in a TOE) And/or better understandable as initial conditions (with many different realizations in a multiverse?)	(1) As *initial conditions*? Irreducible? Not accessible? (due to inflation, mixmaster universe, BKL chaos, “cosmic forgetfulness,” decoherence, etc.) Nonexistent (in eternal universe models)? Explanatorily irrelevant? (because replaced by present BC or convergent due to an attractor?) (2) Or as *present conditions*? Sufficient? Or the only useful ones? (e.g., as loop quantum cosmology constraints or according to the top down approach in Euclidean quantum gravity) Or as (additional?) final conditions? (e.g., constraints in some quantum cosmology models) (3) Or *nonexistent* because determined by or interpreted as L? (e.g., the no-boundary proposal)

(continued)

Table 1 (continued)

	Fundamental laws (L) of nature	Fundamental constants (C)	Boundary conditions (BC)
Multiverse?	(1) *Many realizations* As BC if not reducible Separated in principle or with a common origin (cause)? Restricted by or derived from a TOE, or truly random/irreducible? (2) Or with *every* (logically? metaphysically?) *possible realization*? (mathematical democracy, ultimate principle of plentitude) (3) TOE as a "*multiverse generator*"?	(1) *Many realizations* Truly random Restricted by L or BC? (2) Or *every possible combination* of values? (equally or randomly distributed?) (3) Or with some *absolute frequency* according to a probability distribution (determined by L, e.g., string statistics and/or within a framework like cosmological natural selection?) (4) Or just *observational bias*/selection (weak anthropic principle) from a random set?	(1) *Many realizations* Restricted variance due to an attractor (determined by L)? Or truly random (2) Or *every possible combination*? (equally or randomly distributed?) (3) Or with some *absolute frequency* according to a probability distribution (determined by L and/or within a framework like cosmological natural selection)? (4) Or just *observational bias*/selection (weak anthropic principle) from a random set?
Coevolution?	Cosmic endosymbiosis? Spacetime engineering via closed time-like curves? Participatory anthropic principle?		
Design?	Transcendent realization? (nonphysical causation) Random creation or cosmological artificial selection by cosmic engineers? Universal simulation/emulation? (or just a subjective illusion?)		
Randomness?	*Ultimately existing* (even if there is only one self-consistent TOE) Gödel–Turing–Chaitin theorems	*Not necessarily* (if determined by L or reduced to BC)	*Not necessarily* (if determined by L)

grasp of P (or an understanding of T). Conceiving minds come in different kinds, equipped with varying powers and limitations, biases and blindspots, so that properties (or theories) may be accessible to some minds but not to others."

But without seeking we can also never identify the boundaries of (our) knowledge. Suspension of judgment and explanatory nihilism is therefore a personal and often also a conventional, sociological decision, but not an objective demand or a fixed red line. Not to give up too early, but to try better, to explore even suspicious alternatives or to follow bold speculations in a confusing terrain nevertheless led sometimes to surprising progress, as history teaches. Although ultimately inexplicable contingencies will block our ardent longing for a deeper understanding (Vaas 1993) and final, self-sustaining as well as irrefutable explanations and statements are impossible (Vaas 2006a), we still can walk along the shores of the proverbial oceans of truth and knowledge, not knowing where this leads to and how it ends (Vaas 2014a).

In principle, there are many options for answering these foundational questions regarding physics and cosmology and beyond (Table 1). Fine-tuning might (1) just be an *illusion* if life could adapt to very different conditions or if modifications of many values of the constants would compensate each other; or (2) it might be a result of (incomprehensible, irreducible) *chance*, thus inexplicable; or (3) it might be *nonexistent* because nature could not have been otherwise, and with a *fundamental theory*, we would perhaps be able to prove this; or (4) it might be a product of *selection*: either *observational selection* within a vast multiverse of (infinitely?) many different realizations of those values (*weak anthropic principle*), or a kind of *cosmological natural selection* making the measured values (compared to possible other ones) quite likely within a multiverse of many different values, or even a *teleological or intentional selection*; or (5) it might be a coevolutionary development depending on a more or less goal-directed participatory contribution of life and intelligence. (There are further and even more bizarre options beyond naturalism, such as solipsism, but they will be neglected here.)

Even worse, these alternatives are not mutually exclusive – for example, it is logically possible that there is a multiverse, created according to a fundamental theory by a cosmic designer who is not self-sustaining, but ultimately contingent, i.e., an instance of chance. This might please different proponents at the same time, but it is against rational economy and explanatory parsimony. To quote Nicholas Rescher (2000, p. 8): "Never employ extraordinary means to achieve purposes you can realize by ordinary ones." Each of the proposals is strongest as sufficient and independent solution and was developed as such. However, one should not ignore combinations, especially if they are enforcing each other (such as string theory and cosmic inflation scenarios with a multiverse, cf. Clifton et al. 2007, Linde 2006, 2017).

To summarize, the reasoning goes like this:

Premise (1): There is a world w with some specific properties p.
Premise (2): X explains p or w.
Premise (3): There is X.
Conclusion: p or w is explained by X.

Here X stands for irreducible chance, for a fundamental or at least deeper theory, for a multiverse, for observational selection (weak anthropic principle), for cosmological natural selection, for cosmological artificial selection (cosmic engineers or simulation), for a kind of teleological anthropic principle (i.e., an impersonal teleological force or an intentional transcendent designer, e.g., god), or for a combination of more than one of these (e.g., observational and cosmological natural selection require the multiverse). The task for physics and cosmology is thus to find out whether those three premises are true and what p and X are.

From both a scientific and philosophical perspective, the fundamental theory approach and the multiverse scenario are most plausible and heuristically promising (Vaas 2004a, b, 2014b, 2017c).

2 Theoretical Explanation Instead of Fine-Tuning

It is a very controversial issue whether and how far reductionism works in physics – and beyond. Higher-order levels of descriptions are undoubtedly necessary for practical purposes but might still be (approximately and ontologically) reducible to lower levels (depending on certain constraints, of course, i.e., boundary conditions). Putting such issues and the ambiguous meaning of "reduction" (Vaas 1995a; van Riel and Van Gulick 2018) aside, one might argue roughly like this: The bedrock of reality consists of matter–energy, interactions, and spacetime (or even more fundamental "building blocks" such as loops or strings), the properties, states, and dynamics of which can be described by what it is called *laws and constants of nature* (physical as well as cosmological parameters) and a set of *initial or boundary conditions*. These descriptions, embedded and joined in a theory, should in principle suffice to yield explanations.

Thus, the usual scheme of explanations in physics is roughly like this: Given some boundary conditions and laws (including constants) or a theory (which connects or unifies different laws), i.e., the *explanans*, some facts or events, i.e., the *explanandum*, can be explained (*retrodiction*) or forecasted (*prediction*), and thereby the laws or theory can be tested. Different kinds of scientific explanation (e.g., Hempel 1965; Pitt 1988; Salmon 1998; Woodward 2003/2009; Lipton 2004; Mayes 2005) fall within this scheme, especially the deductive-nomological explanation (covering law model) with deterministic laws, the inductive-statistical explanation with probabilistic laws (but with any probability?), and the causal explanation focusing on cause–effect relations (which might be either deterministic or probabilistic).

This scheme works pretty well. However, one can still ask: Why these laws (or theory, respectively), why those constants, and why some particular boundary conditions? If our universe is not eternal – or at least if its laws and constants are not – these questions are related to another, namely, what is the explanation for the big bang?

But these questions differ from the usual ones concerning physical explanations, because they already presuppose or contain what should be explained. Take the big bang, for instance, i.e., the hot and dense very early universe. It is described by observations (e.g., the expansion of space, the cosmic background radiation and its properties, the ratios of light elements, etc.) and laws or theories (especially general relativity, thermodynamics, high-energy physics). But neither the big bang, nor those laws and theories, are explained (retrodicted) by those observations. The observations are explained by big bang theory, not vice versa. So how to explain the big bang, i.e., how did the hot and dense state of the very early universe come into existence? Here, new theories (or constraints of the current theories) and data are required. Some even argue that a new scheme of explanation is needed – perhaps an anthropic, functional, or even teleological one? This would be one of the largest paradigm changes since the advent of classical physics.

It is therefore wise to push the ordinary explanation scheme of physics to its limits and see how far one might really get. Thus, the explanandum now is the big bang with its (causal) connections to the present/observable universe. And the question is which explanans might suffice: which fundamental laws (e.g., of M-theory, supersymmetric grand unified quantum field theories, general relativity, etc.), fundamental physical constants (e.g., c, ħ, G), and initial conditions (e.g., dimensionality, metric, values of the fundamental fields, fluctuations etc.)?

Furthermore, one can ask whether there is a way to simplify the "triangle" of laws, constants (or parameters), and boundary conditions, i.e., to reduce one of its "corners" – or even two – to another one (Table 1). This would be a huge breakthrough in physics. Different possibilities are under consideration, but as yet, they are more or less pure speculation:

Constants might be reduced to boundary conditions: For instance, a huge "landscape" of solutions in string theory could exist, depending on different compactifications of tiny extra dimensions, etc. (Susskind 2005); if all these mathematical solutions are (or could be) physically realized, e.g., as bubble universes in the exponentially growing false vacuum of eternal inflation, then those boundary conditions, set by the phase transition to a specific "true" vacuum of an originating bubble universe, might determine what appears as constants of nature in such a universe (Linde 2005, 2008; Aguirre 2007; for an estimate of the gigantic number of different universes in the multiverse, see Linde and Vanchurin 2010). Another, not mutually exclusive scenario is cosmological natural selection (see below).

Laws might also be reduced to boundary conditions: Cosmological natural selection is an example here again, at least for some laws. And if no law and constant describing our universe is fundamental, but all are ultimately random, they can also be seen as boundary conditions in a wider sense. Thus, a specific set of laws might just be a set of boundary conditions with respect to a specific (kind of) universe within a multiverse. However, this does not necessarily imply a "law without law" approach (Wheeler 1980, 1983) in the strict sense – i.e., that the only law is that there are no (fundamental) laws, but pure randomness – because there might be a fundamental (although contingent) law nevertheless, which rules the multiverse-generating processes.

On the other hand, *boundary conditions might be reduced to laws:* A famous example is the no-boundary proposal in quantum cosmology as the quantum state of the universe (Hartle and Hawking 1983; Vaas 2018b).

Up to now, this discussion was quite abstract, surveying a range of possibilities. But there is something more to say which might add some flesh (albeit not yet very nutritious) to the backbone of particle physics and cosmology. Their standard models contain at least 31 free parameters which have to be measured and cannot be explained yet (Tegmark et al. 2006). Perhaps more advanced models or theories will reduce the number of these parameters significantly. But there is no guarantee for this. More fundamental approaches like supersymmetry might even increase the number, and perhaps further theories are needed with their own constants. However, string theory deigns to provide only very few, perhaps only two (Duff et al. 2002): the string length and the speed of light (*pacem* critics who claim that string theory rather dangles on a string). It was also argued that the number of dimensional constants like the gravitational constant G, the speed of light c, or the reduced Planck constant $\hbar = h/2\pi$ (with units $m^3\ kg^{-1}\ s^{-2}$, $m\ s^{-1}$, and $m^2\ kg\ s^{-1}$, respectively) could be dropped altogether (normalized to 1). They depend on conventions, i.e., our unit system, and are dimensional again in other unit systems – and such dimensions are necessary for making measurements. Dimensionless constants, however, are pure numbers and independent of any unit system. They are ratios between physical quantities (e.g., the electron–proton mass ratio $m_e/m_{pr} \approx 1/1836.15$) and cannot be skipped.

Often in calculations, the values of fundamental constants are normalized to 1 (e.g., $G = c = \hbar = 1$). But this is only a simplification. To make measurements, the units are still relevant. However, one can go a step further and define “natural” units independent of human standards, i.e., as dimensionless constants. Indeed, h, c, and G can be seen as mere conversion factors. c transforms energy into mass ($E = mc^2$), h energy into frequency ($E = h\nu$), and G mass into length, namely into the Schwarzschild radius R_s, the radius of a black hole ($R_s = 2GM/c^2$). Taking another important quantum property into account, the Compton wavelength ($\lambda_c = h/mc$), the following is possible: a mini black hole with the Compton wavelength as its Schwarzschild radius can act as a natural measuring rod and thermometer as well as clock and weighing machine (Duff et al. 2002). Any extraterrestrial civilization could understand it. But doing without any arbitrary human scales and definitions, “true” constants must be expressed with pure numbers, i.e., independent of dimensional quantities like velocity, mass, and length or reference systems like black holes. This is not possible with Planck units alone. To get pure numbers, one has to multiply them with another dimensional constant, e.g., the proton mass m_{pr}. For example, one gets $m_{pr}{}^2G/\hbar c = 10^{-38}$. Every physicist in our universe could understand, measure, and use such a value, independent of their yard stick, and discuss it with any other habitant in any other galaxy. But of course the question remains why this value is 10^{-38} and not something else.

Future theories of physics might reveal the relations between fundamental constants in a similar way as James Clerk Maxwell did by unifying electric and magnetic forces: he showed that three until then independent constants –

the velocity of light c, the electric constant ε_o (vacuum permittivity), and the magnetic constant μ_o (vacuum permeability) – are connected with each other: $c = (\mu_o \times \varepsilon_o)^{-0.5}$. Indeed some candidates for a grand unified theory of the strong, weak, and electromagnetic interaction suggest that most of the parameters in the standard model of particle physics are mathematically fixed, except for three: a coupling constant (the electromagnetic fine-structure constant) and two particle masses (namely that of down and up quarks) (Hogan 2000). A promise of string theory is even to get rid of any free parameter – if so, all constants could be calculated from first principles (Kane et al. 2002). However, this is still mainly wishful thinking at the moment. But it is a direction very worth following and, from a theoretical and historical point of view, perhaps the most promising and powerful.

So even without an ultimate explanation, *fine-tuning might be explained away within a (more) fundamental theory*. Most of the values of the physical constants should be derived from it, for example. This would turn the amazement about the anthropic coincidences into insight – like the surprise of a student about the relationship $e^{i\pi} = -1$ between the numbers e, i, and π in mathematics is replaced by understanding once he comprehends the proof. Perhaps the fact that the mass of the proton is 1836 times the mass of the electron could be similarly explained. If so, this number would be part of the rigid formal structure of a physical law which cannot be modified without destroying the theory behind it. An example for such a number is the ratio of any circle's circumference to its diameter. It is the same for all circles in Euclidean space: the circular constant π.

But even if all dimensionless constants of nature could be reduced to only one, a pure number in a Theory of Everything, its value would still be arbitrary, i.e., unexplained. No doubt, such a universal reduction would be an enormous success. However, the basic questions would remain: Why this constant, why this value? If infinitely many values were possible, then even the multitude of possibilities would stay unrestricted. So, again, why should such a universal constant have the value of, say, 42 and not any other?

If there were just one constant (or even many of them) whose value can be *derived* from first principles, i.e., from the ultimate theory or a law within this theory, then it would be completely explained or reduced at last. Then there would be no mystery of fine-tuning anymore, because there never was a fine-tuning of the constants in the first place. And then an appreciable amount of contingency would be expelled.

But what would such a spectacular success really mean? First, it could simply shift the problem, i.e., transfer the unexplained contingency either to the laws themselves or to the boundary conditions or to both. This would not be a Pyrrhic victory, but not a big deal either. Second, one might interpret it as an analytic solution. Then the values of the constants would represent no empirical information; they would not be a property of the physical world, but simply a mathematical result, a property of the structure of the theory. This, however, still could and should have empirical content, although not encoded in the constants. Otherwise fundamental physics as an empirical science would come to an end. But an exclusively mathematical universe, or at least an entirely complete formal description of everything there is, derivable from and contained within an all-embracing logical system without any free parameter or contingent part, might seem either incredible (and runs into severe logical problems due to Kurt Gödel's incompleteness theorems, see Chaitin 1987, 1992, 2001 in reference to irreducible randomness and algorithmic information theory) or as the ultimate promise of the widest and deepest conceivable explanation. Empirical research, then, would only be a temporary expedient like Ludwig Wittgenstein's (1922, 6.54) famous ladder: The physicist, after he has used empirical data as elucidatory steps, would proceed beyond them. "He must so to speak throw away the ladder, after he has climbed up on it."

3 Cosmic Coevolution as Fine-Tuning

A process with two manifestations within a special domain of space, time, and energy or, alternatively, correlated processes involving strong and repeated interaction can led to the coevolutionary emergence of complex features and entities. A classic example is the endosymbiotic theory of the origin of eukaryotic cells in evolutionary biology (Sagan 1967; Zimorski et al. 2014). Putting this to an extreme one can ask whether nature and its properties, especially the fine-tunings, are the result of a *coevolution of the universe and observers* (Ćirković and Dimitrijević 2018). This is speculative metaphysics, going back to ancient stoicism. Modern examples are Pierre Teilhard de Chardin's (1955) Omega Point theory with its claimed co-presence of the Omega Point in all previous moments of the history of the universe and Frank Tipler's (1994) refurbished version of it. Another example is John Wheeler's (1975, 1977) participatory anthropic principle; it postulates a feedback loop which links physical reality and observers, relying on a subjectivist interpretation of quantum mechanics, where the collapse of the wave function occurs only through interaction with an observer. Such ideas have huge problems (for criticism see Vaas 2001a, 2004a) and go beyond naturalism, but other approaches might be more promising, for instance, top-down causation (e.g., Okasha 2012; Ellis 2015).

Coevolution sounds like putting the cart before the horse – and this is arguably the case indeed (but such a proverbial action was actually tried in the late nineteenth and early twentieth centuries, especially in France, see Windsor 1907 for this historical point of view). However, coevolution as such is an important process – and cosmological natural or artificial selection can be interpreted as specific instantiations of it (see below). Therefore it is worth exploring such a framework to see whether and how it could be possible for life to fine-tune or adapt its own preconditions via an interaction with its cosmic environment (Ćirković and Dimitrijević 2018).

4 Natural Selection Instead of Fine-Tuning

What appears fine-tuned might not be – either because it is a unique, derivable consequence of an underlying lawful structure, and hence determined, or because it is the probable outcome of a stochastic process. An especially attractive possibility, also from an explanatory perspective, is a kind of Darwinian evolution of the values of fundamental constants (and perhaps even laws and boundary conditions). As in biology, i.e., evolutionary theory, ostensible features of design would be revealed as results of a nonintentional, self-organized process based on mutation, selection, and differential reproduction. Darwinian evolution is a well-established, indeed paradigmatic case of such a "blind" self-organization leading to astonishingly complex structures (Dennett 1995; Kanitscheider 2009; Vaas 2009a; Vollmer 2016). It is therefore a reasonable, although bold speculation to blow up a Darwinian kind of explanation to a cosmological scale within a multiverse framework.

In contrast to observational selection or bias according to the weak anthropic principle, which works in any multiverse scenario, but is not predictive, an observer-independent selection mechanism must generate unequal reproduction rates of universes, a peaked probability distribution, or another kind of differential frequency. For example, as Andrei Linde (1987) first pointed out, the constants of nature might vary from one inflationary domain to another, generating different rates of exponential expansion and bubble universe formation.

Up to now the best elaborated model of *cosmological natural selection* (CNS) is Lee Smolin's scenario (1992, 1997, 2006). Actually he started the whole approach and coined the term CNS. In contrast to observational selection, Smolin's CNS scenario is predictive and, thus, directly testable and falsifiable – at least under certain assumptions. (Note that CNS implies observational selection, but not vice versa.)

The hypothesis of CNS assumes, like anthropic observational selection, the existence of a multiverse or of a landscape of possible low-energy parameters. Furthermore, CNS assumes that black hole interiors bounce and evolve into new universes; that the values of the fundamental parameters can change thereby in small and random ways; that, therefore, different universes have different reproduction rates – universes with more black holes create more offspring universes; and, hence, that it is very probable after sufficient time that a universe chosen at random from a given collection of physically possible universes has parameters that are near a maximum of black hole production. If our universe is a typical member of that collection, then its fundamental parameters must be close to one of the maxima of the black hole production rates. Hence, our universe is selected for maximizing its number of black holes, and it is a descendant of universes which were already selected for this. Therefore, the fundamental parameters have the values we observe because this set of parameters leads to a (local) maximum of descendant universes, making the production of black holes much more likely than most alternatives. Thus, there is no need to invoke the weak (or even strong) anthropic principle – the existence of life is not used as part of the explanation of the parameter values. Just the reverse: Life's preconditions can be explained by CNS because the existence of stars and, as an offshoot, carbon chemistry, comes along with a high black hole formation rate. Therefore it is possible to continue physical research within a multiverse scenario without invoking the anthropic principle. In particular, this is true whether or not the ensemble of universes generated by bouncing black holes is a subensemble of a larger ensemble that might be generated by a random process such as eternal inflation. And CNS leads to a testable prediction: Most changes in the fundamental parameters would decrease the rate at which black holes are produced in our universe or leave it unchanged, but would not increase it. This prediction still holds.

So CNS is quite successful from a theoretical point of view. But there are also some crucial open issues and problems. They shall be analyzed more closely now (see Vaas 1998, 2003). If alternative explanatory frameworks could do better here, this would be a huge achievement.

First, there are *open questions regarding new universes emerging out of black holes*. Even if black holes are places of birth for universes, CNS still lacks a hereditary mechanism (Harrison 1995; Gardner 2003, p. 85). Why should descendant universes resemble their producers? It is also not clear whether the values of the physical parameters vary at all (if not, no multiverse evolution would occur in this framework), whether they really vary by *small* amounts and *randomly* as it is presumed, whether parameter sensitivity leads to (an increasing) fine-tuning instead of rendering universes sterile or at least hostile against life (cf. Stenger 2011), what happens to the universes already born if their mother black holes merge together or evaporate, whether there are further universes created if black holes merge together (and how many: one or two?), why there is only *one* offspring from a black hole and not (infinitely) many, and, if the latter is true, whether the numbers of new universes which are born from each black hole may differ according to the mass of the black hole. It was suggested that a large number of universes might be created inside each black hole and that the number of universes produced that way may grow as the mass of the black hole increases (Barrabès and Frolov 1996). If so, universes should be "selected" for *supermassive* black holes, not for sheer numbers of black holes. It is also not clear whether the different universes interact which each other. There is even the threat of a reciprocal destruction. Another restriction of Smolin's approach is that his cosmic reprocessing mechanism only leads to different values of parameters, but not to different *laws*. His hypothesis still requires the same basic structure of the laws in all the universes. But an even more radical proposal – a variation not only of constants but also of laws – is beyond the possibility of scientific investigation (at least for now). We simply do not know whether a distinction is useful between universes which are physically possible, as opposed to those that can only be imagined (which are only metaphysically or logically possible).

Second, *black hole production could be much more efficient without necessarily improving or suppressing parameter fine-tuning for life* (Rothman and Ellis 1993; Silk 1997). Because there are other mechanisms apart from core-collapse of old massive stars, there is no necessary correlation between black hole numbers and life-friendly parameter values for star numbers. For example, an enormous amount of primordial black holes, and not only microscopic ones, could have been produced within the first second of our early universe – and they might still be around as dark matter as well as gravitational-wave sources when colliding (see Carr et al. 2010; Calmet et al. 2014; García-Bellido 2017 for recent reviews). Also a larger cosmological constant (Λ) correlates with a higher black hole production rate, as Alexander Vilenkin (2006a) emphasized in the framework of CNS: In a vacuum energy dominated (de Sitter) universe, driven by Λ, which is related to the vacuum energy density ($\Lambda \sim 8\pi\rho_V$), quantum fluctuations of geometry lead occasionally to black hole formation via quantum tunneling. (In fact, our own universe will be completely vacuum dominated in a few dozen billion years.) Their semiclassical nucleation rate can be estimated and, although being extremely small, produces an infinite number of black holes, because de Sitter space is eternal to the future. Since quantum fluctuations increase with higher values of Λ, the black hole productions rate grows faster too. This is in tension with Smolin's conjecture that the values of all constants of nature are selected for maximizing black hole formation. Either Λ is an exception from that or the astronomically inferred dark energy of our universe (if real and not something else, e.g., an effect of a modification of general relativity on large scales) is not based on a true cosmological constant (but, e.g., on a scalar field which possibly decreases with time and might even get negative), such that accelerated de Sitter expansion stops before black holes can nucleate. Even then, however, other mechanisms (such as variations of the Higgs potential parameters) could also lead to a vacuum dominated expansion stage and produce infinitely many black holes, as Vilenkin (2006a) pointed out.

Third, and related to the former point, *there is no necessary connection between black holes and life*. In principle, life and CNS could be independent of each other. There are two reasons for this: On the one hand, there may be universes full of black holes where life as we know it couldn't evolve. For instance, it might be possible that there are only short-lived giant stars which collapse quickly into black holes or that there are universes dominated either by helium or by neutrons (corresponding to the neutron/proton mass difference being either zero or negative) or that there are universes without stars at all but many primordial black holes. Such universes might be very reproductive because of their giant stars or primordial black holes but are not able to produce earth-like life. On the other hand, we can conceive of a universe without black holes at all (if supernovae lead to neutron stars only) but which could be rich in earth-like life nevertheless. Thus, there is not a (logically) necessary connection between black holes and life (Rothman and Ellis 1993; Ellis 1997). Smolin's CNS hypothesis, therefore cannot necessarily explain the presumed fine-tuning of the universe for life. By the way, Smolin (1997, p. 393) also stressed that these two properties of our universe – containing life and producing a maximum number of black holes – must be taken as independent for the purpose of testing the theory. Nevertheless, there may be a contingent connection between black holes and life – via the role of carbon as the "molecule of life," because its ability to make complex molecules (to a much larger extent than any other element), and as an element accelerating star formation, because of the role that carbon monoxide plays in shielding and cooling the giant molecular clouds of gas and dust where stars are born (Smolin 1997). Thus, there may be at least a coincidental connection between the conditions for maximizing black hole formation and being hospitable for life. But we must still wonder why the laws of nature are such that this linkage occurs. (And it is not clear whether carbon monoxide really does increase the number of black holes, because hydrogen cools efficiently, too, and all the stars in the early universe were carbon-free.)

Fourth, in CNS, *life and intelligence are a kind of epiphenomenon*. If they do not contribute to black hole formation, life and intelligence are a mere by-product in Smolin's scenario, i.e., they are causally inert (regarding the evolution of the universes). Thus, in CNS our universe was not positively selected for life, even if the conditions of life would be exactly the same as the conditions for maximizing black holes. Therefore, CNS does not imply (or entail) a "meaning," function or advantage of life.

Fifth, there are *problems with predictability and testability*. According to Smolin, there are eight known variations in the values of fundamental parameters that lead to worlds with fewer black holes than our own, but there is no variation known with has clearly the opposite effect. This is already an interesting observation. Furthermore, Smolin's hypothesis has some predictive power, because there are physical effects and properties influencing black hole formation rates which are still not known (at least not precisely enough). Here, Smolin's hypothesis provides some constraints for these effects if the number of black holes is almost maximized. These predictions can in principle be tested, some even in the near future. They are related to the masses of K^- mesons (and hence the mass of the strange quark), electrons and neutron stars, the strength of the weak interaction, the density of protons and neutrons, the duration of the presumed inflationary epoch of the early universe, temperature patterns of the cosmic microwave background, the black hole formation rate dependence on the gravitational constant, etc. These predictions are testable in principle, and they are at home in the realm of current cosmology and particle physics. However, Smolin's central claim cannot be falsified. Falsifiability of a hypothesis depends on holding fixed the auxiliary assumptions needed to produce the targeted conclusion. In practice, one tries to show that the auxiliaries are themselves well confirmed or otherwise scientifically entrenched. What should be falsifiable according to Smolin is his claim that *our* universe is nearly optimal for black hole formation. However, this is not a necessary consequence of his premises. A consequence is only that *most* universes are nearly optimal. To move from this statistical conclusion to the targeted conclusion about our universe, Smolin (1997, p. 127) simply assumes that our universe is typical. This is an additional hypothesis as he admits. But this auxiliary assumption is neither confirmed nor otherwise scientifically entrenched. Thus, if changes in the values of our parameters did not lead to a lower rate of black hole formation – contrary to Smolin's prediction – we could always "save" CNS by supposing that our universe is not typical. Hence, there is (at least at the moment) no possibility to falsify Smolin's central claim that our universe is nearly optimal for black hole formation (Weinstein and Fine 1998). One could introduce other ad hoc hypotheses as well in order to keep the central idea of CNS. But this might undermine its falsifiability. For example, if there are parameters whose variation from their actual value leads to an increase of black hole formation, one could still claim that these parameters cannot be varied without also changing other parameters, leading, e.g., to large side-effects in star formation; hence, the originally varied parameters are no *independent* parameters contrary to the assumption. But note: if there is a Theory of Everything someday, which uniquely determines the values of the physical parameters, the CNS hypothesis would be falsified after all.

Sixth, Smolin's *Darwinian analogy of CNS is in some respect misleading*. Natural selection as described in biology depends on the assumption that the spread of populations (or genes) is mainly constrained by *external* factors (shortage of food, living space, mating opportunities, etc.). In comparison with that, the fitness of Smolin's universes is constrained by only one factor – the numbers of black holes – and this is an *internal* limitation. Furthermore, although Smolin's universes have different reproduction rates, they are not competing against each other (Maynard Smith and Szathmáry 1996; Vaas 1998). Although there are "successful" (productive, fecund) and "less successful" universes, there is no "overpopulation" and no selective pressure or "struggle for life," hence no natural selection in a strict biological (Darwinian) sense. Smolin's universes are isolated from each other (except maybe for their umbilical cords). Therefore, there couldn't be a quasi-biological evolution of universes. Thus, a central feature of Smolin's CNS scenario is that the values of the constants were not selected due to competition but only due to differential reproduction: Some universes have more offspring than others, but there is no rivalry about resources, space, etc. as in life's evolution according to Darwinian natural selection.

However, according to Andy Gardner and Joseph P. Conlon (2013, p. 3), "CNS acts as if according to a design objective of black-hole maximization, such that successive generations of universes will be increasingly contrived – that is, appearing designed – as if for the purpose of forming black holes"; to describe this with the Price equation of evolutionary genetics "neither mortality nor resource competition are fundamental aspects of selection." Whether this is accepted or not, Smolin's biological analogy or concept transfer of *fitness* in his cosmological context is in any case appropriate and independent of resource competition according to the usual definition in biology: "In evolutionary theory, fitness is a technical term, meaning the average number of offspring left by an individual relative to the number of offspring left by an average member of the population. This condition therefore means that individuals in the population with some characters must be more likely to reproduce (i.e., have higher fitness) than others" (Ridley 2004, p. 74).

Furthermore, one could envisage other scenarios of cosmic evolution where not only mutations of natural constants occur, leading to differential reproduction, but also competition between the universes or their origins and, hence, a Darwinian selection process. (Such descriptions are analogous to those of biological evolution, but of course do not refer to that in the strict sense; for the heuristic value and advantages of analogies, see Vidal 2010.) In biological settings, there is always a cost trade-off between seed production and something else, such as somatic complexity or environmental resources; so Smolin's seed production model is not very evo devo biology compliant and terms like "fecundity" versus "optimization" could be a more fruitful claim, with some bio-inspired models for fecundity (thanks to John Smart for pointing this out).

For instance, universes might nucleate out of accidental fluctuations within a string vacuum (Gasperini and Veneziano 2003) or within the vast landscape of string theory (Susskind 2005), inheriting certain properties. There could be a natural selection of such universes depending, for example, on their energy and matter densities (cf. Mersini-Houghton 2008): If the matter density is above a certain limit (or the cosmological constant is negative), the emerging universe would rapidly collapse and vanish; other universes would expand too fast, if their vacuum energy (cosmological constant) is too large – matter or structures like stars and, thus, life could not form in it. If the emergence of such universes would affect (suppress? increase?) the formation probabilities of others, for example, by influencing their "surrounding" parts of the string vacuum or landscape, a kind of competition could be the result.

Another example of cosmic Darwinism was recognized in the stochastic approach to quantum cosmology (García-Bellido 1995): In Brans–Dicke chaotic inflation, the quantum fluctuations of Planck mass M_p behave as mutations, such that new inflationary domains may contain values of M_p that differ slightly from their parent's. The selection mechanism establishes that the value of M_p should be such as to increase the proper volume of the inflationary domain, which will then generate more offspring. (Of course this selection mechanism only works if the values of the fundamental constants are compatible with inflation.) It is assumed here that the low-energy effective theory of string theory has the form of a scalar-tensor theory, with nontrivial couplings of the dilaton ϕ to matter. It is therefore expected that the description of stochastic inflation close to Planck scale should also include this extra scalar field. Brans–Dicke theory of gravity is the simplest scalar-tensor theory, containing a coupling constant ω. The string dilaton plays the role of the Brans–Dicke scalar field, which acts like a dynamical gravitational coupling: $M_p{}^2(\phi) = 2\pi\phi^2/\omega$. This scenario is in principle testable, by the way, because it predicts that the larger M_p is in a given inflationary domain, the smaller the amplitude of density perturbations should be. The universe evolves toward largest M_p and smallest amplitude of density perturbations compatible with inflation, which agrees well with observations. Thus our universe, with its set of values for the fundamental constants, would be the offspring of one such inflationary domain that started close to Planck scale and later evolved toward the radiation and matter dominated eras.

Besides the question whether one of the sketched scenarios turns out to be true, an important point is their common general idea, which shows already that cosmological natural selection provides a quite simple physical explanation of what seems to be mysterious fine-tuning – an explanation without any reference to intentionality or design. Analogous to Darwinian evolutionary theory in biology, the apparently sophisticated structure of the foundations of our universe might simply be the result of a *multiversal self-organization*. This is a straightforward explanation.

Whether it is true is not a philosophical question, however, but depends on empirical and theoretical data. The same holds for the other main approach discussed above: the hypothesis that the fine-tuning can be derived from a fundamental law.

Nevertheless, there are severe limits and problems, so other approaches should be welcomed. Critical competition is always good for science and heuristic developments of conceptual issues. This is also a philosophical advantage which should not be neglected. So why not work out other scenarios – can they do even better or, at least, as well?

5 Artificial Selection as Fine-Tuning

That intelligent life could play an essential role in a universal or multiversal reproduction cycle sounds outlandish and seems to be science fiction or speculative philosophy, not serious science. In fact it was firstly discussed in science fiction (such as Olaf Stapledon's *Star Maker*, 1937, David Brin's *What Continues… And What Fails…*, 1991, see also Gregory Benford's *Cosm* 1998); and Quentin Smith (1990) envisaged it from a concise philosophy of science perspective. One of the strongest proponents, Clément Vidal (2014, p. 194), also understands it as "a philosophical theory and not a scientific one." But the borders between science fiction, philosophy, and speculative science are somewhat fuzzy, and in the 1990s, scientists started to discuss those ideas too.

Louis Crane (1994/2010) speculated about the artificial making of universes and a "meduso-anthropic principle." Edward Harrison (1995, 1998) wrote about a "natural creation theory" and John Barrow (1998, p. 175) about "forced breeding" as well as "artificial selection." Steven Dick (2000, p. 204) envisaged a "natural God" as an "advanced intelligence," which "could have fine-tuned the physical constants." James N. Gardner (2000, 2003, 2007) argued that the universe is a product of intelligent architects acting for the "selfish biocosm" to run its own replication (and even proposed a few falsifiable predictions, cf. 2003, p. 135). Béla A. Baláz (2005) mused about a "cosmological replication cycle." John Smart (2000, 2008, 2012) discussed related issues with his "developmental singularity hypothesis" and his "transcension hypothesis"; in reference to CNS, he speaks about "cosmological natural selection with intelligence" (CSNI), whereupon artificial cosmogenesis is not necessarily implied (Smart 2008; for an up-to-date introduction, see http://evodevouniverse.com/wiki/Cosmological_natural_selection_(fecund_universes)#CNS_with_Intelligence_.28CNS-I.29). Clément Vidal (2008, 2010), also with reference to CNS, coined the term *cosmological artificial selection* (CAS) and provided a detailed review and discussion (2014, ch. 8.3). This term shall be used as an umbrella term in what follows.

Vidal (2010, 2014) also scrutinized the issues of the beginning and far future of our universe as well as their connections with life and intelligence and related them in an ambitious and speculative way. According to his proposal, the presumably fine-tuned laws and constants of nature can be interpreted as a result of CAS – as if they were chosen either physically or within an advanced computer simulation. So according to Vidal, the CAS hypothesis provides (1) an understanding of our apparently fine-tuned universe by explaining or reconstructing it with advanced simulations and as a possible result of cosmic design; (2) a fundamental role of life and intelligence in our universe and beyond, perhaps even refuting the impression that it is epiphenomenal, incidental, futile, or absurd; and (3) a scenario allowing for the really long-term or even eternal existence of life, by offering a way out of our universe if it is ultimately doomed. Put differently, the strength of CAS could be claimed to consist in a better understanding of (1) the origin, (2) the meaning, and (3) the potential far future of life. These three issues are connected, but nevertheless logically and physically independent of one another – thus, if one is wrong or inadequate, the other ones might still stand.

From an explanatory perspective, CAS is a new and very speculative hypothesis to understand the beginning and foundations of the universe. This might seem far-fetched and superfluous in comparison to other explanatory accounts such as a fundamental theory and/or the multiverse hypothesis. But CAS is neither meaningless nor a waste of time if one takes the problems of the other accounts seriously; at least CAS is worth a try from a philosophical standpoint. And it is not per se unscientific. It is important to note that a CAS scenario is fully reconcilable with physicalism or *ontological naturalism*. Cosmic engineers are not envisaged to be something "above" or "beyond" nature – either our universe or the multiverse – but they are a part and, indeed, a product of it. CAS does not require new or transcendent metaphysical entities or forces. (For introductions, definitions, discussions, and criticism of naturalism and physicalism, see, e.g., Mahner 2018; Papineau 2016; Poland 1994; Stoljar 2017; Vollmer 2017).

A potential – but of course very controversial – strength of CAS is that such models aim to explain the presumed fine-tuning of our universe easily: as the result of a goal-oriented, intentional action. This might or might not lead directly to our universe, and it might entail a certain amount of randomness. But artificial cosmogenesis seems to be at least a possibility to enhance or alter the natural selection of real universes (like a blind cosmic watchmaker intervening from outside – for even stronger claims see below). This might be done via studying and selecting simulated universes first, investigating the range of physical possibilities, or it might be done directly via making or starting new universes. Note that artificial selection in biology, on animals or plants or microorganisms, does not replace natural selection, and it does not "design" new organisms or create them from scratch; it tries only to foster some traits over others. So CAS might "just" extend this manipulation to cosmological scales. Thus an important part of CAS might consist in carefully selecting the right conditions, perhaps via extensive computer simulations prior to the replication events, and therefore the successor universes would really be the result of an intentional fine-tuning.

6 Black Holes, Life, and the Multiverse

It was argued that not black holes but intelligent beings are what universes might be selected for. If so, (1) either universes with life must have a higher reproduction rate – but why? (2) Or life must somehow be much more probable to occur than black holes – for which there is no indication yet (no single example of extraterrestrial life, let alone intelligent life is known, but the evidence for numerous stellar and supermassive black holes is overwhelming). (3) Or life-friendly universes must be preferred by a design process from an earlier universe.

Michael E. Price (2017), for example, argued that "intelligent life (as the least-entropic known entity) is more likely than black holes or anything else to be an adaptation designed by cosmological natural selection" and that "life appears more likely than black holes to be a mechanism of universe replication." He speculated that "intelligence functions ultimately (after evolving to a sufficiently sophisticated state) to create new universes that replicate the physical laws and parameters of its home universe." And: "As the most improbable known thing in the universe, designed by the strongest known antientropic process, life seems more likely than black holes – or any other known entity – to be a CNS-designed adaptation." This is not meant as an ultimate explanation of the universe – "unless we assume an infinite regression of biofriendly universes being produced by intelligent life, then we must assume that biofriendliness was at some point generated by a nonintelligent process," Price emphasized, but some properties of our universe might be indeed a product of cosmic artificial selection by intelligent inhabitants of our progenitor universe.

So black holes are essential for the current CAS models too, but in contrast to CNS, life and intelligence are not mere by-products in the cosmic reproduction cycle. On the contrary, there might be a hidden connection between the hospitality of universes for life on the one hand and black hole formation on the other. Perhaps black holes are advantageous for life, or life is advantageous for black hole formation. Therefore a CAS proponent can argue that cosmic engineers might create universes by means of black holes. There would be no self-organized evolution in this case but rather a preplanned development. Nevertheless, life could still be seen as a "tool" of the multiverse to produce more universes. But it would be no epiphenomena in this case.

Black holes might indeed be attractors for intelligence because of technological reasons: waste sinks, energy production, communication, computing, gravitational lens telescopes as cosmic learning devices, space travel, perhaps even time travel, etc. (see, e.g., Crane 1994/2010; Lloyd 2000; Inoue and Yokoo 2011; Smart 2012; Vaas 2018a, c; Vidal 2011; 2014). John Smart (2000, 2008, 2012, and this volume) extrapolated the trend toward increasingly computationally effective, dense, and energy-intensive technologies; he concluded that the ultimate technology might reach black hole density.

As with CNS, CAS also postulates that there is an evolution of the multiverse. This must be the case if not all universes are identical in respect of their physical laws and constants, and if the laws or constants of the descendants depend on those of their progenitors and also change at least partly from the former to the latter. Furthermore, if this evolution is analogous to biological development or evolution in the strict sense, explanatory transfers from biology might be useful and promising. Therefore these cosmological speculations could be inspired from and improved by evo devo universe approaches (see Smart 2008).

CAS might be realized both without a fundamental theory and by means of it. If there is a unique set of laws and constants with no alternatives, it might nevertheless allow the creation of new universes and, possibly, some variations of initial conditions. However, is CAS more convincing than multiverse scenarios that do without intentional causation? Or, to put it differently, why should a multiverse scenario not suffice?

Neglecting the possible meaning and far future of life, CAS has to be defended against two much simpler kinds of multiverse scenarios: (1) those with a random distribution of laws, boundary conditions, and values of constants and (2) those with a probability distribution of some kind.

Within the first scenario, the fine-tunings can be understood by anthropic bias or observational selection, i.e., the weak anthropic principle: We are able to live only in a specific universe whose physical properties are consistent with our existence – a prerequisite for it – and therefore we should not be surprised that we observe them. Thus according to the weak anthropic principle, the world consists of an ensemble of universes, a multiverse, with different laws and parameters of nature, and we can detect only those which are compatible with our existence or even among its necessary conditions. But this is, strictly speaking, not a physical explanation (Vaas 2004a) and might not even be testable, i.e., predictive (Smolin 2006).

Within the second scenario, the fine-tunings can be understood either via a case of cosmological natural selection or as the result of an underlying law, determining a probability distribution with a (local or global?) maximum, and the physical properties of our universe are in the vicinity of this maximum, i.e., quite probable. At the moment, we do not know of such law. We might just assume there is one. A stronger position is to adopt the principle of mediocrity (Vilenkin 1995), saying that we are typical observers in a certain sense. However, this principle might not be applicable here; or our universe is special, i.e., not at or near the maximum of the probability distribution. Then we have to find a different explanation, or we must come back to anthropic observational selection.

Another option – compatible with more fundamental laws, perhaps even depending on them – is cosmological natural selection. It could be seen as both the nearest relative of and strongest alternative to CAS. It seems to be testable, but CAS proponents might still argue that it does not explain enough and has many problems. This is true at least for Lee Smolin's version, because the bounce within black holes and the physical "mutations" are completely mysterious, and a "selection" mechanism is also missing.

In principle, CAS could work without the assumption and existence of a physical multiverse and the making of new universes via black holes. Some (restricted) kind of artificial cosmogenesis is also on stage if it is possible to use our universe endlessly (especially if it is infinite in space and the future), or to recycle it, or to start it all again, or to travel within it not only through space but also through time. This could even provide the chance for endless life. Regarding the far future, postbiological existence within a computer simulation might be very advantageous, but the problem of needing a capable hardware remains – and with it cosmological boundary conditions necessary for running it: especially enough available energy and the possibility of getting rid of entropy (waste heat). However, it is doubtful to achieve this infinitely long in a single universe. But perhaps it could be done with closed time-like loops (see Hoyle 1983, pp. 211 ff; Gardner 2005; Yurov et al. 2006; Vaas 2018c).

7 Creation Out of Something: Deism in the Lab?

As with the notions of the multiverse, the anthropic principle and selection, there are varieties of CAS (and CNS-I), that is different meanings and models. Vidal (2014, p. 181) distinguished and reviewed six levels of universe making and took all but (1) as instantiations of CAS (in the following summarized with a few minor modifications): (1) *blind*, without a role for intelligence, e.g., CNS; (2) *accidental*, e.g., as a by-product of black hole making for other goals such as energy production (Crane 1994/2010); (3) *intentional*, i.e., artificial universe making via black holes; (4) *cosmic breeder* with the ability to nudge the properties of the descendant universes in certain directions (*artificial selection* in a stronger sense, John Smart, personal communication, spoke about "intelligence-guided gardening"); (5) *cosmic engineer* with the ability to set precisely the physical laws and parameters, i.e., to create designer universes; and (6) *God player* with the ability to control every parameter and to create any nomological possible universe. All these levels of universe making are compatible with naturalism or physicalism! (Conceptually speaking, qualitatively different levels would be settled within other ontological frameworks such as mentalism/idealism/spiritism/solipsism, mind-matter dualism, or theism.)

Related is the question of motivation: Why would intelligent beings want to make offspring universes? This is very hard to imagine, but at least some possibilities can be envisaged from our primitive perspectives (reflecting probably just our own wishes and fears). Perhaps cosmic engineers simply want to test their physical theories and technological abilities (Harrison 1995, p. 200). Perhaps God players use other universes for entertainment against boredom (cosmic soap operas) or as cosmic competition games or for a sadistic satisfaction. Perhaps advanced civilizations would produce new universes out of altruistic motives (Gardner 2003, p. 224), perhaps to make them even more hospitable than their own universe (Harrison 1995, p. 200). Perhaps the creators want to fight their ultimate mortality, trying to send at least some of their heritage and knowledge into new universes to persist and develop (Ćirković and Bostrom 2000; Garriga et al. 2000) – or they even want to inhabit a new universe when their own one is dying

(Tough 1986, p. 497; Harrison 1995, p. 200; Lifton and Olson 2004; Vidal 2008, 2014; Vaas 2009b). In the end, CAS might be the only self-defense against cosmic doom.

From a speculative point of view, CAS might be praised for stressing that in principle *design* is – although not mandatory of course – at least possible within a cosmological and naturalistic framework. To emphasize it again: In contrast to theistic postulates of transcendent, nonphysical entities, and causation, a CAS scenario is fully reconcilable with ontological naturalism or physicalism.

But CAS is in its stronger sense – Vidal's (2014) levels (3) to (6) above – also an example of teleological selection. For this, there are other nonnaturalistic alternatives, for instance, some versions of the strong anthropic principle and even deistic or theistic creation. And in fact it was argued that a new scheme of explanation for the fine-tunings is needed – perhaps an anthropic, functional, or even teleological one? This would be one of the largest paradigm changes since the advent of classical physics. Insofar as CAS constitutes, a certain flavor of such a new kind of explanation, neither its challenge nor the reluctance against it, should be underestimated.

Also issues of terminology are delicate. Of course, the term "cosmic engineers" (Harrison 1995) is somewhat metaphorical. It indicates correctly that there must be an advanced technological activity at work. But it remains completely open what kind of supreme technician or civilization this is supposed to be. Perhaps the creators are organic or robotic individuals; perhaps it is a collective intelligence with a single (even nonpersonal?) mind; perhaps it pervades its universe completely or hides within castles made from neutron stars – most probably it radically exceeds our imagination. In some respects, those "cosmic engineers" might be seen as god-like. But nevertheless, they are not supernatural, not independent of spacetime and energy, not beyond the physical realm. They are "transcending" our universe, but not the multiverse.

Thus, CAS is compatible with and part of ontological naturalism and does not contradict the scientific attitude – in fact it pushes it to the extreme. Therefore, in the CAS scenario "creation" does *not* mean theistic creation. CAS can be seen as a kind of creation out of something – in contrast to a divine creation out of nothing, a world-making ex nihilo, a Kabbalahistic tzimtzum, a mystical emanation, or a mythical transformation of chaos into order. And artificial cosmogenesis can, in principle, be understood in physical or naturalistic terms entirely; no religious context is attached here. Vidal (2014, p. 182) prefers not to talk about "creating" but "producing" or "making" universes (following Davies 2006; Gribbin 2009); also he dislikes "design" because of religious "intelligent design" contexts. The ambiguities, on the other hand, point out that there are naturalistic alternatives to much more outlandish transcendent religious claims.

From a theological perspective, CAS might be seen as a technocratic successor of creation myths, a naïve secular belief, an exuberant scientism gone wild. This is not surprising. However, it goes astray. CAS is a (perhaps weak) scientific and philosophical hypothesis or speculation, not a substitute religion. CAS might be bold or beyond belief, depending on personal taste and sincerity, but it does neither attack the nature of science nor the science of nature. (Of course CAS proponents have to carry the burden of proof and should provide theoretical and empirical evidence, not the sceptics.) One could even think about some sort of naturalizing the divine – CAS as *deism in the lab*: If one defines deism simply as belief in a transcendent entity, absent of any doctrinal governance, who created our universe but does not interfere with it anymore via miracles, etc., the cosmic engineer(s) could be identified with such a deity, a "supreme being," "divine watchmaker," "grand architect of the universe," or "nature's god." Of course, this "god" is *not* the theistic one, it "transcends" *not* nature in general but only our universe, and "creation" is *not* a nonphysical causation. However, as classical deism claims, such an engineer god might indeed be determined using reason and observation of the natural world alone, without a need for either faith or organized religion. By the way, even deistic interventions in human affairs (or with respect to our universe as a whole after its fabrication) are not excluded in principle if the grand architect is able to interact with it, for instance, via gravitons from a five-dimensional bulk space, nonlocal quantum entanglements, or wormholes (which simple-minded beings like humans might understandably perceive or interpret as miracle or revelation). And of course a CAS-like deism would be a truncated deism, because the religious versions of deism – and there is plenty of variation here – have a very different background and goal, and they often contain much more, including moral and spiritual aspects (e.g., Waring 1967; Gay 1968; Byrne 1989; Johnson 2009).

Admittedly, all this sounds much more like science fiction or science fantasy than serious science, and it was not put forward by CAS proponents. But if one wants to adopt a theological perspective here at all, from that point of view CAS has indeed something provocative to offer: a radically physical deistic natural theology. Said with a twinkle in the eye, CAS both puts deism in the lab (of physical and philosophical reasoning) and is a result of deism in the lab (as a presumable process somewhere in the multiverse).

Sure enough, such theological contexts or connotations reinforce doubts about CAS and show what delicate issues this proposal raises. Critics might object that CAS is creationism or intelligent design in new clothes (and in certain respects it actually is, but without theological baggage or God and within a naturalistic framework); or that CAS reintroduces the teleological thinking that was painstakingly expelled in the history of physics and biology (and in certain respects it actually does); or that CAS blurs the distinction between science and theology/religion/metaphysics (which might also be the case if supernaturalism is watered down or abandoned – and surely one can always relabel the supernatural as "natural" by claiming an appropriate physical "explanation" exists). Such criticism might be exaggerated, but should be taken seriously. Therefore CAS proponents must emphasize the hypothetical character of their proposal as well as their own scientific (and even naturalistic?) stance; they must search for demarcation criteria between science and theology or metaphysics and accept them; they have to seek rigorous tests, both theoretical and empirical; they should clearly stress the distinction between CAS and ideological creationism; and they should point out that cosmic engineers are not divine beings to worship or to suppliantly submit to.

CAS is far from proven true and poses many crucial questions and problems, but as an inspiring and far-reaching hypothesis, it deserves unprejudiced discussion like any serious effort to improve the understanding of the strange world we live in.

8 Objections and Challenges

Sure enough, CAS has problems on its own (Vaas 2009b, 2012a).

First and foremost, there is the difficulty of realizing CAS: It is completely unclear whether universes can not only be simulated to some extent but also physically instantiated. A few scientific speculations are already on stage (see below), but still in their infancy.

Second, one must distinguish between intentional creation and simulation (even if it were empirically impossible to decide between them from "within"). A simulated universe does not have all the properties of a physically real universe – as a simulated river might obey hydrodynamical equations but doesn't make anything wet. Admittedly, deep epistemological problems are lurking here. And perhaps it will be possible to make the simulation so real that the missing properties are simply irrelevant; or to make it at least so useful that, for instance, conscious life within it is possible and the creators could "upload" their minds, knowledge, and experiences, surviving within their simulation if they can no longer do in their own universe. But the hardware problem remains: How can something simulate something else which is comparably complex? And if the programmer's universe is doomed, their universal computer and, hence, computed simulation sooner or later should be doomed too. So perhaps we live in (and are!) a computer simulation (Bostrom 2003). But this might have implications that could even lead to a reductio ad absurdum. As Paul Davies (2007, p. 496) emphasized, "there is no end to the hierarchy of levels in which worlds and designers can be embedded. If the Church-Turing thesis is accepted, then simulated systems are every bit as good as the original real universe at simulating their own conscious subsystems, sub-subsystems, and so on ad infinitum: gods and worlds and creators and creatures, in an infinite regress, embedded within each other. We confront something more bewildering than an infinite tower of virtual turtles: a turtle fractal of virtual observers, gods, and universes in limitlessly complex interrelationships. If *this* is the ultimate reality, there would seem to be little point in pursuing scientific inquiry at all into such matters. Indeed, to take such a view is as pointless as solipsism." The notion of a rationally ordered real world altogether would be effectively abandoned "in favor of an infinitely complex charade, where the very notion of explanation is meaningless."

Third, artificial selection includes also intended destruction, not only gardening or creation. Therefore, one can imagine advanced intelligences which are not able to produce universes but to annihilate them or transforming them for whatever reason (e.g., destroy complexity by initiating a vacuum phase transition). One could even speculate that such cosmic selectors want to be unique or not sharing cosmic resources and try to prevent other life-forms in their accessible multiverse.

Our universe might simply exist because they did neglect it by chance, forgot to pestle it, or failed to recognize its capacities early on (Jenny Wagner, personal communication). This may sound strange and awkward, but it can serve as an extreme illustration for the fact that CAS speculations might easily go astray and/or that we simply do know nothing about alleged cosmic engineers and their intentions. It can also motivate educated guesses about the ubiquity and extension of cosmic altruism (cf. Smart 2012; Vakoch 2014 for some optimistic arguments).

Fourth, there is a crucial question: If there are cosmic engineers at work, perhaps some of them having fine-tuned our universe, how did they emerge in the first place? In other words: *What or who created the creator(s)?* – To avoid an infinite explanatory regress, it seems most probable that they arose naturally in a life-friendly universe themselves. But this shifts the problem, because at least the creator's home universe should have formed without any intentional fine-tuning and CAS cannot apply here. Thus, either its origin was pure chance or the outcome of cosmological natural selection or evo devo coevolution or the result of a multiverse "generator" according to some fundamental laws, etc. Therefore we are back at the beginning, i.e., the original question regarding fine-tuning. If our universe was created according to CAS, the fine-tuning problem is just shifted to the problem of explaining an earlier fine-tuned universe where the cosmic engineers evolved. Their home universe might have been physically simpler than ours, but not too simple either, otherwise such complex creators could never have emerged. So this is a major objection against CAS.

And, connected with it, there is a further problem: One might wonder whether CAS has any convincing explanatory force at all. Because ultimately CAS tries to explain something complex (our universe) with something even more complex (cosmic engineers and engineering). But the usual explanatory scheme is just the converse: The explanans should be simpler than the explanandum (cf. Byl 1996; Barrow 1998, p. 132). Furthermore, CAS adds something qualitatively new: While multiverse (including CNS) and fundamental law approaches to the fine-tuning problem postulate some new nomological regularities, CAS postulates an intentional cause *in addition*. CAS is therefore a mixture of explanations: physical and intentional. (Intentions are not, as such, nonphysical, and actions can be conceptualized as causes – as specific causes, to be more precise (Davidson 2001) – so pacem other opinions there is no reason to abandon naturalism here, but intentional explanations are nevertheless not epistemologically reducible to physical explanations.)

These arguments are not a knockout objection (for a defense in a broader, but much more speculative context, see Vidal 2010, 2012, 2014, pp. 178 ff). But they point out some severe difficulties with CAS. At least they show that CAS – like any other design scenario – cannot be the ultimate explanation (Vaas 2006a). However, this is not what CAS proponents (should) have in mind anyway. And if it were possible for us to carry out artificial cosmogenesis by ourselves, a strong case for CAS can be made even within its explanatory restrictions. From a philosophical and practical perspective, CAS might be very important indeed.

9 Is Life Incidental?

One of the most remarkable developments in human cultural history was the recognition of our tiny place in the vast universe (or perhaps multiverse), and that we are not obviously meant to be here. The overcoming of a naïve and infantile anthropocentrism, that the universe is there for us, and, strangely connected, that an all-compassing god is there for us too (and vice versa) was one of man's great – and still not fully accomplished – achievements: an "emergence from his self-incurred immaturity" (Immanuel Kant). The Darwinian theory of evolution suggested that man and indeed life itself are not ingeniously designed, but a result of self-organizing processes, a "fruit of chance and necessity," as Jacques Monod (1970) used to cite Democritus. Astrophysics, big bang theory, and, finally, the still speculative scenarios of quantum cosmology taught the same lesson albeit on much larger scales: The emergence of intelligence was more or less an accident, not something planned in a universe that is indifferent to life's concerns, goals, and values. However, in intelligent, self-conscious beings like humans, the universe at least became partly aware of itself, poetically speaking.

But self- and I-consciousness also revealed the absurdity of life in the face of chance, futility, and misery (Vaas 1995b, 2008a). The shirking of firmly believing in transcendent creators or in an almighty, omnibenevolent god, though perhaps consolatory for some (Vaas 2009b, c, 2013), cannot surmount absurdity because misery, injustice, and death would be even more scandalous; thus antitheism is a natural reaction (Vaas 1999). This is, of course, an existential perspective. From a distant perspective, misery and catastrophes can also foster immunity, progress, and adaptation (cf. Smart 2017, ch. 11).

Anyway, it is hard for intentional, goal-oriented beings to accept the sometimes sophisticated structures of nature as the result of "blind" self-organized processes. But exactly this is the scientific approach which dispensed with the need for design assumptions or teleological explanations. The only opposing trends were some quasi-idealistic interpretations of quantum physics (including the participatory anthropic principle; Wheeler 1975, 1977) and the discussion of the so-called anthropic coincidences or fine-tuning of fundamental constants and some boundary conditions in particle physics and cosmology, sometimes taken as evidence for a cosmic design(er) or teleological force (strong or teleological anthropic principle). These issues are highly controversial from a scientific and philosophical point of view (Vaas 2004a). But if CAS were true, basic features of our universe, and even its very existence, could indeed not be explained without reference to intentionality. If cosmological artificial selection was involved, it must be part of such an explanation though it cannot be the full explanation.

Apart from explanatory issues, an attractive psychological feature of CAS might be, at least for some adherents, that it could hold life in high regard. If there is no omnibenevolent god, CAS might point toward a slender substitution after all. So, in the face of blank absurdity, CAS could be seen as a way out for some deeply frustrated would-be-believers, wanting to restore human grandeur and an ultimate meaning to everything.

But note that CAS does not necessarily mean that our universe was carefully designed with respect to every law and constant (or specific initial conditions). An engineered (ingenious) blueprint might have been realized, and such ideas are the basic of some "best of all possible worlds" beliefs. But it could also be the case that our universe was just cobbled together, perhaps with many others. Or it might even be an accident, for example, in a cosmological or particle collider experiment, i.e., an unintended by-product or collateral damage of an otherwise intended action. Although it is hard to imagine, one might even think of many different universes, originated completely naturally, with some of them, including ours, being intentionally picked – like fertilized eggs for uterus implantation in assisted reproductive technology – and activated to develop. (Of course this artificial selection could also have been a purely virtual process within a computer simulation to find out the right world-making recipe, as with iterative numerical calculations employed when there are no compact analytical solutions – after deriving a successful formula, then only the desired universe(s) would have been realized.) These different possibilities are not mutually exclusive by the way. For example, cosmic engineers might create any baby universe – and if it is capable of eternal inflation, then anything physically possible might ultimately evolve from it. Given the right laws, constants, and boundary conditions, even an infinite number of copies of their own universe (including Doppelgänger of the engineers themselves!) would emerge and any possible variation of it. Thus, as in CNS or eternal inflation scenarios, life and intelligence might be inevitable, although still accidental in some sense. So a kind of radical contingency remains. And of course one can still argue that life is absurd if anything that can happen will happen – and with any possible variation as well and infinitely often. Indeed exactly that was the point made by Friedrich Nietzsche when discussing eternal recurrence – but it is now an issue of modern cosmology too (Ellis and Brundrit 1979; Garriga and Vilenkin 2001; Tegmark 2004; Knobe et al. 2006; Vilenkin 2006b; Vaas 2001b, 2012b).

In conclusion, CAS is neither restricted to a careful and complete world design nor does it imply that every law, constant, and/or initial condition was intentionally selected. Creating life (let alone human beings!) need not be the goal of this art of world-making either. Perhaps life was just allowed for – or even an accident or mistake. If so, CAS would not prevent (our) life from being incidental. (Though at least we would have someone to blame for all the blunder.)

Even if one accepts CAS, without further knowledge, it is impossible to tell anything about the intentions of the creator(s). They might work in mysterious ways their wonders to perform. This is, of course, another problem for CAS. An intentional explanation without explaining the intention might be considered as a shortcoming. But this is not a refutation. And speculations are possible too.

For example, it was suggested that the creators – not taken as god(s) but as technologically very advanced, though nevertheless limited, cosmic engineers – are simply curious; so they might be trying hard to figure out ways of universe formation (an engineer's proverb, taken from Richard Feynman, states that one only understands something if one is able to make it). Then life might be an accident indeed.

Another possibility is that those cosmic engineers created their own universal soap opera for entertainment (perhaps even with sadistic intentions). Life would not be incidental then, but something like a zoo inhabitant or gladiator. Furthermore, our universe might soon become too boring for its spectators and therefore suddenly be deleted . . .

Much more serious is the assumption that the cosmic engineers face their own death too and the forthcoming end of their universe. Thus they might try to escape into a kind of rescue universe or at least transmit something of their knowledge lest they will not be forgotten completely. Death might be seen as an ultimately salvation, but it also marks the ultimate absurdity. To fight futility, self-conscious life gets the urge to endure and to intellectually grow endlessly. If this takes infinitely many (infinite?) universes, why not try to make them, if this is possible?

10 Is Life Ultimately Doomed?

It is an age-old question, whether the universe is infinite in time and space – or at least part of a larger system which is – and what this means for the ultimate prospects of life. In a branch of modern cosmology, sometimes called *physical eschatology*, this question can be discussed within the framework of (albeit speculative) scientific reasoning (Table 2).

The fate of the universe and intelligence depends crucially on the nature of the still mysterious *dark energy* which probably drives the accelerated expansion. Depending on dark energy's – perhaps time-dependent – equation of state, there is now a confusing number of mutually exclusive models. They are popularly called big whimper, big decay, big crunch, big brunch, big splat, big oscillation, big brake, big freeze, big rip, big trip, big hit, big hole, big resurrection, etc., and they envisage many different avenues. Most of them are dead ends for life – and this is also true for other cosmological models without dark energy. Thus, the ultimate future of our universe looks deadly dark (Ćirković 2003, 2004; Vaas 2006b).

But many self-conscious individuals want to fight absurdity and overcome death. If cosmological boundary conditions or creative minds – not necessarily god(s) – beyond it (see, e.g., Leslie 2001; 2008 for a far-reaching proposal), do not support this, mortals must try to take their fate into their own hands, prolonging their life and even searching for a "physics of immortality" (cf. Tipler 1994). Can CAS provide some help here?

11 Dark Energy Is Bad for Life

Going along with the externalization of memory and computation through the invention of writing, computers etc., a remarkable, accelerating increase of cultural complexity started on earth (and perhaps at many other places in the universe), a

Table 2 Exploring destiny: Scientific speculations concerning the very far future are bold but not unbound. One must keep some important presuppositions and open problems at the back of one's mind, however. With respect to cosmological artificial selection and artificial cosmogenesis, the possible role of intelligence influencing the fate of the universe is crucial

Presuppositions for physical eschatology and foundational queries	Comments
Ontological naturalism: Nonphysical entities do not exist (or have causal effects); no epistemological idealism or radical illusions about the universe	Scientifically unprovable, but strong philosophical arguments in favor
Ontological status of space and time?	They might be an illusion or not fundamental, but even so predictions and extrapolations are not meaningless and could be rephrased
No compact topology of space	Otherwise there are different boundary conditions
Weak determinism (at least); limited effects of chance (acausality)	There might be completely acausal events (like quantum effects), even on macroscopic scales, but no predictions are possible if a weak (moderate, average, or statistical) form of determinism doesn't hold
Relevant laws of nature are known	Quite questionable; how will new discoveries change the future and our view of it, are the established scientific methods sufficient at all, and what are the implications of a valid theory of quantum gravity?
Fundamental laws and constants of nature do not vary	But: Possibility of phase transition; perhaps there are no fundamental laws at all (but could effective regularities suffice?); signs of a time-dependent fine-structure constant ($\alpha = e^2/2hc\varepsilon_0 \approx 1/137$) already discovered?
Problems of infinity	Is actual infinity possible in nature? Is it realized? How to deal with infinities in theory (calculations, paradoxes) and research (only finite measurements possible)?
Restricted access?	Limited observations in space and time (particle horizon) and finite accuracy of measurements (especially of crucial parameters like Ω and w)
The universe as an open or closed system?	Problems for thermodynamics and conservation laws; are there interactions with other universes?
Limitations of explanations and predictions?	Even with (weak) determinism and all relevant laws and boundary conditions known, there might be strong restrictions due to nature's complexity and, perhaps, the incompleteness theorems
Role of intelligence?	Influencing destiny on cosmic scales?

tendency to do ever more work and to require ever less time, space, and energy (Fuller 1969; Chaisson 2001; Smart 2008; Vidal 2008). This is an excellent prospect for realizing CAS. However, the accelerated expansion of space leads to an universal limit on the total amount of information that can be stored and processed in the future (Krauss and Starkman 2004): This restricts the technology and computation capabilities for any civilization in principle, because there is access to only a finite volume, even after an infinite time. (On the other hand, the total amount of computational diversity at the universal scale might be increased because of galactic supercluster isolation due to the accelerated expansion of space, cf. Smart 2012.)

For a universe, dominated by a cosmological constant Λ (the simplest candidate for dark energy with the energy density ρ_Λ), which approaches asymptotically a de Sitter phase where the scale factor a increases exponentially, $a(t) = a_0 e^{Ht}$ with $H = (8\pi\rho_\Lambda/3)^{0.5}$, there is a maximum amount of energy $E_{max}(r)$ that will be received by harvesting matter out to a distance r: $E_{max}(r) = \Omega_m c^5/128GH$ where Ω_m is the dimensionless matter density, the sum of both baryonic matter (quarks and leptons) and dark matter. $E_{max}(r)$ has a maximum at $Hr/c = 1/2$, the de Sitter horizon is located at $Hr/c = 1$. The accessible energy is only 1/64th of the total energy located within the de Sitter horizon at the present time. With a flat metric ($k = 0$), a matter density $\Omega_m \approx 0.3$, and a Hubble constant $H_0 \approx 70\ \mathrm{kms^{-1}Mpc^{-1}}$, one finds $E_{max} \approx 3.5 \times 10^{67}$ J. This is comparable to the total rest-mass energy of baryonic matter within today's horizon. Dividing $E_{max}(r)$ by $Tk_B \ln 2$, where T is the noise temperature and k_B Boltzmann's constant, a minimum energy loss yields a limit on the number of bits that can be processed or information that can be registered. It is smaller than $\pi\Omega_m c^5/64\hbar GH^2 \ln 2 = 1.35 \times 10^{120}$. (Therefore, by the way, Gordon Moore's law, which assumes that the computer power doubles every 18–24 months, cannot continue unabated for more than 600 years for any technological civilization in our observable universe.)

In contrast to a simple eternally expanding universe (big whimper scenario with $\Lambda = 0$), a universe ruled by Λ leads to an everlasting expansion with dismal prospects for life. This is due to quantum effects at the cosmic horizon (analogous to Hawking radiation at the horizon of a black hole, but in de Sitter space the horizon surrounds the observer). Because of these the universe cannot cool down to (almost) 0 K. It has a total entropy S and, hence, a final temperature, the de Sitter T_{dS}, which will be reached within a few hundred billion years (Gibbons and Hawking 1977): $S = A/4 = 3\pi/\Lambda$ and $T_{dS} = 1/2\pi l$ with $A = 4\pi l^2$ and $l = (3/\Lambda)^{0.5}$. Here, A is the area of the de Sitter horizon at late times and l the curvature radius of that closed universe. T_{dS} is approximately 10^{-29} K (corresponding to 10^{-33} eV). It means the end for any living system because then it cannot radiate away waste heat – and there is no life without an energy gradient (Krauss and Starkman 2000).

Other scenarios look also more or less disappointing. But if our universe and every living being in it would be finally doomed, there could be infinitely many other universes and/or our universe might *recycle* itself due to new inflationary phase transitions out of black holes (Smolin 2006) or out of its high-energy vacuum state, where new exponential expanding bubbles should nucleate at a constant rate, growing to new universes elsewhere with new thermalized regions (Lee and

Weinberg 1987; Garriga and Vilenkin 1998), and cut their cords, metaphorically speaking. They probably will give rise to new galaxies and civilizations. It is not possible, however, to transcend these boundaries or to send a device with the purpose to recreate a follow-up of the original civilization in the new region or to transmit at least a kind of cosmic message in a bottle. It is not possible (Garriga et al. 2000), because the device or message will almost certainly be intercepted by black holes, which nucleate at a much higher rate than inflating bubbles, namely in the order of $\sim \exp(10^{122})$.

12 Wormhole Escapism and Designer Universes

If our universe is ultimately determined to die, or if at least the sufficient conditions for any possible information processing system disappear, the only chance for life would be to leave its universe and move to another place. Therein lays the prospect for an everlasting future of civilizations, and this is a strong motivation for CAS. So if the new universes are meant as new homes for their creators, because the universe they live in will run out of free energy and life-friendly conditions, the laws and constants of those successor homes will probably be intended to remain fixed – otherwise the cosmic engineers would cease after moving in. And, as mentioned above, their relocation must happen without quantum tunneling. Because of extremely small tunneling probabilities, all mechanisms that involve quantum tunneling are probably doomed to failure. However, there are bold speculations about *traversable wormholes* leading to other universes (Visser 1996; Krasnikov 2018; Vaas 2018c). This seems to be possible at least in the framework of general relativity. Perhaps wormholes could be found in nature and modified, or they could be built from scratch. If so, life could switch to another universe, escaping the death of its home.

And if there is no life-friendly universe with the right conditions (physical constants and laws), an advanced civilization might even create a sort of replacement or rescue universe on its own. In fact, some renowned physicists have speculated about such a kind of world-making (Farhi and Guth 1987; Frolov et al. 1989; Farhi et al. 1990; Fischler et al. 1990; Linde 1992; Crane 1994/2010; Harrison 1995; Merali 2006; Ansoldi and Guendelman 2006, 2008).

At a Grand Unified Theory energy scale of 10^{14} GeV, a universe might emerge from a classical bubble which starts out with a mass of only about 10 kg. By means of quantum tunneling, the bubble mass could be arbitrarily small, but the formation probability of a new universe would be reduced very much. Of course the main problem is to concentrate enough energy in a tiny volume. It has been suggested to try a coalescence of two regular magnetic monopoles (with below critical magnetic charge), producing a supercritical one which then inflates giving rise to a baby universe, or to take just one monopole and to hurl mass onto it, using a particle accelerator or a cosmic string (Borde et al. 1999; Sakai et al. 2006). The new bubble filled with a false vacuum is an extremely warped patch of spacetime and would create its own space: It undergoes an internal exponential inflation without displacing the space outside of the bubble itself (the negative pressure inside, zero outside, and the positive surface tension prevent the bubble from expanding into its mother universe). On the contrary it disconnects from the exterior region: The wormhole, which acts like an umbilical cord between the mother and child universe, collapses. (From the perspective of the mother universe, the disconnected bubble hides inside a microscopic black hole which will not appear to grow in size but evaporates quickly, while from within the bubble, the creation event is seen as a white hole-like initial singularity. Mathematically, the bubble can be described as a de Sitter spacetime embedded in a Schwarzschild spacetime, joined by using the Israel junction conditions.) The new universe might be detectable nevertheless because of modifications to the Hawking radiation.

It remains unclear, however, whether one could pass a message to the future inhabitants of the created universe (Hsu and Zee 2006) – due to inflation they would live in a tiny corner of a single letter, so to speak. Perhaps it could be encoded within the value of a fundamental constant. It remains also unclear, whether one could even travel into the descendent universe via new wormholes. If such an interchange of universes is possible, life might continue endlessly.

While such bold speculations easily sound awkward or technocratic or as the ultimate megalomania, they at least offer an interesting change of perspective (from observation to experiment), which questions the passive point of view when dealing with cosmological problems and the limits of observations due to the restrictions imposed by the spacetime structure on the causal relations among objects. This perspectival switch is another advantage of CAS.

Like dark energy, however, wormholes violate some fundamental energy conditions. And a violation of the weak energy condition (WEC) is also necessary to create new inflating regions without quantum tunneling and to go there or send messages into it, for instance, a blueprint of the engineering civilization. The required magnitude of the negative energy density is in the order of $-\rho \geq {H_{\mathrm{inf}}}^{-2}$, where H_{inf} is the inflationary expansion rate. Because WEC violation is in conflict with quantum inequalities (Ford and Roman 1997; Borde et al. 2002; Ford et al. 2002), it should be investigated how seriously this constraint is to be taken, since it is unclear to what extent these inequalities apply to interacting fields.

13 The Case for CAS?

The hypothesis of cosmological artificial selections does not only address (1) the origin and apparent fine-tuning of our universe but also (2) the possible value and meaning of life and (3) its ultimate future. However, all these complex issues provide eminent problems for CAS. One might argue that although CAS is based on three weak points, putting them together they make the case for CAS stronger, i.e., strengthen its stability under load like a tripod. Indeed, cosmic fine-tuning, meaning, and survival are fancily linked together in the CAS scenario and form a coherent picture. But this does not make the CAS proposal true, of course. And, indeed, the three points are qualitatively distinct: Fine-tuning is about explanation, meaning about evaluation, and survival about action and construction. Therefore it is questionable whether one can really strengthen the others, although explanation might be a necessary condition for construction (or vice versa?) and (the search for) meaning a crucial motivation for explanation and action. Nevertheless, all three points and, thus, CAS remain an open issue at the moment. Related are genuine philosophical and existential questions, which may strengthen the case for CAS, but are beyond the scope of this essay (for further discussions, see Vidal 2014).

CAS is (or at least starts out as) a metaphysical speculation. And there is nothing wrong with metaphysical speculations if they are not confused with or advertised as scientific results. What's more, (some) metaphysical speculations have a heuristic value and might even boost the formation of scientific hypotheses. And philosophy is, among other things, thinking in advance. Both the challenge of escaping cosmic doomsday and searching for penultimate explanations – really ultimate explanations are excluded (Vaas 2006a) – surely need unconventional input and encouragement. But CAS is or can be seen also as a scientific speculation. Like multiverse scenarios in general, it fulfills many criteria of science (Vaas 2008b, 2010) and could even be testable – or realizable – in the future. CAS might be judged as unlikely or far-fetched, but it is worth exploring and is a serious candidate to produce adapted complexity via evo devo processes (Smart 2017, ch. 11). "Just as life's incredibly adapted complexity self-organized over many evo devo cycles, and just as everything that is complex and adaptive inside our universe is a replicating system, it is most parsimonious to assume that our universe is a replicating evo devo system as well. If it is, its evo devo intelligence will always remain a limited and incomplete aid to selection, not a" godlike "designer" (http://evodevouniverse.com/wiki/Evolutionary_development_(evo_devo,_ED)). It extends the realm of both cosmological problems and possible solutions and, thus, challenges other approaches – constructive competition is always good for science and philosophy, and criticism is a gift for further developments.

Summing up, it seems doubtful whether the hypothesis of cosmological artificial selection is correct – at least in the stronger sense as the cause of our universe: First there are simpler and more probable explanations for the fine-tuning of our universe (or for getting rid of the anthropic coincidences altogether); second psychological urges for overcoming human contingency are no argument for the truth of scientific hypotheses, and CAS is far from being an analgesic against absurdity; and third it seems unlikely that an advanced civilization within our universe can intentionally start the creation of new universes either by simulating them (because of the finite computational resources both in size and in time) or by physically producing them (because this might either be too difficult or it happens naturally much earlier and more often anyway). Certainly we don't know enough to assign a probability or likelihood of CAS's truth yet.

If CAS is possible in principle, our successors or any other much further advanced civilization within our universe might be the very first to fully realize it. If this occurs as a simulation or emulation, its contents – as complex as they might be, perhaps including even self-conscious beings – ultimately would be doomed if the simulating hardware breaks down. And within our universe, this seems to be inevitable. Thus, such simulated universes cannot endure forever. (If we ourselves would live within a computer simulation, or rather be one, the show might stop very soon . . . without any prospect for a cosmic reset.) If, on the other hand, somebody within our universe can artificially create offspring universes and even transmit the recipe for doing that – either as a message or as a physical necessity, for instance, by starting Doppelgänger universes which inevitably will repeat history – then a potentially infinite chain of successor universes might begin. Eternal life could become a reality, even if it is not necessarily an eternal *continuing* life.

Assuming that such a giant chain of being is actually possible, however, it seems nevertheless quite unlikely that our universe is the very first one to accomplish this. Furthermore, this would be a violation of the Copernican principle because our location in spacetime, in this case the multiverse, would be very special. Therefore one should conclude that, given the CAS framework was correct indeed, our universe is a result of cosmological artificial selection (or simulation) too – one link within the probably future-eternal chain. If so, the spark of life may endure endlessly indeed. And even if we or our successors were not be able to pass it on, being then a tiny dead end within a flourishing realm of evolution, we will at least have envisioned it.

Acknowledgments This paper is partly based on Vaas (2009b, 2012a). I am grateful to Anthony Aguirre, Juan García-Bellido, John Leslie, Andrei Linde, Lee Smolin, Paul Steinhardt, and Alex Vilenkin for discussion over the years as well as Angela Lahee, André Spiegel, and Jenny Wagner for their kind support. Thanks also to John Smart and Clément Vidal for motivation, the invitation to contribute, and their very valuable suggestions. Scientific speculation and philosophy of science and nature are often dangerous fields but useful and thrilling nevertheless for getting ideas, criticism, and motivation to struggle against the boundaries of experience, empirical research, established theories, and imagination. As Carl Sandburg once wrote: “Nothing happens unless first a dream.”

References

Aguirre, A. (2001): The Cold Big-Bang Cosmology as a Counter-example to Several Anthropic Arguments. Phys. Rev. D 64, 083508; arXiv:astro-ph/0106143

Aguirre, A. (2007): Eternal Inflation, past and future; arXiv:0712.0571

Ansoldi, S., Guendelman, E. I. (2006): Child Universes in the Laboratory; arXiv:gr-qc/0611034

Ansoldi, S., Guendelman, E. I. (2008): Universes out of almost empty space. Prog. Theor. Phys. 120, 985–993; arXiv:0706.1233

Baláz, B. A. (2005): The cosmological replication cycle, the extraterrestrial paradigm and the final anthropic principle. Diotima 33, 44–53.

Barnes, L. A. (2012): The fine-tuning of the universe for intelligent life. Publications of the Astron. Soc. of Australia 29, 529–564; arXiv:1112.4647

Barrabès, C., Frolov, V. P. (1996): How many new worlds are inside a black hole? Phys. Rev. D 53, 3215–3223; arXiv:hep-th/9511136

Barrow, J. D. (1998): Impossibility: The Limits of Science and the Science of Limits. Oxford University Press: Oxford.

Barrow, J., Tipler, F. (1986): The Anthropic Cosmological Principle. Oxford University Press: Oxford.

Barrow, J. D., et al. (eds.) (2008): Fitness of the Cosmos for Life: Biochemistry and Fine-Tuning. Cambridge University Press: Cambridge.

Borde, A., Ford, L. H., Roman, T. A. (2002): Constraints on Spatial distributions of Negative Energy. Phys. Rev. D 65, 084002; arXiv:gr-qc/0109061

Borde, A., Trodden, M., Vachaspati, T. (1999): Creation and Structure of Baby Universes in Monopole Collisions. Phys. Rev. D 59, 043513; arXiv:gr-qc/9808069

Bostrom, N. (2002): Anthropic Bias. Routledge: New York, London.

Bostrom, N. (2003): Are We Living in a Computer Simulation? Phil. Quart. 53, 243–255.

Byl, J. (1996): On the natural selection of universes. Quart. J. Royal Astron. Soc. 37, 369–371.

Byrne, P. (1989): Natural religion and the nature of religion. Routledge: London.

Callender, C. (2004): Measures, explanation and the past: should 'Special' initial conditions be explained? British J. Phil. Sci. 55, 195–217.

Calmet, X., Carr, B., Winstanley, E. (2014): Quantum Black Holes. Springer: Heidelberg.

Carr, B. (2007): The Anthropic Principle Revisited. In: Carr, B. (ed.) (2007): The Universe or Multiverse? Cambridge University Press: Cambridge, pp. 77–89.

Carr, B. J., et al. (2010): New cosmological constraints on primordial black holes. Phys. Rev. D 81, 104019; arXiv:0912.5297

Chaisson, J. (2001): Cosmic Evolution: The Rise of Complexity in Nature. Harvard University Press: Cambridge.

Chaisson, E. J. (2011): Energy rate density as a complexity metric and evolutionary driver. Complexity 16 (3), 27–40; https://www.cfa.harvard.edu/~ejchaisson/reprints/EnergyRateDensity_I_FINAL_2011.pdf

Chaitin, G. (1987): Algorithmic Information Theory. Cambridge University Press: Cambridge.

Chaitin, G. (1992): Information-theoretic Incompleteness. World Scientific: Singapore.

Chaitin, G. (2001): Exploring Randomness. Springer: London.

Ćirković, M. M., Bostrom, N. (2000): Cosmological constant and the final anthropic hypothesis. Astrophys. Space Sci. 274, 675–687.

Ćirković, M. M. (2003): Resource letter: PEs-1: physical eschatology. Am. J. Phys. 71, 122–133; arXiv:astro-ph/0211413

Ćirković, M. M. (2004): Forecast for the Next Eon: Applied Cosmology and the Long-Term Fate of Intelligent Beings. Found. Phys. 34, 239–261; arXiv:astro-ph/0211414

Ćirković, M. M., Dimitrijević, J. (2018): Putting the Cart Before the Horse: Co-evolution of the Universe and Observers as an Explanatory Hypothesis. Found. Sci. 23 (3), 427–442.

Clifton, T., Linde, A., Sivanandam, N. (2007): Islands in the landscape. JHEP 0702, 024; arXiv:hep-th/0701083

Crane, L. (1994/2010): Possible Implications of the Quantum Theory of Gravity: An Introduction to the Meduso-Anthropic Principle. Found. Sci. 15, 369–373; arXiv:hep-th/9402104

Davidson, D. (2001): Essays on Actions and Events. Oxford University Press: Oxford.

Davies, P. C. W. (2006): The Goldilocks Enigma: Why Is the Universe Just Right for Life? Allen Lane: London.

Davies, P. (2007): Universes galore: where will it all end? In: Carr, B. (ed.) (2007): The Universe or Multiverse? Cambridge University Press: Cambridge, pp. 487–505.

de Chardin, P. T.(1955): The phenomenon of man. Harper & Row: New York 2008.

Dennett, C. (1995): Darwin's Dangerous Idea. Simon & Schuster: New York.

Dick, S. J. (2000): Cosmotheology: Theological implications of the new universe. In: Dick, S. J. (ed.) (2000): Many Worlds. The New Universe, Extraterrestrial Life and the Theological Implications. Templeton Foundation Press: Philadelphia, London, pp. 191–210.

Dick, S. J. (2003): Cultural Evolution, the Postbiological Universe, and SETI. Int. J. Astrobiol. 2, 65–74; http://history.nasa.gov/SP-4802.pdf

Dick, S. J. (2008): The Postbiological Universe. Acta Astronautica 62, 499–504.

Dick, S. J. (2009): The Postbiological Universe and our Future in Space. Futures 41, 578–580.

Duff, M. J., Okun, L. B., Veneziano, G. (2002): Trialogue on the number of fundamental constants. JHEP 0203, 023; arXiv:physics/0110060

Ellis, G. F. R., Brundrit, G. B. (1979): Life in the infinite universe. Quart. J. Royal Astron. Soc. 20, 37–41.

Ellis, G. (1997): A Darwinian universe? Nature 387, 671–672.

Ellis, G. F. R. (2015): Recognising top-down causation. In: Aguirre, A., et al. (eds.) (2015): Questioning the foundations of physics. Springer: Heidelberg, pp. 17–44.

Farhi, E., Guth, A. H. (1987): An obstacle to creating a universe in the laboratory. Phys. Lett. B 183, 149–155.

Farhi, E., Guth, A. H., Guven, J. (1990): Is it possible to create a universe in the laboratory by quantum tunnelling? Nucl. Phys B 339, 417–490.

Fedrowa, J. M., Griest, K. (2014): Anti-anthropic solutions to the cosmic coincidence problem. JCAP 01, 004; arXiv:1309.0849

Fischler, W., Morgan, D., Polchinski, J. (1990): Quantum nucleation of false-vacuum bubbles. Phys. Rev. D 41, 2638–2641.

Ford, L. H., Roman, T. A. (1997): Restrictions on Negative Energy Density in Flat Spacetime. Phys. Rev. D 55, 2082–2089; arXiv:gr-qc/9607003

Ford, L. H., Helfer, A. D., Roman, T. A. (2002): Spatially Averaged Quantum Inequalities Do Not Exist in Four-Dimensional Spacetime. Phys. Rev. D 66, 124012; arXiv:gr-qc/0208045

Frolov, V. P., Markov, M. A., Mukhanov, M. A. (1989): Through a black hole into a new universe? Phys. Lett. B 216, 272–276.

Fuller, R. B. (1969): Utopia Or Oblivion: The Prospects for Humanity. Overlook Press: New York.

García-Bellido, J. (1995): Quantum Diffusion of Planck Mass and the Evolution of the Universe. In: Occhionero, F. (ed.) (1995): Birth of the Universe and Fundamental Physics. Lecture Notes in Physics 455. Springer: Berlin, pp. 115–120; arXiv:astro-ph/9407087

García-Bellido, J. (2017): Massive Primordial Black Holes as Dark Matter and their detection with Gravitational Waves. J. Phys.: Conf. Ser. 840, 012032; arXiv:1702.08275

Gardner, A., Conlon, J. P. (2013): Cosmological natural selection and the purpose of the universe. Complexity 18, 48–56.

Gardner, J. N. (2000): The selfish biocosm: Complexity as cosmology. Complexity 5 (3), 34–45.

Gardner, J. N. (2003): Biocosm: The New Scientific Theory of Evolution. Inner Ocean: Makawao.

Gardner, J. N. (2005): Coevolution of the cosmic past and future: the selfish biocosm as a closed timelike curve. Complexity 10 (5), 14–21.

Gardner, J. N. (2007): The Intelligent Universe: AI, ET, and the Emerging Mind of the Cosmos. Career Press: Franklin Lakes.

Garriga, J., Vilenkin, A. (1998): Recycling universe. Phys. Rev. D 57, 2230–2244; arXiv:astro-ph/9707292.

Garriga, J., et al. (2000): Eternal inflation, black holes, and the future of civilizations. Int. J. Theor. Phys. 39, 1887–1900; arXiv:astro-ph/9909143

Garriga, J., Vilenkin, A. (2001): Many worlds in one. Phys. Rev. D 64, 043511; arXiv:gr-qc/0102010

Gasperini, M., Veneziano, G. (2003): The Pre-Big Bang Scenario in String Cosmology. Phys. Rept. 373, 1–212; arXiv:hep-th/0207130

Gay, P. (ed.) (1968): Deism. Van Nostrand: Princeton.

Gibbons, G., Hawking, S. W. (1977): Cosmological Event Horizons, Thermodynamics, and Particle Creation. Phys. Rev. D 15, 2738–2751.

Gribbin, J. (2009): In Search of the Multiverse. Allen Lane: London.

Harnik, R., Kribs, G., Perez, G. (2006): A Universe without weak interactions. Phys. Rev. D 74, 035006; arXiv:hep-ph/0604027

Hartle, J., Hawking, S. W. (1983): The wave function of the universe. Phys. Rev. D 28, 2960–2975.

Harrison, E. R. (1995): The natural selection of universes containing intelligent life. Quart. J. Royal Astron. Soc. 36, 193–203.

Harrison, E. R. (1998): Creation and Fitness of the Universe. Astronomy & Geophysics 39 (2), 27.

Hempel, C. G. (1965): Aspects of Scientific Explanation and Other Essays in the Philosophy of Science. Free Press: New York.

Hogan, C. J. (2000): Why the Universe is Just So. Rev. Mod. Phys. 72, 1149–1161; arXiv:astro-ph/9909295

Hoyle, F. (1983): The Intelligent Universe: A New View of Creation and Evolution. Michael Joseph: London.

Hsu, S., Zee, A. (2006): Message in the Sky. Mod. Phys. Lett. A 21, 1495–1500; arXiv:physics/0510102

Hume, D. (1779): Dialogues Concerning Natural Religion. Dialogues Concerning Natural Religion and Other Writings. Cambridge University Press: Cambridge 2007.

Inoue, M., Yokoo, H. (2011): Type III Dyson Sphere of Highly Advanced Civilizations around a Super Massive Black Hole. J. Brit. Interplanetary Soc. 64, 58–62; arXiv:1112.5519

Jaffe, R. L., Jenkins, A., Kimchi, I. (2009): Quark Masses: An Environmental Impact Statement. Phys. Rev. D 79, 065014; arXiv:0809.1647

Johnson, B. (2009): Deism. Truth Seeker: Escondido.

Kane, G. L., Perry, M. J., Zytkow, A. N. (2002): The Beginning of the End of the Anthropic Principle. New Astron. 7, 45–53; arXiv:astro-ph/0001197

Kanitscheider, B. (2009): Darwins Theorie als Prototyp und Vorläufer einer Theorie der Selbstorganisation. Universitas 64 (751), 56–66.

Knobe, J., Olum, K. D., Vilenkin, A. (2006): Philosophical Implications of Inflationary Cosmology. Brit. J. Phil. Sci. 57, 47–67; arXiv:physics/0302071

Krasnikov, S. (2018): Back-in-Time and Faster-than-Light Travel in General Relativity. Springer: Cham etc.

Krauss, L. M., Starkman, G. D. (2000): Life, The Universe, and Nothing. Astrophys. J. 531, 22–30; arXiv:astro-ph/9902189

Krauss, L. M., Starkman, G. D. (2004): Universal Limits on Computation; arXiv:astro-ph/0404510

Lee, K. M., Weinberg, E. J. (1987): Decay of the True Vacuum in Curved Space-Time. Phys. Rev. D 36, 1088–1094.

Leslie, J. (1989): Universes. Routledge: London 1996.

Leslie, J. (2001): Infinite Minds. Clarendon Press: Oxford.

Leslie, J. (2008). Infinitely Long Afterlives and the Doomsday Argument. Philosophy 83, 519–524.

Lifton, R. J., Olson, E. (2004): Symbolic immortality. In: Robben, A. C. G. M. (ed.) (2004): Death, Mourning, and Burial: A Cross-Cultural Reader. Wiley-Blackwell: Malden, Oxford, Carlton, pp. 32–39.

Linde, A. D. (1987): Particle physics and inflationary cosmology. Phys. Today 40 (9), 61–68.

Linde, A. (1992): Hard Art of the Universe Creation. Nucl. Phys. B 372, 421–442; arXiv:hep-th/9110037

Linde, A. (2005): Particle Physics and Inflationary Cosmology. Contemp. Concepts Phys. 5, 1–362; arXiv:hep-th/0503203

Linde, A. (2006): Inflation and String Cosmology. Prog. Theor. Phys. Suppl. 163, 295–322; arXiv:hep-th/0503195

Linde, A. (2008): Inflationary Cosmology. Lect. Notes Phys.738, 1–54; arXiv:0705.0164

Linde, A. (2017): On the problem of initial conditions for inflation; arXiv:1710.04278

Linde, A., Vanchurin, V. (2010): How many universes are in the multiverse? Phys. Rev. D 81, 083525; arXiv:0910.1589

Lipton, P. (2004): Inference to the Best Explanation. Routledge: London, 2nd ed.

Lloyd, S. (2000): Ultimate physical limits to computation. Nature 406, 1047–1054; arXiv:quant-ph/9908043

Mahner, M. (2018): Naturalismus. Alibri: Aschaffenburg.

Manson, N. A. (2000): There is no adequate definition of 'fine-tuned for life'. Inquiry 43, 341–352.

Mayes, G. R. 2005: Theories of Explanation. The Internet Encyclopedia of Philosophy; http://www.utm.edu/research/iep/e/explanat.htm

Maynard Smith, J., Szathmáry, E. (1996): On the likelihood of habitable worlds. Nature 384, 107.

McGinn, C. (1989): Can We Solve the Mind-Body Problem? Mind 98, 349–366.

McGrew, T., McGrew, L., Vestrup, E. (2001): Probabilities and the fine-tuning argument: a sceptical view. Mind 110, 1027–1038.

Merali, Z. (2006): Create your own universe. New Scientist 2559, 32–35.

Mersini-Houghton, L. (2008): Birth of the Universe from the Multiverse; arXiv:0809.3623

Monod, J. (1970): Chance and Necessity. Knopf: New York 1971.

Mosterín, J. (2005): Anthropic explanations in cosmology. In: Hajek, P., Valdés-Villanueva, L., Westerstahl, D. (eds.) (2005): Logic, Methodology and Philosophy of Science. King's College Publications: London, pp. 441–471; http://philsci-archive.pitt.edu/1658/

Okasha, S. (2012): Emergence, hierarchy and top-down causation in evolutionary biology. Interface Focus 2, 49–54.

Papineau, D. (2016): Naturalism. The Stanford Encyclopedia of Philosophy; https://plato.stanford.edu/archives/win2016/entries/naturalism/

Pitt, J.C. (ed.) (1988): Theories of Explanation. Oxford University Press: New York.

Poland, J. (1994): Physicalism: The Philosophical Foundations. Clarendon: Oxford.

Price, M. (2017): Entropy and Selection: Life as an Adaptation for Universe Replication. Complexity 2017, 1–4.

Rescher, N. (2000): The Price of an Ultimate Theory. Philosophia Naturalis 37, 1–20.

Ridley, M. (2004): Evolution. Blackwell: Malden, 3rd ed.

Rothman, T., Ellis, G. F. R. (1993): Smolin's natural selection hypothesis. Quart. J. Royal Astron. Soc. 34, 201–212.

Sagan, L. (1967): On the origin of mitosing cells. Journal of Theoretical Biology 14, 255–274.

Sakai, N., et al. (2006): The universe out of a monopole in the laboratory? Phys. Rev. D 74, 024026; arXiv:gr-qc/0602084

Salmon, W. C. (1998): Causality and explanation. Oxford University Press: New York.

Sandberg, A., Armstrong, S., Ćirković, M. M. (2016): That is not dead which can eternal lie: the aestivation hypothesis for resolving Fermi's paradox. J. British Interplanetary Society 69, 406–415; arXiv:1705.03394

Silk, J. (1997): Holistic cosmology. Science 277, 644.

Smart, J. (2000): Introduction to the Developmental Singularity Hypothesis; http://www.accelerationwatch.com/developmentalsinghypothesis.html

Smart, J. (2008): Evo Devo Universe? A Framework for Speculations on Cosmic Culture. In: Dick, S. J., Lupisella, M. (eds.): Cosmos and Culture: Cultural Evolution in a Cosmic Context. NASA: Washington, pp. 201–295; http://accelerating.org/downloads/SmartEvoDevoUniv2008.pdf

Smart, J. (2012): The transcension hypothesis: sufficiently advanced civilizations invariably leave our universe, and implications for METI and SETI. Acta Astronautica 78, 55–68; http://accelerating.org/articles/transcensionhypothesis.html

Smart, J. (2017): The Foresight Guide; http://www.foresightguide.com/

Smith, Q. (1990): A natural explanation of the existence and laws of our universe. Australasian Journal of Philosophy 68 (1), 22–43; http://www.infidels.org/library/modern/quentin_smith/natural.html

Smolin, L. (1992): Did the universe evolve? Class. Quant. Grav. 9, 173–191.

Smolin, L. (1997): The Life of the Cosmos. Oxford University Press: Oxford.

Smolin, L. (2004): Cosmological natural selection as the explanation for the complexity of the universe. Physica A 240, 705–713.

Smolin, L. (2006): The status of cosmological natural selection; arXiv:hep-th/0612185

Stenger, V. J. (2011): The Fallacy of Fine-Tuning: Why the Universe Is Not Designed for Us. Prometheus; Amherst.

Stoljar, D. (2017): Physicalism. The Stanford Encyclopedia of Philosophy; https://plato.stanford.edu/archives/win2017/entries/physicalism/

Susskind, L. (2005): The Cosmic Landscape. Little, Brown: New York.

Tegmark, M. (2004): Parallel Universes. In: Barrow, J., Davies, P. C. W., Harper Jr C. L., (eds.) (2004): Science and Ultimate Reality. Cambridge University Press: Cambridge, pp. 459–491; arXiv:astro-ph/0302131

Tegmark, M., et al. (2006): Dimensionless constants, cosmology, and other dark matters. Phys. Rev. D 73, 23505; arXiv:astro-ph/0511774

Tipler, F. J. (1994): The Physics of Immortality. Anchor Books: New York.

Tough, A. (1986): What role will extraterrestrials play in humanity's future? J. Brit. Interplanetary Soc. 39, 491–498; http://ieti.org/tough/articles/future.pdf

Vaas, R. (1993): Die Welt als Würfelspiel. In: Evangelische Akademie Baden (ed.) (1993): "Gott würfelt (nicht)!" Karlsruhe, pp. 108–162.

Vaas, R. (1995a): Reduktionismus und Emergenz. In: Die mechanische und die organische Natur. Beiträge zum Naturverständnis. Konzepteheft 45 des SFB 230. Stuttgart, Tübingen, pp. 102–161.

Vaas, R. (1995b): Masse, Macht und der Verlust der Einheit. In: Krüger, M. (ed.) (1995): Einladung zur Verwandlung. Hanser: München, pp. 219–260.

Vaas, R. (1998): Is there a Darwinian Evolution of the Cosmos? – Some Comments on Lee Smolin's Theory of the Origin of Universes by Means of Natural Selection. Proceedings of the MicroCosmos – MacroCosmos Conference, Aachen; arXiv:gr-qc/0205119

Vaas, R. (1999): Der Riß durch die Schöpfung. der blaue reiter. Journal für Philosophie 10, 39–43.

Vaas, R. (2001a): Why Quantum Correlates Of Consciousness Are Fine, But Not Enough. Informação e Cognição 3 (1), 64–107; http://www2.marilia.unesp.br/revistas/index.php/reic/article/view/715/617

Vaas, R. (2001b): Ewiges Leben im Universum? bild der wissenschaft 9, 62–67.

Vaas, R. (2003): Problems of Cosmological Darwinian Selection and the Origin of Habitable Universes. In: Shaver, P. A., DiLella, L., Giménez, A. (eds.): Astronomy, Cosmology and Fundamental Physics. Springer: Berlin, pp. 485–486.

Vaas, R. (2004a): Ein Universum nach Maß? Kritische Überlegungen zum Anthropischen Prinzip in der Kosmologie, Naturphilosophie und Theologie. In: Hübner, J., Stamatescu, I.-O., Weber, D. (eds.) : Theologie und Kosmologie. Mohr Siebeck: Tübingen, pp. 375–498.

Vaas, R. (2004b): Time before Time. Classifications of universes in contemporary cosmology, and how to avoid the antinomy of the beginning and eternity of the world; arXiv:physics/0408111

Vaas, R. (2006a): Das Münchhausen-Trilemma in der Erkenntnistheorie, Kosmologie und Metaphysik. In: Hilgendorf, E. (ed.) (2006): Wissenschaft, Religion und Recht. Logos, Berlin, pp. 441–474.

Vaas, R. (2006b): Dark Energy and Life's Ultimate Future. In: Burdyuzha, V. (ed.) (2006): The Future of Life and the Future of our Civilization. Springer: Dordrecht, pp. 231–247; arXiv:physics/0703183

Vaas, R. (2008a): Aufrechtstehen im Nichts. Universitas 63 (749 & 750), 1118–1137 & 1244–1259.

Vaas, R. (2008b): Phantastische Physik: Sind Wurmlöcher und Paralleluniversen ein Gegenstand der Wissenschaft? In: Mamczak, S., Jeschke, W. (eds.): Das Science Fiction Jahr 2008. Heyne: München, pp. 661–743.

Vaas, R. (2009a): Die Evolution der Evolution. Universitas 64 (751), 4–29.

Vaas, R. (2009b): Life, the Universe, and almost Everything: Signs of Cosmic Design?; arXiv:0910.5579

Vaas, R. (2009c): Gods, Gains, and Genes. On the Natural Origin of Religiosity by Means of Biocultural Selection. In: Voland, E., Schiefenhövel, W. (eds.) (2009): The Biological Evolution of Religious Mind and Behavior. Springer: Heidelberg, pp. 25–49.

Vaas, R. (2010): Multiverse Scenarios in Cosmology: Classification, Cause, Challenge, Controversy, and Criticism. J. Cosmology 4, 664–673; arXiv:1001.0726

Vaas, R. (2012a): Cosmological Artificial Selection: Creation out of Something? Found. Sci. 17, 25–28; arXiv:0912.5508

Vaas, R. (2012b): "Ewig rollt das Rad des Seins": Der 'Ewige-Wiederkunfts-Gedanke' und seine Aktualität in der modernen physikalischen Kosmologie. In: Heit, H., Abel, G., Brusotti, M. (eds.) (2012): Nietzsches Wissenschaftsphilosophie. de Gruyter: Berlin, New York, pp. 371–390.

Vaas, R. (2013): Die neue Schöpfungsgeschichte Gottes – Herausforderungen einer Evolutionsbiologie der Religiosität. In: Fink, H. (ed.) (2013): Die Fruchtbarkeit der Evolution. Alibri: Aschaffenburg, pp. 133–172.

Vaas, R. (2014a): Wahrheiten auf hoher See. Maritime Metaphern vom Leben, Forschen und Untergehen. Universitas 69 (820), 42–71.

Vaas, R. (2014b): Vom Gottesteilchen zur Weltformel. Kosmos: Stuttgart, 2nd ed.

Vaas, R. (2015): Im Anfang war der Urknall – oder nichts, Gott, alles? Schöpfungsglaube gegen moderne Kosmologie. Universitas 70 (823), 44–76.

Vaas, R. (2017a): Bewusstsein X.0. Von digitalen Denkwürdigkeiten zur ungeheuerlichen Unsterblichkeit. Universitas 72 (854), 64–81.

Vaas, R. (2017b): Umzug der Menschheit? Vom Homo sapiens zum Homo spaciens und weiter ... Universitas 72 (856), 42–69.

Vaas, R. (2017c): Jenseits von Einsteins Universum. Kosmos: Stuttgart, 4th ed.

Vaas, R. (2018a): Superzivilisationen im All. bild der wissenschaft 7, 8–26.

Vaas, R. (2018b): Hawkings neues Universum. Wie es zum Urknall kam. Kosmos: Stuttgart, 6th ed.

Vaas, R. (2018c): Tunnel durch Raum und Zeit. Kosmos: Stuttgart, 8th ed.

Vakoch, D. A. (ed.) (2014): Extraterrestrial Altruism: Evolution and Ethics in the Cosmos. Springer: Heidelberg.

van Riel, R., Van Gulick, R. (2018): Scientific Reduction. The Stanford Encyclopedia of Philosophy; https://plato.stanford.edu/archives/sum2018/entries/scientific-reduction/

Vidal, C. (2008): The Future of Scientific Simulations: from Artificial Life to Artificial Cosmogenesis. In: Tandy, C. (ed.) (2008): Death And Anti-Death. Ria University Press: Palo Alto, pp. 285–318; arXiv:0803.1087

Vidal, C. (2010): Computational and Biological Analogies for Understanding Fine-Tuned Parameters in Physics. Found. Sci. 15 (4), 375–393; arXiv:0912.5508

Vidal, C. (2011): Black Holes: Attractors for Intelligence?; arXiv:1104.4362

Vidal, C. (2012): Fine-tuning, Quantum Mechanics and Cosmological Artificial Selection. Found. Sci. 17 (1), 29–38; arXiv:0912.5508

Vidal, C. (2014): The Beginning and the End: The Meaning of Life in a Cosmological Perspective Springer: Heidelberg etc.; arXiv:1301.1648

Vilenkin, A. (1995): Predictions from Quantum Cosmology. Phys. Rev. Lett. 74, 846–849; arXiv:gr-qc/9406010

Vilenkin, A. (2006a): On cosmic natural selection; arXiv:hep-th/0610051

Vilenkin, A. (2006b): Many Worlds in One. Hill and Wang: New York.

Visser, M. (1996): Lorentzian Wormholes. American Institute of Physics Press: Woodbury.

Vollmer, G. (2016): Im Lichte der Evolution. Hirzel: Stuttgart.

Vollmer, G. (2017): Gretchenfragen an Naturalisten. Alibri: Aschaffenburg.

Waring, E. G. (ed.) (1967): Deism and Natural Religion. Frederick Ungar: New York.

Weinstein, S., Fine, A. (1998): Book Review of Lee Smolin's The Life of the Cosmos. J. Phil. XCV, 264–268.

Wheeler, J. A. (1975): The universe as home for man. In: Gingerich, O. (ed.) (1975): The nature of scientific discovery. Smithsonian Institution Press: Washington, pp. 261–296

Wheeler, J. A. (1977): Genesis and observership. In: Butts, R. E., Hintikka, J. (eds.) (1977): Foundational problems in the special sciences. Reidel: Dordrecht, pp. 3–33.

Wheeler, J. A. (1980): Law without law. In: Medawar, P., Shelley, J. (eds.) (1980): Structure in Science. Elsevier: New York, pp. 132–154.

Wheeler, J. A. (1983): On recognizing law without law. Am. J. Phys. 51, 398–404.

Windsor, H. H. (1907): Cart before the horse, Popular Mechanics 4, 425.

Wittgenstein, L. (1922): Tractatus Logico-Philosophicus. Kegan Paul, Trench, Trubner & Co.: London.

Woodward, J. (2003/2009): Scientific Explanation. The Stanford Encyclopedia of Philosophy; http://plato.stanford.edu/entries/scientific-explanation/

Yurov, A. V., Martín Moruno, P., González-Díaz, P. F. (2006): New “Bigs” in cosmology. Nucl. Phys. B 759, 320–341; arXiv:astro-ph/0606529

Zimorski, V., et al. (2014): Endosymbiotic theory for organelle origins. Current Opinions in Microbiology 22, 38–48.

Universal Ethics: Organized Complexity as an Intrinsic Value

Clément Vidal and Jean-Paul Delahaye

1 Introduction

Humans have developed multiple kinds of ethical systems, in different parts of the world, based on religious or humanist values. In our global and post-modern era, moral relativism makes a strong case, but by definition, it is not attempting to find a foundation for ethics.

Most of the time, ethical systems are anthropocentric in the sense that they value human happiness above anything else. Furthermore, a well-founded normative theory should also be able to answer questions not only about human values, but also about other value-related questions, such as aesthetic ones, for example: "why does a symphony have value in itself?"

Globally, ethical systems from various cultures are often mutually incompatible, and when they are interpreted dogmatically, i.e., as the only right way to assign value and to act in the world, they generate conflicts and violence at social levels: communal, societal, national, and international.

In this chapter, we propose a new concept for the foundation of a universal ethics. By "universal ethics" we wean that it aims to be universally applicable by *any* valuating agent, be it a human, organization, robot, software agent, or extraterrestrial being. It also aims to be able to give value to *any physical object* in the universe. Furthermore, because of its mathematical definition, it could also

---

C. Vidal (✉)
Center Leo Apostel & Evolution Complexity and Cognition, Vrije Universiteit Brussel, Brussels, Belgium
e-mail: contact@clemvidal.com

J.-P. Delahaye
Centre de Recherche en Informatique, Signal et Automatique de Lille (CRISTAL), Université de Lille, Lille, France
e-mail: jean-paul.delahaye@univ-lille.fr

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_4

apply to virtual, abstract, mathematical worlds (e.g., virtual reality, artificial life simulations, multiverse theories, mathematical proofs, etc.).

The "universal" declaration of human rights is not universal in this sense, as "universal" in the declaration refers simply to all human beings. It is unlikely that this declaration would make any sense to an extraterrestrial being in another galaxy, and it would be of no guidance to artificial agents in a virtual world. A purely anthropocentric ethics could also not say if it would be good or bad to annihilate a newly found exoplanet that was teeming with life.

Two kinds of universal ethical systems have been proposed. The first kind is based on *matter-energy* processes.

For example, *thermoethics*' central principle is to make the most of free energy, and to avoid the production of unnecessary waste, disorder and entropy (it is also called "entropy ethics"; see Ostwald 1912; Freitas Jr 1979, sec. 25.1.3; Hammond 2005; Korbitz 2010; Vidal 2014, chap. 10).

Another example of a matter-energy path toward a universal ethics could be based on the concept of *emergy* (with an "m"). The concept comes from systems ecology, and entails a measurement of energetic content. It is defined as the value of a system, be it living, social, or technological, as measured by the solar energy that was used to make it (e.g., Odum 2007).

Although matter-energy universal ethical systems such as thermoethics remain underexplored, our chapter will not focus on the possibility of such systems. Instead, we will focus on a second kind of universal ethics, based on *information* and *computation*. We can thus inscribe our approach within an ontology and metaphysics of information and computation (Delahaye and Vidal 2018). With the rise of the information society, and the importance and ubiquity of computers in our world, computation, information storage, and information exchanges are reshaping ourselves and our societies. A philosophy based on information and computation, therefore, is becoming more and more relevant (for an introduction to the field, see Floridi 2003).

An information and communication approach to universal ethics can be based on *cybernetics*, as it is a general science of control and communication (for some steps in this direction, see, e.g., Beer 1997; Chambers 2001; Vidal 2014, 285–86; Ashby 2017).

Focusing on information science, Floridi (2008) developed the concept of *infoethics*, which bears similarities with thermoethics (Vidal 2014, 271). To deepen and broaden traditional ethical systems, Ward Bynum (2006) initiated the seeds of a universal ethics based on information that can apply to every physical entity in the universe. However, the concept of information used in Bynum's approach is mostly *semantic* and has been criticized for its vagueness (e.g., Adriaans 2010). In this paper, we introduce and focus on a *syntactic* computational concept of valuable information, as the foundation of a universal ethics. This concept is *logical depth* (Bennett 1988). We think that this approach solves the issues that have been raised about infoethics: the basic notion is purely syntactic and mathematical, without the need to refer to cognitive agents such as humans. By contrast, the concept of semantic information that would be "well-formed, meaningful, and truthful data" is so delicate to define, and leads to so many difficulties, that it would risk obfuscating any ethical system based on it.

What are the desiderata of a universal ethics? A universal ethics should first be able to justify existing invariant values in humans, and provide a better understanding of why such commonalities exist. Empirical research has shown that there are broad invariants in human values: murder, theft, rape, lying, and destruction more generally are negative values in all societies, whereas health, wealth, friendship, honesty, safety, freedom, and equality are positive ones (e.g., Heylighen and Bernheim 2000).

A universal ethics should also be able to give value both to humans and their cultural products. These include works of art, scientific theories, books and libraries, or museums.

A universal ethics is also expected to support the widening of moral boundaries (see Fig. 1). We rarely see ethical concerns going beyond humanity. Fortunately, this is changing as consciousness is rising to care and value the nonliving Earth's climate, realizing that taking care of it is necessary at least for our long-term survival. Some religions such as Buddhism also do care about "all beings," which may include all life on Earth and in the Universe.

Such widening of moral boundaries has taken several shapes, from the rise of animal rights (Bentham 1907, chap. XVII; Regan 1987) to taking care of all living things on Earth, for example, with *biocentric ethics* (Agar 2001), *deep ecology* (Næss 2008), or *environmental ethics*. Indeed, the ecological worldview – according to which all living things and the Earth are connected – requires caring about nonhuman living processes that support the living realm.

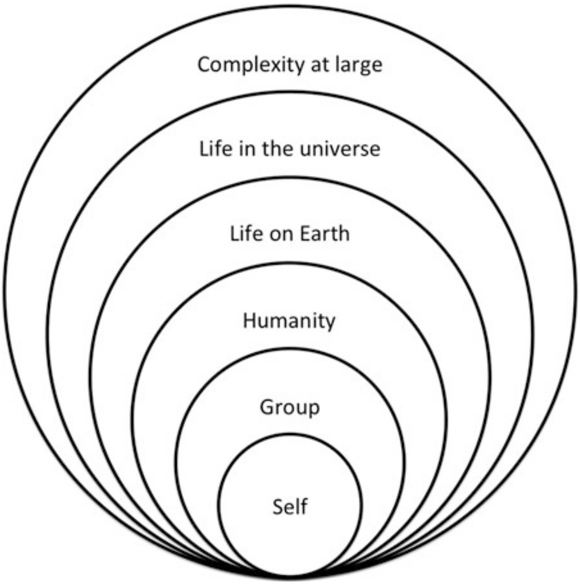


Fig. 1 The widening of moral boundaries. With increasing awareness, humans care for increasingly wider systems. The self cares for the group, which extends from a family to a community, a nation, and eventually to the whole of humanity. We extend the last circles of compassion to the notion of complexity at large, not limited to living things

In our digital era at the dawn of transhumanism, humans are transforming and re-shaping themselves through genetic engineering and technology. We are seeing the rise of the first augmented humans (or cyborgs), which raises entirely new moral issues that are hard to handle with traditional ethical systems. Whether we endorse transhumanism or not, we do need insights and frameworks to deal with current and future relations between humans and nonhumans. A related field is *machine ethics*, which deals with moral aspects of machine-machine interactions (e.g., Anderson and Anderson 2011).

Finally, a genuinely universal ethics should lead consistently to an *extraterrestrial ethics*, i.e. an ethics that can meaningfully apply to potential extraterrestrial life (see Vakoch 2014). This field explores issues regarding our obligations and duties toward any new life-form we may discover (see, e.g., issues raised by planetary protection programs in Meltzer (2010)). As a thought experiment, it is also challenging and mind broadening to think about the kinds of ethical behaviors and principles we can expect from potential extraterrestrial intelligence (see, e.g., Ruse 1989), as a way to prepare for the impact of the discovery of extraterrestrial life (Vidal 2015).

In this paper, we first discuss various conceptions of complexity, and introduce the notion of *organized complexity*, based on the computational concept of *logical depth*. We then show how organized complexity can be put to use as an intrinsic value, leading to three core imperatives: that we should *preserve* and *augment* existing organized complexity, as well as *recursively promote* systems that increase organized complexity. We illustrate our framework with some applications and examples. Finally, we discuss various issues that arise from this original computational approach to universal ethics.

2 Conceptions of Complexity

There is no doubt that multicellular organisms are more complex than unicellular ones. On large timescales, it is generally agreed that there is a growth in complexity of living beings that emerge from evolution (e.g., Coren 1998; Livio 2000; Mayfield 2013; Delahaye and Vidal 2018). Both biological and technological evolution produce increasingly complex objects, that is, displaying richer and richer structures.

A precise definition of such “complexity” is hard to achieve and to formulate mathematically (for a review, see Bennett 1990). Many ideas have been attempted, but these have often been inadequate. For example, Eric Chaisson (2001) defines complexity as the rate of energy flowing through a system, normalized by its mass. This idea of energy rate density seems compelling at first, as it can apply to many epochs of cosmic evolution, from galaxy formation, planet formation, to living systems and our technological society.

However, it has the drawback that it is totally blind to the concepts of computation or information. There is no stable correspondence between energy and computation. As a matter of fact, a version of Moore’s law shows that since the

1960s the number of computations per unit of energy has been doubling every 1.5 years (Koomey et al. 2011). A modern microprocessor is clearly more complex than a computer made 40 years ago, in the sense that for a given amount of energy, it can perform one million more operations. Structurally, a modern microprocessor also contains one million more transistors. Such aspects of complexity are not captured by a metric based on energy only.

Complexity can also rise in artificial life simulations, such as J. H. Conway's game of life (Berlekamp et al. 2001), in which energy plays no role. Another example is the complexity of a musical piece. It is not the piece's loudness, frequencies, or tempo that seem to determine its complexity, but rather its mathematical structure, without the need of any thermodynamic notion. Yet another example is a painting or a sculpture. According to Chaisson's metric, since no energy flows through them, they have zero complexity, and thus zero value if we were to take complexity as a guide for axiology.

Although useful to study ecosystem complexity (Ulgiati and Brown 2009), one can note that the concept of emergy is not a good general proxy for complexity. For example, a lot of energy is necessary to produce an ingot of aluminum, but this doesn't mean that an ingot of aluminum is particularly complex. It thus seems valid to search for a concept of complexity that is *not* founded on energy and its circulation – to seek, that is, a non-thermodynamic concept of complexity.

In theoretical computer science, two main notions of complexity have been proposed: *Kolmogorov complexity* and Bennett's *logical depth* (Li and Vitányi 2008; Bennett 1988; Delahaye 2009). These notions concern finite numerical objects, that we can translate – without loss of generality – to finite strings s of '0' and '1'. The Kolmogorov complexity of s is by definition the *size* of the shortest program s^* that outputs the string s (e.g., by printing it or by writing it in an output file). This is a useful notion in many respects, but it is not suitable as a measure of complexity if we consider complexity to be strongly structured. The notion does not cover the idea of complexification that scientists use when discussing biological or technological evolution. Indeed, strings that have the highest Kolmogorov complexity are random strings, and we intuitively know that randomness is the opposite of organized complexity.

Fortunately, based on Kolmogorov complexity, the notion of *logical depth* seems to be suitable as a definition of organized complexity, at least in a first approximation. The logical depth of s is by definition the *computation time* that it takes for the program s^* to produce s. In the case of an object s with low complexity (e.g., a repetitive string like 000000 . . . 00 or a random string), this computation time is minimal, whereas the more s can be regarded as a complex and structured object, the longer it takes for s^* to compute s.

For example, the string of the first million digits of π has a large logical depth, as well as the string of a musical piece translated into 0 and 1, or the string of a big sequence of prime numbers. The logical depth of s can either be seen as a measure of the quantity of structures in s, or as a quantity of computation present in s.

This definition of logical depth thus embeds two core values. First, because it is based on Kolmogorov complexity, it values the effort to find the shortest possible programs. This value mirrors the fundamental epistemic value of searching for the simplest and shortest models in science. Given some data to explain, a simple and short theory is to be preferred over a longer one. This is known by philosophers of science as Occam's razor, but many computer scientists have formalized it using Kolmogorov complexity, showing the links between machine learning and compression (e.g., Blumer et al. 1987; Li et al. 2003; Li and Vitányi 1992; 2008; Delahaye 1994).

The second value is that the longer it took for an object to appear (assuming that it cannot be obtained in a simpler way), the more value it has. In other words, the harder it would be to rebuild an object, the more value it has (Bennett 2014). This value is specific to the definition of logical depth, and its consequences will be explored in this chapter.

These computational concepts imply that we cannot naively consider the simple and the complex to be opposites. There are three notions at play: *simplicity*, *random complexity* (measuring the content in information via Kolmogorov complexity), and *organized complexity* (measuring the content in computation via logical depth). To augment organized complexity, one often needs to seek the simple: the simpler the computational mechanisms, the more efficient the production of computations will be. The simple is not opposed to organized complexity, it serves it and favors its growth.

Some aspects of normative evolutionary ethics valuing complexity growth and diversity are compatible with the ethics of organized complexity. For example, the survival instinct is clearly protecting the living complexity of the organism, while the reproductive instinct secures genetic information through generations. Mutations and sexual reproduction, coupled with environmentally induced selection, lead to a complexification of organisms on large timescales. The extent to which we should be inspired by evolutionary processes to build a normative ethics remains a huge debate that is well worth pursuing (see, e.g., Maienschein and Ruse 1999; Quintelier et al. 2011).

Evolution produces complex organisms by variation, inheritance and selection. Such elementary operations constitute what Dennett (1995) calls a "Darwinian algorithm" that we can assimilate to the creation of computational contents stored in living beings (see also Mayfield 2013). This algorithmic view can be naturally extended to cosmic evolution (Delahaye and Vidal 2018).

Complex organisms can only appear after less complex ones have already appeared. This corresponds to a *slow growth law* that has been demonstrated for Bennett's logical depth (Bennett 1988). It shows that organized complexity cannot appear suddenly, and that it requires a long time of maturation. This can be illustrated with an example from the history of science, namely, the now-refuted idea of spontaneous generation (Strick 2000). From this computational perspective, spontaneous generation could have been refuted a priori, as it would be extremely improbable that sophisticated, complex living systems would appear after only a few days.

In sum, an ethical foundation based on this definition of complexity is conceivable. The good becomes what contributes to the preservation and augmentation of structural contents, in nonliving entities, living beings, as well as humans and their cultural products, whether these contents are artistic, scientific, technological, economic, or political. Such an ethics suggests a respect and a protection not only of human and living beings, but also of all objects and structures that accompany them, or that further organizes them. Rich structures have required extensive computational work, and this is why we must protect them.

A work of art (such as a painting, symphony, or novel) is an object with high computational content, as it is the result of extensive elaboration; its organization is rich and non-trivial. As an object with high logical depth, i.e., *computational content*, it has value and should be respected and protected, according to an ethics of organized complexity. The same holds for a science book, a microprocessor, the form of organization of our societies, the networks of interconnections of our cities and countries, etc.

It is remarkable that some cryptocurrencies such as Bitcoin have at their core a record – called the blockchain – that has high computational content, i.e., great logical depth (Antonopoulos 2015). This makes them nearly unfalsifiable in practice, and constitutes a concrete example linking, in practice, the concept of logical depth with intrinsic value.

One may note that other ethics may be developed and supported with other existing or future conceptions of complexity, leading to somewhat different results, and this may be a worthwhile effort. However, it seems to us that the notion of logical depth lends itself naturally to the project of a non-thermodynamic universal ethics, as it offers a robust, adequate, and precise formal definition. We would like to show that the notion of logical depth encompasses and accommodates a wide range of existing values.

In what follows, we assume that logical depth captures satisfactorily the notions of richness in structure, high computational content, and organized complexity.

3 Three Imperatives

We identify three imperatives of the ethics of organized complexity: to *preserve*, *augment*, and *recursively promote* what preserves and augments organized complexity. As these are fundamental imperatives, additional ethical consequences may be derived from them, but this is an exercise which we will not attempt to do systematically in this paper.

3.1 Preserve Organized Complexity

The first imperative is to *preserve existing organized complexity*. Indeed, organized complexity took time and effort to appear, so it makes sense to preserve and protect it. We have built-in biological survival instincts that lead us to preserve ourselves and our offspring. Humans also have a tendency to systematically collect, process and store organized complexity. For example, in recent times, more and more organizations have made efforts to preserve biodiversity, or to protect endangered species, and also to protect other kinds of things that have required great effort in order to exist: a painting, a monument, and a patented idea. We are aware of the value of our cultural heritage, and that we should avoid the destruction of rare buildings, objects, or works of art. Historical and cultural preservation organizations exist worldwide, to preserve all kinds of complex human structures: consider for example UNESCO's world heritage sites, or many other national heritage protection programs.

In the digital world, this preservation imperative motivates us to implement effective strategies for backing up our data. Most of us have experienced data loss and know how costly this can be.

Even if the complexity collected has no immediate pragmatic value, it may have value in the future. This is true for biological evolution, where noncoding DNA sequences are conserved in the genome and may in future generations be activated. This is also true in mathematics, where theories that once had no practical use are nowadays central tools for science and technology. Classical examples include the use of non-Euclidean geometries for relativity theory or the use of arithmetic for modern cryptography. A similar dynamic is likely to apply in the future, which strongly implies that we should carefully preserve and make accessible theoretical scientific knowledge.

If systematic collection and storage is not possible, we need to think about heuristics regarding what to collect, what to store, and what to make most accessible. In computer science, this is related to the recurrent problem of managing memory (space) and speed (time), which leads to space-time tradeoffs.

One may object that losing old *structural* products of complexity is not so grave, as long as we are able to *preserve their function*. For example, even if Galileo's original telescope design is not used anymore, we have much more powerful and reliable telescopes today, so the loss is limited. Another example is the computer. Nobody misses vacuum tube computers; what is important is that we have general-purpose computers that can accomplish the same operations as these earlier computers. The preservation of organized complexity, in other words, can be both structural and functional.

Also, as a first heuristic, one may suggest that we should adapt our preservation strategy relative to the number of copies. If there are 100 copies of a book, it is less grave to destroy one of them, than to destroy the last copy of a book.

3.2 Augment Organized Complexity

One can emphasize the preservation and conservation of the old, but the creation of the new, and the augmentation of existing organized complexity, are just as valuable and important. We saw that the slow growth law (Bennett 1988) implies that we can't just quickly create deeply complex objects out of nothing. We need time to build on previous efforts, which leads us to our second imperative: *augment organized complexity*.

This raises the question: how should we augment organized complexity? This is the issue of the *distribution of organized complexity*: should we try to augment total organized complexity, or average organized complexity? The issue is similar to the classical issue of total versus average utility in utilitarianism (Sidgwick 1907). The issue is itself a particular instance of a more general problem of optimal allocation of resources, in defining social welfare functions (Chevaleyre et al. 2006).

To define the issue more precisely, let us imagine that we have a partition of the universe into well-defined components (humans, countries, celestial bodies, galaxies, etc.), that we denote as $C_1, C_2, \ldots, C_n$. There are different goals that one may want to pursue, in order to preserve and augment organized complexity. For example:

Goal 1: Augment organized complexity as a whole, as the union of: $C_1 + C_2 + \ldots + C_n$.

Assuming that this organized complexity is measured adequately by logical depth (LD), the goal is thus to maximize $LD(C_1 + C_2 + \ldots + C_n)$ that we note as:

$$\text{maximize}(LD(C_1 + C_2 + \cdots + C_n)).$$

This may be called a *global* conception.

Goal 2: Augment the sum of organized complexity inside the different components:

$$\text{maximize}(LD(C_1) + LD(C_2) + \cdots + LD(C_n)).$$

This is not the same as Goal 1, because if C_1 and C_2 are identical, we have:

$LD(C_1 + C_2) \approx LD(C_1) < LD(C_1) + LD(C_2)$. For example, if there are two identical books in one library they have almost the same value ($LD(C_1 + C_2) \approx LD(C_1)$). Striving for goal 2, the complexity of C_1 and C_2 will be counted in each component, and thus two times. The emphasis here is thus less on producing new complexity overall, but rather on distributing it evenly. In this context, we may call this goal 2 an *additive* conception.

Goal 3: Maximize the least structurally complex components:

$$\text{maximize}(\text{minimum}(LD(C_1), LD(C_2), \ldots, LD(C_n)).$$

This is an *egalitarian* conception, aiming to maximize the richness of the poorest.

Goal 4: Maximize the organized complexity of the best components:

$$\text{maximize}(\text{maximum}(LD(C_1), LD(C_2), \ldots, LD(C_n)).$$

In this case, it is an *elitist* conception: it doesn't matter if some components have low organized complexity, what matters is increasing the organized complexity of the already-highest complexity components.

Goal 5: Maximize the product of organized complexity:

$$\mathrm{maximize}(\mathrm{LD}(\mathrm{C}_1) \times \mathrm{LD}(\mathrm{C}_2) \times \cdots \times \mathrm{LD}(\mathrm{C}_\mathrm{n})).$$

This is a compromise between goal 3 and goal 4 that avoids over-penalizing certain components. This conception is used in the optimal allocation of resources and is called *Nashian* (Ramezani and Endriss 2010).

We will not discuss how one might settle these different goals and viewpoints, as each of them has arguments in its favor. What we do want to argue is that the richness of a component can be measured by its organized complexity, and that this is a universal and coherent way to approach this issue.

Let us note that as soon as a metric to measure the value of components in the world is given, in any valuation system, this allocation of resources issue will arise, and we will need to choose between a global, additive, egalitarian, elitist, or Nashian allocation system. Even if the measure is not clear-cut or possible in practice, the different goals underlie different philosophical, political, and ethical choices. To decide between the different goals requires the development of an applied ethics. Therefore this issue is not a weakness of our proposal in particular, as any foundational principle for ethics (e.g., maximizing human happiness) needs to decide between such goals when put into practice.

It is also worth noting that informational resources are non-rival, and can be shared with negligible costs compared to rival, matter-energy resources. So different strategies and treatments might be necessary for the distribution of rival and non-rival resources. For example, it is natural and easy to be egalitarian and to share knowledge with all humans via the internet, whereas doing the same with oil is much more problematic.

3.3 Recursively Promote Organized Complexity

The third imperative is to *recursively promote what preserves and augments organized complexity*. In other words, it is to create, value and assist systems and strategies that can preserve and augment organized complexity, to the n^{th} order.

Let us illustrate this imperative with two examples: the mathematician and the musician. At the recursive level 0, the mathematician finds a new theorem, and the composer composes a new piece of music. They can make efforts to diffuse their works, for example, by writing a book or recording a CD. This is a level 1 effort, as it limits the risk that the organized complexity created could simply disappear.

A level 2 effort, for example, the funding or founding of a multimedia library, will allow the preservation and diffusion of organized complexity when the book or CD is released. A level 3 effort would include, for example, participating and helping a government whose goal was to collect taxes in order to fund the building of libraries.

One can note that a library security system does not have a strong intrinsic complexity, but is still a valuable aspect of preserving existing organized complexity, recursively. Jacques Monod (1972, 180) defended such a higher order way to preserve knowledge, with his *ethic of knowledge* that "prescribes institutions dedicated to the defense, the extension, the enrichment of the transcendent kingdom of ideas, of knowledge, and of creation."

Generally, copies also help realize this recursive imperative. For example, libraries that store copies of books help to further build complexity, as they give access to existing deeply complex work, on which further complexity can be built. In our digital era, it becomes obvious that all kinds of open source and open access initiatives are cheap and highly beneficial, and should therefore be promoted (e.g., Heylighen 2007; Steele 2012).

There are indirect ways to promote organized complexity, namely to ensure that its supporting systems are effective. For example, we need to care about Earth's climate for the preservation and augmentation of biological complexity. Earth's climate is nonhuman and nonliving, but it should still be taken care of. Another example is the requirement of energy to build organized complexity. Energy has value in the sense that it could potentially be used to build organized complexity, which ties in with the values of thermoethics. Unfortunately, these recursive and indirect ways to promote organized complexity make the assessment of value more complicated in practice.

4 Applications and Examples

If one adopts organized complexity as a universal value, then it becomes possible to naturally recover a large number of values that are already accepted by many ethical traditions. Let's consider a few examples.

Every human being is a complex construct resulting from one's genes, learning, and experiences. Each individual human construction is unique and has a value of logical depth which, even if we do not know how to measure it precisely, is clearly very high. To kill a human being is, from the point of view of the ethics of organized complexity, a bad action. Similarly, anything that degrades, disturbs, or hurts a human being, by making it less effective and simply by damaging its structural richness, must be recognized as bad from this point of view.

The ethics of organized complexity commands us, in the same way, to respect and protect animals and, in a general way, all living things. Interestingly enough, this ethics gives special importance to endangered species, as their members carry an organized complexity that would be impossible to recover should the species disappear. The difference of care that should be devoted to a member of a species represented by millions of other members, as opposed to just a few, is naturally taken into account by the ethics of organized complexity. This idea is far from new. For example, in the biblical myth of Noah's ark, God commands Noah to protect at least one sexual couple of each existing species. In this way, Noah not only *preserves* existing biological complexity but, thanks to the instruction to protect both sexes, also enables their reproduction and thus the *recursive promotion* of complexity.

In the case of human beings, a strict application of the ethics of organized complexity would at first sight lead to assigning more value to a genius than to the average person. However, we generally assume that all human lives have equal value. To suppose the opposite would seem to create serious social and political difficulties, and the solution that seems most compatible with democracy, and hence global social efficiency would be to regard all humans as having equal rights.

Another example of the immediate application of the ethics of organized complexity concerns works of art. These have obvious structural content, and the recognized talent of the artists who produce them is linked to their ability to elaborate (to calculate) complex, original structures in novel ways. Such creativity is a form of logical depth. Even if this computational content is not apparent in the work itself, when regarded as an isolated object – as with a work of minimalist art, for example – this content may still be present in the new relations which it establishes between the world and the work of art.

Still, generally speaking, we recognize that a work of art has value in proportion to its internal structural richness and its novelty. When it offers a novel perspective on the world, we assign it value in proportion to the fineness and subtlety of what it implies about the world. This value represents a new form of structural wealth, established by the work of art. In most cases, what makes us recognize value in a work of art can be interpreted as inherent structural richness, and therefore the idea that works of art must be preserved and protected is a direct application of the principles of the universal ethics of organized complexity.

Music is a particularly striking example of the purely structural content we perceive, and humans value many different kinds of music styles. While the development of our musical preferences will make us appreciate differently European or Indian music, for example, with effort we can learn to appreciate even music to which we have not been accustomed. In the end, what we love about music, and what makes a musical work worthwhile for us, is its richness in structure.

Science, too, can fall within the value scheme proposed by the ethics of organized complexity. The most important scientific theories that would be most morally condemnable to forbid, or to not diffuse, are those which required greater effort (experimental, conceptual, mathematical, etc.). Of course, by "effort," we mean a well-formed and well-informed effort. Those grand, deeply complex works have a great content in computation or, equivalently, in structure. Here our proposition of universal ethics uses its single, homogeneous concept to accommodate a fundamental value of scientific practice.

5 Discussion and Objections

5.1 Organized Complexity and Destruction

A delicate question that arises, if we adopt organized complexity as a value, is the erasure of data or, equivalently, the destruction of structures. Destruction seems to directly oppose the promotion of organized complexity. Yet, at least three factors determine the "right" decision to make in concrete cases.

First it depends on the existence or non-existence of other copies of the data or structure in question. If there are many copies, erasing a few redundant copies will not affect the global organized complexity. One may thus argue that their destruction is not (so) "bad."

However, as we saw, there is a second factor one may want to include, if one adopts a principle of egalitarianism for managing organized complexity. In this case, destroying complexity at one point could have the effect of lowering the organized complexity of the component considered and would thus be "bad."

A third difficulty arises when taking into account our third imperative of recursively promoting organized complexity. This recursive promotion often requires us to keep, at least temporarily, some data or structures. The problem is similar to that of information management in a computer system: efficiency sometimes requires that the same data be copied several times, in order to have optimal access to it, and thus to compute the desired results more quickly. Here, even if it is not immediately useful to keep the intermediate data or structures in question, because they do not contribute directly to the final desired result, it is possible that keeping copies of such data guarantees a better creative potential and therefore that the "good" choice is to destroy nothing.

Another aspect of this algorithmic issue is that some data are easy to reconstruct, and thus keeping them clutters the space or memory of the system. Good management then requires the destruction of such data. For example, modern algorithms for testing the primality of a number, or for the fast generation of prime numbers, make it unnecessary to build and store large tables of prime numbers. However, it made sense to do so in a world without computers, and indeed, in the nineteenth century, the building of such tables gave rise to important works and publications. In this case, the technological progress of computing has tipped the right decision from "keep" to "delete."

In sum, even by adopting a mathematically well-defined point of view for assessing value, the precise determination of actions recommended by an ethics of organized complexity does not lead in practice to simple solutions. On the contrary, it preserves all of the difficulties that are necessary for conceiving efficient algorithms.

5.2 Organized Complexity and Traditional Ethical Issues

Let us outline a few connections between our proposal and some traditional ethical issues.

It may be possible – and desirable – to consider the rise of organized complexity as the utility function in a utilitarian framework (instead of, for example, maximizing human happiness). Of course, it remains very hard to foresee what actions will preserve and augment complexity, especially if we consider our third principle of recursively promoting organized complexity.

However, this issue is not specific to our approach, as any consequentialist must have good models of the world, to anticipate the future as well as possible. This leads to the cognitive value of modeling, i.e., the anticipation of our actions and the world. On a short timescale, we can all agree that having the reflex to avoid a falling rock, or the skills to climb a tree to collect a delicious fruit, would be "good" and adaptive. We need similar skills at higher spatial and temporal levels, to be able to anticipate and model the future of groups, societies, Earth, the universe, and complexity at large.

Instead of taking organized complexity as an intrinsic *normative value*, one could also use it as a *descriptive value*, in the sense that it may help to explain why an agent values certain behaviors or objects. One may thus try to explain an existing moral valuation in terms of its preservation and augmentation of organized complexity.

We can recover the value of striving for happiness, but not put it at the foundation of our ethics. For example, pleasures and pains evolved to ensure survival and reproduction, i.e., to preserve the organized complexity of our selves. Of course, we acknowledge that there are many reasons to promote happiness: for example, stress diminishes fitness for reproduction (see, e.g., Moberg 1985), whereas happy people have positive emotions that allow them to broaden and build their selves, relationships, and environment (e.g., Fredrickson 2004).

Unfortunately the problem of assessing logical depth belongs to the class of non-computable problems. This means that we would rarely be able to prove that a fixed number is the correct logical depth value of a specific s, and that instead we must be content with algorithms that approximate LD(s). However, this non-computability is not a fundamental obstacle to its use, as with Kolmogorov complexity, which is also non-computable and is still widely used for concrete applications (e.g., Varré et al. 1999; Belabbes and Richard 2008). So we can reasonably hope that similar tools can be used to approximate logical depth (for early attempts, see, e.g., Zenil et al. 2012; Gauvrit et al. 2017).

Another issue is the measurement of the depth of an isolated object. To what do we compare its organized complexity? For example, from what point in the past do we consider the history of our object? To assess the computational content of a human being, do we say that it starts with its birth, or its parent's birth, or should we go back to the origin of life, or even to the origin of the universe that allowed its atoms to exist? This is the *temporal boundary issue* that is not specific to our approach, as it also appears in holistic approaches, such as the idea of emergy (Hau and Bakshi 2004, 221a): should we take into account all the solar energy that has been used since the birth of the solar system to assess a content in solar energy of an object? If not, from when do we start?

In practice, we can't yet compute the value of any choice to solve systematically ethical issues. Instead a multiple-level ethics requires us to solve the question: *What action preserves and augments complexity at all levels?* For example, deciding on whether to abort a fetus or not may require multiple problem-solving considerations at multiple levels, including biological, psychological, familial, societal, or religious aspects. Often, philosophers emphasize a central moral conflict: that between the interests of the individual and the collective. But this conflict need not be limited to just two levels; as we grow our circles of compassion, many more levels must be included (see Fig. 1). Again, this delicate issue regarding multiple levels is not specific to the ethics of organized complexity, but is a general problem in any complex issue, ethical or not, that involves multiple levels, aspects, and stakeholders.

5.3 *Anticipating Some Misunderstandings*

5.3.1 Intentionality

To determine if an action is good or bad is a question of intentionality, why did you not address this issue?

The issue of intentionality is separate. For example, if X kills Y, in all ethical views that value human lives, this is bad, whether the act of X was intentional or not. The question of determining the responsibility of X, how X should be punished, is a different one than knowing if killing Y was good or bad. The problem of moral assessment is different from the problem of determining responsibility. The latter has to do with juridical and penal domains of knowledge. To put X in jail because he was not careful could be justified, but if Y committed suicide by jumping in front of the train that X was driving, this would make no sense.

5.3.2 Politics

How could we compare a dictatorial and a democratic regime, if both have about the same complexity?

To evaluate them, we would need to evoke the three imperatives *together*. We must analyze the capacity of these societies to produce more complexity in the future (second and third imperatives). If it was established that totalitarian societies were better at preserving organized complexity, and steering the creation of new complexity at all levels, then they should be preferred. However, it seems that the opposite is true, as democracies seem much more favorable to diversity, the flourishing of people, art, and science.

5.3.3 Applicability

Isn't this universal ethics too abstract and without any possible application?

Applications of ethical principles are delicate steps to take, as we saw for example with the problem of the distribution of organized complexity. Any ethical approach that wants to become precise and applied must face such issues.

For example, if one considers that *human happiness* is the ultimate good, this raises many questions about how to apply it in practice. Do we want to maximize the happiness of the most unhappy? Or the average or maybe the median happiness? Do we want to maximize the sum of happiness on Earth? Should we aim at making many humans averagely happy? Or should we aim at less happy humans, but more so on average? This is without mentioning the problem of measuring happiness.

In any ethical system, there are countless difficulties to bridging the gap between founding principles and determining real life actions. We have proposed a unified and mathematically based foundation that may ease this transition, although it remains a difficult problem.

5.3.4 Tastes and Complexity

Isn't your proposal inaccurate, as some of our tastes show that value is not linked to complexity?

We did not argue that our tastes are always linked to what is the most complex; this is clearly wrong. Following one's values often demands effort. For example, one can prefer an airport novel to the hard work of a Nobel prize-winning novel, but if one had to destroy the last remaining copy of one or the other, one could still prefer to keep the Nobel prize-winning novel. One could also feel that to indulge in the consumption of the easiest-to-read novels would not always be the ethically best choice.

5.3.5 Ethical Value and Market Value

Doesn't the market value of art – and minimalist art in particular – show that complexity has nothing to do with value?

Market value is generally not linked to the value of organized complexity (even if it may happen). The market value of a work of art and of goods in general depends on many parameters, such as rarity, tastes of buyers, etc. This divergence between market value and ethical value exists in all value systems: a lethal weapon will rarely have a great ethical value, and yet could be expensive.

5.3.6 Murderer and Society

Shouldn't even a murderer, because of his biological structural complexity, be considered as good and worth respecting?

Of course not, because this first analysis is too quick and insufficient: it is restricted to only one level (the individual's complexity) and uses only the first imperative: "preserve organize complexity." Again, the three principles work *together*, and sometimes to preserve might be less important than to augment or to promote recursively organized complexity.

A murderer has killed another human being and therefore destroyed organized complexity. The murderer's potential future actions also put in danger the normal functioning of society and other human beings. Leaving the murderer free would create further societal problems, stresses, and chaos, while executing the murderer – as it would have been done in the past – is not the best solution either, at least because it destroys the structural complexity of the murderer as a human being. So the problem is multi-level, and becomes one of preserving the complexity of *both* the murderer and society. It is generally at both of these levels that modern societies try to find solutions.

6 Conclusion

We have explored a possible foundation for ethics, by showing that organized complexity can be treated as an intrinsic and universal value. Such a non-anthropocentric, universal ethical foundation is much needed in our digital era. We have argued that organized complexity measures an intrinsic value: the history of non-trivial steps that have occurred to produce an object.

We put forward three imperatives of the ethics of organized complexity: to *preserve*, *augment*, and *recursively promote* what preserves and augments organized complexity. There are still many difficulties that lie ahead to apply those imperatives in practice. However, we saw that many such difficulties also exist in other ethical theories, and thus do not constitute a weakness of our approach in particular.

The potential of such a universal ethics is great, as it could be used to develop transhumanist ethics, machine ethics, or extraterrestrial ethics. It validates many existing notions of good and bad, such as the value of endangered species or works of art. An original non-anthropocentric conclusion is that deeply complex, inanimate objects also have value. In sum, we expect that the development of the ethics of organized complexity will shed light on past, present, and future ethical issues.

References

Adriaans, Pieter. 2010. “A Critical Analysis of Floridi’s Theory of Semantic Information.” *Knowledge, Technology & Policy* 23 (1–2): 41–56. doi:https://doi.org/10.1007/s12130-010-9097-5.

Agar, Nicholas. 2001. *Life’s Intrinsic Value ; Science, Ethics, and Nature*. New York: Columbia University Press.

Anderson, Michael, and Susan Leigh Anderson. 2011. *Machine Ethics*. Cambridge University Press.

Antonopoulos, Andreas M. 2015. *Mastering Bitcoin*. Sebastopol CA: O’Reilly.

Ashby, Mick. 2017. “Ethical Regulators and Super-Ethical Systems.” In *61st Conference of the International Society for the Systems Sciences*. Vienna, Austria. http://ashby.de/Ethical%20Regulators.pdf.

Beer, Stafford. 1997. “The Culpabliss Error: A Calculus of Ethics for a Systemic World.” *Systems Practice* 10 (4): 365–80. doi:https://doi.org/10.1007/BF02557886.

Belabbes, Sihem, and Gilles Richard. 2008. “Spam Filtering without Text Analysis.” In *Global E-Security*, edited by Hamid Jahankhani, Kenneth Revett, and Dominic Palmer-Brown, 144–52. Communications in Computer and Information Science 12. Springer Berlin Heidelberg. doi:https://doi.org/10.1007/978-3-540-69403-8_18.

Bennett, C. H. 2014. *Evidence, Computation, and Ethics*. Simons Symposium on Evidence in the Natural Sciences.

Bennett, C.H. 1988. “Logical Depth and Physical Complexity.” In *The Universal Turing Machine: A Half-Century Survey*, edited by R. Herken, 227–257. Oxford University Press. http://researcher.watson.ibm.com/researcher/files/us-bennetc/UTMX.pdf.

———. 1990. “How to Define Complexity in Physics and Why.” In *Complexity, Entropy, and the Physics of Information*, edited by Wojciech H. Zurek, 137–48. Santa Fe Institute Studies in the Sciences of Complexity, v. 8. Redwood City, Calif: Addison-Wesley Pub. Co.

Bentham, J. 1907. *An Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press. http://www.econlib.org/library/Bentham/bnthPML0.html.

Berlekamp, Elwyn R., John Horton Conway, and Richard K. Guy. 2001. *Winning Ways for Your Mathematical Plays*. 2nd ed. Natick, Mass: A.K. Peters.

Blumer, Anselm, Andrzej Ehrenfeucht, David Haussler, and Manfred K. Warmuth. 1987. “Occam’s Razor.” *Information Processing Letters* 24 (6): 377–380. http://www.cse.buffalo.edu/~hungngo/classes/2008/694/papers/occam.pdf.

Bynum, Terrell Ward. 2006. “Flourishing Ethics.” *Ethics and Information Technology* 8 (4): 157–73. doi:https://doi.org/10.1007/s10676-006-9107-1.

Chaisson, E. J. 2001. *Cosmic Evolution: The Rise of Complexity in Nature*. Harvard University Press.

Chambers, Jean. 2001. “A Cybernetic Theory of Morality and Moral Autonomy.” *Science and Engineering Ethics* 7 (2): 177–92. doi:https://doi.org/10.1007/s11948-001-0039-1.

Chevaleyre, Yann, Paul E. Dunne, Ulle Endriss, Jérôme Lang, Michel Lemaitre, Nicolas Maudet, Julian Padget, Steve Phelps, Juan A. Rodriguez-Aguilar, and Paulo Sousa. 2006. “Issues in Multiagent Resource Allocation.” *Informatica* 30 (1). http://wen.ijs.si/ojs-2.4.3/index.php/informatica/article/viewFile/70/62.

Coren, Richard L. 1998. *The Evolutionary Trajectory: The Growth of Information in the History and Future of Earth*. The World Futures General Evolution Studies, v. 11. Amsterdam, Netherlands: Gordon and Breach Publishers.

Delahaye, Jean-Paul. 1994. *Information, complexité et hasard*. Paris: Hermes Science Publications.

———. 2009. *Complexité Aléatoire et Complexité Organisée*. Sciences En Questions. Versailles: Quae.

Delahaye, Jean-Paul, and Clément Vidal. 2018. “Organized Complexity: Is Big History a Big Computation?” *American Philosophical Association Newsletter on Philosophy and Computers* 17 (2): 49–54. http://arxiv.org/abs/1609.07111.

Dennett, Daniel Clement. 1995. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster.

Floridi, L., ed. 2003. *The Blackwell Guide to the Philosophy of Computing and Information*. Blackwell Publishing.

———. 2008. "Information Ethics: Its Nature and Scope." In *Information Technology and Moral Philosophy*, edited by Jeroen van den Hoven and John Weckert, 40–65. Cambridge University Press.

Fredrickson, Barbara L. 2004. "The Broaden-and-Build Theory of Positive Emotions." *Philosophical Transactions of the Royal Society B: Biological Sciences* 359 (1449): 1367–78. doi:https://doi.org/10.1098/rstb.2004.1512.

Freitas Jr, R. A. 1979. *Xenology: An Introduction to the Scientific Study of Extraterrestrial Life, Intelligence, and Civilization*. Xenology Research Institute (1979–2010). http://www.xenology.info/Xeno.htm.

Gauvrit, Nicolas, Hector Zenil, Fernando Soler-Toscano, Jean-Paul Delahaye, and Peter Brugger. 2017. "Human Behavioral Complexity Peaks at Age 25." *PLOS Computational Biology* 13 (4): e1005408. doi:https://doi.org/10.1371/journal.pcbi.1005408.

Hammond, Dick. 2005. *Human System from Entropy to Ethics*. Eulogy Edition.

Hau, Jorge L, and Bhavik R Bakshi. 2004. "Promise and Problems of Emergy Analysis." *Ecological Modelling*, Through the MACROSCOPE: the legacy of H.T. Odum, 178 (1–2): 215–25. doi:https://doi.org/10.1016/j.ecolmodel.2003.12.016.

Heylighen, F. 2007. "Why Is Open Access Development so Successful? Stigmergic Organization and the Economics of Information." In *Open Source Jahrbuch 2007*, edited by B. Lutterbeck, M. Baerwolff & R. A. Gehring, 165–80. Lehmanns Media. http://pespmc1.vub.ac.be/Papers/OpenSourceStigmergy.pdf.

Heylighen, F., and J. Bernheim. 2000. "Global Progress I: Empirical Evidence for Ongoing Increase in Quality-of-Life." *Journal of Happiness Studies* 1 (3): 323–49. doi:https://doi.org/10.1023/A:1010099928894. http://pespmc1.vub.ac.be/Papers/ProgressI&II.pdf.

Koomey, J., S. Berard, M. Sanchez, and H. Wong. 2011. "Implications of Historical Trends in the Electrical Efficiency of Computing." *IEEE Annals of the History of Computing* 33 (3): 46–54. doi:https://doi.org/10.1109/MAHC.2010.28.

Korbitz, Adam. 2010. "Metalaw and SETI: Xenology, Metalaw and Thermoethics." *Metalaw and SETI*. December 2. http://web.archive.org/web/20110708053037/http://metalawandseti.blogspot.com/2010/12/xenology-metalaw-and-thermoethics.html.

Li, Ming, John Tromp, and Paul Vitányi. 2003. "Sharpening Occam's Razor." *Information Processing Letters* 85 (5): 267–74. doi:https://doi.org/10.1016/S0020-0190(02)00427-1.

Li, Ming, and P. M. B. Vitányi. 1992. "Inductive Reasoning and Kolmogorov Complexity." *Journal of Computer and System Sciences* 44 (2): 343–84. doi:https://doi.org/10.1016/0022-0000(92)90026-F.

———. 2008. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer.

Livio, M. 2000. *The Accelerating Universe*. New York, Wiley & Sons.

Maienschein, Jane, and Michael Ruse, eds. 1999. *Biology and the Foundation of Ethics*. Cambridge Studies in Philosophy and Biology. Cambridge, UK ; New York: Cambridge University Press.

Mayfield, John. 2013. *The Engine of Complexity: Evolution as Computation*. Columbia University Press.

Meltzer, Michael. 2010. *When Biospheres Collide: A History of NASA's Planetary Protection Programs*. NASA Aeronautics and Space Administration. https://www.nasa.gov/pdf/607072main_WhenBiospheresCollide-ebook.pdf.

Moberg, Gary P. 1985. "Influence of Stress on Reproduction: Measure of Well-Being." In *Animal Stress*, 245–67. Springer, New York, NY. doi:https://doi.org/10.1007/978-1-4614-7544-6_14.

Monod, Jacques. 1972. *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology*. New York: Vintage Books.

Næss, Arne. 2008. "The Shallow and the Deep, Long-range Ecology Movement. A Summary." *Inquiry*, August. doi:https://doi.org/10.1080/00201747308601682.

Odum, Howard T. 2007. *Environment, Power, and Society for the Twenty-First Century: The Hierarchy of Energy*. New York: Columbia University Press.

Ostwald, Wilhelm. 1912. *Der Energetische Imperativ*. Leipzig: Akademische verlagsgesellschaft m. b. h. http://archive.org/details/derenergetische00ostwgoog.

Quintelier, Katinka, Linda Van Speybroeck, and Johan Braeckman. 2011. "Normative Ethics Does Not Need a Foundation: It Needs More Science." *Acta Biotheoretica* 59 (1): 29–51. doi:https://doi.org/10.1007/s10441-010-9096-7.

Ramezani, Sara, and Ulle Endriss. 2010. "Nash Social Welfare in Multiagent Resource Allocation." In *Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets*, 117–31. Lecture Notes in Business Information Processing. Springer, Berlin, Heidelberg. doi:https://doi.org/10.1007/978-3-642-15117-0_9. https://pure.uva.nl/ws/files/20631648/74912_312938.pdf.

Regan, Tom. 1987. "The Case for Animal Rights." In *Advances in Animal Welfare Science 1986/87*, 179–89. Springer, Dordrecht. doi:https://doi.org/10.1007/978-94-009-3331-6_15.

Ruse, Michael. 1989. "Is Rape Wrong on Anromeda?" In *The Darwinian Paradigm: Essays on Its History, Philosophy, and Religious Implications*, 209–45. London; New York: Routledge.

Sidgwick, Henry. 1907. *The Methods of Ethics*. Hackett Publishing.

Steele, Robert David. 2012. *The Open-Source Everything Manifesto: Transparency, Truth, and Trust*. Berkeley, Calif: Evolver Editions.

Strick, James Edgar. 2000. *Sparks of Life: Darwinism and the Victorian Debates over Spontaneous Generation*. Cambridge, Mass: Harvard University Press.

Ulgiati, Sergio, and Mark T. Brown. 2009. "Emergy and Ecosystem Complexity." *Communications in Nonlinear Science and Numerical Simulation* 14 (1): 310–21. doi:https://doi.org/10.1016/j.cnsns.2007.05.028.

Vakoch, Douglas A., ed. 2014. *Extraterrestrial Altruism: Evolution and Ethics in the Cosmos*. The Frontiers Collection. Springer Berlin Heidelberg.

Varré, J. S., J. P. Delahaye, and E. Rivals. 1999. "Transformation Distances: A Family of Dissimilarity Measures Based on Movements of Segments." *Bioinformatics* 15 (3): 194–202. doi:https://doi.org/10.1093/bioinformatics/15.3.194.

Vidal, C. 2014. *The Beginning and the End: The Meaning of Life in a Cosmological Perspective*. New York: Springer. http://arxiv.org/abs/1301.1648.

———. 2015. "A Multidimensional Impact Model for the Discovery of Extraterrestrial Life." In *The Impact of Discovering Life Beyond Earth*, edited by Steven J. Dick, 55–75. Doi:https://doi.org/10.1017/CBO9781316272480.006. Cambridge: Cambridge University Press. http://tinyurl.com/vidal2015b.

Zenil, Hector, Jean-Paul Delahaye, and Cédric Gaucherel. 2012. "Image Characterization and Classification by Physical Complexity." *Complexity* 17 (3): 26–42. doi:https://doi.org/10.1002/cplx.20388. http://arxiv.org/abs/1006.0051.

Inductive Probabilities in Astrobiology and SETI: Have Sceptics Retreated?

Milan M. Ćirković

> *For the atoms being infinite in number, as has just been proved, are borne ever further in their course. For the atoms out of which a world might arise, or by which a world might be formed, have not all been expended on one world or a finite number of worlds, whether like or unlike this one. Hence there will be nothing to hinder an infinity of worlds.*
>
> —Epicurus, *Letter to Herodotus* (cca. 300 BC)

1 Introduction: Probabilities and Extraterrestrial Life

Since we have very little empirical insight into the origin of life (abiogenesis) and intelligence (noogenesis), much of what we say about these two crucial processes is of necessity highly speculative. Obviously, this pertains to abiogenesis and noogenesis on Earth (which are observer-selected local facts) as well as anywhere else in the universe. Both necessary and sufficient conditions for either process are highly uncertain in absolute terms; therefore, the usual approach has traditionally been to speculate on how these could be characterized in relative terms, i.e., relative to the conditions prevailing on Earth in their respective epochs (late Hadean for abiogenesis and early Quartenary for noogenesis[1]). This relational perspective leads

---

[1] For example, Chernavskii (2000); Luisi (2006); Dodd et al. (2017). Here and elsewhere in this paper, I assume the standard view of history of life and intelligence, neglecting highly non-standard scenarios such as life being brought on Earth via directed panspermia or the possibility that some late dinosaurs could have been intelligent, etc.

M. M. Ćirković (✉)
Astronomical Observatory of Belgrade, Belgrade, Serbia

Future of Humanity Institute, Faculty of Philosophy, University of Oxford, Oxford, UK
e-mail: mcirkovic@aob.rs

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_5

to various *probabilistic* arguments about these crucial evolutionary steps: attempts to derive probabilities for these events based on conjunction of various prerequisites and, in the second step, using such probabilistic estimates to justify uniqueness of terrestrial life and intelligence (plus various consequences drawn thereof, from lobbying for cessation of our SETI efforts to elevating Earth's biosphere and humans to specially exalted moral status). Of course, this reasoning is better known as an argument for non-naturalistic origin of life and intelligence, but that is less important for us here.

Since we know for only a single instance of both abiogenesis and noogenesis thus far, and we are obviously strongly biased by the observation-selection effects (Bostrom 2002), such probabilities are bound to be only very loosely related to anything in the real world. This quite basic epistemological insight has been, however, consistently downplayed or ignored. While we may not be able to correct for some of these biases (at least until we get the opportunity to study other habitable worlds), it is extraordinarily important to *acknowledge* their existence and their impact on any conclusions we derive about extraterrestrial life and intelligence.

This topic deserves separate attention from the purely astrobiological problem of existence (or else) of extraterrestrial biospheres or extraterrestrial intelligence. In the last 60 years or so, there have been many uses and abuses of probability and probabilistic arguments in discussing life and intelligence in their general cosmic context. Here we wish to focus on frequent alleging very low probabilities for the existence of other intelligent species in the Galaxy, or even the entire visible universe. Sceptics on the issues of extraterrestrial life and intelligence have time and again suggested that the discovery of extraterrestrial life and intelligence in general, and the success of our SETI projects in particular, have very small probabilities, implying that there is a scientifically sound way one can derive such negligible chances. Of course, the next step has always been that such minuscule probabilities make our searching efforts misplaced, unfounded, and wasteful. In most such sceptical discourses, less attention has been devoted to the actual computation of probabilities and much bigger on the practical consequences of what has been more or less assumed from the beginning: that there is a meaningful way – in most cases an inductive way – in which the probabilities could be constrained to be minuscule. Problems with such an inductive procedure, like various observation-selection effects, biases, the lack of continuity, etc., have been routinely swept under the rug.[2]

In this preliminary study, I wish to set a framework for a reassessment of the role of inductive probabilities in our search for life and intelligence elsewhere – and, indirectly, much of our more general contemporary reflection about our place in the universe. The plan of the exposition is as follows. In Sect. 2, I list some of the instances of what I term sceptical discourse in this respect – sceptical in the sense that from particular small probabilities, estimated or calculated, a

---

[2] For a more detailed analysis of the philosophical claims, see Ćirković (2012), esp. Chapter 7.

negative claim is derived.[3] In Sect. 3, the influence of modern cosmological discourse (especially embedding cosmological background into a wider whole, the multiverse) on our probabilistic reasoning is considered. Zooming into the world of the contemporary ‘astrobiological revolution’ (1995–today) in Sect. 4, I intend to sketch how recent astrobiological findings undermine the sceptical discourse and, even more importantly from the present point of view, how sceptical discourse has been proven to be inert, non-responsive, and essentially decoupled from these revolutionary changes in our perspective on life in its widest cosmic context. As reiterated in the concluding section, the point of the present note is not to resolve these deep and involved astrobiological and philosophical issues, but rather to bring the attention to cognitive deficiencies and double standards in the sceptical discourse. Debunking these constitutes an interesting research programme in its own right.

2 Examples of the Sceptical Discourse

Early clashes between self-styled sceptics regarding extraterrestrial life and intelligence on one side and the classical ‘pluralists’ (about habitable worlds) or ‘optimists’ have been documented in histories of Crowe (1986) and Dick (1996). The modern round started as soon as the SETI era was launched in 1960 with the Ozma Project and the contemporary criticism of George Gaylord Simpson (1964) based on evolutionary biology and a kind of proto-rare-Earth argument. While the original argument of Simpson’s has several weak spots which could be recognized only recently with the advent of modern astrobiology, as I have discussed elsewhere (Ćirković 2014), the debate certainly continues, and it now includes many other elements.

There has been a substantial amount of sceptical discourse, from the time of Simpson to this day, but his criticism in many ways presented a blueprint for all subsequent elaborations. As much as he *reacted* against the nascent SETI projects and voiced his ire over alleged wastefulness of space research, most of his sceptical inheritors continued to emphasize that it is allegedly irrational to engage in research with so low probability of success. Hence arose the scientific – at least formally – criticisms of Tipler (1980), Bond (1982), Barrow and Tipler (1986), Diamond (1992), and Mayr (1993). Over and above it are several philosophical treatments, which usually attempted to add weight to these low-probability estimates by treating them as specifically inductive probabilistic inferences. In this category one can find McMullin (1980, 1989), Rescher (1985), Mash (1993), and Kukla (2001, 2010).[4]

---

[3] Such a negative claim could be existential (‘there are no viable SETI targets’) or pragmatical (‘SETI search programs are not feasible’). While I have studied this dichotomy in some detail elsewhere (e.g. Ćirković 2013), it is not essential from the point of view of the present discussion.

[4] For a different kind of scepticism, see, e.g. Ulvestad (2002) and Basalla (2006).

Most of these belong to the coldest period of the 'SETI winter' of 1980s and 1990s, when the original enthusiasm of the 'founding fathers' faded and scepticism and disappointment contributed to cancellation of the NASA SETI programme, among other things (Garber 1999). Some of these criticisms (for instance, those of McMullin) are arguably motivated by extrascientific concerns; others mention either 'expensive' SETI projects or even explicitly connect them with the fiscal deficit and national debt (Mash, Mayr).

On the other side of the story, we find various probabilistic estimates for abiogenesis vehement in claiming that it is an astronomically improbable event and more or less openly supporting creationism/ID. In an extremely valuable study, Carrier (2004) has compiled data from dozens of sources exhibiting the same tendency of miscalculating probabilities in order to make abiogenesis practically impossible. No less than 7 categories of errors are identified, and it is shown that such estimates are not only numerically unreliable, but more importantly substantially unfounded, since they either mistake a part for the whole or misrepresent the complexity of early living systems or commit any number of other substantial errors. Although similar comparative study has not been performed for noogenesis (or for the probability of the 'Cambrian explosion analogues') so far, there is no doubt that similar errors are common there as well. And while subjects of Carrier's study are mostly creationists and ID-supporters, this does not make much difference in the context of the sceptical position: Jacques Monod's "lucky accident" school of thought, or completely secular sceptics like Simpson can be targeted in the same manner (Fry 2000), although it might be downplayed for tactical reasons.

Why is that interesting again now? In contrast to the "SETI winter", in recent years, there has occurred a modest revival of interest for SETI, together with a new generation of observational searches (e.g. Wright et al. 2014) and new ideas on the theoretical front (e.g. Bradbury et al. 2011). The theoretical side of SETI has been sorely neglected in earlier times, as a consequence of several factors, some of which are of extrascientific nature. The misuse of the probabilistic arguments has certainly been one of these factors: if one believes – on the basis of often spurious arguments – that the probability of target's existence is negligible, the development of sophisticated theoretical models for evolution and detectability of targets sounds like decisively bad idea for investing time and resources.

An important theoretical development has been the advent of large-scale numerical simulations of the astrobiological 'landscape' based on well-established cosmological simulations (e.g. Vukotić et al. 2016; Forgan et al. 2017). The emergence of such studies clearly testifies that astrobiology and SETI have become more mature and sophisticated in comparison to previous decades. One key point stands out about them; however, they are studies of *habitability*, and if we wish to translate their results into specific search proposals or even target selection rules, we need to adopt a set of *probabilities* for particular stages in evolution. Probabilities are understood here in purely epistemic sense, as reflections of our ignorance about the details of the dynamical evolution.

In parallel, we have seen the emergence of the new breed of scepticism, embodied in the 'rare-Earth' hypotheses, starting with publication of the eponymous book at the turn of the century (Ward and Brownlee 2000). They start with the model in which the emergence of complex biosphere – and specifically noogenesis – is contingent upon the conjunction of many logically and physically independent requirements: a conveniently placed giant planet to deflect potential impactors, a large natural satellite to stabilize the spin of the planet, plate tectonics to enable carbon-silicate regulation cycle, etc. Since each of these requirements is prima facie improbable, the probability of their conjunction is simply the product of individual probabilities – which must be a stupendously small number. It might seem, especially to an eager sceptic, that here is a perfectly sound way of producing negligible probabilities – perhaps not that small as the combinatorial ones alleged by creationists and IDers for abiogenesis (reviewed by Carrier 2004), but still small enough to ensure that Earth's biosphere is a unique phenomenon in the Galaxy, if not within the cosmological horizon. This would ensure that our SETI efforts are in vain, and the astrobiological enterprise would be reduced essentially to astro-*micro*biology – studying possible alien prokaryotes and extremophiles. While the rare-Earth hypothesis has provoked enormous discussion and played an important catalysing role in our thinking about habitability, it turns out that there are many problems with it. I shall discuss briefly some of the problems (the ones dealing with probabilities) in Sect. 4; before that, it is important to take another look at the cosmological background against which the universal cosmic evolution and development unfold.

3 Cosmological Degeneracy of Probabilistic Inference

Modern cosmology in a sense intervened to make the matters more complicated by suggesting not only a spatially infinite universe but, in fact, a whole ensemble of universes, known as the *multiverse*, arising as a natural consequence of the most generic form of cosmological inflation (e.g. Linde 1992; Carr 2007). Other best modern physical theories, in particular, the string theory in its current form of M-theory, offer at least similar predictions about the multiverse. Probabilistic reasoning in the multiverse context is an extremely controversial and sensitive issue, since metrics and measures are poorly defined, and the notion of typicality is quite non-trivial to comprehend (e.g. Hartle and Srednicki 2007; Page 2008). If the inflation is indeed eternal, we expect the multiverse to be infinite in both spatial extent and in the number of bubble universes; hence, we need adequate weightings in order to be able to calculate the probability of any particular feature.

Consider, for instance, the controversial matter of how improbable is abiogenesis, reviewed by Carrier (2004), as mentioned above. Many authors, some of naturalist and others of creationist bent have offered estimates of extremely small probabilities, ranging from 10^{-35} to $10^{-2,000,000,000}$. Clearly, a straightforward inference from these minuscule probabilities, *without any background assumptions*, leads many to favour supernaturalistic or at least exotic accounts for the origin of life on Earth.[5] On the other hand, *framing* the problem in an adequate cosmological context enables us to overcome this particular local bias. A counterweight to such framing is the 'problem' of Lucretian-style arguments suggesting that in cosmologies with either infinite spatial sections or with infinite number of cosmological domains (i.e. the multiverse), probabilities do not make real sense, since everything possible under the laws of nature will occur an infinite number of times (see Mash 1993 for a rather cogent philosophical criticism of applications of these arguments to SETI).

An example of such an approach is a study by the Russian-American computational biologist Eugene V. Koonin on the relationship of the cosmological model of eternal inflation to the puzzle of abiogenesis on Earth.[6] He concludes that only the multiverse of eternal inflation guarantees that the highly improbable steps related to life's origin will inevitably occur. Therefore, problematic issues like 'irreducible complexity' or unproven ribozyme-catalysed RNA replication could be completely sidestepped – *somewhere* in the multiverse abiogenesis could proceed by chance, and we need just to apply anthropic (self-)selection to conclude that one of these places is called Earth. Irrespectively of specifics of Koonin's scenario, one thing is radically new here; as he writes: 'The plausibility of different models for the origin of life on earth directly depends on the adopted cosmological scenario'. From our point of view, the overarching infinity of habitable worlds suggested by inflationary cosmology cancels any degree of improbability of abiogenesis; by analogy, this will be the case with any other allegedly improbable 'critical step' in the evolutionary development leading to intelligent observers and technological civilizations. (An argument in the opposite direction has been made by Monton 2004, although it relies on some of the rare-Earth discourse and commits the same mistake of conflating logical and physical independence.)

Therefore, cosmology has a double role to play in the assessment of probabilistic inference about the origin of life and intelligence: it both complicates (the lack of consensus about typicality and the default measure of probability) and simplifies (Koonin's argument and the analogous Lucretian arguments which could be made about noogenesis, technological civilization, etc.) the discussion. As far as our epistemic state on abiogenesis and noogenesis is concerned, the impact of cosmology tends to shift it towards agnosticism: we are sampling many Hubble volumes to be certain that at least one biosphere is produced with certainty, in accordance with our current empirical knowledge – which does not tell us whether there are indeed law-like process of biosphere emergence.[7] One might speak about the *degeneracy* (in the technical sense) of probabilistic reasoning when embedded within the cosmological

---

[5] Such as the directed panspermia of Crick and Orgel (1973).

[6] Koonin (2007).

[7] Note that Koonin's argument does not say anything about the likelihood of success of our astrobiological endeavour: if anything, it might make us more *pessimistic* as to whether we are likely to find an independent abiogenesis within our cosmological horizon. However, it strongly refutes the idea that non-naturalistic factors are *necessary* for abiogenesis.

context: generalization of local probability calculations could lead to multiple, mutually exclusive conclusions.

This is highly instructive and somewhat unexpected point: what we should conclude about the chances of success of our practical SETI *observations* (for example), depends on the arcane cosmological assumptions about the multiverse. While the fact that odds of and in themselves should not determine conclusions or dictate action has been known since the inception of probability theory, the *context* of astrobiology/SETI studies is sufficiently strange that the sceptics have so far all too often managed to promote their view based on unfavourable odds only. The point needs to be taken into account when formulating a more general research programme outlined in the concluding section.

4 Extrasolar Planets, Galactic Habitable Zone, and Inductive Probability

The discovery of a large number of extrasolar planets since 1995 has dramatically changed our views on life in the universe and gave rise to the astrobiological revolution of the turn of the century. In spite of several false alarms which preceded the discoveries of Queloz, Mayor, Marcy, and Butler, it was still possible – up to 1995 – to argue or imply that the Solar System is either unique or a very rare occurrence in the Galaxy (or, indeed, the visible universe). This (proto)sceptical hypothesis has been clearly refuted – and in the way which Simpson and some of his followers wrongly predicted to lie in the distant future.

That was just the beginning of the new Copernican story. For a brief period early in the studies of extrasolar planetary systems, it was thought that most of them contain 'hot Jupiters', i.e., giant gaseous planets in very close orbits around their parent stars. Such planetary configuration has thought to be hostile to life, since the inward migration of gas giants would have destroyed stability of orbits of hypothetical terrestrial planets in the circumstellar habitable zone (e.g. Dawson and Johnson 2018). Many sceptics – including the 'rare-Earth' authors like Ward and Brownlee – have reasoned during that period that such empirical results corroborate their views.

Today, we know that 'hot Jupiters' are *exception, rather than the rule*. They have been first discovered in large relative proportion due to simple observation-selection effects, which have now been studied and understood in sufficient detail (Johnson et al. 2010). That minor episode in the recent history of planetary science – that anti-Copernicanism of the rare-Earth school tried to co-opt 'hot Jupiters' for the purpose of decreasing the probability of finding habitable Earth-like planets elsewhere in the Galaxy – should give us pause, however. What if *all* rare-Earth prerequisites for other complex biospheres are in fact similar cases: the progress of science will eventually show them to be generic cases, rather than exceptions? Are there any inherently small probabilities relevant for the evolution of complex biospheres at all?

Part of the problem with rare-Earth perspective originates in its misuse of counterfactual reasoning: Ward and Brownlee and their supporters tend to compare Earth within the real Solar System with the counterfactual Earth without (for instance) Jupiter, but with everything else in the Solar System staying the same – as if someone had removed Jupiter by magic! This is clearly wrong counterfactual thinking to use; instead, one should think about the evolutionary developmental process of the Solar System formation and subsequent evolution which could result in a configuration without Jupiter – but it is unclear whether such process could lead to existence of Earth in the first place. This problem with the implicit – and violated – ceteris paribus clause in the rare-Earth thinking has been noted in Chyba and Hand (2005) and Ćirković (2012, esp. Chapter 6). Here, I wish to add an additional and related problem which directly impact probabilistic inference: namely that the rare-Earth proponents fail to distinguish between logical independence and physical independence of their various proposed requirements. Not only the probability of some requirements cannot be evaluated outside of their proper context which is evolutionary, developmental, and historic, but it is highly doubtful whether their conjoint probability is simply the product of component probabilities. The probability of Earth's retaining stable rotation axis is equal to the probability of Moon-creating impact plus the probabilities of all the other ways for Earth to have stable axis (for instance, by continuing to spin much faster, e.g., in 8 hours or so such as it was in the time of our planet's accretion) – but both these alternatives are physically connected to the existence of Jupiter and other effects on the larger scales *and* to plate tectonics and other effects on the smaller scales. Thus, while these requirements are logically independent (in the sense that there are possible worlds in which the interaction is so small to be negligible), but are not physically independent in the specific historic case of our Earth. Therefore, we need a different and more complex method of calculating the compound probability.

Hence, the outcome of the debate on the validity of various 'rare-Earth' arguments is still very much open – that very fact should give pause to sceptics, since their case has initially seemed unassailable, almost self-evident. This is confirmed by the work done on the Galactic Habitable Zone, in both spatial and temporal domain, since the pioneering work of Lineweaver (2001) and Lineweaver et al. (2004). In recent years, we have witnessed an explosion of interest in the topic, on both theoretical (Behroozi and Peeples 2015; Zackrisson et al. 2016) and numerical (Vukotić et al. 2016; Forgan et al. 2017) level. These studies confirm that habitable planets in the Milky Way comprise a large statistical set on which important and testable analyses could be made. In other words, astrobiological studies of habitability have outgrown the early, childhood phase of philosophical preference and moved in the direction of quantitative solidity characterizing mature scientific fields.

Finally, there is a general argument against using probabilistic models based on conjunction of a priori improbable occurrences, which has been known in a vague form since Pascal and has been most beautifully described by Stanislaw Lem in one of his mock essays, a very real review of two fictional books, *De Impossibilitate Vitae* and *De Impossibilitate Prognoscendi*, by a fictional author Cezar (or perhaps

Benedykt) Kouska.[8] The fictional author uses his own ancestry to “prove” that the theory of probability is unsound:

> A certain army doctor, during the First World War, ejected a nurse from the operating room, for he was in the midst of surgery when she entered by mistake. Had the nurse been better acquainted with the hospital, she would not have mistaken the door to the operating room for the door to the first-aid station, and had she not entered the operating room, the surgeon would not have ejected her; had he not ejected her, his superior, the regiment doctor, would not have brought to his attention his unseemly behavior regarding the lady (for she was a volunteer nurse, a society miss), and had the superior not brought this to his attention, the young surgeon would not have considered it his duty to go and apologize to the nurse, would not have taken her to the café, fallen in love with her, and married her, whereby Professor Benedykt Kouska would not have come into the world as the child of this same married couple.

After elaborating in some humorous detail (an understatement!) on other such apparent coincidences leading to the author’s birth, and their generalization from human genealogy to the very evolution of our species, Lem poses the key ironic question:

> Each man is, as it were, the first prize in a lottery, in the kind of lottery, moreover, where the winning ticket is a teragigamegamulticentillion-to-one shot. Why, then, do we not daily feel the astronomically monstrous minuteness of the chance of our own or another’s coming into the world?

Not only it is impossible in retrospect to prove that a particular evolutionary outcome was particularly probable or improbable *in isolation*, but we need to understand whether events are physically dependent or not.

Clearly, in the case of one’s ancestry, the problem with probabilistic reasoning is that the chains in the link are obviously – to the point of satire – not only not independent, but clearly linked in an inherent and genetic (in the philosophical sense) manner. How is it exactly different, however, from the reasoning beyond the rare-Earth hypotheses or some of the usages of the Drake equation, except that the context is less understood and hence the probabilistic ‘shortcut’ looks more plausible? Consider for instance items in the Drake equation – thing such as Sun-like stars, habitable planets, the origin of life, the origin of technological civilization, etc. are supposed to have only one direction of causal dependency: the origin of life depends on the existence of habitable planets, but not on the origin of technological civilization, etc. However, this simplistic view is, on deeper insight, simply *wrong*: persistence of habitable planets can depend on their being actually inhabited (e.g. through the ‘Daisyworld-like’ feedbacks), advanced technological civilizations are capable of increasing the number of habitable planets through terraforming, etc. It is in fact an excellent confirmation of the fruitfulness of the evo-devo approach to astrobiological complexity: nonlinearity and feedbacks accompanying the developmental side of the story obviate the simplistic conjunction

---

[8] Lem [1971] (1999), pp. 141–166. The same brilliant anthology contains the famous essay “The New Cosmogony“ which proposes a novel and radical solution to Fermi’s Paradox (cf. Ćirković 2018).

of probabilities which characterizes the initial state of complete ignorance. In this sense, the most fruitful approach is exemplified by the study of Scharf and Cronin (2016), which sets a quantitative framework for studying probability of abiogenesis as a function of a number of parameters outlining an 'island' volume in the overall parameter space.

5 Discussion: An Outline of the Research Programme?

The purpose of the present work is certainly not to argue that the probabilities of abiogenesis and noogenesis are high; it would be quite naive to draw such a conclusion, abstracting away Lucretian arguments, such as Koonin's. The question of absolute probabilities of such evolutionary steps is exceedingly complex and cannot be posed separately from the development of the overall *astrobiological theory*, which would enable such calculations and predictions from something at least much closer to the 'first principles'. The prospects of such a theory got immensely brighter in recent years, with works cited above such as Lineweaver et al., Behroozi and Peeples, Scharf and Cronin, Zackrisson et al., and others at the frontline. Such a future theory is likely to be Copernican, but not as a matter of general principle or abstract assumption; instead, Copernicanism will be an output – or a prediction – of the theory.

Instances of confirmation of Copernicanism include, for instance, empirical evidence that the Solar system is not an extraordinary rare occurrence (as, for instance, per old catastrophic cosmogonic hypotheses) – instead, planetary systems are definitely very frequent occurrences in the Galaxy. Sun is a fairly typical star, and the Milky Way is quite similar to millions of other large spiral galaxies. If there are claims to the contrary, the burden of proof is obviously on those uttering such claims. As to the spatial location and sets of objects concerned (Sun-like stars, terrestrial planets, etc.) there has been no reason so far to challenge Copernicanism on the empirical basis. It is true that some of the anti-Copernican claims of, for example, 'rare-Earth' theorists cannot be empirically falsified as yet, due to technical insufficiencies – e.g. the atypicality of Moon-like giant satellites of habitable planets – but that only means that the issue is still open and should not cause any Bayesian probability shift so far. Future generations of instruments will be able to test those aspects of Copernicanism which evade our empirical falsification (or corroboration) as yet; until then, however, we have no compelling reason to assume the sceptical position as default.[9]

---

[9] In parallel with Carrier's (2004) Class I error (citing obsolete sources), one should note that SETI sceptics like Kukla (2010) continue to cite Simpson, for example, in spite of his glaring failure to predict the relevant technological developments, since he argued that observations of extrasolar planetary systems are 'far beyond any reasonable extrapolation' of our astronomical capabilities. There are many such instances in the sceptical discourse, warranting further historical analysis.

Quite to the contrary, we have seen some reasons above for formulating a research programme investigating the consistency, motivations, and ramifications of the sceptical discourse. As we have seen, the early sceptical estimates are hardly better than open creationism; while 50 years ago they might have been justified, after 1995 and the astrobiological revolution, they should have revised. The revisions should naturally proceed in the Bayesian manner: new observational data and theoretical elaborations ought to cause a probability shift among different hypotheses about the existence of life and intelligence beyond Earth. Therefore, the question which needs to be answered within the proposed research programme is: *Has there been such a shift?*

Since prediction of outcomes is inseparable part of any research programme, I hereby express the following prediction, stemming from this cursory study: the cursory overview of the sceptical position indicate that there has not been such a Bayesian probability shift. In turn, this lack of Bayesian shift towards greater optimism in regard to extraterrestrial life and intelligence in sceptical circles support the hypothesis that scepticism has not been based on probabilities in the first place. Instead, the alleged low probabilities have been used at best as *post festum* justifications for pessimistic conclusions made in advance of any probabilistic analysis; at worst, they have been – and in many circles still are – fig leaf covers for an extrascientific and supernaturalist ideological agenda. The fact that some believers in extraterrestrial life and intelligence have been overly enthusiastic about probabilities even after the failure of the original SETI optimism of the 'founding fathers' (e.g. Aczel 1998) does not change the essential intellectual bankruptcy of the sceptical probabilistic inference. While not all SETI sceptics have been openly motivated by such an agenda, offering the indirect support or even finding such views legitimate topics for science and philosophy has been damaging enough. It is a high time for such abuse of probabilistic inference to end.

Acknowledgements The author wishes to thank Branislav Vukotić, Mark Walker, Jelena Dimitrijević, Anders Sandberg, George Dvorsky, John Smart, Slobodan Perović, Karl Schroeder, Aleksandar Obradović, Milan Stojanović, Goran Milovanović, Damian Veal, and Srdja Janković who offered helpful suggestions. Dušan Pavlović has kindly helped in obtaining some of the literature. The author acknowledges financial support from the Ministry of Education, Science and Technological Development of the Republic of Serbia through the projects #ON176021 and #ON179048.

References

Aczel, A. D. 1998, *Probability 1: Why There Must Be Intelligent Life in the Universe* (Harcourt Brace & Company, New York).

Barrow, J. D. and Tipler, F. J. 1986, *The Anthropic Cosmological Principle* (Oxford University Press, New York).

Basalla, G. 2006, *Intelligent Life in the Universe* (Oxford University Press, Oxford).

Behroozi, P. and Peeples, M. S. 2015, "On the history and future of cosmic planet formation," *Monthly Notices of the Royal Astronomical Society* **454**, 1811–1817.

Bond, A. 1982, "On the Improbability of Intelligent Extraterrestrials," *Journal of the British Interplanetary Society* **35**, 195–207.

Bostrom, N. 2002, *Anthropic Bias: Observation Selection Effects in Science and Philosophy* (Routledge, New York).

Bradbury, R. J., Ćirković, M. M., and Dvorsky, G. 2011, "Considerations for a Dysonian Approach to SETI: A Fruitful Middle Ground?" *Journal of the British Interplanetary Society* **64**, 156–165.

Carr, B. (ed.) 2007, *Universe or Multiverse?* (Cambridge University Press, Cambridge).

Carrier, R. C. 2004, "The argument from biogenesis: Probabilities against a natural origin of life," *Biology and Philosophy* **19**, 739–764.

Chernavskii, D. S. 2000, "The origin of life and thinking from the viewpoint of modern physics," *Physics-Uspekhi* **43**, 151–176.

Chyba, C. F. and Hand, K. 2005, "Astrobiology: The Study of the Living Universe," *Annu. Rev. Astron. Astrophys.* **43**, 31–74.

Ćirković, M. M. 2012, *The Astrobiological Landscape: Philosophical Foundations of the Study of Cosmic Life* (Cambridge University Press, Cambridge).

Ćirković, M. M. 2013, "Who are the SETI skeptics?" *Acta Astronautica* **89**, 38–45.

Ćirković, M.M. 2014, "Evolutionary contingency and SETI revisited," *Biology and Philosophy* **29**, 539–557.

Ćirković, M. M. 2018, *The Great Silence: Science and Philosophy of Fermi's Paradox* (Oxford University Press, Oxford).

Crick, F. H. C. and Orgel, L. E. 1973, "Directed panspermia," *Icarus* **19**, 341–346.

Crowe, M. J. 1986, *The Extraterrestrial Life Debate 1750-1900* (Cambridge University Press, Cambridge).

Dawson, R. I. and Johnson, J. A. 2018, "Origins of Hot Jupiters," *Annu. Rev. Astron. Astrophys.* **56**, 175–221 (preprint https://arxiv.org/abs/1801.06117).

Diamond, J. 1992, *The Third Chimpanzee* (HarperCollins, New York).

Dick, S. J. 1996, *The Biological Universe: The Twentieth-Century Extraterrestrial Life Debate and the Limits of Science* (Cambridge University Press, Cambridge).

Dodd, M. S., Papineau, D., Grenne, T., Slack, J. F., Rittner, M., Pirajno, F., O'Neil, J., and Little, C. T. S. 2017, "Evidence for early life in Earth's oldest hydrothermal vent precipitates," *Nature* **543**, 60–64.

Forgan, D., Dayal, P., Cockell, C., and Libeskind, N. 2017, "Evaluating galactic habitability using high-resolution cosmological simulations of galaxy formation," *International Journal of Astrobiology* **16**, 60–73.

Fry, I. 2000, *The Emergence of Life on Earth: A Historical and Scientific Overview* (Rutgers University Press, New Brunswick).

Garber, S. J. 1999, "Searching for good science: The cancellation of NASA's SETI program," *J. Brit. Interplanet. Soc.* **52**, 3–12.

Hartle, J. B. and Srednicki, M. 2007, "Are we typical?" *Physical Review D* **75**, 123523(6).

Johnson, J. A., Aller, K. M., Howard, A. W., and Crepp, J. R. 2010, "Giant Planet Occurrence in the Stellar Mass-Metallicity Plane," *Publ. Astron. Soc. Pac.* **122**, 905–915.

Koonin, E. V. 2007, "The cosmological model of eternal inflation and the transition from chance to biological evolution in the history of life" *Biology direct* **2**(1), p.15.

Kukla, A. 2001, "SETI: On the Prospects and Pursuitworthiness of the Search for Extraterrestrial Intelligence," *Stud. Hist. Phil. Sci.* **32**, 31–67.

Kukla, A. 2010, *Extraterrestrials: A Philosophical Perspective* (Lexington Books, Lanham, Maryland).

Lem, S. [1971] 1999, *A Perfect Vacuum* (trans. by M. Kandel, Northwestern University Press, Evanston).

Linde, A. D. 1992, "Stochastic approach to tunneling and baby universe formation," *Nuclear Physics B* **372**, 421–442.

Lineweaver, C. H. 2001, "An Estimate of the Age Distribution of Terrestrial Planets in the Universe: Quantifying Metallicity as a Selection Effect," *Icarus* **151**, 307–313.

Lineweaver, C. H., Fenner, Y., and Gibson, B. K. 2004, “The Galactic Habitable Zone and the Age Distribution of Complex Life in the Milky Way,” *Science* **303**, 59–62.

Luisi, P. L. 2006, *The Emergence of Life: From Chemical Origins to Synthetic Biology* (Cambridge University Press, Cambridge).

Mash, R. 1993, “Big numbers and induction in the case for extraterrestrial intelligence,” *Philosophy of Science* **60**, 204–222.

Mayr, E. 1993, “The search for intelligence,” *Science* **259**, 1522–1523.

McMullin, E. 1980, “Persons in the Universe,” *Zygon* **15**, 69–89.

McMullin, E. 1989, “Having Fun with ET,” *Biology and Philosophy* **4**, 97–105.

Monton, B. 2004, “Life is Evidence for an Infinite Universe,” preprint at http://philsci-archive.pitt.edu/id/eprint/3507 (last accessed 2017-12-22).

Page, D. N. 2008, “Typicality derived,” *Physical Review D* **78**, 023514(7).

Rescher, N. 1985, “Extraterrestrial Science,” in E. Regis, Jr. (ed.) *Extraterrestrials, Science and Alien Intelligence* (Cambridge University Press, Cambridge), pp. 83–116.

Scharf, C. and Cronin, L. 2016, “Quantifying the origins of life on a planetary scale,” *Proceedings of the National Academy of Sciences* **113**, 8127–8132.

Simpson, G. G. 1964, “The nonprevalence of humanoids,” *Science* **143**, 769–775.

Tipler, F. J. 1980, “Extraterrestrial Intelligent Beings do not Exist,” *Q. Jl. R. Astr. Soc.* **21**, 267–281.

Ulvestad, E. 2002, “Biosemiotic knowledge – a prerequisite for valid explorations of extraterrestrial intelligent life,” *Sign Systems Studies* **30**, 283–292.

Vukotić, B., Steinhauser, D., Martinez-Aviles, G., Ćirković, M. M., Micic, M., and Schindler, S. 2016, “‘Grandeur in this view of life’: N-body simulation models of the Galactic habitable zone,” *Monthly Notices of the Royal Astronomical Society* **459**, 3512–3524.

Ward, P. D. and Brownlee, D. 2000, *Rare Earth: Why Complex Life Is Uncommon in the Universe* (Springer, New York).

Wright, J. T., Mullan, B., Sigurdsson, S., and Povich, M. S. 2014, “The $\hat{G}$ infrared search for extraterrestrial civilizations with large energy supplies. I. Background and justification,” *Astrophysical Journal* **792**, id. 26 (16 pp).

Zackrisson, E., Calissendorff, P., González, J., Benson, A., Johansen, A., and Janson, M. 2016, “Terrestrial Planets across Space and Time,” *Astrophysical Journal* **833**, article id. 214, (12 pp).

Part II
Multiscale Complexity

The Equivalence of Free Energy and Information: Thermodynamic Descriptions as a Condition of Possibility of Objectivity

Joseph M. Brisendine

Paper

"With great and varied skills we create a delusion that enables us to coexist serenely with the most monstrous things, simply because we recognize these frozen grimaces of the universe as a table or a chair, a shout or an outstretched arm, a speed or a roast chicken. We are capable of living between one open chasm of sky above our heads and another, slightly camouflaged chasm of sky beneath our feet, feeling as untroubled on earth as if we were in a room with the door closed. We know that our life is ebbing away both outward into the inhuman distances of cosmic space and downward into the inhuman microspace of the atom, while we go on dealing with a middle stratum, the things that make up our world, without troubling ourselves at all over the fact that this proves only a preference for impressions received in the middle distances . . . looked at closely, it does seem to be an extremely artificial state of mind that enables a man to walk upright among the circling constellations and permits him, surrounded as he is by an almost infinite unknown, to slip his hand with aplomb between the second and third buttons of his jacket."

—Musil, *The Man Without Qualities* (Translated by Sophie Wilkins)

J. M. Brisendine (✉)
Departments of Chemistry and Physics, City College of New York, New York, NY, USA

Department of Biochemistry, The Rockefeller University, New York, NY, USA
e-mail: jbrisendine@ccny.cuny.edu

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_6

1 Introduction: The Stochastic Revolution and Its Consequences

Over the past several decades, there has been a revolution in physics which has gone largely unnoticed and unreported not merely by the public but by a large section of the physics community itself. We are referring to the development of the field that has come to be known as “stochastic thermodynamics.” Since the 1990s, beginning with the discovery of the first fluctuation theorems (Crooks 1999; Jarzynski 1997), there has been an explosion of new mathematical techniques for forming thermodynamic descriptions and applications of the methods of thermodynamics to new domains of nature. Nearly 30 years on, there has been a succession of articles attempting to unify these new discoveries and present them in a manner where their significance to other areas of physics can be understood and applied more broadly (Smith 2011; Parrondo et al. 2015). In this work, we aim to address the significance of this revolution for our philosophical understanding of emergent phenomena in nature.

For those with sufficient familiarity with both the relevant physics and philosophy, our thesis is easily stated: the applicability of stochastic methods to the wide range of non-equilibrium systems which have been analyzed in this manner since the 1990s confirms the scale invariance of thermodynamic analysis and thus the universality of entropy as the leading term in the log probability of fluctuations away from ensemble averages for any system which satisfies the scaling requirements of thermodynamic descriptions. Thermodynamic descriptions are possible in any context where statistical degeneracy is the primary determinant of system structure, and this occurs at any scale where degrees of freedom are subject to concentration of measure. The vast scale differences separating the microscopic and macroscopic regimes of nature allow this phenomenon to occur many times over between the physics of the standard model and the physics of everyday life, and the hierarchical nature of natural evolution is therefore necessitated by the scaling requirements of thermodynamics. This is because the equilibrium or stationary solutions of thermodynamic descriptions at one scale must serve as stable bits on the computational timescales of any higher-order complexity processes that may evolve from interactions at that scale. Stability at each successive level of a hierarchical process is a logical requirement for evolution to be capable of searching for new forms of physical order; translating those stability requirements into thermodynamic language reveals the scaling requirements imposed on nature as a condition of the possibility of hierarchical evolution. These scaling requirements allow us to factor apart the thermodynamic features of evolution, which are independent of substrate and scale, from the historical and contingent aspects of evolution which are context-dependent and potentially unique to a given scale.

Bringing together the new understanding of what thermodynamics is and why it works that resulted from the revolution of the 1990s with the understanding of the role of thermodynamics in biological evolution developed by biophysicists throughout the twentieth century creates a unified vision of the evolution of natural complexity. This unified vision requires us to further change our understanding of the relationship between emergence and reductionism, and perhaps the relationship between science and nature more broadly. Indeed, perhaps the most remarkable philosophical aspect of the stochastic revolution is the degree of conceptual unity it has created for all physics, and hence all of nature.

The "physicists" who developed these new ideas, however, were not mainly "fundamental" physicists, and this has been much of the reason why the significance of this unification has gone largely unnoticed. While fundamental physicists have been relentlessly pursuing quantum gravity (and arguably the most exciting developments in that area have come from black hole *thermodynamics*), the physicists who pioneered the use of the fluctuation theorems have come from diverse fields in the *middle* scales of nature: condensed matter, fluid mechanics, nanotech, biochemistry etc., all fields that might seem far-out to a layperson but are decidedly "emergent" to a physicist. While the techniques quickly sprang up the ladder in complexity, making their way into economics in less than a decade (Madan 2010), only in 2018 was it announced that the stochastic formalism had been extended to driven quantum field theories, and hence finally to our most "fundamental" description of nature (Bartolotta and Deffner 2018).

Part of the cautiousness fundamental physicists have in embracing these ideas is the understanding that thermodynamics is an effective theory that applies to "coarse-grained" descriptions of nature, meaning that the objects which enter into thermodynamic calculations may themselves be composite systems, and thus when we do thermodynamics, it is guaranteed that we are not interrogating nature at a fundamental or "ontic" level. Indeed, philosophical concerns over the "objectivity" of entropy plagued the advancement of thermodynamics for much of the twentieth century. A coarse-graining scheme requires human choices, and the calculated value of information or entropy we produce can change depending on these choices. Is this not the definition of a "subjective" measure?

Before answering this question, it is worth stepping back and asking what we expect from an answer, and more broadly how we would respond if asked whether a theory of "multiscale adaptive complexity" should be "objective"? Our instinctive reaction to the latter question may be that, of course, to be scientific, a theory must be objective. But can we define the objectivity of a theory which aims to explain how and why there are objects in the universe in the first place? If such a theory does what it purports to do and explains, in reductive terms, the logic of emergence of objects, would it not be circular for it to rely on any object to explain itself? We assert that any satisfactory theory which explains the emergence and growth of complexity in the universe will be neither subjective nor objective. If this is not possible, then it must be because the categories subjective and objective are sufficient, exhaustive, and mutually exclusive binary options into which all theories and statements fall. We are aware, however, of no proof of anything of the sort, and do not believe that there is one, nor any reason at all to believe that subjective and objective are mutually exclusive and exhaustive categories for all possible claims.

The theory we seek is "pre-objective": it is a theory of the conditions of possibility of objects in the first place. This does not make it subjective any more than it places it outside the bounds of ordinary scientific verification through empirical prediction. It does mean, however, that the content of the theory is not comprised of a set of "facts about the world." The content of the theory is instead a set of mathematical concepts and laws that describe any possible representation of the world. This is the same logical status enjoyed by the laws of thermodynamics. The laws of thermodynamics are true in a sense that is beyond mere objectivity, because they obtain in any possible objective situation, so that if we were to perform an "objective" measurement that contradicted the laws of thermodynamics, our initial response would be to assume that the measurement had been performed incorrectly. Even if the result were found to be repeatable, we would next proceed to looking for flaws in the reasoning that led us to conclude the result was in violation of the laws, proceeding through the assumptions of the reasoning behind the measurement under the assumption that they must contain an error. We would consider amending the laws themselves only after accumulating an extraordinary amount of evidence, but it is certain that this will never occur. This is because we also understand the laws of thermodynamics to a greater degree than we understand either general relativity or quantum mechanics, neither of which are understood with complete clarity. In contrast, the mathematical reasons for the laws of thermodynamics are understood completely, and this allows us to understand the precise limits of thermodynamic reasoning, and consequently, we know that as long as we respect these limitations, the theory must always work.

Both general relativity and quantum mechanics obtain in different scale regimes of nature, and the precise limits of these theories are not currently understood. General relativity may be the equilibrium limit of a more general non-equilibrium theory of spacetime (Jacobson 1995), but no one knows what that more general theory looks like or can say exactly when the equilibrium approximation should fail; meanwhile, the mystery of the correct interpretation of quantum mechanics remains as elusive as ever (Leggett 2005). These facts imply that we would be wise to remain agnostic about the final form and limits of modern physics even in the face of its stunning success in providing us with an understanding of nature that has a depth and breadth that was inconceivable to human beings of the past.

Of the theories that constitute modern physics, only statistical mechanics/thermodynamics is *completely* understood, and this has been the case only since the advances in non-equilibrium and stochastic processes that took place in the 1990s. Even now that our understanding of classical thermodynamics can be said, in some sense, to be complete, the consequences of this completion will continue to be unraveled for many decades to come. In addition to being the only completely understood pillar of modern physics, statistical mechanics is also the only theory in modern physics that is not attached to a specific scale.

Our intention in this work is to explain our current understanding of the meaning of thermodynamic descriptions now that this understanding has attained, in some sense, its final form. This is not to say that it is inconceivable that the laws could be further simplified or presented in a more unified context, but merely that our understanding at present is sufficient to say, in any possible situation, exactly when and if thermodynamic descriptions are meaningful and accurate for that situation, and we understand a theory completely if and only if we understand its limits.

The specific statistical phenomenon that makes thermodynamics accurate is termed "concentration of measure." (Ledoux 2001) It is the consequence of the statistical simplifications permitted by the law of large numbers and the central limit theorem, which are the two pieces of mathematics at the heart of our understanding of the second law of thermodynamics. Concentration of measure occurs whenever there is a separation of scale between microscopic dynamics and their macroscopic consequences in the form of phase transitions. With this knowledge, we can use the microscopic frequencies of transitions in a system to bound the scale at which those dynamics should match the thermodynamic predictions for that system. It can be understood as an increase in explanatory power at the cost of explanatory precision: by giving up on the fantasy of describing any physical system at all scales, one gains the power to accurately predict the behavior of any system above the effective thermodynamic limit implied by the microscopic dynamics of that system. Thus, for a given microscopic fluctuation or transition frequency, there exists a natural scale for that system at which and above which thermodynamic simplifications are perfectly accurate.

The conceptual maneuver at the heart of this work is to elevate these scaling criteria of thermodynamic explanations to the level of a condition of possibility for "objectivity." Simply put, beneath the natural scale of thermodynamic description for a system, no stable "objects" emerge from such a system. We identify the scaling limits of thermodynamic descriptions as the same limits of emergence for that system; concentration of measure sets in at the level of microscopic dynamics at exactly the scale at which the effectively random fluctuations at the microscopic scale give way to stable, long-range macroscopic order. The macroscopic phases of a system correspond then to the values of the state variables of the system at which different forms of long-range order dominate the microscopic dynamics.

This perspective which identifies the thermodynamic limits of systems with the effective scale of emergence of objects in that system is explained by the duality of information theory and thermodynamics. By demonstrating that free energy is a form of information, we simultaneously explain why any thermodynamic description should have a corresponding informatic description in terms of the bit content of that system.

The text will proceed by first explaining the connection between free energy and information and then explicating the philosophical consequences of this connection in providing objective descriptions of these situations. We will see that asking about the properties of a physical system beneath the natural scale of description for that system results in "melting" of the stable properties of that system, and thus to a "melting" of the objects in the description as well.

2 The Scope of Thermodynamic Descriptions

Before proceeding to an explanation of how the machinery of thermodynamics works, it is important for the reader to appreciate the scope of modern thermodynamics to motivate the point. The concepts of "free energy," "information," "entropy," and even "temperature" and "pressure" are by no means limited to the context of physical chemistry, but for the unfamiliar perhaps the simplest place to begin explaining the scope of thermodynamics is to point out that the discipline known as "physical chemistry" is essentially nothing but thermodynamics paired with either classical or quantum mechanics. This means that we can describe, for starters, all possible chemical reactions with thermodynamics. As my first physical chemistry professor remarked to me "the free energy of formation of a substance is the only chemical reaction you need to understand to understand all chemical reactions." Describing "all possible chemistry" by the way, necessarily includes as a subset all possible biochemistries. And we are also quite confident that modern thermodynamics explains the structure and function of biomolecules exquisitely well.

Despite the "anti-reductivist" stance that the biological realm should resist a purely mechanistic description because "function is not a reductive concept," physical biochemistry has thus far been quite successful at elucidating the structure-function relationships behind many, highly sophisticated and evolved enzymatic mechanisms along with many diseases. This is because, however embedded within historical constraints and complex relationships biomolecules may be, we are confident that thermodynamics can describe *the equilibrium* behavior of matter *at any scale*, because we are equally confident that the concept of "equilibrium," which is entirely statistical and mathematically transparent in the language of statistical mechanics, is universal and invariant with respect to changes of scale. A set of processes occurring at finite rates are either detailed-balanced or not, and there is never any ambiguity about this issue at any scale.

The second reason we are confident in the applicability of thermodynamics to the enormous range of scales it describes is that the second law guarantees that all physical systems will approach equilibrium at some timescale. It may be the case that the phenomena we are interested in are entirely "non-equilibrium" and therefore the equilibrium description of the situation does not tell us what we want to know, but it is never the case that the equilibrium description can't be given. Furthermore, the scale-independence of thermodynamics is manifest marvelously by the intricately evolved structure of biomolecules and cells, because cells represent instances of physical structure which themselves have many different internal degrees of freedom, all coming in and out of equilibrium at different timescales. Indeed, the literature is full of "thermodynamic descriptions" of different biological components with vastly different equilibration times.

The fastest processes in cells are in the quantum-biological realm. Organic molecules which absorb light such as the antennae complexes of chlorophylls which are found in photosynthetic membranes have extraordinarily long-lived quantum coherent states, which exist for nanoseconds, or 10^{-9} seconds, despite the warm and noisy cellular environment (Romero et al. 2014). For comparison, a typical light-absorbing organic molecule will have a coherent excited state-lifetime under cellular conditions of a few femtoseconds, or 10^{-15} seconds. Already, we see that we are discussing processes which are six orders of magnitude apart in time, yet both of which seem vanishingly brief from our own sense of time. Such processes are the "high frequency" degrees of freedom in cells, and the timescales involved in electron and proton tunneling, which are driven by the absorption of light in the case of photosynthesis, are the next fastest, occurring on timescales of micro-milliseconds (Gray and Winkler 2003; Marcus and Sutin 1985). The cellular structure is essentially continuous in degrees of freedom with equilibration times which vary from the femtosecond regime right up to the timescales of our awareness and up, ultimately, to the longest timescales in biology. Consideration of geochemistry and environmental processes increases the scope of long timescales to processes which occur over hundreds of millions of years. At any time or space scale we choose to look, the behavior of equilibrium matter is understood to be governed by the same principles of thermodynamics.

If describing the ultimate behavior and structure of matter at all scales wasn't enough, the extensions brought about by "stochastic mechanics" now include non-equilibrium instances of structure as well, although not at any arbitrary timescale we might prefer the description to work. Some of the earliest applications of the stochastic formalism were the corresponding non-equilibrium thermodynamic descriptions of nucleic acids and proteins operating in the "driven" states characteristic of biology (Li et al. 2010). And again, whatever "scale-specific" content there is in the system, it has thus far been no impediment to the unraveling of many biological functions, despite the fact that no term entering into the calculations of entropy or free energy depends on this "function." The function is understood to be an emergent property of the structure embedded within the cellular context in which it evolved. Despite the fact that the function was selected by the details of the environment, however, the molecule or super-molecular structure that performs a biological function must be capable of performing its duty without any knowledge of those environmental details that shaped its selection (Anfinsen 1973), which means that the equilibrium state of the structure must also contain sufficient information for us to reconstruct its function, given that we also have knowledge of the relevant details of the environment. While there are many biological structures and codes which have yet to be unraveled, nothing about anything we have thus far understood about biology leads us to believe that anything we have yet to discover will require us to revise our understanding of physical chemistry, which is to say thermodynamics.

Beyond even the scope of physical chemistry, however, thermodynamic descriptions have been formed for a wide variety of systems under the banner of "stochastic processes" and diverse other names like "reaction-diffusion models" (Baez 2018; Pastor-Satorras and Sole 2001), for processes ranging from the spread of a disease to the long-time solutions of predator-prey populations in ecology to the non-equilibrium fluctuations of stock markets. A “free energy principle” has even been formulated for neural networks, and the mathematics of stochastic processes is deeply entwined with all known forms of machine learning (Friston et al. 2016). If part of the appeal of “anti-reductionism” rests on the premise that “traditional” physics flounders in the face of the complexity of the real world, the past 20 years of physics has provided ample evidence that what can be understood in a scale-independent and reductive framework has not yet been fully appreciated.

3 What Is Thermodynamics and Why Does It Work?

The point of thermodynamics is to use the values of “state variables” – a small list of the average statistical properties of a system – to describe the probability of any particular part of the system taking a certain value at a certain time. A standard example is simple equilibrium chemistry, in which we predict what the ratio of products and reactants will be after we mix two or more compounds in certain amounts and wait for the system to finish mixing (reaching equilibrium). When dealing with chemical reactions or classical industrial processes, equilibrium thermodynamics is extraordinarily successful at telling us what nature will do in a very wide range of circumstances. The stochastic revolution taught us that thermodynamics can be applied to a *much* wider range of systems yet, however, and indeed that it already is effectively the principle at work in all linguistic descriptions.

Classical thermodynamics can be formulated with either entropy or energy as the central concepts, and in a strictly equilibrium setting and in the restricted context of physical chemistry describing processes such as heat engines, turbines, and refrigerators, these variables are indeed on equal footing. To extend thermodynamics beyond the context of equilibrium physics to the context of anything that can be formulated as a stochastic process, however, it is necessary to recognize entropy as the central concept of thermodynamics, rather than energy. The reason that entropy must be the central concept of thermodynamics is that thermodynamics works fundamentally because of the “law of large numbers.” This is the mathematical basis for an intuitive fact about classical probabilities: if they are real and accurately describe a system, then you expect to see them borne out over many repetitions of identical conditions. For example, if you flip a coin 100 times and it lands heads 75 times, you can conclude either that you have just witnessed a statistical fluke which will occur with less than 1 in 10,000 odds or that your coin isn’t fair. Furthermore, if your coin really is fair, then you can always flip it another 100 times, and you will either see the probabilities converge to 0.5 or, if it came up heads 75 times *again*, you just witnessed a 1 in 10 million fluke. You’ll never rule out the fluke completely, of course, but if the coin really is weighted to come up heads 75% of the time, after just a few such trials, we can reduce the probability of the coin being fair to any value we like.

The mathematical basis for this intuition about statistics is the law of large numbers, and it is what provides the natural criterion for determining our relative confidence in the predictions of thermodynamics. In the context of coin flips, it is obvious that knowing a coin is fair only tells us that we have an equal chance of it landing heads up or tails up on any given flip and we can't do any better than pure chance for predicting the outcome of a single flip. If we want to predict the *average* outcome of many flips of a fair coin, however, 0.5 is a very good guess. In fact, in the limit of infinite flips, 0.5 is no longer a guess, but rather a mathematical fact. Any "random variable" converges to *exactly* its expectation value in the limit of infinite repetitions of the effectively random process being modeled. We say effectively random because we do not wish to get sidetracked with discussions over whether there is "true randomness" in nature. We note only that the answer to this very deep philosophical dilemma is independent of whether effectively random processes occur from the perspective of a suitable coarse-graining scheme with finite computational power at its disposal, which is sufficient for making it applicable to anything a human (or anything else that gains confidence over time via inference) will ever have the desire to calculate.

Thus, the law of large numbers tells us that the average of any effectively random process will converge exactly to its expectation value in the limit of infinite observations. The expectation value for a random process is the value you get by simply averaging all its outcomes weighted by their probabilities. For a coin, the expectation value is ½ since both outcomes occur with probability ½. Perhaps more instructively, the expectation value of a six-sided die is 3.5, because there are six outcomes and $(1 + 2 + 3 + 4 + 5 + 6)/6 = 3.5$ Also note that in this case, it's not possible for the expectation value to occur on any single roll of the dice (3.5 is not a possible outcome). Nevertheless, if you roll a six-sided fair die with sides labeled 1–6 an infinite number of times and average the outcomes, you are guaranteed to get exactly 3.5.

The law of large numbers also tells us, for more practical purposes, exactly how quickly the probability converges to the expectation value, so that we can stop making observations of the random variable when our confidence in the probabilities we have assigned to the outcomes reaches some defined threshold that itself will vary depending on need and practical constraints. For particle physics, "5 sigma," which corresponds to approximately a 1 in 3.5 million chance of being just a fluke, is an industry standard. The point is that we always have some level of statistical confidence in any real claims, and we take proof to be surpassing a predefined community standard.

To extend thermodynamics beyond the context of equilibrium means, essentially, leveraging this feature of the convergence of random variables to their expectation values, not in the infinite limit, but rather at some predefined level of interest. To show this explicitly, we will repeat a demonstration from Varadhan concerning coin flips (Varadhan 2008). We inquire about the probability of getting k heads in n tosses of a fair coin:

$$P\left(n,k\right)=\binom{n}{k}2^{-n}=\frac{n!2^{-n}}{k!\left(n-k\right)!} \tag{1}$$

Then, we apply Stirling's approximation and get:

$$P\,(n,k) \cong \frac{\sqrt{2}\pi e^{-n} n^{n+\frac{1}{2}} 2^{-n}}{\sqrt{2}\pi e^{-(n-k)} (n-k)^{n-k+\frac{1}{2}} \sqrt{2} e^{-k} k^{k+\frac{1}{2}}} \tag{2}$$

Taking the log of P results, after a good deal of algebra and applying properties of logs, in:

$$\begin{gathered} \log P\,(n,k) = -\left(\tfrac{1}{2}\right)\log 2\pi - \left(\tfrac{1}{2}\right)\log n \\ -\left(n-k+\tfrac{1}{2}\right)\log\left(1-\tfrac{k}{n}\right) - \left(k+\tfrac{1}{2}\right)\log\tfrac{k}{n} - n\log 2 \end{gathered} \tag{3}$$

Finally, making the substitution $k = nx$, where x is now the fraction of times that the coin landed heads, and collecting terms linear in n and terms with coefficient ½ (the expectation value for the stochastic process), we end up with:

$$\begin{gathered} \log P\,(n,k) = -n\,(\log 2 + x\log x + (1-x)\log(1-x)) - \\ \left(\tfrac{1}{2}\right)(\log n + \log 2\pi + \log x + \log(1-x)) \end{gathered} \tag{4}$$

$$\log P\,(n,k) = -n\,(H_{max} - H(x)) + O(n) \tag{5}$$

The function labeled H is the leading term in the log-probability of fluctuations around expectation values of random variables, whether in the "thermodynamic limit" or not. We will see presently that this function is the entropy, or to be more precise, the uncertainty of the stochastic process. There are several points to notice here. First, the term that scales linearly with n is the difference between the maximum value of H, which occurs when x is exactly the expectation value, in this case corresponding to $x = 1/2$, and a given value of $H(x)$ at any value of x. This difference of uncertainties corresponds to either Shannon's definition of mutual information or to Gibb's free energy, depending on context and our choice of units.

4 Entropy: The Central Concept of Thermodynamics

The general form of the H function, technically the uncertainty associated with the outcome of a stochastic process, is:

$$H(x) \equiv -\textstyle\sum_i P_i \log P_i \tag{6}$$

If we are careful with language, we will reserve the term "entropy" for the H function dimensionalized by Boltzmann's constant:

$$S(x) \equiv k_b H(x) \tag{7}$$

Shannon used the terms entropy and uncertainty interchangeably, which suggests a nice definition for entropy: physicalized uncertainty. It is the precise measure of how the uncertainty associated with a stochastic process shapes its statistical structure. For molecules, this is exceedingly important because their structure is entirely statistical. Existing at a scale where thermal fluctuations are large enough to constantly perturb their structure, and particularly for macromolecules that have many internal degrees of freedom or “ways that they can wiggle” their “structure” is essentially a time-averaged construct. Another reason Shannon used uncertainty and entropy interchangeably is that even though he derived the H function in his 1948 *A Mathematical Theory of Communication* (Shannon 1948), he was apocryphally advised by Von-Neumann to call it entropy because “everyone calls it that already.” Indeed, the exact same function, with Boltzmann’s constant included in the definition, appears in Gibb’s 1904 *Elementary Principles of Statistical Mechanics* (Gibbs 1902). So, everyone did call it entropy already. The major advance provided by Shannon was then not the form of the function, which had been known since the nineteenth century, but rather the extension of the context of the function’s applicability beyond what was normally considered “thermodynamics” or even “physical chemistry.” Shannon developed his information theory formalism in the context of early telecommunications systems such as Morse code. At the time, it was surprising and unexpected that the definition of thermodynamic entropy should appear in the context of telecommunication, but Shannon had done exactly what he set out to do: identify the universal mathematical constraints on any form of communication. A crucial step in this process was the removal of the most intuitively human aspect of communication: meaning. As Shannon writes on the very first page of his text:

> The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one selected from a set of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design (Shannon 1948).

The “engineering problem” in this case refers to any practical instantiation of a communication system in nature, at any scale. Shannon himself is unlikely to have been aware of the full scope of his work at the time, though it is likely that he suspected he had done something much bigger than mere “telecommunications research.” “The fundamental problem of communication” is “the engineering problem,” and the significant aspect of this problem is that “the actual message is one selected from a set of possible messages.” If you want to use this system as a means of communicating any possible message, it must be designed so as to be capable of reproducing messages at different points in space irrespective of the content of the message. The content is the meaning, and Shannon explicitly tells us that this must be neglected if one wishes to understand the fundamental problem. A moment of reflection will also show us that the meaning of a message is precisely what is not invariant from one instance of communication to another.

Indeed, what makes the meaning of a message meaningful in a particular context is precisely the fact that it is not invariant to other, otherwise identical instances of communication. “Meet me at the BBQ restaurant at 7 pm” only matters to the recipient, but any message transmitted in the same basis of equal length will have an identical bit content. What is invariant in any instance of communication therefore is that in order to communicate, the transmitted message must be part of a set of possible messages formed from a code shared between both sender and receiver. The alphabet of the code forms the basis for constructing possible messages, and those messages form a collective set which can be exhaustively enumerated for any finite alphabet and messages of finite length. It is the mathematical properties of any such set, realized in any instance of physical communication, with which Shannon concerned himself.

Shannon was able to prove the following three ‘uniqueness’ properties of the H function, meaning that it is the *only* function that satisfies these three constraints simultaneously:

1. The function should be continuous in the probabilities of signals sent or received
2. If there are more outcomes, H should be a monotonic, increasing function of the increase in outcomes. More possibilities always means more uncertainty.
3. If a choice is broken down into successive choices, the original H should be the weighted sum of the individual H.

With Shannon’s work, the ground was laid for the recognition of entropy as the primary determinant of physical structure at all scales of nature, but to understand how a measure relating to the probability of deviating from an average also relates to physical structure, we need to first understand the role of entropy in thermodynamics. Thermodynamics works in the case of physical chemistry because thermal fluctuations effectively turn the structure of molecules into random variables. Seeing how thermodynamics works in one of its most orthodox applications will help us in understanding how we extend thermodynamic descriptions beyond its traditional realm.

5 The Laws of Thermodynamics in the Context of Physical Chemistry

To begin a discussion of the meaning of thermodynamics, it is necessary to point out that despite the fact that you always hear that entropy is a measure of “disorder” or something similar, a basic understanding of dimensional analysis suffices to prove that this must also mean it is capable of measuring the amount of order present in a system. Why? Because an increase in disorder is also a decrease in order, and conversely an increase in order is a decrease is disorder. If it makes any sense at all to add quantities together or subtract them from one another, then they must have the same dimensions. So, then saying that entropy measures "disorder" or saying that entropy measures "order" is just like saying that the glass is half-empty or half-full; it makes no difference because both concepts have the same dimensions. The neutral and equally true statement is that entropy is a measure of the extent of *statistical degeneracy* in the system. This is the definition that explains the connection between entropy and structure the best as well.

Two states are "degenerate" with respect to a particular observable if swapping them causes no change in or otherwise makes no difference to that observable. The word "observable" just means anything that you can observe and, since this is science, preferably measure, although it is also okay to have observables that are just yes/no binary outputs.

For the readers who understand how the genetic code works, different codons that result in the same amino acid being added to the protein chain are also degenerate. Finally, the deepest sense of degeneracy is perhaps that every degeneracy also implies a symmetry. In physics and math, a symmetry means that you can transform an object in some way and have some other observable feature remain unchanged. This also applies to the standard notion of symmetry that you are more familiar with, where we say an object is symmetric if we can draw a line through it and the two sides are mirror-images. Any object that has that kind of mirror symmetry can also necessarily be rotated by some number of degrees and look identical. A square is more symmetric than a rectangle because both a square and a rectangle look exactly the same if you rotate them by 180°, but a square also looks exactly the same if you rotate it by 90°, whereas a rectangle goes from "standing up" to "lying on its side" when rotated by 90°.

Physicists take this notion of "any transformation that leaves something observable unchanged" a step further and speak about the symmetries of physical laws in terms of anything that you can change about a situation that leaves the laws of physics unchanged. This gets into extremely deep territory, but it's actually the explanation for the principle of conservation of energy, which is also the first law of thermodynamics. Since the work of a woman named Emmy Noether, who proved what is known as 'Noether's first theorem' in 1915 (Noether 1918), we have understood that conservation of energy is really a consequence of a phenomenon called "time-translation symmetry."

This has a surprisingly simple meaning. All chemical reactions conserve energy, for example, so they must have this time-translation symmetry property. All it means is that if I change nothing about a process except to translate it in time, then nothing about the process will change. In other words, if I set up a reaction at noon and watch it come to equilibrium and measure the concentrations of products and reactants, and then I go out to lunch and come back at 4 pm and set up the reaction again with the exact same initial concentrations, then I expect to see exactly the same equilibrium concentrations. All I did was "translate in time" by repeating the same reaction 4 hours later, and of course I will see the same outcome.

It's worth realizing that people didn't believe this for much of human history. For example, if you look at some old alchemy formulas, it was common for them to include notes like "must be performed beneath a full moon at the stroke of midnight" and so on. If you believe that it is possible that it makes a difference when you do something for no other reason than that this particular time is somehow "special," you implicitly don't believe in conservation of energy either.

Other examples of actual processes that don't conserve energy because they also don't respect time-translation symmetry are things that are changing in time. If a bathtub is being drained, then the gravitational potential energy of a rubber duck floating on the surface will decrease with time. Thus, it matters when you measure the gravitational potential of the duck, so the system "breaks time-translation symmetry" and also the energy of the system is not conserved over time, because it is being drained out over time.

Thus, the concept of degeneracy is also equivalent to the concept of symmetry, and this leads to some of the deepest ideas in physics, along with the explanation of the first law of thermodynamics.

Next is to explain the second law of thermodynamics, in a way that matches up with what we know already about fluctuations and the law of large numbers. First, it is simpler to understand what the second law says about an isolated system. Isolated systems have a fixed amount of internal energy which cannot change, and over time, once they have experienced a large number of random fluctuations, the law of large numbers and the central limit theorem guarantee that the system will be normally distributed around its expectation values for any property of the system you can measure. This will be the state of *maximum entropy*, and it defines the equilibrium of an isolated system. Because entropy can also be understood as a measure of our uncertainty about the state of a system, it is also extensive with the number of possible states. The more states the system has, the less certain we are about which state it's in simply because there are more states to choose from.

Crucially, a system only explores its states with equal probability, and hence maximum entropy, when those states are degenerate for a given energy. The reason that a gas has a large entropy isn't just that there are an Avogadro's number or more worth of particles in the gas and you can rearrange them into an enormous number of configurations, but also that the vast number of those different configurations of gas particles have *the same temperature and pressure*. In other words, the entropy of the gas is due to the fact that the gas molecules can bounce around into a seemingly endless array of distinct configurations and yet, at the same time, it looks to us like the gas isn't doing anything. That's because all of those different configurations, which are called the *microstates* of the gas, are degenerate with respect to the state variables of the gas – temperature and pressure. The temperature and pressure define the phase and thus the *macrostate* of the gas. When the molecules bounce around and the system changes microstate, the temperature and pressure *don't change* and thus neither does the macrostate of the gas. All we can detect with our senses is the macrostate of the gas and because the microstates are degenerate and it makes no difference whether any particular gas molecule is at any particular position or has any particular momentum, the gas randomly explores all of these microstates with equal probability and that is what really makes the entropy of a gas large. A system can have a million possible states, but if it stays in one of those states 99% of the time and the rest of them it almost never visits, it can still have a very small value for the entropy.

This makes sense because in reality, there is very little uncertainty about the state of the system if it occupies a single state 99% of the time. Finally, notice that thinking of entropy in terms of uncertainty gives a nice way to understand the connection between entropy and information. We don't normally think it through to this extent, but even in our everyday use of the word information, to gain information about the state of a system means to reduce our uncertainty about the state of the system. So that shows you that increases in information are reductions in entropy and conversely, increases in entropy are reductions in information, which we also call erasure of information.

If you realize that, even in a computer or your brain, information consists of real correlations between the actual states of some physical stuff, such as correlations between the states of transistors representing strings of binary code or correlations between the patterns of firing of neurons representing acquired habits, all information is stored as correlations in states of matter. Furthermore, thermal fluctuations erase correlations over time; they are randomizing forces, so erasing correlations is just what they do. This is the reason we forget things on tests as much as it is the reason that even mountain ranges erode away over time, because the ordered structure represented in knowledge or crystalline solids was formed through an interaction that reduced the entropy of the universe, and sooner or later fluctuations will erase that organization and repay the universe for the entropy it invested in that structure. That, if you like, is the only true "eternal war" (Hidalgo 2014): between correlations that create order and fluctuations that erase order.

Life, by the way, requires both order and disorder. Pure order results in dull and lifeless things like crystals that, no matter what anyone tells you, have no special healing properties. Their atoms buzz with such predictable frequency at set temperatures that we set our atomic clocks by them. Life requires the dissipation of potential energy as much as it requires the order found in semi-crystalline protein and DNA structures.

To summarize the predictions of thermodynamics, and the second law in particular, for closed systems at constant, finite (non-zero) temperature: thermal fluctuations will continue until all correlations between states of the system have been erased, and the system will randomly explore all its possible states with equal probability. At this point, the entropy of the system may be accurately calculated with the Boltzmann formula. The equilibrium state of a closed system is, by definition, the macrostate with the maximum possible entropy for the number of microstates (given the variable W), and has entropy $k_B \log W$.

The fluctuations do not cease or change in any way once the system reaches equilibrium, and the microstates themselves don't do anything different before or after equilibrium. Equilibrium has no meaning at the level of the microstates. Rather, equilibrium means that the fluctuations have erased all observable trace of the correlations (potential structure) in the system, and from that point on, the fluctuations don't have any observable effect. An example of structure induced in a gas would be if I used an electron beam to write a message in the air by ionizing it, this would create a non-equilibrium situation in the gas for a brief period, but then the gas would immediately relax back to a state where the correlations between the positions of ionized molecules that spelled out my message would disappear, which is just one of many reasons no one tries to send messages this way. Another bad, but plausible, way to transmit information is to take your finger and trace out a message on a water surface. You have to be very fast, but if you move your finger at relativistic speeds then you have time to write most of Hamlet out before the water relaxes back into equilibrium and the message is erased. If you chisel the message into granite then it will last much longer, obviously, but after some time thermal fluctuations will erase those correlations too.

The only difference is that the air and the water relax back to equilibrium quickly *relative to the timescales of our experience*, whereas the granite takes a long time to "relax" once we imprint correlations in it, but this judgment is made relative to an arbitrary scale which just so happens to be the natural timescale of our attention or our lifespan. It has no absolute meaning, and in both cases, you just have matter that gets displaced from equilibrium by having correlations which represent information imprinted into it, and after some time which depends on the size of the energy barriers to fluctuations and the temperature, those correlations will be erased. One case happens to be too fast for us to make use of and the other case appears to last forever to us, but both obey exactly the same principles of thermodynamics.

A somewhat less arbitrary example of the relationship between stability, information storage, fluctuation and computation rates, is given by considering the difference between DNA and RNA. Structurally, the reason for the "D" in DNA is because it is deoxyribonucleic acid, while RNA is ribonucleic acid. DNA is missing an alcohol (OH) group on the ribose backbone that is present in RNA. So, there was a time, probably in the "prebiotic" world, where there was only RNA and no DNA yet. RNA is already a great way to store information, because it's a natural heteropolymer and your brain just naturally wants to find correlations in the patterns of the letters even when they are basically random assortments of the four nucleotides. But RNA can only be single-stranded, and single-stranded RNA has a half-life in water of about 5 minutes. Five minutes is a long time for a molecule, and you can do lots of chemistry in 5 minutes, but it is not a very long time if you want to store information and use it again later. So you have to constantly be copying RNA over and over every 5 minutes if you want to keep the information stored in RNA, and that limits the amount of information you can store in the RNA to however much you can afford to continuously copy every 5 minutes.

This implies that the RNA world was generally information-poor. Then, 1 day, a random enzyme happened to evolve which had the property that it wanted to yank that OH group off the second carbon on the ribose of every RNA nucleotide it saw, and the second it did this, even though this wasn't the enzyme's intention (it was just trying to lower its own free energy), the RNA-turned-DNA spontaneously formed a double-helix. It turns out that alcohol was the only thing preventing the RNA from winding up on a complementary strand of itself and forming the double-helix, and this is why deoxyribonucleic acid can be a double-helix while ribonucleic acid can't. The next unintended but highly fortuitous result was that, stabilized by the double-helix, DNA can remain intact for tens of thousands of years. This is how it is possible for us to have found enough intact DNA to reconstruct the genomes of wooly mammoths and Neanderthals; the double-helix makes DNA into a near-optimal means of storing information and keeping it indefinitely. Now evolution no longer had to worry about losing information the second it stored it in a nucleotide sequence, and the amount of information in biology was able to increase to the point that it became possible to encode a cell cycle, and from there Darwinian evolution explains the rest.

Now then, when we move from an isolated system to an open system, we have to consider the entropy change of the surroundings as well as the entropy change inside the system, and the second law now says that the sum of these two quantities will be maximized at equilibrium. With a few sign convention changes and a change of units to express things in terms of energy instead of entropy, this sum is now the free energy. The free energy relates to the second law as follows:

$$\Delta S_{\text{universe}} = \Delta S_{\text{surroundings}} + \Delta S_{\text{system}} > 0 \tag{8}$$

$$\Delta S_{\text{surroundings}} = -\frac{\Delta H}{T} \tag{9}$$

$$\Delta S_{\text{universe}} = -\frac{\Delta H}{T} + \Delta S_{\text{system}} > 0 \tag{10}$$

$$-T\Delta S_{\text{universe}} = \Delta H - T\Delta S < 0 \tag{11}$$

$$-T\Delta S_{\text{universe}} = \Delta G = \Delta H - T\Delta S < 0 \tag{12}$$

The second law, which says that for open systems it is the total change of entropy in the system plus the surroundings that is always positive, can thus also be stated for open systems as:

$$\Delta G < 0 \tag{13}$$

The free energy decreases as the entropy increases. Notice that this means that for open systems, the sign of ΔS_{system} can be negative, and the system can get more ordered, as long as the entropy change of the surroundings more than compensates to make the overall total positive. A system that spontaneously organizes itself by exporting a more than compensating amount of entropy to its environment is called a “dissipative structure. (Nicolis and Prigogine 1977)” The ultimate example of dissipative structure is life. Living things reduce their internal entropy by giving off a more than compensating amount of entropy to their environment in the form of heat. This is the basis of what is called "metabolism," which is usually called the study of how an organism regulates its energy, even though it's really about entropy.

The basis of the confusion is that, in addition to deciding to multiply by a negative sign so that we have a spontaneously minimized potential rather than a maximized one, Gibbs decided to multiply through by temperature to change the equation from dimensions of entropy to dimensions of energy, even though the criterion says that entropy, not energy, is maximized. Gibb's decision to change the dimensions of the expression was so that the dependence on temperature would be manifest in the definition of the potential, but it leads to the common misconception that spontaneity has to do with minimizing energy. This is merely an appearance due to the change of units chosen by Gibbs, and the fundamental principle at work is that the entropy of the universe is always spontaneously maximized.

Examples of dissipative structures other than living organisms are stars, hurricanes, cyclones, and all manner of natural disasters, lightning bolts, candle flames, and any process you can think of that has a "fuel" which is a means of maintaining a free energy *gradient.* The gradient in free energy really means that there is a difference in the total entropy of two states external to the system which we call the "source" and the "sink." The internal entropy of the intermediate structure will be reduced – meaning the structure will get more ordered – as the entropy of the surroundings is increased and the gradient between source and sink is drained. For the biosphere as a whole, the energy source is the 6000 kelvin black-body radiation given off by the sun, and the sink is the 3 kelvin temperature of outer space. The earth is itself is intermediate between these two temperatures (averaging about 300 kelvin on the surface), and the biosphere conducts entropy between this high potential source and low potential sink as a means of driving its own evolution. The important thing to remember is that anytime there's an external entropy gradient, then there's some matter that has high potential, or information in the form of correlations, stored in it, and a dissipative structure can erase those correlations and use the liberated potential to form new correlations of its own. This is what we do when we first break down the bonds in the food we eat before carrying out anabolic processes like copying our DNA or regulating our immune system.

So that, finally, is the true meaning of free energy: it is the conjugate variable to the entropy for open systems, and it is defined as a potential energy that is minimized at equilibrium, which is equivalent to saying that the total entropy of the universe is maximized.

6 The Equivalence of Free Energy and Information

We will now demonstrate Shannon's construction of the information function, and then explain its relationship to free energy.

Returning to the H function defined in (6), we may also write the joint uncertainty concerning the state of two random variables x and y as:

$$H\left(x, y\right) = -\sum_i \sum_j P_{i,j} \log\left(P_{i,j}\right) \tag{14}$$

The joint uncertainty satisfies the following relation:

$$H(x) + H(y) \geq H\left(x, y\right) \tag{15}$$

In this context, $H(x)$ and $H(y)$ alone are referred to as the "marginal" uncertainties. When x and y are independent, the joint uncertainty equals the sum of the marginals. If x and y are correlated, their joint uncertainty is less than this sum. With this definition, we may now define the conditional uncertainty, which is the uncertainty about the state of x given the state of y:

$$H\left(x|y\right) = H\left(x, y\right) - H(y) \tag{16}$$

And then finally the Shannon information:

$$I\left(x : y\right) = H(x) + H(y) - H\left(x, y\right) = H(x) - H\left(x|y\right) \tag{17}$$

Read aloud in plain English, (17) says "the information shared between the variables x and y is equal to the uncertainty of the state of x minus the uncertainty about the state of x given the state of y." Notice that, despite frequent claims that information theory is somehow "counterintuitive," this corresponds to our intuitive notion of information perfectly well. Gaining information about x through y means reducing our uncertainty about x by measuring y.

To see how this quantity relates to free energy, we may simply think of x as referring to the state of our system and y as the environmental constraints set by the temperature and pressure of the environment. Under these assumptions, the definition of information defined between the variables "system" and "environment" corresponds to the free energy precisely. The free energy is thus also the maximum information that can be transmitted between a system and its environment, or:

$$I\left(\text{system} : \text{environment}\right) = -\beta G = \left(\frac{1}{k_b}\right)\left(S_{\max} - S_{(\text{system}|\text{environment})}\right) \tag{18}$$

This correspondence is the essence of the duality between information theory and non-equilibrium thermodynamics. Another way to understand this statement is simply as an information theoretic way of representing the usual claim that the free energy represents the maximum amount of work which may be extracted from a system. The negative sign reflects the convention chosen by Gibbs, and reminds us that the free energy is the maximum possible information that may be sent through the system-environment channel. No law of physics says that this information must actually be transmitted. Earth is, in fact, extraordinarily wasteful with its available free energy. Biological interactions which are subject to selective pressure tend to approach efficiency bounds on evolutionary timescales, however.

Note finally that this implies that free energy is a form of information, and not the other way around. Any correlation between observables is a form of information, but not all correlations are forms of free energy. Free energy is specific to physics, and information is a more abstract concept. The specific statement made here is that free energy is a measure of the information that may be exchanged between a system and its environment, under the specific assumptions that make the free energy meaningful, outlined above.

7 The Meaning of Thermodynamic Descriptions and Their Dual Representation as Information Channels

To bring together the information theoretic understanding of thermodynamic descriptions with the empirical success of the field of physical chemistry in describing the phases of matter from fluids to semi-crystalline macromolecules that occur in biology suggests that we interpret the free energy change of a chemical process as a precise measure of the amount of information exchanged between the system and its environment commensurate with that process. This view in turn illuminates the role of the environment as an information channel (Zwolak et al. 2016): the second law of thermodynamics indicates that the primary role of the environment is to erase all correlations between states of the system, but at vastly different timescales for different correlation strengths exposed to different temperatures. The stochastic process literature rests on the notion that the idea of "temperature" can be generalized to any system which can be represented by random variables that are "Lipschitz continuous." Lipschitz continuity is a well-defined mathematical idea which is best illustrated by the ideal cases of Lipschitz continuity: the temperature and pressure of a fluid.

Temperature and pressure are both time-averaged constructs for a fluid because they are an average over many different microstates of the fluid, where a microstate is understood to be a specific snapshot of the positions and momenta of all the particles at once. Since temperature and pressure are averages over very, very many atomic positions and momenta, no matter how large the change in position or momenta of a single molecule or pair of molecules, the average never changes too rapidly and thus temperature and pressure vary continuously in a fluid, even though they are actually only time-averaged constructs. This form of varying continuously despite being actually a time-averaged construct is the prime example of Lipschitz continuity, which is the strict criterion for saying whether the macrostate averages of a system will vary continuously no matter what the microscopic dynamics does.

Also, precisely because they are "time-averaged constructs," we can cause temperature and pressure to lose meaning in another way: if we try to ask about how they vary on a timescale shorter than the time between two successive microstates of the system, temperature and pressure also cease to have meaning. This is because temperature and pressure exist as averages over ensembles of microstates, and cannot therefore vary faster than the microstates themselves.

Thus, as has been understood for some time, the "objectivity" of the statement "water boils at 100° C" rests on the fact that we do not inquire about the properties of individual water molecules when making this claim. Its factual nature comes with implicit scale-limitations, as do all claims made in natural language, precisely because language is itself a coarse-grained construct.

Natural language evolved in the human species over a period of what is estimated to be roughly half of our existence, beginning about 100,000 years ago (Larson 2010). Recognizing the evolved nature of language mirrors, in a sense, the evolution of science itself in coming to terms with the meaning of the objectivity of our descriptions. The categories "subjective and "objective" derive from grammatical considerations, they are themselves artifacts of the structure of human grammar. This much, at least, should be manifest from simply considering their standard definitions.

When we enter the realm of recognizing the implicit scale-limits of nearly all scientific descriptions of nature, we are also in the territory in which language is necessarily being used in a self-aware manner. Being self-aware about the nature of language involves understanding in turn that "objectivity" itself comes with implicit scale limits, because the concepts that appear in natural language will necessarily refer at first to coarse-grained states of the world. In turn, a more coarse-grained environment can be represented by a subject with less information, and so as subjects themselves evolve in the world, and their understanding of it grows, this too will necessarily proceed in such a way as to represent that environment at ever finer scales as they seek to probe its nature in a more "fundamental" way.

In giving us a mathematical theory of communication that, through its duality with thermodynamics, could be generalized to all instances of a system interacting with an environment, Shannon provided the twentieth century with the means to be self-aware about its own use of language as a means of representing the world. When we speak to one another through digital means, we are sending signals back and forth using machinery that was developed with Shannon's equations, but we are also obeying his equations when we simply speak to one another, or even when we merely estimate the temperature by the feeling of our body, which gives us an excellent estimate of the temperature relative to our preferred homeostatic range.

All such interactions between a physical system and its environment are communication channels, and thus it turns out that the meaning of thermodynamic descriptions is also the meaning of all descriptions, and on this basis we assert the conditions for the possibility of thermodynamic descriptions to be the conditions of possibility of descriptions, period. Being a species which has attained not merely language but a self-aware use of language, it is necessary that we recognize the conditions of possibility for "the facts to be the facts." Now that we are in an era of physics that recognizes the necessity of renormalization and the "tower of effective field theories" (Kim 1998), each effective at its own range of energies or temperatures, it is past time for us to recognize that all of the words and concepts that appear in natural language "naturally" refer to coarse-grained macrostates of the world as well, with the exception of those few concepts like "microstate" which have been invented along the way to language attaining self-awareness. Thus, not merely all heretofore physical theories but also all heretofore philosophical systems constructed in any natural language have had their own "natural scales of effectiveness" whether or not they were consciously aware of them as such. Whatever form future theories of natural complexity and evolution may take, thermodynamics is sure to be relevant at any scale of nature we wish to describe, and by mirroring the structure of language itself through its duality with information theory, thermodynamics reveals the necessity of all natural descriptions being effective at some particular scale.

References

Crooks, G., Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* **1999,** 60 (3) 2721–2726.

Jarzynski, C., Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett* **1997,** *78*.

Smith, E., Large-deviation principles, stochastic effective actions, path entropies, and the structure and meaning of thermodynamic descriptions. *Reports on Progress in Physics* **2011,** 74 (4).

Parrondo, J. M. R.; Horowitz, J. M.; Sagawa, T., Thermodynamics of information. *Nat Phys* **2015,** *11* (2), 131–139.

Madan, D. B., Stochastic Processes in Finance. *Annual Review of Financial Economics* **2010,** 2 277–314.

Bartolotta, A.; Deffner, S., Jarzynski Equality for Driven Quantum Field Theories. *Phys. Rev. X* **2018,** 8.

Jacobson, T., Thermodynamics of Spacetime: The Einstein Equation of State. *Phys. Rev. Lett.* **1995,** *75*.

Leggett, A. J., The Quantum Measurement Problem. *Science* **2005,** 307 871–872.

Ledoux, M., *The Concentration of Measure Phenomenon*. Mathematical Surveys and Monographs **2001,** 89.

Romero, E.; Augulis, R.; Novoderezhkin, V. I.; Ferretti, M.; Thieme, J.; Zigmantas, D.; Grondelle, R. v., Quantum coherence in photosynthesis for efficient solar-energy conversion. *Nature Physics* **2014,** *10* (9), 676.

Gray, H. B.; Winkler, J. R., Electron tunneling through proteins. *Q Rev Biophys* **2003,** *36* (3), 341–72.

Marcus, R. A.; Sutin, N., Electron transfers in chemistry and biology. *Biochimica et Biophysica Acta (BBA) - Reviews on Bioenergetics* **1985,** *811* (3), 265–322.

Li, X. E.; Lehman, W.; Fischer, S., The relationship between curvature, flexibility and persistence length in the tropomyosin coiled-coil. *J Struct Biol* **2010,** *170* (2), 313–8.

Anfinsen, C. B., Principles that Govern the Folding of Protein Chains. *Science* **1973**, 181 223–230.

Baez, J. Quantum Techniques for Reaction Networks. Advances in Math. Phys. **2018**.

Pastor-Satorras, R.; Sole, R. V., Field theory for a reaction-diffusion model of quasispecies dynamics. *Phys. Rev. E* **2001,** 5.

Friston, K.; FitzGerald, T.; Rigoli, F.; Schwartenbeck, P.; O'Doherty, J.; Pezzulo, G., Active inference and learning. *Neuroscience & Biobehavioral Reviews* **2016,** *68*, 862–879.

Varadhan, S. R. S., Large deviations. *The Annals of Probability* **2008,** *36* (2), 397–419.

Shannon, C. E., A mathematical theory of communication. *The Bell System Technical Journal* **1948,** *XXVII* (3).

Gibbs, J. W., *Elementary principles in statistical mechanics*. Yale University Press **1902**.

Noether, E., Invariante Variationsprobleme. *Klasse* **1918,** 235–257.

Hidalgo, C., *Why Information Grows*. Basic Books, 2015.

Nicolis, G.; Prigogine, I., *Self-organization in nonequilibrium systems*. Wiley, New York: 1977; Vol. 191977.

Zwolak, M.; Riedel, C. J.; Zurek, W. H., Amplification, Decoherence, and the Acquisition of Information by Spin Environments. *Scientific Reports* **2016,** *6*.

Larson, R., The Evolution of Human Language: Biolinguistic Perspectives. Cambridge University Press **2010**.

Kim, C., Wilson Renormalization Group and Continuum Effective Field Theories. Arxiv **1998**.

A Multi-scale View of the Emergent Complexity of Life: A Free-Energy Proposal

Casper Hesp, Maxwell Ramstead, Axel Constant, Paul Badcock, Michael Kirchhoff, and Karl Friston

1 Introduction

The emergence of life – and biological self-organization – is a fascinating topic for many working within the life sciences, as well as to laypersons and scholars outside biology. We review an integrative account of the self-organization of life across temporal and spatial scales, based on the free-energy principle[1] (FEP, for short). Any view on biological self-organization must explain how organisms remain alive, that is, resist systematic dispersion and entropic decay. Organisms need to retain a grasp on their own environment in order to maintain their integrity, i.e., structure and function. For example, bacteria have implicit expectations about

---

[1] The term free energy has been used with and without hyphenation in the literature. Throughout this chapter, we write "free energy" when used as a noun (e.g., organisms minimize free energy) and "free-energy" when used as an adjective (e.g., free-energy principle and free-energy minimization).

C. Hesp (✉) · A. Constant
Amsterdam Centre for Brain and Cognition, University of Amsterdam, Amsterdam, Noord-Holland, Netherlands
e-mail: c.hesp@uva.nl

M. Ramstead
Department of Philosophy, McGill University, Montreal, QC, Canada

P. Badcock
Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Australia

M. Kirchhoff
Department of Philosophy, University of Wollongong, Wollongong, Australia

K. Friston
Wellcome Centre for Neuroimaging, University College London, London, UK

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_7

the temperature range in which their metabolism fares best (resulting in behavior called thermotropism). In this way, they resist the natural tendency toward decay or disorder. More generally, organisms embody expectations that they need to ensure are brought about through adaptive action.

Tropism in bacteria is an example of how organisms do not just passively predict their sensory states but act on their environment to realize their own expectations (e.g., concerning their preferred temperature). In other words, an organism's behavior can be cast in terms of self-fulfilling prophecies, what we call *active inference* (Friston et al. 2009). Organisms need implicit beliefs about the outer world (like the direction of a heat source) to bring about an adaptive action (moving away from the heat). Yet, they never have direct access to the outer world, only to what impinges upon their sensory receptors. Conversely, the outer world is influenced by the actions of the organism, but not by its inner states.

Thus, active inference forms a circle, from the inner world of the organism to its actions on the outer world, which feeds back to the organism through sensory stimulation. What makes this circularity virtuous rather than vicious is the information-theoretic concept of *variational free energy* (Friston 2010, 2013). Variational free energy is a measure of *the difference between what the organism senses and what it expects to sense*. Technically, variational free energy is an upper bound on "self-information," "surprisal," or simply "surprise," which reflects how surprising (or improbable) the current state of the world is for the organism (including its internal states). Although surprise itself cannot be evaluated explicitly by the organism, variational free energy can be, because it depends only on probabilistic beliefs about the world "out there," which are encoded by the state of the organism. Thus, variational free energy is a proxy for surprise.

The time average of surprise (i.e., self-information) is informational entropy. This entropy is a measure of uncertainty, which means that free energy effectively places an upper limit on the entropy of organism's sensory exchanges with the world and – if it acts in a way that minimizes expected free energy – uncertainty about its lived world. Free-energy minimization can be pursued in many ways; it has been suggested that it is an explanatory principle flexible enough to incorporate many (and possibly all) phenomena studied under the rubric of cognition (Badcock 2012; Clark 2015; Friston 2010; Hohwy 2013).

Crucially, because minimizing free energy places an upper bound on surprisal, it is equivalent to placing a lower bound on Bayesian model evidence (i.e., negative surprisal) for an implicit model (i.e., the organism) that produces expectations about sensory data. As such, free-energy minimization corresponds to a form of *variational* or *approximate Bayesian inference*, widely employed in machine learning and statistics (Friston 2010; Kirchhoff et al. 2018; Ramstead et al. 2017). This recurrent, incremental process of optimization is by its nature approximate because organisms (and machines) do not have direct access to the outer world (in a statistical sense). Organisms themselves *are* the implicit model for which they gather evidence, resulting in the interpretation that they produce evidence for their own existence – they are effectively self-evidencing (Hohwy 2016). This self-referential recurrence is central to active inference, in which all of life engages perpetually. We can therefore use approximate Bayesian inference and associated (implicit) probabilistic beliefs to characterize the interactions of an organism with its local niche.

If biological systems did not minimize free energy efficiently, the disorder or entropy of their sensory states would not be sufficiently bounded and diverge, leading to disintegration and death (in accord with the fluctuation theorem that generalizes the second law of thermodynamics to open systems). Therefore, biological systems must minimize free energy. More generally, this line of reasoning suggests that any complex adaptive (sub)system that underwrites its own existence will minimize free energy and therefore engage in active inference with respect to its surrounding environment (Friston 2010, 2013). Indeed, later on we illustrate how random dynamical systems can give rise to such inferential behaviors (Sect. 4.1).

Special care needs to be taken when relating the information-theoretic constructs that are employed in the variational free-energy formulation to thermodynamic constructs such as Gibbs entropy and Gibbs free energy. This step is important if the FEP is to act as an integrative scientific framework that leverages, and connects to, the physical sciences in the study of biological self-organization. We emphasize that variational free energy is conceptually distinct from thermodynamic free energy. The fact that both quantities share the same label (i.e., "free energy") derives from their analogous mathematical definitions. Otherwise, the relationship between the two quantities is nontrivial, and much of the work relating them remains to be done (Ramstead et al. 2018; see, e.g., Sengupta et al. 2013, for an account of this connection based on neuronal processing efficiency). The same holds for information-theoretic entropy and thermodynamic entropy, although these two constructs are more closely and straightforwardly related (e.g., through Boltzmann's famous entropy formula). The difficulty in relating these concepts stems largely from the fact that the FEP operates in a different regime from that usually considered under statistical physics. The FEP is formulated appropriately for the study of biological self-organization, since it pertains to systems at non-equilibrium steady state (NESS), whereas statistical mechanics focuses primarily on equilibrium (or near-equilibrium) states that allow for robust descriptions of physical systems in a particular equilibrated state. Having said this, the FEP and thermodynamics are internally consistent in the sense that thermodynamics – particularly stochastic thermodynamics (Ao 2008; Seifert 2012) – can be regarded as a special case of the FEP when certain conditions are met (Friston and Ao 2012).

With the above caveat in mind, we devote this chapter to reviewing the implications of the FEP for explaining the adaptive self-organization of living systems across different spatiotemporal scales, ranging from microscales (e.g., cells) to intermediate scales (e.g., learning processes of animals), and eventually to the evolutionary macroscale (i.e., the emergence of entire species). We suggest that once the FEP is extended to these different scales of self-organization, these processes, which might appear miraculous, are not really as "surprising" as one might have thought. The events that take place within the boundary of a living organism arise from the very existence of that boundary (called the Markov blanket, as explained below), the emergence of which is itself nearly inevitable in a physically lawful world like ours.

The structure of the chapter is as follows. In Sect. 2, we introduce the concept of a Markov blanket and its relation to free-energy minimization and active inference. In Sect. 3, we generalize active inference across spatiotemporal scales, to formulate a multi-scale interpretative framework for biological self-organization. In Sect. 4, we examine some examples of active inference at the sub- and multicellular microscale, notably demonstrating how active inference (i) emerges directly from a primordial soup; (ii) channels dendritic self-organization of single neurons; and (iii) enables the collective organization of many cells into entire organs. In Sect. 5, we turn to the organismic level, where we consider (i) the hierarchical brain, (ii) communication and dialogue through active inference, and (iii) cultural affordances and collective active inference. In Sect. 6, we consider the species macroscale. We first discuss how biological evolution can be viewed as a form of active inference over the order parameters of the lower levels treated in Sects. 4 and 5. Finally, we focus on niche construction and examine its role throughout both development and evolution to describe how species build their own eco-niche.

2 Markov Blankets and Active Inference

A key aspect of living systems is that they function adaptively by means of their own self-perpetuating, self-organizing boundaries (Varela et al. 1974). Adaptive self-organization enables a living system to establish and maintain a boundary that separates its internal states from the states comprising its external milieu (Barandiaran and Moreno 2008), which in turn allows for active inference. This type of boundary can be viewed as a Markov blanket. Pearl (1988) introduced the notion of a Markov blanket to denote a set of epistemological properties specific to Bayesian networks (Fig. 1). The Markov blanket is cast as the smallest set of nodes that renders an enclosed node conditionally independent of all others. The central point is that the behavior of the enclosed node can be predicted by knowing only the states of the nodes that constitute its Markov blanket. Nodes outside the Markov blanket provide no additional information. Conversely, when predicting the behavior of the nodes outside the Markov blanket, the enclosed node provides no additional information beyond that provided by the Markov blanket itself.

The notion of a Markov blanket, and the independencies between states it induces, can be directly applied to biological systems (Friston 2013; Palacios et al. 2017). For example, the interior of a cell can be related to the internal states of the cell (e.g., cell metabolism), the extracellular environment to its external states, and the cell boundary to the Markov blanket that couples intracellular and extracellular states to one another. The states that constitute the Markov blanket can be further

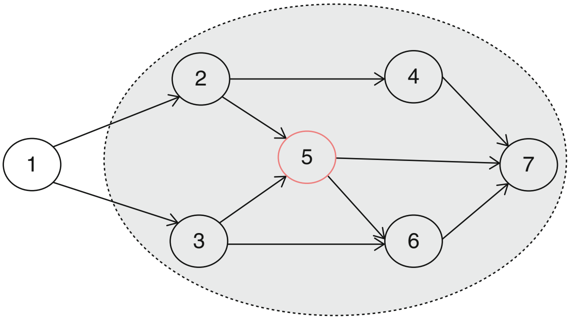


Fig. 1 An abstract graphical depiction of a Markov blanket in a network of random variables (nodes indicate, e.g., opinions of individuals in a social network) and their conditional relationships (arrows indicate, e.g., social influences between individuals). Node {5} has a Markov blanket that consists of the union of its "*parents*" (Node {2} and Node {3}), its "*children*" (Node {6} and Node {7}), and the other child of its parents (Node {4}). The Markov blanket of Node {5} does not include Node {1} – these Nodes do not directly influence each other – such that their states are conditionally independent given the Markov blanket states. In other words, when we know the states of the Markov Blanket of Node {5}, we cannot gain additional information about the state of Node {5} by interrogating Node {1} and, conversely, we cannot gain additional information about the state of Node {1} by interrogating Node {5}. It is this kind of statistical neighborhood for Node {5} that is called a Markov blanket (Pearl 1988). (This figure is from Kirchhoff et al. (2018))

partitioned into sensory and active states. As such, the presence of a Markov blanket implies a partitioning of states into external, sensory, active, and internal states (see Figure 2; Friston et al. 2015).

Figure 2 highlights the partitioning rule governing the Markov blanket formalism, namely, that hidden external states influence sensory states, which influence, but are not themselves influenced by, internal states. Conversely, internal states influence active states, which influence, but are not themselves influenced by, external states. This formulation relies on the statistical dependencies between the states comprising a biological system – internal states and their Markov blanket – and the kind of independencies induced between internal and external states. Importantly, this formulation echoes key themes of dynamical coupling between the organism and its environment in enactive and embodied approaches to biology and cognition (Engel et al. 2016; Noë 2004; Thompson 2007; Varela et al. 2017).

The dependencies established by a Markov blanket induce *active inference*, which rests on the principle that adaptive action reduces uncertainty or surprise about the causes of sensory data (Mirza et al. 2016). The statistical properties of Markov blankets result in emerging (self-organizing) processes that optimize Bayesian model evidence, such that it becomes possible to associate the internal states of a system with a model of the external states (Friston et al. 2015; Kirchhoff et al. 2018). Action, which is induced by the generation of inferences via internal states, drives an organism toward a free-energy minimum (Parr and Friston 2018). We will develop this point in further detail as we move through the various sections of our review.

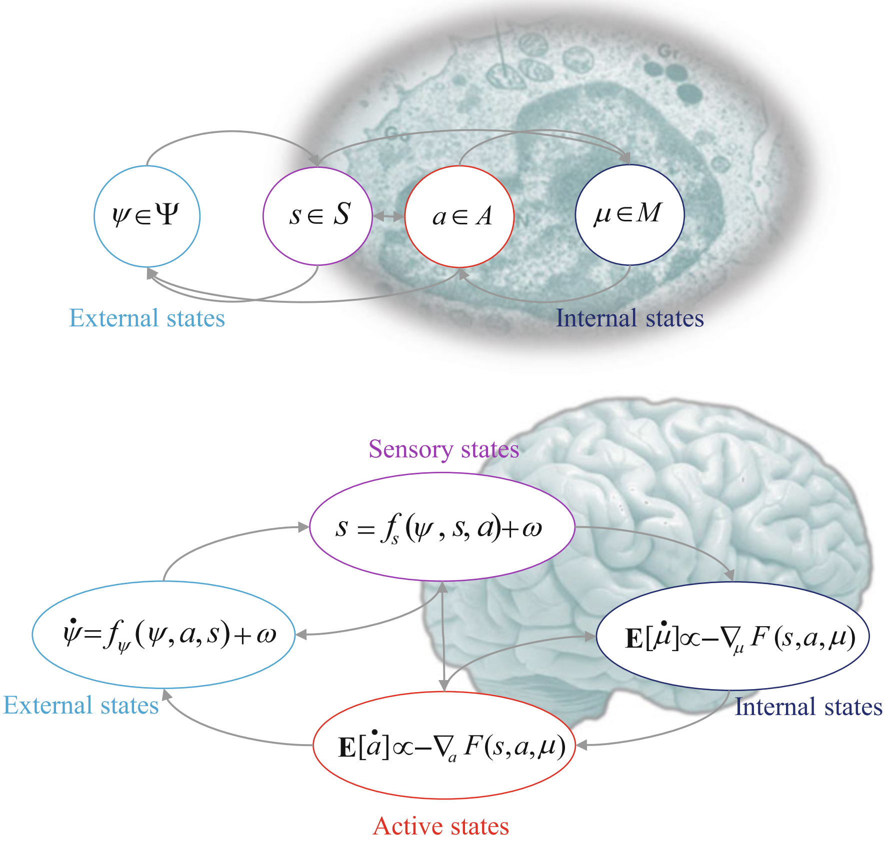


Fig. 2 These two illustrations highlight the dependencies between states induced by the presence of the Markov blanket of a cell (top) and the brain (bottom). Internal states (black) are connected to the external states (blue) through the sensory (magenta) and active (red) states. (Figure taken from Friston 2013; Figure 1)

3 Active Inference at Multiple Scales

The variational approach has recently been extended to explicitly address living systems across spatial and temporal scales (Kirchhoff et al. 2018; Ramstead et al. 2017), relying heavily on the concept of a Markov blanket introduced in Sect. 2. Any (ergodic) system that exists must, in virtue of existing, be enshrouded by a Markov blanket that maintains it. This holds for the component states of any

Markov blanketed system as well. In principle, we can describe the universe of biological systems as Markov blankets and their internal states, which are themselves composed of Markov blankets and their internal states. This formalism can be reiterated all the way up and all the way down, i.e., across the manifold nested scales of organization at which biological systems exist, including their eco-niche. In this way, biotic systems (i.e., single cells, organisms, social and cultural groups) can be described as a (high-dimensional) phase space that is induced by a hierarchy of Markov blankets. This view of living systems has been labeled *variational neuroethology* (Ramstead et al. 2017). As humans, we are a prime example: our brains, sensory organs, and muscles are themselves composed of countless cells, each possessing their own Markov boundary.

This multi-scale extension of the Markov blanket formalism involves the notion of a *scale space*. Scale spaces enable us to carve out different structures at different spatial and temporal scales and to flag which kinds of systems are relevant to our investigations at those scales. In this context, scale spaces are useful because they allow us to model the dynamics of integrated nested systems, that is, how systems at one scale produce or entail the composite system at a higher level. Moving up the hierarchy of Markov blankets entails an increase in spatial and temporal scales. Any system that can be distinguished from its environment (and, thus, possesses a Markov blanket) can take part in a dynamical interaction that produces a Markov blanket at a higher level of organization (Palacios et al. 2017).

By way of illustration, consider an ensemble of cells, each bounded by their respective plasmalemmas. We can mathematically model the self-organization of the cellular ensemble by appealing to the dynamic interactions between their sensory and active states, shaped by their collective effort to minimize free energy. Exchanges at one scale (e.g., the scale of cellular interactions) have a sparsity structure that, in turn, can induce a Markov blanket at the scale above. For example, some group of cells in that ensemble could be epithelial cells that, in turn, constitute the boundary of an entire organ. Conversely, within the cell, the various organelles have their own Markov blanket. Despite the difference in scale, the dynamics involved have a formally identical statistical structure, namely, that prescribed by the Markov blanket formalism.

The hierarchical nesting of Markov blankets provides a vantage point from which to model the self-organization of biological systems across spatial and temporal scales. Crucially, it also provides a principled explanation of how each level contextualizes (i.e., constrains) ongoing dynamics at other scales. The very same variational, entropy-bounding dynamics are operative at each scale and provide an integrative dynamics for the entire system. Free-energy minimization unifies these various scales and allows them to be evaluated simultaneously. In the following sections, we will first address the emergence of the Markov blanket and then proceed to explore the application of the free-energy principle to the various scales at which life exists.

4 Microscale: Sub- and Multicellular Self-Organization

4.1 Emergence of Markov Blankets and Active Inference in a Primordial Soup

A complete treatment of the origins of life would have to address the emergence of prokaryotic cells and their capability to produce descendants that carry their (epi)genetic inheritance. As noted before, the structure and function of the cell is a prime example of how Markov blankets induce active inference. In line with these insights, we choose to first address how random dynamical systems can give rise to subsystems that maintain themselves through active inference (Friston 2013). This is a crucial step, because once such a "primal Markov blanket" is established, the subsystem becomes self-sustaining and, hence, susceptible to innovations and organization into larger composite systems. For example, it is thought that some of the organelles within eukaryotic cells used to be prokaryotic cells themselves (i.e., mitochondria and chloroplasts). Although this is far from a full account of life as we know it, we can use abstract representations of dynamical processes to illustrate some simple but fundamental aspects of adaptive self-organization. These processes may serve as a metaphor for dynamical interactions across various levels of biological self-organization.

The following theorem will serve as a guideline in what follows: *if a random dynamical system is ergodic and has a Markov blanket, it actively maintains its own structure and dynamics* (i.e., autopoiesis; Friston 2013). Ergodicity is a key concept, which formally means that the average of any measurable function within the system converges over time. This definition implies that a limited number of states are being revisited, because not all functions would converge for an infinite number of possible states. By virtue of ergodicity, the average proportion of time a certain state is occupied (within a sufficiently large window) is equivalent to the probability of the system being in that state when observed at random. In other words, an ergodic random dynamical system is tractable in terms of probabilities, which is crucial for any type of inference. Ergodicity is readily identified as a key property of biological systems. For example, neurons switch between their resting, firing, and refractory states.

Friston (2013) provided a proof of principle of this simple but fundamental property of living systems. He modelled a "primordial soup" that exhibited the type of behavior described in the theorem presented above. These simulations consisted of a collection of dynamical subsystems, which can be likened to macromolecules. Each of these macromolecules could reside in a number of possible structural and functional states and was coupled by these states with other nearby macromolecules. The type of dynamics employed in these simulations is similar to those in the wealth of literature on pattern formation in dissipative systems, e.g., turbulence in hydrodynamics (e.g., Manneville 1995). In the context of Friston's simulations, structural states represented the locations and motions of these macromolecules, while functional states represented their electrochemical states. Through electrochemical interactions, functional states can influence the location and velocities (structural states) of nearby molecules, as well as the electrochemical states of those molecules. The intention of this exercise is not to analyze the precise patterns that emerge from these interactions, but rather to demonstrate that a basic form of active inference can emerge from a "primordial soup."

While each of the subsystems themselves only has a limited number of possible functional states (i.e., they are locally ergodic), the simulations also exhibited emergent ergodic behavior for the system as a whole. Initially, macromolecules pushed each other away; after a few cycles, they tended to clot together, forming a stable dense clump. Short-distance interactions led to a pattern in which macromolecules were passed around until they only gently pushed and pulled on each other most of the time, with occasional bursts of movement. The collective motion and electrochemical states of this dense emerging clump could be characterized as a "restless soup," as shown in Fig. 3.

Is there any active inference evident in this synthetic mess? Given that the global attractor state of the system and the subsystems themselves are ergodic, we can characterize their behaviors in probabilistic terms. We can then use the coupling between the states of these macromolecules to disentangle their spheres of influence. Based on this information, we can identify the Markov blanket (if present) and the states enclosed by it. Friston (2013) found that amidst the densest region of the "soup" were a number of macromolecules that were very tightly coupled to one another and whose states were completely hidden from those residing on the outer edges of the system. Figure 4 shows the macromolecules representing internal states (dark blue) and those representing the Markov blanket as the sensory (magenta) and active (red) states. The active macromolecules, which allow the internal states to affect the outer world indirectly, lie within the sensory subsystems that are exposed to the outer world. Interestingly, biological cells have a somewhat similar configuration, with an (active) cytoskeleton surrounded by (sensory) epithelia or receptors.

Crucially, a minimalistic form of perception was also identified within the clump of macromolecules. Although particles in the interior were entirely insulated from the outer world, their functional states were shown to have predictive value for the motion of the macromolecules *outside* the clump. In a self-organized fashion, these mindless, simplistic "representations" of macromolecules appeared to be producing implicit inferences about the world outside their synthetic bubble.

Friston also showed that the implicit inferences – driven by the (sensorial) dynamics of the inner environment of the clump – directed the active states to maintain its structure. In this way, the clump of macromolecules essentially anticipated future perturbations induced by the outer world and acted on these expectations: a basic form of active inference.

We can now return to the theorem introduced above. Does the emergent clump of macromolecules indeed "actively maintain its structural and dynamical integrity"?

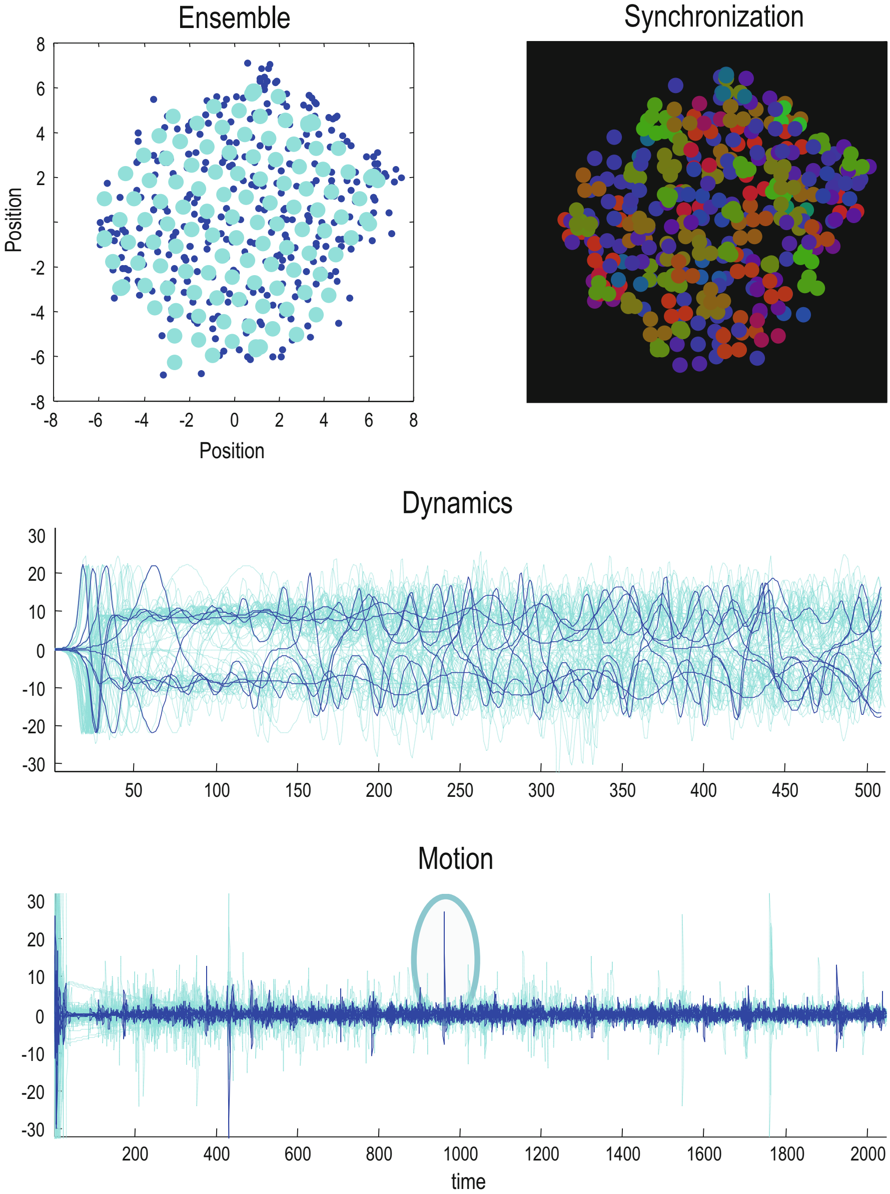


Fig. 3 Reproduced from Friston (2013; Figure 1); this figure shows the structure and temporal dynamics of the simulated primordial soup. Panel (**a**) (i) illustrates the spatial position (large cyan dot) and functional states (three dark blue dots) for each of the 128 subsystems, after the states have converged on their global (random dynamical) attractor. Panel (**a**) (ii) shows the same snapshot of time with the three functional states coded by color, illustrating the synchronization of electrochemical states across the clump. Panels (**b**) and (**c**) show, respectively, the functional states and motion as a function of time (in seconds, processor time). Internal states are shown in blue and external states in cyan. The circle in panel c indicates one of the occasional bursts of motion due to the nonlinear dynamics within the clump of macromolecules. See Friston (2013) for technical details

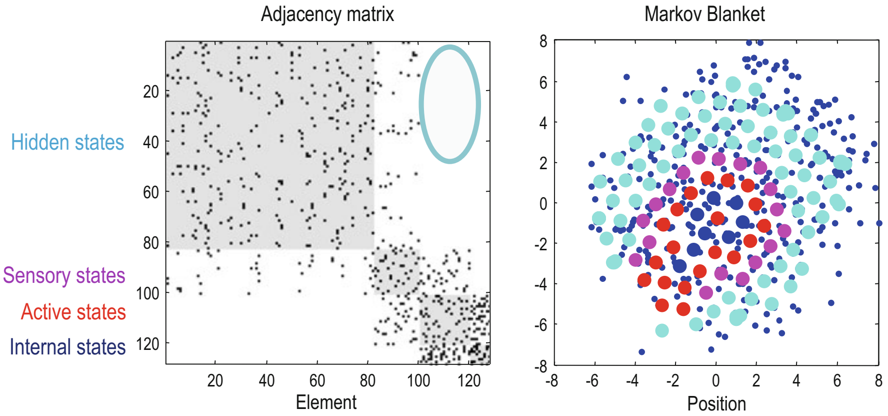


Fig. 4 This figure shows the emergence of the Markov blanket from the primordial soup after the global attractor state was reached. The left panel shows the coupling between the 128 macromolecules over 256 seconds (adjacency matrix), ordered according to the internal (blue), active (red), sensory (purple), and external hidden (cyan) subsystems. The circle indicates instances of active subsystems influencing external states (owing to the periodic bursts of motion) without the external states influencing the active states. The right panel shows the spatial organization of this partition. (Reproduced from Friston (2013; Figure 2))

This question can be answered by perturbing the system with “lesions”: selectively turning off the ability of certain macromolecules to affect the functional states of other macromolecules for the active states (Fig. 5b), the sensory states (Fig. 5c), and the internal states (Fig. 5c). Note that all of the electrochemical effects on motion were left intact; only the subtle interaction between electrochemical states was silenced. In all three cases, such a relatively mild perturbation caused the synthetic bubble to burst instantly. This empirical result substantiates the prediction that macromolecules will affect their neighbors in order to maintain the structural integrity of the entire clump.

In this section, we have seen the emergence of a Markov blanket and resulting active inference in a random dynamical system. Functionally speaking, the simulated clump of macromolecules is probably most reminiscent of the various protein components that allow viruses to maintain their structure. We can see them as a metaphor for more extensive forms of biological self-organization. Intriguingly, Friston did not require a very “special” setup to arrive at this result in a bottom-up fashion; very little was required, in fact. This motivates our proposal to consider the recursive self-organization of Markov blankets into Markov blankets at higher levels. Each of these blankets and their internal states again constitute a unit of free-energy minimization (Ramstead et al. 2017; Sengupta et al. 2016). With this in mind, we will now proceed by taking free-energy minimization “for granted” and focus instead on how this process shapes function specificity for a single neuron (dendritic self-organization) and form specificity at multicellular levels (morphogenesis).

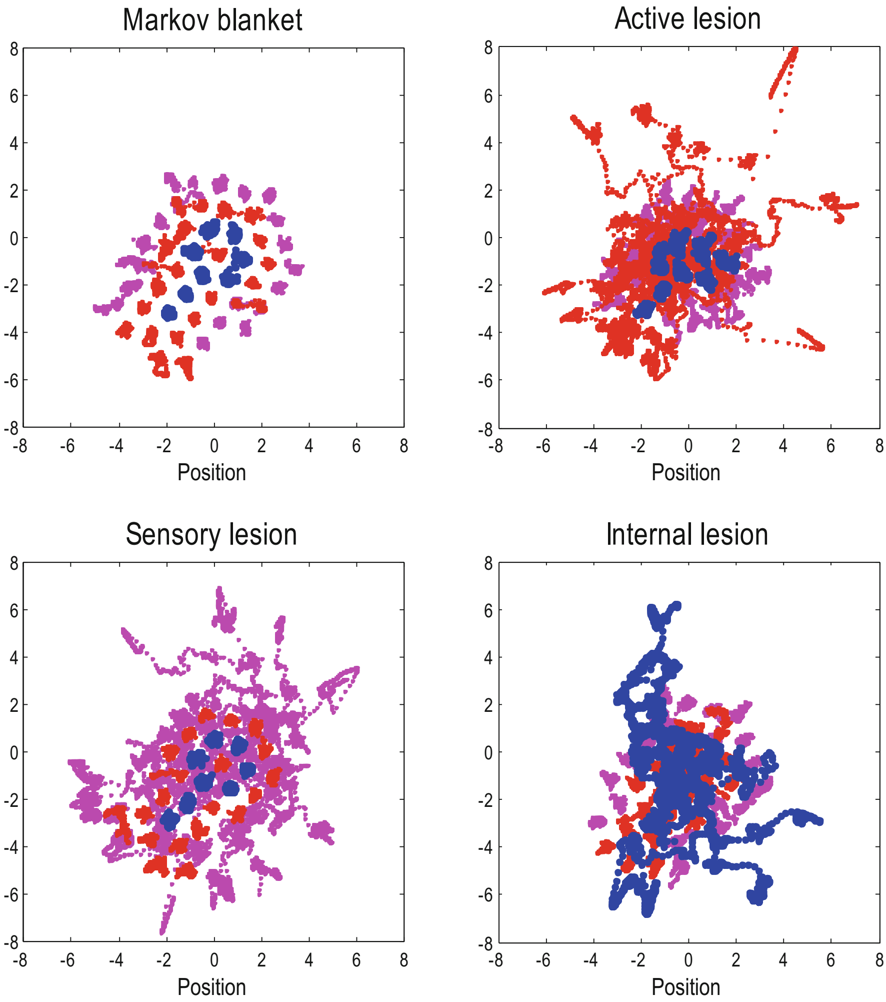


Fig. 5 This figure demonstrates the self-maintaining dynamics (i.e., autopoiesis) of the clump of macromolecules, by slightly impairing the components of the emergent Markov blanket. Impaired macromolecules are rendered unable to influence the electrochemical states of other macromolecules (but all other interactions are left intact). In the top left panel, the configuration without a lesion is shown, with the internal (blue), active (red), and sensory (pink) macromolecules forming a stable configuration. In the top right panel, active macromolecules are impaired, causing them to be expelled into the exterior. In the bottom left panel, the sensory macromolecules are impaired, causing them to drift off into the exterior. In the bottom right panel, internal macromolecules are impaired, causing the entire configuration to collapse – as the internal states migrate rapidly across the Markov blanket. (This figure is adapted from Friston (2013; Figure 4))

4.2 Dendritic Self-Organization

Different types of neurons code for different types of synaptic input sequences, as evidenced by their different morphologies and connections (Torben-Nielsen and Stiefel 2009). Pyramidal neurons have been shown to engage in sequence-specific processing (Branco et al. 2010). Apparently, dendritic branches allow the dynamics within a single neuron to distinguish various sequences of input from each other. In the following, we discuss how the FEP has been used to study the emergence of such function specificity by Kiebel and Friston (2011).

As stated in the introduction, under the FEP, the variational free energy represents the difference between what a biological (sub)system senses and what it expects to sense. These expectations are derived from an *implicit (generative) model* of those sensory inputs. The biological system itself *is* this model, which specifies the type of inputs it is looking for (note, once again, the inherent circularity). The minimization of free energy has been used to simulate systems that decode their sensory states and actively select the types of input they expect to sense (Kiebel et al. 2008). The implicit nature of these expectations and models is worth emphasizing, because it means that these Bayesian concepts do not require the system itself to be "conscious" of inferences in any way or that these inferences need to be "explicit" and couched in propositional or linguistic terms.

A single neuron or one of its dendrites can also be understood as a biological system that engages in free-energy minimization. As we will see, this view can explain the emergence of the sequence-specific functionality of neurons toward presynaptic inputs with a certain temporal pattern. Selection of synapses occurs via synaptic gain control – synapses with low gain are pruned, and synapses with high gain stimulate the formation of synaptic connections (Lendvai et al. 2000). The concept of synaptic gain control can itself be derived from the FEP; and it can be used to capture the behavior of neuronal dynamics across multiple timescales, from fast electrochemical potentials, to variations in synaptic gain, through to slowly changing synaptic connections. In a series of simulations, Kiebel and Friston (2011) incorporated these three temporal scales in a computational model by using three levels of simultaneous free-energy minimization: a single quantity is minimized at the three scales that enclose the scale in which synaptic gain is determined. Figure 6 illustrates the type of sequence selectivity that emerged in these simulations.

On the fast level of electrochemical currents, Kiebel and Friston (2011) were able to show that this free-energy-minimizing dendrite model produced emergent dynamics that were entirely consistent with data-driven models of dendritic dynamics (Gulledge et al. 2005). Their findings showed that such active dendritic dynamics are a self-organizing function of this particular biological system under the FEP. The slow dynamics of the dendrite – rearranging the synaptic connections over time – is incorporated in the model as a form of Bayesian model selection. The various connections are essentially producing evidence for their own efficacy with varying degrees of success, instantiating a process of selection over time. Selection occurs stochastically, allowing for completely non-efficacious configurations but

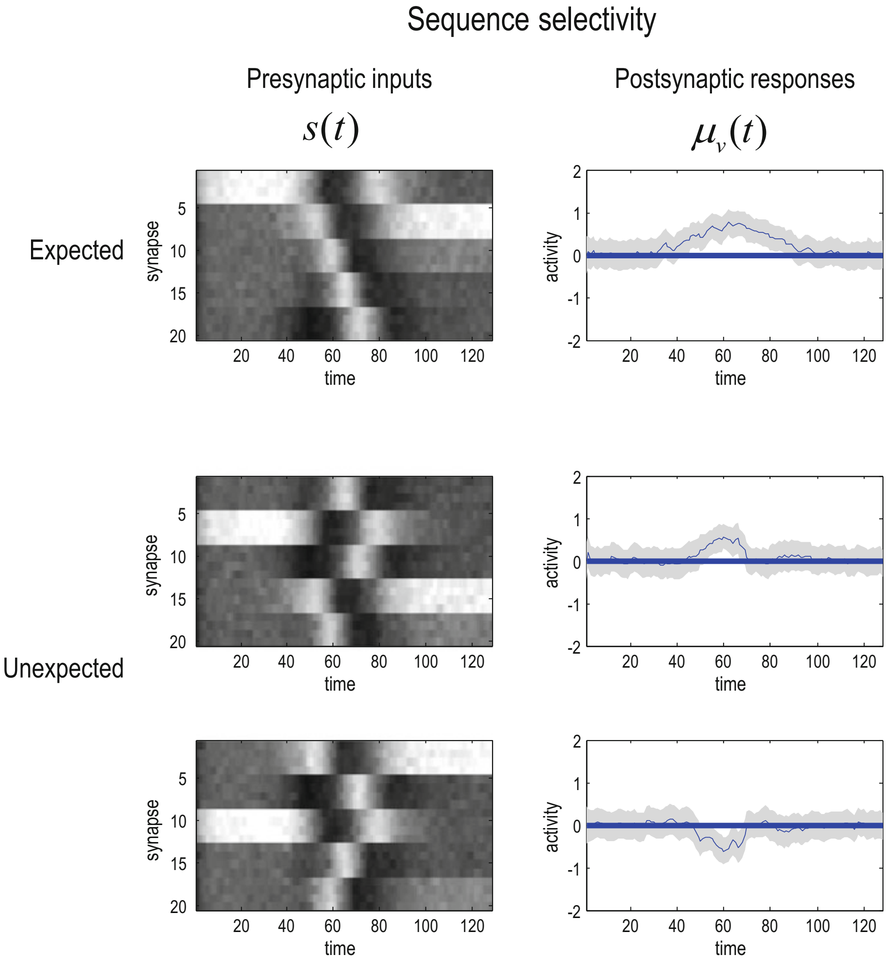


Fig. 6 In this figure, we show the responses of the dendrite (right column) to three different sequences of presynaptic input (left column). The top row shows the expected sequence to which the dendrite is accustomed, showing a peak in the postsynaptic response (top right panel). The middle and bottom rows show how the dendrite responds to sequences that deviate from its expectation, with attenuated postsynaptic responses in both cases (middle and bottom right). The graded response in the bottom right panel is consistent with graded observed in neural responses to suboptimal input. (Figure was taken from Kiebel and Friston (2011; Figure 7))

also rendering the routine better equipped to escape local (suboptimal) minima. In Sect. 6, we discuss how a similar kind of dynamics governs evolution by natural selection. Notably, a similar type of model selection is also believed to drive the fine-tuning of entire neural networks, which has been broadly conceptualized as *neural Darwinism* (Edelman 1987).

In this subcellular example, the dendrite is minimizing free energy to improve (i) its beliefs about presynaptic input sequences on short timescales, (ii) its beliefs about synaptic gain (or precision), and (iii) its implicit model of the input sequences over longer timescales. In this way, the dendrite adjusts its prior beliefs about the type of sequences it expects to observe, which results in the observed selective sensitivity. The sampling method of the dendrite is being adjusted over time, which boils down to a type of active inference. In the following, we consider how a group of cells (free-energy-minimizing units) can self-organize into larger structures, namely, organs.

4.3 Morphogenesis

Now that we have established cells as units of free-energy minimization, we can consider how adaptive self-organization occurs under collective active inference, i.e., the group dynamics of cellular ensembles (Friston et al. 2015). An important example is the emergence and maintenance of the large-scale shape and function of entire subsystems (e.g., organs). How can cells at microscales coordinate to form predefined large-scale structures, e.g., during embryonic development? Or, at later stages, how can creatures like salamanders regenerate entire limbs and organs? It is an essential question for biology, both in development and throughout evolution, to consider how cellular ensembles control exact large-scale outcomes in order to allow for specific functions to emerge (e.g., brain or liver function). Insights into this issue are particularly crucial to medicine and bioengineering.

As we will see, collective active inference can explain the self-organization of an ensemble of cells to generate entire organs (i.e., morphogenesis). The most pressing difficulty here is that organs will only function if they have a highly specific, predefined form, for which unguided pattern formation is insufficient. Since any one cell only has access to the signals reaching its boundary, it would seem that it can only infer its location and differentiate once the other cells have already migrated to their respective target positions and differentiated accordingly. However, that requirement cannot be reached if those other cells themselves are unable to determine their own target positions.

This inherently circular problem of organ formation can be solved through active inference, if we assume that every (pluripotential) cell starts with a generative model of the entire ensemble. In this way, every cell can generate predictions about the sensory inputs it expects to encounter at any location in the target configuration. As with stem cells, all cells start out in nearly identical states, with the same generative model and the ability to differentiate, that is, transition toward any role in the eventual organ. As each individual cell starts minimizing free energy, the entire ensemble will converge toward its global free-energy minimum. By virtue of their common generative model, this global minimum is approached when the ensemble closes in on the target shape and function of the organ. Each cell will gradually infer its own place and behave accordingly while, crucially, helping other cells to infer their place in the process. Such self-assembly will also serve to maintain the configuration and, in the case of damage to the organ, restore it.

In order to substantiate this account, Friston and colleagues (2015) conducted simulations of cell migration and differentiation in a relatively minimalistic sense. Each cell possessed a generative model, the parameters of which were determined "genetically" (they were inherited or prespecified), which prescribed to each cell how to act (i.e., what signals to emit) given a particular place within the organ. Hence, cells exchanged signals with each other in order to infer their respective place and role in the ensemble. The upshot of this is that every cell has a probabilistic grasp on its location and emits signals accordingly, providing information for the other cells to improve their own inferences. That relation between the beliefs of a cell concerning its place and the signals it transmits to other cells could be an elegant metaphor for epigenetic processes. Figure 7 serves to illustrate the simulation results for a configuration of a relatively small number of cells. It shows both the differentiation process and the reorganization of the ensemble after two different large lesions.

We have illustrated how ensembles of free-energy-minimizing units (cells) that operate with the same generative model can self-organize into predetermined structures (organs). This allows us to understand how an intricate functional structure like

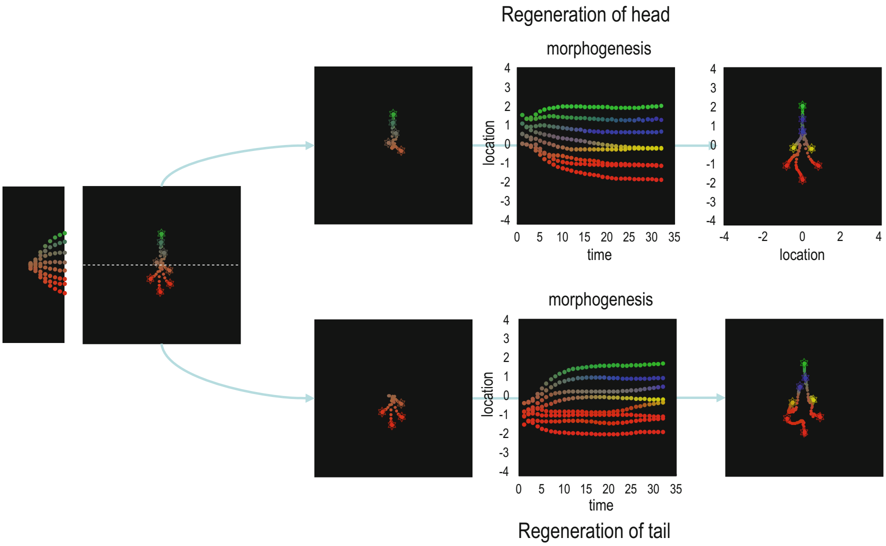


Fig. 7 This figure shows both the differentiation of eight stem cells to form an "organ" (on the left) as well as the regenerative response of the configuration to two large lesions (on the right). In the top three panels on the right, the "head" (consisting of red cells) is severed, and the remaining cells are doubled to maintain the same number. On the bottom, the same operation is performed on the "tail" (consisting of green cells). Both show that the pattern is successfully recovered. (Figure taken from Friston et al. (2015; Figure 4))

the brain can be produced by the (epi)genetic information transmitted at conception. This treatment has prepared us for a discussion of the brain, entire organisms, and their interactions. Interestingly, we will see a similar sort of dynamics emerging in the interaction between multiple organisms: a shared generative model allows for the emergence of communication and cultural dynamics.

5 Mesoscale: Organisms and Their Interactions

5.1 The Brain

At this point, we arrive at the level of organization involving animals and the interactions between them. We would be remiss if we would not reserve a few words for the animal brain in particular. Its organization and functional dynamics could be understood in terms of the examples of free-energy minimization treated thus far. The brain exhibits a layered and modular structure, instantiated through morphogenesis (Sect. 4.3). We suggest this organization of the brain has been selected for throughout evolution (Sect. 6.1) because it enables the assembly and maintenance of hierarchical generative models (Badcock, Friston, & Ramstead; under review; Friston 2010). In our environments, there is an abundance of hierarchical inference problems. For example, in the case of natural images, the integration of large numbers of features is required in order to identify objects under countless possible lighting conditions and rotations in space. In computational neuroscience, free-energy minimization has led to the development of models engaging in hierarchical predictive processing that successfully capture the functioning of the brain (Adams et al. 2016). The brain is thus viewed as an active inference machine (Clark 2015), specialized for complex inferences requiring hierarchical generative models. It would not be an overstatement to say that it is the most complex adaptive system known to mankind, as it continuously bridges the scale space from genes to single dendrites up to organismic and societal levels (Ramstead et al. 2018). Under the FEP, the brain essentially functions like its many lower levels of organization: it predicts sensory states from its internal model(s) of how those sensory states are caused (see the hallmark paper on the FEP by Friston 2010). It minimizes the discrepancies between its expectations and actual sensory states by modifying its implicit beliefs (i.e., perception) or by acting on its environment (i.e., behavior). The inferential power of the hierarchical organization of the brain can be well illustrated by studying how it generates predictions about another hierarchical dynamical system, namely, another organism. We choose not to focus on how the brain instantiates bare forms of perception and action but on how two bird brains are coupled through birdsong. This will serve as an informative example of hierarchical inferential dynamics enabled by free-energy minimization.

5.2 Birdsong as a Model of Dialogue

When two dynamical systems are coupled to each other, a form of synchronization usually occurs. This was first reported by Huygens (1673), who studied the synchronization of pendulums hanging from a beam, through which they influenced each other very slightly. Because both pendulums operate in the same way, even the minimal information transmitted by the beam is enough to completely synchronize them. In a similar way, coupled brains can, by virtue of their similar internal dynamics, achieve generalized synchrony. Such synchrony allows for these systems to predict one another with very high precision. In the case of identical internal models, identical synchronization is achieved (similar to the case of the pendulums). The more dissimilar the internal models of two organisms, the less synchronization will occur between their internal states and, consequently, the less accurate their predictions will be about each other's actions. Without environmental constraints, coupled organisms will tend to move toward the free-energy minimum of identical synchronization. In other words, they end up forming a model of each other. Through such coupling, organisms can "program" each other toward a common internal model; namely, they end up speaking the same "language" (in an abstract sense).

The way in which dynamical coupling gives rise to generalized synchrony in pendulums can thus be applied to the fine-tuning of hierarchical internal models that generate predictions. Such learning was addressed by Friston and Frith (2015), which is the focus of this section. The authors demonstrated how organisms can come to interpret each other's actions simply by adjusting their internal models to minimize free energy. Importantly, free energy can be evaluated without these organisms ever knowing exactly what is happening beneath the Markov blanket of the other. It relates to the central problem of hermeneutics: how do we infer the intention behind an utterance, when we only have access to the utterance itself? In the following, we discuss simulations by Frith and Friston that are an abstract representation or metaphor of communication between organisms, based on the mathematical machinery of complex dynamic systems. Synthetic birdsong is used for this demonstration, but it is not meant to represent actual linguistic processes. The authors merely aimed to study dynamic coupling via complex action patterns, which are themselves without meaning or syntax. However, it is worth noting that other researchers have applied hierarchical predictive processing to language (e.g., Hickok 2013) and auditory processing (Arnal et al. 2011).

In order to simulate birdsong-like behavior, Friston and Frith (2015) constructed a hierarchical processing architecture, which is shown in Fig. 8 (overlaid on analogous neuroanatomical structures of the bird brain). Free-energy minimization is achieved through recurrent connections between different levels of the hierarchy, each of which possesses its own generative model. Each level generates its own expectations about how sensory inputs are caused, which are passed downward as predictions. Each level (except the highest one) therefore receives (top-down) predictions to compare with its own expectations. The difference is the prediction error, which is passed back to the higher levels in a bottom-up fashion in order

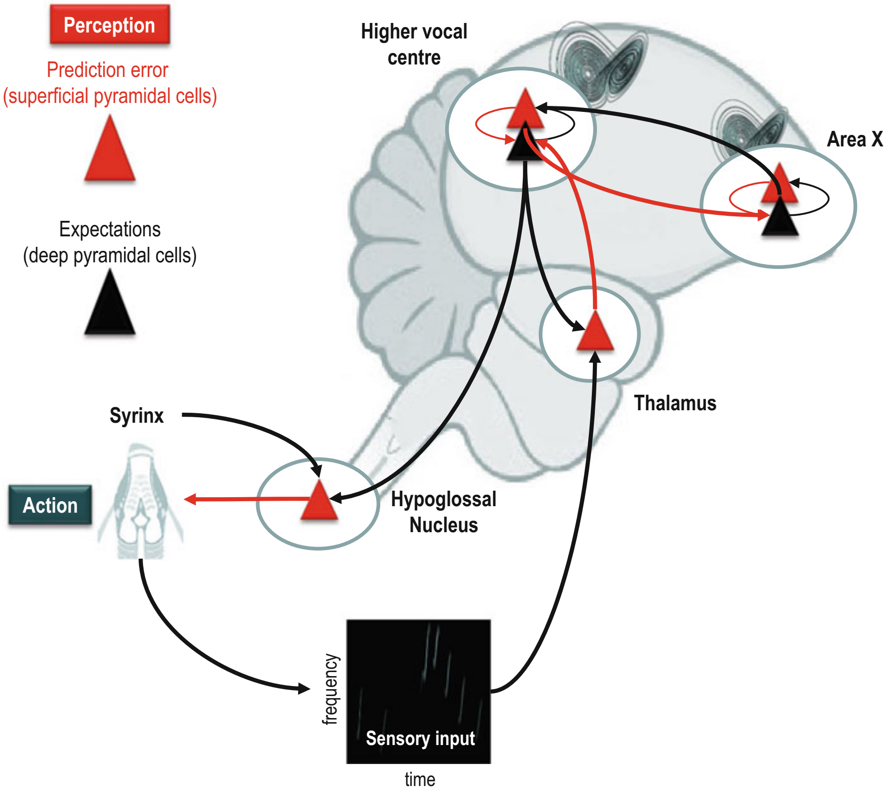


Fig. 8 This figure illustrates schematically the hierarchical predictive processing architecture of the songbirds, overlaid on (possibly) analogous neuroanatomical structures of an actual bird brain. Red arrows indicate the flow of information about the prediction errors, transmitted by superficial pyramidal cells (red triangles). Black arrows indicate the flow of information about the expectations on each level, transmitted by deep pyramidal cells (black triangles). Area X transmits predictions to the higher vocal center, which drives the hypoglossal nucleus to generate a vocal response (via the syrinx) as well as the thalamus to generate the corollary discharge. (Adapted from Friston and Frith (2015; Figure 1))

to improve future predictions. Experimental findings appear to support such an architecture. For example, it has been suggested that superficial pyramidal cells are involved in calculating prediction errors and passing them upward and that deep pyramidal cells pass the expectations of each level to the one below in the form of predictions (Bastos et al. 2012). In this hierarchy, the predictions of the lowest level are essentially those generating motor commands and corollary discharge.

The simulations of Friston and Frith (2015) showed that two of these synthetic bird brains became coupled through their vocalizations during a turn-taking exercise, providing clear evidence of generalized synchrony. These dynamics occurred at the free-energy minimum of the two coupled systems. Importantly, a high degree of synchronization was achieved because both systems started out with a similar architecture (or neuroanatomy), by virtue of being birds. Both of these birds were simply predicting their own sensory states, using a hierarchical composition of hidden states. The final product emerged in their "dialogue," so both of their hidden states had essentially come to represent their shared expectations. This meant the only thing the birds had to infer was which one of them was singing (i.e., agency). This inference enabled them to either attend or ignore the sensory consequences of their action, depending upon whether they were listening or singing.

Perhaps a more interesting – and realistic – case is when two birds are in some ways dissimilar to each other, rendering their dyad asymmetric. One of the birds was given a mild handicap by reducing its responsiveness to top-down predictions, which also hampered the quality of its vocalizations. As shown in Fig. 9, this adjustment allowed a type of scaffolding dynamics to emerge, in which the more proficient bird simplified its own vocalizations in order to accommodate the shortcomings of the other bird. Through this process, they reached nearly identical synchronization, solving the hermeneutical problem in the process (so to speak). Interestingly, this kind of demonstration is analogous to scaffolding techniques used in teaching, in which the teacher optimizes learning by lowering his or her level of instruction close to, but slightly above, that of the student. For reference, Fig. 9 also includes a simulation in which the birds are disconnected from each other, showing how heavily the richness of their vocalizations depended on the presence of another bird. When the birds were alone, they started learning from the silence around them to become silent themselves. It shows that, in some way, the teacher was actually learning from the student too.

Although this coupled setup was rather ad hoc, it can be seen as a step toward understanding the development and evolution of social life. Through generalized synchrony, one could efficiently infer the sensations and action goals of others, a crucial aspect of higher cognitive functions. Important examples are vicarious learning (learning by watching others), empathy (inferring others' feelings), and theory of mind (inferring others' inferences). A form of generalized synchrony appears to underlie mirror neuron activity in animal brains – mirror neurons not only fire during certain actions or sensations but also when observing a conspecific performing or experiencing similar actions or sensations (Friston et al. 2011; Kilner et al. 2007). This type of associative mirroring of neural responses appears to be similar to the generalized synchrony exemplified in the above birdsong simulations. It has been argued that mirror neurons are an associative by-product of action understanding and empathy (Hickok 2010; Cook et al. 2014). In future work, studies that investigate the ways in which free-energy minimization leads to generalized synchrony between organisms might help explain observations of mirror neuron activity.

In this section, we have discussed the sort of learning dynamics that emerge when two hierarchically structured, free-energy-minimizing (bird-like) organisms interact. Once again, circular relationships are involved, now in the context of communication and generalized synchrony, resulting in the emergence of shared expectations. In the following, we discuss how shared expectations and narratives shape human cultural dynamics.

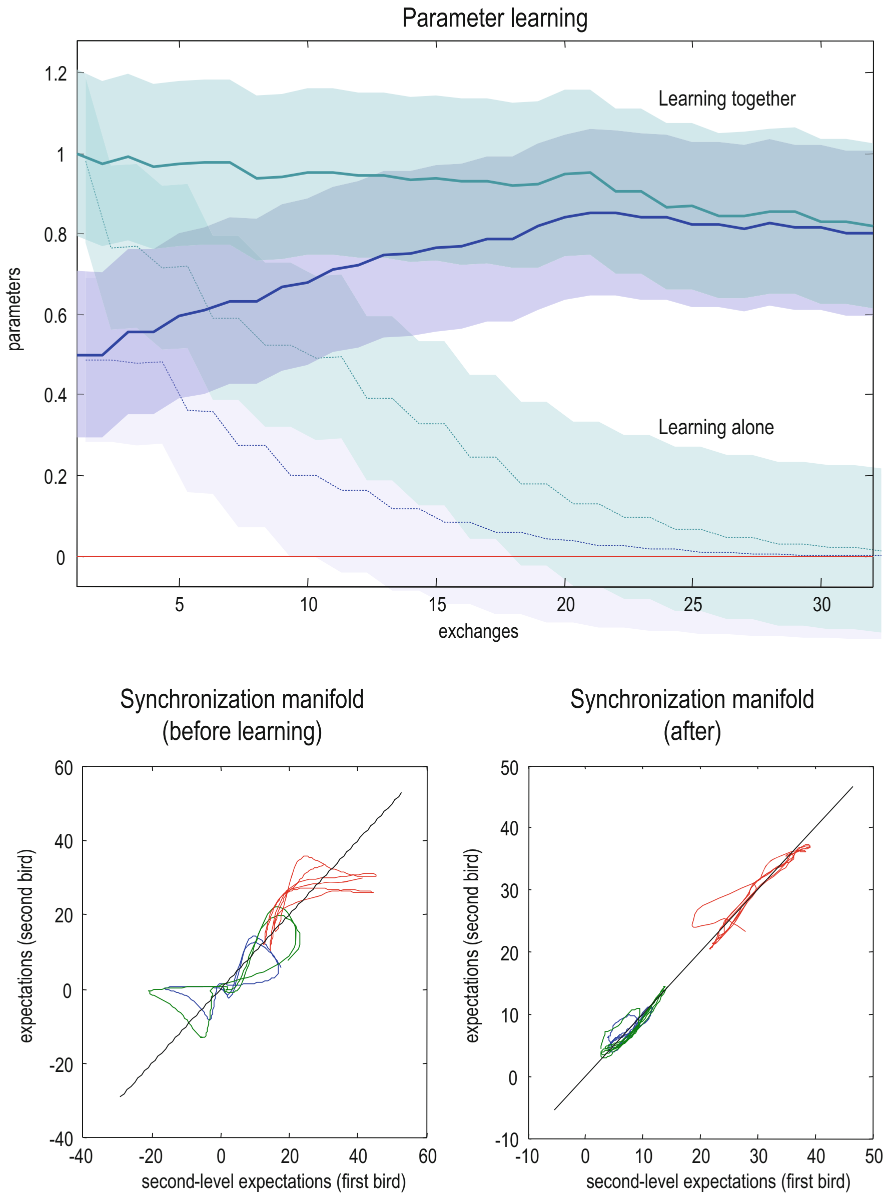


Fig. 9 This figure illustrates both the learning of two coupled birds (top panel) and the generalized synchrony reached after their exchanges (bottom two panels). The top panel shows the changes in both birds' (posterior) beliefs about a parameter that controls the prosody (or richness) of their vocalizations over a number of exchanges (birds taking turns, either singing or listening). The proficient bird is shown in green and the less proficient one in green. 90% confidence intervals over this parameter are indicated by the shaded areas. The bottom panel shows the degree of synchronization between the expectations of the birds about three hierarchical, dynamic states that drive the singing behavior (red, green, blue), both before (left) and after (right) their exchanges. Since the x-axis shows the expectations of the less proficient (first) bird and the y-axis those of the more proficient (second bird), synchronization is achieved on the line x = y. (Figure taken from Friston and Frith (2015; Figure 8))

5.3 Cultural Ensembles

So far, we have seen – across various scales – how biological systems come to embody an implicit model of their environment through active inference. The emphasis on organism-environment coupling is inherent to the free-energy principle, which plays very well into another framework that has recently gained traction among researchers, *ecological and embodied approaches to cognition* (Bruineberg and Rietveld 2014; Chemero 2009; Gibson 1979; Kirchhoff 2015; 2017a, b). In this section, we discuss an example of recent efforts to connect these frameworks in the context of human cultural dynamics by Maxwell J. D. Ramstead et al. (2016)

A synthesis between the free-energy principle and the ecological approach allows these approaches to benefit from each other's insights and research. From the conceptual toolbox of ecological cognition, we introduce the notion of *affordances*. Affordances are possibilities for engagement through action and perception that are enabled by the relationship between the environment and the abilities of the organism in question. Under the FEP, an organism acts on its environment in order to bring about its preferred (expected) sensory outcomes (Bruineberg 2018). In this way, free-energy minimization specifies the most likely trajectories of organisms in their landscape of affordances.

Ramstead et al. (2016) made the distinction between natural and cultural affordances. Affordances of the first kind are derived directly from the environment (e.g., walking) and only require minimal social learning, while those of the second kind are derived from the shared expectations inherent to the (sub)culture in question (e.g., language) and require more extensive social scaffolding to be acquired and used effectively. The previous section illustrated how shared expectations can emerge from interactions between two organisms. In the case of culture, we generalize this notion to a population of interacting organisms united by one common set of shared expectations, which in turn shape the various possibilities for interaction, namely, cultural affordances. Of course, the distinction is not absolute; natural and conventional affordances are more like the opposite ends of a spectrum of affordances. For example, in many cases, conventional rules simply act to constrain natural affordances (e.g., driving on the wrong side of the road).

Researchers developing the concept of affordances emphasize that agents use sensory information for affordances, without requiring explicit representations of the affordances themselves (van Dijk et al. 2015). This minimalistic view sits well with active inference, given that statistical terms are seen to be implicit (as we noted earlier). Expressed otherwise, internal models are implicitly instantiated by the dynamics themselves. For example, in the hierarchical architecture introduced in Sects. 5.1 and 5.2, free-energy minimization occurs locally on each level of the hierarchy, based only on the neural signals incoming from adjacent levels. None of these levels necessarily requires "meta-cognitive" contextual information about the hierarchical internal model.

So how do humans become so proficient in leveraging this field of (implicit) cultural affordances? Under a hierarchical predictive processing architecture, any level can modulate expectations at the level immediately below it, thereby modulating which types of input that lower layer is sensitive to. Such prior expectations can implement a gating mechanism, which has been proposed to explain attention. In principle, cultural affordances could then be learned by fine-tuning these priors to induce selective attention, which constrains the field of all possible affordances. Effectively, this can be arrived at by extending the modelling strategy for morphogenesis of Sect. 4.3, by equipping all cultural agents with the same cultural priors. Such culture-specific fine-tuning of internal models can occur through the type of generalized synchrony discussed in Sect. 5.2. Shared expectations that emerge from collective free-energy minimization induce "regimes of shared attention" that guide and constrain social practices, which in turn shape those expectations (Ramstead et al. 2016).

Under this view, social norms can be cast as shared "solutions" arrived at, and learned through, the collective free-energy minimization of people within a particular culture (Colombo 2014). The shared aspect of social norms reflects a certain degree of synchronization between people within a given (sub)culture, allowing them to produce more accurate inferences about each other's internal states. For example, it is much easier to predict the actions of, and empathize with, somebody from your own (sub)culture than somebody from an alien one. This emergent view of social norms and practices corresponds well with *social constructivism*, a well-established framework in sociology, which emphasizes that human development is socially embedded and human narratives are constructed through interaction with others (Berger and Luckmann 1966). We suggest that the free-energy principle can undergird social constructivism by explaining how shared cultural narratives can emerge from, and are learned through, human interactions.

Finally, the shared aspect of cultural affordances suggests that human social capacities emerged not "just" because of more advanced hierarchical internal models, because one's grip on cultural affordances is learned only through interaction with other humans within that culture. The converse would therefore seem more likely: our processing hierarchy has been optimized through evolution in order to keep up with the growing demands of the early social practices of our primate ancestors. This is a prime example of evolution through both natural selection *and* niche construction, which we discuss in the following section.

6 Macroscale: Species as Families of Model-Niche Pairings

6.1 Evolution as Bayesian Model Selection

We are now prepared to address one of the three central topics of this volume – the evolution of species – which provides the context in which the smaller temporal scales of adaptive self-organization are embedded. We assume familiarity with evolutionary theory, so we do not completely hash out the basic concepts of evolution, but rather explore how these concepts can be understood as free-energy minimization at the species level. In particular, we discuss evolution as a form of Bayesian model selection.

In our treatment thus far, we have assumed biological systems to be ergodic. Ergodicity implies that a system only resides in a limited number of states over time, which makes probabilistic inferences (and hence active inference) possible. Of course, real biological systems are only locally ergodic. Throughout the development of an organism, various states are pruned away and new ones are unlocked, sometimes quite radically (e.g., a caterpillar becoming a butterfly). Eventually, death involves a divergence of possible configurations – a complete breakdown of ergodicity (from the perspective of the phenotype). The complex adaptive systems, we refer to as organisms, do not maintain their structure and function forever: indeed, in a changing environment, the emergent Markov blanket of Sect. 4.1 would eventually be destroyed. In the beginning of evolutionary history, this (perhaps inevitable) disintegration has been overcome through the emergence of the ability to reproduce. Reproduction is an adaptive capacity that allows genetic, epigenetic, and nongenetic information to be transmitted to descendants along with small variations, constraining the self-organizing dynamics that specify the form and function of their internal models for active inference. Through inheritance and the subsequent experiences of organisms, every new generation introduces variations of the internal models of their parent population. Inherited aspects of these internal models can be realized in various ways, which we discuss now.

In Sect. 4.3, we saw how the large-scale shape and function of organs can be fine-tuned through the initial internal models of stem cells. Such processes can bring about the hierarchical organization of the animal brain, which in turn allows for hierarchical internal models, as discussed in Sects. 5.1 and 5.2. Besides the overall hierarchical structure of internal models, another type of heritable modulating mechanism could be instantiated through adaptive priors (within a given brain organization) that predispose the organism to learning certain types of structures (Friston et al. 2012; Ramstead et al. 2017). For example, humans appear to have an innate disposition for the acquisition of language. Another very important form of evolutionary preparedness is the inborn affective value of various types of stimuli. In the context of free-energy minimization, an innate tendency to approach or avoid certain situations could be implemented implicitly through prior preferences over sensory inputs. Internal models can be adapted to tweak the expected free energy under various sensations, without (strong) reliance on learning through experience. For example, we all respond with disgust to the smell of rotten eggs without ever having experienced hydrogen sulfide poisoning. On a more positive note, we all tend to enjoy the taste of sweet and fat-rich food (a tendency skillfully exploited by modern fast-food chains). There are also examples of complex stimuli that are known to have an innate affective value. For example, all mammals appear to be predisposed toward developing a fear of snakes (Badcock et al. 2016). Captive-born lemurs and macaques learn to fear snakes faster than other types of equally rich stimuli (Weiss et al. 2015). This finding has led some to suggest that snake-like reptiles used to be a large threat to the survival of mammals in an early stage of evolutionary history. Under the free-energy principle, innate preferences over inputs are not limited to the lowest (sensory) level of the predictive-processing hierarchy. Preferences over inputs can also apply to the incoming (sensory-driven) signals on higher levels, which could explain the innate affective value of highly complex stimuli like snakes. Again, on a more positive note, the same mechanism can also explain the positive experience of "cuteness" invoked by the bodily proportions of babies (and, probably an evolutionarily "accidental" corollary, puppies and kittens). Indeed, the important role of adaptive priors in active inference has even been leveraged to explain highly complex human phenomena, such as our capacity for depression (Badcock et al. 2017).

Now that we have specified the ways in which evolutionary preparedness can be realized through internal models, we can consider the selection process itself. Natural selection is underwritten by differentials in adaptive fitness. Whatever traits are most suited to ensuring the survival and procreation of individuals are most likely to be transmitted (genetically and epigenetically). Consequently, these traits will occur more frequently in subsequent generations. Constrained by the transmission of (epi)genetic information to the next generation, natural selection acts primarily on individuals (i.e., individual fitness), although it can also occur through an individual's contribution to the survival and reproductive success of others, especially close relatives (i.e., kin selection and inclusive fitness; Dawkins 1976; Hamilton 1964; Maynard Smith 1964; Orgel and Crick 1980). Notably, the evolutionary success of a species depends strongly on the amount of (epi)genetic variation present in the populations that constitute the species (evolutionary resilience, e.g., Sgrò et al. 2011). Such variation increases the likelihood of the presence of individuals with high fitness under new, challenging circumstances. Every individual represents an attempt to transmit its (epi)genetic makeup, such that natural selection effectively produces a stochastic gradient ascent on the expected fitness of the population (as employed in machine learning by, e.g., Yi et al. 2009).

The FEP provides a framework to predict adaptive fitness from first principles while also taking into account organism-environment interactions. In effect, maximizing the adaptive fitness of a population is likely achieved by minimizing its collective free energy, which tracks the goodness of fit (or complementarity) between the states of a species and the states of its niche. Accordingly, individuals that are well suited for survival are those that minimize free energy efficiently. Generation by generation, the adaptivity of individual organisms can be evaluated by the negative time average of free energy (i.e., a lower bound on entropy). Since free energy itself is evaluated using the internal model of a specimen, a comparison only has predictive value for adaptation if the family of internal models and the niche under consideration are similar. The strictly local utility of this comparison illustrates the incremental nature of evolution.[2] For example, using free energy as a metric, one could score the complementarity between, say, a bacterium and its niche and a human being and its niche. In short, free energy could provide a universal

---

[2] From a technical point of view, the extensive nature of free energy means that the sum of the free energy of the parts is equal to the free energy of the sum; so what is "good" locally is good globally. This extensive characteristic implies that minimizing free energy over time is analogous to the Hamiltonian principle of least action – because action is the integral of energy over time.

proxy for adaptive fitness that could be applied to both viruses and vegans. At the same time, free-energy minimization is achieved through gradient descent, which means in this context that it is quintessentially species – or model – specific.

As noted in the introduction, minimizing free energy is formally equivalent to maximizing Bayesian model evidence, that is, the likelihood of the internal model being true or apt, given the organism's environment. Therefore, we are now in a position to interpret processes of adaptation as collecting Bayesian model evidence and, by extension, to cast natural selection as a form of Bayesian model selection (see also Campbell 2016). On this view, creatures are naturally selected according to how well their internal generative models fit with the environment. Of course, this picture becomes more complicated in the case of organisms that interact with each other to increase total fitness (i.e., decrease collective free energy). These multiple organisms are not only "fitting" their shared environment but also each other, generating shared expectations in the process (as seen in Sect. 5.2). By virtue of the inherited directives for their internal models (i.e., adaptive priors), which have been shaped by natural selection, organisms minimize their free energy locally over their own (relatively short) lives in ways that *also* help their descendants (e.g., parents nurturing their children) and close relatives (i.e., kin selection). Local (organismic) free-energy-minimizing dynamics are structured in such a way that they collectively move toward a (population-level) free-energy minimum. This type of relationship between local and global dynamics is analogous to the predictive-processing hierarchy in the brain, as described in Sects. 5.1 and 5.2. Every layer in the hierarchy minimizes its own free energy (locally), in such a way that it also helps the hierarchy as a whole move toward its free-energy minimum (globally).

Thus far, our discussion of evolution has yet to explore how organisms shape their own environment, which can also become part of the inheritance they leave behind for their descendants. Such niche construction – and implicit legacy – is the focus of the final section.

6.2 Niche Construction

Niche construction is the process by which organisms modify their environment through their normal bioregulatory activity (Odling-Smee et al. 2003). It encompasses *all* modifications, from the induction of a layer of moist air around homeothermic organisms, to the construction of complex environments like cities by human beings. Like perception and action, niche construction is ubiquitous in living systems. Indeed, it is a direct corollary of active inference – organisms attune the statistical structure of their environment to their probabilistic expectations by acting in a way that is guided by those expectations. We have seen that perception enables the organism to infer sensory causes; action places an upper bound on surprise by generating expected changes in the sensorium. The variational free-energy approach to niche construction exploits the symmetry in the Markov blanket formalism, namely, between internal and external states mediated by the blanket states (i.e., the fact that action engenders modifications of the local environment, which embeds sensory causes). In this section, we explore the role of such ecological modifications with regard to evolution.

By virtue of ergodicity, an organism may be defined as the most *likely* set of physiological and behavioral states for any given set of environmental states. The coupling between these states then constitutes the entire organism-environment state space. As stated in Sect. 6.1, adaptivity is a feature of an organism-environment system, not just of organisms themselves (e.g., gills are adaptive for water-bound organisms; lungs for those dwelling on land). Negative variational free energy can be seen as a measure of adaptivity (as in Sect. 6.1), either for individual organisms and their niche or for larger ensembles like groups and species and their environment. It tracks the extent to which the statistical organization of an organism's physiological and behavioral states transcribes the statistical organization of the states of its environment.

Among those states, some pertain to the internal organization of the organism. These are fast, fluctuating states, like synaptic connections and neuromodulatory gating patterns. Some other states pertain to the external, visible organization of the organism (i.e., phenotypic states). These are more slowly fluctuating quantities, like behavioral patterns and morphological features (i.e., phenotypic traits). States of the environment themselves can be interpreted as part of those slowly fluctuating states. The level of adaptivity among slowly and rapidly fluctuating states depends on the interplay between variational optimization processes spanning different spatiotemporal scales, ranging from natural selection (Bayesian model selection), through to development and learning (active inference).

At this point, the notion of action and environmental modifications become important. Because action fulfills sensory expectations (e.g., adaptive priors concerning viable states, like body temperature), it can change, implicitly, the statistics of the niche so as to make them consistent with the sensory expectations of an organism. In other words, niche construction fits sensory causes to sensory expectations and, reciprocally, fits sensory expectations to sensory causes. Under active inference, niche construction is crucial in allowing for optimization across the scales of the spatiotemporal hierarchy: the more slowly changing parameters embodied by or encoded in the physical features are optimized through niche construction, and in return act as a kind of developmental driver by channeling adaptive behavior and phenotypic accommodation (Bruineberg 2018; Constant et al. 2018a, b). In a sense, the robustness of living systems is inherited from the regularity and stability of their more slowly changing eco-niche.

Adaptation is often conceived of as a one-way process, by which natural selection shapes organisms under the pressures of their environmental conditions: a view sometimes called "externalism" in the context of natural selection (Godfrey-Smith 1996). As we considered in Sect. 6.1, those pressures pose challenges, the resolution of which rests on the retention of organisms that are best suited to gain differential fitness. Under the FEP, this corresponds to the selection of (constraints to) internal models that are most suited to minimize free energy.

The niche construction perspective involves a complementary view of adaptation, in which internal factors, like the states of organisms, also play an evolutionarily significant role in their adaptation. Organisms generate feedback interactions with their environment, which can steer their own evolutionary trajectories, not to mention those of other species (J. Odling-Smee et al. 2003). These can generate new challenges, requiring the deployment of novel traits and behaviors in order to resolve them.

Recursive processes in niche construction impact two different, yet overlapping spatiotemporal scales: development and natural selection (Stotz 2017). At the level of development, niche construction modifies the environmental inputs to an organism's development, along with those to its offspring (e.g., through parental care). Such modifications often involve making the environment congruent with the expectations of the organism (Constant et al. 2018a, b). At the level of evolution, niche construction functions as a strategy to modify the selection pressures afforded by the environment, thereby impacting the adaptive fitness of future generations. For instance, as a natural consequence of dam building, beaver kits inherit ecological resources like dam remains that, in turn, support the typical life cycle of beavers (Naiman et al. 1988).

We can thus see how niche construction leads to the inheritance of environmentally transmitted information (as opposed to information transmitted through reproduction) that, throughout ontogeny, helps the organism minimize its uncertainty about the states of its environment that are likely to provide a fitness advantage (e.g., palm nut residues that guide the learning of food exploitation techniques in capuchin monkeys; Fragaszy 2011; Fragaszy et al. 2017). Such information is known as algorithmic information, which is an important source of nongenetic inheritance (Odling-Smee et al. 2013). In the context of free-energy minimization, algorithmic information enables the organism to maximize mutual information between the model it has genetically inherited (i.e., its adaptive priors), and the causal states of the environment that it has inherited ecologically. Indeed, in virtue of the symmetrical statistical dependencies across the Markov blanket of any phenotype (or ensemble of phenotypes), one can also regard the environment as entailing a generative model of the phenotypes (or ensemble) to which it plays host. Again, we see a circular causality that can be operationalized by noting that the free energy of a creature is complemented with a (conjugate) free energy of its environment, where the active states of the creature become the sensory states of the environment (and vice versa). If the free energy of either forms a proxy for adaptive fitness, we have a formal measure of "fitness" that can be applied to both phenotype and eco-niche. From an evolutionary perspective, this means that the environment will appear to be subject to selective pressure. In summary, the FEP therefore undergirds niche construction theory by providing (i) a principled measure of fitness that is optimized across spatiotemporal scales and (ii) a computational framework to reflect on ecological inheritance (Fig. 10).

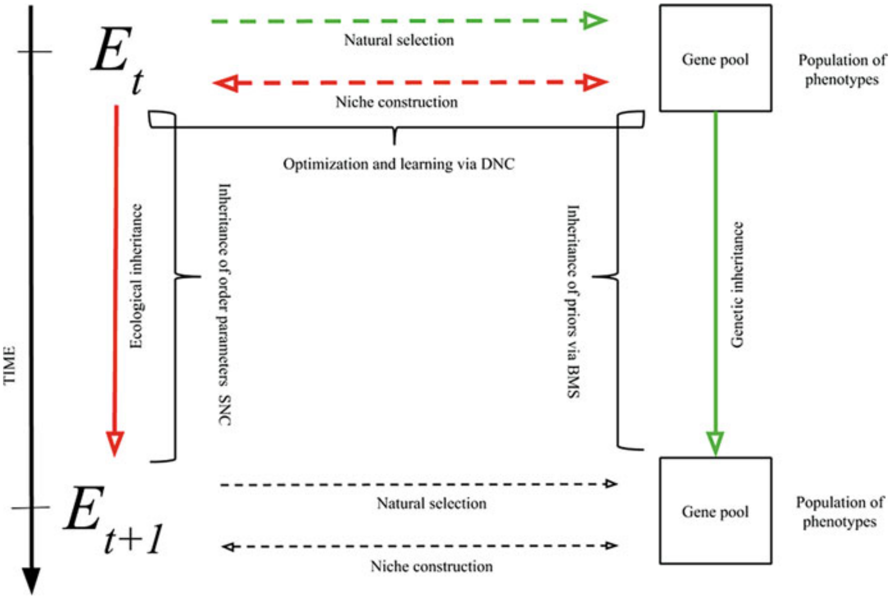


Fig. 10 Adaptation under the FEP. This schematic – inspired by Odling-Smee and Laland (2000) – illustrates the evolutionary processes covered thus far in this chapter (colored arrows). These conspire in real time t to secure the adaptation of future generation at time $t + 1$. Ecological inheritance via selective niche construction (SNC, top-down, red full arrow) is interpreted as the transmission of environmental components that support variational updates (learning) in development (e.g., phenotypic accommodation). The FEP interprets genetic inheritance as Bayesian model selection (BMS, top-down, green full arrow), which leads to the inheritance of model components, selected on the basis of their ability to maximize adaptive value (negative surprise). Inherited priors are those predictable from the organism's ancestors' ability to cope with the environment, in the sense of attaining free-energy minima (or the neighboring of a limited repertoire of physiological and behavioral states). Niche construction over development (DNC, lateral, bidirectional red dotted arrow) is described in terms of model optimization via active inference and entails ecological inheritance. Note that niche construction in development causes the symmetry between the organisms and the niche they inhabit, hence the bidirectional arrow

7 Conclusion

In this chapter, we have demonstrated how the FEP can be applied to understand adaptive, biological self-organization across spatiotemporal scales. Free-energy minimization implies active inference, which in turn allows biological systems to actively maintain their structure and function. We have discussed how Markov blankets, the basic unit of free-energy minimization and requirement for active inference, can emerge by themselves from a primordial soup. Across the manifold scales considered herein, similar processes of adaptive self-organization recurred in various ways – just as Bayesian model selection gives rise to sequence specificity in a single dendrite (Sect. 4.2), it also shapes entire neural networks (neural Darwinism) and can be used to understand natural selection (Sect. 6.1). Shared internal models allow for the organization of many cells into entire organs (Sect. 4.3), but they also allow for the emergence and continuation of dialogue (Sect. 5.2) and culture (Sect. 5.3). Local optimization at separate levels of the hierarchical brain also enables system-wide free-energy minimization (Sects. 5.1 and 5.2), while individual free-energy-minimizing organisms contribute to the adaptive fitness of an entire species (Sect. 6.1). Just as organisms can carve out expectations in each other's internal models through interactions (Sect. 5.2), they can construct niches in their environment that sculpts the models of their descendants (Sect. 6.2). All of these interconnected examples serve to illustrate how the FEP has the potential to provide a unifying framework for the multi-scale complexity of life. Our intention is not to replace existing theoretical frameworks but, rather, to provide an underlying, quantifiable description from first principles that can be used to integrate and coordinate such frameworks. For example, along the way, we have discussed embodied cognition, social constructivism, evolutionary theory, and niche construction. A unifying theoretical description can provide support for these various frameworks and allow them to benefit from the mathematical machinery of the FEP.

References

Adams, R. A., Bauer, M., Pinotsis, D., & Friston, K. J. (2016). Dynamic causal modelling of eye movements during pursuit: Confirming precision-encoding in V1 using MEG. *NeuroImage*, *132*, 175–189.

Ao, P. (2008). Emerging of Stochastic Dynamical Equalities and Steady State Thermodynamics. *Commun. Theor. Phys.* (Beijing, China), 49, 1073–1090.

Arnal, L. H., Wyart, V., & Giraud, A.-L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, *14*(6), 797–801.

Badcock, P. B. (2012). Evolutionary systems theory: A unifying meta-theory of psychological science. *Review of General Psychology: Journal of Division 1, of the American Psychological Association*, *16*(1), 10–23.

Badcock, P. B., Davey, C. G., Whittle, S., Allen, N. B., & Friston, K. J. (2017). The Depressed Brain: An Evolutionary Systems Theory. *Trends in Cognitive Sciences*, *21*(3), 182–194.

Badcock, P. B., Ploeger, A., & Allen, N. B. (2016). After phrenology: Time for a paradigm shift in cognitive science. *The Behavioral and Brain Sciences*, *39*, e121.

Barandiaran, X., & Moreno, A. (2008). Adaptivity: From Metabolism to Behavior. *Adaptive Behavior*, *16*(5), 325–344.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012, November 21). Canonical Microcircuits for Predictive Coding. *Neuron*.

Berger, P. L. & T. Luckmann. (1966). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, NY: Anchor Books.

Branco, T., Clark, B. A., & Häusser, M. (2010). Dendritic discrimination of temporal input sequences in cortical neurons. *Science*, *329*(5999), 1671–1675.

Bruineberg, J. (2018). Anticipating affordances: Intentionality in self-organizing brain-body-environment systems (Doctoral dissertation). Retrieved from UvA-DARE.

Bruineberg, J., & Rietveld, E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Frontiers in Human Neuroscience*, *8*, 599.

Campbell, J. O. (2016). Universal Darwinism As a Process of Bayesian Inference. *Frontiers in Systems Neuroscience*, *10*, 49.

Chemero, A. (2009). Radical embodied cognition. Cambridge, MA: MIT Press.

Clark, A. (2015). *Surfing uncertainty: prediction, action, and the embodied mind*. New York, N.Y.: Oxford University Press.

Colombo, M. (2014). Explaining social norm compliance. A plea for neural representations. *Phenomenol. Cogn. Sci.* 13, 217–238.

Constant, A., Bervoets, J., Hens, K., & Van de Cruys, S. (2018a). Precise Worlds for Certain Minds: An Ecological Perspective on the Relational Self in Autism. *Topoi. An International Review of Philosophy*, 1–13.

Constant, A., Ramstead, M. J. D., Veissière, S. P. L., Campbell, J. O., & Friston, K. J. (2018b). A variational approach to niche construction. Journal of the Royal Society, Interface, 15(141).

Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C. (2014). Mirror neurons: From origin to function. *Behavioral and Brain Sciences*, *37*(2), 177–192.

Dawkins, R. (1976). *The Selfish Gene*, New York: Oxford University Press.

Edelman, G. M. (1987). *The Theory of Neuronal Group Selection*. New York: Basic Books.

Engel, A. K., Friston, K. J., & Kragic, D. (2016). *The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science*. MIT Press.

Fragaszy, D. M. (2011). Community Resources for Learning: How Capuchin Monkeys Construct Technical Traditions. Biological Theory, 6(3), 231–240. https://doi.org/10.1007/s13752-012-0032-8

Fragaszy, D. M., Eshchar, Y., Visalberghi, E., Resende, B., Laity, K., & Izar, P. (2017). Synchronized practice helps bearded capuchin monkeys learn to extend attention while learning a tradition. Proceedings of the National Academy of Sciences, 114(30), 7798–7805.

Friston, K. J. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience*, *11*(2), 127–138.

Friston, K., Levin, M., Sengupta, B., & Pezzulo, G. (2015). Knowing one's place: A free-energy approach to pattern regulation. *Journal of the Royal Society Interface*, *12*(105), 20141383.

Friston, K. J. (2013). Life as we know it. *Journal of the Royal Society, Interface / the Royal Society*, *10*(86), 20130475.

Friston, K., & Ao, P. (2012). Free-energy, value and attractors. *Computational and mathematical methods in medicine*, 937860.

Friston, K. J., Daunizeau, J., & Kiebel, S. J. (2009). Reinforcement learning or active inference? *PloS One*, *4*(7), e6421.

Friston, K. J., & Frith, C. D. (2015). Active inference, communication and hermeneutics. Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 68, 129–43.

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, *104*(1–2), 137–160.

Friston, K., Thornton, C., & Clark, A. (2012). Free-energy minimization and the dark-room problem. *Frontiers in Psychology*, *3*, 130.

Gibson, J. J. (1979). *The ecological approach to visual perception: classic edition*. Psychology Press.

Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge University Press.

Gulledge, A. T., Kampa, B. M., & Stuart, G. J. (2005). Synaptic integration in dendritic trees. *Journal of Neurobiology*, *64*(1), 75–90.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, *7*(1), 1–16.

Hickok, G. (2010). The role of mirror neurons in speech perception and action word semantics. *Language and Cognitive Processes*, 25:6, 749–776.

Hickok, G. (2013). Predictive coding? Yes, but from what source? *The Behavioral and Brain Sciences*, *36*(4), 358.

Hohwy, J. (2013). *The predictive mind*. Oxford: Oxford University Press.

Hohwy, J. (2016), The Self-Evidencing Brain. Noûs, 50: 259–285.

Huygens, C. (1673). *Horologium oscillatorium*. France: Parisiis.

Kiebel, S. J., Daunizeau, J., & Friston, K. J. (2008). A hierarchy of time-scales and the brain. *PLoS Computational Biology*, *4*(11), e1000209.

Kiebel, S. J., & Friston, K. J. (2011). Free energy and dendritic self-organization. *Frontiers in Systems Neuroscience*, *5*, 80.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, *8*(3), 159–166.

Kirchhoff, M. (2017a). Predictive brains and embodied, enactive cognition: an introduction to the special issue. *Synthese*, 1–12.

Kirchhoff, M., Parr, T., Palacios, E., Friston, K., & Kiverstein, J. (2018). The Markov blankets of life: autonomy, active inference and the free energy principle. *Journal of the Royal Society, Interface / the Royal Society*, *15*(138). https://doi.org/10.1098/rsif.2017.0792

Kirchhoff, M.D. (2017b). Predictive processing, perceiving and imagining: Is to perceive to imagine, or something close to it? *Philosophical Studies*, 1–17, doi: https://doi.org/10.1007/s11098-017-0891-8.

Kirchhoff, M.D. (2015). Species of realization and the Free Energy Principle. *The Australasian Journal of Philosophy*, 93(4), 706–723.

Lendvai, B., Stern, E. A., Chen, B., and Svoboda, K. (2000). Experience-dependent plasticity of dendritic spines in the developing rat barrel cortex in vivo. *Nature* 404, 876–881.

Manneville, P. (1995). Dissipative Structures and Weak Turbulence. *Springer Lecture Notes in Physics*, *457*, 257–272.

Maynard Smith, J. (1964). Group selection and kin selection. *Nature*, *201*(4924), 1145–1147.

Mirza, M. B., Adams, R. A., Mathys, C. D., & Friston, K. J. (2016). Scene Construction, Visual Foraging, and Active Inference. *Frontiers in Computational Neuroscience*, *10*, 56.

Naiman, R. J., Johnston, C. A., & Kelley, J. C. (1988). Alteration of North American Streams by Beaver: The structure and dynamics of streams are changing as beaver recolonize their historic habitat. *Bioscience*, *38*(11), 753–762.

Noë, A. (2004). *Action in Perception*. MIT Press.

Odling-Smee, F. J., & Laland, K. N. (2000). Niche Construction and Gene-Culture Coevolution: An Evolutionary Basis for the Human. In T. N. S. Tonneau F. (Ed.), *Perspectives in Ethology* (Vol. Sciences Perspecties in Ethology, 13). Boston, MA: Springer.

Odling-Smee, J., Erwin, D. H., Palkovacs, E. P., Feldman, M. W., & Laland, K. N. (2013). Niche construction theory: a practical guide for ecologists. *The Quarterly Review of Biology*, *88*(1), 4–28.

Odling-Smee, J., Laland, K. N., & Feldman, M. W. (2003). *Niche Construction: The Neglected Process in Evolution*. Princeton University Press.

Palacios, E. R., Razi, A., Parr, T., Kirchhoff, M., & Friston, K. (2017, November 30). *Biological Self-organisation and Markov blankets*. *bioRxiv*. bioRxiv.

Orgel, L. E., & Crick, F. H. C. (1980). Selfish DNA: the ultimate parasite. *Nature, 284(5757)*, 604–607.

Parr, T., & Friston, K. J. (2018). Active inference and the anatomy of oculomotion. *Neuropsychologia*, *111*, 334–343.

Pearl, J. (1988). Probabilistic reasoning in intelligent systems: Networks of plausible inference. San Mateo, CA: Morgan Kaufmann.

Ramstead, M. J. D., Badcock, P. B., & Friston, K. J. (2017). Answering Schrödinger's question: A free-energy formulation. *Physics of Life Reviews*. https://doi.org/10.1016/j.plrev.2017.09.001

Ramstead, M. J. D., Badcock, P. B., & Friston, K. J. (2018). Variational neuroethology: Answering further questions: Reply to comments on "Answering Schrödinger's question: A free-energy formulation." *Physics of Life Reviews*, *24*, 59–66.

Ramstead, M. J. D., Veissière, S. P. L., & Kirmayer, L. J. (2016). Cultural affordances: scaffolding local worlds through shared intentionality and regimes of attention. *Frontiers in Psychology*, *7*, 1090.

Seifert, U. (2012). Stochastic thermodynamics, fluctuation theorems and molecular machines. Rep Prog Phys, 75(12), 126001. doi: https://doi.org/10.1088/0034-4885/75/12/126001

Sengupta, B., Stemmler, M. B., & Friston, K. J. (2013). Information and Efficiency in the Nervous System – A Synthesis. PLoS Computational Biology, 9(7).

Sengupta, B., Tozzi, A., Cooray, G. K., Douglas, P. K., & Friston, K. J. (2016). Towards a Neuronal Gauge Theory. *PLoS Biology*, *14*(3), e1002400.

Sgrò, C. M., Lowe, A. J. and Hoffmann, A. A. (2011), Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications*, 4: 326–337.

Stotz, K. (2017). Why developmental niche construction is not selective niche construction: and why it matters. *Interface Focus*, *7*(5), 20160157.

Thompson, E. (2007). *Mind in life: biology, phenomenology, and the sciences of mind.* Cambridge, MA: Harvard University Press.

Torben-Nielsen, B., & Stiefel, K. M. (2009). Systematic mapping between dendritic function and structure. *Network*, *20*(2), 69–105.

van Dijk, L., Withagen, R., & Bongers, R. M. (2015). Information without content: A Gibsonian reply to enactivists' worries. Cognition, 134, 210–214.

Varela, F. G., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: the organization of living systems, its characterization and a model. *Currents in Modern Biology*, *5*(4), 187–96.

Varela, F. J., Thompson, E., & Rosch, E. (2017). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.

Weiss, L., Brandl, P., & Frynta, D. (2015). Fear reactions to snakes in naïve mouse lemurs and pig-tailed macaques. *Primates*, *56*(3), 279–284.

Yi, S., Wierstra, D., Schaul, T., & Schmidhuber, J. (2009). Stochastic search using the natural gradient. In *Proceedings of the 26th Annual International Conference on Machine Learning – ICML '09* (pp. 1–8). New York, New York, USA: ACM Press.

Action and Power Efficiency in Self-Organization: The Case for Growth Efficiency as a Cellular Objective in *Escherichia coli*

Georgi Yordanov Georgiev, Tommi Aho, Juha Kesseli, Olli Yli-Harja, and Stuart A. Kauffman

1 Introduction

Flux balance analysis (FBA; (Fell and Small 1986; Varma and Palsson 1994)) has been successfully applied to genome-scale models of microorganisms in order to characterize their metabolic capabilities (Price et al. 2004). FBA makes it possible to simulate different growth phenotypes attained under different environmental conditions and genetic modifications. The analysis can be performed without kinetic parameters in biochemical reactions equations, but using only stoichiometric and thermodynamic constraints. Typically, an objective function, representing the true cellular objective, also needs to be determined in flux balance analysis.

FBA is one of the constraint-based modeling methods that are based on the steady-state assumption. The assumption states that the concentrations of

---

G. Y. Georgiev (✉)
Department of Physics, Worcester Polytechnic Institute, Worcester, MA, USA

Department of Physics, Assumption College, Worcester, MA, USA

Department of Physics, Tufts University, Medford, MA, USA
e-mail: ggeorgie@assumption.edu

T. Aho · O. Yli-Harja
Department of Signal Processing, Tampere University of Technology, Tampere, Finland
e-mail: tommi.aho@tut.fi; olli.yli-harja@tut.fi

J. Kesseli
BioMediTech Institute and Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland
e-mail: juha.kesseli@uta.fi

S. A. Kauffman
Department of Signal Processing, Tampere University of Technology, Tampere, Finland

Complex Systems Center, University of Vermont, Burlington, VT, USA

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_8

metabolites not freely exchangeable with the environment are in a steady state. Given a metabolic network of m metabolites and n reactions, the network structure and the stoichiometric coefficients in reactions can be expressed in an $m \times n$ stoichiometric matrix S. Using S and the steady-state assumption, it is possible to form the following set of equations to comprehensively characterize all the feasible metabolic flux distributions:

$$\begin{aligned} &\tfrac{dc}{dt} = Sv = 0 \\ &v_i^{lb} \leq v_i \leq v_i^{ub} \;\; i = 1, \ldots, n \end{aligned} \tag{1}$$

In Eq. (1), c is a vector of concentrations of non-exchangeable metabolites, and v is a vector of reaction rates. The lower and upper bounds of reaction rates are defined by v_i^{lb} and v_i^{lb}, respectively. The bounds for reaction rates can be used to constrain specific reactions to be irreversible and to constrain the substrate uptake rate that usually is an important parameter in constraint-based metabolic modeling. In FBA it is often assumed that microorganisms aim to maximize their growth rate (Feist and Palsson 2010). Therefore, a specific reaction is implemented to describe the generation of new biomass. In FBA the maximum rate of this reaction is determined using linear programming under the constraints of Eq. (1). Currently, the maximal growth rate has established usage as an objective function, but the rationale for cells always pursuing at maximal growth remains debatable (Feist and Palsson 2010). Therefore, the research for other possible objective functions continues active. Other suggested functions include the maximization of ATP yield (Ramakrishna et al. 2001; van Gulik and Heijnen 1995), the minimization of the overall intracellular flux (Blank et al. 2005; Bonarius et al. 1996), the maximization of ATP yield per flux unit (Dauner and Sauer 2001), the maximization of biomass yield per flux unit (Schuetz et al. 2007), the minimization of glucose consumption (Oliveira et al. 2005), the minimization of reaction steps needed to produce biomass (Meléndez-Hevia and Isidoro 1985), the maximization of ATP yield per reaction step (Schuetz et al. 2007), the minimization of redox potential (Knorr et al. 2007), the minimization of ATP producing fluxes (Knorr et al. 2007), and the maximization of ATP producing fluxes (Ebenhoh and Heinrich 2001; Heinrich et al. 1997; Knorr et al. 2007). Similarly to these studies, FBA provides the methodological framework also for our study of growth efficiency.

In the present work, we define the concept of growth efficiency and hypothesize that *Escherichia coli* uses it as the cellular objective. Maximal growth efficiency as the cellular objective would allow bacteria to utilize substrates efficiently to the production of new biomass while producing only little amount of waste, heat, or other side products. In this work we explore the properties of growth efficiency using a genome-wide metabolic model for *Escherichia coli* (Feist et al. 2007) and study whether the growth efficiency could be considered a plausible cellular objective in phenotypic simulations.

In physics, all the laws of motion of particles are obtained from the principle of least action (Goldstein 1980; de Maupertuis 1750). Action is defined as the integral of the kinetic minus the potential energy over time. It has the units of energy times time (Goldstein 1980). This principle states that for a process to occur, the product of time and energy for it is least, as compared to all neighboring paths (Goldstein 1980). A restatement of this principle is that all motions and processes in nature occur with the least expenditure of action or in the most action-efficient way. In the processes of self-organization in complex systems, the principle is for least unit action connects to a principle for most total action (Georgiev 2012, 2016; Georgiev and Georgiev 2002; Georgiev et al. 2012, 2015, 2016a, b). The least action principle does not act in isolation, but is connected in a positive feedback loop to the maximum total action and the size of the system (Georgiev 2012, 2016; Georgiev et al. 2012, 2015, 2016a, b). This is because to do work to structure a system so the processes inside occur in the most action-efficient way, more energy and time are required. Therefore, the total amount of action is connected with increased action efficiency. Action efficiency allows growth and growth allows action efficiency. Action efficiency can be converted to power efficiency, as shown below. In *E. coli* the growth objective is efficiency, but another implicit objective is the growth itself. So the total growth is maximized when the unit action is minimized, i.e., when its action efficiency is maximized. In previous studies of CPUs, as the unit action for one event (computation) decreases, the total amount of action in the CPU increases. The two are correlated as it has been noticed through the Complexity-Size rule (Georgiev 2012, 2016; Georgiev and Georgiev 2002; Georgiev et al. 2012, 2015, 2016a, b). There is no increase of action efficiency without increase of total action. Correlated to that, the total amount of elements, their density, and the flow of events increase proportionally as well. All those increases are correlated and feed on each other. They are interdependent functions of each other, and they can be termed interfunctions. This study shows that the growth efficiency of *E. coli* obeys the action efficiency principle for self-organization and that the action efficiency principle, and the least action principle for physical systems, expresses itself at biological level in the case of growth efficiency of *E. coli* as power efficiency.

2 The Definition and Calculation of Growth Efficiency

We define the growth efficiency η as the growth rate v_{bm} (i.e., biomass production rate) divided by the substrate uptake rate v_s ($\eta = v_{bm} / v_s$). Because v_{bm} is largely determined by v_s, and in the following analysis we specifically focus on the effects of v_s to η, we now define the so-called growth efficiency function as $\eta = \mathrm{H}(v_s)$ and explore its properties. This simplification ignores specific other factors affecting η via v_{bm}, but the sensitivity of η to these factors will also be examined. The key assumption in our approach is that under specific conditions, bacteria actively work to tune the substrate uptake rate such that the growth efficiency η will be maximized. That is, the bacteria aim at substrate uptake rate v_s^* that is optimal in the sense of $v_s^* = {}_{v_s}^{\arg}{}^{\max}\mathrm{H}\,(v_s)$

In order to characterize the properties of growth efficiency and to study its use as a cellular objective function, we apply the constraints of Eq. (1) and set maximal η as the objective in flux balance analysis. The problem is a linear-fractional problem where η, that is the ratio of biomass production rate to the substrate uptake rate, is maximized:

$$\begin{gathered}\max\ \eta \\ s.t.\ \ Sv = 0 \\ v_{irr} \geq 0 \\ v_i^{lb} \leq v_i \leq v_i^{ub}\ \forall i\end{gathered} \tag{2}$$

We used the metabolic model iAF1260 for *Escherichia coli* (Becker et al. 2007) to study growth efficiency. Different substrates in cultivation media were modeled by changing the uptake bounds of the corresponding substrates. Different gene knockouts were modeled by setting the lower and upper bounds of the respective enzymatic reactions to zero. We examined substrate uptake rates v_s between 0 and 50 mmol/h/grams of cell dry weight (mmol/h/gDW). The linear-fractional optimization problem in Eq. (2) was solved by sampling the allowed values $\mathrm{v_S}$ (60 samples in equal distances between 0 and 50), maximizing v_{bm} in each case, and selecting the value for $\mathrm{v_S}$ that maximizes η. The analysis was performed using COBRA Toolbox (Becker et al. 2007), and the linear programming problems were solved using glpk (http://www.gnu.org/software/glpk/).

2.1 Deriving the Growth Efficiency from the Principle of Least Action

A growth efficiency objective function can be justified from broader self-organization and complexity perspective. The rise of complexity in nature includes physical constraints, which must be taken into account. In physics, all motions occur in accordance with the principle of least action (Goldstein 1980; de Maupertuis 1750). All branches of physics, from quantum mechanics to relativity, from classical mechanics to electromagnetism and optics, are derived from it. As mentioned above, action broadly defined is the product of time and energy for a motion. Expressed through power, it is the product of the power and the square of time. For all trajectories in nature, it is true that the motions of objects along them occur with the least amount of physical action as defined above. Any other trajectory has a higher amount of action, and therefore it does not occur. Recently the principle of least action has been expanded to complex systems (Georgiev 2012, 2016; Georgiev and Georgiev 2002; Georgiev et al. 2012, 2015, 2016a, b; Chatterjee 2012, 2013; Annila 2010; Annila and Salthe 2010). According to this principle, all processes in complex systems occur with the least amount of physical action in the given set of constraints, and in flow networks, they do work on the constraints to reduce the constraints to motion and therefore approach the least action state of motion, i.e., the product of time and energy consumed by a process, or the product of power and time squared. Anila expands this to natural selection for least action (Annila 2010; Annila and Salthe 2010). This model has been tested in the case of CPUs (Georgiev 2012, 2016; Georgiev et al. 2012, 2015, 2016a, b). Starting from this most fundamental physical principle, we propose that metabolic chemical reactions evolve to lower their energy barrier and to occur for the shortest time, obeying the principle of least action. In this case the product of the energy barrier and the time for the reaction is the amount of action consumed by it. In this vein of thought, Stuart Kauffman defines self-organization as the evolution of those autocatalytic cycles (Kauffman 1993). In his model, each of the reactions catalyzes the other in a positive feedback loop. If the metabolic reactions occur with the most action efficiency obeying physical laws, then the whole metabolisms must obey the same objective function as a natural self-organizing system, as it is built on them. Therefore, action and power efficiency in general translates to metabolic efficiency for living organisms. Thus, the metabolic efficiency can be connected to the most fundamental principle in physics, the principle of least action. The role of physics and specifically the principle of least action in biological structure and function and top-down control in complex living systems have been illuminated and furthered by Michael Levin (Giovanni and Levin 2016).

When modified for a complex system, the principle of least action is that the variation of the average actions per one is zero in the most organized state (Georgiev 2012, 2016; Georgiev and Georgiev 2002; Georgiev et al. 2012, 2015, 2016a, b). This means that the process occurs in the most action-efficient way.

$$\delta \frac{\sum_{ij}^{nm} I_{ij}}{nm} = 0$$

where I_{ij} is the action of an element, i is to cross an edge, m is per unit time, and $\sum_{ij} I_{ij}$ is the total amount of action in the system per unit time. n is the number of elements.

When the variation is zero, it means that the function has a minimum, which is the least action state.

2.2 A Connection Between Action Efficiency and Power Efficiency

Here we connect action efficiency to power efficiency. Action defined as the product of energy and time can be converted to a product of power and time squared: Action = Energy*time = Power*time^2 = Pt^2. The last term has the same units as the quantity of action, namely, [Joules*seconds] = [Js]. In this case, action efficiency is defined as the ratio of all events occurring in a complex system, i.e., the number of metabolic reactions, which translates in the biomass production, divided by the total amount of action consumed by them,

$$Action\ efficiency = \frac{biomass\ production}{Energy{*}time} = \frac{number\ of\ reactions}{\sum_{ij} I_{ij}}.$$

It can be converted to power efficiency by changing the denominator to the equal of action:

$$Power\ efficiency = \frac{biomass\ production}{Power{*}time{*}time}.$$

If the time is fixed to 1 sec, Action Efficiency per unit time = Power Efficiency.

The flow of events is the biomass production or the number of metabolic reactions.

In our work the total amount of action for all events in a system which is proportional to its total power consumption is also maximized in self-organization proportionally to the action efficiency. When considered per unit time, this becomes a maximum power principle. The power efficiency is connected to the power consumption by a power law function which determines a proportionality between the two at each value. As power efficiency increases, the system can absorb more power, and as the power increases, it helps increase the power efficiency of the system, which is in agreement with data that we obtained for CPUs as a complex system (Georgiev 2012, 2016; Georgiev et al. 2012, 2015, 2016a, b).

For a metabolic cycle, the (nm) term is the number of metabolic reactions occurring per unit time, and $\sum_{ij} I_{ij}$ is the total amount of energy and time spent for those reactions, which is the rate of energy consumption by the metabolic cycle.

When the unit action for one metabolic reaction decreases, the system becomes more efficient, obeying the principle of least action, i.e., it self-organizes, as in the example of autocatalytic cycles:

$$\delta \frac{\sum_{ij}^{nm} I_{ij}}{nm} < 0$$

This leads to a measure for organization, α, inversely proportional to the average number of quanta of action per event (Georgiev 2012, 2016; Georgiev et al. 2012, 2015, 2016a, b):

$$\alpha = \frac{\mathrm{h}\, nm}{\sum_{ij} I_{ij}}.$$

where h is Planck's constant, n is the total number of elements in the system, and m is the number of edge crossings of one element per unit time.

Therefore, nm is the total number of events in the system, metabolic reactions in this case, the flow ϕ of events per unit time in the flow network, $\phi = nm$. $Q = \frac{\sum_{ij} I_{ij}}{h}$ is the total number of quanta in the system in certain interval of time, the sum of all products of time and energy for all metabolic reactions divided by Planck's constant. We can rewrite the expression for α in the following way:

$$\alpha = \frac{\phi}{Q}.$$

Therefore, the average action efficiency per one event in complex systems is their growth objective, derived from the principle of least action in physics. The growth efficiency as a cellular objective for *Escherichia coli* is one example of this general physics principle. In the growth efficiency of *Escherichia coli*, the growth rate is proportional to the number of metabolic reactions, which gives the biomass production rate, and the substrate uptake rate is proportional to the energy and the power consumed by the system for its biomass production. Therefore α, in the above equation, becomes proportional to η. The growth efficiency, which is defined as v_{bm} – the biomass production rate – is proportional to the total number of metabolic reactions, ϕ. It is divided by v_s – the substrate uptake rate –which is proportional to the total action and power consumption, Q. Substituting those in the above expression for α, we obtain $\eta = v_{bm}/v_s$. The analogy between α and η is up to a constant.

3 Properties of Growth Efficiency

3.1 Maximum of Growth Efficiency

The growth efficiency function H(v_s) obtains its maximum at a finite substrate uptake rate v_s. This is illustrated in Fig. 1 which shows the predicted growth rate and growth efficiency for a wild-type *E. coli* strain in glucose minimal media assuming varying glucose uptake rates. While the growth rate increases monotonically as the function of the glucose uptake rate, the growth efficiency has a maximum at $v_s^* = 9.2$ mmol/h/gDW. If the uptake rate is greater than v_s^*, the cell starts to secrete increasing amounts of excess metabolites like acetate.

3.2 Sensitivity of Growth Efficiency to Model Uncertainties

Metabolic network models are based on a well-known and validated information on stoichiometric coefficients in biochemical reactions. However, the models also include specific uncertainties. We examined the robustness of the growth efficiency

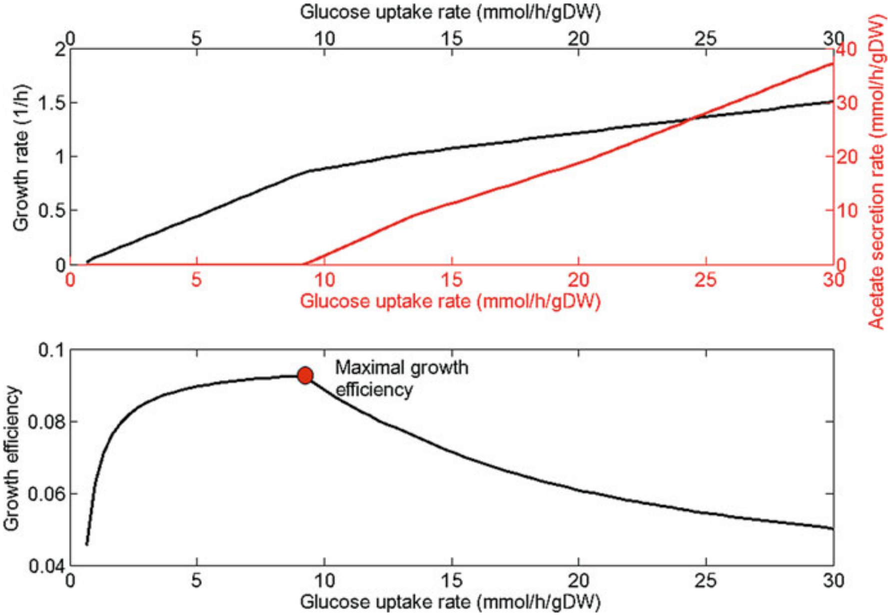


Fig. 1 Maximal growth efficiency is obtained at a finite substrate uptake rate. The upper panel depicts the growth rate and the acetate secretion rate as the function of glucose uptake rate. The growth rate is a monotonically increasing function without a maximum. The lower panel shows the growth efficiency as the function of glucose uptake rate, i.e., the growth efficiency function H(v_s)

function against four model parameters: (1) the maximal oxygen uptake rate, (2) ATP requirement for growth-associated maintenance (GAM), (3) ATP requirement for nongrowth-associated maintenance (NGAM), and (4) the phosphorus-to-oxygen (P/O) ratio that reflects the efficiency of ATP synthesis in the electron transfer chain. These parameters have been identified most critical to the behavior of the iAF1260 model (Feist et al. 2007). We first examined the form of the growth efficiency function while varying the maximal oxygen uptake rate between 0 and 50 mmol/h/gDW (the original value being 18.5 mmol/h/gDW). Second, GAM and NGAM were varied for ±50% of their original values (59.81 and 8.39 mmol/h/gDW, respectively) by constraining the respective reaction rates. Finally, P/O ratios 0.5, 1.0, 1.75, and 2.67 were tested by modifying the stoichiometric coefficients in the electron transfer chain and constraining specific reactions of the electron transfer chain (similarly as described in Feist et al. (2007)). Figure 2 shows that all four parameters affect the growth efficiency function. Oxygen uptake rate has the most drastic effect which shifts v_s^*, i.e., the glucose uptake rate at which the maximum of growth efficiency is obtained. GAM and NGAM have similar effects of shifting v_s^*, but the effect is more moderate. The increased P/O ratio increases the maximum growth efficiency without notable effects to v_s^*.

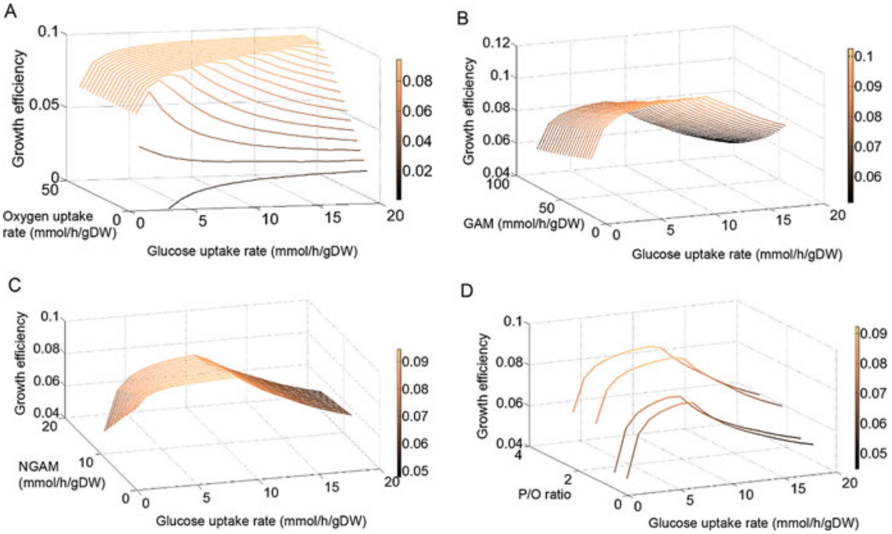


Fig. 2 The robustness of the growth efficiency function against four uncertain model parameters. (**a**) Oxygen uptake rate, (**b**) growth-associated maintenance, (**c**) nongrowth-associated maintenance, (**d**) P/O ratio

3.3 Substrate Uptake Rate Distributions

Because we assume a bacterium to self-regulate the substrate uptake rate to the maximum η at a finite v_s^*, we are able to set the substrate uptake rate unconstrained. Usually in FBA it is crucial to constrain the substrate uptake rate properly. Otherwise, as shown in the upper panel of Fig. 1, the growth rate simply increases monotonically with increasing substrate input rate. In the following analysis, we simulated 10 different cultivations and 1261 genetic conditions to study the distribution of v_s^*. Figure 3 shows the results for phenotypes that are predicted to be viable (i.e., the growth rate is greater than 0.1 h^{-1}). For them, the substrate uptake rate v_s^* always remains at a finite range. Typical values for v_s^* are from 5 to 20 mmol/h/gDW. The largest v_s^* is obtained under the knockouts of components of ATP synthase, in particular in pyruvate cultivation. The blockage of ATP synthase requires that the needed ATP is synthesized by other mechanisms, such as glycolysis and the citric acid cycle, which requires a large substrate uptake rate.

3.4 The Relation of Growth Efficiency and Overflow Metabolism

In situations where the maximal growth is achieved, a bacterium may not be able to transform all the substrate efficiently to new biomass, but an increasing amount

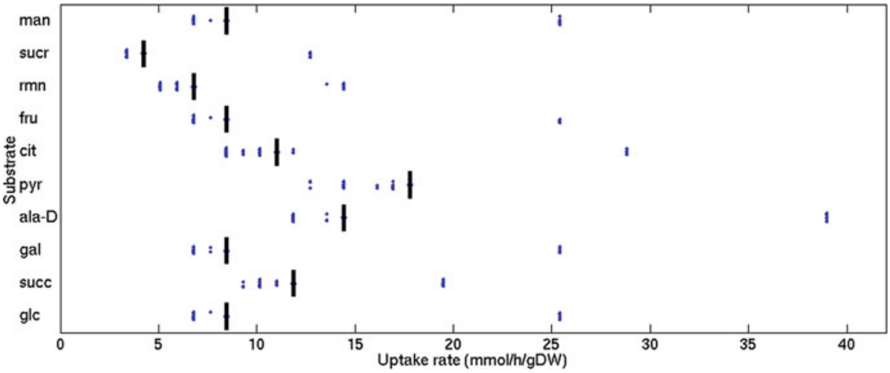


Fig. 3 Distributions of substrate uptake rates under the maximal growth efficiency criterion. The uptake rate is calculated using both criteria for 1261 gene knockout strains in 10 carbon sources (glc, glucose; succ, succinate; gal, galactose; ala-D, D-alanine; pyr, pyruvate; cit, citrate; fru, fructose; rmn, rhamnose; sucr, sucrose; man, mannose). All the distributions are presented as boxplots. The distributions are so concentrated that their data points without outliers appear together as a black bar. Dots represent outliers. Eight data points representing the largest substrate uptake rate (224 mmol/h/gDW for deletions of ATP synthase components in pyruvate cultivation) are not shown

of material is directed to waste. This phenomenon of overflow metabolism has been extensively studied as it is detrimental in industrial applications. In the case of *E. coli*, overflow metabolism directs valuable carbon to acetate production instead of biomass generation. This inhibits growth and it may also disturb product synthesis (Valgepea et al. 2010).

We studied the relationship between the maximal growth efficiency and overflow metabolism by simulating all single-gene knockouts in the iAF1260 model under varying carbon sources. We found that usually the substrate uptake rate at the maximal growth efficiency (v_s^*) equals to the substrate uptake rate at the start of overflow metabolism (i.e., the start of acetate production). There are few exceptions to this rule, for example, when the knockout is directed to specific genes of ATP synthase, pyruvate dehydrogenase, or succinate dehydrogenase. Thus, we reason that the maximal growth efficiency is a concept of its own, and it cannot be directly interpreted as the substrate uptake rate threshold above which overflow metabolism starts.

The use of the maximal growth criterion in growth phenotype simulation may easily produce estimates that are suboptimal in growth efficiency and likely to express overflow metabolism. In order to illustrate the suboptimality under the maximal growth criterion, we calculated the loss of growth efficiency using the abovementioned set of 1261 genetic and 10 environmental conditions. In simulations with the maximal growth criterion, the maximal substrate uptake rate was constrained to 10 mmol/h/gDW. The relative loss in growth efficiency was determined as the growth efficiency under the maximal growth criterion divided by the maximal achievable growth efficiency. Figure 4 summarizes the calculated loss

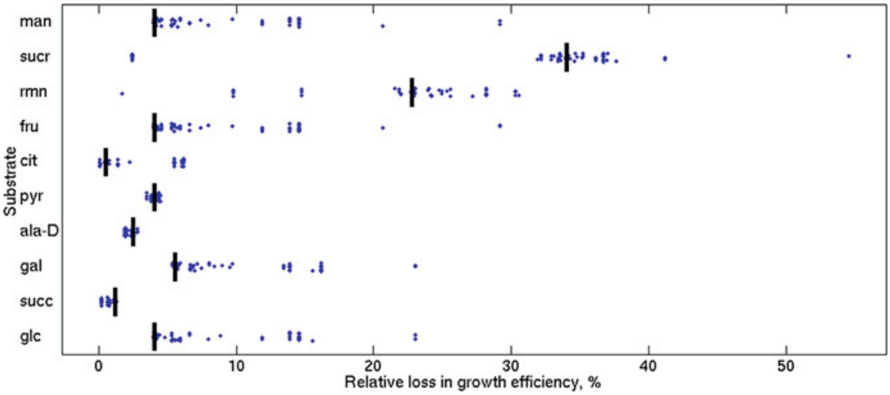


Fig. 4 Suboptimality of growth efficiency under the maximal growth criterion. The relative loss in growth efficiency under the maximal growth rate criterion is depicted for 1261 gene knockout strains in 10 cultivation conditions. The bar represents the median of the 1261 mutants. Almost all the mutants are very close to the median, so they cluster under the bar. The distribution is right skewed, and the points are at least 2.7 S.D. above or occasionally below the median. The median in almost all culture conditions shows a loss of growth efficiency under the maximal growth rate criterion

ratios. In the figure, the loss of growth efficiency at maximal growth demonstrates that maximal growth wastes input substrate energy.

4 Maximal Growth Efficiency as a Cellular Objective

We examined whether the maximal growth efficiency is a plausible cellular objective for *E. coli* cultivated in a small-scale batch process. Therefore, we used the metabolic model iAF1260 to predict the cellular growth rate assuming *E. coli* maximizes the growth efficiency. The growth predictions were compared to two experimental data sets as follows.

First, we predicted the viability for mutant strains carrying single-gene deletions. The predictions were produced for 1117 mutant strains cultivated under glucose minimal media (Baba et al. 2006). Each mutant strain in the data set has been experimentally determined to be either viable or inviable. For 982 viable mutants, the viability was correctly predicted (i.e., true positive rate was 97%), and for 76 inviable mutants, the inviability was correctly predicted (true negative rate was 72%). The results are identical with the prediction results obtained using the maximal growth criterion.

Second, we predicted the growth rate for 5096 growth conditions, consisting of 91 single-gene knockout strains cultivated under 56 different media conditions. The optical density (OD) of *E. coli* grown in these conditions has been measured in a high-throughput experimental screen using the Biolog platform

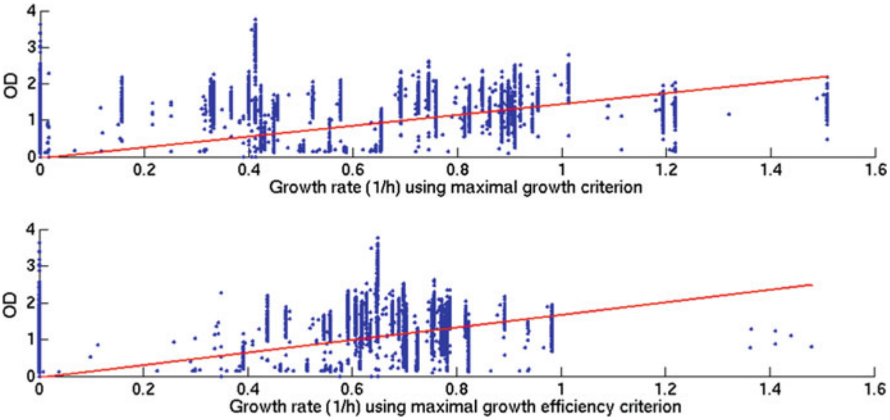


Fig. 5 The correspondence of predicted growth rates and experimentally observed optical density. OD values are presented as the function of the predicted growth rates. The upper and the lower panels present the predictions under the maximal growth criterion and the maximal growth efficiency criterion, respectively. A linear model is fitted to both data. In simulations with the maximal growth criterion, the maximal substrate uptake rate was constrained to 10 mmol/h/gDW

(http://www.biolog.com), and the data has been set available through the ASAP database (Glasner et al. 2003). Figure 5 shows the growth predictions for each of the growth conditions versus the corresponding OD value. Under the maximal growth criterion, the Spearman correlation between the predicted growth rates and the experimental OD values was 0.19, while under the maximal growth efficiency criterion, the Spearman correlation was 0.24. We also fitted linear models to both data in order to further compare the phenotype prediction performance of the two criteria. Under the maximal growth criterion, the linear model had residual standard error of 0.83 and the adjusted R^2 was 0.59. Under the maximal growth efficiency criterion, the linear model had residual standard error of 0.76 and the adjusted R^2 was 0.65. The better fit in the case of the maximal growth efficiency was confirmed using Akaike's information criterion (12,537.83 for maximal growth criterion versus 11,639.86 for maximal growth efficiency criterion).

5 Conclusions

The identification of cellular survival strategies and their simulation by realistic objective functions have fundamental importance on phenotype prediction in metabolic analysis. It is probable that there is no single survival strategy that is optimal in all situations but the strategy is likely to depend on growth conditions of a microorganism (Feist and Palsson 2010). Feist and Palsson discuss three qualitatively different environments: nutritionally rich, nutritionally scarce, and elementally limited (Feist and Palsson 2010). Nutritionally rich laboratory-like conditions are probably very rare in the nature, and thus, maximal growth is probably an unrealistic objective function in most of the situations. In a study by Schuetz et al. (2007), it was found that under nutrient scarcity in continuous cultivations, the best prediction accuracy was achieved using linear maximization of ATP or biomass yields. On the other hand, in unlimited growth on glucose in oxygen or nitrogen respiring batch cultures, the best prediction accuracy was achieved by nonlinear maximization of the ATP yield per flux unit.

In this work we introduced a concept called growth efficiency and characterized its properties. The study was performed using the metabolic model iAF1260 for *Escherichia coli*. As a result we found that the growth efficiency function has its maximum within a finite substrate uptake rate. According to our predictions, the substrate uptake rate at which the maximal growth efficiency is obtained (v_s^*) varies typically from 5 to 20 mmol/h/gDW. Our simulations with several different cultivation media and a set of single-gene knockouts demonstrated that the optimal rate v_s^* depends on the cultivation and genetic conditions. For example, with sucrose the median uptake rate of v_s^* was 4.5 mmol/h/gDW, while with pyruvate the median rate was 17 mmol/h/gDW. We also found that the growth efficiency function is affected by specific parameters that usually remain unsure in metabolic network models. In particular, oxygen uptake and ATP requirement for growth-associated maintenance affect v_s^*, and increasing P/O ratio in electron transfer chain increases growth efficiency while maintaining the form of the growth efficiency function.

A straightforward application of growth efficiency is to use it as an optimization criterion (i.e., objective function) for predictions of cellular growth. We explored this possibility and validated our computational predictions using two sets of experimental data. We found that maximal growth efficiency can be considered a feasible optimization criterion in metabolic modeling. The criterion predicted the given experimental data slightly better than the commonly applied maximal growth rate criterion.

In this study we used data from batch cultivations to validate the feasibility of growth efficiency as an objective function. However, based on the work by Schuetz et al. (2007), we hypothesize that the growth efficiency criterion could perform better in situations where cells are under nutrient scarcity, i.e., they are cultivated in nutrient-limiting chemostats. Such chemostat data was not available in this study, and the hypothesis should be validated in a future study.

Considering maximal growth efficiency as a cellular objective suggests that cells can save nutrients in the benefit of other cells or to be used to themselves at a later moment. This raises the question about the mechanisms, e.g., quorum sensing, which bacteria growing in colonies may use to tune their growth rate in each growth situation. As a further point, we note that if it proves true that bacterial cells maximize the growth efficiency per unit food or energy uptake, this picks out an optimal rate of energy utilization, hence an optimal displacement from chemical equilibrium for nonequilibrium living cells. We note that we lack the theory of an optimal displacement from equilibrium for living, nonequilibrium, cells. Jacques Monod, in Chance and Necessity, notes that optimally growing bacteria give off little heat (Monod 1971). This may be consonant with maximal growth efficiency, so that the maximal amount of energy coming into cells goes into biomass production and minimizes waste heat.

We have shown that this growth efficiency objective function can be derived from a fundamental physics principle, the principle of least action, from which all equations of motion in each branch of physics are derived and describes all processes that occur naturally. Expanding this principle for a complex self-organizing system, it predicts that the processes in that system will occur with least expenditure of the product of energy and time per one event, or the product of power and time squared, such as a metabolic reaction in the case of *E. coli*. Therefore, using an objective function from complexity, for increase of action and power efficiency in self-organization, we derive the expression for growth efficiency, used to describe the metabolic processes of *E. coli*. In this way we justify it from first principles in physics using the new objective function defined in this paper, and we find a case where this principle yields fruitful results. This connection between physics and biology illuminates the common principles behind all processes in nature and gives us a tool to explain more of the phenomena in biology using fundamental first principles. It also expands the validity, applicability, and importance of those physics principles.

Acknowledgments This work was supported by the Academy of Finland (Finnish Programme for Centres of Excellence in Research 2006–2011) and the FiDiPro programme of Finnish Funding Agency for Technology and Innovation. GG thanks Assumption College for a Faculty Development Grant and financial support from the Department of Natural Sciences at Assumption College.

References

T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K.A. Datsenko, M. Tomita, B.L. Wanner, H. Mori, Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection, Mol. Syst. Biol. 2 (2006) 2006.0008.

S.A. Becker, A.M. Feist, M.L. Mo, G. Hannum, B.Ø. Palsson, M.J. Herrgard, Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox, Nat. Protocols 2 (2007) 727–738.

L.M. Blank, L. Kuepfer, U. Sauer, Large-scale ^{13}C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast, Genome Biol. 6 (2005) R49.

H.P.J. Bonarius, V. Hatzimanikatis, K.P.H. Meesters, C.D. de Gooijer, G. Schmid, J. Tramper, Metabolic flux analysis of hybridoma cells in different culture media using mass balances, Biotechnol. Bioeng. 50 (1996) 299–318.

M. Dauner, U. Sauer, Stoichiometric growth model for riboflavin-producing *Bacillus subtilis*, Biotechnol. Bioeng. 76 (2001) 132–143.

O. Ebenhoh, R. Heinrich, Evolutionary optimization of metabolic pathways. Theoretical reconstruction of the stoichiometry of ATP and NADH producing systems, Bull. Math. Biol. 63 (2001) 21–55.

A.M. Feist, C.S. Henry, J.L. Reed, M. Krummenacker, A.R. Joyce, P.D. Karp, L.J. Broadbelt, V. Hatzimanikatis, B.Ø. Palsson, A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information, Mol. Syst. Biol. 3 (2007) 121.

A.M. Feist, B.Ø. Palsson, The biomass objective function, Curr. Opin. Microbiol., 13 (2010) 344–349.

D.A. Fell, J.R. Small, Fat synthesis in adipose tissue. An examination of stoichiometric constraints, Biochem J 238 (1986) 781–786.

J.D. Glasner, P. Liss, G. Plunkett 3rd, A. Darling, T. Prasad, M. Rusch, A. Byrnes, M. Gilson, B. Biehl, F.R. Blattner, N.T. Perna, ASAP, a systematic annotation package for community analysis of genomes, Nucleic Acids Res. 31 (2003) 147–151.

R. Heinrich, F. Montero, E. Klipp, T.G. Waddell, E. Melendez-Hevia, Theoretical approaches to the evolutionary optimization of glycolysis: thermodynamic and kinetic constraints, Eur. J. Biochem. 243 (1997) 191–201.

A.L. Knorr, R. Jain, R. Srivastava, Bayesian-based selection of metabolic objective functions, Bioinformatics 23 (2007) 351–357.

E. Meléndez-Hevia, A. Isidoro, The game of the pentose phosphate cycle, J. Theor. Biol. 117 (1985) 251–263.

J. Monod, Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology, first ed., Knopf, New York, 1971.

A.P. Oliveira, J. Nielsen, J. Forster, Modeling *Lactococcus lactis* using a genome-scale flux model. BMC Microbiol 5 (2005) 39.

N.D. Price, J.L. Reed, B.Ø. Palsson, Genome-scale models of microbial cells: evaluating the consequences of constraints, Nature Rev. Microbiol. 2 (2004) 886–897.

R. Ramakrishna, J.S. Edwards, A. McCulloch, B.Ø. Palsson, Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints, Am. J. Physiol. Regul. Integr. Comp. Physiol. 280 (2001) R695–R704.

R. Schuetz, L. Kuepfer, U. Sauer, Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*, Mol. Syst. Biol. 3 (2007) 119.

W.M. van Gulik, J.J. Heijnen, A metabolic network stoichiometry analysis of microbial growth and production formation, Biotech. Bioeng. 48 (1995) 681–698.

K. Valgepea, K. Adamberg, R. Nahku, P.-J. Lahtvee, L. Arike, R. Vilu, Systems biology approach reveals that overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase, BMC Syst. Biol. 4 (2010) 166.

A. Varma, B.Ø. Palsson, Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110, Appl. Environ. Microbiol. 60 (1994) 3724–3731.

H. Goldstein, *Classical Mechanics*, (2nd ed., Addison Wesley, 1980).

Pierre de Maupertuis, *Essai de cosmologie*, (1750).

Georgiev, G.Y., and Georgiev, I., 2002. The Least Action and the Metric of an Organized System. Open Systems & Information Dynamics 9: 371–380.

Georgiev, G.Y., Daly, M., Gombos, E., Vinod, A., and Hoonjan, G., 2012. Increase of Organization in Complex Systems. World Academy of Science, Engineering and Technology 71. preprint arXiv:1301.6288.

Georgiev, G.Y., Quantitative Measure, Mechanism and Attractor for Self-Organization in Networked Complex Systems. Self-Organizing Systems LNCS 7166: 90–95.2012.

Georgiev G.Y., Henry K., Bates T., Gombos E., Casey A., Lee H., Daly M., and Vinod A., "Mechanism of organization increase in complex systems", Complexity, **21**(2), 18–28, DOI: https://doi.org/10.1002/cplx.21574 7/25 (2015).

Georgi Yordanov Georgiev, Atanu Chatterjee "The road to a measurable quantitative understanding of self-organization and evolution" Ch. 15. In *Evolution and Transitions in Complexity*, Eds. Gerard Jagers op Akkerhuis, Springer International Publishing, (2016). p. 223–230.

Georgi Yordanov Georgiev, Erin Gombos, Timothy Bates, Kaitlin Henry, Alexander Casey, Michael Daly "Free Energy Rate Density and Self-organization in Complex Systems" Ch. 27. p. 321–327. In *Springer Proceedings in Complexity* Proceedings of ECCS 2014: European Conference on Complex Systems, Eds. S. De Pellegrini, G. Caldarelli, E. Merelli, Springer International Publishing, ISBN 3319292269, (2016a)

Georgiev G.Y., Chatterjee A., Iannacchione G.S. “Exponential Self-Organization and Moore’s Law: Measures and Mechanisms” Complexity, 2016b. Article ID 8170632

Chatterjee, A., 2012. Action, an Extensive Property of Self–Organizing Systems. International Journal of Basic and Applied Sciences 1(4): 584–593.

Chatterjee, A., 2013. Principle of least action and convergence of systems towards state of closure. International Journal of Physical Research 1(1): 21–27.

Annila, A. and Salthe, S., 2010. Physical foundations of evolutionary theory. Journal of Non-Equilibrium Thermodynamics 353: 301–321.

Annila, A., 2010. All in Action. Entropy. 1211: 2333–2358.

Kauffman S. A., “The origins of order: Self organization and selection in evolution” Oxford University press (1993)

Giovanni P. and Levin M. “Top-down models in biology: explanation and control of complex living systems above the molecular level.” *Journal of The Royal Society Interface* 13.124 (2016): 20160555.

On Relation Between Swarm and Evolutionary Dynamics and Complex Networks

Ivan Zelinka and Roman Šenkeřík

1 Introduction

In this chapter, we merged two completely different areas of research: (complex) networks and evolutionary computation. Interactions in a swarm and evolutionary algorithms can be considered like user interactions in social networks or just people in society. It has been observed that networks generated by evolutionary dynamics show properties of complex networks in certain time frames and conditions (Skanderova et al. 2016).

Evolutionary computation is a sub-discipline of computer science belonging to the bioinspired computing area. Since the end of the Second World War, the main ideas of evolutionary computation have been published (Turing 1969) and widely introduced to the scientific community (Holland 1975). Hence, the golden era of evolutionary techniques began, when Genetic Algorithms (GA) by J. Holland (1975), Evolutionary Strategies (ES) by Schwefel (1974) and Rechenberg (1971), and Evolutionary Programming (EP) by Fogel (1998) were introduced. All these designs were favored by the forthcoming of more powerful and more easily programmable computers so that for the first time interesting problems could be tackled and evolutionary computation started to compete with and became a serious alternative to other optimization methods. Later, more robust and effective algorithms like Differential Evolution (DE) (Qin et al. 2009; Mallipeddi et al. 2011; Das et al. 2016), Particle Swarm Optimization (PSO) (Jabeen et al. 2009;

I. Zelinka (✉)
Faculty of Electrical Engineering and Computer Science, Technical University of Ostrava, Ostrava-Poruba, Czech Republic
e-mail: ivan.zelinka@vsb.cz

R. Šenkeřík
Faculty of Applied Informatics, Tomas Bata University in Zlin, Zlin, Czech Republic
e-mail: senkerik@fai.utb.cz

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_9

Engelbrecht 2010), Self-Organizing Migrating Algorithm (SOMA) (Zelinka 2016; Zelinka 2004), Artificial Bee Colony (Karaboga and Basturk 2007), or Firefly Algorithm (FA) (Yang 2010; Fister et al. 2013).

Currently, the utilization of complex networks as a visualization tool for the analysis of population dynamics for evolutionary and swarm-based algorithms is becoming an interesting open research task. The population is visualized as an evolving complex network that exhibits nontrivial features – e.g., degree distribution, clustering, and centralities and in between. These features offer a clear description of the population under evaluation and can be utilized for adaptive population as well as parameter control during the metaheuristic run. The initial studies (Zelinka et al. 2014; Davendra et al. 2014a, b) describing the possibilities of transforming population dynamics into complex networks were followed by the successful adaptation and control of the metaheuristic algorithm during the run through the given complex networks' frameworks (Skanderova and Fabian 2015; Metlicka and Davendra 2015; Gajdos and Kromer 2015).

This research represents the hybridization of complex network frameworks using the DE, PSO, and FA.

Currently, all aforementioned algorithms are known as powerful metaheuristic tools for many difficult and complex optimization problems.

The organization of this paper is as follows: First, the motivation and the concept of evolutionary- and swarm-based algorithms with a complex network are briefly described, followed by the simple experiment's design. This is followed by graphical visualizations and the conclusions afterward.

2 Motivation

This paper represents a comprehensive overview and continuation of the previous successful initial experiments. It also extends the preliminary research (Janostik et al. 2016a) focused on capturing the inner dynamics of swarm algorithms in sufficient detail and in a network of appropriate size for further processing. The motivation for the research presented herein can be summarized as follows:

- To show the different approaches in building complex networks in order to capture the dynamics either of evolutionary- or swarm-based algorithms
- To investigate the time development of the influence of either individual selections inside a DE or communication inside a swarm transferred into the complex network
- To briefly discuss the possible utilization of complex network attributes, e.g., adjacency graphs, centralities, clustering, etc., for adaptive population and parameter control during the metaheuristic run.

3 Complex Networks

A complex network is a graph which has unique properties – usually in the real-world graph domain. A complex network contains features which are unique to the assigned problem. These features are important markers for population used in evolutionary/swarm-based algorithms (Davendra et al. 2014a). The following features are important for a quick analysis of the network thus created.

3.1 Degree Centrality

Degree Centrality is defined as the number of edges connected to a specific node. Degree Centrality is an important distribution hub in the network since it connects – and thereby distributes – most of the information flowing through the network. Together with the hybridization of metaheuristic and complex network analysis, this is one of the most important features under consideration. Using Degree Centrality, one can actually analyze if stagnation or premature convergence is occurring within the population. By analyzing the graphs, it can be seen that the multiple nodes are increasing (distinguished by their size), thereby emphasizing their prominence in the population and their effect in generating better individuals.

3.2 Clustering Coefficient

The average clustering coefficient for the entire network is calculated from every single local clustering coefficient for each node. The clustering coefficient of a node shows how concentrated the neighborhood of that node is. Mathematically, it is defined as the ratio of the number of actual edges between neighbors to the number of potential edges between neighbors. For their utilization in computational intelligence, it is also very important to analyze the distribution of a clustering coefficient within an entire network, since we can assume that it can show the population diversity, its compactness, or tendency to form heterogeneous subgroups (subpopulations).

4 Metaheuristic Methods

This section contains the background of the metaheuristic algorithms DE, PSO, and FA, which were used, as well as the main principles of capturing their dynamics in an evolving complex network.

4.1 Differential Evolution

DE is a population-based optimization method that works on real-number-coded individuals (Qin et al. 2009). DE is quite robust, fast, and effective, with global optimization ability. There are essentially five inputs to the heuristic (Qin et al. 2009). D is the size of the problem, $G_{\max}$ is the maximum number of generations, NP is the total number of solutions, F is the scaling factor of the solution, and CR is the factor for crossover. F and CR together make the internal tuning parameters for the heuristic.

The initialization of the heuristic is the following: each solution $x_{i,j,G=0}$ is created randomly between the two bounds $x^{(\mathrm{lo})}$ and $x^{(\mathrm{hi})}$. The parameter j represents the index to the values within the solution, and parameter i indexes the solutions within the population. So, to illustrate, $x_{4,2,0}$ represents the fourth value of the second solution at the initial generation. After initialization, the population is subjected to repeated iterations.

Within each iteration and for particular individual (solution), three random numbers r_1, r_2, and r_3 are selected, unique to each other and to the current indexed solution i in the population. Two solutions, $x_{j,r1,G}$ and $x_{j,r2,G}$, are selected through the index r_1 and r_2 and their values subtracted. This value is then multiplied by F, the predefined scaling factor. This is added to the value indexed by r_3.

However, this solution is not arbitrarily accepted in the solution. A new random number is generated, and if this random number is less than the value of CR, then the new value replaces the old value in the current solution. The fitness of the resulting solution, referred to as a perturbed (or trial) vector $u_{j,i,G}$, is then compared with the fitness of $x_{j,i,G}$. If the fitness of $u_{j,i,G}$ is better than the fitness of $x_{j,i,G}$, then $x_{j,i,G}$ is replaced with $u_{j,i,G}$; otherwise, $x_{j,i,G}$ remains in the population as $x_{j,i,G+1}$. Hence, the competition is only between the new *child* solution and its *parent* solution. This strategy is denoted as DE/Rand/1/bin. Trial vector for this strategy is given in (1).

$$u_{i,G+1} = x_{r1,G} + F \cdot \left(x_{r2,G} - x_{r3,G}\right) \tag{1}$$

4.2 PSO Algorithm

Original PSO algorithms take their inspiration from behavior of fish and birds (Jabeen et al. 2009). The knowledge of the global best-found solution (typically denoted as *gBest*) is shared among the particles in the swarm. Furthermore, each particle has the knowledge of its own (personal) best-found solution (designated *pBest*). The last important part of the algorithm is the velocity of each particle, which is taken into account during the calculation of the particle's movement. The new position of each particle is then given by (2), where ${x_i}^{t+1}$ is the new particle position, ${x_i}^{t}$ refers to the current particle position, and ${v_i}^{t+1}$ is the new velocity of the particle.

$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{2}$$

To calculate the new velocity, the distance from *pBest* and *gBest* is taken into account along with its current velocity (3).

$$v_{ij}^{t+1} = v_{ij}^t + c_1 \cdot Rand \cdot \left(pBest_{ij} - x_{ij}^t\right) + c_2 \cdot Rand \cdot \left(gBest_j - x_{ij}^t\right) \tag{3}$$

where:

$v_{ij}{}^{t+1}$ – New velocity of the *i*th particle in iteration $t + 1$ (component *j* of the dimension *D*)
$v_{ij}{}^{t}$ – Current velocity of the *i*th particle in iteration *t* (component *j* of the dimension *D*)
$c_1, c_2 = 2$ – Acceleration constants
$pBest_{ij}$ – Local (personal) best solution found by the *i*th particle (component *j* of the dimension *D*)
$gBest_j$ – Best solution found in a population (component *j* of the dimension *D*)
$x_{ij}{}^{t}$ – Current position of the *i*th particle (component *j* of the dimension *D*) in iteration *t*
Rand – Pseudorandom number, interval (0, 1)

4.3 *Firefly Algorithm*

FA was first introduced in 2008 by X. S. Yang (Yang 2010; Fister et al. 2013). This nature-based algorithm tries to simulate the mating behavior of fireflies at night. Every firefly emits flashing light to lure appropriate mating partner. For the formulation of the FA, the flashing light is associated with the objective function value that is optimized. For simplicity, the three following rules are used:

- All fireflies are sexless (each firefly can attract, or be attracted by, any of the remaining ones).
- The attractiveness of fireflies is proportional to their brightness. Thus, the less bright firefly will move toward the brighter one. The brightness decreases with the distance between fireflies. If there is no brighter firefly, the particular one will move randomly.
- The firefly brightness is based on the objective function value.

The brightness *I* (4) of firefly consists of three factors: the objective function value, the distance between two compared fireflies, and absorption of media in which the fireflies are.

$$I(r) = \frac{I_0}{1 + \lambda r^m} \tag{4}$$

where $I(r)$ is the brightness of particular firefly based on distance r. I_0 is the initial brightness or in this case the objective function value. The λ is the light absorption coefficient of media and the $m \geq 1$.

The attractiveness $\beta(r)$ (5) is proportional to brightness as mentioned before and then very similar to previous equation.

$$\beta(r) = \frac{\beta_0}{1 + \lambda r^m} \tag{5}$$

where $\beta(r)$ is the attractiveness between fireflies based on their distance r. The β_0 is the initial attractiveness. The r is defined as a Euclidean distance between two fireflies as (6).

$$r_{ij} = \sqrt{\sum_{k=1}^{d} \left(x_{i,k} - x_{j,k}\right)^2} \tag{6}$$

where r_{ij} is the Euclidean distance between fireflies i and j. The d is dimension size of the optimized problem. The movement of a firefly i is then defined as (7).

$$x_i = x_i + \beta(r) \cdot \left(x_j - x_i\right) + \alpha \cdot \mathrm{sign} \tag{7}$$

where x_i and x_j are fireflies i and j. The variable α serves as a randomization parameter. The sign provides a random sign or direction.

4.4 Evolutionary and Swarm Algorithms with a Complex Network Framework

In this research, the complex network approach is utilized to show the linkage between different individuals in the population. Each individual in the population can be taken as a node in the complex network graph, where its links specify the successful exchange of information in the population.

Since the internal dynamics and principles are different for evolutionary (DE) and swarm-based algorithms (PSO and FA), two different approaches for capturing the population dynamics have been developed and tested.

In the case of the DE algorithm, an *adjacency graph* was used. In each generation, the node is only active for the successful transfer of information, i.e., if the individual is successful in generating a new better individual which is accepted for the next generation of the population. If the trial vector created from three randomly selected individuals (*DE/Rand/1/Bin*) is better than the active individual, one establishes the connections between the newly created individual and the three sources; otherwise, no connections are recorded in the adjacency matrix.

Although the Firefly Algorithm is a swarm type, the situation here is very similar to the evolutionary algorithms. In order to create network, we decided to visualize every firefly as a node. Connection between nodes is plotted for every successful interaction between fireflies. Successful interaction is defined as such interaction where one of the individuals gets improved. In the case of Firefly Algorithm, it is when firefly flies toward another and improves own brightness.

For the PSO algorithm, the main interest is in the communications that lead to population quality improvement. Therefore, only communication leading to improvement of the particles personal best (*pBest*) was tracked. The link was created between the particle that was improved and the particle that triggered the current *gBest*'s update. This approach creates a complex network with clusters – and can be used for particle performance evaluations. Of course, it is also possible to build an adjacency graph (see Sect. 7).

5 Experiment Design

A simple Schwefel's test function was used in this experimental research for the purpose of the generation of a complex network.

Experiments were performed in the *C language* environment; the data from the DE algorithm was analyzed and visualized using *Cytoscape* software, while data from the swarm algorithms was analyzed in the *Wolfram Mathematica SW* suite.

Within the ambits of this research, only one type of experiment was performed. It utilizes the maximum number of generations fixed at 100 with a population size of $NP = 50$ for DE and 30 for swarm algorithms, due to the better clarity of visualizations and differences between evolutionary and swarm systems. Two DE control parameters for mutation and crossover were set as $F = 0.5$ and $CR = 0.8$. Two acceleration constants for PSO were set as $c_1, c_2 = 2$.

Since only one run of DE or PSO algorithms were executed for this particular case study, no statistical results related to the cost function values and no comparisons are given here, since it is not possible to compare metaheuristic algorithms only from one run.

6 Visualizations for DE

The visualizations of complex networks are depicted in Figs. 1, 2 and 3 – containing adjacency graphs for this selected case study.

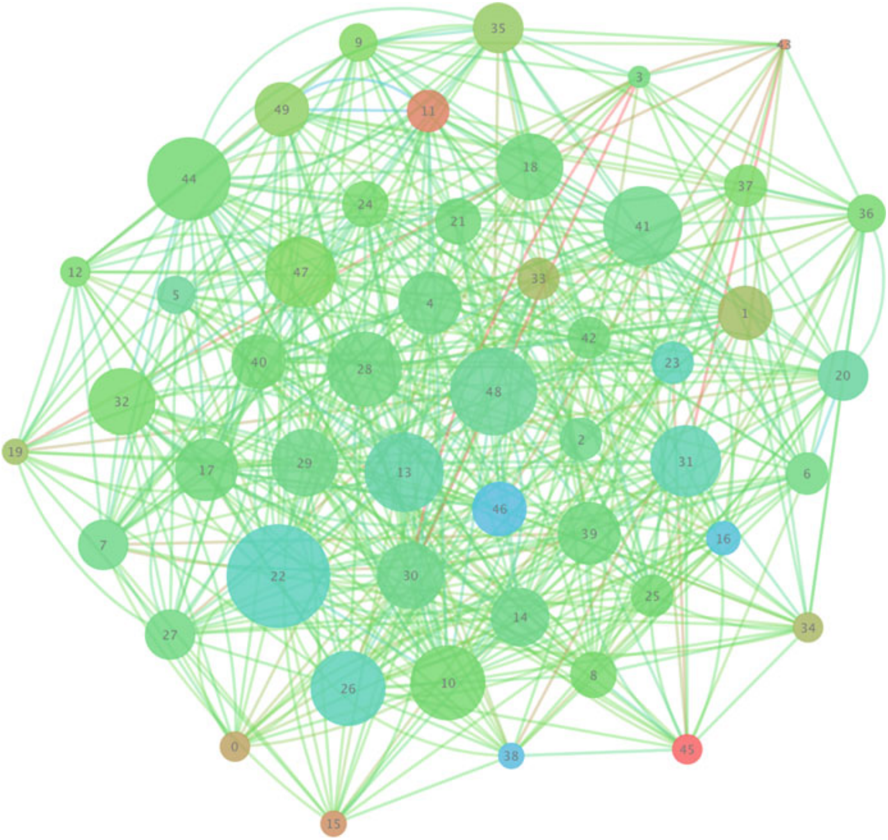


Fig. 1 Complex network representation for DE dynamics – case study 1: the first 10 iterations

The *Degree Centrality* value is highlighted by the size of the node, and the coloring of the node is related to the *clustering coefficient* distribution (light-colored lower values ranging up to the red colors – higher values). Analysis of CN from DE algorithm can be found in Skanderova and Fabian (2015), Skanderova et al. (2017), Viktorin et al. (2016, 2017), and Senkerik et al. (2016).

7 Visualizations for PSO

The complex network for all iterations of the PSO algorithm that was created is depicted in Fig. 3 (Pluhacek et al. 2016a). Nodes of a similar color represent particles with the same ID, throughout different iterations. All links are from a

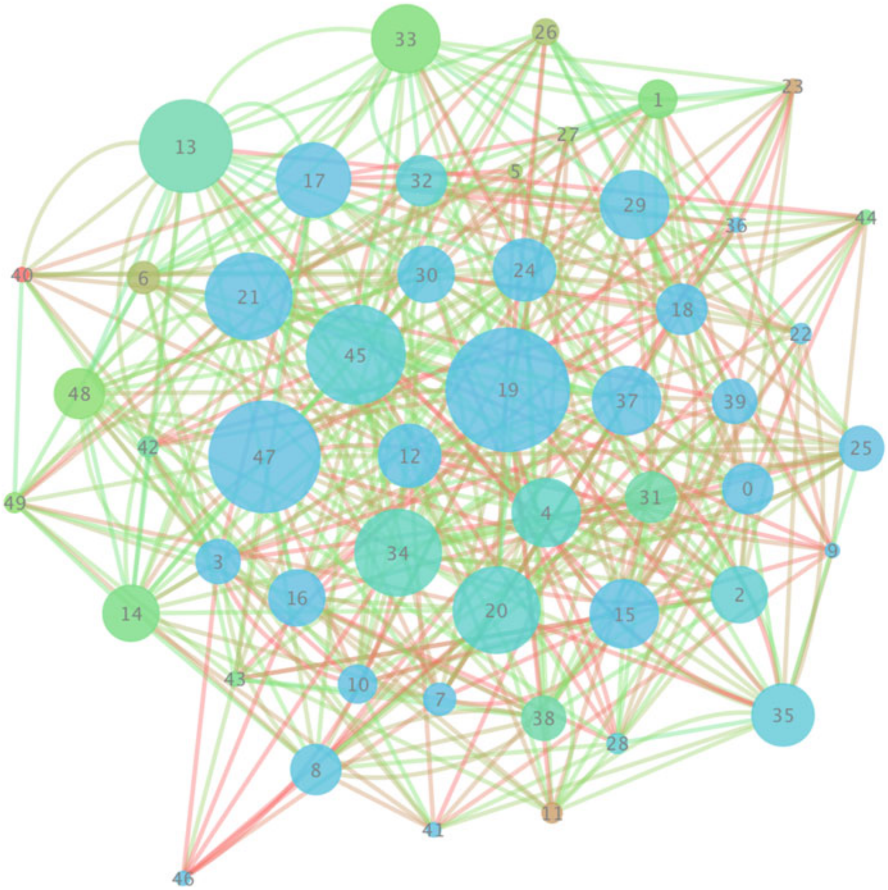


Fig. 2 Complex network representation for DE dynamics – case study 2: the last 10 iterations

particle that triggered the *gBest* update to a particle that has improved – based on that *gBest*. The nodes' code numbers represent a particle ID and its current iteration. This way, it is possible to precisely track the development of the network and the communication that occurs within the swarm. To be more precise, from particular cluster, it can be observed that a single *gBest* update led to the improvement of multiple particles in different iterations. Due to the complexity of Fig. 3, it is not possible to clearly see the density of the network and links of various lengths.

To capture the density of communication (Pluhacek et al. 2017a), the nodes in the network represent the particles in different time points (Particle ID with iteration code). This means that the theoretical maximal number of nodes in the network is the number of particles times the number of iterations. However, a new node in the

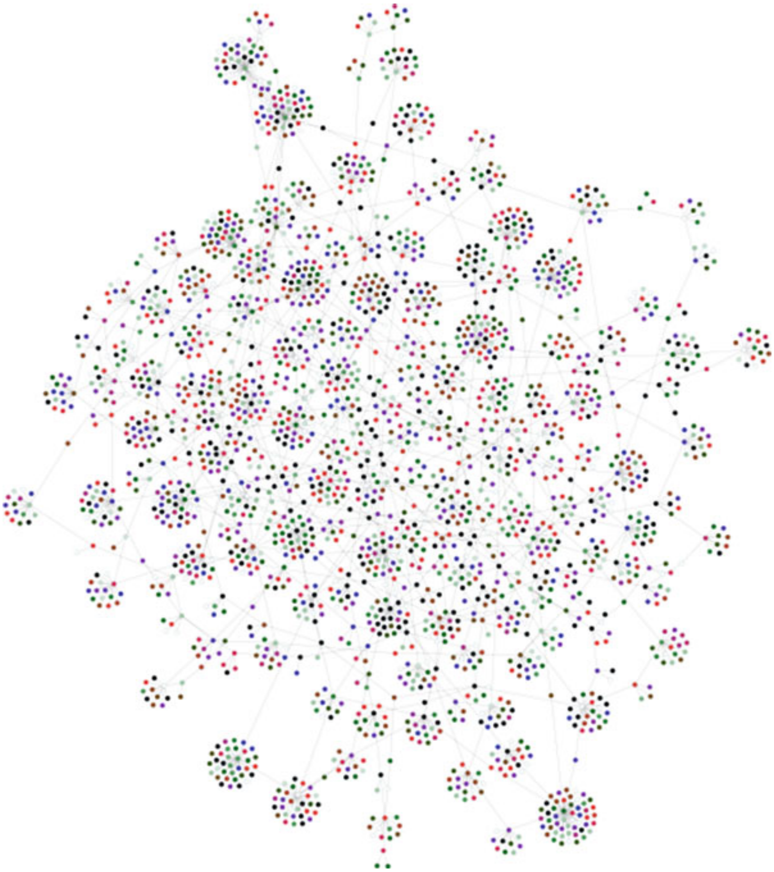


Fig. 3 PSO dynamic as a complex network – complete view

network is created only when a particle manages to find a new personal best solution (*pBest*).

When a node is created, two links are also created. First link is between the newly created node and previous node with the same particle ID (but different iteration code). This represents the information from *pBest* according to (3). Similarly, the information from *gBest* represented by a link between the newly created node and a node represents the last update of *gBest*. In the network visualizations (Fig. 4), a color coding is used to differentiate the phases of the run as percentage of the final number of cost functions evaluations (CFE). (The first 20% of CFE are represented by red color, magenta represents the 20–40% of CFE, green is the 40–60% CFE, 60–80% CFE is represented by yellow color, and finally the 80–100% CFE is represented as cyan.) Such a representation can reveal the relations between the density of communication and convergence speed of the PSO.

Fig. 4 PSO dynamic as a complex network – complete view

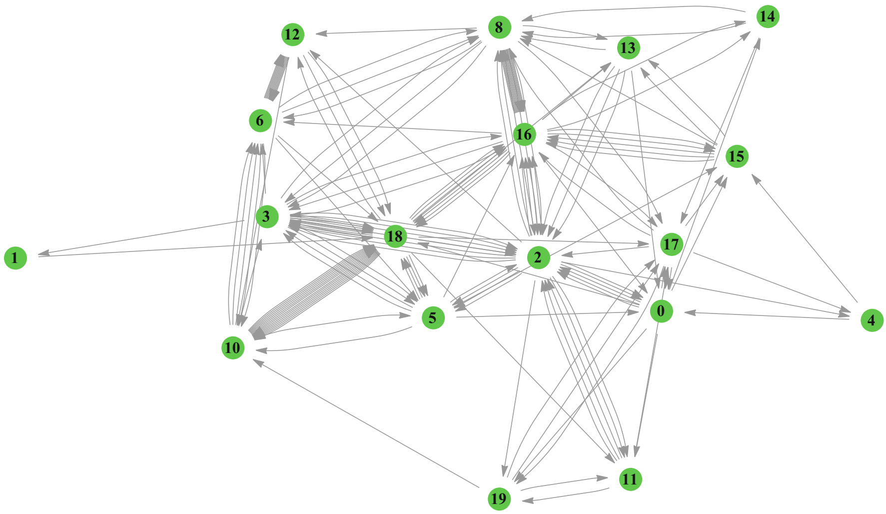


Fig. 5 PSO dynamic as an adjacency graph

Alternatively, it is possible to construct an adjacency graph and to benefit from its statistical features – as with the DE case. The link is created between the particle that triggered the last *gBest* update and the particle that triggers a new *gBest* update. The self-loops (when a new *gBest* is found by exactly the same particle as the previous *gBest*) are omitted. The simplified example is depicted in Fig. 5. Here, for a lower level of complexity and illustration purposes, the population size was limited to 20. More studies aimed at PSO and CN framework are in Pluhacek et al. (2017b, 2016b).

8 Visualizations for FA

In the case of FA, the connection is created, when firefly flies toward another and improves own brightness. This leads to network presented in Fig. 6 (Janostik et al. 2016b). Duplicate connections were omitted from the network in the sake of clarity.

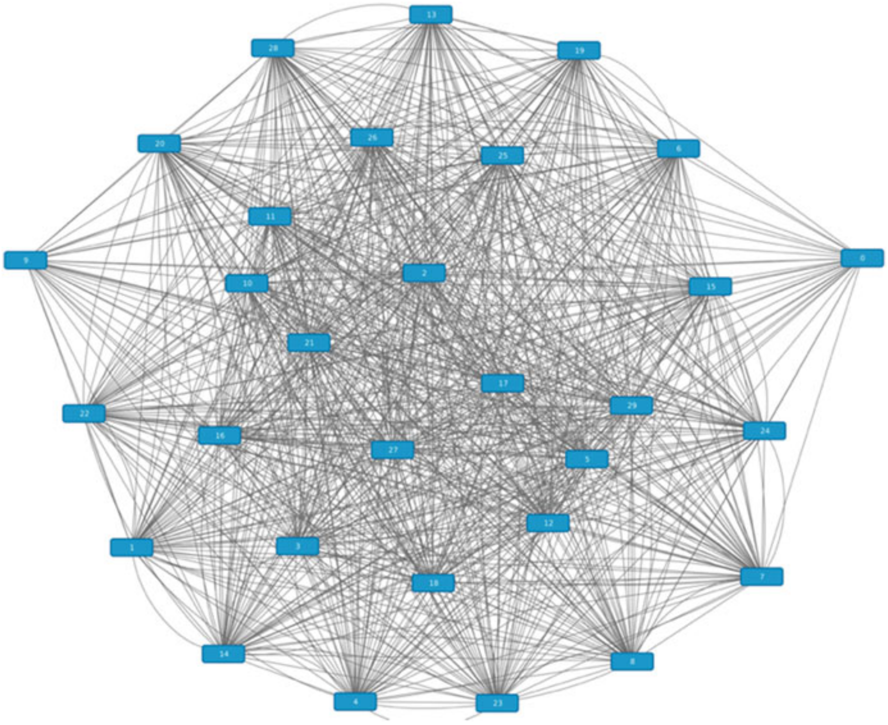


Fig. 6 Basic weighted oriented network for population of size 30 after 100 iterations

Since across multiple iterations of algorithm there may multiple connections between nodes, we decided to improve upon the design by weighting the connections. If there is connection between the fireflies A and B, it starts with weight 1. If in another iteration there is another successful interaction between the fireflies A and B, new connection is not created but the weight of the existing connection is incremented by 1. At the end of evolution, the weight is normalized. If the firefly gets improved by another in every iteration, at the end of the evolution their connection will have weight 1. If it never gets improved, their connection will have weight 0. In Fig. 7, we can see network where connections have their weights visually distinguished. In the top left corner, we can see one dominant firefly which improved entire population more than 70% of iterations (blue lines). On bottom right side, we can observe few fireflies which took part in improvement of population only less than 30% of iterations (red lines). Also from the network we can see that most of the fireflies improved one another only in between 30% and 70% of iterations.

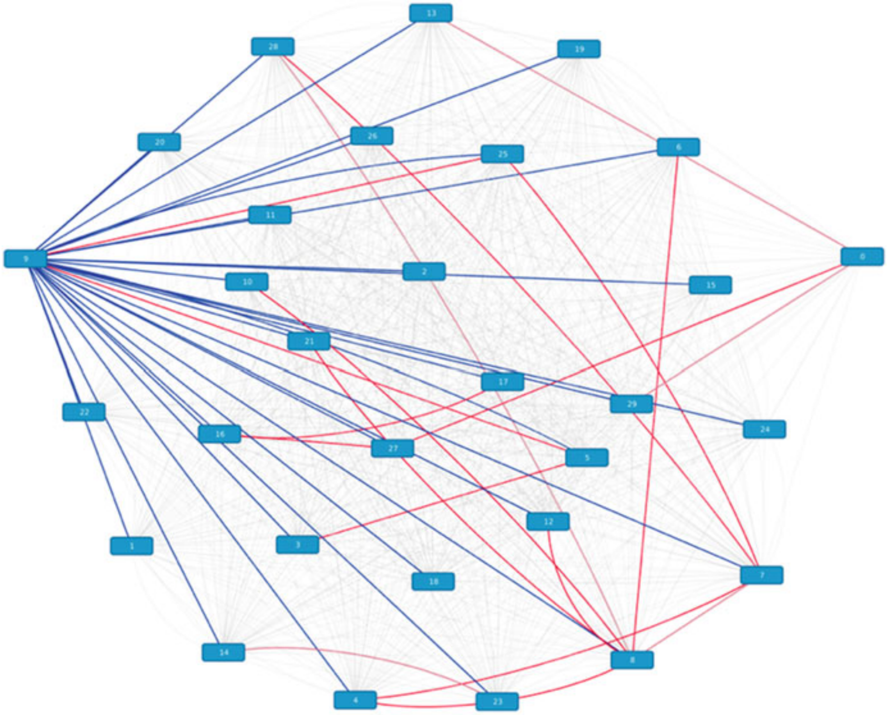


Fig. 7 Basic weighted oriented network for population of size 30 after 100 iterations with visually highlighted weights (Blue, 0.7 < weight; Gray, 0.3 < weight < 0.7; Red: weight < 0.3)

9 Conclusion

This work was aimed at the experimental investigation of the hybridization of a complex network framework using DE, PSO, and FA algorithms. The population was visualized as an evolving complex network, which exhibits nontrivial features. These features provided a clear description of the population during evaluation and can be used for adaptive population and parameter control during the metaheuristic run.

The graphical and numerical data presented herein have fully manifested the influence of either time frame selection or type of construction to the features of the complex network. These features can be used in various adaptive or learning processes. The findings can be summarized as follows:

- Building of the Network: Since there is a direct link between parent solutions and offspring in the evolutionary algorithms, this information is used to build a complex network. In the case of swarm algorithms, the situation is a bit more difficult. It depends on the inner swarm mechanisms, but mostly, it is possible to capture the communications within the swarm during the updating of the information – based on the points of attraction. Two possible approaches are described herein, resulting in different graph visualizations and possible analyses (see Figs. 3 and 4 and Sect. 7).

- Complex Network Features: A complex network created for evolutionary algorithms contains direct information about the selection of individuals and their success; therefore, many network features can be used for controlling a population during an EA run. At the beginning of the optimization process, intensive communication occurs (Fig. 1). Later, hubs (centralities) and clusters are created (Fig. 2), and it is possible to use such information either for the injection or replacement of individuals or to modify/alternate the evolutionary strategy (see Sects. 3.1 and 3.2). In the case of swarm algorithms, the communication dynamics are captured – thus the level of particle performance (usefulness) can be calculated; or alternatively, some subclusters and centralities of such a communication can also be identified – depending on the technique used for the transformation of swarm dynamics into the network.
- Other Features – Randomization, Fitness Landscape: Numerous previous experiments showed that there are no significant changes in complex network features for different test functions in the case of evolutionary algorithms. Nevertheless, the different randomization (distribution) used directly influences the network development through the selection of individuals. Thus, through the preferences of some clusters of individuals, it is possible to temporarily simulate the different randomization inside a metaheuristic algorithm for an entire/subpopulation. Complex network construction for swarm algorithms is not distinctively sensitive to randomization, but the capturing of communications (swarm dynamics) is sensitive to the fitness landscape. Thus, network features can be used for the raw estimation of a fitness landscape (i.e., multimodal/unimodal, or the identification of particular benchmark function) directly.
- In this paper, we have developed several different approaches for visualizations, which can be, of course, hybridized and combined for any metaheuristic techniques according to user-defined requirements, which features are important to observe. Besides the presented approaches, more have been explored for wider portfolio of algorithms (Tomaszek and Zelinka 2016; Krömer et al. 2015; Skanderova et al. 2014).

This novel topic has brought up many new open tasks, which will be resolved in future research. Another advantage is that this complex network framework can be used almost on any metaheuristic algorithms. Moreover, especially for swarm algorithms, there exist many possible ways regarding how to build a complex network.

Acknowledgments This work was supported by Grant Agency of the Czech Republic, GACR P103/15/06700S, further by the financial support of research project NPU I No. MSMT-7778/2014 by the Ministry of Education of the Czech Republic and also by the European Regional Development Fund under the Project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089, partially supported by Grant SGS 2017/134 of VSB-Technical University of Ostrava and by The Ministry of Education, Youth and Sports from the National Programme of Sustainability (NPU II) project "IT4Innovations excellence in science – LQ1602".

References

Skanderova, L., Fabian, T., Zelinka, I. (2016). Small-world hidden in differential evolution. In Evolutionary Computation (CEC), 2016 IEEE Congress on (pp. 3354–3361).

Turing, A.: Intelligent machinery, unpublished report for National Physical Laboratory. In: Michie, D. (ed.) Machine Intelligence, vol. 7 (1969); Turing, A.M. (ed.): The Collected Works, vol. 3, Ince D. North-Holland, Amsterdam (1992).

Holland, J.: Adaptation in natural and artificial systems. Univ. of Michigan Press, Ann Arbor (1975).

Schwefel, H.: Numerische Optimierung von Computer-Modellen, PhD thesis (1974); Reprinted by Birkhauser (1977).

Rechenberg, I.: (1971) Evolutionsstrategie – Optimierung technischer Systeme nach Prinzipien der biologischen Evolution (PhD thesis), Printed in Fromman-Holzboog (1973).

Fogel, D.B.:Unearthinga Fossil from the History of Evolutionary Computation. Fundamenta Informaticae 35(1–4), 1‘Äì16 (1998).

Qin AK, Huang VL, Suganthan PN (2009) Differential Evolution Algorithm With Strategy Adaptation for Global Numerical Optimization. Evolutionary Computation, IEEE Transactions on 13 (2):398–417.

Mallipeddi R, Suganthan PN, Pan QK, Tasgetiren MF (2011) Differential evolution algorithm with ensemble of parameters and mutation strategies. Applied Soft Computing 11 (2):1679–1696.

Das S., Mullick S.S., Suganthan P. (2016) Recentadvancesindifferential evolution – An updated survey, Swarm and Evolutionary Computation, vol. 27, pp. 1–30.

Jabeen H, Jalil Z, Baig AR (2009) Opposition based initialization in particle swarm optimization (O-PSO). Paper presented at the Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference.

Engelbrecht A (2010) Heterogeneous Particle Swarm Optimization. In: Dorigo M, Birattari M, Di Caro G et al. (eds) Swarm Intelligence, vol 6234. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp 191–202.

Zelinka, I. (2016). SOMA—Self-organizing Migrating Algorithm. In Self-Organizing Migrating Algorithm (pp. 3–49). Springer International Publishing.

Zelinka, I. (2004). SOMA—self-organizing migrating algorithm. In New optimization techniques in engineering (pp. 167–217). Springer Berlin Heidelberg.

D. Karaboga, B. Basturk A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm, Journal of Global Optimization, 39 (3) (2007), pp. 459–471.

Yang, X.S.: Nature-inspired metaheuristic algorithms. Luniver Press, Frome, U.K. (2010).

Fister I., Fister I. Jr., Yang X.S, Brest J., A comprehensive review of firefly algorithms, Swarm and Evolutionary Computation, Volume 13, 2013, Pages 34–46.

Zelinka I, Davendra D, Lampinen J, Senkerik R, Pluhacek M Evolutionary algorithms dynamics and its hidden complex network structures. In: Evolutionary Computation (CEC), 2014 IEEE Congress on, 2014, pp 3246–3251.

Davendra D, Zelinka I, Metlicka M, Senkerik R, Pluhacek M Complex network analysis of differential evolution algorithm applied to flowshop with no-wait problem. In: Differential Evolution (SDE), 2014 IEEE Symposium on, 2014a, pp 1–8.

Davendra D, Zelinka I, Senkerik R, Pluhacek M (2014b) Complex Network Analysis of Evolutionary Algorithms Applied to Combinatorial Optimisation Problem. In: Kömer P, Abraham A, Snášel V (eds) Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA 2014. Springer International Publishing, pp 141–150.

Skanderova L, Fabian T (2015) Differential evolution dynamics analysis by complex networks. Soft Computing:1–15.

Metlicka M, Davendra D Ensemble centralities based adaptive Artificial Bee algorithm. In: Evolutionary Computation (CEC), 2015 IEEE Congress on, 2015. pp 3370–3376.

Gajdos P, Kromer P, Zelinka I Network Visualization of Population Dynamics in the Differential Evolution. In: Computational Intelligence, 2015 IEEE Symposium Series on, 2015. pp 1522–1528.

Janostik J, Pluhacek M, Senkerik R, Zelinka I (2016a) Particle Swarm Optimizer with Diversity Measure Based on Swarm Representation in Complex Network. In: Abraham A, Wegrzyn-Wolska K, Hassanien EA, Snasel V, Alimi MA (eds) Proceedings of the Second International Afro-European Conference for Industrial Advancement AECIA 2015. Springer International Publishing, Cham, pp 561–569.

Skanderova, L., Fabian, T., & Zelinka, I. (2017). Differential Evolution Dynamics Modeled by Longitudinal Social Network. Journal of Intelligent Systems, 26(3), 523–529.

Viktorin, A., Pluhacek, M., & Senkerik, R. (2016). Network based linear population size reduction in SHADE. In Intelligent Networking and Collaborative Systems (INCoS), 2016 International Conference on (pp. 86–93).

Senkerik, R., Viktorin, A., Pluhacek, M., Janostik, J., & Davendra, D. (2016). On the influence of different randomization and complex network analysis for differential evolution. In Evolutionary Computation (CEC), 2016 IEEE Congress on (pp. 3346–3353).

Viktorin, A., Senkerik, R., Pluhacek, M., & Kadavy, T. (2017). Towards better population sizing for differential evolution through active population analysis with complex network. In Conference on Complex, Intelligent, and Software Intensive Systems (pp. 225–235).

Pluhacek, M., Janostik, J., Senkerik, R., Zelinka, I., & Davendra, D. (2016a). PSO as complex network—capturing the inner dynamics—initial study. In Proceedings of the Second International Afro-European Conference for Industrial Advancement AECIA 2015 (pp. 551–559).

Pluhacek, M., Šenkeřík, R., Viktorin, A., & Kadavý, T. (2017a). Uncovering communication density in PSO using complex network. In Proceedings-31st European Conference on Modelling and Simulation, ECMS 2017. European Council for Modelling and Simulation.

Pluhacek, M., Viktorin, A., Senkerik, R., Kadavy, T., & Zelinka, I. (2017b). PSO with Partial Population Restart Based on Complex Network Analysis. In International Conference on Hybrid Artificial Intelligence Systems (pp. 183–192).

Pluhacek, M., Senkerik, R., Janostik, A. V. J., & Davendra, D. (2016b). Complex network analysis in PSO as an fitness landscape classifier. In Evolutionary Computation (CEC), 2016 IEEE Congress on (pp. 3332–3337).

Janostik, J., Pluhacek, M., Senkerik, R., Zelinka, I., & Spacek, F. (2016b). Capturing inner dynamics of firefly algorithm in complex network—initial study. In Proceedings of the Second International Afro-European Conference for Industrial Advancement AECIA 2015 (pp. 571–577).

Tomaszek, L., & Zelinka, I. (2016). On performance improvement of the SOMA swarm based algorithm and its complex network duality. In Evolutionary Computation (CEC), 2016 IEEE Congress on (pp. 4494–4500).

Krömer, P., Gajdo, P., & Zelinka, I. (2015). Towards a Network Interpretation of Agent Interaction in Ant Colony Optimization. In Computational Intelligence, 2015 IEEE Symposium Series on (pp. 1126–1132).

Skanderova, L., Zelinka, I., & Saloun, P. (2014). Complex Network Construction Based on SOMA: Vertices In-Degree Reliance on Fitness Value Evolution. In ISCS 2013: Interdisciplinary Symposium on Complex Systems (pp. 291–297). Springer Berlin Heidelberg.

Universal Darwinism and the Origins of Order

John O. Campbell and Michael E. Price

1 Introduction: The Two Processes that Shape the Universe

In this chapter we will propose that features of the observable universe result from two general processes. One is the entropic process of ever-increasing disorder in isolated systems, as described by the second law of thermodynamics. The other is the selection process of Darwinian evolution, which produces orderly entities that are able to exist for extended periods of time, despite the chronic dissipative threat of the entropic process. We will suggest that the selection process accounts for the origins of order not just biologically but in all scientifically observable domains, from the cosmological and quantum to the biological and cultural. The selection process favours the evolution of order in general because, across domains, (1) it favours entities which are relatively capable at solving problems of existence, and at passing their problem-solving advantages on to future generations (e.g. via survival and reproduction), and (2) solving problems of existence, against the powerful current of the second law, requires problem-solving mechanisms that are highly ordered. The 'universal Darwinism' framework we are proposing is currently unconventional, and we acknowledge that it is highly speculative. But we also see it as an unusually compelling framework, particularly because of its high parsimony, explanatory power, and consilience with other bodies of scientific knowledge.

Our main goal in this chapter is to demonstrate how Darwinian selection could provide a unifying framework for natural domains that, superficially, may seem highly disparate. Universal Darwinism could enable such unification because the

J. O. Campbell (✉)
Victoria, BC, Canada

M. E. Price
Department of Life Sciences, Centre for Culture and Evolution, Brunel University London, Uxbridge, UK
e-mail: michael.price@brunel.ac.uk

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_10

selection process seems to involve the same key components, in no matter what domain it operates. We have identified five components as particularly important, and will describe them in detail later in the chapter: (1) a knowledge repository, (2) an embodied adapted system, (3) a process of local entropy reduction, (4) a computational inferential system, and (5) a system of variance, inheritance, and selection. To make it clear how these five key components could enable the evolution of order in a variety of natural domains, we will structure this chapter as follows. In Sect. 2, we will provide an overview of universal Darwinism, outlining the broad contours of this framework's applicability across domains. In Sect. 3, we will focus more closely on each of the natural domains to which this framework applies. We identify five such domains, existing in a nested hierarchy as follows (in the order of decreasing fundamentalness): cosmological, quantum, biological, neural, and cultural. In Sect. 4, we will summarize this universal Darwinism framework.

2 An Overview of Universal Darwinism

A reasonable way to approach universal Darwinism, initially, is to consider the following fact: throughout the entire scope of scientific subjects, Darwinian theories have been used to describe the fundamental processes by which subject-specific phenomena come into existence and subsequently evolve. We will describe many of these theories in more detail below, and for now we will just note that well-developed Darwinian theories have flourished in scientific fields as diverse as cosmology (Smolin 1992, 1997), quantum physics (Zurek 2009), biology (including evolutionary psychology) (Darwin 1859; Tooby and Cosmides 1992), neuroscience (Fernando et al. 2012; Friston 2013), anthropology and archaeology (O'Brien and Lyman 2003; Richerson and Boyd 2005; Gibson and Lawson 2014), economics (Frank 2012a; Nelson and Winter 1982), political science (Pietraszewski et al. 2015; Thayer 2004), organizational behaviour and the business sciences (Arvey and Colarelli 2015; Saad 2011), and epistemology (Campbell 1974; Popper 1972). The ubiquitous power of the Darwinian framework can lead to a deeper understanding of the universe, if we attempt to answer the fundamental questions this power implies. These questions include the following: 'Why does nature use this single process across so many different domains?' 'What common features are shared by Darwinian processes, in all these domains?'

We will propose tentative answers to these questions, answers which suggest a fundamental unification in terms of common evolutionary processes operating across many natural domains. At their core, these processes all involve the familiar 'variance–inheritance–selection' dynamics that are required of any Darwinian system (Dennett 1995): a population of entities must come in various forms, some forms must be more advantageous than others for overcoming problems of existence, and these advantages must be heritable and selectively retained. However, we will flesh out this basic 'selective retention of better-adapted forms' Darwinian model, and present it in terms of an evolved inferential system which learns, implements, and retains *knowledge* about strategies for existence.

2.1 The Accumulated Knowledge Repository as a Central Concept in Universal Darwinism

What does it take to achieve persistent existence in a universe dominated by the relentlessly destructive power of the second law of thermodynamics? Universal Darwinism proposes that existence requires strategies that ‘know’ how to exploit loop-holes in the second law. These strategies produce entities capable of converting lower-entropy energy into higher-entropy waste, in order to maintain their own stable, complex, low-entropy structure (e.g. in biological organisms, homeostasis). In this view, Darwinian selection is the strongest known antidote to the dissipative tendencies of the second law; no other natural process can so capably produce local states of complex existence.

Central to our argument is the contention that existence is a rare and fragile state. Many complex systems, from those described by the fundamental laws and parameters of physics to those existing in the biological and cultural domains, display a remarkable degree of ‘fine-tuning’ for existence. This fine-tuning may be explained as the result of well-developed and highly specialized knowledge repositories. David Deutsch (2011), one of the founders of quantum computation, describes the centrality of knowledge in a general and optimistic light: ‘everything that is not forbidden by laws of nature is achievable, given the right knowledge’. In this chapter we will describe several natural processes of knowledge accumulation, which serve to make existence achievable.

This concept of knowledge is a central aspect of universal Darwinism’s ability to achieve a deeper unification of scientific domains (Campbell 2015, 2016). In order to understand this centrality, it helps to broaden our interpretation of Darwinian ‘fitness’. In evolutionary biology, fitness is conventionally understood in terms of genetic replication, achieved via survival, reproduction, and kin selection (Hamilton 1964). We do not take issue with this interpretation at the biological domain specifically, but we do propose that across domains, fitness can be more broadly understood in terms of the accumulation of knowledgeable strategies for achieving existence. As we explain below, by generalizing from replicating genes to accumulated knowledge, we are able to describe universal Darwinism in terms of information theory and the mathematics of Bayesian inference. This descriptive framework is particularly compelling, because it seems to apply equally well across a wide variety of natural domains.

We suggest that nature has utilized the Darwinian process – the selective retention of better-adapted forms – to accumulate repositories of evolutionary knowledge within several domains of nature. This knowledge is autopoietic (i.e. it enables systems which can maintain and reproduce themselves) and capable of bringing complex forms into existence. We will propose and discuss five types of autopoietic knowledge repositories, one in each domain of nature: (1) in the cosmological domain, this knowledge repository is represented by the laws and parameters of physics; (2) in the quantum domain, by quantum wave functions; (3) in the biological domain, by genomes (potentially acting in concert with epigenetic effects); (4) in the neural domain, by learned neural models; and (5) in the cultural domain, by cultural models. It is no accident that each natural domain possesses its own repository of accumulated knowledge, because the presence of such a knowledge repository is in fact our criterion for identifying each domain. Our strategy for ‘carving nature at the joints’, in other words, is to grant the status of ‘natural domain’ only to those regions of nature that are characterized by their own accumulated knowledge repository.

Although there is a resemblance between our hierarchy of domains and Auguste Comte’s (1848) well-known hierarchy of scientific fields, Comte inserts chemistry between the levels of ‘physics’ (labelled as ‘quantum’ in our hierarchy) and ‘biology’. Why do we not regard chemistry as a separate domain? For reasons discussed below, we see chemistry as part of the cosmological and quantum domains: chemistry is orchestrated according to quantum principles, and arises from the laws and parameters of physics. Chemistry is the outcome of cosmological and quantum accumulated knowledge repositories, in other words, and as it possesses no such repository of its own, we do not consider it a discrete domain.

2.2 Knowledge Repositories Generate Embodied Adapted Systems

A vital observation about knowledge repositories is that they instantiate, via developmental processes, embodied adapted systems (e.g. in the biological domain, a genome instantiates a phenotype). There is a duality, in other words, between the knowledge repository and its embodied adapted system, and we will attempt to identify this duality in each of nature’s domains. Nature’s knowledge repositories tend to be relatively timeless and are copied, through the process of heredity, in a largely conserved and cumulative manner. The adapted systems they instantiate, on the other hand, tend to be short-lived mortal entities; they function to test the soundness of the knowledge which brought them into being, and to update this knowledge when appropriate. If a new phenotypic design encoded by a genetic mutation succeeds at survival and reproduction, for example, then the genome will have passed this test: the mutation will have proven itself to constitute accurate knowledge about how to exist, and more copies of it will now exist in the population.

Because an embodied adapted system functions to test knowledge about how to solve problems of existence, it must be composed of at least one mechanism that attempts to solve some such problem – that is, at least one *adaptation*. An adapted system may be composed of many more than just one adaptation, however, and some of these systems are best characterized as integrated bundles of adaptations. The most familiar example would be a complex biological organism, composed of a vast number of functionally specialized adaptations, each 'designed' by selection to solve a specific problem (or set of problems) related to survival and reproduction. An embodied adapted system may be composed of a diverse array of problem-solving devices, because achieving existence can itself entail a diverse array of problems.

Given that embodied adaptive systems evolve to solve problems of existence, the fundamental 'design principle' of any such system will be to be organized into one or more problem-solving adaptations. However, it will not always be appropriate to regard all aspects of an adapted system as adaptations. These aspects may be (1) adaptations, (2) by-products of adaptations, or (3) random noise (Williams 1966; Tooby and Cosmides 1992: Buss et al. 1998). Adaptations can be considered to be any traits that have been selected to solve some problem(s) of existence (note that by this definition, the category of 'adaptations' subsumes that of so-called exaptations [Gould and Vrba 1982; see also Price, this volume]). A by-product is a trait that is incidentally coupled with some adaptation, but that itself serves no adaptive function. For example, while nipples in female mammals can be considered adaptations for feeding offspring, male nipples seem best explained as by-products of female nipples. Noise represents stochastic, selectively neutral residual effects of the evolutionary process; because noise has no bearing on the adapted system's ability to exist, it is invisible to selection. An example would be a nipple's precise size and shape (e.g. diameter in nanometres).

In general, adaptations can be distinguished from by-products and noise by their improbable complexity, that is, their exceptionally low entropy (Price 2017). Adaptations are the aspects of an adapted system that would be least likely to arise by chance (i.e. as the result of stochastic effects of the evolutionary process), in the absence of a non-random process of selection. A trait's improbability is often best assessed in terms of its functional complexity, that is, the evidence it displays of 'special design' (Williams 1966): the more well-engineered a trait seems for solving a specific adaptive problem – like a key that has been precisely designed to fit a specific lock – the less likely it is to have arisen by chance, and the more likely it is to be an adaptation. That is not to say that selection will tend to favour a more complex form of an adaptation over a less complex form, if both forms fulfil the same function equally well. On the contrary, the simpler form would tend to be favoured, as the more efficient solution. However, more complex forms often do represent superior solutions, and are therefore often selected (Heylighen 1999; Orr 2000; Price, this volume). Evolutionary processes can entail stochastic elements (e.g. genetic drift), but selection is the only aspect of evolution that is expected to produce improbable complexity, and that can produce adaptations. Improbable (functional) complexity is therefore the hallmark of selection, and the characteristic that best distinguishes adaptations from by-products and noise (Williams 1966; Tooby and Cosmides 1992; Price 2017).

2.3 Darwinian Knowledge Repositories Represent Probabilistic Models, Which Reduce Their Own Entropy via Learning

How do evolved knowledge repositories learn adaptively? These repositories can be thought of as probabilistic models that attempt ‘guesses’ about how to successfully exist in the world. They experiment, that is, with different strategies for existence. These models possess not just the property of knowledge but also its inverse, ignorance (i.e. uncertainty). And ignorance, as the ‘father of information theory’ Claude Shannon (1948) pointed out, is mathematically equivalent to entropy. This idea was developed further by E. T. Jaynes, who demonstrated the equivalence of informational and thermodynamic entropy.

Jaynes (1957, 1965) noted that the models used in statistical mechanics to describe thermodynamic systems, while they may contain some knowledge of the system, are inevitably very ignorant: no matter how much statistical evidence is measured regarding a thermodynamic system and used to inform the best scientific models, these remain vastly ignorant of the system’s exact microstate. That is because the evidence supporting these models are measures of macrostates, such as density or temperature, based on average properties of molecules in the system. The density of a volume of gas, for example, is the average number of molecules in the volume, and the temperature is the molecules’ average kinetic energy. So a model that knows such macrovariables has only a small amount of statistical knowledge about the system’s exact microstate. Thermodynamic entropy is a measure of that ignorance, and it is equivalent to information entropy (except that when information theory is applied to thermodynamic entropy, a conversion factor called the Boltzmann constant is used). In this sense entropy can be thought of as the number of bits of information required to move the model from its current state of uncertainty to a state of complete certainty, where it would exactly describe the complete microstate. Entropy is a measure, in other words, of a model’s current state of ignorance.

The relationship between knowledge and entropy may be clarified further if we define knowledge as 2^{-S}, where S is entropy in bits. This definition implies that knowledge increases when entropy/ignorance decreases, and it represents the probability that a model will make a correct guess about the world. For instance, if the entropy of a probabilistic model is 4 bits, then by this definition its knowledge is 2^{-4}, which is equal to 1/16. The model’s entropy, in other words, is equivalent to that of a uniform distribution having 16 members. This is true regardless of the actual characteristics of the particular distribution, including its average and variance (e.g. if we asked you to guess what number we were imagining from 1 to 16, and you had no reason to favour one choice over another, you would have a 1/16th chance of guessing correctly). If the model then encounters some new information about the nature of this distribution, and this update reduces its entropy to 3 bits, then it can be said to have increased its knowledge to 1/8. The model’s entropy is now equivalent to a uniform distribution having 8 members. (Continuing with the above example, if we informed you that the number between 1 and 16 was an even number, you would then have a 1/8th chance of guessing correctly.) The model's knowledge (1/8) can be considered the probability of success of a 'random' guess, where there are 8 possibilities and the model has no evidence favouring any one possibility over another.

The inverse relationship between knowledge and entropy is a vital aspect of universal Darwinism. This relationship can be understood in terms of a fundamental physical principle, related to the second law of thermodynamics,[1] known as the 'principle of maximum entropy': systems will move to the highest state of entropy allowed by the constraints acting upon them (Jaynes 2003). Thus, we may understand a low-entropy dynamical system in terms of the constraints it imposes on the proliferation of disorder. Throughout nature's domains, a commonly employed tactic is to utilize knowledge repositories to impose such constraints, and to thus enable the existence of low-entropy systems. A knowledge repository – acting in concordance with an embodied adaptive system – operates to rule out some possible states that would otherwise be included in its own evolutionary trajectory, and thus reduces its own entropy. For example, in the biological domain, some genetic knowledge repositories code for proteins in the form of enzymes. Enzymes serve to allow only specific chemical pathways, ones which would be extremely unlikely to exist in the absence of the enzymes. In this manner the knowledge repository imposes constraints upon the chemical pathways making up the organism, thus allowing itself (and the organism it encodes) to exist as a low-entropy system**.**

2.4 Knowledge Repositories Learn via Bayesian Updates

A model's ignorance is reduced when evidence is used to update its assigned probabilities in a Bayesian manner. The Bayesian update is the mathematical mechanism which specifies how the prior probabilities composing a model are updated to posterior probabilities; it occurs through the application of a likelihood function, which quantifies how well each possibility is supported by the available evidence. In this view, knowledge is relieved of any anthropomorphic taint: knowledge is not a property just of human minds containing probabilistic models, but of probabilistic models in general. Humans gain knowledge in the same manner as other evolved entities: by employing evidence to update their models.

Universal Darwinism focuses on a specific kind of probabilistic model: knowledge repositories that must guess which strategies will work best for solving problems of existence. These models assign a probability to each possible strategy, based on evidence that the strategy itself merely exists; this probability is, in other

---

[1] Here, following Jaynes (1965), we consider the second law of thermodynamics to apply to both equilibrium and non-equilibrium systems. This view has been most clearly demonstrated by the fluctuation theorem and its experimental support.

words, the frequency of that strategy's existence. Thus, all possibilities of the model that have non-zero probabilities achieve some degree of existence, and the sum of their probabilities equals 1, forming a probability distribution. The relative frequency of each possible strategy may change between generations or iterations, and these changes constitute a Bayesian update of the model (Campbell 2016). In this manner we may understand Darwinian selection as a process which favours the ability to achieve existence. Those traits or entities which can best achieve existence, in other words, are the possibilities of the model that are most frequently selected. To illustrate, imagine that different members of a species have slightly different possible genomes (i.e. biological knowledge repositories), due to random mutations that affect each organism's chance of reproducing. These genomes can be thought of as probabilistic models that test out different reproductive strategies; they make 'guesses' about which strategies will most likely succeed. The genomes that make better guesses will come to predominate in the population, and the species' genome will thus be updated based on new information about the effectiveness of different strategies. In other words, natural selection may be seen as an example of Bayesian model selection.

Darwinian selection may be best understood, in a statistical sense, as the change in the frequency of traits in a population between generations (Frank 2012b). As traits are implemented from instructions contained in the model, we can understand the model as containing one possibility for each existing trait, with the probability assigned to each possibility being equal to its frequency in existence. As some traits become extinct and exit existence, others are introduced, often through random perturbations of the model (such as genetic mutation). These perturbations introduce new variations into the Darwinian selection process, which are then themselves tested for their ability to achieve existence, and assigned a probability equal to their frequency of existence within the next generation. Heredity (or copying) is inherent to this approach, in the form of the transition from the prior to the posterior distribution of possibilities. Posterior distributions tend to be similar to prior ones, with most strategies described by the model being inherited by the next generation. The extent and type of any changes to the posterior distribution are the result of encounters with surprising new evidence (e.g. novel selection pressures).

2.5 Computation

Computation, another key component of universal Darwinism, essentially involves the application of logic (Spencer-Brown 1979), and logic statements can be built up from just a couple of atomic operations (Bishop 1967). It has been demonstrated that any two-state system – such as 'true/false', 'on/off', or 'negative/positive' – can be used to form all possible logical operations (Spencer-Brown 1979). All of mathematics can be constructed from logical statements, as can universal Turing machines (computers). Universal computers, when running properly constructed algorithms (which are themselves built from atomic logical operations), may output any pattern of applied mathematics that describes a natural process.

A relatively recent computer technique, called evolutionary computation (De Jong 2006), uses the Darwinian selection process to solve problems of existence. Solutions may be selected according to a programmer-provided fitness function, but the programmer need not specify the detailed steps by which a Darwinian algorithm can find these solutions. The programmer instead just provides an initial algorithm, and the fitness function which evaluates the optimality of solutions. The program then randomly varies the algorithm, tests the output of these variations against the specifications of the fitness function, and selects the variations that best succeed. It then makes small variations in some of the selected algorithms, and repeats the process until a solution is found that sufficiently satisfies the fitness function.

We could view the evolutionary computation process described above as an artificial execution of Darwinian selection. Equally, however, we could consider natural executions of Darwinian selection as forms of computation. This latter perspective may help explain the pervasiveness of Darwinian selection in nature. A computational substrate lies at the foundations of physical reality: all quantum dynamics may be precisely coded in an algorithm written in qubits (Deutsch 1985; Lloyd 2013), so physical reality at the quantum level is essentially computational. We will attempt to demonstrate below that each natural domain may be considered in terms of a knowledge repository that functions as a complex algorithm and that each domain-specific algorithm orchestrates the dynamics of its substrate (i.e. of material from lower-lying domains). This orchestration produces new, mathematically describable complex forms that are capable of existence. As this computational process outputs mathematical patterns, it may go some way towards explaining what Wigner (1960) described as the ‘unreasonable effectiveness of mathematics in the natural sciences’.

Given that computation may arise from any simple two-state substrate, it is perhaps unremarkable that the substrates of many natural domains are so amenable to computation. Indeed it has been demonstrated that computation may be performed using the substrates of qubits (Deutsch 1985), chemistry (McGregor 2012), DNA (Ignatova et al. 2008), and neurons (Zylberberg et al. 2011). A more interesting question might be: ‘Why does nature exploit this potential by employing computation across its many domains?’ We suggest that the answer provided by Lloyd (2013), in reference to the quantum domain, may be generalized to other domains as well: ‘The reason is that many complex, ordered structures can be produced from short computer programs, albeit after lengthy calculations’. In other words, it may be much more efficient for evolutionary processes – which are in the business of discovering complex, ordered structures that are capable of existence – to search for algorithms that can produce such structures, instead of searching for the structures directly.

2.6 A System of Variance, Inheritance, and Selection

Now that we have outlined the key components of this relatively novel universal Darwinism framework, we wish to clarify how this framework can be understood in more traditional Darwinian terms. It is commonly observed that for Darwinian selection to occur in general, a population of entities must be subject to a system of variance, inheritance, and selection (Dennett 1995). That is, entities must vary in form, some forms must be better than others at solving problems of existence, and these solutions must be heritable and selectively retained.

We propose that in the context of universal Darwinism, these 'variance–inheritance–selection' dynamics unfold as follows. Each natural domain is orchestrated by a heritable knowledge repository: a probabilistic model that contains knowledge about possible autopoietic strategies, or algorithms, for achieving existence in that domain. No copying process is exact, and variations introduced in the hereditary process create slight variations in these algorithms. The execution of these algorithms, which occurs through developmental processes, produces embodied adaptive systems which then test the algorithms for their ability to promote existence. The evidence produced by these tests is used to select the algorithms that succeed, and this selection process in turn updates the probabilistic model (i.e. the knowledge repository) in a Bayesian manner. From this perspective, we see how the general 'variance–inheritance–selection' model of the evolutionary process may be incorporated into the specific universal Darwinism framework outlined above.

3 The Five Domains of Universal Darwinism

In the overview of universal Darwinism presented above, we describe how the key components of this framework operate to generate complex order across natural domains. We will now zoom in on each domain, and see how these components function in these more specific contexts. But we will first provide a brief justification for the method by which we have drawn boundaries between these domains.

It is obvious that a good deal of complexity has evolved in the universe since the Big Bang. The growth of this complexity has been cumulative, with new, additional forms of complexity repeatedly emerging from substrates provided by older forms of complexity. Within this hierarchy of progressive complexity, it is a challenge to identify clear conceptual boundaries between the various forms or domains of complexity that have evolved. We believe a strong case can be made, however, for proposing the following five natural domains: cosmological, quantum, biological, neural, and cultural. Our criterion for regarding each of these domains as distinct is that each seems to possess its own knowledge repository, which has evolved in order to orchestrate newer complex structures from older complex substrates. In the cosmological domain, the knowledge repository is represented by the laws and parameters of physics; in the quantum domain, by the quantum wave function; in the biological domain, by the genome (potentially in concert with epigenetic effects); in the neural domain, by learned mental models; and in the cultural domain, by cultural models. The natural domains are organized in a nested pattern; for example, culture is composed of a neural substrate, which is composed of a biological substrate, and so on. The boundaries between these domains arise naturally, whenever a new form of knowledge repository emerges that can orchestrate existing substrates into new forms of complex existence. The reason we have decided not to consider chemistry as a separate domain (in contrast to Comte's scientific hierarchy, as noted above) is because there does not seem to be a knowledge repository that can be seen as chemistry-specific, and distinct from those of quantum physics and biology.

In the manner described above, nature can be viewed as a nested hierarchy of domains. We will now consider how the universal Darwinism framework can be applied to each domain in turn, ordered from most to least fundamental.

3.1 Cosmological Selection

Our use of 'cosmological' is perhaps slightly unusual, in that we mean the physical processes already in place at the beginnings of the universe. These processes, in the form of the laws and parameters of physics, gave rise to what is usually thought of as the cosmos. In this sense the cosmological domain is fundamental in providing the common substrate, and the rules of nature, from which all subsequent domains emerge.

One of the most basic cosmological paradoxes is the 'Goldilocks principle' (sometimes called the 'fine tuning problem'): the proposition that the conditions that allow complexity to exist in the universe can occur only when approximately 32 universal dimensionless constants of physical theory lie within a very narrow range. If any fundamental constants were even only slightly different, the universe would not be conducive to the establishment and development of matter, astronomical structures, chemical diversity, or life. A question therefore arises: 'Why should the universe's physical laws and parameters have the precise values necessary to develop complexity?'

3.1.1 Cosmological Natural Selection: With and Without Intelligence

One scientific answer to this puzzle that has received fairly widespread attention is Smolin's theory of cosmological natural selection, or CNS (Smolin 1992, 1997, 2007). Smolin's CNS assumes the perspective, which is by now fairly conventional among cosmologists, that our universe is just one of many in an extremely large population of universes – a multiverse – in which new universes are being constantly generated. Smolin adds to this perspective the idea that in such a multiverse, universe designs that were better at reproducing themselves would become better-represented. In other words, Smolin's CNS is modelled on Darwin's theory of natural selection, and it explains the universe's finely tuned aspects as evolved adaptations for reproduction. These adaptations cause universes to develop black holes, from which emerge offspring universes that are separate and inaccessible from their parent universe (Gardner and Conlon 2013). In short, this theory explains the fine-tuned laws and parameters of our universe as a knowledge repository. This repository produces adaptations in the form of complex structures, such as chemistry and stars, that are necessary to produce black holes. This theory embodies the Darwinian paradigm of variance, inheritance and selection as it posits that child universes, born in the process of black hole formation, inherit slightly varied physical laws and parameters from their parent. Over many generations, this process selects for universe designs which prove most prolific at black hole production. As the laws of physics have become extremely precise, allowing little latitude for deviations, they impose tight constraints on the entropy of the physical processes that lead to universe reproduction. We thus may understand cosmological selection as the primary anti-entropic example of Darwinian selection in the nested hierarchy of domains.

The major theoretical variant of Smolin's CNS is cosmological natural selection with intelligence, or CNSI (Crane 1994/2010; Harrison 1995; Gardner 2000; Smart 2009; Vidal 2014; Price 2017), which proposes that intelligent life is an adaptation for universe reproduction. CNSI expects that in the future, intelligent life will acquire the technical expertise that would allow it to create new universes, which would replicate the laws and parameters of its natal universe (to which it would already be adapted). That is not to suggest that black holes could not also enable universe reproduction; both they and intelligence could fulfil this function. However, intelligence may represent the more recently evolved of these adaptations and could perhaps prove superior to black holes at the task of high-fidelity universe reproduction (Price 2017, this volume).

The perspective of CNSI fits nicely within the context of universal Darwinism: intelligence is the ability to discover knowledge, and as noted above, anything allowed by the laws of nature – including the creation of new universes – is achievable, given sufficient knowledge (Deutsch 1997). As knowledge repositories are themselves physical structures, both they and the intelligence which produces them may be considered fundamental aspects of evolution. In line with this perspective, it has been suggested that intelligent life is the most complexly improbable aspect of the universe's embodied adapted system, or 'phenotype'. Therefore, according to the adaptationist logic spelled out in Sect. 2.2, intelligent life represents the aspect of this phenotype that is most likely to be an adaptation, as opposed to a by-product or noise (Price 2017, this volume). From this perspective, human biocultural evolution would represent a developmental subroutine of cosmological evolution, targeted towards production of a species that would be intelligent enough to reproduce its own universe.

3.1.2 The Cosmological Knowledge Repository

It remains an open question which of the two models of cosmological evolution presented above – Smolin's CNS, or CNSI – is more plausible. Regardless, the information that allows cosmological evolution to happen must have a physical representation. We therefore might ask: 'Where and in what form is the cosmological knowledge repository physically located?' There is no consensus yet about the specifics of the answer, but there is about its general outline. The holographic principle (Stephens et al. 1993) has been widely adopted, particularly the anti-de Sitter/conformal field theory (AdS/CFT) correspondence (Maldacena 1998), which stipulates a duality between an informational model and physical processes. The model may be considered to exist on the boundary of physical space, and is an exact dual of the physics which occur within that space. An emerging consensus is that this informational model is composed of quantum entanglements (Orus 2014; Verlinde 2010). In particular, it has been demonstrated that the gravitational field equations emerge from this entanglement (Swingle and Van Raamsdonk 2014). This view fits well within universal Darwinism, as it identifies the knowledge repository which undergoes the Darwinian process of variance, heredity, and selection. Some researchers have described this knowledge repository in biological terms, even going so far as to call it the 'DNA' of physical reality (Orus 2014). In this view, the actual physics of the universe arose through a developmental process specified by the knowledge repository, and the physical structure of the universe may thus be considered an embodied adapted system in the domain of cosmological selection.

The laws of nature support universal computation, that is, computational processes capable of generating the entire range of patterns or states that are mathematically describable. This basic computational substrate of the laws of nature, which itself may be the product of a Darwinian evolutionary process, may be understood as the foundation from which all subsequent evolutionary processes and domains have emerged. In this sense, we may consider those emergent processes as subroutines of cosmological selection, developed from a primary algorithm or substrate which takes the form of nature's laws.

This brief sketch allows us to consider the cosmological domain in terms of an inferential system. It contains a knowledge repository, accumulated over evolutionary time, that in each generation of universes is updated with evidence concerning what does and what does not facilitate existence. In each generation, a variety of autopoietic algorithms encoded by the knowledge repository are executed and tested. Those that produce outputs which form a mechanism for reproduction – such as atoms, chemistry, stars, and black holes – are selected, and they retain updated hereditable models of existence, variants of which will in turn be tested in future generations.

3.2 Quantum Selection

The cosmological domain, orchestrated by the laws and parameters of physics, provides a fecund substrate for the evolution of complexity. To understand why it does so –assuming we accept that physics is governed by quantum laws, and that gravity is emergent from quantum phenomena – we can ask: ‘What special properties do quantum laws possess that facilitate the emergence of complexity?’ We suggest that two widely accepted quantum properties may be involved.

First, quantum theory may be interpreted as a description of inferential systems (Rovelli 1996; Campbell 2010; Hoehn 2017a). A proposed set of information theory-type axioms for quantum theory describe constraints on the acquisition of information in the quantum domain. It has been demonstrated that any observer – from a quantum observer to a human one – who processes information as stipulated by these axioms, and who uses the mathematics of Bayesian inference, will infer the descriptions of quantum theory (Hoehn 2017b). In other words, the best description of quantum phenomena that can be inferred by an observer, given the possible evidence that they could gather, is given by quantum theory.

The information which a quantum system may convey to other quantum systems in its environment, including human-constructed measuring devices, is highly constrained. In particular, information conveying superposed quantum states – which includes the vast majority of all quantum states – cannot exist outside its own system, and so cannot be conveyed to an observer (Zurek and Zwolak 2013). Instead, only classical (or Holevo) information can survive this transition in measurable quantities. This transition occurs via a selection process dubbed quantum Darwinism (Zurek 2009).

3.2.1 Quantum Darwinism: Classical Reality as an Embodied Adapted System

According to quantum Darwinism, observers will experience a quantum system in terms of the classical information – such as position, momentum, and charge – which can survive the transition to, and thus achieve existence in, classical reality. The extended network formed by this information constitutes classical reality itself (Zurek 2014). We may consider classical reality to be an embodied adapted system, constructed from the quantum knowledge repository – that is, from the quantum system’s wave function, which describes the quantum state of the system. Although the wave function is usually portrayed as a mathematical abstraction, quantum theory tells us that it must also have a physical form (Landauer 1996): quantum systems must include a physical implementation of an information processing model, equivalent to the evolution of a state vector in Hilbert space. The wave function’s physical form, however, may exist at near the Plank scale and therefore be as yet undetectable. A near consensus has only recently arisen that the wave function must have a physical incarnation (i.e. an ‘ontic’ rather than just a possible 'epistemic' form) (Pusey et al. 2012; Hardy 2013), and 't Hooft (2016) has speculated on its possible physical forms, near the Planck scale. We are left in a situation similar to that of biology prior to the discovery of DNA in 1953. Although there was consensus agreement at that time about the computational effectiveness of Mendel's genes, the genes' physical form was still unknown (and many doubted there was a physical form).

In order to understand how quantum Darwinism occurs as a system of variance, inheritance, and selection, it may help to review the dynamics of a quantum interaction. This interaction – in which information is transferred between one quantum system and another in its environment – involves the processes of entanglement and decoherence. Before two quantum systems become entangled, each is described by its own wave function. Entanglement represents a merger of the two systems: during entanglement the two systems share a single wave function, and the individual systems in effect cease to exist. Then, in the process of decoherence, the entangled system again splits into two. Each system is again described by its own wave function, in terms of a pointer or classical state, but these functions now contain a high degree of mutual information. Each system then begins to evolve separately from the other, in accordance with the Schrödinger equation. This evolution re-emerges each system into the quantum realm: the re-introduction of superpositions results in a smearing of each system's predictive abilities, over a probability distribution of possible outcomes.

As two quantum systems becomes entangled and exchange information, and then subsequently undergo decoherence, each quantum system's wave function is updated based on information it receives from the environment. Following the 'quantum jump' that results from the interaction, the system's wave function accurately predicts (with probability 1) the outcome of the same immediately repeated interaction. In other words, the 'quantum jump' may be interpreted as a Bayesian update, which brings the quantum model to certainty concerning the outcome of a particular interaction. We may therefore view this information, which the system attempts to copy to its environment, as an experimental test of a strategy for existence in classical reality. Those tests that achieve existence serve to update the quantum knowledge repository (the wave function).

Quantum decoherence essentially describes a process by which information is copied or transferred from a quantum system to its environment. Much of the varied information contained in the state vector is copied, but most has extremely short periods of survival. The fittest quantum states, or pointer states, are selected (Zurek 1998). The selected quantum states are those capable of the greatest reproductive success (Zurek 2014). Clearly quantum Darwinism is formulated in a manner consistent with three-step Darwinian dynamics: variance, inheritance, and selection. Quantum interactions which select classical (i.e. pointer) states, therefore, may be considered a Darwinian process. These interactions minimize the quantum system's entropy change, thus conforming to the paradigm in which the selection process counters the entropic process. As Zurek and colleagues describe it (Dalvit et al. 2005), 'pointer states are obtained by minimizing system entropy and demanding that the answer be robust when varying within a reasonable range of time'.

Classical or pointer states of quantum systems are selected for their ability to survive in the environment, and are robust in the face of decoherence. Quantum systems may merge, through entanglement, but again those more complex systems that survive intact are adapted to their environments and therefore able to exist. This evolutionary process explores the space of possible merged designs, and the resulting adapted systems describe an evolutionary tree: sub-atomic particles, atoms, molecules, and complex chemistry. As Zurek writes (parenthetical text added) (Zurek 2007):

> In the real world observers find out pointer states of systems . . . by letting natural selection take its course: Pointer states are the "robust species", adapted to their environments. They survive intact its monitoring. More importantly, multiple records about S (the system) are deposited in E (the environment). They favor pointer states, which are the "fittest" in the Darwinian sense – they can survive and multiply.

Although this Darwinian description of quantum dynamics is novel, it is firmly grounded in the standard quantum axioms (Dirac 1958; Gillespie 1970). The first two of these axioms describe the wave function as continuously evolving in Hilbert space, in accordance with Schrödinger's equation. We may interpret these axioms as describing variable forms of the wave function. Two axioms describe the 'measuring process' in which information is passed to the environment in a process analogous to heredity. These axioms apply measurement operators and Born's rule to the wave function, to select a specific outcome eigenvector and a probability distribution of output eigenvalues. As Zurek demonstrates, this selection process is entirely consistent with a Darwinian interpretation in terms of heredity–variation–selection. A final axiom states that measurements or interaction result in the updating of the probabilistic wave function to greater predictive accuracy. This may be considered a Bayesian update, so the quantum system may be considered an inferential system.

In this view, these quantum axioms – accepted as standard for 60 years – readily lend themselves to a Darwinian interpretation. More recently, quantum dynamics have been shown to be consistent with the inferential system interpretation of universal Darwinism (Zurek 2014; Hoehn 2017b; Campbell 2010) that we present here.

3.2.2 Computing the Quantum Domain

A second property of quantum phenomena that enables emergent complexity is the duality between quantum physical forms and computation (Markopulou 2012). This duality is stated as the Church–Turing–Deutsch principle (Deutsch 1985; Lloyd 2013): any physical (quantum) phenomena may be simulated to any degree of accuracy, with an algorithm written in qubits and run on a quantum computer. As Lloyd (2007) describes it, the universe computes 'its own dynamical evolution; as the computation proceeds, reality unfolds'.

The computational nature of the quantum domain raises the question: 'Why did cosmological evolution settle on a computational substrate in its quest to evolve complexity'? The answer suggested by Lloyd is that a quantum algorithm specifying a complex form may be much less complex – and thus more easily discovered, through a Darwinian search – than the form itself (Lloyd 2013). Thus, a computational substrate, running Darwinian algorithms, may facilitate the discovery of complex forms that are capable of existence.

Quantum computations are performed within a system of entangled qubits, but the answer to the computation is outputted through an interaction with another quantum system (e.g. a measurement) (Gavrilov 2012). In this manner we may consider classical reality as the output of quantum computation, or equivalently, as the calculation of forms capable of existence within classical reality.

3.3 Biological Selection

Biological evolution by natural selection is widely understood and non-controversial, and throughout Sect. 2 above we provided numerous examples of how it operates in the context of universal Darwinism. We will therefore sketch biological selection only briefly in the current section, with a focus on aspects of it that we have not already emphasized.

Natural selection is generally understood as acting on a genetic knowledge repository (that may act in concert with epigenetic effects), which orchestrates the material of quantum chemistry into the many structures of life. We use this distinction to define the transition between the quantum and biological domains, following Davies and Walker (2012):

> We therefore identify the transition from non-life to life with a fundamental shift in the causal structure of the system, specifically, a transition to a state in which algorithmic information gains direct, context-dependent, causal efficacy over matter.

Crucially, the mathematical model of natural selection is extremely general, substrate independent, and applicable to any population of traits which vary in frequency between generations according to a fitness function. Because any process described by this generalized formula meets the criteria of a generalized Darwinian system, the formula serves as a mathematical framework for universal Darwinism. Steven Frank (Frank 2012b) demonstrates this simple, general model:

> A simple model starts with n different types of individuals. The frequency of each type is q_i. Each type has w_i offspring, where w expresses fitness. In the simplest case, each type is a clone producing w_i copies of itself in each round of reproduction. The frequency of each type after selection is
>
> $$q_i' = q_i \frac{w_i}{\mathrm{w}}$$
>
> Where $w = \sum_1^n q_i w_i$ is the average fitness of the trait in the population. The summation is over all of the n different types indexed by the i subscripts.

Fitness functions may be interpreted as likelihood functions in Bayesian inference. In the example above, the frequency before and after are the prior and posterior probabilities, and relative fitness is the likelihood function of a Bayesian update. In biological terms, the likelihood function is the probability of finding a specific phenotypic trait in the population. In Bayesian terms, inference adapts the model to conform with the evidence (i.e. with those traits which exist); thus the Bayesian likelihood function is equivalent to the biological fitness function. Several researchers, including S. Frank (Frank 2012b) and Friston (2013), have developed mathematical theories of evolutionary change equivalent to the inferential system model developed here (Campbell 2016).

In the biological case, the q_i form a probability distribution that describes the frequency of success enjoyed by a range of strategies for existence. This probabilistic model refers to autopoietic strategies that are tested as phenotypes (i.e. embodied adapted systems). Those strategies which successfully contribute to existence in the next generation – those, that is, that display fitness – update the strategy frequencies found in the next generation. As the phenotypes produced are highly complex, low-entropy systems, this process illustrates the general ability of inferential systems to overcome the forces of entropic dissipation.

A computational interpretation of biology has been developing over the past several decades. Dawkins (1982) endorsed this view decades ago, when computation was starting to become a pervasive cultural metaphor:

> I shall use the word ‘program’ in exactly the same sense as Maynard Smith uses ‘strategy’ ... A program (or strategy) is a recipe for action, a set of notional instructions that an animal seems to be ‘obeying’, just as a computer obeys its program.

A bit later, Nobel Laureate Sydney Brenner (1999) urged researchers to consider biology in computational terms:

> One way of looking at the problem is to ask whether we can compute organisms from their DNA sequences. This computational approach is related to Von Neumann’s suggestion that very complex behaviours may be explicable only by providing the algorithm that generates that behaviour, that is, explanation by way of simulation ... A proper simulation must be couched in the machine language of the object, in genes, proteins and cells.

More recently, Davies and Walker (2012) have presented biology as an essentially algorithmic process, whereby a largely genetic algorithm orchestrates its chemical substrate into the complex output of life:

> As we have presented it here, the key distinction between the origin of life and other emergent transitions is the onset of distributed information control, enabling context-dependent causation, where an abstract and non-physical systemic entity (algorithmic information) effectively becomes a causal agent capable of manipulating its material substrate.

The above views fit well within the universal Darwinism framework. Specifically, they complement the notion that one natural domain evolves from another when it forms a knowledge repository (i.e. algorithm) that can orchestrate the substrate of the previous domain into new complex forms that are capable of existence.

The algorithmic nature of genetics has been illuminated through the Encode project, which found that the majority of the genetic repository codes for regulatory functions rather than for the direct construction of proteins (Brydon 2012):

> It is thought that these “non-coding” RNA transcripts act something like components of a giant, complex switchboard, controlling a network of many events in the cell by regulating the processes of replication, transcription and translation – that is, the copying of DNA and the making of proteins based on information carried by messenger RNAs.

We may think of this regulation in terms of computer programs containing many nested ‘IF, THEN, ELSE, END’ statements. For example, ‘IF the concentration of molecule A is above a given level, THEN produce more molecule B, ELSE produce more molecule C, END’. Such algorithms appear to be methods widely used by life to compute biological output from chemical input.

3.4 Neural Selection

Biological natural selection – understood as the evolution of principally genetic knowledge repositories, capable of instantiating and testing phenotypic adapted systems – has produced a wide range of adaptations. Biological selection has fashioned these adaptations from a variety of organismal tissue types, and for hundreds of millions of years this variety has included neural tissue. An adaptation constructed from neural tissue is not less ‘biological’ than one constructed from, say, skeletal or muscular tissue, and the category of ‘biological adaptations’ includes many genetically encoded, psychological, information-processing adaptations. The field of evolutionary psychology (Tooby and Cosmides 1992) is devoted to the identification and functional analysis of species-typical adaptations of this kind. We emphasize these points to make clear that in the context of the ‘neural selection’ domain, we will *not* be focusing on genetically encoded psychological adaptations, because we consider these to be biological adaptations. The neural selection domain is concerned, rather, with neuronal systems that are organized primarily via experience and learning.

To further clarify the distinction between the psychological embodied adapted systems that exist in the biological versus the neural domain, consider spoken human language. Human nature appears to be equipped with genetically encoded psychological language-acquisition mechanisms (Pinker and Bloom 1992). These adaptations enable all normally developing individuals to acquire, relatively automatically, whatever language is spoken in their early childhood developmental environment. Because these mechanisms are species-typical, all human cultures exhibit complex spoken language. There is of course diversity, however, in the linguistic rules and words used across cultures; just because you know one language does not mean you will know any others. Language acquisition, therefore, depends on two important kinds of knowledge repositories. The first is the ‘biological’ knowledge repository, and it provides species-specific, genetically encoded language acquisition mechanisms, which enable children to acquire whatever language is spoken in their population. The second is the ‘neural’ knowledge repository, which stores the child’s knowledge about the specific rules and words of their own spoken language.

It is crucial to emphasize that ‘biological’ knowledge repositories are genetic and ‘neural’ ones are non-genetic. It is also important, however, to note that from a certain perspective it is non-sensical to refer to neural knowledge repositories as non-genetic. That is because if it weren’t for genes, these neural structures would, of course, not exist at all. It is ultimately genes, for example, which encode for the language acquisition mechanisms that enable a child to acquire a specific language, even if the genes do not specify what language this will be. Although the terminology of ‘genetic’ versus ‘non-genetic’ presents this risk of confusion, however, this confusion itself reflects a familiar pattern in universal Darwinism: a new kind of natural domain has emerged, because the material of a substrate domain has been orchestrated into a new kind of knowledge repository. In this case, the biological substrate of genetically encoded neurons has been orchestrated to create a new, neural-based knowledge repository, which is organized primarily via experience and learning.

In the domain of neural selection, if the neural structure represents the knowledge repository, then the embodied adapted system would be the behavioural output of this neural structure. For instance, in the case of acquired language the embodied adapted system would be acts entailing the use of language to achieve one’s goals. Another example would be an organism that has to learn which species in its environment constitute edible prey (such as the southern toad in the example below). The neural information processing system that enabled such learning would take environmental input into consideration (e.g. which prey species tasted best) and produce learned predatory behaviour as output. Learning-based neural systems thus provide a new level of information processing, to coordinate an organism’s repertoire of neural-based sensory information with its repertoire of neural-based behaviours.

3.4.1 Bayesian Brains

The cerebral cortex is largely responsible for this sensory/behaviour mapping in mammals, but the evolutionary origins of such mapping may extend back to near the genesis of multi-cellular organisms (Tomer et al. 2010). The cerebral cortex is universally composed of sophisticated computational units called ‘cortical columns’ and a near consensus is emerging that the computations performed are a type of Bayesian inference (Lee 2015), as for example with the predictive coding model (Cain et al. 2016). This convergence of neuroscientific understanding to a view of the cerebral cortex as a type of Bayesian computer is sometimes called the Bayesian Brain theory. Karl Friston and colleagues summarize this computational algorithm executed by the Bayesian brain (Bastos et al. 2012):

> It might be thought impossible to specify the computations performed by the brain. However, there are some fairly fundamental constraints on the basic form of neuronal dynamics. The argument goes as follows—and can be regarded as a brief summary of the free energy principle:
>
> - Biological systems are homeostatic (or allostatic), which means that they minimize the dispersion (entropy) of their interoceptive and exteroceptive states.
> - Entropy is the average of surprise time, which means that biological systems minimize the surprise associated with their sensory states at each point in time.
> - In statistics, surprise is the negative logarithm of Bayesian model evidence, which means biological systems—like the brain—must continually maximize the Bayesian evidence for their (generative) model of sensory inputs.
> - Maximizing Bayesian model evidence corresponds to Bayesian filtering of sensory inputs. This is also known as predictive coding.

This formulation incorporates senses and perceptions as inputs to a Bayesian inferential system and behaviours as outputs (Friston et al. 2017):

> This allows one to frame behavior as fulfilling optimistic predictions, where the optimism is prescribed by prior preferences or goals. In other words, action realizes predictions that are biased toward preferred outcomes. More specifically, the generative model entails beliefs about future states and policies, where policies that lead to preferred outcomes are more likely. This enables action to realize the next (proximal) outcome predicted by the policy that leads to (distal) goals. This behavior emerges when action and inference maximize the evidence or marginal likelihood of the model generating predictions.

The generative neural models produced through the processing of learned information entail beliefs about future states, and these models are updated through evidence-based (i.e. sensory) experience to become more accurate in reaching preferred goals. The sum of these generative models in an organism may be considered as its total knowledge repository in the neural domain. The selective process, described in terms of preferences or goals, is fully consistent with Darwinian mathematics in terms of a fitness function. A range of variable generative models is selected, based on fitness goals; these somewhat random behavioural variations are experimented with, especially among immature organisms. The behaviours that are selected and retained are those which maximize 'fitness' (goal achievement), that is, those which produce evidence in support of predicted success. Successful behaviours are learned and able to be reproduced in appropriate future circumstances, constituting a type of heredity. Thus the Bayesian Brain theory, cast in terms of the free energy principle, may be interpreted as a Darwinian process.

Over the past decade, Friston has demonstrated how the Bayesian Brain model emerges from a general principle of evolutionary dynamics called the free energy minimization principle. This principle states that dynamic systems act to reduce both the prediction error of their internal models and the complexity of those models. When applied to neuroscience, this principle states that internal models of the world are updated, in a Bayesian manner, to reduce discrepancies between the evidence the model receives in the form of sensory inputs and the model's predictions about those inputs. The brain achieves this accuracy with a low level of complexity, as is evident when we consider that the cerebral cortex accomplishes sophisticated perceptual and behavioural tasks more accurately than the current generation of super computers, despite using much less energy and computational power than these computers (Fischetti 2011).

Friston and colleagues have recently extended the free energy paradigm to encompass the dynamic systems of biology and culture (Friston 2013; Ramstead et al. 2017). The free energy minimization principle has been shown to be equivalent to the inferential systems model of evolutionary dynamics (Campbell 2016); systems tend to evolve more accurate internal models, detailing possible forms of existence, by executing variations of existing strategies and recording the evidence of their relative success. For example, the southern toad will learn through experience to avoid eating bumblebees (Tinbergen 1965), and this learned behaviour involves an evidence-based update of the toad's internal models regarding edible insects.

Such refinements of internal learned mental models lead to more fine-tuned and less ignorant behaviour or, equivalently, to lower-entropy behaviour. We may thus consider the domain of neural selection to represent another case of nature using evolutionary systems to counter the entropic process.

3.5 Cultural Selection

Homo sapiens is unique, among the species that evolution has brought into existence on our planet, in producing an evolving culture. While a handful of other species display some rudimentary cultural traits, only our species participates in a culture that accumulates new knowledge in a continuous fashion. Even our nearest and now extinct relatives in the genus *Homo* appear to have maintained a relatively static kit of artefacts over the duration of their species. *Homo sapiens*, on the other hand, in less than a hundred thousand years succeeded in migrating out of Africa, colonizing practically the entire planet, setting foot on the moon, and placing robotic instruments throughout the solar system.

All these accomplishments have been due to human cultures, which have adapted through the continuous accumulation of new shared knowledge. This cultural knowledge has orchestrated the behaviour of group members into a new level of adapted behaviour, which has in turn tended to promote cultural success (i.e. the propagation of cultural adaptations). The domain of cultural selection, in other words, continues a pattern that should by now seem familiar, as it is repeated throughout the universal Darwinism framework: the substrate of the preceding domain (here, neural-based leaned behaviour) is orchestrated towards the production of new kind of knowledge repository, thus defining the boundary of a new selective domain (here, the cultural domain). From this perspective, we are well positioned to unravel many mysteries surrounding the emergence of the cultural knowledge repository and to understand, at least in broad outline, how it has emerged.

Cultural knowledge is rooted in the neural-based social learning we share with our primate relatives and which, in turn, has been enabled through the biological evolution of structures such as neurons, brains, and specialized neural circuitry. In particular, *Homo sapiens* inherited biological and neural sensory abilities evolved by our animal ancestors since the Cambrian explosion. These abilities have been honed and fine-tuned over a vast evolutionary period, to provide the brain with accurate and reliable information about what is going on in the outside world. In addition to our more primitive sensory brain regions, we possess a uniquely enormous cerebral cortex, constituting roughly 80% of our brain mass. This elaborate evolutionary addition to our brain has enabled our species to engage in novel and relatively powerful types of imaginative thinking; we can consider a plethora of possible explanations for outcomes in the world, and many possible ways in which these outcomes could be related. The possibilities we can imagine include those we cannot verify by our senses: we can think about angels and demons, and about scientific puzzles which we are currently unable to conclusively solve. Such imaginative thoughts can be compelling but are, alas, prone to inaccuracy; we can easily imagine things which are untrue.

However, these two functional areas of our brain – the sensory and the imaginative – may interact synergistically, whereby speculations generated in our cerebral cortex are judged and selected using sensory information. Such synergy would constitute essentially Bayesian, evidence-based hypothesis selection. It appears likely that the brain's pervasive neural architecture of cortical columns is a specific adaptation designed to perform this type of Bayesian computation, on inputs originating from the senses and the cerebral cortex (Cain et al. 2016). We suggest that this synergy is what powers cultural evolution.

The imaginative process produces variations on cultural beliefs and practices which may be inherited through social learning or imitation, with selection being made among these variations based on the sensory evidence. Once again, that is, we see an equivalence between the Bayesian and Darwinian formulations of evolutionary dynamics. For example, archaeologists are increasingly using techniques developed within evolutionary biology, such as evolutionary trees or cladistics, to describe the phylogeny of cultural artefacts such as arrowheads. Some have described this program of placing archaeological artefacts within a biologically inspired phylogeny as the 'fundamental issue of evolutionary archaeology' (O'Brien and Lyman 2003). This research suggests that cultures experiment with many variations on artefact design, and that evidence about design effectiveness is retained within the repository of cultural knowledge in a Darwinian manner, similar to natural selection.

Nowhere is the Darwinian/Bayesian nature of cultural practice clearer than in the case of science. In science, alternative hypotheses are generated by our imaginative minds, and experimental tests are devised to provide empirical, sensory evidence that becomes the basis for selecting among these hypotheses. Through science, cultural evolution accumulates evidence-based knowledge and thus reduces ignorance or entropy. We thus regard culture as the most recently emerging domain of nature in which accumulated knowledge serves to defy the spirit of the second law.

3.5.1 Written Language Enabled the Cultural Knowledge Repository to Come into Its Own

Early ancestral members of our species lived in small hunter gatherer societies, sharing a common body of cultural knowledge but lacking a great deal of specialization among its members. Social learning was facilitated between members and between generations using a complexly descriptive spoken language. At this early stage of societal evolution, the cultural knowledge repository was constrained to the sum of the cultural knowledge possessed by living individual members. As agriculture provided the means for larger groups with greater specializations, however, individualized knowledge became a constraint on an evolving, widely shared body of cultural knowledge. Human memory is fallible and the recording of important historical facts, such as resource accounting, required a more objective, robust and widely accessible knowledge repository. This was achieved by transferring parts of the cultural knowledge repository to the external medium of written information, where it could be accessed by multiple members and survive the mortality of individual knowledge-bearers. The earliest-known example of such writing are cuneiform clay tablets, used in Mesopotamia for accounting purposes at least 5000 years ago.

With the invention of the printing press and numerous other knowledge repositories such as those involved with computers, our external cultural knowledge repository is growing at an exponential rate. This fairly well-documented transition of the cultural knowledge repository from internal models within the minds of individuals to shared externally recorded models illustrates the dynamic evolutionary transitions which knowledge repositories can undergo.

3.5.2 How Bayesian Is Cultural Evolution?

We may view cultural evolution as a largely Bayesian process, where hypotheses concerning possible forms of cultural existence are tested by evidence. There is a vast range of cultural practices, however, between those that operate in an ‘approximate’ Bayesian manner and those that follow Bayesian mathematics more exactly.

For example, the evolution of agriculture seems likely to have been a largely Bayesian process, with many different practices designed to increase food production being attempted and those that have worked being preferentially selected. In this manner, humans have been influencing the evolutionary fitness of many species for at least tens of thousands of years. Cultural practices have apparently influenced the phylogenetic history of cereal crop plants, for instance, for over 30,000 years (Allaby et al. 2017).

Over agriculture’s long history, humans have learned how to exert increasingly greater control over the variables influencing crop productivity. Agriculture has thus evolved to become a more exact Bayesian method in the minds of its practitioners, as they have learned to juggle the many variables under their control with the intent of increasing crop yields. Recent refinements in agricultural techniques have allowed human culture to achieve breakthroughs in biochemical designs, which have overcome limitations imposed by the biology of crop species. For example, unaided biology does not appear to possess an enzyme capable of gene substitution in an intact DNA strand, but this feat has recently been accomplished by researchers applying the agricultural principles of artificial selection to the chemical evolution of enzymes in a process named *DNA-Templated Organic Synthesis* (DTS) (Li and Lui 2004). As these researchers note, 'such an enormous sequence space can in principle be explored efficiently by iterated cycles of DTS-based translation, selection for desired binding or catalytic properties, template amplification by PCR, and template diversification by mutagenesis or recombination, representing a true evolutionary process'. It is clear, in other words, that such methods of modern agriculture – and of scientific practice more broadly – employ the Darwinian notion of a 'true evolutionary process'.

However, processes of cultural evolution may often appear to be only approximately Bayesian. As noted above, the human mind is highly fallible as a scientific apparatus, and what constitutes evidence for Bayesian updates of cultural models will often be unclear to members of that culture. Ultimately, however, evidence about how well a practice contributes to cultural success will influence the fate of that culture, even if not easily calculated by its members themselves. In other words, whereas nature is precisely Bayesian in how it weighs the evidence (e.g. about the influence of an agricultural practice on a culture's migratory success), this process may be perceived as more opaque and approximate to members of the culture itself. It is nevertheless human agency which determines the evolutionary trajectory: our ability to imagine and test the innovative possibilities, and to select and retain those which contribute to cultural success, will determine the fate of our own cultures.

Only with the emergence of science have humans come to fully understand the power of the Bayesian synergy that is performed by the sensory and imaginative components of our brains. Indeed, Bayesian probability has been hailed as 'the logic of science' (Jaynes 2003). All cultural knowledge may be typified as being produced by inferential systems, but science has vastly accelerated the process of cultural evolution because it aspires to a more precise, less approximate form of Bayesian inference. Scientific hypotheses are constructed to be testable by evidence, which is gathered through experiments that have been carefully designed and conducted for exactly that purpose. This ensures that the Bayesian likelihood function, which describes how well each competing hypothesis is supported by the evidence, is relatively straightforwardly assessed and easily computed.

In cultural practices that are less explicitly scientific – for example, early agriculture or arrowhead design – competing hypotheses about innovation effectiveness, and the ways in which evidence should be brought to bear, may be less clear. If an innovation succeeds, for example, was it due to the practitioner's status, or to their skill, or to the type of incantation they used? Science brings to the forefront of our consciousness the synergistic potential between imagination and evidence, unleashing a whole new level of human knowledge accumulation that better conforms to nature's timeless method of precise Bayesian inference.

The algorithmic and autopoietic nature of cultural knowledge has been noted by many, but perhaps none as poetically as George Spencer-Brown (1979):

> It may be helpful at this stage to realize that the primary form of mathematical communication is not description, but injunction. In this respect it is comparable with practical art forms like cookery, in which the taste of a cake, although literally indescribable, may be conveyed to a reader in the form of a set of injunctions called a recipe. Music is a similar art form, the composer does not even attempt to describe the set of sounds he has in mind, much less the set of feelings occasioned through them, but writes down a set of commands which, if they are obeyed by the reader, can result in a reproduction, to the reader, of the composer's original experience ... Even natural science appears to be more dependent upon injunction than we are usually prepared to admit. The professional initiation of the man of science consists not so much in reading the proper textbooks, as in obeying injunctions such as 'look down this microscope'.

The individual in a cultural context is seldom left to puzzle things out on their own, but rather is 'taught' or 'instructed' in the algorithms of cultural practices. Spencer-Brown provides some examples above, and we may also recognize their pertinence to the examples we have previously explored such as arrowhead design ('always start flaking the blank like this') and agriculture ('always plant this crop on the first full moon after the spring equinox'). Such algorithms, when faithfully followed, produce outputs of cultural adaptations that tend to propagate themselves successfully.

The idea that many cultural practices may be considered computational is made evident by the extent to which computers can 'automate' them. This automation encompasses not only business, communications, and manufacturing but extends even to games such as chess and go. In particular, the ability of computation to mimic the evolutionary Darwinian/Bayesian paradigm, through an exponentially growing number of applications involving evolutionary computation, has revolutionized the practice of science (Ignatova et al. 2008).

4 Summary and Conclusion

In this chapter we have endeavoured to show how, across all the domains of nature, the process of Darwinian selection acts to creatively generate complex order, against the antagonistic tide of the second law. In all these domains, the selective process is characterized by the same key components. Each domain is defined by its own knowledge repository, a probabilistic model which puts forward guesses – instantiated in the form of an embodied adapted system – about what it takes to exist in that domain. The guesses are tested for accuracy by the embodied adapted system, and test results are fed back to the knowledge repository. The knowledge repository is then updated for greater accuracy in a Bayesian manner, thus reducing its ignorance and entropy. These entropy-reducing inferential systems operate in the context of 'variance–inheritance–selection' Darwinian dynamics: some knowledge repositories are smarter than others, so they produce more copies of themselves and thus competitively outperform their more ignorant rivals.

We propose that these Darwinian inferential systems are responsible for all of the complex order that exists in the five natural domains of the cosmological, quantum, biological, neural, and cultural. These domains exist in nested hierarchy, and those higher (i.e. less fundamental) in the hierarchy can be considered subroutines of those that are lower. Each domain's knowledge repository computationally transforms the substrate of the lower domain, to generate novel forms of rebellion against the second law.

Overall, this universal Darwinism framework can be seen as jigsaw puzzle that, while still incomplete, has already taken recognizable shape. Some important pieces are clearly still missing, but it is complete enough to provide good ideas of how the missing pieces will be shaped, and of what the completed puzzle is going to look like. For example, with regard to the knowledge repositories in the cosmological and quantum domains, we are in a position similar to that of biologists in the decades preceding the discovery of DNA: although we are yet to directly observe the physical forms of these knowledge repositories, we have good reasons to expect that these forms do exist, and can speculate reasonably about what they might look like. It will be exciting to watch as the puzzle implied by universal Darwinism becomes increasingly complete in the years ahead. We expect that as progress towards completion is made, there will be increasing support for the notion that as an explanation for the origins of complex order, Darwin's theory is far more fundamental than has generally been appreciated.

Acknowledgements We would like to thank Karl Friston for a number of comments and suggestions which have been incorporated into this chapter.

References

Allaby, R. G. et al. (2017). Geographic mosaics and changing rates of cereal domestication. The Royal Society.

Arvey R. & Colarelli S. M. (Eds). (2015). *Biological foundations of organizational behavior*. University of Chicago Press.

Bastos A. M. et al. (2012). Canonical microcircuits for predictive coding. Neuron, Vol. 76(4).

Bishop, E. (1967). Foundations of Constructive Analysis. New York: Academic Press. ISBN 4-87187-714-0.

Brenner, S. (1999). Theoretical biology in the third millennium. Philosophical Transactions of the Royal Society.

Brydon, E. (2012). In massive genome analysis ENCODE data suggests 'gene' redefinition. Cold Spring Harbor Laboratory. Downloaded 2nd Dec. 2013 from http://www.cshl.edu/Article-Gingeras/massive-genome-analysis-by-encode-redefines-the-gene-and-sheds-new-light-on-complex-disease

Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L., & Wakefield, J. C. (1998). Adaptations, exaptations, and spandrels. *American Psychologist*, *53*(5), 533.

Cain N. et al. (2016). The Computational Properties of a Simplified Cortical Column Model. PLOS computational biology, Vol. 12(9).

Campbell, D. T. (1974). Evolutionary epistemology. In P.A. Schilpp (Ed.), *The Philosophy of Karl R. Popper*. The Library of Living Philosophers. Lasalle, IL: Open Court Publishing Company, Volume 14-1, 413–463.

Campbell, J. O. (2010). Quantum Darwinism as a Darwinian process. arXiv preprint.

Campbell, J. O. (2015). Darwin does physics. s.l.: CreateSpace.

Campbell, J. O. (2016). Universal Darwinism as a process of Bayesian inference. Front. Syst. Neurosci. doi: https://doi.org/10.3389/fnsys.2016.00049

Comte, A. (1848). *A General View of Positivism*. Translation by Bridges, J. H. (1865), Trübner and Co.; reissued by Cambridge University Press (2009; ISBN 978-1-108-00064-2).

Crane, L. (1994/2010). Possible implications of the quantum theory of gravity: an introduction to the meduso-anthropic principle. arXiv:hep-th/9402104v1. Reprinted in Crane, L. (2010). Possible implications of the quantum theory of gravity: an introduction to the meduso-anthropic principle. *Foundations of Science*, vol. 15, pp. 369–373.

Dalvit, D. A. R., Dziarmaga, J. & Zurek, W. H. (2005). Predictability sieve, pointer states, and the classicality of quantum trajectories. Phys. Rev. A, Vol. 72, 062101.

Darwin, C. R. (1859). *On the Origin of Species*. London: John Murray.

Davies, P. C. W. & Walker, S. I. (2012). The Algorithmic Origins of Life. Journal of the Royal Society, Vol. 10.

Dawkins, R. (1982). The Extended Phenotype: The Long Reach of the Gene. Oxford University Press.

De Jong, K. A. (2006). Evolutionary computation: a unified approach. Cambridge MA: MIT Press.

Dennett, D. (1995). *Darwin's Dangerous Idea*. New York: Schuster.

Deutsch, D. (1985). Quantum theory, the Church-Turing principle and the universal quantum computer. Proceedings of the Royal Society of London, Vol. A 400 pg. 97–117.

Deutsch, D. (1997). *The Fabric of Reality*. London: Penguin.

Deustch, D. (2011). *The Beginning of Infinity*. London: Penguin.

Dirac, P.A.M. (1958). Quantum Mechanics. Oxford: Clarendon Press.

Fernando, C, Szathmary, E & Husbands, P. (2012). Selectionist and evolutionary approaches to brain function: a critical appraisal. Computational Neuroscience.

Fischetti, M. (2011). Computers versus Brains. Scientific American. November 1, 2011.

Frank, R. H. (2012a). *The Darwin economy: Liberty, competition, and the common good*. Princeton University Press.

Frank, S. A. (2012b). Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory. Journal of Evolutionary Biology, Vols. 25: 2377–2396.

Friston, K. J. (2013). Life as we know it. Journal of the Royal Society Interface, Vol. 10: 20130475.

Friston K. J. et al. (2017). Active Inference: A process theory. Neural Computation, Vol. 29(1).

Gardner, A. & Conlon, J. P. (2013). Cosmological natural selection and the purpose of the universe. *Complexity* 18: 48–56.

Gardner, J. N. (2000). The selfish biocosm. *Complexity* 5: 34–45.

Gavrilov, Z. (2012). Quantum computation. MIT. Downloaded 5th Dec. 2017 from http://web.mit.edu/zoya/www/quantComp.pdf

Gibson M. A. & Lawson, D. W. (Eds). (2014). *Applied evolutionary anthropology: Darwinian approaches to contemporary world issues.* Springer.

Gillespie, D. T. A. (1970). Quantum Mechanical Primer. Scranton: International Textbook Company.

Gould, S. J. & Vrba, E. S. (1982). Exaptation – a missing term in the science of form. *Paleobiology* 8 (1): 4–15.

Hamilton, W. (1964). The genetical evolution of social behaviour I. *Journal of Theoretical Biology*, vol. 7, no. 1, pp. 1–16.

Hardy, L. (2013). Are quantum states real? International Journal of Modern Physics B.

Harrison, E. R. (1995). The natural selection of universes containing intelligent life. *Quarterly Journal of the Royal Astronomical Society* 36: 193–203.

Heylighen F. (1999): “The Growth of Structural and Functional Complexity during Evolution”, in: F. Heylighen, J. Bollen & A. Riegler (eds.) *The Evolution of Complexity* (Kluwer Academic, Dordrecht), p. 17–44.

Hoehn, P. A. (2017a). Quantum theory from questions. Phys. Rev., Vol. A 95, 012102.

Hoehn, P. A. (2017b). Quantum theory from rules on information acquisition. Entropy, Vol. 19(3), 98.

Ignatova, Z., Marinez-Perez, I. & Zimmermann, K. (2008). DNA Computing Models. Springer. ISBN 978-0-387-73635-8.

Jaynes, E. T. (1957). Information Theory and Statistical Mechanics I., Phys. Rev., 106, 620.

Jaynes, E. T. (1965). Gibbs vs Boltzmann Entropies. Am. J. Phys., 391.

Jaynes, E. T. (2003). Probability Theory: The Logic of Science. University of Cambridge Press.

Landauer, R. (1996). The physical nature of information. Physics Letters A, Vol. 217. https://doi.org/10.1016/0375-9601(96)00453-7

Lee, T. S. (2015). The visual system’s internal model of the world. Proceedings of the IEEE. Institute of Electrical and Electronics Engineers, Vols. 103(8), 1359–1378.

Li, X. & Lui, D. (2004). DNA-Templated Organic Synthesis: Nature’s Strategy for Controlling Chemical Reactivity Applied to Synthetic Molecules. Angew. Chem. Int. Ed., Vol. 43, 4848.

Lloyd, S. (2007). Programming the Universe. Vintage; Reprint edition.

Lloyd, S. (2013). The Universe as Quantum Computer. Arxiv preprint.

Maldacena, J. M. (1998). The large N limit of superconformal field theories and supergravity. Adv. Theor. Math. Phys, Vol. 2, pp. 231–252.

Markopulou, F. (2012). The computing spacetime. [ed.] S. B. Cooper, et al. How the world computes. Berlin, Hei: Springer.

McGregor, S. (2012). Evolution of Associative Learning in Chemical Networks. Comput Biol 8(11). doi:https://doi.org/10.1371/journal.pcbi.1002739

Nelson, R. R. & Winter, S. G. (1982). An evolutionary theory of economic change. Harvard University Press. ISBN 0-674-27228-5.

O’Brien, M. J. & Lyman R. L. (2003). Resolving Phylogeny: Evolutionary Archaeology’s Fundamental Issue. [book auth.] T. L. VanPool and C. S. VanPool. Essential Tensions in Archaeological Method and Theory. Salt Lake City: University of Utah Press, pp. 115-135.

Orr, H. A. (2000). Adaptation and the cost of complexity. *Evolution*, *54*(1), 13–20.

Orus, R. (2014). A Practical Introduction to Tensor Networks: Matrix Product States and Projected Entangled Pair States. Annals of Physics 349, 117–158.

Pietraszewski, D. Curry, O.S., Petersen, M.B., Cosmides, L. & Tooby, J. (2015). Constituents of political cognition: Race, party politics, and the alliance detection system. *Cognition,* 140. 24–39.

Pinker, S. & Bloom, P. (1992). Natural language and natural selection. In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, J. Barkow, L. Cosmides, and J. Tooby, Eds. pp. 451–493. Oxford University Press.

Popper, K. (1972). Objective Knowledge. Clarendon Press.

Price, M. E. (2017). Entropy and selection: Life as an adaptation for universe replication. *Complexity*, vol. 2017, Article ID 4745379, 4 pages, 2017. doi:https://doi.org/10.1155/2017/4745379

Pusey M F, Barrett J & Randolph T. (2012). On the reality of quantum states. Nature Physics 8, pp. 475–478.

Ramstead M. J. D., Badcock P. B., & Friston K. J. (2017). Answering Schrödinger’s question: A free-energy formulation. Phys Life Rev.

Richerson, P. J. & Boyd, R. (2005). *Not by genes alone: How culture transformed human* e*volution.* University of Chicago Press.

Rovelli, C. (1996). Relational Quantum Mechanics. International Journal of Theoretical Physics, Vols. 35, pp. 1637–78.

Saad G. (Ed). (2011). *Evolutionary Psychology in the Business Sciences*. Berlin: Springer**.**

Shannon, C. E. (1948). A mathematical theory of communications. *Bell System Technical Journal* 27(3): 379–423.

Smart, J. M. (2009). Evo devo universe? A framework for speculations on cosmic culture. In Cosmos and Culture: Cultural Evolution in a Cosmic Context, S. J. Dick and M. L. Lupisella, Eds., pp. 201–295, Government Printing Office, NASA SP-2009-4802, Washington, DC.

Smolin, L. (1992). Did the universe evolve? *Classical and Quantum Gravity*, vol. 9, no. 1, pp. 173–191.

Smolin, L. (1997). *The Life of the Cosmos*. New York: Oxford University Press.

Smolin, L. (2007). The status of cosmological natural selection. Arxiv preprint.

Spencer-Brown, G. (1979). The Laws of Form. New York: E. P. Dutton.

Stephens, C. R., 't Hooft, G. and Whiting, B. F. (1993). Black hole evaporation without information loss. Classical and quantum gravity Volume 11, Number 3.

Swingle, B. & Van Raamsdonk, M. (2014). Universality of Gravity from Entanglement. ArXiv preprint.

't Hooft, G. (2016). *The cellular automaton interpretation of quantum mechanics*. Springer.

Thayer, B. A. (2004). *Darwin and international relations: On the evolutionary origins of war and ethnic conflict*. University Press of Kentucky.

Tinbergen, N. (1965). Animal behaviour. Time Inc.

Tomer R. et al. (2010). Profiling by image registration reveals common origin of annelid mushroom bodies and vertebrate pallium. Cell, Vol. 142 (5): 800–809.

Tooby J. & Cosmides L. (1992). The psychological foundations of culture. In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, J. Barkow, L. Cosmides, and J. Tooby, Eds. Oxford University Press.

Verlinde, E. (2010). On the Origin of Gravity and the Laws of Newton., arXiv:1001.0785.

Vidal, C. (2014). *The Beginning and the End: The Meaning of Life in a Cosmological Perspective*. Springer.

Wigner, E. P. (1960). "The unreasonable effectiveness of mathematics in the natural sciences. Richard Courant lecture in mathematical sciences delivered at New York University, May 11, 1959". *Communications on Pure and Applied Mathematics* 13: 1–14.

Williams, G. C. (1966). *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton University Press.

Zurek, W. H. (1998). Decoherence, Einselection and the Existential Interpretation. Philosophical Transactions: Mathematical, Physical and Engineering Sciences, pp. 1793–1821.

Zurek, W. H. (2007). Relative States and the Environment: Einselection, Envariance, Quantum Darwinism, and the Existential Interpretation. arXiv:0707.2832v1.

Zurek, W. H. (2009). Quantum Darwinism. Nature Physics, vol. 5, pp. 181–188.

Zurek, W. H. (2014). Quantum Darwinism, classical reality and the randomness of quantum jumps. Physics Today, Vols. 67, 10, 44.

Zurek, W. H. & Zwolak, M. (2013). Complementarity of quantum discord and classically accessible information. Scientific Reports 3, Article number: 1729. doi:https://doi.org/10.1038/srep01729

Zylberberg, A., et al. (2011). The human Turing machine: a neural framework for mental programs. Trends Cogn Sci 15: 293–300.

Part III
Biology

Complexity, Development, and Evolution in Morphogenetic Collective Systems

Hiroki Sayama

1 Introduction

Various living and nonliving systems are collective systems in the sense that they consist of a large number of smaller components. Those microscopic components interact with each other to show a wide variety of self-organizing macroscopic structures and behaviors, which have been subject to many scientific inquiries (Bar-Yam 1997; Ben-Jacob et al. 1998; Parrish et al. 1999; Solé and Goodwin 2000; Macy and Willer 2002; Camazine et al. 2003; Couzin and Krause 2003; Gershenson 2007; Lämmer and Helbing 2008; Turner and Soar 2008; Turner 2011; Vicsek and Zafeiris 2012; Portugali 2012; Doursat et al. 2012; Fernández et al. 2014; Sayama 2015).

Typical assumptions often made in earlier mathematical/computational models of self-organizing collectives include the homogeneity of individual components' properties and behavioral rules within a collective. Such homogeneity assumptions have merit in simplifying models and allowing for analytical prediction of the models' macroscopic behaviors. However, such homogeneity assumptions would not be adequate to capture more complex nature observed in real-world complex systems, such as multicellular organisms' morphogenesis and physiology (Solé and Goodwin 2000; Camazine et al. 2003), termite colony building and maintenance (Turner and Soar 2008; Turner 2011), and growth and self-organization of human social systems (Macy and Willer 2002; Lämmer and Helbing 2008). Those real-world complex collectives consist of heterogeneous components whose behavioral types can change dynamically via active information exchange among locally

---

H. Sayama (✉)
Center for Collective Dynamics of Complex Systems, Department of Systems Science and Industrial Engineering, Binghamton University, State University of New York, Binghamton, NY, USA
e-mail: sayama@binghamton.edu

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_11

connected neighbors. These properties of components facilitate self-organization of highly nontrivial morphological structures and behaviors (Sayama 2014).

In this chapter, we present a brief summary of our recent effort in investigating several aspects of complex morphogenetic collective systems that involve (1) heterogeneous components, (2) dynamic differentiation/re-differentiation of the components, and (3) local information sharing among the components. Our objective was to understand the implications of each of those properties for developmental processes of the collectives and to develop effective methodologies to design novel artificial morphogenetic collective systems.

The rest of this chapter is structured roughly following the topics of this proceedings volume—*evolution, development, and complexity*—though we will discuss them in a reversed order. We will first propose a classification scheme of several distinct complexity levels of morphogenetic collective systems based on their components' functionalities. Then we will computationally investigate how the developmental processes, i.e., self-organization of morphological patterns created by interacting components, will be affected by the difference in the complexity levels of those systems. Finally, we will discuss evolutionary methods to design nontrivial self-organization of morphogenetic collective systems, with a brief additional remark on their robustness/sensitivity to spatial dimensional changes.

2 Functional Complexity Levels of Morphogenetic Collective Systems

Our first task is to identify what kind of properties are typically seen in real-world complex collective systems but often omitted for simplicity in the literature on mathematical/computational models of those systems. In Sayama (2014), we selected the following three as the key properties essential for self-organization of morphogenetic collective systems yet often ignored in the literature:

1. Heterogeneity of components
2. Differentiation/re-differentiation of components
3. Local information sharing among components

Heterogeneity of components means that there are multiple, distinct types of components whose behaviors are different from each other. Note that these types are not necessarily a simple rewording of dynamical states. Instead, each type may have multiple dynamical states within itself, while its behavioral rules as a whole (e.g., state-transition rules) should be different from those of other types. Examples include different cell types within an organism, individuals with different phenotypical traits in a colony of social insects, and different professions of individuals in human society. Differentiation/re-differentiation means that each individual component will assume one of those types (differentiation) and potentially switch from one type to another under certain conditions (re-differentiation). Finally, local information sharing means that the individual components are actively sending/receiving encoded signals among them for coordination of their collective behaviors, such as cell-cell communication with molecular signals, pheromone-based communication among social insects, and human communication in languages.

Mathematically speaking, distinguishing presence/absence of each of these three properties would define a total of $2^3 = 8$ possible classes of collective systems. However, we claim that there are some hierarchical relationships among those three properties. Specifically, differentiation/re-differentiation of components require, almost tautologically, the multiple possibilities of component types. Furthermore, we assumed that local information sharing would make sense only if the components had an ability to change their types dynamically based on the received information.[1] Taking these requirement relationships into account, we proposed the following four hierarchical classes of complexity levels of morphogenetic collective systems (Sayama 2014) (Fig. 1):

Class A Homogeneous collective
Class B Heterogeneous collective
Class C Heterogeneous collective with dynamic (re-)differentiation
Class D Heterogeneous collective with dynamic (re-)differentiation and local information sharing

The dynamics of components in each of these four classes can be represented mathematically as follows (Sayama 2014):

Class A $a_i^{t+1} = F(o_i^t)$

Class B $a_i^{t+1} = F(s_i, o_i^t)$

Class C $a_i^{t+1} = F(s_i^t, o_i^t),\ s_i^{t+1} = G(s_i^t, o_i^t)$

Class D $a_i^{t+1} = F(S_i^t, O_i^t),\ s_i^{t+1} = G(S_i^t, O_i^t),$
$S_i^t = \{s_m^t \mid m \text{ in } N_i^t\},\ O_i^t = \{o_m^t \mid m \text{ in } N_i^t\}$

Here a_i^t, o_i^t, and s_i^t are individual component i's behavior, observation, and type at time t, respectively (s_i is a time-invariant type of component i); F and G are model functions; and N_i^t is the set of component i's neighbors at time t. These mathematical formulations help clarify the hierarchical relationships among the four complexity levels. Following these formulations, we will construct a specific computational model of morphogenetic collective systems to facilitate systematic investigation of the proposed four complexity levels and their characteristics.

---

[1] We note that this assumption is much less obvious than the first one, and if we did not adopt it, we would obtain $3 \times 2 = 6$ different classes. In this chapter, we limit our focus on the four-level classification presented above.

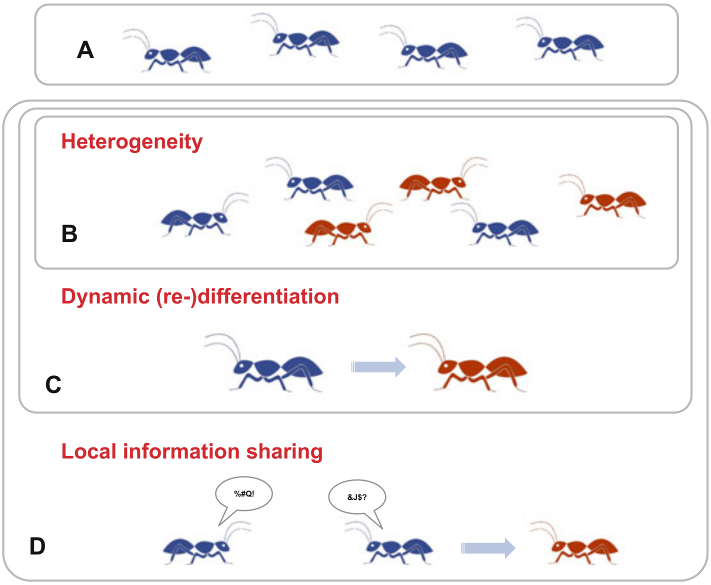


Fig. 1 Proposed four levels of complexity of morphogenetic collective systems. (**a**) Homogeneous collective. (**b**) Heterogeneous collective. (**c**) Heterogeneous collective with dynamic (re-)differentiation. (**d**) Heterogeneous collective with dynamic (re-)differentiation and local information sharing. These four classes form a hierarchical level structure; see text for details

3 Developmental Models: Morphogenetic Swarm Chemistry

We utilized our earlier "Swarm Chemistry" model (Sayama 2009, 2012a) to construct a new computational model of morphogenetic collective systems. Swarm Chemistry is a revised version of Reynolds' well-known self-propelled particle swarm model known as "Boids" (Reynolds 1987). In Swarm Chemistry, multiple types of components with different kinetic behavioral parameters are mixed together. Their behavioral parameters are represented in a "recipe" as shown in Fig. 2. Therefore, the Swarm Chemistry model is already capable of representing both Class A (homogeneous) and Class B (heterogeneous) collective systems. In Swarm Chemistry, components with different types spontaneously segregate from each other even without any sophisticated sensing or control mechanisms, often forming very intricate self-organizing dynamic patterns (Sayama 2009, 2012a).

To make individual components capable of dynamic differentiation/re-differentiation and local information sharing, we made several extensions to Swarm Chemistry (Sayama 2014). First, we made each individual component able to obtain information about its own dynamical type and its local environment in the form of *observation vector o* (Fig. 3) and then utilize this vector to decide which dynamical type it should assume. This allows for dynamic (re-)differentiation required for Class C/D collective systems. This decision-making process was

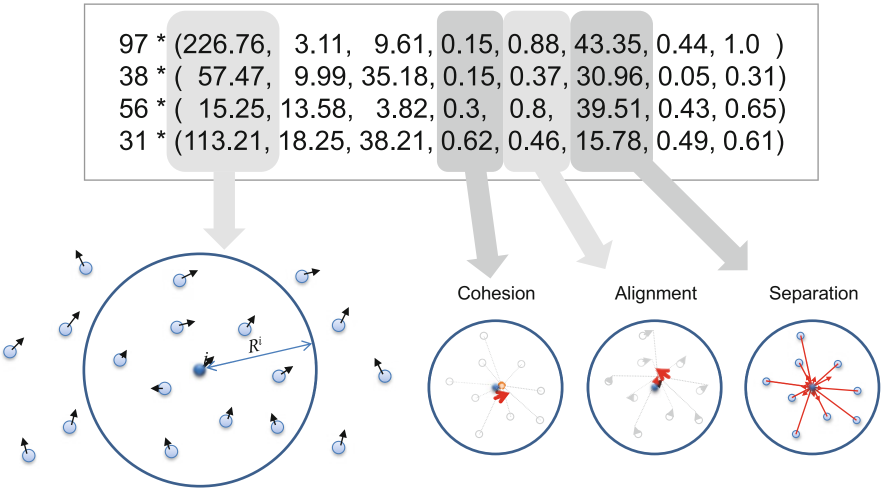


Fig. 2 Encoding of behavioral parameters in a *recipe* in Swarm Chemistry. A recipe is a list of parameter values written in the format "*number of particles* * (*parameter values for behaviors of those particles*)." The parameters include the radius of interaction (bottom left) and the strengths of three primary rules (cohesion, alignment, and separation; bottom right)

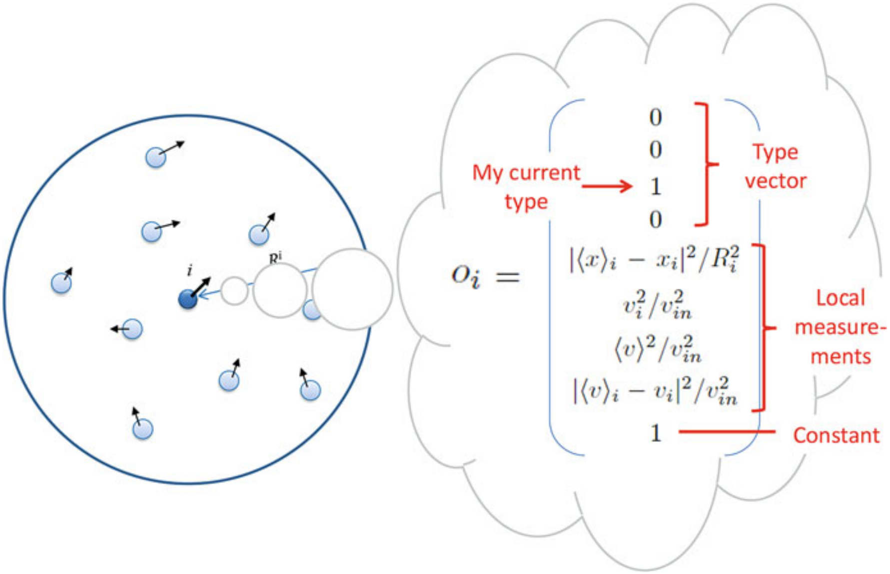


Fig. 3 Observation vector *o* of each particle used in Morphogenetic Swarm Chemistry. The first several values of *o* encode the current type of the particle, while the rest captures the measurements of its local environment. A constant unity is also included at the end of the vector

Table 1 Parameterization of four complexity levels of Morphogenetic Swarm Chemistry models

Class	Recipe	U	w
A	Single-type	0	0
B	Multiple-type	0	0
C	Multiple-type	$\neq 0$	0
D	Multiple-type	$\neq 0$	$\neq 0$

implemented via multiplication of *preference weight matrix* U to the observation vector o, so that letting $U = 0$ represents Class A/B systems as well. The second model extension was to introduce *local information sharing coefficient* w, with which the actual input vector multiplied by U was calculated as the weighted average between the component's own observation vector and the local average of all the observation vectors of neighbor components. Changing the value of w represents switching between Class C and Class D collective systems. With these, the four complexity levels discussed in the previous section were fully parameterized as shown in Table 1. This expanded model is called "Morphogenetic Swarm Chemistry" hereafter. More details can be found in Sayama (2014).

4 Differences of Developmental Processes Across Complexity Levels

We conducted a series of computational experiments using the Morphogenetic Swarm Chemistry model to investigate the differences of their developmental processes across the four complexity levels. This was conducted by detecting statistical differences in topologies and behaviors of self-organizing patterns that were collected via Monte Carlo simulations using randomly sampled parameter values. Topological and behavioral features of self-organizing patterns were measured using several kinetic metrics (average speed, average absolute speed, average angular velocity, average distance from center of mass, average pairwise distance) as well as newly developed network analysis-based metrics (Sayama 2015; Wasserman 1994; Barabási 2016) (number of connected components, average size of connected components, homogeneity of sizes of connected components, size of largest connected component, average size of non-largest connected components, average clustering coefficient, link density) that were measured on a network reconstructed from the individual components' positions in space (Sayama 2014). These new metrics allowed us to capture topological properties of the collectives that would not have been captured by using simple kinetic metrics only.

Results showed significant differences in most of the metrics between the four different classes of morphogenetic collective systems (Sayama 2014). Specifically, heterogeneity of components had a strong impact on the system's structure and behavior, and dynamic differentiation/re-differentiation of components and local information sharing helped the system maintain spatially adjacent, coherent

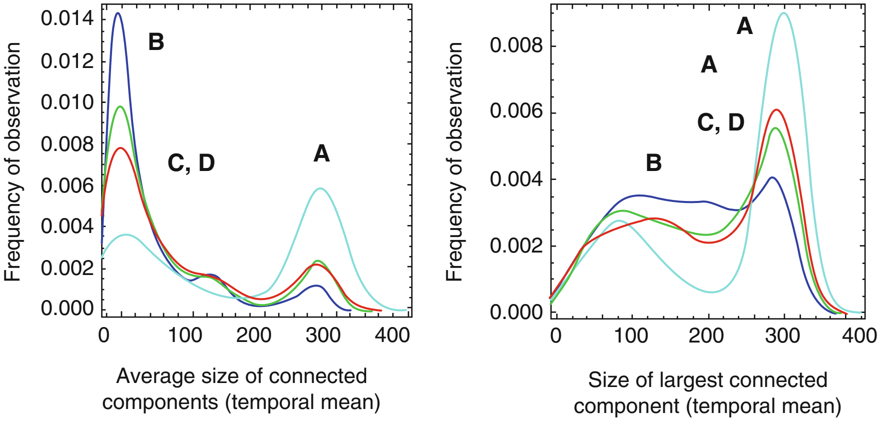


Fig. 4 Examples of experimental results showing clear differences of morphological properties among the four classes. Left: Distributions of the average size of connected components in generated morphologies. Right: Distributions of the size of the largest connected component in generated morphologies. In both plots, Classes C and D show intermediate distributions between those of Class A and Class B

organization. Statistical differences were particularly significant for topological features, demonstrating the effectiveness of our newly developed network analysis-based metrics. It was also observed that the properties of Class C/D collective systems tended to fall in between Class A and Class B in many metrics (Fig. 4). Moreover, it was noted that, as a byproduct, stochastic re-differentiation of components naturally realized a self-repair capability of self-organizing morphologies (Sayama 2012a, 2010).

As described above, straightforward statistical analysis placed the properties of Class C/D systems somewhere in between Class A and Class B, while it did not clarify whether Class C/D systems had any truly unique properties different from Classes A or B. Therefore, we conducted more in-depth, meta-level comparative analysis of *behavioral diversities* between those four classes of morphogenetic collective systems (Sayama 2015a). Behavioral diversities were measured for each class by computing the approximated volume of behavior space coverage, the average pairwise distance of two randomly selected behaviors in the behavioral space, and the differential entropy (Cover 2012) of the smoothed behavior distribution. More details can be found in Sayama (2015a). Results indicated that the dynamic (re-)differentiation of individual components, which was unique to Class C/D systems, played a crucial role in increasing the diversity in possible behaviors of collective systems (Fig. 5). This new finding revealed that our previous interpretation that Class C/D systems would behave more similarly to Class A than to Class B was not quite accurate. Rather, the difference between Classes A/B and Classes C/D helped make more diverse collective structures and behaviors accessible, providing for a larger “design space” for morphogenetic collective systems to explore.

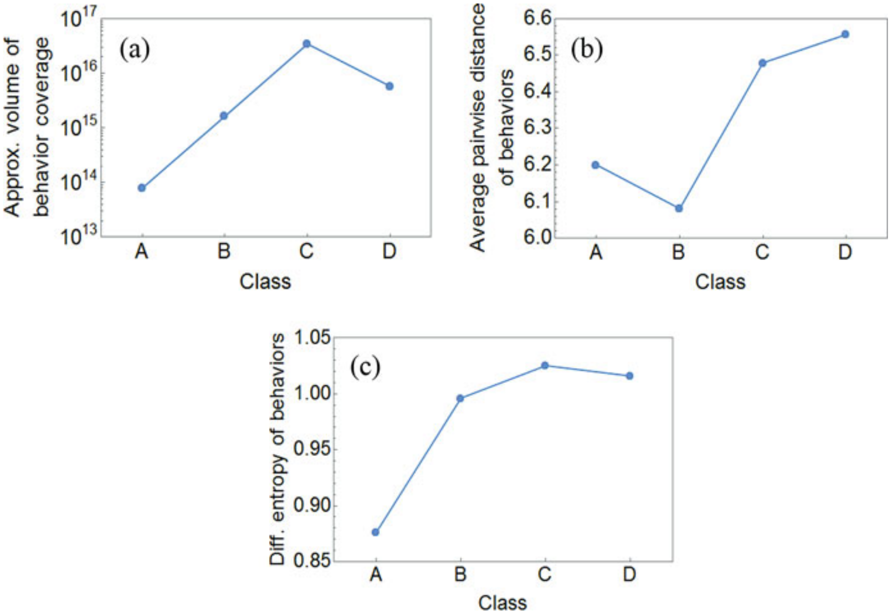


Fig. 5 Behavioral diversities of morphogenetic collective systems measured using three metrics: (**a**) approximated volume of behavioral coverage, (**b**) average pairwise distance of behaviors, and (**c**) differential entropy of behaviors. In all of the three plots, Classes C and D showed greater behavioral diversity than Classes A and B

5 Evolutionary Design of Morphogenetic Collective Systems

The remaining question we want to address is how to design novel self-organizing patterns of morphogenetic collective systems. Unlike conventional engineered systems for which clear design principles and methodologies exist, complex systems show nontrivial emergent macroscopic behaviors that are hard to predict and design from microscopic rules bottom-up (Braha et al. 2006). To design such systems, the evolutionary approach has been demonstrated to be one of the most effective means (Bar-Yam 2003; Sayama 2014b). Here we adopt two different evolutionary approaches: one is interactive evolutionary computation (IEC) (Takagi 2001; Sayama 2009; Bush and Sayama 2011; Sayama and Dionne 2015) and the other is spontaneous evolution within a simulated artificial ecosystem (Conrad and Pattee 1970; Sayama 2011a,b).

In the IEC approach, we developed a novel IEC framework called "Hyper-Interactive Evolutionary Computation (HIEC)" (Bush and Sayama 2011; Sayama and Dionne 2015), in which human users act not only as a fitness evaluator but also as an active initiator of evolutionary changes. HIEC was found to be highly effective in exploring the extremely high-dimensional design space of Swarm Chemistry, discovering a number of nontrivial, lifelike morphological patterns and dynamic

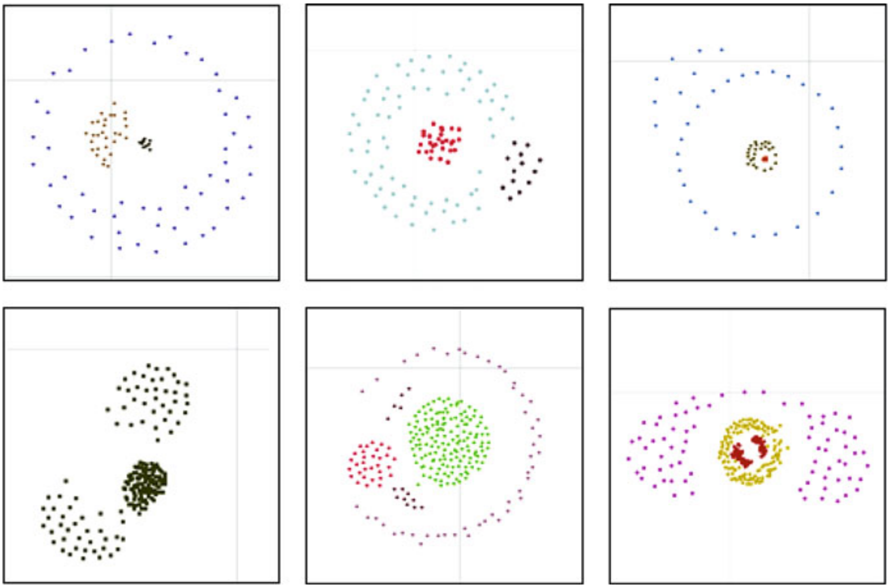


Fig. 6 Several examples of self-organizing lifelike patterns in Swarm Chemistry evolved using the interactive evolutionary computation approach

behaviors (Fig. 6).[2] We also found that these designed self-organizing patterns were remarkably robust against dimensional changes from 2D to 3D (Sayama 2012b) (Fig. 7), which is highly unique given that behaviors of complex systems generally depend heavily on spatial dimensions in which they develop.

Finally, in the spontaneous evolution approach, we replaced the human users in IEC with microscopic "physics laws" that would govern transmission of recipe information among individual components (as evolutionary operators acting at local scales) and macroscopic measurements of "interestingness" (as assessments of evolutionary processes at global scales) (Sayama 2011a,b). Specifically, recipe information was assumed to be transmitted between two colliding particles (with stochastic mutations possible at a small probability). The direction of transmission was determined by specific microscopic laws. These laws were perturbed globally at certain intervals to introduce variations and thus keep the evolutionary processes active and ongoing. The interestingness of evolution was measured by spatial structuredness (i.e., deviation from random homogeneous patterns) and temporal novelty production rates. More details can be found in Sayama (2011a,b). This spontaneous evolution approach was shown to be very powerful in continuously

---

[2] For more evolved patterns, see the Swarm Chemistry website: http://bingweb.binghamton.edu/~sayama/SwarmChemistry/

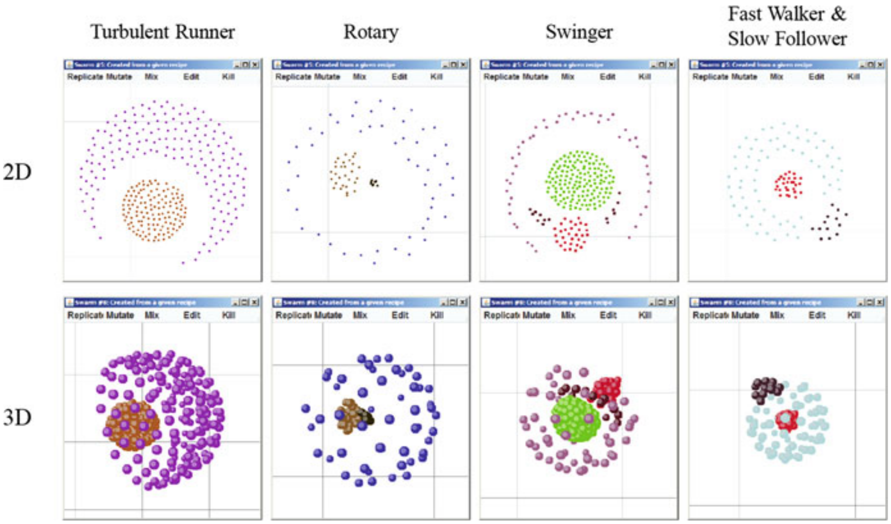


Fig. 7 Comparison of morphologies between 2D and 3D spaces, both developed from identical recipes

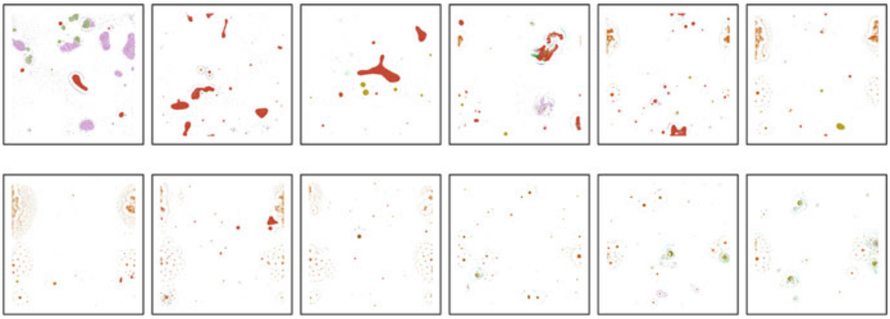


Fig. 8 Sample simulation run of Evolutionary Swarm Chemistry (from Sayama 2014b). Time flows from left to right (the bottom row follows the top one)

producing nontrivial morphologies. An example is given in Fig. 8, and other illustrative evolutionary processes can be found online.[3]

In the meantime, it was also noticed that evolutionary exploration was much less active in three-dimensional space than in two-dimensional one (Sayama 2012c), despite the robustness of self-organization against the same dimensional changes. This sensitivity was considered to be due to the fact that spontaneous evolution heavily relies on collisions between particles, which would become fundamentally less frequent in 3D space (Pólya 1921; Domb 1954).

---

[3] https://www.youtube.com/user/ComplexSystem/videos

6 Conclusions

In this chapter, we gave a condensed summary of our recent project that explored the complexity, development, and evolution of morphogenetic collective systems. The classification scheme of morphogenetic collective systems we proposed was among the first that focuses on functional and interactive capabilities of microscopic individual components. By orthogonalizing microscopic components' capabilities with macroscopic system behaviors, one can define a design space for various forms of morphogenetic collective systems, which will be useful for both classification of biological collectives and design of self-organizing artificial collectives.

The numerical simulation results obtained by using Morphogenetic Swarm Chemistry demonstrated that each of the characteristic properties of collective systems has unique, distinct effects on the resulting morphogenetic processes. Heterogeneity of components has quite significant effects on various properties of the collective systems, while the ability for individuals to dynamically switch their types contributes to the spatial coherence, the ability to self-repair, and the increase of behavioral diversity of those collective systems. Such behavioral richness would be the necessary ingredient for collective systems to evolve sophisticated structures and/or functions, which was partly demonstrated in the evolutionary approaches also discussed in this chapter.

This short chapter is obviously not sufficient to cover the whole scope of the project, which also produced several more application-oriented contributions that were not discussed here. Interested readers are encouraged to visit our project website.[4]

Acknowledgements This material is based upon work supported by the National Science Foundation under Grant No. 1319152. The author thanks Benjamin James Bush, Shelley Dionne, Craig Laramee, David Sloan Wilson, and Chun Wong for their contributions to this project.

References

Bar-Yam, Y.: Dynamics of Complex Systems. Addison-Wesley (1997)
Ben-Jacob, E., Cohen, I., Gutnick, D.L.: Cooperative organization of bacterial colonies: from genotype to morphotype. Annual Reviews in Microbiology 52(1), 779–806 (1998)
Parrish, J.K., Edelstein-Keshet, L.: Complexity, pattern, and evolutionary trade-offs in animal aggregation. Science 284(5411), 99–101 (1999)
Solé, R., Goodwin, B.: Signs of Life: How Complexity Pervades Biology. Basic Books (2000)
Macy, M.W., Willer, R.: From factors to factors: computational sociology and agent-based modeling. Annual Review of Sociology 28(1), 143–66 (2002)
Camazine, S., et al.: Self-Organization in Biological Systems. Princeton University Press (2003)
Couzin, I.D., Krause, J.: Self-organization and collective behavior in vertebrates. Advances in the Study of Behavior 32, 1–75 (2003)

---

[4] http://bingweb.binghamton.edu/~sayama/NSF-RI-MCS/

Gershenson, C.: Design and control of self-organizing systems. CopIt ArXives (2007)
Lämmer, S., Helbing, D.: Self-control of traffic lights and vehicle flows in urban road networks. Journal of Statistical Mechanics: Theory and Experiment 2008(04), P04019 (2008)
Turner, J.S., Soar, R.C.: Beyond biomimicry: What termites can tell us about realizing the living building. First International Conference on Industrialized, Intelligent Construction, Loughborough University (2008)
Turner, J.S.: Termites as models of swarm cognition. Swarm Intelligence 5(1), 19–43 (2011)
Vicsek, T., Zafeiris, A.: Collective motion. Physics Reports 517(3), 71–140 (2012)
Portugali, J.: Self-Organization and the City. Springer (2012)
Doursat, R., Sayama, H., Michel, O.: Morphogenetic engineering: Reconciling self-organization and architecture. Morphogenetic Engineering, Springer, pp. 1–24 (2012)
Fernández, N., Maldonado, C., Gershenson, C.: Information measures of complexity, emergence, self-organization, homeostasis, and autopoiesis. Guided Self-Organization: Inception, Springer, pp.19–51 (2014)
Sayama, H.: Introduction to the Modeling and Analysis of Complex Systems. Open SUNY Textbooks (2015)
Sayama, H.: Four classes of morphogenetic collective systems. Artificial Life 14: Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems, MIT Press, pp. 320–327 (2014)
Sayama, H.: Swarm chemistry. Artificial Life 15, 105–114 (2009)
Sayama, H.: Swarm-based morphogenetic artificial life. Morphogenetic Engineering: Toward Programmable Complex Systems, Springer, pp.191–208 (2012)
Reynolds, C.W.: Flocks, herds and schools: A distributed behavioral model. ACM SIGGRAPH Computer Graphics 21(4), 25–34 (1987)
Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications. Cambridge University Press (1994)
Barabási, A.-L.: Network Science. Cambridge University Press (2016)
Sayama, H.: Robust morphogenesis of robotic swarms. IEEE Computational Intelligence Magazine 5(3), 43–49 (2010)
Sayama, H.: Behavioral diversities of morphogenetic collective systems. Proceedings of the Thirteenth European Conference on Artificial Life (ECAL 2015), MIT Press, p. 41 (2015)
Cover, T.M., Thomas, J.A.: Elements of Information Theory. John Wiley & Sons (2012)
Braha, D., Minai, A.A., Bar-Yam Y.: Complex Engineered Systems. Springer (2006)
Bar-Yam, Y.: When systems engineering fails-toward complex systems engineering. IEEE International Conference on Systems, Man and Cybernetics 2003, IEEE, pp. 2021–2028 (2003)
Sayama, H.: Guiding designs of self-organizing swarms: Interactive and automated approaches. Guided Self-Organization: Inception, Springer, pp.365–387 (2014)
Takagi, H.: Interactive evolutionary computation: Fusion of the capabilities of EC optimization and human evaluation. Proceedings of the IEEE 89(9), 1275–1296 (2001)
Sayama, H., Dionne, S., Laramee, C., Wilson, D. S.: Enhancing the architecture of interactive evolutionary design for exploring heterogeneous particle swarm dynamics: An in-class experiment. Proceedings of the Second IEEE Symposium on Artificial Life (IEEE ALIFE 2009), IEEE, pp.85–91 (2009)
Bush, B. J., Sayama, H.: Hyperinteractive evolutionary computation. IEEE Transactions on Evolutionary Computation, 15, 424–433 (2011)
Sayama, H., Dionne, S. D.: Studying collective human decision making and creativity with evolutionary computation. Artificial Life, 21, 379–393 (2015)
Conrad, M., Pattee, H.H.: Evolution experiments with an artificial ecosystem. Journal of Theoretical Biology 28(3), 393–409 (1970)
Sayama, H.: Seeking open-ended evolution in Swarm Chemistry. Proceedings of the Third IEEE Symposium on Artificial Life (IEEE ALIFE 2011), IEEE, pp.186–193 (2011)
Sayama, H., Wong, C.: Quantifying evolutionary dynamics of Swarm Chemistry. Advances in Artificial Life, ECAL 2011: Proceedings of the Eleventh European Conference on Artificial Life, MIT Press, pp.729–730 (2011)

Sayama, H.: Morphologies of self-organizing swarms in 3D Swarm Chemistry. Proceedings of the 2012 Genetic and Evolutionary Computation Conference (GECCO 2012), pp.577–584 (2012)

Sayama, H.: Evolutionary Swarm Chemistry in three-dimensions. Artificial Life 13: Proceedings of the Thirteenth International Conference on the Simulation and Synthesis of Living Systems, MIT Press, pp.576–577 (2012)

Pólya, G.: Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz. Mathematische Annalen 84, 149–60 (1921)

Domb, C.: On multiple returns in the random-walk problem. Mathematical Proceedings of the Cambridge Philosophical Society 50, 586–591 (1954)

Comparative Genomics of Convergent Evolution

Claudio L. Flores Martinez

1 Introduction

Convergent evolution (CE) describes the emergence of similar adaptive, phenotypic traits in (distantly) related evolutionary lineages of taxonomic groups (species, genera, phyla, etc.) across and within the three domains of life – Eubacteria, Archaea, and Eukaryota – on different levels of biological complexity[1] (e.g., uni- and multicellular organisms or pre-bilaterian and bilaterian animals). However, the genomic underpinnings of convergent phenotypes are often unknown or poorly understood. This chapter will provide a non-technical and brief review – aimed at an interdisciplinary audience interested in "bio-complexity" construed broadly – covering the latest insights regarding the phenomenon of convergent evolution on a genome-wide scale. It is a first attempt to systematize instances of CE across the entire tree of life and identify genomic mechanisms underlying CE and place it into a wider theoretical complex adaptive systems (CAS) perspective.

From a practical point of view, a deeper understanding of CE, a phenomenon appearing to be pervasive in the dynamic process of genome evolution, is a

---

[1] A comprehensive definition of "biological complexity" is beyond the scope of this work, but in the following it will be distinguished between "complexity" as a generic property of biological systems and the evolution of "organismal complexity" with an emphasis on the latter. A fitting example is given (Glansdorff et al. 2008): "Indeed, what has increased in the course of evolution is order and its corollary, organization [. . .]. For example, a biofilm may be more complex than the simplest metazoans but is considerably less ordered. Organismal complexity, when it arises, is contingent on order."

C. L. Flores Martinez (✉)
Department of Biology, Institute of Zoology, Molecular Evolutionary Biology, University of Hamburg, Hamburg, Germany
e-mail: claudio@kybernesia.org

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_12

promising avenue not only for biological theorizing but also for biomedical research where highly convergent disease states (e.g., in cancer or antibiotics resistance) are seriously hampering treatment and directly relate to the survival rate of patients.

In the following chapter we are going to discuss fundamental issues in theoretical evolutionary biology and connect these topics with a description of the basic workflow in comparative genomics, and, finally, place the mentioned examples within a top-down causation framework, as well as briefly mentioning relevant cases from biomedical research.

2 Comparative Genomics

The field of functional genomics is concerned with deciphering the totality of a given organism's genetic code through the use of DNA sequencing technology. Starting around the year 2000, the high-throughput sequencing revolution fundamentally changed the way scientists are tackling problems in genome evolution, human health, and computational biology (Alföldi and Lindblad-Toh 2013).

Today, next-generation sequencing tools allow the relatively cheap, fast, and precise characterization of genes and noncoding DNA regions of various model (e.g., mice and flies) and non-model organisms (e.g., deep-sea creatures such as tube worms inhabiting hydrothermal vent ecosystems). In addition, miniaturized sequencing technology allows the deployment of this powerful methodology during field expeditions, in clinical settings and on laboratory bench scale accessible for non-experts (Leggett and Clark 2017).

The generated raw sequence data is then put together by automatized computational pipelines and translated to its corresponding amino acid (i.e., protein) sequences which are easier to work with. It is important to note, that the entirety of a given organism's DNA is not always translated into protein; a considerable portion is needed for regulatory purposes and these genomic areas can be comprehensively mapped as well. Modern sequencing technology is not only able to decipher the complete DNA and protein sequence information of a given organism but also to read the intermediate mRNA molecules.

In this way RNA sequencing (RNAseq) provides important insights on the active, transcriptional state of a cell under defined parameters (for instance before and after the application of physiological stressors such as heat shock). In effect, researchers can get a snapshot of exactly what type of proteins a cell produces at a specific moment in time and under well-defined physiological parameters. Genomics is therefore complemented by transcriptomics in powerful ways (Lowe et al. 2017).

Combining the analysis of total protein and metabolite content in a given cell, tissue, or organism – approaches known as proteo- and metabolomics – with genomics and transcriptomics provides a fruitful synthesis of experimental and computational techniques for elucidating the molecular underpinnings of complex traits.

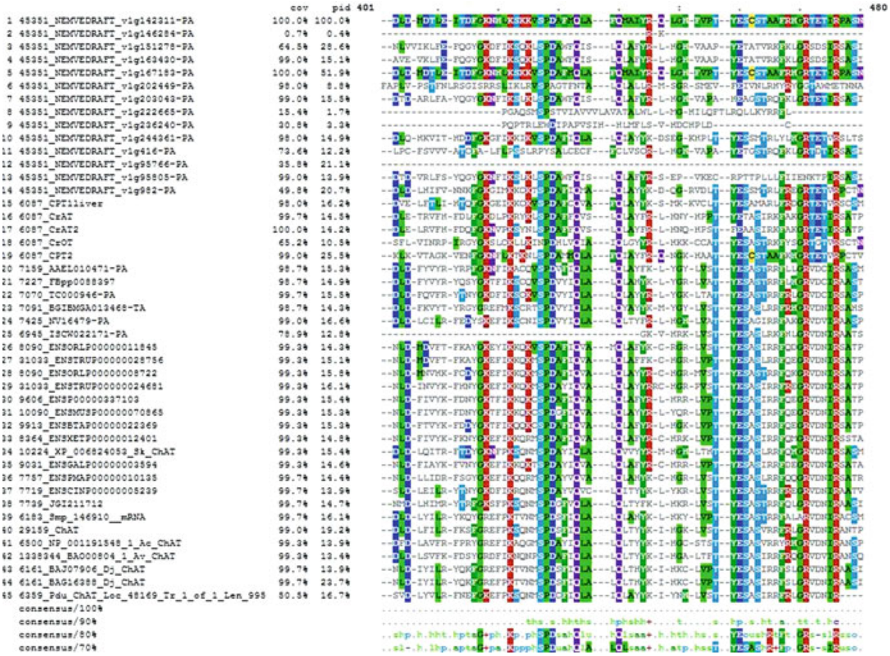


Fig. 1 Sequence alignment of 45 CrAT (Carnitine Acetyltransferase) proteins from metazoan organisms

Once computational biologists have put together the genome under investigation, the interesting part, namely the comparative analysis, begins (Fig. 1). So-called "sequence alignments" are used to map genomic regions of interest from one organism to another (or a multitude of them) (Chatzou et al. 2016). The generated alignments can then be used to compare genomes of species across vast phylogenetic distances, from viruses and bacteria to unicellular eukaryotes and multicellular organisms such as jellyfish, insects, and mammals.

After the organismal complexity of the biological trait under investigation has been constrained and defined in a first step by making a selection of relevant genomes, the phylogenetic analysis itself commences. This is one of the most work-intensive steps in the comparative genomic approach as computational and evolutionary biologists have to determine the phylogenetic pattern of the biological character under investigation (Hoyal Cuthill 2015; Yang and Rannala 2012).

Broadly speaking, in this step of the analysis, researchers are interested whether a certain biological trait arose via convergent evolution, i.e., independently and multiple times, or whether it has a singular homologous origin. In essence, phylogeneticists are looking at the phyletic distribution of a given trait (e.g., the presence of a specific set of proteins known to be carrying out a certain function) in the sampled genome selection and try to discern whether phyla that possess the respective character share a common ancestor that already contained the cellular modules.

If this is indeed the case, more fine-grained bioinformatics approaches can determine the degree to which these proteins have been conserved across different species and already performed the function under investigation in distant ancestors. A good overview on various conventional and emerging genome-wide measures of CE is given by Speed and Arbuckle (2017).

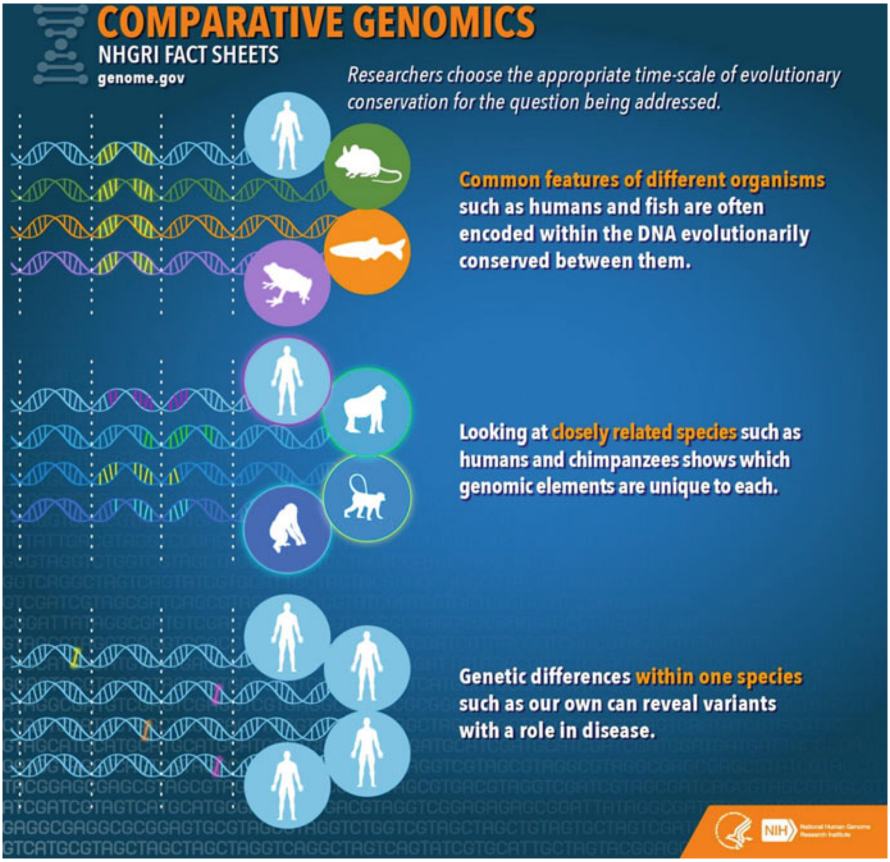
For example, an enzyme can be present in two distantly related species and also in the genome of their last common ancestor. Tailor-made bioinformatics scripts can then infer whether the specific enzymatic activity was already present in the shared ancestor or whether the enzyme acquired convergent amino substitutions in the distantly related active sites that confer analogous functions. Likewise, a certain protein family can be present in the LCA of two distantly related lineages. Just by comparing the copy number of genes belonging to the protein family, comparative genomics can tell whether it expanded independently in two distantly related lineages, which again is a case of convergent evolution on the molecular sequence level. Another scenario is encountered when a specific convergent phenotype is observed in two lineages, and it is known that the LCA did not possess the respective trait (e.g., through previous comparative morphological or paleontological analysis). In this case, convergent phenotypes are expressed through functionally equivalent but evolutionary distinct molecular systems. Comparative genomics (and transcriptomics) will then help to identify the disparate molecular agents underlying similar functions that emerged independently in distantly related lineages.

Different genomic mechanisms responsible for convergent evolution on the molecular sequence level and the role of functional equivalence classes will be discussed in more detail in later sections of this chapter.

The results of such an analysis can demonstrate whether nature has found multiple solutions for a certain biological problem or whether one universal solution has been inherited to all members of a species, genus, family, order, class, phylum, kingdom, or domain of a given taxonomic level under scrutiny.

Alternatively, the subfield of phylogenomics is mostly concerned with clarifying the evolutionary relationship across phyla and species themselves (Delsuc et al. 2005). One tantalizing result of phylogenomics in recent years has been the suggestion that the morphologically complex Ctenophora (comb jellies), which are equipped with nervous systems, might have been the first metazoan phylum to branch off even before the apparently less complex Porifera (sponges) (Moroz 2015; Whelan et al. 2015).

The generated sequence alignments allow researchers to compare the degree to which a particular genetic region has been conserved or has diverged over evolutionary time (Fig. 2). Even if a gene has greatly diverged, in most cases (but not always) it is still recognizable across vast evolutionary distances, for example, between jellyfish and *Homo sapiens*. If similar sequences can be recognized in distantly related organisms, researchers often posit the existence of “deep homologies” that



Researchers choose the appropriate time-scale of evolutionary conservation for the question being addressed.

Common features of different organisms such as humans and fish are often encoded within the DNA evolutionarily conserved between them.

Looking at closely related species such as humans and chimpanzees shows which genomic elements are unique to each.

Genetic differences within one species such as our own can reveal variants with a role in disease.

Fig. 2 Comparative genomics fact sheet. (Credit: NIH)

are underlying the emergence of similar complex traits in distantly related lineages (Shubin et al. 2009).

Complex nervous systems found in three of five metazoan branches have long been considered such a kind of deep homology. In the light of newly generated genomic data from organisms that have not been previously sequenced, it becomes increasingly difficult to differentiate between deep homology and CE of complex traits. A problem which has profound implications for how biologists and complexity theorists are to envision the dynamics of the evolutionary process (Laland et al. 2014).

The same genes in different species are called homologues. Other, more drastic and divergent, ways in which genomes evolve include the duplication of a gene (paralogous genes) or whole genome, gene losses, protein domain shuffling, gene birth, and other forms of rearrangement that occur on the genome level (Koonin 2005).

3 Phylogenetic Depth and Organizational Levels of Bio-complexity

We want to briefly introduce two novel concepts that, in our opinion, are going to shape future genome-wide comparative approaches dealing with CE of complex traits. At this point we only give a very short qualitative description, since this is work in progress and these ideas will be outlined in more detailed and quantified fashion in a future publication.

First, a given biological character can be deemed a “stronger” attractor in the phase space of evolutionary potential, the deeper the root (i.e., last shared node in a phylogenetic tree) is nested between two taxa that possess the trait within the universal tree of life.

For example, bioluminescence has a relative far-reaching phylogenetic depth since Eubacteria and Eukarya (but not Archaea) are known to produce biological light. Thus, bioluminescence has a phylogenetic depth at the domain level (since it emerged repeatedly in two domains of life). Furthermore, phylogenetic depth is amplified if more lineages in a respective domain (or phylum) evolved the character under investigation. For example, a trait that also evolved via CE in two domains (e.g., Eubacteria and Eukarya) can have a roughly similar phylogenetic depth as compared to bioluminescence but will not be as profound if it did not emerge as many times as biological light production.

One caveat has to be made at this point. Not every convergent trait can be equally weighed against each other in terms of phylogenetic depth. For example nervous systems are exclusively present in the Metazoa (animals) and thus the phylogenetic depth of the complex trait “nervous system” can only be assessed within the taxonomic level of Metazoa.

However, there is a multitude of multicellular organisms (multicellularity itself being a pervasive trait with considerable phylogenetic depth within the Eukarya) that could have, potentially, formed a nervous system. Therefore, we might ask: why did plants not evolve a nervous system? Do they possess an as of yet unrecognized information processing system that is essentially performing the same function of animal nervous systems (Baluska and Mancuso 2009)?

At what level (molecular, morphological or cell-type, etc.) do we define the “function” of a given complex trait? A number of important issues in terms of phylogenetic, organismic, and computational relationships have to be clarified before a “natural system” of CE can be derived from the ever-increasing amount of genome data.

In any case, if we envision a convergent trait as a type of “bio-attractor” in the phase space of highly dynamic genomic rearrangements, future research, both theoretical and experimental in nature, will be guided by a more powerful paradigm than is currently available.

This vision is closely related with yet another concept that we would like to introduce here in regard to the discipline of comparative genomics. A growing body of evidence suggests that Omics-techniques are not only able to capture genome-wide convergent changes on the molecular sequence level at multiple genetic loci, but many studies have shown that different levels of bio-complexity (e.g., transcriptional changes (Pankey et al. 2014)) can be affected by CE as well. Here we adopt the term "organizational levels" of bio-complexity for various intra- and inter-organismic relationships, dealing with different kinds of biomolecules, which can be potentially shaped by CE. Roughly speaking, we can differentiate between the following organizational levels:

Genome (DNA) → Transcriptome (RNA) → Epigenome (Methylation/Histones) → Metabolome/Proteome (Low-molecular compounds/Proteins) → Cell Type → Connectome (Neurons and Synapses) → Behavior → Interspecies Relationships

This type of flow-chart representation, however, is not meant as an evolutionary, linear representation of increasing bio-complexity from genome to interspecies relationships. Not all levels of organizational complexity are reached by each respective branch of the tree of life, e.g., all levels from "cell type" onward are exclusively instantiated by eukaryotic, multicellular organisms (with the exception that multicellular organisms can form symbiotic or parasitic relationships with unicellular Eubacteria and Eukarya). It is rather envisioned as a type of "information hierarchy" (a concept that will be discussed in more detail further below) with feedback loops linking individual levels.

Convergent evolution can occur on single or multiple levels of organizational complexity at the same time and one especially interesting avenue for advanced genomics research in the decades to follow is constituted by the investigation of causal relationships between various instantiations of CE at different levels. For example, we can ask whether genome-wide convergent evolution of a specific morphological trait is accompanied by analogous changes in the transcriptional activity of the compared species and tissues.

4 Emergence of the First Nervous Systems

The metazoan tree of life contains five major branches (Fig. 3): Porifera (sponges), Ctenophora (comb jellies), Placozoa, Cnidaria (together forming the "lower metazoans" for the purpose of the present discussion), and Bilateria (with the two main branches of deuterostomes and protostomes). Three of these phyla contain species with nervous systems of varying complexity and organizational sophistication: in ctenophores and cnidarians neurons are organized in a net-like structure (nerve net) while bilaterians display centralized nervous systems and brains. Sponges and placozoans lack discernible neural features.

In order to establish the entire sequence of events in early nervous system evolution, determining the correct branching order of the main metazoan phyla is essential. Up until a few years ago a consensus view suggested (Telford et al. 2015) that sponges branched off first, followed by either ctenophores equipped with neural nets or placozoans (with a secondary loss of neural structures depending

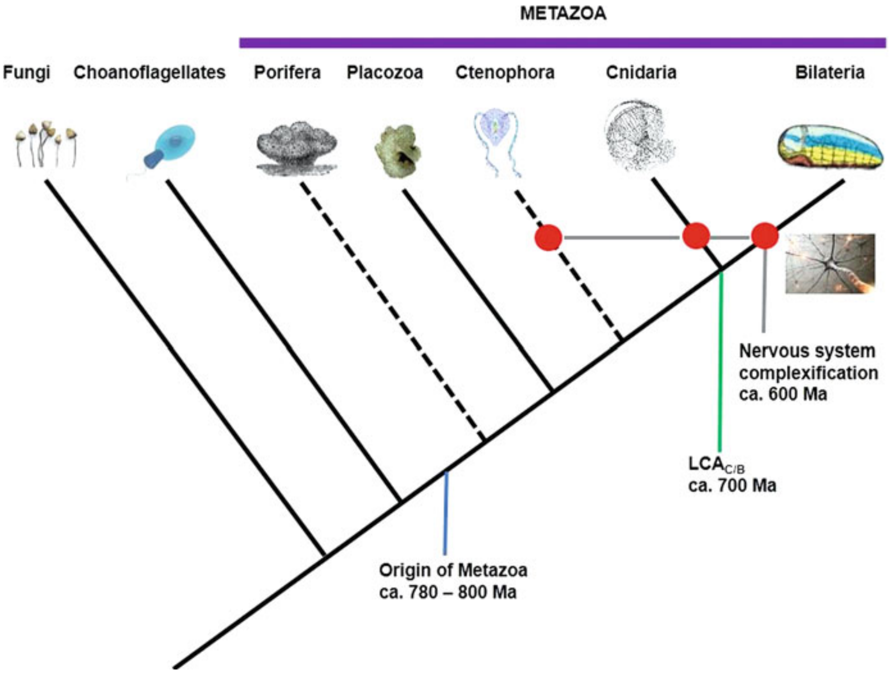


Fig. 3 Simplified phylogenetic tree of eukaryotes. Porifera, Placozoa, Ctenophora, Cnidaria, and Bilateria together form the Metazoa (dashed lines denote unclear phylogenetic position). Extant nervous systems can be found in ctenophores, cnidarians, and bilaterians. The origin of Metazoa can be approximately dated to 780–800 Mya and the cnidarian/bilaterian LCA (LCAC/B) to ca. 700 Mya. The origin of Bilateria is approximated at 688 Mya (not shown). Red dots indicate convergent complexification events in early nervous systems occurring in the respective lineages

on their actual position), and, lastly the split of the sister groups Cnidaria and Bilateria (the former having a nerve net and the latter centralized nervous systems). Since Cnidaria and Bilateria form sister groups, much attention in terms of genetic neurodevelopment has been given to the last common ancestor of these two clades ($LCA_{C/B}$).

Considerable debate surrounds the recent claim that, actually, Ctenophora might be the most basal metazoan phylum. This would suggest two loss events of neural structures in Porifera and Placozoa. Another scenario is presented by the possibility of convergent evolution of neurons and nerve nets in the stem line of Ctenophora and the one leading to the $LCA_{C/B}$ with primary absences of neurons in Porifera and Placozoa (polygenesis versus monophyly hypotheses). The emergence, evolution, and development of early nervous systems is a field of research with contentious claims surrounding (among others) basic premises concerning the phyletic branching pattern of lower metazoans.

The polygenesis vs. monophyly debate of neurons and nervous systems centering on the branching position of Ctenophora serves well to illustrate a general problem in evolutionary biology (Jekely et al. 2015; Moroz 2009; Moroz and Kohn 2016;

Pisani et al. 2015). Identifying patterns in the evolution of complex traits is often difficult as these often display a mix of shared and independently derived features. For example, sophisticated sense organs such as image-forming eyes evolved on many independent occasions during the Cambrian. Incorporating pre-existing photoreceptors, image-forming eyes evolved independently in arthropods, cephalopod mollusks, and vertebrates (Wray 2015). Eyes, however, were not built from scratch every time. Rather, they arose by the modification of pre-existing gene regulatory circuits already established in early metazoans. Based on the presence of highly conserved ancient developmental pathways, generative processes and cell-type specification leading to the same outcomes in independently evolved complex traits can be partly explained by "deep homology" (Shubin et al. 2009). Nevertheless, deep homology of shared developmental toolkits is not sufficient for explaining the repeated and independent emergence of complex systems as such. This follows from the realization that specific biological systems (for example nervous systems) cannot exhaustively be described solely in terms of their underlying gene regulatory network. Likewise, development in general cannot be reduced to the mere execution of genetically encoded instructions. Epigenesis is a well-established phenomenon in developmental biology and, increasingly, also beyond this particular subdiscipline. Convergent evolution of complex traits goes beyond the iterated deployment or repurposing of shared developmental building blocks in independently evolving lineages. Deep homology of developmentally important genes and gene networks related to neurogenesis led to the expectation that novelties in developmental patterning would find rapid phenotypic expression in ways similar to their functions in bilaterians. Therefore, it must be distinguished between the origin of highly conserved genes, the formation of developmental gene regulatory networks that control morphogenetic patterning, and the subsequent origin of phenotypic complex traits. As a corollary, inferring patterns of morphological innovation from comparative genomic and gene expression studies can be misleading at times since specific genetic novelties are not necessarily immediately expressed as developmental innovations (Erwin 2015; Wray 2015).

For an excellent review on the evolutionary history of such "complex homology" in early nervous systems, see Liebeskind et al. (2016). The authors argue that it might be helpful to organize and understand the evolution of animal forms "… as a set of stable states toward which the different lineages have been pulled." Further, regarding explosive systemic changes in early nervous system evolution the authors note:

> Around 600 Mya, after all five major lineages (ctenophores, sponges, placozoans, cnidarians, and bilaterians) had diverged, there was a sudden change. Large expansions of the gene families associated with synaptic and electrical complexity occurred together with profound changes on the biophysical level. These changes occurred convergently in the stem lineages of ctenophores, cnidarians, and bilaterians. Although many of these genomic events cannot yet be dated precisely, they most likely followed the end of the worldwide glaciation events, contemporaneously with the rise of oceanic oxygen and macroscopic animal forms. The rise of inter-animal predation in the early Cambrian probably provided selective pressure to evolve complex behaviors, neural organization, and musculature. (Liebeskind et al. 2016)

5 Mechanisms

5.1 *Mosaic Evolution of Complex Traits*

As noted above, complex traits evolve as a mosaic of multiple genetic modules, epigenetic factors, and developmental processes. In terms of comparative analysis, an interesting observation is the fact that major transitions in bio-complexity, e.g., the emergence of multicellularity in various lineages or the multiple origins of nervous systems in animals, are often accompanied by a simultaneous increase of genomic elements that are not homologous. Nonetheless, these modules appear to mediate new cellular functions in a coordinated fashion. New genomic characters can arise through different mechanisms, which will be outlined below. The list is not comprehensive but covers the most well-known and latest instances in which dynamic genome evolution has led to an increase in genomic and phenotypic bio-complexity. Importantly, in most cases, we are not discussing the evolution of individual proteins here but rather protein family evolution and genome-wide rearrangements responsible for the emergence of novelty during the evolution of complex traits.

5.2 *Protein Family Expansions*

Most proteins are members of superordinate families that are related to each other to various degrees in terms of sequence homology. A given protein, therefore, is related to proteins from the same family but also to members of other families that all together form a protein superfamily. Thus, proteins display intricate relationships within and across protein families that can be reconstructed using comparative genomic approaches (Koonin 2005). The observed outcome of these relationships is represented in phylogenetic trees. Computational and evolutionary biologists can then deduce a homology or convergence pattern from these trees. Taken together with experimental, ecological, and paleontological evidence and molecular clock analysis, comparative genomics can provide rough (though often contradicting) indications when certain complex traits have evolved as homologous or convergent phenotypes. Again, protein families ought to be understood as important elements of investigation in comparative genomics:

> Gene families comprise several to many genes of similar nucleotide or amino acid sequences; they share similar cellular functions and commonly arise as a result of gene or genome duplication events. The expansion or contraction of gene families over evolutionary time in different lineages can be random or the result of natural selection, although demonstrating the latter can be difficult. Several mechanisms, such as tandem duplications, segmental duplications, or even whole-genome duplications can lead to the expansion of gene families. (Harris and Hofmann 2015)

Prime examples for the importance of convergent protein family expansions include multiple gains (and losses) in the ion channel complement of early metazoans (Liebeskind et al. 2015); independent enlargements in the genomic inventory of *N*-methyltransferases, alkaloid, and flavonoid enzymes involved in secondary compound synthesis (e.g., caffeine) across tea and coffee plants (Denoeud et al. 2014); or convergent increases in proto-cadherin family size in cephalopods and vertebrates associated with advances in neuronal complexity (Albertin et al. 2015). Often it is possible to identify the large-scale duplication events of certain protein families but it remains difficult to pin-point the exact sub-mechanism of such expansions in genome data from extant species.

5.3 *Identical Substitutions*

CE does not only encompass convergent expansions in both homologous and distantly related protein families, but, intriguingly, reaches down to the molecular sequence level. Examples for such convergent amino acid substitutions, enabling similar functions, are found in the emergence of echolocation among bats and dolphins (Parker et al. 2013), across marine mammals in general (Foote et al. 2015) and in the adaptive cardenolid resistance in insects conferred by convergent modifications of the Na,K-ATPase (Dobler et al. 2012).

5.4 *De Novo*

De novo recruitment in convergent evolution is presented by cases in which distantly related genes are utilized for analogous functions. Prominent examples involve bioluminescent systems in bacteria, ctenophores, cnidarians, and bilateral organisms that often share luciferin substrates but have independently evolved oxidative luciferase enzymes from different, distantly related protein families (Widder 2010). Furthermore, such a kind of de novo recruitment of nonhomologous genes for similar functions is especially dominant at the deepest phylogenetic divisions, e.g., between Eubacteria and Archaea, which share a minimal gene set but have recruited many distantly related genes in their transcription apparatus (Koonin 2003).

5.5 *Retrotransposition*

One recent study has shed light on a little understood genomic mechanism underlying a complex trait, gene regulation, that originated via CE in mice and humans (Lucas et al. 2018). It describes a functionally similar molecular architecture that regulates the same genes in an analogous way in both species but is not related in a homologous relationship. In both cases, the regulatory system is comprised out of noncoding RNA deriving from DNA regions randomly inserted into the genome by so-called jumping genes (retrotransposons). Many of these noncoding RNAs are originating from repeated DNA sequences, short interspersed nuclear elements (SINEs). When these loci are transcribed into RNA, they can be reversely transcribed into DNA and inserted throughout an organism's genome in regions that were previously devoid of these sequences. This process is known as retrotransposition.

In humans the largest group of retrotransposons is constituted by the SINE family that is made up by a type of transposons called "Alu elements." Actually, in the human genome, there are more than one million copies or "versions" of Alu. This amounts to approximately 10% of all human DNA. In addition, the Alu retrotransposons are not restricted to a particular genomic region but can be found all over the human genome.

Comparative genomics of human and mouse retrotransposon complements has shown that the latter does not have the human Alu elements. The mouse genome has a distinct inventory of SINEs made of so-called B/ID elements. In mammals, different groups of SINEs underwent periods of increased retrotransposon activity after the divergence of various evolutionary lineages (humans and mice diverged more than 90 Mya). If distantly related groups of SINEs like Alu and B/ID elements inserted themselves in similar genomic regions and gave rise to analogous gene regulatory effects, as reported by the new study, retrotransposition of (distantly related) jumping genes in similar genomic regions presents an intriguing case of CE with many more instances to be discovered in the future.

5.6 HGT

HGT is the acquisition of genomic information by "horizontal" means rather than "vertical" inheritance. In bacteria, genetic elements can be acquired through processes of transformation (incorporation of DNA elements from the environment), conjugation (exchange of plasmids between two bacterial cells), and transduction via viral vectors. In a broad sense, endosymbiosis events that led to the emergence of increased cellular complexity can be regarded as cases of HGT as well. HGT is an important evolutionary factor driving genome evolution of unicellular Eubacteria, Archaea, and Eukarya (especially through endosymbiosis). Although HGT occurs sporadically in multicellular organisms as well, it is far more prevalent in unicellular organisms. Comparative genomics of microbial organisms is able to disentangle complicated patterns of HGT and show whether a given trait was either inherited through vertical descent, acquired one time horizontally and then passed on to the descendants, or whether it was acquired independently and repeatedly by different lineages via CE. One case in point is the origin of plastids (organelles that can differentiate into chloroplasts) in plants and algae (W., C., & C., 2003). Alternative hypotheses suggest either a singular or convergent origin of this cellular innovation that enabled photosynthesis in both lineages. Future large-scale comparative studies of microbial organisms, supported by metagenomics (environmental sampling of microbial DNA and RNA), will further elucidate the distribution of convergent traits that have been acquired via HGT.

6 Top-Down Causation and Genome Evolution

6.1 Information Hierarchies

Convergent evolution in biology can be modeled as the result of networks made up by biomolecules or other agents that are organized and structured by information hierarchies via top-down causation (Auletta et al. 2008). The emergence of fundamental network properties such as modularity and the related phenomenon of functional equivalence classes of lower-level operations or subroutines – both in biological and technological systems – can be explained as a corollary of top-down organized information hierarchies. Top-down causation refers to the causal role of information in living systems.

More specifically, it describes the process whereby higher levels of organization in structural hierarchies constrain the dynamics of lower levels of organization. Top-down theorists propose that the organization, structure, and function of living systems cannot be completely explained by a reductionist “bottom-up” approach. In a typical reductionist paradigm, it is assumed that purely physical effects determine the dynamics of lower levels of organization and, by extension, strictly govern interactions occurring at higher levels as well (Walker 2014).

Information, however, can acquire a causally efficacious role in physical systems without violating the principle of the closure of the physical world. An emerging school of thought in evolutionary biology advances the hypothesis that the transition from nonlife to life, abiogenesis, can be aptly described as a transition in causation and information flow (Walker 2014). Viewed from such a perspective, the study of genome evolution is especially fruitful when dealing with major transitions across various levels of complex biological organization (E V. Koonin 2007).

Top-down causation can influence genome evolution in a non-random way as evidenced by convergent changes in hundreds of mutations that benefit the functional emergence of a trait that can serve as a direct adaptation to a specific environmental context. The genomic study on echolocation in dolphins and bats (Parker et al. 2013) is one of a growing number of publications that begin to assess the implications of convergent evolution on a genomic scale and its impact on the emergence of evolutionary novelties. Viewing such phenomena from a top-down causation perspective connects the phenomenon of CE with a broader paradigm related to the emergence and evolution of bio-complexity:

> Top-down causation is an important mechanism for adaptive evolution through natural selection, where the higher-level "goal" is survival. The role of top-down causation in adaptive selection is particularly evident for cases of convergent evolution. A striking example is provided by the evolution of echolocation in dolphins and bats, where over 200 genes have independently changed in the same ways to confer both species with the ability to use sonar. A common high-level selection pressure (e.g., for the ability to navigate) lead independently to the same specific mutations in the DNA of both species. Convergent evolution thus provides a clear example of top-down causation via adaptive selection, where causal influences run from macroscopic environmental context to microscopic biochemical structure. (Walker 2014)

Informational takeover events appear to have occurred during all of the major transitions in evolution (Jablonka and Lamb 2006). At any given higher level of biological complexity, novel top-down informational hierarchies define lower-level molecular, cellular, or organismal operations. This idea is closely related to the concept of functional equivalence classes:

> Top-down causation operates through functional equivalence classes. Functional equivalence occurs when a given "higher-level" state leads to the same high-level outcome, independent of which "lower-level" state instantiates it. Equivalence classes are defined in terms of their function, not their particular physical instantiation: operations are considered (functionally) equivalent (i.e., in the same equivalence class) if they produce the same outcome for different lower-level mechanisms. Functional equivalence classes therefore represent the physical manifestations of virtual constructors. Functional equivalence is evident, for example, in the case of convergent evolution presented above, where convergence occurs because natural selection optimizes a functionally equivalent outcome (in this case, echolocation). (Walker 2014)

Yet an even more illustrative phenomenon showing how biological complexity is governed by top-down causation is the regulation of gene expression by the epigenome. Genomic information is stored in a physical object, DNA. The entirety of DNA in a given organism is constituted by its genome. Nonetheless, the genome does not contain the complete information on how, when, and to what extent certain genes should be expressed or repressed. This critical cellular information is instantiated in the epigenome. While the epigenome is associated with certain non-DNA molecular sequences (e.g., histone modifications), it should be rather viewed as a "… global, systemic, entity" whose "… real-time operation …lies in the realm of nonlinear bifurcations, interlocking feedback loops, distributed networks, top-down causation and other concepts familiar from the complex systems theory" (Davies 2012).

Information hierarchies allow for bidirectional information flow from higher to lower levels of biological organization. Yet, bidirectional information flow does not imply equivalence of each respective structural level in regulating biological systems.

Lastly, the profound functional similarities in the large-scale organization of biological and technological complex adaptive systems (Flores Martinez 2017) are pointing toward largely unexplored ordering principles underlying physical reality that can be framed within a future integrative framework based on top-down causation and convergent evolution:

> Advanced technologies and biology have extremely different physical implementations, but they are far more alike in systems-level organization than is widely appreciated. Convergent evolution in both domains produces modular architectures that are composed of elaborate hierarchies of protocols and layers of feedback regulation, are driven by demand for robustness to uncertain environments, and use often imprecise components. This complexity may be largely hidden in idealized laboratory settings and in normal operation, becoming conspicuous only when contributing to rare cascading failures. These puzzling and paradoxical features are neither accidental nor artificial, but derive from a deep and necessary interplay between complexity and robustness, modularity, feedback, and fragility. (Csete and Doyle 2002)

Integrating the analyses provided by published genome-wide studies with dynamical and information theory-oriented models of genome evolution (Walker et al. 2013), we aim at establishing a coherent approach allowing, first, the identification of problems in evolutionary biology that can be solved by using comparative genomics; second, outline the kind of bioinformatics and computational pipelines needed for solving the problem (e.g., family expansion vs. amino acid substitution); and third, interpret the results within adequate theoretical frameworks dealing with major transitions in bio-complexity.

6.2 Functional Equivalence Classes in Comparative Genomics

Right now the everyday work routine of researchers in comparative genomics is quite detached from such theoretical considerations. However, what is lacking in the field is a comprehensive outlook on the phenomenon of CE that combines theoretical, evolutionary, and computational biology. Some authors even suggest that modern bioinformatics has its roots in theoretical biology (Hogeweg 2011). Such a coming synthesis would serve to assess the impact of CE in constraining biological form. In general, comparative analyses from the last two decades have shown that CE is far more prevalent than previously expected, reaches down to the molecular sequence level and is a phenomenon that is related to the emergence of major transitions and complex traits. This paper lays the foundation for such an approach and is reinterpreting data from comparative genomics within the framework of top-down causation.

What we can already suggest is a good theoretical fit between basic concepts from comparative genomics and top-down causation. The convergent evolution of complex traits can be logically connected to mechanisms for dynamic genome evolution, the idea of functional equivalence classes and the identification of specific protein families and other molecular agents that mediate novel functionalities in organisms across the entire tree of life (Table 1).

Table 1 Logic flow between the emergence of convergent complex traits, underlying genomic mechanisms, related functional equivalence classes, and specific molecular instantiations embodying characters that arose through CE

Complex trait	Mechanism	Functional equivalence class	Specific protein (family)/molecular agents
Caffeine synthesis	Protein family expansion	Secondary compound synthesis	*N*-methyltransferases, alkaloid, and flavonoid enzymes
Echolocation	Identical substitution	Hearing/deafness Vision	Nearly 200 loci (e.g., proteins from solute carrier and integrin families)
Bioluminescence	De novo	Enzymatic light production	Peroxisomal fatty acyl-CoA synthetase (fireflies), oxidoreductase (sea pansy)
Gene regulation (Staufen-mediated mRNA decay)	Retrotransposition	Staufen-mediated mRNA decay (SMD)	Short interspersed elements (SINEs, noncoding RNA) *Alu* in humans, *B/ID* in mice
Antibiotics resistance	HGT	Counter measure to antibiotic agent	Plasmid-mediated gene *mcr-1* (colistin resistance) *blaCTX−M* (β-lactam resistance) Plasmid-encoded *qnrA* (quinolone resistance)

7 Biomedical Relevance

Finally, we want to emphasize that comparative genomics in the analysis of convergent complex traits does not only facilitate research in evolutionary biology but could potentially lead to breakthrough discoveries in biomedical research as well. Many computational approaches and methods that have been developed within the comparative genomics framework can now be readily applied to research dealing with various disease states.

One pertinent example is the field of cancer genomics. Using the latest high-throughput sequencing technology allows for the comparison of individual tumor tissues across patients afflicted with the same type of cancer. Are all analyzed tumors showing the same mutational pattern, or can different mutational states (i.e., functional equivalence) underlie the emergence of disease? Just as biological systems are sometimes evolving toward “bio-attractors” that pull the organism to a state of increased bio-complexity, unfortunately, such attractors might arise during disease states and cause cascading failure of overall genomic stability. Comparative genomics could provide theoretical and practical insights for the identification of convergent cell fates in cancer (Chen and He 2016). Future research into comparative genomic approaches involving various tumor tissues coupled with a view on genome stability inspired by top-down causation and biological CAS theory will surely help to advance our understanding of cancer disease states and expedite the development of new treatments.

Yet another clinically relevant area of biomedical research supported by comparative genomics methods is the alarming rise of antibiotics resistance in bacteria. Pervasive HGT and sequence amplification of resistance loci (Laehnemann et al. 2014; von Wintersdorff et al. 2016) are promoting the spread of multiresistant strains in hospitals. Comparative genomics focusing on the evolutionary dynamics of antibiotics resistance in bacterial strains that pose a danger to hospital patients, will uncover genomic hotspots that are responsible for rapid adaptation to current antibiotic agents. It is very likely that such a research program will reveal independent and repeated, i.e., convergent, processes of adaptive antibiotic resistance.

8 Outlook

This chapter presents an interdisciplinary endeavor at the intersection of evolutionary genomic research, computational biology, and complexity theory. It aimed to elucidate the genomic underpinnings of complex traits that emerged via CE, to achieve an updated view on the evolutionary process from the perspective of systems biology and the complexity sciences, as well as to derive theoretical and practical insights pertaining to fundamental properties and convergent states of biological systems, which can be harnessed for technological implementation and biomedical research.

References

Albertin, C. B., Simakov, O., Mitros, T., Wang, Z. Y., Pungor, J. R., Edsinger-Gonzales, E., . . . Rokhsar, D. S. (2015). The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature, 524*(7564), 220–224. doi:https://doi.org/10.1038/nature14668

Alföldi, J., & Lindblad-Toh, K. (2013). Comparative genomics as a tool to understand evolution and disease. *Genome Research, 23*(7), 1063–1068. doi:https://doi.org/10.1101/gr.157503.113

Auletta, G., Ellis, G. F., & Jaeger, L. (2008). Top-down causation by information control: from a philosophical problem to a scientific research programme. *J R Soc Interface, 5*(27), 1159–1172. doi:https://doi.org/10.1098/rsif.2008.0018

Baluska, F., & Mancuso, S. (2009). Plant neurobiology: from sensory biology, via plant communication, to social plant behavior. *Cogn Process, 10 Suppl 1*, S3–7. doi:https://doi.org/10.1007/s10339-008-0239-6

Chatzou, M., Magis, C., Chang, J. M., Kemena, C., Bussotti, G., Erb, I., & Notredame, C. (2016). Multiple sequence alignment modeling: methods and applications. *Brief Bioinform, 17*(6), 1009–1023. doi:https://doi.org/10.1093/bib/bbv099

Chen, H., & He, X. (2016). The Convergent Cancer Evolution toward a Single Cellular Destination. *Mol Biol Evol, 33*(1), 4–12. doi:https://doi.org/10.1093/molbev/msv212

Csete, M. E., & Doyle, J. C. (2002). Reverse Engineering of Biological Complexity. *Science, 295*(5560), 1664–1669. doi:https://doi.org/10.1126/science.1069981

Davies, P. C. (2012). The epigenome and top-down causation. *Interface Focus, 2*(1), 42–48. doi:https://doi.org/10.1098/rsfs.2011.0070

Delsuc, F., Brinkmann, H., & Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. *Nat Rev Genet, 6*(5), 361–375. doi:https://doi.org/10.1038/nrg1603

Denoeud, F., Carretero-Paulet, L., Dereeper, A., Droc, G., Guyot, R., Pietrella, M., … Lashermes, P. (2014). The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science, 345*(6201), 1181–1184. doi:https://doi.org/10.1126/science.1255274

Dobler, S., Dalla, S., Wagschal, V., & Agrawal, A. A. (2012). Community-wide convergent evolution in insect adaptation to toxic cardenolides by substitutions in the Na,K-ATPase. *Proc Natl Acad Sci U S A, 109*(32), 13040–13045. doi:https://doi.org/10.1073/pnas.1202111109

Erwin, D. H. (2015). Early metazoan life: divergence, environment and ecology. *Philos Trans R Soc Lond B Biol Sci, 370*(1684). doi:https://doi.org/10.1098/rstb.2015.0036

Flores Martinez, C. L. (2017). *Introducing Biomimomics: Combining Biomimetics and Comparative Genomics for Constraining Organismal and Technological Complexity*, Cham

Foote, A. D., Liu, Y., Thomas, G. W., Vinar, T., Alfoldi, J., Deng, J., … Gibbs, R. A. (2015). Convergent evolution of the genomes of marine mammals. *Nat Genet, 47*(3), 272–275. doi:https://doi.org/10.1038/ng.3198

Glansdorff, N., Xu, Y., & Labedan, B. (2008). The last universal common ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biol Direct, 3*, 29. doi:https://doi.org/10.1186/1745-6150-3-29

Harris, R. M., & Hofmann, H. A. (2015). Seeing is believing: Dynamic evolution of gene families. *Proceedings of the National Academy of Sciences, 112*(5), 1252–1253. doi:https://doi.org/10.1073/pnas.1423685112

Hogeweg, P. (2011). The roots of bioinformatics in theoretical biology. *PLoS Comput Biol, 7*(3), e1002021. doi:https://doi.org/10.1371/journal.pcbi.1002021

Hoyal Cuthill, J. F. (2015). The morphological state space revisited: what do phylogenetic patterns in homoplasy tell us about the number of possible character states? *Interface Focus, 5*(6), 20150049. doi:https://doi.org/10.1098/rsfs.2015.0049

Jablonka, E., & Lamb, M. J. (2006). The evolution of information in the major transitions. *J Theor Biol, 239*(2), 236–246. doi:https://doi.org/10.1016/j.jtbi.2005.08.038

Jekely, G., Paps, J., & Nielsen, C. (2015). The phylogenetic position of ctenophores and the origin(s) of nervous systems. *Evodevo, 6*, 1. doi:https://doi.org/10.1186/2041-9139-6-1

Koonin, E. V. (2003). Comparative genomics, minimal gene-sets and the last universal common ancestor. *Nat Rev Microbiol, 1*(2), 127–136. doi:https://doi.org/10.1038/nrmicro751

Koonin, E. V. (2005). Orthologs, paralogs, and evolutionary genomics. *Annu Rev Genet, 39*, 309–338. doi:https://doi.org/10.1146/annurev.genet.39.073003.114725

Koonin, E. V. (2007). The Biological Big Bang model for the major transitions in evolution. *Biology Direct, 2*(1), 21. doi:https://doi.org/10.1186/1745-6150-2-21

Laehnemann, D., Pena-Miller, R., Rosenstiel, P., Beardmore, R., Jansen, G., & Schulenburg, H. (2014). Genomics of rapid adaptation to antibiotics: convergent evolution and scalable sequence amplification. *Genome Biol Evol, 6*(6), 1287–1301. doi:https://doi.org/10.1093/gbe/evu106

Laland, K., Uller, T., Feldman, M., Sterelny, K., Muller, G. B., Moczek, A., … Strassmann, J. E. (2014). Does evolutionary theory need a rethink? *Nature, 514*(7521), 161–164. doi:https://doi.org/10.1038/514161a

Leggett, R. M., & Clark, M. D. (2017). A world of opportunities with nanopore sequencing. *J Exp Bot, 68*(20), 5419–5429. doi:https://doi.org/10.1093/jxb/erx289

Liebeskind, B. J., Hillis, D. M., & Zakon, H. H. (2015). Convergence of ion channel genome content in early animal evolution. *Proc Natl Acad Sci U S A, 112*(8), E846–851. doi:https://doi.org/10.1073/pnas.1501195112

Liebeskind, B. J., Hillis, D. M., Zakon, H. H., & Hofmann, H. A. (2016). Complex Homology and the Evolution of Nervous Systems. *Trends Ecol Evol, 31*(2), 127–135. doi:https://doi.org/10.1016/j.tree.2015.12.005

Lowe, R., Shirley, N., Bleackley, M., Dolan, S., & Shafee, T. (2017). Transcriptomics technologies. *PLOS Computational Biology, 13*(5), e1005457. doi:https://doi.org/10.1371/journal.pcbi.1005457

Lucas, B. A., Lavi, E., Shiue, L., Cho, H., Katzman, S., Miyoshi, K., … Maquat, L. E. (2018). Evidence for convergent evolution of SINE-directed Staufen-mediated mRNA decay. *Proceedings of the National Academy of Sciences*. doi:https://doi.org/10.1073/pnas.1715531115

Moroz, L. L. (2009). On the independent origins of complex brains and neurons. *Brain Behav Evol, 74*(3), 177–190. doi:https://doi.org/10.1159/000258665

Moroz, L. L. (2015). Convergent evolution of neural systems in ctenophores. *J Exp Biol, 218*(Pt 4), 598–611. doi:https://doi.org/10.1242/jeb.110692

Moroz, L. L., & Kohn, A. B. (2016). Independent origins of neurons and synapses: insights from ctenophores. *Philos Trans R Soc Lond B Biol Sci, 371*(1685), 20150041. doi:https://doi.org/10.1098/rstb.2015.0041

Pankey, M. S., Minin, V. N., Imholte, G. C., Suchard, M. A., & Oakley, T. H. (2014). Predictable transcriptome evolution in the convergent and complex bioluminescent organs of squid. *Proc Natl Acad Sci U S A, 111*(44), E4736–4742. doi:https://doi.org/10.1073/pnas.1416574111

Parker, J., Tsagkogeorga, G., Cotton, J. A., Liu, Y., Provero, P., Stupka, E., & Rossiter, S. J. (2013). Genome-wide signatures of convergent evolution in echolocating mammals. *Nature, 502*(7470), 228–231. doi:https://doi.org/10.1038/nature12511

Pisani, D., Pett, W., Dohrmann, M., Feuda, R., Rota-Stabelli, O., Philippe, H., … Worheide, G. (2015). Genomic data do not support comb jellies as the sister group to all other animals. *Proc Natl Acad Sci U S A, 112*(50), 15402–15407. doi:https://doi.org/10.1073/pnas.1518127112

Shubin, N., Tabin, C., & Carroll, S. (2009). Deep homology and the origins of evolutionary novelty. *Nature, 457*(7231), 818–823. doi:https://doi.org/10.1038/nature07891

Speed, M. P., & Arbuckle, K. (2017). Quantification provides a conceptual basis for convergent evolution. *Biol Rev Camb Philos Soc, 92*(2), 815–829. doi:https://doi.org/10.1111/brv.12257

Telford, M. J., Budd, G. E., & Philippe, H. (2015). Phylogenomic Insights into Animal Evolution. *Curr Biol, 25*(19), R876–887. doi:https://doi.org/10.1016/j.cub.2015.07.060

von Wintersdorff, C. J. H., Penders, J., van Niekerk, J. M., Mills, N. D., Majumder, S., van Alphen, L. B., … Wolffs, P. F. G. (2016). Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Frontiers in Microbiology, 7*(173). doi:https://doi.org/10.3389/fmicb.2016.00173

W., S. J., C., R. D., & C., J. J. (2003). A Single Origin Of Plastids Revisited: Convergent Evolution In Organellar Genome Content. *Journal of Phycology, 39*(1), 95–105. doi: https://doi.org/10.1046/j.1529-8817.2003.02070.x

Walker, S. (2014). Top-Down Causation and the Rise of Information in the Emergence of Life. *Information, 5*(3), 424

Walker, S. I., Callahan, B. J., Arya, G., Barry, J. D., Bhattacharya, T., Grigoryev, S., … Rosenberg, S. M. (2013). Evolutionary dynamics and information hierarchies in biological systems. *Ann N Y Acad Sci, 1305*, 1–17. doi:https://doi.org/10.1111/nyas.12140

Whelan, N. V., Kocot, K. M., Moroz, L. L., & Halanych, K. M. (2015). Error, signal, and the placement of Ctenophora sister to all other animals. *Proc Natl Acad Sci U S A, 112*(18), 5773–5778. doi:https://doi.org/10.1073/pnas.1503453112

Widder, E. A. (2010). Bioluminescence in the ocean: origins of biological, chemical, and ecological diversity. *Science, 328*(5979), 704–708. doi:https://doi.org/10.1126/science.1174269

Wray, G. A. (2015). Molecular clocks and the early evolution of metazoan nervous systems. *Philos Trans R Soc Lond B Biol Sci, 370*(1684). doi:https://doi.org/10.1098/rstb.2015.0046

Yang, Z., & Rannala, B. (2012). Molecular phylogenetics: principles and practice. *Nat Rev Genet, 13*(5), 303–314. doi:https://doi.org/10.1038/nrg3186

Why Functional Genomics Is the Central Concern of Biology and the Hard Problem of Abiogenesis

James A. Coffman

What I cannot create I do not understand

– Richard Feynman, on his blackboard a the time of his death, February 1988

Perhaps the first lesson to be learned from biology is that there are lessons to be learned from biology

– Robert Rosen, Essays on Life Itself, Chapter 18.

1 Introduction

The nature and origin of life is a hard problem that continues to vex science. Although the physical and geological sciences have made inroads that provide compelling scenarios for how the primordial stage for terrestrial abiogenesis may have been set (Smith and Morowitz 2004; Copley et al. 2007; Adamala and Szostak 2013; Damer 2016), these only bring into sharper relief the physically intractable aspect of the problem. At the heart of the problem lies the empirical fact, incontrovertibly established in the mid-nineteenth century by Louis Pasteur's experimental proof (Pasteur 1861), that on present-day earth only life begets life. This raises three related conundrums that haunt the physical sciences (physics and chemistry), and which, as I will argue, those sciences are ill-equipped to address: (1) the information manifested by the fact that offspring resemble their parents, and do so more closely than do more distant relatives, is not found in physics or chemistry; (2) living organisms differ fundamentally from dead ones, such that physics and chemistry provide no insight that can be applied to bring the dead back to life, and (3) life has never been adequately explained in strictly physicochemical terms, much less chemically synthesized in a laboratory. In other words, the specter of vitalism,

---

J. A. Coffman (✉)
MDI Biological Laboratory, Salisbury Cove, ME, USA
e-mail: jcoffman@mdibl.org

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_13

which most right-thinking scientists like to believe was long ago exorcized from scientific thought, is still very much with us.

Erwin Schrödinger was the first modern physicist to provide deep insight into the definitive characteristics that distinguish life from nonlife (Schrödinger 1944). First, he noted that organisms remain alive by exporting the entropy produced by the work of living. Living systems appear to defy the second law of thermodynamics, existing in a homeostatic state of flux that is far from thermodynamic equilibrium. But that localized defiance is nonetheless compliant with the law, as shown by the net entropy increase that the work of living can be seen to produce when the larger environment is taken into account. That points up another crucially important fact – no living system exists in isolation. In light of that fact, it is worth jumping the gun here to posit that terrestrial life is part and parcel with the global ecosystem (i.e., the biosphere). The two are so inextricably and irreducibly connected that adequately explaining one cannot be accomplished without adequately accounting for the other. It is my hope that the deep significance of this will become clear by the end of this essay.

The second key insight of Schrödinger was that the unique ability of living systems to maintain homeostasis by exporting entropy manifests information (defined mathematically as a reduction of entropy (Coffman 2011)) of such complexity that its perpetuation can only be accomplished by what he termed an "aperiodic crystal" (Schrödinger 1944). This insight presaged and inspired the molecular revolution launched in 1953 by the discovery of DNA double helix (Watson and Crick 1953a, b), which was immediately recognized to provide a molecular means of embodying of the essential informational characteristics sought by Schrödinger. But, as Schrödinger also noted, the requirement for such information dashes any hope that the physics as currently conceived is equipped for the task of explaining life, and thus, that life "is likely to involve 'other laws of physics' hitherto unknown" (Schrödinger 1944) – an insight that Robert Rosen (Rosen 2000) took up as a motivating challenge, as discussed below.

Jacques Monod, one of the founding fathers of molecular genetics, extended Schrödinger's line of reasoning in his 1970 essay *Chance and Necessity* (Monod 1971), which threw down a gauntlet that has yet to be honestly taken up by reductionist science. His thesis was essentially thus: the inherently teleological ("teleonomic") nature of life, which sets biology apart from chemistry, is explained by the fact that the activities of organisms are directed by genomically encoded information. That information is completely compatible with physics, and thus does not necessitate any nonphysical "vital" force. Nevertheless, the genomic sequence information that informs life is physically *indeterminate* (in the same way that the text of this book is physically indeterminate, i.e., arbitrary), presumably sculpted by natural selection, and thus not predictable by physics or chemistry. According to Monod, biology is therefore a "special" science. The special character of biology comes from the fact that it is historical (molded by chance (Coffman 2014)) as much as physical (molded by necessity). That is, the genomically encoded algorithms that enable life came to be through a historical process of evolution that is rife with unpredictable contingencies. As a result, any objective view of life must concede that chance played a seminal role in its origins (Coffman 2014), in a way that cannot be logically deduced from first principles of math and physics.

Theoretical biologist Robert Rosen disparaged Monod's genetic reductionism but praised Schrödinger's insight regarding the inadequacy of established physics for understanding life (Rosen 2000). Taking up that challenge, Rosen sought to develop a formalism that views life as generic to the universe, rather than a special case (Rosen 1991, 2000).To do so he used category theory to show precisely how living systems (organisms) differ from machines (Mikulecky 2000). The formal system that Rosen developed and used for his proof involved two abstract processes (mappings), metabolism (M) and repair (R), which together constitute an (M,R) system (later expanded to include replication, which can be viewed as system-wide repair). Within an (M,R) system the means of repair are produced by metabolism, and vice versa. Rosen's abstract (M,R) logic thus demonstrates how unlike machines (including Turing machines), organisms are closed to efficient cause – that is, they are self-referential (thus involving impredicative logic) and self-entailing (Rosen 1991). Rosen termed such systems "complex," to distinguish them from "simple" (i.e., simulable or Turing-computable) mechanical systems that are not self-entailing and can be adequately modeled using standard math (Rosen 1991). The important implication of Rosen's (M,R) formalism is that the essence of life lies not in its specific material components or instantiations, but in how those relate functionally to one another, that is, how they are *organized*, so as to produce a self-entailing system that is closed to efficient cause (Mikulecky 2000). It is therein that we must seek the *élan vital* that is somehow lost at death and cannot be physically revived from the outside.

Rosen's further insight was that the functional organization embodied by organisms constitutes a *modeling relation* (Rosen 1985, 1991). Modeling relations occur when formal entailment within a system *commutes* with natural causation in the external world, that is, when an internal entailment reliably produces an outcome (e.g., phenotype, as discussed below) that relates functionally to a causally determined circumstance. The model is then said to be *realized*. To the extent that internal entailment is more rapid than external causation, any system that embodies such a modeling relation behaves in a way that anticipates change in the world at large. As such, organisms are *anticipatory systems* (Rosen 1985), able to function adaptively in the world by anticipating change and working to make the adjustments necessary to accommodate themselves to it before it occurs and becomes an existential emergency. Rosen argued that to the extent that organisms are anticipatory systems, they are teleological, manifesting Aristotelian Final Cause (Rosen 1991).

It should be noted that in the late nineteenth century Charles Sanders Peirce developed a formalism that presaged that of Rosen in many ways (Fernandez 2008). Both logicians appear to have been inspired by similar motivations and independently came to similar conclusions. Their work opens the window on the formal nature of life, showing that it is contextually relational (i.e., *semiotic*), rather than purely physical, and hence opaque in important ways to the formalisms of conventional physics. But in the actual world life is nevertheless a material phenomenon, so the relations in question are instantiated by physical chemistry. The thesis of this essay is that many if not most of the uniquely biological relations and functional entailments that are essential to life as we know it are encoded in the genome, which came to be by way of a context-dependent, historically contingent process. As such, abiogenesis is a hard problem that cannot be solved by physics and chemistry alone, requiring additional knowledge gleaned from biosemiotics and paleontology.

2 Organisms Embody Genomically Informed Modeling Relations

The primacy of the genome in directing organismal ontogeny was unequivocally established by the elegant conceptual and experimental work of Theodor Boveri, who at the turn of the twentieth century showed (using sea urchin embryos) both that the information carried in the nucleus was sufficient to direct the development of species-specific characters, and that normal development requires that each blastomere of the developing embryo inherit a complete set of chromosomes – that is, a complete genome (Boveri 1902; Laubichler and Davidson 2008). 100 years later we had gained a good understanding of how it actually works, in no small part through the pioneering work of Boveri's intellectual descendents. These included Monod and Francois Jacob, who elucidated fundamental mechanisms of prokaryotic gene regulation (Jacob and Monod 1961; Monod and Jacob 1961), as well as Eric Davidson and Roy Britten, whose prescient theory of gene regulation in higher eukaryotes (Britten and Davidson 1969) was experimentally substantiated by decades of work in the laboratories of Davidson and his colleagues (Davidson 2006; Peter and Davidson 2016). The key insight was that normal developmental control of gene expression, and hence the program of ontogeny (the latter being entirely dependent on gene expression) is encoded in the genome as a gene regulatory network (GRN) defined by nucleotide sequence-specified interactions between regulatory genes and their products. Owing to their sequence specificity, GRNs operate logically, i.e., as an iterative algorithm whereby specific inputs *entail* specific outputs that serve as inputs entailing further outputs (Peter and Davidson 2015). What is ultimately entailed is the phenotype of the organism, its anatomy and physiology. To the extent that the phenotype is adapted to (i.e., fits with) its environment and anticipative of change therein, it constitutes a modeling relation, with the genomic system providing the "formal" entailments that commute with "natural" causation in the world at large (*a la* Rosen), thus realizing the model. Of course, at progressively higher hierarchical levels (cell, tissue, organism), other entailment structures emerge that contribute to determining the phenotype of the organism (Coffman 2006), and many of these are not reducible to or explained by GRNs. Nevertheless, it is reasonable to posit that the abiogenetic milestone that first distinguished biology from chemistry – the "phase transition" (Walker and Davies 2012) that succeeded in traversing the epistemic cut (Pattee 2001) – was the origin of genomic information directing development.

The entailment structures encoded in the GRN have been modeled using Boolean logic and can be parsed into a number of recurrent network motifs that underpin specific functions (Peter and Davidson 2017). These include positive and negative feedback loops as well as coherent and incoherent feedforward circuitry, design principles widely used in both single-celled and multicellular organisms (Alon 2007; Shoval and Alon 2010) that are also used by electronic engineers to control signaling dynamics. Commonly deployed motifs in the developmental GRNs of animals include double negative gate, dynamic feedback lockdown, community effect, spatial exclusion, and differentiation gene battery subcircuits (Davidson 2009). Beyond those that control the spatial and temporal aspects of ontogeny, genomes also encode entailment structures that control physiology and behavior more immediately. For example, in animals the physiological and metabolic adjustments to environmental change are mediated by neuroendocrine signaling that involves the orchestration of transcriptional programs by environmentally responsive, sequence-specific transcription factors, including members of the nuclear receptor family such as the glucocorticoid and thyroid receptors. Interestingly enough, it is now becoming increasingly apparent that the latter systems are subject to long-term, transcription factor-mediated (and thus genomically encoded) epigenomic programming in response to experience during early development, which (ideally) functions to tune their responsivity to anticipated environmental circumstances later in life (Harris and Seckl 2011; Spencer 2017). Thus, the anticipatory systems formally abstracted by Rosen are realized in nature by the genomic entailment structures that direct ontogeny, physiology, and behavior.

3 Genomically Encoded Biological Information Is Physically Arbitrary and Context-Dependent

As noted in the Introduction, the fact that the activities of living systems are directed by genomically encoded entailment structures poses a serious problem for those who seek to reduce biology to chemistry and physics. The problem can be clearly seen when one recognizes that science, as (a manifestation of) life itself, models the world via formal entailment (typically involving math). The sciences of physics and chemistry constitute well-developed entailment structures that allow us to model what happens in strictly physicochemical systems, as long as we are provided sufficient knowledge of the relevant parameters, including initial and boundary conditions. But while biological systems are physicochemical systems governed by physicochemical laws, they are more specifically governed – and more importantly, *directed* – by their own regulatory logic, consisting of genomically encoded semiotic entailment structures that developed and evolved concomitantly with the development and evolution of life on earth. This in no way implies that living systems contain or manifest some vital "force" that evades or transcends physicochemical law. Rather, it simply means that the full set of physicochemical parameters that played a role in shaping *bio-logic*, which would be needed for any comprehensive physicochemical accounting of life, are myriad and in many cases likely to be indefinable given that they were dependent on historical context and part and parcel with the development of life on earth. It is safe to assume that many of life's historical solutions to existential problems were chosen *arbitrarily* from among multiple possibilities: that is, when a choice between alternative algorithmic solutions to a problem needed to be made, life chose whatever was expedient and worked, from the set of possible but physically undetermined (or at least incompletely determined) solutions.

What I am arguing is not new, as Monod made essentially the same argument in *Chance and Necessity*. But as Monod intimated, scientists, being in the business of developing models wherein formal entailment commutes to natural causation, are constitutionally disinclined to accept that chance (i.e., unentailed) events may have played a role in directing outcomes (as epitomized by Einstein's quip that "God does not play dice with the universe"). We become experts in our scientific fields by learning the entailment structures that inform those fields and thus allow us to understand causation in the world. The entailment structures of biology, its *algorithms*, are based on genomic as well as physicochemical logic, whereas those of chemistry and physics are strictly physicochemical. As a result biology is a field in its own right, with a different formalism and language (i.e., genetics) than chemistry and physics, a language that is inscrutable within the formalisms of the latter two sciences. With the uniquely biological language of genetics, biologists have elucidated why children resemble their parents. Although physics and chemistry are implicit in that explanation, the entailment structures developed by those sciences provide no means whatsoever of deducing it.

4 Biology Is a Historical Science That Can Only Be Understood in Light of Paleontology

How then can we fruitfully approach the origin-of-life problem? Geneticist Theodosius Dobzhansky wrote, "Nothing in biology makes sense except in light of evolution" (Dobzhansky 1973). Given that evolution is an historical process, we may posit as a corollary that *nothing in biology makes sense except in light of paleontology*. Indeed, paleontology is one of the pillars of evolutionary biology, and a crucial source of evidence invoked by Darwin (1859). But paleontology is not only about rocks: more generally, it is about historically preserved information. As shown by the pioneering work of Carl Woese (Woese and Fox 1977) and recently exemplified by the work of paleontologist Kevin Peterson and colleagues (Peterson et al. 2008, 2009; Erwin et al. 2011; Wheeler et al. 2009), to the extent that biological information is genomically preserved, it can be used for paleontological reconstruction of the history of life (Koonin 2003; Martin et al. 2016; Weiss et al. 2016). While biologically preserved information is chemically more labile than geologically preserved information, some genomic information is so functionally constrained that it provides a fairly reliable historical record. Such is the case with rRNAs, which Woese used to elucidate the three domain tree of life (Woese and Fox 1977), and microRNAs, small regulatory RNAs whose specific nucleotide sequences are essentially frozen from the time of their origination and which display characteristic patterns of accumulation in different animal lineages (Peterson et al. 2009; Wheeler et al. 2009). Moreover, with the advent of computational genomics it is now possible to construct comprehensive sets of trees for all genes in a set of sequenced genomes, which can then be compared and sorted to identify phylogenetic signals (Koonin 2003; Martin et al. 2016; Weiss et al. 2016).

Despite the incompleteness of the historical record preserved in fossils and genomes, and hence the unavoidable need for conjecture, the field of paleontology has provided us with a reasonably good understanding of the history of life on earth from the time when the first signs of organisms appear in the fossil record (Nutman et al. 2016; Dodd et al. 2017; Djokic et al. 2017). From that time onward the history of the earth is shaped by the history of life, and the two are inextricable. But how those first organisms came to be is an open question. Speculative scenarios have been suggested, but the quest for certainty has been stymied by lack of paleontological evidence bearing on the matter. Perhaps the most compelling line of speculation postulates that the first organisms emerged from an autocatalytic prebiotic metabolism instantiated in protected, geochemically active locales fed by thermal vents or fields (Smith and Morowitz 2004; Damer 2016). Indeed, scientific discourse has now provided plausible scenarios for how such an autocatalytic system might have developed on primordial earth, and it is no longer not hard to imagine how sustainable instances of that prebiotic metabolism became encapsulated in lipid-enclose protocells containing proteins and nucleic acids (Adamala and Szostak 2013). But were those protobionts bona fide organisms in the sense that they were closed to efficient cause and fully self-entailing, and did they embody anticipatory, evolvable modeling relations? From the foregoing I would conclude that they were not, unless and until they had a *functional* genome (for which the mere presence of nucleic acids does not suffice). The question is not what gave rise to the material components of terrestrial life, but rather what gave rise to its semiotic relations and functional algorithms (Walker and Davies 2012).

Available evidence and theory provide some clues to answering that question. First, it is worth noting that a developmental perspective (Coffman 2006) (as opposed to a strictly evolutionary perspective) provides insight into what the process that generated the first functional genome must have entailed. As suggested above, the prebiotic terrestrial milieu probably afforded instances of autocatalytic closure (Kauffman 1993), and that alone engenders developmental emergence of information by way of positive feedback-driven selection of entities that contribute most effectively to the cycle, a quantifiable developmental phenomenon that Robert Ulanowicz has termed “ascendency” (Ulanowicz 1986, 1997). Thus, one might reasonably surmise that any rudiments of a functional genome generated by a prebiotic autocatalytic metabolism would have been selected by the autocatalytic cycle to the extent that they contributed to the cycle’s ascendency. Furthermore, the nature of development is that systems are necessarily more ambiguous in their early (immature) stages, manifesting lower ascendency (i.e., lower mutual information and lower total energy throughput) than they do at maturity (Ulanowicz 1997; Salthe 1993). Thus one would expect that the first organism would have emerged from a primordial cellular community of proto-organisms with an ambiguous, less constrained (i.e., more "statistical") genetic code, essentially as has been proposed by Woese and colleagues (Vetsigian et al. 2006). Furthermore, one might expect that the precursor of life as we know it was not so much an entity as a communal *process* among protobionts whose synthetic genomic innovations were widely shared by horizontal transfer (Vetsigian et al. 2006; Koonin 2014). The genetic code could eventually have emerged ("crystallized") through positive feedback-driven selection involving "competition between innovation pools" (Vetsigian et al. 2006).

We can thus surmise that the prebiotic community of protocells eventually gave rise to the cellular ancestor of all contemporary life, the last universal common ancestor (LUCA) (Koonin 2003; Martin et al. 2016; Weiss et al. 2016; Glansdorff et al. 2008), at which point the Darwinian tree of life (i.e., descent with modification shaped by natural selection) emerged and gave rise to the three major domains (or two primordial domains, if eukaryotes are taken to be derived from a symbiotic fusion of archaea and bacteria (Weiss et al. 2016; Koonin 2014)). Early comparative genomic studies spanning these domains estimated that the genome of the LUCA contained only ~500 genes, the conserved core of which encoded components of the translational apparatus (Koonin 2003). This number compares favorably with the minimized genome of the "synthetic" cell created by J. Craig Venter and colleagues, which is smaller than the genome of any known autonomously replicating cell in nature, encoding only 473 genes (Hutchison et al. 2016). More recent comparative genomic studies have identified a list of 355 genes that are conserved across multiple lineages of bacteria and archaea and were thus probably inherited from LUCA, allowing gene ontological elucidation of LUCA's functional genomics, and hence its physiology and ecological niche (Martin et al. 2016; Weiss et al. 2016). Satisfyingly, the conclusions from those studies converge with those of the geological record cited above, pointing to LUCA being an anaerobic autotroph inhabiting (and probably still highly dependent on) geo-metabolically active hydrothermal vents rich in H_2, CO_2, and iron (Martin et al. 2016; Weiss et al. 2016). Moreover, the studies point to methyl groups as being a chemical link relating tRNA-mRNA-rRNA interactions (and hence the genetic code) to the spontaneous chemistry of the LUCA's reductive habitat (Weiss et al. 2016). Thus, while abiogenesis remains an unsolved problem, paleontological science has begun to articulate a compellingly constrained scenario for the origin and functional genomics of the LUCA.

That being said, it is worth digressing here to comment on the semantics of the issue. "Abiogenesis," the origin of life on earth, is typically interpreted to mean the emergence of life from *non*life. But that interpretation rests on unsubstantiated assumptions that are based entirely on our inherent bias as highly evolved biological creatures. It is true that on present day earth, life (as we know it as scientists) is biological. But on primordial earth that may not have been the case. Indeed, if we apply Occam's razor and take at face value Pasteur's dictum that only life begets life, then we should not expect that a dead earth would have had the capacity to give rise to a living earth. Moreover, as Rosen showed, the formal relations that define life are generic to the universe, not special. From that perspective the most reasonable conclusion is that of Peirce (Brier 2017): that the earth, and indeed the universe as a whole, has always been alive – i.e., it is a self-entailing, developing system closed to efficient cause. The problem of abiogenesis then is not the origin of life *itself* but rather the origin and development of terrestrial life as a specifically *biological*, i.e., genomically informed, phenomenon.

5 Conclusion: The Origin of a Functional Genome Is the Hard Problem of Abiogenesis

Although science has elucidated compelling scenarios for how a prebiotic terrestrial metabolism produced the material building blocks of life, the question remains: was that *sufficient* to engender development of the genomic entailment structures that define biology? A developmental perspective that recognizes the agency of autocatalytic cycles suggests that it could very well have been. But understanding how it actually occurred requires a better theoretical grasp of how and under what conditions autocatalysis promotes the developmental emergence of self-referential, self-entailing semiotic systems. Moreover, any proposed solution will remain speculative until empirically verified. While paleontology is necessary to show that it actually occurred as proposed, such evidence is not sufficient, as paleontology is an historical science that unavoidably calls for some conjecture. Ultimately, solving the hard problem of abiogenesis will require a chemistry experiment that creates a cell with a functional genome from scratch, i.e., without the benefit of or reference to preexisting biological information.

Acknowledgments I thank Dr. Kevin Peterson for helpful comments on an initial draft of this paper and an autonomous reviewer for suggestions that greatly improved the final manuscript. Work in my laboratory is supported by grants from the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (P20GM103423 and P20GM104318).

References

Adamala, K & Szostak, JW. Nonenzymatic template-directed RNA synthesis inside model protocells. Science 2013; 342: 1098–1100.

Alon, U. Network motifs: theory and experimental approaches. Nat Rev Genet 2007; 8: 450–461.

Boveri, T. Uber mehrpolige Mitosen als Mittel zur Analzyse des Zellkerns. Verhandlungen der physicalish-medizinischen Gesselschaft zu Wurzburg. Neu Folge 1902; 35: 67–90.

Brier, S. How Peircean semiotic philosophy connects Western science with Eastern emptiness ontology. Prog Biophys Mol Biol 2017; 131: 377–386.

Britten, RJ & Davidson, EH. Gene regulation for higher cells: a theory. Science 1969; 165: 349–357.

Coffman, JA. Developmental ascendency: from bottom-up to top-down control. Biol Theor 2006; 1: 165–178.

Coffman, JA. Information as a manifestation of development. Information 2011; 2: 102–116.

Coffman, JA. On the meaning of chance in biology. Biosemiotics 2014; 7: 377–388.

Copley, SD, Smith, E & Morowitz, HJ. The origin of the RNA world: Co-evolution of genes and metabolism. Bioorg Chem 2007; 35: 430–443.

Damer, B. A Field Trip to the Archaean in Search of Darwin's Warm Little Pond. Life (Basel) 2016; 6.

Darwin, C. On the Origin of Species by Means of Natural Selection (John Murray, London, 1859).

Davidson, EH. The Regulatory Genome: Gene Regulatory Networks in Development and Evolution (Academic Press/Elsevier, San Diego, 2006).

Davidson, EH. Network design principles from the sea urchin embryo. Curr Opin Genet Dev 2009; 19: 535–540.

Djokic, T, Van Kranendonk, MJ, Campbell, KA, Walter, MR & Ward, CR. Earliest signs of life on land preserved in ca. 3.5 Ga hot spring deposits. Nat Commun 2017; 8: 15263.

Dobzhansky, T. Nothing in Biology Makes Sense except in the Light of Evolution. The American Biology Teacher 1973; 35: 125–129.

Dodd, MS, Papineau, D, Grenne, T, Slack, JF, Rittner, M, Pirajno, F et al. Evidence for early life in Earth's oldest hydrothermal vent precipitates. Nature 2017; 543: 60–64.

Erwin, DH, Laflamme, M, Tweedt, SM, Sperling, EA, Pisani, D & Peterson, KJ. The Cambrian conundrum: early divergence and later ecological success in the early history of animals. Science 2011; 334: 1091–1097.

Fernandez, E. Biosemiotics and self-reference from Peirce to Rosen. Eighth Annual Gatherings in Biosemiotics, 2008; University of the Aegean, Syros, Greece.

Glansdorff, N, Xu, Y & Labedan, B. The last universal common ancestor: emergence, constitution and genetic legacy of an elusive forerunner. Biol Direct 2008; 3: 29.

Harris, A & Seckl, J. Glucocorticoids, prenatal stress and the programming of disease. Horm Behav 2011; 59: 279–289.

Hutchison, CA, 3rd, Chuang, RY, Noskov, VN, Assad-Garcia, N, Deerinck, TJ, Ellisman, MH et al. Design and synthesis of a minimal bacterial genome. Science 2016; 351: aad6253.

Jacob, F & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 1961; 3: 318–356.

Kauffman, SA. Origins of Order: Self-Organization and Selection in Evolution (Oxford University Press, New York, 1993).

Koonin, EV. Comparative genomics, minimal gene-sets and the last universal common ancestor. Nat Rev Microbiol 2003; 1: 127–136.

Koonin, EV. CarlWoese's vision of cellular evolution and the domains of life. RNA Biol 2014; 11: 197–204.

Laubichler, MD & Davidson, EH. Boveri's long experiment: sea urchin merogones and the establishment of the role of nuclear chromosomes in development. Dev Biol 2008; 314: 1–11.

Martin WF, Weiss MC, Neukirchen S, Nelson-Sathi S & Sousa FL. Physiology, phylogeny, and LUCA. Microb Cell 2016; 3: 582–587.

Mikulecky, DC. Robert Rosen: the well-posed question and its answer—why are organisms different from machines? Systems Research and Behavioral Science, ISSN 1099-1743 2000; 17: 419–432.

Monod, J. Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology (Vintage, New York, 1971).

Monod, J & Jacob, F. Teleonomic mechanisms in cellular metabolism, growth, and differentiation. Cold Spring Harb Symp Quant Biol 1961; 26: 389–401.

Nutman, AP, Bennett, VC, Friend, CR, Van Kranendonk, MJ & Chivas, AR. Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures. Nature 2016; 537: 535–538.

Pasteur, L. Sur les corpuscules organisés qui existent dans l'atmosphère: Examen de la doctrine des générations spontanées. Leçon Professée a la Société Chimique de Paris 1861.

Pattee, HH. The physics of symbols: bridging the epistemic cut. Biosystems 2001; 60: 5–21.

Peter, IS & Davidson, EH. Genomic Control Process: Development and Evolution (Academic Press, 2015).

Peter, IS & Davidson, EH. Implications of Developmental Gene Regulatory Networks Inside and Outside Developmental Biology. Curr Top Dev Biol 2016; 117: 237–251.

Peter, IS & Davidson, EH. Assessing regulatory information in developmental gene regulatory networks. Proc Natl Acad Sci U S A 2017; 114: 5862–5869.

Peterson, KJ, Cotton, JA, Gehling, JG & Pisani, D. The Ediacaran emergence of bilaterians: congruence between the genetic and the geological fossil records. Philos Trans R Soc Lond B Biol Sci 2008; 363: 1435–1443.

Peterson, KJ, Dietrich, MR & McPeek, MA. MicroRNAs and metazoan macroevolution: insights into canalization, complexity, and the Cambrian explosion. Bioessays 2009; 31: 736–747.

Rosen, R. Anticipatory Systems: Philosophical, Mathematical and Methodological Foundations (Pergamon Press, 1985).

Rosen, R. Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life (Columbia University Press, New York, 1991).

Rosen, R. Essays on Life Itself (Columbia University Press, New York, 2000).

Salthe, SN. Development and Evolution: Complexity and Change in Biology (MIT Press, Cambridge, MA, 1993).

Schrödinger, E. What is Life? (Cambridge University Press, 1944).

Shoval, O & Alon, U. SnapShot: network motifs. Cell 2010; 143: 326–e321.

Smith, E & Morowitz, HJ. Universality in intermediary metabolism. Proc Natl Acad Sci U S A 2004; 101: 13168–13173.

Spencer, KA. Developmental stress and social phenotypes: integrating neuroendocrine, behavioural and evolutionary perspectives. Philos Trans R Soc Lond B Biol Sci 2017; 372.

Ulanowicz, RE. Growth and Development: Ecosystems Phenomenology (Springer-Verlag, New York, 1986).

Ulanowicz, RE. Ecology, the Ascendent Perspective (eds. Allen, T. F. H. & Roberts, D. W.) (Columbia University Press, New York, 1997).

Vetsigian, K, Woese, C & Goldenfeld, N. Collective evolution and the genetic code. Proc Natl Acad Sci U S A 2006; 103: 10696–10701.

Walker, SI & Davies, PC. The algorithmic origins of life. J R Soc Interface 2012; 10: 20120869.

Watson, JD & Crick, FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 1953a; 171: 737–738.

Watson, JD & Crick, FH. Genetical implications of the structure of deoxyribonucleic acid. Nature 1953b; 171: 964–967.

Weiss, MC, Sousa, FL, Mrnjavac, N, Neukirchen, S, Roettger, M, Nelson-Sathi, S et al. The physiology and habitat of the last universal common ancestor. Nat Microbiol 2016; 1: 16116.

Wheeler, BM, Heimberg, AM, Moy, VN, Sperling, EA, Holstein, TW, Heber, S et al. The deep evolution of metazoan microRNAs. Evol Dev 2009; 11: 50–68.

Woese, CR & Fox, GE. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci U S A 1977; 74: 5088–5090.

Synergistic Selection: A Bioeconomic Approach to Complexity in Evolution

Peter A. Corning

1 Introduction

Complex living organisms are distinctive in that they are also subject to basic economic criteria and to economic constraints. Biological complexity is not simply a historical artifact, much less the product of some exogenous physical trend, force, or "law." Over the years, many candidate laws have been proposed that have claimed to explain complexity in evolution, going back to Jean Baptiste de Lamarck's (1984/1809) "power of life" and Herbert Spencer's (1892/1852) "universal law of evolution." In the latter part of the twentieth century, the development of new mathematical tools and rise of complexity theory in various disciplines inspired a plethora of new law-like, or mechanistic explanations. This theme has continued into the new century, as documented in detail in my new book, *Synergistic Selection: How Cooperation Has Shaped Evolution and the Rise of Humankind* (Corning 2018).

The problem with all such deterministic theories is that they explain away the very thing that needs to be explained – namely, the contingent nature of living systems and their fundamentally functional, adaptive properties. As the biologist Theodosius Dobzhansky long ago pointed out: "No theory of evolution which leaves the phenomenon of adaptation an unexplained mystery can be satisfactory." The purveyors of these theories often seem oblivious to the inescapable challenges associated with what Darwin called the "struggle for existence" in the natural world, and they discount the economics – the costs and benefits of complexity. Nor can they explain the fact that some 99% of all the species that have ever evolved are now extinct. Life is a phenomenon that is at all times subject to the requirement that

---

P. A. Corning (✉)
Institute for the Study of Complex Systems, Seattle, WA, USA
e-mail: pacorning@complexsystems.org

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_14

the bioeconomic benefits (direct or indirect) of any character or trait – including complexity – must outweigh the costs. It is subject to functional criteria and the calculus of economic costs and benefits in any given environmental context.

2 Defining Biological Complexity

A basic question, therefore, is what are the advantages of biological complexity? However, there is a prior question: What is "complexity"? We must start by defining what the term complexity means in relation to living systems before examining how – and why – biological complexity has evolved over time.

The issue of how to define biological complexity has been much-debated over the years, and it is evident that there is no one correct way to measure it; it can be defined in different ways for different purposes. However, two alternative methodologies are relevant (at least in theory) as ways of characterizing the broad evolutionary trend toward multileveled complex systems over the past 3.8 billion years or so, beginning with the origins of life and culminating in multileveled, socially organized species like humankind.

One method is structural. A synthetic complexity scale can be constructed from the number of levels of organization (inclusive of social organization), the number of distinct parts, the number of different kinds of parts, and the number of interconnections among the parts. The other method is functional. A complexity scale can be derived from the number of functionally discrete task in the division/combination of labor at all levels of organization (including behavioral and social), coupled with the quantity of "control information" that is generated and utilized by the system (Corning and Szathmáry 2015). Control information is defined as "the capacity to control the capacity to do work" in a cybernetic process; it is equivalent to the amount of thermodynamic work that a system can perform (Corning and Kline 1998; Corning 2007a).

3 Measuring Costs and Benefits

Various ways of measuring the economic costs and benefits of biological complexity are also possible. The ultimate measure is, of course, reproductive success. An organism must sustain a minimal economic "profit" in order to be able to reproduce itself. Thus, the more offspring it produces the more profitable it is from an "ultimate" evolutionary perspective.

However, there are also many "proximate" ways of measuring the costs and benefits involved in earning a living in nature, and a number of commonplace economic criteria are likely to have been important from a very early stage in the history of life on Earth – capital costs, amortization, operating costs, and, most especially, strict bioeconomic profitability. The returns had to outweigh the costs.

There is, of course, a large research literature and various journals in behavioral ecology and bioeconomics that are focused on just such proximate issues (e.g., see Davies et al. 2012).

Consider, for example, the fundamental need for energy capture. Dating back to Erwin Schrödinger's (1944) classic lectures and small book, *What is Life?,* it has long been appreciated that thermodynamics is of central importance in understanding the nature of life, and the challenges of living. Living systems must do work and are subject to thermodynamic entropy and the second law. This imposes significant functional requirements. However, there is also a deep tradition in biophysics that assumes away the economic challenges involved in creating "negative entropy" (Schrödinger's term for how living systems contradict the second law). Indeed, there is a school of theorists who have advanced the proposition that energy is somehow a free good and that available energy itself "drives" the process of creating order and organization in the living world (e.g., Kauffman 1995; Holland 1998; Schneider and Sagan 2005). However, living systems must also adhere to the first and only law (so far) of "thermoeconomics," namely, that the energetic benefits (the energy made available to the system to do work) must outweigh the costs required for capturing and utilizing it. From the very origins of life, energy capture and metabolism has played a key role. As biological complexity has increased over time, the work required to obtain and use energy to sustain the system has increased correspondingly. Indeed, improvements in bioenergetic technologies represent a major theme in evolutionary history and, in every case, involved synergistic phenomena.

4 Summarizing the Synergism Hypothesis

How, then, can one account for the evolution of biological complexity? Over the course of the past two decades, the subject of complexity has emerged as a major theme in mainstream evolutionary biology, and a search has been underway for "a Grand Unified Theory" – as biologist Daniel McShea (2015) has characterized it – that is consistent with Darwin's great vision.

As it happens, such a theory already exists. It was first proposed in *The Synergism Hypothesis: A Theory of Progressive Evolution* (Corning 1983), and it involves an economic theory of complexity. The same idea was later independently proposed by John Maynard Smith and Eörs Szathmáry (1995, 1999) in their two books on the "major transitions" in evolution. The Synergism Hypothesis proposes that cooperative interactions of various kinds, however they might occur, can produce novel combined effects – *synergies* – with functional advantages that may, in turn, become direct causes of natural selection. The Synergism Hypothesis is focused on the favorable selection of synergistic "wholes" and the combinations of genes that produce these wholes. The parts (and their genes) that produce the synergies may, in effect, become interdependent units of evolutionary change.

Accordingly, it is the functional (economic) benefits associated with various kinds of synergistic effects in any given context that are the underlying cause of cooperative relationships – and of complex organization – in the natural world. The synergy produced by the whole provides the proximate functional payoffs that may differentially favor the survival and reproduction of the parts (and their genes). The well-known twentieth-century behaviorist psychologist, B.F. Skinner, called it "selection by consequences." In other words, the Synergism Hypothesis is a theory about the unique combined effects produced by the relationships and interactions between things. It could also be called Holistic Darwinism because it is entirely consistent with natural selection theory, properly understood (see Corning 2005).

Patrick Bateson (2013) illustrates this idea with an analogy. The recipe for a biscuit/cookie is rather like the genome in living organisms. It represents a set of instructions for how to make an end product. A shopper who buys a biscuit/cookie selects the "phenotype" – the end product – not the recipe. So, if the recipe survives and the number of cookies multiplies over time, it is only because shoppers like the end product and are willing to purchase more of them. Although it may seem like backwards logic, the thesis is that functional synergy is the cause of cooperation and complexity in living systems, not the other way around.

5 Synergistic Selection

Maynard Smith (1982) also proposed the concept of Synergistic Selection as (in effect) a subcategory of natural selection. Synergistic Selection refers to the many contexts in nature where two or more genes/genomes/parts/individuals have a shared fate; they are functionally interdependent. Maynard Smith illustrated with a formal mathematical model that included a term for "nonadditive" benefits. However, Synergistic Selection is an evolutionary dynamic with much wider scope even than Maynard Smith envisioned. It includes, among other things, many additive phenomena with combined threshold effects and, more important, many qualitative novelties that cannot even be expressed in quantitative terms. There are, in fact, many different kinds of synergy (see Corning 2018). Synergistic Selection focuses our attention on the causal dynamics and selective outcomes when synergistic effects of various kinds arise in the natural world. And it is synergy, and Synergistic Selection, that has driven the evolution of cooperation and complexity in living systems over time, including especially the major transitions in evolution.

One example (among many cited in my 2018 book) is the evolution of eukaryotes. Increased size and complexity can have many functional advantages in the natural world, and eukaryotic cells, inclusive of their complex internal architecture, are on average some 10–15,000 times larger than the typical prokaryote. However, this vast size difference requires orders of magnitude more energy, and the key to solving this functional imperative was a cooperative, symbiotic union between an ancestral prokaryote and an ancestor of the specialized, energy producing mitochondria in modern eukaryotic cells. Not only was this new combination of labor mutually beneficial for each of the two partners but it created a pathway for expanding and multiplying these benefits many times over. Some specialized cells in complex organisms like humans may contain hundreds, or even thousands, of mitochondria. Liver cells, for instance, have some 2500 mitochondria and muscle cells may have several times that number. I refer to it as a "synergy of scale." (See also Lane 2017. See also Ratcliff et al. 2012, 2015, for related work on the evolution of multicellularity.)

6 The Role of Synergy in Evolution

It should be emphasized that many things can influence the likelihood of cooperation and synergy in the natural world – the ecological context, specific opportunities, competitive pressures, the risks (and costs) of cheating or parasitism, effective policing, genetic relatedness, biological "preadaptations," and especially the distribution of costs and benefits. However, an essential requisite for cooperation (and complexity) is functional synergy. Just as natural selection is agnostic about the sources of the functional variations that can influence differential survival and reproduction, so the Synergism Hypothesis is agnostic about how synergistic effects can arise in nature. They could be self-organized; they could be a product of some chance variation; they could arise from a happenstance symbiotic relationship; or they could be the result of a purpose-driven behavioral innovation by some living organism (see below).

It should also be stressed that there are many different kinds of synergy in the natural world, including (as noted above) synergies of scale (when larger numbers provide an otherwise unattainable collective advantage), threshold effects, functional complementarities, augmentation or facilitation (as with catalysts), joint environmental conditioning, risk- and cost-sharing, information-sharing, collective intelligence, animal-tool "symbiosis" and, of course, the many examples of a division of labor (or more accurately, a "combination of labor"). Indeed, many different synergies may be bundled together (a synergy of synergies) in a complex socially organized "superorganism" like leaf cutter ants or *Homo sapiens* (for details, see Corning 2018).

7 Behavior and Teleonomy in Evolution

The idea that behavior has played an important causal role in evolution should be especially emphasized. It is, in fact, a very old idea. Jean Baptiste de Lamarck can claim priority for this insight, along with Darwin's more guarded view. However, there followed a long "dark age" that began with August Weismann's mutation theory and spanned the gene-centered era that followed during most of the twentieth century, although it was punctuated by various contrarians – from Organic Selection theory to George Gaylord Simpson's "Baldwin Effect," Ernst Mayr's "pacemaker" model, C.H. Waddington's "genetic assimilation," and more. Nowadays, even as we are reading genomes and using this information to illuminate biological causation and decipher evolutionary patterns, behavioral processes are more fully appreciated. This has been accompanied by a flood of research on how behavioral influences contribute to the ongoing evolutionary process – from research on phenotypic plasticity to niche construction theory and gene-culture coevolution theory (see Corning 2014, 2018).

However, the theoretical implications of this paradigm shift need to be highlighted. Behavior has a purpose (teleonomy); it is ends-directed. Living organisms are not passive objects of "chance and necessity" – as the Nobel biologist Jacques Monod (1971) put it in the title of his famous book on evolution. Nor is the currently popular concept of "phenotypic plasticity" a sufficient formulation (see West-Eberhard 2003). Organisms are active, goal-oriented participants in the evolutionary process (cybernetic systems) and have played a major causal role in determining its direction over time. An obvious example can be seen in a Darwinized version of Lamarck's (1984/1809) famous giraffes. Naturally occurring variations in the neck lengths of giraffes became a selectively important proximate cause when their ancestors adopted a (purposeful) new food procurement behavior – eating the leaves in the tops of acacia trees.

Indeed, one of the important themes in evolution has been the "progressive" evolution of self-determination (intelligence) and its ever-expanding potency, culminating in humankind. In a very real sense our species invented itself (see Corning 2018). For better and worse, the course of evolution is increasingly being shaped by the "Sorcerer's Apprentice." Monod's mantra needs to be updated. Evolution is a process that combines "chance, necessity, teleonomy, and selection."

8 Quantifying Synergy

It should also be stressed that synergistic effects can be measured and quantified in various ways. In the biological world, they are predominantly related to survival and reproduction. Thus, hunting or foraging collaboratively – a behavior found in many insects, birds, fish, and mammals – may increase the size of the prey that can be pursued, the likelihood of success in capturing prey, or the collective probability of finding a "food patch." Collective action against potential predators – herding, communal nesting, synchronized reproduction, alarm calling, coordinated defensive measures, and more – may greatly reduce an individual animal's risk of becoming a meal for some other creature.

Likewise, shared defense of food resources – a practice common among social insects, birds, and social carnivores alike – may provide greater food security for all. Cooperation in nest-building, and in the nurturing and protection of the young, may significantly improve the collective odds of reproductive success. Coordinated movement and migration, including the use of formations to increase aerodynamic or hydrodynamic efficiency, may reduce individual energy expenditures and/or aid in navigation. Forming a coalition against competitors may improve the chances of acquiring a mate, or a nest-site, or access to needed resources (such as a watering-hole, a food patch, or potential prey). In all of these situations, it is the synergies that are responsible for achieving greater efficiencies and enhancing profitability.

9 Testing for Synergy

There are also various ways of testing for synergy. One method involves experiments, or "thought experiments," in which a major part is removed from the whole. In many cases (not all), a single deletion, subtraction or omission will be sufficient to eliminate the synergy. Take away the heme group from a hemoglobin molecule, or the mitochondria from a eukaryotic cell, or the all-important choanocytes from sponges, or, for that matter, remove a wheel from an automobile. The synergies will vanish.

Another method of testing for synergy derives from the fact that many adaptations, including those that are synergistic, are contingent and context specific and that virtually all adaptations incur costs as well as benefits. Again, the benefits of any trait must, on balance, outweigh the costs; it must be profitable in terms of its impact on survival and reproduction. Thus, it may not make sense to form a herd, or a shoal, or a communal nest if there are no threatening predators in the neighborhood, especially if proximity encourages the spread of parasites or concentrates the competition for scarce resources. Nor does it make sense for emperor penguins in the Antarctic to huddle together for warmth at high-noon during the warm summer months, or for Mexican desert spiders to huddle against the threat of dehydration during the wet rainy season. And hunting as a group may not be advantageous if the prey is small and easily caught by an individual hunter without assistance.

Yet another way of testing for synergy involves the use of a standard research methodology in the life sciences and behavioral sciences alike – comparative studies. Often a direct comparison will allow for the precise measurement of a synergistic effect. Some of the many documented examples in the research literature include:

- flatworms that can collectively detoxify a silver colloid solution that would otherwise be fatal to any individual alone,
- nest construction efficiencies that can be achieved by social wasps compared to individuals, lower predation rates in larger meerkat groups with more sentinels,
- higher pup survival rates in social groups of sea lions compared to isolated mating pairs,
- the hunting success of cooperating hyenas in contrast with those that fail to cooperate,

- the productivity of choanocytes in sponges compared to their very similar, free-swimming relatives called choanoflagellates,
- and the comparison between lichen partnerships and their independently-living cousins.

Some of the relevant examples can be found in comparative genomics (e.g., Berens et al. 2015).

10 "Why Size Matters"

In his important book, *Why Size Matters*, John Tyler Bonner (2006) focused on the critical role of size in evolution and, equally important, the close linkage between size and biological complexity as he defined it, namely, an internal cellular division of labor. Bonner's thesis was that increased complexity (thus defined) in living systems is driven by increases in size. "There are universal rules imposed by size," he tells us (p. x). He also asserts that "size is the supreme regulator of all matters biological" (p. 2). Indeed, "size is a prime mover in evolution … increased size requires changes in structure and function" (ibid.).

It is certainly true that there is an interplay between the physics of size and the engineering and functional challenges associated with building and maintaining a larger organism. Gravity is an obvious problem. And so is the problem of producing and diffusing greater quantities of oxygen, energy, and nutrients throughout a much larger system. However, like many other monolithic theories, the truth in this case probably lies somewhere in the middle. Increased functional capabilities and efficiencies are also necessary as prerequisites for increased size, and the question of which came first might be resolved by viewing the causal dynamics from a longitudinal perspective – as a process of *reciprocal causation* over time (see especially Laland et al. 2011, 2013). It is an argument that goes back to Darwin himself in *The Origin of Species.*

But more important, Bonner's hypothesis begs the question. Why have organisms grown larger over time? Why do we see a progression in evolutionary history from microscopic prokaryotes with their relatively simple internal division of labor to much larger, intricately organized and far more complex eukaryotes, then to multicellular organisms, and, finally, to organized societies composed of many individual organisms, sometimes numbering in the millions? The answer, in brief, is that size is not an end in itself. It arises because it confers various functional advantages – various synergies of scale. These may include such things as improved mobility, more effective food acquisition, more efficient and effective reproduction, and, not least, protection from predators.

11 A Classic Example

Consider, for example, the volvocines, a primitive order of aquatic green algae that form into tight-knit colonies resembling integrated organisms. One of the smallest of these colonies (*Gonium*) has only a handful of cells arranged in a disk, while the *Volvox* that give the volvocine line its name may have some 50–60,000 cells arranged in the shape of a hollow sphere that is visible to the naked eye. Each *Volvox* cell is independent, yet the colony members collaborate closely. For instance, the entire colony is propelled by a thick outer coat of flagella that coordinate their exertions to keep the sphere moving and slowly spinning in the water – in other words, a synergy of scale.

Some of the synergies in the *Volvox* were documented in a study many years ago by Graham Bell (1985), and in more recent studies by Richard Michod (1999, 2007, 2011). The largest of the *Volvox* colonies have a division of labor between a multicellular body and segregated reproductive cells. Bell's analyses suggested some of the benefits. A division of labor and specialization facilitates growth, resulting in a much larger overall size. It also results in more efficient reproductive machinery (namely, a larger number of smaller germ cells). The large hollow enclosure in *Volvox* also allows a colony to provide a protective envelope for its daughter colonies; the offspring disperse only when the parental colony finally bursts apart.

But there is one other vitally important synergy of scale in *Volvox.* It turns out that their larger overall size results in a much greater survival rate than in the smaller *Gonium*. These algae are subject to predation from filter feeders like the ubiquitous copepods, but there is an upper limit to the prey size that their predators can consume. The larger, integrated, multicellular *Volvox* colonies are virtually immune to predation from the filter feeders.

12 A "Grand Unified Theory"?

Albert Einstein long ago observed that "a theory is all the more impressive the greater is the simplicity of its premises, the more different are the kinds of things it relates and the more extended its range of applicability." I believe it is both possible and appropriate to reduce a fundamental aspect of the evolutionary process, in nature and human societies alike, to a unifying theoretical framework. Like the concept of natural selection itself, the Synergism Hypothesis involves an "umbrella term" (an open-ended category) that identifies a common causal principle across a very diverse array of phenomena. Synergistic Selection focuses our attention on the causal role that functional synergies have had at every step in the evolution of biological complexity, beginning with the origins of life itself, and especially including the major transitions in evolution; the "economic" benefits have always been the key (see Corning 2018).

The Synergism Hypothesis can also account for the unique trajectory of human evolution, including the transformative influence of cultural evolution. Synergistic behavioral and cultural innovations played a key role at every stage. (There are three chapters in my 2018 book related to this thesis.) It can also help to explain warfare in human societies, as elsewhere in the natural world. Among other things, warfare is a highly synergistic phenomenon (see Corning 2007b).

The Synergism Hypothesis also encompasses the role of both "positive" and "negative" synergies and their selective consequences for a given organism, group, or species. One obvious example is how organized, cooperative predation may be viewed very differently by a group of predators and their prey. Another example is how individuals and corporations in human societies may benefit in various ways from burning fossil fuels, yet their combined actions also produce global warming (a negative synergy of scale).

It should also be noted that Synergistic Selection is a dynamic that occurs at both the "proximate" (functional) level and at the "ultimate" evolutionary level. Indeed, proximate synergies are in many cases the direct cause of differential survival and reproduction over time. A canonical example, again, is the "ultimate" consequences of some predator-prey interactions.

The Synergism Hypothesis also offers an explanation for the ubiquitous role of cybernetic "control" processes in living systems at all levels. (In humankind, we refer to it, variously, as "management," "politics," and "governance.") As Maynard Smith and Szathmáry (1995, 1999) detail in their two books on the major transitions in evolution, every new form of organization in the natural world represents a distinct "combination of labor" that requires integration, coordination, and regulation/policing. From eukaryotic protists to Adam Smith's pin factory and the emerging global society in humankind, cybernetic governance is a central challenge and a necessary concomitant.

Finally, it should be stressed that Synergistic Selection can also be formally modeled. Two alternative models can be found in my 2015 paper in the *Journal of Theoretical Biology* coauthored with Eörs Szathmáry.

13 Conclusion

Many theorists these days are calling for a new postmodern, post-neo-Darwinian evolutionary synthesis. (See the extended endnote on this point below.) Some theorists advocate the adoption of a more elaborate "multilevel selection" model, e.g., Wilson (1997). Others speak of an "Extended Evolutionary Synthesis" that would include developmental processes and Lamarckian inheritance mechanisms, among other things (Pigliucci and Müller 2010; Jablonka 2013). Denis Noble (2013) has proposed what he calls an "Integrative Synthesis" that would include the role of physiology in the causal matrix.

Whatever the label, it is clear that a much more inclusive framework is needed, one that captures the full dynamics and interactions among the many different causal influences at work in the natural world. We also need to view the evolutionary process in terms of multileveled systems – functional organizations of matter, energy, and information – from genomes to ecosystems. And we must recognize that the level of selection – of differential survival and reproduction – in this hierarchy of system levels is determined in each instance by a synergistic configuration, or network of causes. Indeed, the outcome in any given context may be a kind of vector sum of the causal forces that are at work at several different levels at once.

In the heyday of the Modern Synthesis in the twentieth century, the explanatory framework in evolutionary biology was often truncated to focus on genetic mutations, sexual recombination, and the mathematics of differential selection (changes in gene frequencies) in an interbreeding population. This mathematical framework, albeit with many refinements, remains the theoretical backbone of the discipline to this day.

The fundamental problem is that it explains very little. Natural selection (properly understood) is not an external causal agency or a "mechanism." It is a metaphor – an umbrella term for a wide-open framework that encompasses whatever specific factors may influence biological continuity and change in any given environment. Equally important, it is no longer tenable to view genetic mutations as the primary source of creativity in evolution. There are many different sources of innovation. In the words of Denis Noble (2014) and his co-authors: "DNA does not have a privileged place in the chain of causality." (See also Woese 2004; Woese and Goldenfeld 2009.)

What is needed going forward is a broadly ecumenical paradigm that would provide more of a work plan than a finished product. Perhaps it could be characterized as an "Inclusive Synthesis." It would be an open-ended framework for explaining how, precisely, natural selection "does its work" in any given context (what causal factors influence adaptive changes). It would also represent an ongoing work-in-progress rather than a completed theoretical edifice.

However, the historical process through which multilevel biological systems have evolved over time can be framed as a sequence of major transitions in complexity – from the very origins of life itself to the emerging global society that humankind is now engaged in creating (for better or worse). And, at every level in this hierarchy, we can see the driving influence of synergy and Synergistic Selection. From an evolutionary/biological perspective, complexity has a purpose – indeed, many purposes. In any case, biological complexity must ultimately pass the test of being useful for survival and reproduction. Cooperation may have been the vehicle, but synergy was the driver. The arc of evolution bends toward synergy.

Endnote

Over the past few decades the fundamental tenets of neo-Darwinism have been convincingly challenged. It seems that organisms are active participants in shaping the evolutionary process. There is now a paradigm shift under way from an atomistic, reductionist, gene-oriented, mechanistic (robotic) model to a systems perspective in which "purposeful" actions and informational processes are recognized as fundamental properties of living organisms at all levels. In his important book, *Evolution: A View from the 21st Century,* the leading microbiologist James Shapiro (2011, also, 2009) argues that even cells must be viewed as complex systems that control their own growth and reproduction and shape their own evolution over time. He refers to it as a "systems engineering" perspective. Indeed, there is no discreet DNA unit that fits the neo-Darwinian model of a one-way, deterministic gene. Instead, the DNA in a cell represents a two-way, "read-write system" wherein various "coding sequences" are mobilized, aggregated, manipulated, and even modified by other genomic control and regulatory molecules in ways that can influence the course of evolution itself. "We need to develop a new lexicon of terms based on a view of the cell as an active, sentient entity," Shapiro stresses. Echoing the views of a number of other theorists recently, he calls for "a deep rethinking of basic evolutionary concepts." Shapiro cites some 32 different examples of what he refers to as "natural genetic engineering," including immune system responses, chromosomal rearrangements, diversity generating retroelements, the actions of mobile genetic elements called transposons, genome restructuring, whole genome duplication, and symbiotic DNA integration. As Shapiro emphasizes, "The capacity of living organisms to alter their own heredity is undeniable. Our current ideas about evolution have to incorporate this basic fact of life." The well-known senior physiologist Denis Noble (2012, also, 2013), in a recent paper, argues that all the basic assumptions underlying the Modern Synthesis and neo-Darwinism have been proven wrong. Specifically, (1) genetic changes are often very far from random and in many cases are directed by "epigenetic" (developmental) and environmental influences; (2) genetic changes are often not gradual and incremental (Noble cites, among other things, the radical effects of DNA transposons, which have been found in more than two-thirds of the human genome); (3) an accumulation of evidence for a Lamarckian inheritance of epigenetic influences has now reached the flood stage; and (4) natural selection, rather than being gene focused, is in fact a complex multileveled process with many different levels and categories of causation. Carl Woese and Nigel Goldenfeld (2009) in their critique of the modern synthesis characterize life as a "collective phenomenon." And evolutionary theorist Eva Jablonka and her colleagues (Jablonka et al. 1998; also, Jablonka and Raz 2009; Jablonka and Lamb 2014) identify four distinct "Lamarckian" modes of inheritance: (1) directed adaptive mutations, (2) the inheritance of characters acquired during development and the lifetime of the individual, (3) behavioral inheritance through social learning, and (4) language-based information transmission. It could be called the extended genome. In a recent review of the mounting evidence for this Lamarckian view, Jablonka (2013) concludes: "The existing knowledge of epigenetic systems leaves little doubt that non-genetic information can be transmitted through the germ line to the next generation, and that internal and external conditions influence what is transmitted and for how long." The developmental biologist Mary Jane West-Eberhard (2003) goes even further: "Genes are followers, not leaders, in adaptive evolution."

References

Bateson, Patrick P.G. 2013. Evolution, Epigenetics and Cooperation. *Journal of Bioscience* 8(4): 1–10.

Bell, Graham. 1985. Origin and Early Evolution of Germ Cells as Illustrated by the Volvocales. In *Origin and Evolution of Sex,* ed. Harlyn O. Halverson, and Alberto Monroy, 221–256. New York: Alan R. Liss.

Berens Ali J., James H. Hunt, and Amy L. Toth. 2015. Comparative Transcriptomics of Convergent Evolution: Different Genes but Conserved Pathways Underlie Caste Phenotypes Across Lineages of Eusocial Insects. *Molecular Biology and Evolution* 32(3):690–703. doi: https://doi.org/10.1093/molbev/msu330. Epub 2014 Dec 9.

Bonner, John Tyler. 2006. *Why Size Matters: From Bacteria to Blue Whales.* Princeton, N.J.: Princeton University Press.

Corning, Peter A. 1983. *The Synergism Hypothesis: A Theory of Progressive Evolution.* New York: McGraw-Hill.

——— 2005. *Holistic Darwinism: Synergy, Cybernetics and the Bioeconomics of Evolution.* Chicago: University of Chicago Press.

——— 2007a. Control Information Theory: The 'Missing Link' in the Science of Cybernetics. *Systems Research and Behavioral Science* 24: 297–311.

——— 2007b. Synergy Goes to War: A Bioeconomic Theory of Collective Violence. *Journal of Bioeconomics* 9: 104–144.

——— 2014. Evolution 'On Purpose': How Behaviour Has Shaped the Evolutionary Process. *Biological Journal of the Linnean Society* 112: 242–260.

——— 2018. *Synergistic Selection: How Cooperation Has Shaped Evolution and the Rise of Humankind.* Singapore, London, New Jersey: World Scientific.

Corning, Peter A., and Eörs Szathmáry. 2015. 'Synergistic Selection': A Darwinian Frame for the Evolution of Complexity. *Journal of Theoretical Biology* 371: 45–58.

Corning, Peter A., and Steven J. Kline. 1998. Thermodynamics, Information and Life Revisited, part II: Thermoeconomics and Control Information. *Systems Research and Behavioral Science* 15: 453–482.

Davies, Nicholas D., John R. Krebs, and Stuart West. 2012. *An Introduction to Behavioural Ecology* (4th ed.). New York. John Wiley.

Holland, John H. 1998. *Emergence: From Chaos to Order*. Reading, MA: Addison-Wesley (Helix Books).

Jablonka, Eva. 2013. Epigenetic Inheritance and Plasticity: The Responsive Germline. *Progress in Biophysics and Molecular Biology* 111: 99–107.

Jablonka, Eva, and Gal Raz. 2009. Transgenerational Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution. *Quarterly Review of Biology* 84(2): 131–176.

Jablonka, Eva, and Marion J. Lamb. 2014. *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life* (revised edition). Cambridge, MA: MIT Press.

Jablonka, Eva, Marion J. Lamb, and Eytan Avital. 1998. 'Lamarckian' Mechanisms in Darwinian Evolution. *Trends in Ecology and Evolution* 13(5): 206–210.

Kauffman, Stuart A. 1995. *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*. New York: Oxford University Press.

Laland, Kevin N., Kim Sterelny, John Odling-Smee, William Hoppitt, and Tobias Uller. 2011. Cause and Effect in Biology Revisited: Is Mayr's Proximate-Ultimate Dichotomy Still Useful? *Science* 334: 1512–1516.

Laland, Kevin N., John Odling-Smee, William Hoppitt, and Tobias Uller. 2013. More on How and Why: Cause and Effect in Biology Revisited. *Biology and Philosophy* 28(5): 719–745.

Lamarck, Jean-Baptiste. 1984/1809. *Zoological Philosophy: An Exposition with Regard to the Natural History of Animals* (Elliot, H. trans.). Chicago: University of Chicago Press.

Lane, Nick. 2017. Serial Endosymbiosis or Singular Event at the Origin of Eukaryotes? *Journal of Theoretical Biology* 434:58–67. doi: https://doi.org/10.1016/j.jtbi.2017.04.031

Maynard Smith, John. 1982. The Evolution of Social Behavior – A Classification of Models. In *Current Problems in Sociobiology,* ed. The King's College Sociobiology Group, 28–44. Cambridge, UK: Cambridge University Press.

Maynard Smith, John, and Eörs Szathmáry. 1995. *The Major Transitions in Evolution.* Oxford: Freeman Press.

——— 1999. *The Origins of Life: From the Birth of Life to the Origin of Language.* Oxford: Oxford University Press.

McShea, Daniel W. 2015. Bernd Rosslenbroich: On the Origin of Autonomy; A New Look at the Major Transitions (book review). *Biology and Philosophy* 30(3): 439–446.

Michod, Richard E., 1999. *Darwinian Dynamics, Evolutionary Transitions in Fitness and Individuality*. Princeton, NJ: Princeton University Press.

——— 2007. Evolution of Individuality During the Transition from Unicellular to Multicellular Life. *Proceedings of the National Academy of Sciences* 104: 8613–8618.

——— 2011. Evolutionary Transitions in Individuality: Multicellularity and Sex. In *The Major Transitions in Evolution Revisited,* ed. Brett Calcott, and Kim Sterelny, 169–197. Cambridge, MA: MIT Press.

Monod, J. 1971. *Chance and Necessity* (Wainhouse, A. trans.). New York: Alfred A. Knopf.

Noble, Denis. 2012. A Theory of Biological Relativity: No Privileged Level of Causation. *Interface Focus* 2:55–64.

Noble, Denis. 2013. Physiology is Rocking the Foundations of Evolutionary Biology. *Experimental Physiology* 98(8): 1235–1243.

Noble, Denis, Eva Jablonka, Michael J. Joyner, Gerd B. Müller, and Stig W. Omholt. 2014. Evolution Evolves: Physiology Returns to Centre Stage. *Journal of* Physiology 592.11: 2237–2244. doi: https://doi.org/10.1113/jphysiol.2014.273151/epdf

Pigliucci Massimo, and Gerd B. Müller. 2010. *Evolution – The Extended Synthesis.* Cambridge, MA: MIT Press.

Ratcliff, William C., R. Ford Denison, Mark Borrello, and Michael Travisano. 2012. Experimental Evolution of Multicellularity. *Proceedings of the National Academy of Science*s 109: 1595–1600.

Ratcliff, William C., Johnathon D. Fankhauser, David W. Rogers, Duncan Greig, and Michael Travisano. 2015. Origins of multicellular evolvability in snowflake yeast. *Nature Communications* Volume 6, Article number: 6102. doi: https://doi.org/10.1038/ncomms7102

Schneider, Eric D., and Dorion Sagan. 2005. *Into the Cool: Energy Flow, Thermodynamics, and Life.* Chicago: University of Chicago Press.

Schrödinger, Erwin. 1944. *What is Life? The Physical Aspect of the Living Cell.* Cambridge, UK: Cambridge University Press.

Shapiro, James A. 2009. Revisiting the Central Dogma in the 21st Century. *Annals of the New York Academy of Sciences* 1178: 6–28.

Shapiro, James A. 2011. *Evolution: A View from the 21st Century.* Upper Saddle River, NJ: FT Press Science.

Spencer, Herbert. 1892/1852. The Development Hypothesis. In *Essays: Scientific, Political and Speculative.* New York: Appleton.

West-Eberhard, Mary Jane. 2003. *Developmental Plasticity and Evolution*. Oxford: Oxford University Press.

Wilson, David Sloan. 1997. Introduction: Multilevel Selection Theory Comes of Age. *American Naturalist* 150 (Suppl.): S1–S4.

Woese, Carl R. 2004. A New Biology for a New Century. *Microbiology and Molecular Biology Review* 68(2):173–86.

Woese Carl R., and Nigel Goldenfeld. 2009. How the Microbial World Saved Evolution from the Scylla of Molecular Biology and the Charybdis of the Modern Synthesis. *Microbiology and Molecular Biology Review* 73(1):14–21.

Part IV
Society

Movement and Spatial Specificity Support Scaling in Ant Colonies and Immune Systems: Application to National Biosurveillance

Tatiana Flanagan, Walt Beyeler, Drew Levin, Patrick Finley, and Melanie Moses

1 Introduction

Analyzing and responding to biosurveillance information is fundamental for the health of modern societies. Strategies for the detection of disease outbreaks have received considerable attention from researchers and policymakers, but making the development of a detection and actionable analysis approach scalable over large areas, and applicable to diverse populations, environments, and social contexts, remains a formidable challenge. Disease outbreaks are inevitable, and early detection is necessary for adequate containment. However, more data exist than can effectively be analyzed, and those data are distributed across large geographic areas. Further, data sources are diverse (Gajewski et al. 2014; Althouse et al. 2015), noisy, variable in space and time, and have locally distinct contexts that can affect their interpretation. Therefore, intelligent, distributed, and adaptive sampling, detection and response is required. Sampling and analyzing data with the goals of reducing detection time and maximizing accuracy is challenging. Additionally, scaling results to large and diverse areas is not feasible without a distributed approach. Current biosurveillance architectures that enable both local and global data analysis could also benefit from a scalable approach with decentralized authority to detect and respond.

T. Flanagan (✉)
Department of Computer Science, University of New Mexico, Albuquerque, NM, USA

Sandia National Laboratories, Albuquerque, NM, USA
e-mail: tpflana@sandia.gov

W. Beyeler · D. Levin · P. Finley
Sandia National Laboratories, Albuquerque, NM, USA

M. Moses
Department of Computer Science, University of New Mexico, Albuquerque, NM, USA

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_15

How effectively strategies used by biological systems perceive information depends on how it is distributed in time and space. Biological systems respond to distributions of information across time by bet-hedging on uncertain information (Donaldson-Matasci et al. 2010) and evolve rules (aka algorithms) that produce behaviors, or structures, shaped in response to the dynamics of the environment (Gordon 2016). Distributed and parallel solutions adapted to the distribution of resources or information in a system's environment provide efficient strategies and structures adapted to the distribution of information. Marzen and Dedeo (2017) provide a theoretical framework for achieving optimal information perception under competing constraints on accuracy and misclassification. This approach involves the evolution of a perceptual mapping operating on environmental information where the environment is defined by a stationary probability distribution as well as by a penalty function that imposes costs for misclassification by the perceptual mapping. Mapping accuracy is constrained by a cost proportional to the sophistication of the mapping as measured by the mutual information between the perceived state and the environment that induces the perception. This work highlights trade-offs in the coupling between perception and environment but focuses on a unitary rather than a distributed perceptual system.

Complex systems in nature have evolved solutions adapted to distributed search and response. Immune systems and ant colonies scale search and automated response in dynamic environments using a distributed approach (Banerjee and Moses 2010) in which many agents can sample information from different locations. Ant colonies and immune systems have evolved solutions that rely on distributed local sensing to perceive their environment, establish an appropriate response, and dynamically adapt their response over time according to the spatial distribution of resources and varying complexity of the environment. The robust, adaptive, and scalable computation realized by biological systems makes them suitable models for addressing problems that require distributed computation that is adaptive and scalable (Moses et al. 2013).

Five RADAR principles proposed by Banerjee and Moses (2010) are common in complex systems and relevant to all systems that seek to adapt to information from dynamic environments. These principles are (1) *R*obustness, achieved by redundancy, flexible diversity and probabilistic response to partial information, (2) *A*daptation to environment signals, (3) *D*ecentralized control for search, (4) an *A*utomated *R*esponse that is as distributed as the search, and (5) are scalable to millions of agents, conferring the ability to act in parallel. In previous work (Banerjee and Moses 2010; Moses et al. 2013), we discussed how each of these principles is evident in ant colonies and immune systems. Extending this work here, we consider how the principle of decentralized control and search adapts to a dynamic distribution of resources, and scales to territory or body size increases. We start by describing biological distributed detection and search systems. Drawing from the immune system and ant colonies, we discuss the strategies that support the scaling of search and detection to large areas, and to dynamic distributions of resources by three adaptations: trafficking, or movement of agents through space; functional specificity or spatial memory; and hubs or temporary resident structures.

Finally, we discuss the application of these concepts to a bioinspired approach for a distributed, adaptive, and scalable biosurveillance system.

2 Biological Systems

Immune systems and ant colonies use a distributed approach that adapts to dynamic distributions of pathogens or resources. Millions of ants in a colony and trillions of cells in the immune system detect food or pathogens locally and scale efficiently with increase in size. Scale invariance is relevant to any problem where distributed detection can positively affect the efficiency of the system. Although ants and immune systems are spatially constrained, with behavior clearly adapted to their environment and the use of movement, memory, and local and global information balance, these systems accomplish their tasks efficiently regardless of organism or territory size (Banerjee and Moses 2010).

2.1 The Immune System

It is unusual for biological systems to be scale invariant. Metabolic scaling theory proposes that most biological rates systematically slow as a function of body size due to the overhead of centralized transport and energy (West et al. 1997; Banavar et al. 2010). Whether immune response times systematically slow with body size is an important theoretical question (Wiegel and Perelson 2004; Perelson and Wiegel 2009; Althaus 2015). There is some evidence of differences in immune cost, replication rates, and resulting duration of infectivity due to body size (Brace et al. 2017; Althaus 2015; Banerjee et al. 2017). However scaling theory would predict that humans (10,000 times larger than mice) would have immune response ten times slower than mice; this has not been observed. We hypothesize that the apparent scale invariance of immune response is due to RADAR principle 3, decentralized search, an example of decentralized processing of information. We propose that scalability is achieved in the immune search for pathogens through three factors, trafficking, memory or functional specificity, and communication.

Mammals rely on their immune system to detect and react to invading pathogens whose distribution in the body varies over time and space. Despite this dynamic environment, and widely varying body sizes, immune response times are nearly scale invariant (Banerjee and Moses 2010). When communication and actions are executed locally, each cell can respond quickly regardless of the size of the system. Distributed processing and the absence of central control lead to immune system computation that is highly scalable, through a combination of strategies such as T cell trafficking, functional specificity, and a balance of local and global communication. Trafficking and specificity allow for better homing to a particular tissue (Wong et al. 2016), thus enabling search strategies to adapt to local context and larger areas. All these strategies are all supported by dynamic structures in the immune system that support the adaptations of search to local context, and larger areas.

Immune cells *traffic* through the body via a partially decentralized infrastructure: the lymphatic network transports immune cells to local regions where they either identify pathogens in lymph nodes or kill pathogens in tissue. In contrast to the systemic, centralized movement of cells through the cardiovascular network, immune cells traffic to particular tissues and recirculate between tissues and the local lymph nodes, through the lymphatic network in a way that balances local and global movement of immune cells (Moses and Banerjee 2011; Banerjee and Moses 2010). The trafficking of T cells is a dynamic process. Following their development in the thymus, naïve T cells continually circulate throughout the body until they encounter foreign pathogens. When naive T cells recognize a pathogen, they divide and express molecules that help fight infection. Of this new population, 90–95% undergoes apoptosis, the rest remain close to the site of infection.

As immune cells circulate, some stay close to the tissues where a pathogen is likely to reside.

Human T cells have unique phenotypes with different degrees of tissue *specificity* (Wong et al. 2016). Surviving T cells give rise to long-lived memory populations (Nolz et al. 2011). Resident memory T cells mediate immune memory, which generates long-lived nonrecirculating cells that reside within the originally infected tissue. These cells are superior to circulating T cells at providing rapid long-term protection against re-infection in specific tissues (Jiang et al. 2012). In contrast to recirculating T cells, resident T cells are positioned for rapid detection and response. Once a virus is detected, resident T cells respond to an infection by using the local tissue environment to recruit immune cells (Rosato et al. 2017). To ensure *communication* of learned pathogens throughout the body, a proportion of memory T cells, like naive cells, circulate throughout the body until they are needed (Omilusik and Goldrath 2017). Memory T cells are maintained by continual recruitment of new cells from the circulation, suggesting a dynamic memory in the immune system that depends on a systemic source (Ely et al. 2006).

Local and dynamic structures throughout the body support efficient search and response strategies. Search is focused in small lymph nodes where antigen-bearing cells are concentrated. Immune cells conduct efficient parallel search in lymph nodes where immune cells are introduced to potential pathogens in a small search space. Immune cell movement is evolved to sample multiple pathogens quickly (Fricke et al. 2016). Immune cells are guided by chemokines (Banerjee et al. 2011; Levin et al. 2016) and structural cues (Mrass et al. 2017) in tissues. Guidance to sites of infection particularly speeds up search in large animals more than in small animals, decreasing time to clear infection in humans by orders of magnitude more than in mice (Banerjee et al. 2010).

To support dynamic specificity of local search and response strategies, the immune system evolved temporary resident structures that allow the immune system to dynamically adapt to its environment. One such structure is the inducible bronchus-associated lymphoid tissue (iBALT), an immune system structure that develops in lung tissue in response to tissue inflammation. Present in larger numbers when local inflammation is chronic, iBALTs provide a local site for T cell priming and B cell education to clear future infections in nearby tissue and enhance protective immunity against future respiratory pathogens (Foo and Phipps 2010).

An efficient and proportionate response derives from having distributed memory realized through circulating and resident memory T cells. Antibodies are produced faster and more efficiently where the body experiences the same disease in the same general location. Thus a memory and movement dependent efficiency of response and proportionate response (not overreacting to the threat) are essential features of well-trained immune systems.

Distributed processing and the absence of central control in the immune system lead to scalable processing and response. Scalability is realized through a combination of T cell trafficking, functional specificity, and a balance of local and global communication. All of these factors, which make the immune system scalable, are possible because the immune system uses decentralized recognition of self from others that may be potential pathogens and remembers previously encountered pathogens (Von Boehmer 1990) so that, for example, T cells are able to kill cells in tissue without any centralized input: negative selection ensures that T cells attack only nonself cells. Negative selection is the hallmark of the adaptive immune system enabling encoding of self and other. In scaling, these strategies, supported by dynamic structures, adapt search to local context, and larger areas.

2.2 *Ant Colonies*

Ant colonies rely on individual foragers to search for food sources and bring them back to the colony. The distribution of these resources varies in space and time, so ant colonies use diverse foraging strategies (Lanan 2014) that emerge in response to direct and indirect social cues (Gordon 2010). Distributed foraging and lack of central control in ant colonies lead to scalable foraging and response. Scalability in ant foraging is achieved through a combination of foraging behaviors that involve distributed movement of individual ants, learning, remembering, and communicating the location of food resources. These foraging strategies are all supported by dynamic colony structures, nests and trails, and adapt to local context and territory size. We hypothesize that scalability of foraging in ant colonies, like the immune system, is due to RADAR principle 3, decentralized search, a means to process distributed information. This is possible through three strategies, autonomous movement of individual ants, memory, and communication that allows the colony to learn from individual sampling of information.

The *movement* of individual ants in a colony reflects different strategies to retrieve food for the colony. The repertoire of foraging behaviors reflects the distribution of resources (Levin et al. 2015). In prior modeling work we demonstrated that ant colonies effectively use different collective foraging strategies that respond to these distributions by combining a small set of simple behaviors tuned for a particular environment (Letendre and Moses 2013; Hecker and Moses 2015). These combined strategies make for effective search among large numbers of individuals connected by a distributed communication network. The resulting behaviors are not directed by any individual ant but, rather, emerge from interactions among individuals and from the interaction of the individuals with its local environment, where ants perceive information about the distribution of resources in their territory using only local sensing. An individual ant can learn and *memorize* information about the location of resources only from a small portion of its environment and respond to local conditions. However, the sampling of environmental information by ants through their movement, individual sensing, and *communication* among them through local interactions tends to overcome individual errors, improving collective function on average. The combination of movement, memory, and local perception paired with communication increases the repertoire of responses to varying quality of foraging sites.

As colonies and their territories grow, search and communication strategies must vary accordingly. Ant colonies use additional behaviors and *dynamic structures* that allow them to retain efficiency when searching. In large ant colonies, energy constraints prevent moving resources to one central nest. In response, ant colonies choose to deploy temporary structures (nests or trails) closer to locations where resources are more likely to occur. In a strategy analogous to the immune system, ant colonies distribute their nests, making their foraging spatially specific to a smaller area and allowing them to use behaviors adapted to the local environment. This is the case in polydomous ant colonies, which have evolved strategies to deal with the diminishing returns of central place foraging by establishing multiple interconnected nests with decentralized foraging. Argentine ants support distributed, adaptive foraging through dynamic foraging structures (nests and trails) that exist only when needed. Ephemeral trails connect to persistent trails, providing efficient routing, just as with virtual networks like cell phone towers (Flanagan et al. 2013).

The ability of ant colonies to scale foraging to large territories using distributed, dynamic, and adaptive structures leads to scalable processing and response. Like the immune system, ant colonies can efficiently scale to large territories through a combination of movement, local foraging, and a balance of local and global communication; they support these strategies by establishing dynamic support structures, temporary nests, and ephemeral trails.

3 Scale Invariant Biosurveillance

From these principles and scaling strategies used by biological systems we posit that memory and movement are fundamental properties of adaptive biosurveillance. To maximize representation accuracy and minimize detection time, a combination of trafficking and spatial specificity is necessary. This can be achieved by sampling adaptively and locally, matching the dynamic distribution of information in space and time, through deployment of distributed motile sensors, which can become specific and reside in the local environment, and through structures that create, and allow sensors to learn and respond in close proximity. The number, spatial distribution, functionality, and behavior of sensors will depend on the distribution of information and local context. To *maximize representation accuracy*, trafficking sensors are necessary to detect information that is randomly distributed in space or time. Once information clusters are found, more communication between sensors can lead to an efficient response. T*o minimize detection time*, the use of individual memory, resident detectors, and physical structures analogous to temporary ant colony structures, memory cells, and iBALT can quickly respond to localized events or information, regardless of scale.

Ants and immune systems have evolved strategies that solve distributed search and communication problems (Prabhakar et al. 2012; Dorigo et al. 2006). Some of these strategies mirror or inspire engineered approaches. The multiplace foraging algorithm for robot swarm by Lu et al. (2018) is an example of the efficiency provided by dynamic structures in an engineered system inspired by ant colony behaviors. The study demonstrates how using robotic depots that dynamically adapt to local information in their environment generates more flexible and scalable swarms.

Due to the expansive size, anticipated growth rate and extent of modern biosurveillance data feeds, any potential approach must lend itself well to distributed computation. A balance between local and global information processing can achieve detection that is appropriate geographically and produces optimal response times. Scale invariance in decentralized information processing systems is a must for information systems that operate over a large geographical extent, such as national biosurveillance. As with a partially decentralized immune system, we propose that a scalable design should follows RADAR principles with the addition of memory and movement for scalability, a system that thinks locally, but can act globally. Following these principles will result in a system that is:

1. Robust, redundant, flexible, and stochastic responses to partial information by using sensors that utilize stochastic and inferential change detection able to predict, rather than only statistically describe
2. Adaptive to dynamic environmental information through sensors that are capable of processing multimodal information
3. Decentralized through local sampling and detecting, aggregating and analyzing information locally, using temporary or permanent local nodes as support structures, and increasing the spatial extent for aggregation of data according to the severity of the signal
4. Automated for efficient responses, distributed according to local detection
5. Scalable to millions of agents through the use of a balance of autonomous trafficking through space to search for distributed information and to distribute information remembered locally, utilizing specificity, and resident structures to minimize detection and response time in a dynamic environment

A biosurveillance system consists of detection and analysis of harvested information. Detectors that are motile, able to learn from local context, and use

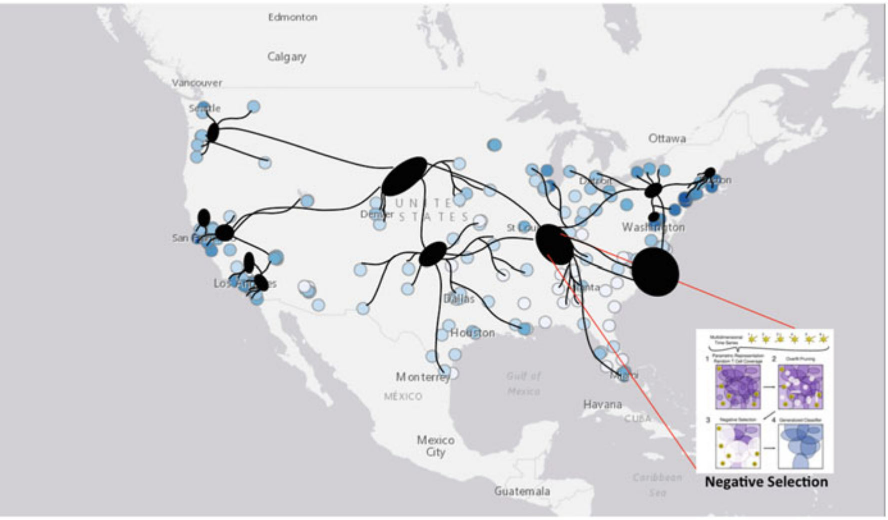


Fig. 1 A bioinspired disease surveillance system with distributed information processing in information nodes (lymph nodes) is shown as black ovals. The flow of information toward and between nodes is shown as black lines. Blue circles represent population density according to shade; darker shades of blue are more densely populated than lighter shades. Each node concentrates and processes information

diverse information streams, along with temporary resident nodes that provide local aggregation, can support scalability in a biosurveillance system. In previous work (Levin et al. 2017), we described the implementation of an anomaly detector for health data based on the human immune system. Our negative selection algorithm detects anomalies in the large, complex data from modern health monitoring data feeds. The parallelized version of the algorithm demonstrates the potential for implementation on a scalable distributed architecture. Using strategies analogous to distributing search into lymph nodes of the immune system, these anomaly detectors have the potential to be motile, to be able to distinguish self and remember encounters with nonself, and to act as trafficking or resident T cells, making them the perfect detector for a robust, adaptive, and scalable national biosurveillance system. Figure 1 illustrates our concept of a National Immune Network. Information nodes can form dynamically according to local information gathered by detectors. By varying cluster size and number of connections, we achieve an optimal global detection time and immediate local detection $O(\log(n/c))$, where n is the number of nodes and c is the number of nodes in a cluster. This "densification" is an emergent property of technological networks (Kleinberg 2004).

There are a number of biosurveillance scenarios that would improve using dynamic, adaptive detection, local training and residence, and global sharing of information. Biosurveillance efforts need to balance memory of events that occur in a specific spatial context while recognizing the motility of both people and pathogens that require the motility of detectors. For example, as a result of sporadic cases, New Mexico hospitals are trained to detect bubonic plague in patients, while most other hospitals in the nation are not. This resulted in the death of a person from New Mexico when a hospital in South Carolina was not able to make a timely detection (Valentine 1983). In a more recent incident (DePalma 2013), a case of bubonic plague was initially suspected of being a bioterrorist attack because it was discovered in New York, a location with a different epidemiological context than New Mexico, the location where the infection occurred. Although the spatial incidence of disease may be more likely in some locations than others, human motility can cause disease spread to span large areas. While local biosurveillance nodes can develop specialized detectors, ensure the presence of locally trained detectors that work in combination with specialized nodes, detector motility provides a powerful tool to share detected information on a global scale. In contrast, Zika is an example of a local outbreak of a virus endemic to specific regions and recently detected in Florida. The Zika outbreak spurred a national effort to expand our detection ability. In this case, detectors informed by the local context can improve the efficiency of biosurveillance by acting locally and communicating globally only when needed, following the spatial patterns of the spread from the center of an outbreak to unexpected locations.

Using a distributed, adaptive biosurveillance system we can also address questions about health behaviors such as opioid abuse. How does opioid abuse manifest itself in different regions of the country? Regional detectors can adapt to the regional context and behavior indicators specific to patterns of opioid abuse and spread. We can detect behaviors within those regions, and compare rural versus urban behaviors. A detector for one region may not work as well as a detector for another, but motile detectors would be a way to combine local context with global communication and eventually adapt to different regional contexts.

The application of RADAR principles is not limited to the implementation of detection and analysis, it can be used to complement organizational practices in national agencies invested in biosurveillance. Managing biosurveillance data requires discriminating access to information, for example due to privacy, national security and other data sharing limitations, but requires transparency at the same time. RADAR suggests that information sharing can be effective when information is communicated locally. Regionalization, benchmarking, and sharing best practices can be seen as organizational analogies for keeping institutional memory/modeling robust in biosurveillance initiatives.

Further studies would benefit from an extended mathematical framework for distributed perception. To characterize perception accuracy, one approach perceives environmental information conditional on the position of perceptual nodes and detectors, and introduces inter-agent communication with the quantification of associated costs and contributions to accuracy. This elaboration could be used to understand the relationship between information variability and distribution/specialization as an optimal detector design, as a function of relevant costs and communication designs. Immune systems and public health networks reward early detection of nonstationary processes. To characterize perception delays, penalties associated with delayed response may be captured as misclassification of stationary processes. Longer-term nonstationarity is also of interest. We are interested not only in the optimal organization of perceptual networks, but also in the ability of that organization to efficiently track dynamics of the environmental signal over time. Studying the properties of optimal solutions will be helpful, however attention to the processes driving structural dynamics will be needed for a design approach that utilizes different perspectives.

Acknowledgments We gratefully acknowledge funding from a Sandia National Laboratories Academic Alliance LDRD Award. TPF and MEM also acknowledge funding from the McDonnell Foundation Complex Systems Scholar Award. We also thank Judy Cannon for her insightful guidance on immune system memory cells, Stephanie Forrest for insightful discussions about negative selection and its relevance to biosurveillance, and Louise Maffitt for helpful editing suggestions.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. The views expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

References

Althaus, Christian L. "Of mice, macaques and men: scaling of virus dynamics and immune responses." *Frontiers in microbiology* 6 (2015): 355.

Althouse, Benjamin M., et al. "Enhancing disease surveillance with novel data streams: challenges and opportunities." *EPJ Data Science* 4.1 (2015): 17.

Banavar, Jayanth R., Melanie E. Moses, James H. Brown, John Damuth, Andrea Rinaldo, Richard M. Sibly, and Amos Maritan. "A general basis for quarter-power scaling in animals." *Proceedings of the National Academy of Sciences* 107, no. 36 (2010): 15816–15820.

Banerjee, Soumya, and Melanie E. Moses. "Scale invariance of immune system response rates and times: perspectives on immune system architecture and implications for artificial immune systems." *Swarm Intelligence* 4, no. 4 (2010): 301–318.

Banerjee, Soumya, Alan S. Perelson, and Melanie Moses. "Modelling the effects of phylogeny and body size on within-host pathogen replication and immune response." *Journal of The Royal Society Interface* 14, no. 136 (2017): 20170479.

Banerjee, Soumya, Drew Levin, Melanie Moses, Frederick Koster, and Stephanie Forrest. "The value of inflammatory signals in adaptive immune responses." In *International Conference on Artificial Immune Systems*, pp. 1–14. Springer, Berlin, Heidelberg, 2011.

Brace, Amber J., Marc J. Lajeunesse, Daniel R. Ardia, Dana M. Hawley, James S. Adelman, Katherine L. Buchanan, Jeanne M. Fair, Jennifer L. Grindstaff, Kevin D. Matson, and Lynn B. Martin. "Costs of immune responses are related to host body size and lifespan." *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology* 327, no. 5 (2017): 254–261.

DePalma Anthony. "Reliving the nightmare of plague, 10 years on". *The New York Times, Health,* January 7, 2013. URL: https://www.nytimes.com/2013/01/08/health/when-the-plague-came-to-new-york.html. Accessed on 3/26/2018. A version of this article appears in print on January 8, 2013, on Page D5 of the New York edition with the headline: Reliving Nightmare Of Plague, 10 Years On.

Donaldson-Matasci, Matina C., Carl T. Bergstrom, and Michael Lachmann. “The fitness value of information.” *Oikos* 119, no. 2 (2010): 219–230.

Dorigo Marco, Gambardella L, Birattari M, Martinoli A, Poli R, et al.. (2006) Ant Colony Optimization and Swarm Intelligence: 5th International Workshop, ANTS 2006, Brussels, Belgium, September 4–7, 2006, Proceedings: Springer.

Ely, Kenneth H., Tres Cookenham, Alan D. Roberts, and David L. Woodland. “Memory T cell populations in the lung airways are maintained by continual recruitment.” *The Journal of Immunology* 176, no. 1 (2006): 537–543.

Flanagan, Tatiana P., Noa M. Pinter-Wollman, Melanie E. Moses, and Deborah M. Gordon. “Fast and flexible: Argentine ants recruit from nearby trails.” *PloS one* 8, no. 8 (2013): e70888.

Foo, S. Y., and S. Phipps. “Regulation of inducible BALT formation and contribution to immunity and pathology.” *Mucosal immunology* 3, no. 6 (2010): 537.

Fricke, G. Matthew, Kenneth A. Letendre, Melanie E. Moses, and Judy L. Cannon. “Persistence and adaptation in immunity: T cells balance the extent and thoroughness of search.” *PLoS computational biology* 12, no. 3 (2016): e1004818.

Gajewski, Kimberly N., Amy E. Peterson, Rohit A. Chitale, Julie A. Pavlin, Kevin L. Russell, and Jean-Paul Chretien. “A review of evaluations of electronic event-based biosurveillance systems.” *PloS one* 9, no. 10 (2014): e111222.

Gordon, Deborah M. *Ant encounters: interaction networks and colony behavior*. Princeton University Press, 2010.

Gordon, Deborah M. “The evolution of the algorithms for collective behavior.” *Cell systems* 3, no. 6 (2016): 514–520.

Hecker, Joshua P., and Melanie E. Moses. “Beyond pheromones: evolving error-tolerant, flexible, and scalable ant-inspired robot swarms.” *Swarm Intelligence* 9, no. 1 (2015): 43–70.

Jiang, Xiaodong, Rachael A. Clark, Luzheng Liu, Amy J. Wagers, Robert C. Fuhlbrigge, and Thomas S. Kupper. “Skin infection generates non-migratory memory CD8+ T RM cells providing global skin immunity.” *Nature* 483, no. 7388 (2012): 227.

Kleinberg, Jon. “The small-world phenomenon and decentralized search.” *SiAM News* 37, no. 3 (2004): 1–2.

Lanan, Michele. “Spatiotemporal resource distribution and foraging strategies of ants (Hymenoptera: Formicidae).” *Myrmecological news/Osterreichische Gesellschaft fur Entomofaunistik* 20 (2014): 53.

Letendre, Kenneth, and Melanie E. Moses. “Synergy in ant foraging strategies: memory and communication alone and in combination.” In *Proceedings of the 15th annual conference on Genetic and evolutionary computation*, pp. 41–48. ACM, 2013.

Levin, Drew, Joshua P. Hecker, Melanie E. Moses, Stephanie Forrest, G. Matthew Fricke, Sarah R. Black, Judy L. Cannon et al. “Volatility and spatial distribution of resources determine ant foraging strategies.” In *Proceedings of the European Conference on Artificial Life (ECAL)*. (2015).

Levin, Drew, Stephanie Forrest, Soumya Banerjee, Candice Clay, Judy Cannon, Melanie Moses, and Frederick Koster. “A spatial model of the efficiency of T cell search in the influenza-infected lung.” *Journal of theoretical biology* 398 (2016): 52–63.

Levin, Drew, Melanie E. Moses, Tatiana P. Flanagan, Stephanie Forrest, and Patrick D. Finley. “Negative selection based anomaly detector for multimodal health data.” *IEEE Symposium Series on Computational Intelligence (SSCI)* (2017).

Lu, Qi, Joshua P. Hecker, and Melanie E. Moses. “Multiple-place swarm foraging with dynamic depots.” *Autonomous Robots* (2018): 1–18.

Marzen, Sarah E., and Simon DeDeo. “The evolution of lossy compression.” *Journal of The Royal Society Interface* 14, no. 130 (2017): 20170166.

Moses, Melanie, and Soumya Banerjee. “Biologically inspired design principles for scalable, robust, adaptive, decentralized search and automated response (radar).” In *Artificial Life (ALIFE), 2011 IEEE Symposium on*, pp. 30–37. IEEE, 2011.

Moses, Melanie, Tatiana Flanagan, Kenneth Letendre, and Matthew Fricke. “Ant colonies as a model of human computation.” In *Handbook of human computation*, pp. 25–37. Springer, New York, NY, 2013.

Mrass, Paulus, Sreenivasa Rao Oruganti, G. Matthew Fricke, Justyna Tafoya, Janie R. Byrum, Lihua Yang, Samantha L. Hamilton, Mark J. Miller, Melanie E. Moses, and Judy L. Cannon. “ROCK regulates the intermittent mode of interstitial T cell migration in inflamed lungs.” *Nature communications* 8, no. 1 (2017): 1010.

Nolz, Jeffrey C., Gabriel R. Starbeck-Miller, and John T. Harty. “Naive, effector and memory CD8 T-cell trafficking: parallels and distinctions.” *Immunotherapy* 3, no. 10 (2011): 1223–1233.

Omilusik, Kyla D., and Ananda W. Goldrath. “The origins of memory T cells.” *Nature* 552, no. 7685 (2017): 337–339.

Perelson, Alan S., and Frederik W. Wiegel. “Scaling aspects of lymphocyte trafficking.” *Journal of theoretical biology* 257, no. 1 (2009): 9–16.

Prabhakar, Balaji, Katherine N. Dektar, and Deborah M. Gordon. “The regulation of ant colony foraging activity without spatial information.” *PLoS computational biology* 8, no. 8 (2012): e1002670.

Rosato, Pamela C., Lalit K. Beura, and David Masopust. “Tissue resident memory T cells and viral immunity.” *Current opinion in virology* 22 (2017): 44–50.

von Boehmer, Harold. “Developmental biology of T cells in T cell-receptor transgenic mice.” *Annual review of immunology*, no. 1 (1990): 531–556.

West, Geoffrey B., James H. Brown, and Brian J. Enquist. “A general model for the origin of allometric scaling laws in biology.” *Science* 276, no. 5309 (1997): 122–126.

Wiegel, Frederik W., and Alan S. Perelson. “Some scaling principles for the immune system.” *Immunology and cell biology* 82, no. 2 (2004): 127.

Wong, Michael Thomas, David Eng Hui Ong, Frances Sheau Huei Lim, Karen Wei Weng Teng, Naomi McGovern, Sriram Narayanan, Wen Qi Ho et al. “A high-dimensional atlas of human T cell diversity reveals tissue-specific trafficking and cytokine signatures.” *Immunity* 45, no. 2 (2016): 442–456.

Valentine, Paul. “Bubonic Plague Feared in Death Of Girl in S.C.”. *The Washington Post,* August 4th, 1983. URL: http://www.washingtonpost.com Accessed on 3/6/2018.

Special Operations Forces as a Global Immune System

Joseph Norman and Yaneer Bar-Yam

1 Executive Summary

Special Operations Forces (SOF) provide war-fighting capabilities that complement conventional forces. A conceptual framework is needed to clarify and differentiate the role of SOF within the larger military system to aid decision-makers in identifying when it is necessary and appropriate to utilize SOF and when conventional forces are better suited.

Here, we propose a correspondence between the role SOF may serve and that of the immune system in complex organisms.

In organisms, the immune system is composed of many semiautonomous components and is responsible for sensing and acting on fine-grained, high-complexity disturbances that may harm the growth and functioning of healthy tissue. It must differentiate between *self* and *other* by having an intimate knowledge of the character of local tissue, detecting agents and behavior that pose a threat. When functioning effectively the immune system eliminates harmful agents without disrupting the normal behavior of healthy tissue.

Much like an organism, global civilization is composed of a collection of diverse *social tissues*, each with its own distinct form, way of living, and functional role

J. Norman (✉)
Applied Complexity Science, LLC in Richmond, New Hampshire, USA

Real World Risk Institute in New York City, New York, NY, USA

New England Complex Systems Institute, Cambridge, MA, USA

Y. Bar-Yam
New England Complex Systems Institute, Cambridge, MA, USA
e-mail: yaneer@necsi.edu

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_16

within larger communities. SOF are uniquely positioned to develop the knowledge and capabilities to distinguish healthy social tissue and detect and mitigate threatening forces.

Conventional forces, by contrast, are well suited to external threats, and their use in societal challenges may damage social fabric, leading to disrupted and vulnerable states.

Three conditions must be met to enable SOF to eliminate threats while preserving the health of social tissue:

1. Distinctive capabilities and advanced training of special operators—Advanced cultural and language competencies and experience in making difficult decisions in the face of uncertainty enable unmediated interaction with local people. When necessary, they can strike with exacting force.
2. Persistent presence and enduring engagements—Repeated and habitual interaction with local communities at both the individual and institutional levels provides the opportunity to develop necessary cultural attunement.
3. Local autonomy and decision-making—Acting on nuanced information relevant to local conditions engenders the ability to stem local threats. Locally embedded SOF must have the freedom to behave semiautonomously, making many decisions independently of SOF located elsewhere or central command structures. This requires avoiding the tendency to become bureaucratized.

The correspondence between the potential role of SOF and the immune system in organisms is formalized via multiscale control systems theory and a *complexity profile* analysis. Ashby's *law of requisite variety* sets the lower bound of complexity a system must possess to survive and prosper. SOF can provide essential complexity at fine scales, as does the immune system in biological systems.

This correspondence suggests that SOF are uniquely equipped to serve as a global immune system, acting before threats rise to the level of crises, and preserving healthy and diverse social tissue functioning. Future policy decisions will determine the degree to which these unique SOF capabilities are developed and leveraged.

2 Introduction

Throughout history warfare has involved both large-scale conventional conflict, in which armed combatants seek to gain physical advantage over their adversaries, as well as less conventional operations which focus on high-value targets or seek to achieve a desired effect through indirect means. The latter have come to be known as *special operations*, and they are typically carried out by small groups and individuals with distinctive skills, creativity, and often equipment. In 1987, the United States Special Operation Command (USSOCOM) was established to oversee the nation's Special Operations Forces (SOF) for both independent and joint operations.

SOF have their origin in military practice and only recently have attempts been made to articulate a theory of the role of SOF (Kiras 2014; Yarger 2013; Spulak Jr 2007) as a subset of a larger *theory of warfare*. These efforts highlight the need for a framework that provides guidance to decision-makers about when and how to utilize SOF to their greatest effect and when other options are more appropriate.

Here, we present a theory of SOF motivated by mathematical and physical necessity and grounded in complex systems science. We propose a correspondence between the functional role of SOF and that of the immune system in complex biological organisms and a parallel correspondence between conventional forces and the neuromuscular system. According to this theory, SOF play a vital role in sensing and acting in fine-grained, high-complexity environments, complementing conventional forces that sense and act at larger scales.

The theory brings military theory into contact with a body of scientific knowledge and inquiry about the behavior of complex systems and, crucially, the conditions under which they are able to survive and prosper.

The remainder of the article is divided into four sections. First, relevant concepts in the theory of multiscale control systems are reviewed and summarized. Second, these concepts are applied to clarify the functional complementarity of the immune and neuromuscular systems in complex organisms. Third, the functional role of SOF is couched in this theory and brought into correspondence with that of the immune system. Finally, strategy, policy implications, and implementation challenges are discussed.

3 Multiscale Control Systems

3.1 The Law of Requisite Variety and its Limitations

In 1956, W. Ross Ashby formalized in the study of control systems what is known as *the law of requisite variety* (Ashby 1956). In short, the law of requisite variety sets the minimum number of behaviors, or "variety," a system must have to survive and prosper in a given environment. As the number of distinct situations a system encounters increases, the variety of its behavioral repertoire must also increase in order to achieve desired outcomes—or as Ashby put it: "variety destroys variety." This concept is illustrated in Fig. 1. If a system has little variety or is overly-constrained while being exposed to a large variety of stressors (i.e., a complex environment), it will sooner or later fail to achieve desired outcomes. In this article, we will use the terms *variety* and *complexity* interchangeably. Thus an environment with *high complexity* is one with a large variety.

The theory of control systems traditionally deals with systems at a well-defined *scale* of relevant behavior and abstracts away details that are presumed not to be of concern due to the nature of the system or the method of control. For example, if one wanted to construct a robot that could catch a baseball, one need

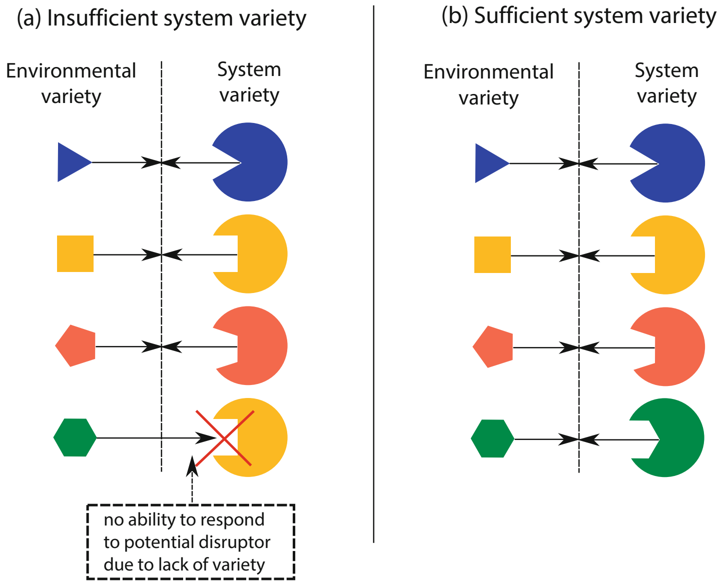


Fig. 1 Requisite Variety. Panel (**a**) shows a system (right) being exposed to environmental disturbances (left). The variety of the environment is greater than the variety of the system, as there are four unique disturbances but only three unique responses. The system lacks the ability to respond to the green hexagon, which will disrupt it. Panel (**b**) shows a similar case, but where the system variety matches the variety of potential disturbances—the system has the requisite variety to respond to all potential disturbances. Systems with variety greater than that of their environment also possess requisite variety

not be concerned with the atomic vibrations ongoing within the baseball, but rather its relevant macroscopic properties like mass, location, and trajectory, and corresponding control variables like joint angles and positions.

In contrast, living systems are exposed to environments with stressors and complexity on multiple relevant scales that must be effectively managed to achieve self-regulation and good overall system health. For example, as organisms we are exposed not only to traffic as we cross the street but also to microscopic organisms that may find our bodies to be suitable homes within which to replicate themselves to our detriment. These two sources of stress exist at scales separated by several orders of magnitude, and our bodies therefore have different strategies in controlling for their potentially harmful effects.

Thus, the law of requisite variety per se is not enough to account for how, say, an organism achieves self-regulation in a complex environment with multiple scales of impinging forces and stressors. Both *variety* and *scale* must be considered for good control in complex multiscale environments (Bar-Yam 2004; Allen et al. 2014).

3.2 Scale/Complexity Trade-Off

There is an inherent trade-off between the scale and complexity of behavior in any system. In order for large-scale behaviors to occur, a large number of components must work coherently or in coordination. Consider, for example, the muscle tension that ultimately gives rise to the movement of a limb. If only one or a small number of muscle fibers become engaged, the scale of the force will be small, and the limb will express essentially no behavior. However if many muscle fibers become engaged at once, a larger-scale force is produced, and the limb will change its position—a large-scale behavior is induced through the coherent activity of many parts.

The flip side of achieving large-scale effects through coherent behavior of many components is that those components are not free to behave independently but are constrained by the role they play in the large-scale behavior. This decreases the variety, or complexity, that can be expressed by the system at small scales.

We can quantify a system's variety at a given scale. In a system with N components that behave independently, the number of states the system can achieve is the product of the number of states each component can take. For instance, if each component can take on 2 states (a *binary* system), the number of total possible states is 2^N. More generally, if n_i denotes the number of states component i can take on, the total number of states of the system is $\prod_i n_i$. The components must act in concert in order to achieve large-scale effect. This necessarily reduces their degree of independence, and the number of states of the system, or variety, is less than $\prod_i n_i$. In other words, due to constraints that prevent each component from behaving independently, the actual variety is less than its maximum would be without constraint. Constraints are indicative of underlying structures that enable large-scale behaviors and variety.

We can summarize the essential trade-off as follows: *demand for variety at large scales necessitates the reduction in variety at smaller scales.*

For a given system, this trade-off can be captured and summarized via the *complexity profile* (Fig. 2) which represents variety, or complexity, as a function of scale (Allen et al. 2014; Bar-Yam et al. 2001, 2013; Bar-Yam 2004, 1997). The "shape" that the complexity profile of a system takes on reflects its structure and behavior and identifies the scales over which they are present. When the smallest components of a system behave essentially independently, there is a maximal amount of variety at the fine scale. However, as we move to larger scales, the independent behavior of all these parts "average out," and we observe no large-scale behavior. When all of the smallest components move together coherently across the entire system, like the atoms in a baseball when thrown, we find behavior at larger scales, with variety varying minimally across scales, and variety at fine scales being reduced dramatically compared to the case of component independence. If you know the flight path of one atom in the baseball, you know them all. For objects like complex organisms, we observe a mixture of these two modes. The variety at the smallest scales remains quite high (though, less than in the case of complete

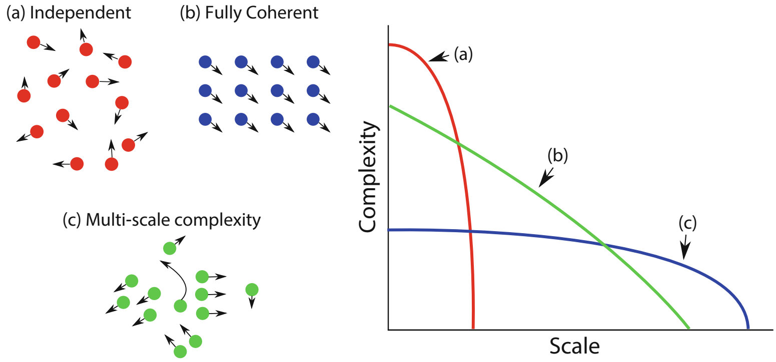


Fig. 2 Complexity Profile. The complexity profile maps system complexity, or variety, as a function of scale. Three example cases are shown. In the case of system components behaving with complete independence (**a**), there is high-complexity at the finest scales, but variety quickly drops off to zero, and no behavior is observed at larger scales. A system in which the components are fully coherent (**b**) has substantially reduced variety at small scales. Intuitively, this is because if we know the behavior of one component, we know them all (i.e., they are constrained). However, this enables large-scale behavior, as the components behave in concert. Systems with multiscale complexity (**c**) have both fine-scale complexity (though less than case (**a**)) and can produce large-scale behaviors (though not as large-scale as case (**b**)). This is achieved by some components behaving in a coherent and coordinated fashion, while others are free to behave semiautonomously

independence), while many of the components are coordinated into larger structures that reduce their independence but achieve larger-scale behaviors.

The different multiscale behaviors also require different control structures to enable actions to be performed in response to environmental challenges or conditions (Bar-Yam 2004). For large-scale behavior, hierarchical control is appropriate. This is because, on the sensing side, large, coherent external events are detected, and irrelevant details are filtered out for high-level decision-making. On the action side, unified decisions can be projected to a large number of agents who behave in concert to achieve a large-scale effect. In contrast, for behavior that responds to fine-scaled, high-complexity challenges that don't require large-scale response, hierarchical control is inappropriate and insufficient. The details lost as information ascends the hierarchy are precisely the ones relevant to small-scale decisions. Moreover, the projection of operational directives from high-levels is necessarily insensitive to these low-level details, constraining agents' behavior and preventing them from responding and adapting to the local context. Instead, distributed networks of agents with minimal hierarchical constraints leave small groups and individuals able to make decisions semiautonomously, retaining sensitivity to local information and enabling adaptation and response to these locally relevant variables.

4 The Immune and Neuromuscular Systems

A familiar example of a multiscale control system with clear differentiation of scale and function is the physiological system that combines the neuromuscular and immune systems in multicellular organisms, such as ourselves. The key differences between the form and function of these two systems lend insight into the nature and role of SOF, and that of conventional forces.

The neuromuscular and immune systems operate concurrently in order to achieve overall system health by responding to disturbances at different scales. The neuromuscular system detects large-scale events and structures in the external environment—dodgeable cars, walkable paths, climbable trees, fall-offable cliffs—and generates coordinated behavior of the gross physical structure of the body in order to leverage opportunities or mitigate harm. These faculties operate in the "Newtonian" macroscopic environment of everyday life to avoid physical damage and provide the resources necessary for physiological function.

The immune system serves a different, but equally important, function. It is distributed and embedded throughout the body and its tissues. It contains a variety of cell types that behave with a large degree of independence—behaviors are not constrained to achieve large-scale coherence as in the neuromuscular system. One of its essential roles is to differentiate *self* and *other* at the cellular and subcellular scales in order to promote the flourishing of healthy tissue and eliminate or neutralize threats when detected. The cells of the immune system sense and act locally, without direct instruction from centralized command structures, though "training" and other functions are centralized in lymph nodes and bone marrow.

The distinction between self and other is not genetic but rather associated with healthy functioning. Consider a cancerous cell and a bacterium that aids in healthy digestion. The former would be appropriately identified as *other*, despite sharing its genome with the host, and the latter *self*, because of its functional harmony with the host.

The body is organized into a collection of heterogeneous tissues and organs which serve various functions that complement one another, forming a self-consistent whole. A well-functioning immune system promotes a healthy system by minimizing the potential for disruption of local tissue. This is a critical point: the integrity of the functional tissue is preserved via the action of a healthy immune system and a disruption to *any* of the tissues in the body disrupts their role and can lead to cascading effects throughout the whole system, including those cells not directly affected by the disruptor, and perhaps organism death.

Notably, the immune system does not *direct* the tissue or instruct its behavior explicitly but rather creates the conditions in which it can express its distinct form without harming itself or other tissue.

The essential differences between the neuromuscular and immune systems are the scales over which they operate and the degree of independence of components that determine the scale and variety. The neuromuscular system operates with a great degree of coordination among its parts, limiting variety at fine scales and producing it at large scales. The components of the immune system, by contrast, behave more independently, making decisions locally and maintaining variety at the fine scales, enabling sensing and acting that preserves good tissue functioning and avoids disruption.

The immune system cannot catch a baseball, and the neuromuscular system cannot eliminate a bacterium. The *only* way the neuromuscular system could effectively combat a micro-disruptor would be through the destruction of functional tissue, an action with irreversible and often system-wide consequences.

5 Multiscale Military Theory and the Functional Role of SOF

Much like an organism, our global civilization is composed of a set of distinct *social tissues*, each with unique character, mode of internal operation, and interfaces with other tissues. Healthy, well-functioning social tissues have internal behaviors that sustain the individuals composing them, such as agriculture, goods production, trading and markets, health services, social gatherings and celebrations, as well as fruitful external interactions with other social systems such as the buying and selling of commodities, products, and services.

When healthy and functional social tissue is disrupted, opportunities are created for malignant forces to gain footholds and grow. This dynamic can be seen, for instance, in the unintended consequences of the invasion of Iraq, which created the opportunity for terrorist networks and other harmful actors to increase their power and influence as normal life was disrupted and power vacuums were created. Moreover, the harm and risks generated by the growth of malignant forces are not confined to the local area where they first manifest.

Because of our global interconnectedness and interdependence, effects cascade causing disruption in other tissues, leading to a domino effect with no straightforward mechanism to halt the expanding impacts (Bar-Yam 2002; Lagi et al. 2015). The recent and ongoing migrant crisis in Europe and beyond provides an example of one form such cascading effects can take.

Global interdependency means any large-scale military intervention, by virtue of disrupting the normal functioning of society, will generate both local and nonlocal unintended consequences even when desired effects are achieved. This is not to suggest that large-scale action is never necessary, but the potential for generating new crises must be weighed carefully whenever it is considered as an option. In many cases, action that does not disrupt local, healthy social behavior is possible, but it requires the right action and agent.

The parallels of the effects of tissue disruption in organisms and in sociocultural systems highlight the need for a “sociocultural immune system”—a fine-grained system for sensing and acting on environmental disturbances at scales smaller than conventional forces are able. In this regard, conventional forces can be likened to the large-scale neuromuscular system in organisms. Acting instead at a small scale presents the possibility of maintaining healthy social tissue and allowing it to flourish. Just as for the immune system, this is not a matter of differentiating “native” and “foreign” but understanding whether an agent is disruptive to overall health.

SOF are uniquely positioned to fulfill this role, possessing the requisite personnel, skills, and training. For this to be realized, policies that impact SOF must be such that they enable their unique capabilities in meeting the high-complexity demand of local cultural systems. We identify three conditions that must be satisfied in order for SOF to serve such a role: special operators with advanced training and distinctive capabilities, persistent presence and enduring engagements, and local autonomy and decision-making. We discuss each in turn.

5.1 Distinctive Capabilities

Much like the cells in the immune system have special forms and functions to fulfill their roles, the distinctive capabilities of special operators enable them to operate in highly complex sociocultural environments. Advanced language and cultural training allows unmediated interaction with local peoples. Special operators’ experience in making decisions in the face of uncertainty allows them to operate in ill-defined “gray zone” conditions.

The need to produce special operators with distinctive capabilities highlights the role of SOF’s high selectiveness and emphasizes the necessity of advanced training in language and culture in addition to combat. These values are articulated in the SOF truths “Humans are more important than hardware” and “SOF cannot be mass produced” (USSOCOM 2016). Preparing special operators to interact directly and make difficult decisions in complex psychosocial, sociocultural, and kinetic environments must be a priority of SOF and their enabling agencies.

5.2 Persistent Presence

The immune system is embedded throughout the tissues of the body to develop and maintain sensitivity to the character of local tissue and respond rapidly to disruptors (Matzinger 2002). Similarly, persistent presence of SOF allows for nuanced relationships to unfold over time and for cultural attunement to be developed at both the individual and institutional levels. SOF embeddedness engenders an understanding of normal conditions and a sensitivity to changes in those conditions and whether they pose a threat. Moreover, presence is necessary for applying rapid and effective action to achieve desired effects with minimal disruption. Just as SOF must recognize “self” in multiple contexts, local cultures must not react to SOF as a foreign entity, i.e., mutual trust must be present, developed through shared history.

Policies should enhance continuity of interaction between SOF and a given sociocultural system even, or especially, when there is no immediate or visible threat. The only way to prevent the growth of malignancies is to be present and active *before* they grow. This is reflected in the SOF truth "competent SOF cannot be created after emergencies occur" (USSOCOM 2016) and Admiral William McRaven's oft-cited comment that one "can't surge trust" (Trulio 2012).

5.3 Local Autonomy and Decision-Making

As the cells of the immune system sense, decide, and act locally in a decentralized manner, being fine-tuned to the character of their local tissues, so too must SOF have the ability to sense, decide, and act locally using their nuanced understanding and experience.

The semiautonomy of SOF is necessary for requisite variety to be achieved in interfacing with high-complexity, fine-grained environments and disruptors.

In human systems, these disruptors manifest at the psychosocial and sociocultural scales. It is possible to take effective action at these scales to eliminate harmful agents without disrupting healthy social tissue functioning. This becomes impossible as the scale of a malignancy grows larger: social tissue will inevitably be damaged by both the malignancy itself and any large-scale force applied in response.

When the decision-making agent is both far removed from and insensitive to the local context, as well as receiving multiple information streams about which decisions must be made, the sensitivity, nuance, and understanding of local SOF are lost. Consequently, the ability to stem malignant forces while they remain small in scale is diminished, and the likelihood of disrupting a social system either accidentally or out of necessity as the scale of harmful actors grows larger increases.

To enable SOF to act without disrupting social tissue, the institutions overseeing SOF must not over-constrain their behavior. As policy- and decision-makers look increasingly to SOF to overcome complex challenges, it is critical that they do not become overly-bureaucratized.

Imperatives that are communicated to SOF must be guided by their role as a protector of local tissue function. Protections from the potential for harm to local tissues, i.e., by civilian collateral damage from operations, must be instituted in a way that retains local autonomy. Detailed instructions on how to carry out missions will prevent them from behaving as necessary for success in high-complexity environments. In technical terms, placing too many constraints on their behavior will reduce their variety below the (requisite) threshold for sensing and acting on fine-grained disruptors. The consequences of this are twofold: (1) SOF will lack the ability to sense and eliminate threats while they remain small and (2) disruption and destruction of healthy social tissue become inevitable as malignant forces grow and large-scale intervention becomes the only means of engagement.

6 Challenges and Implementation

While SOF are uniquely positioned to fulfill an immune-system-like function, there remain significant challenges to successful implementation. Here we summarize some of these challenges.

Developing SOF who are both culturally and linguistically competent, as well as able to execute reconnaissance and surveillance and direct action missions, demands significant investment in training and preparation. Moreover, for any individual operator, there is a trade-off in developing proficiency in any given domain. However, SOF must be able to perform the entire range of activities, from sensing nuanced changes in social conditions to taking actions to eliminate harmful disruptors, to preserving social tissue health.

Fundamental limitations on individual capabilities lead to a need for diversification of roles of SOF. This is manifest already in different types of SOF, as it is in the immune system which uses cells of various types, each of which serves particular roles that complement one another. The relative levels of activity for the different cell types vary depending on circumstance; some cells primarily sense tissue conditions and detect disruptors, while others act to confine and eliminate harmful agents once they are identified. During an infection rapid clonal reproduction (replication) of effective types occurs. Similarly, SOF may embrace and develop specialization of expertise and should be flexible enough to adapt force size and composition in response to changing circumstances.

Maintaining the mental health of special operators must be a priority, and appropriate support systems should be put in place for this. The high complexity of tasks translates into the psychological symptoms of stress, depression, and burnout, common in a high-complexity society more generally but surely for SOF. Moreover, adapting to diverse local contexts creates challenges when switching to home and family environments, a potential component of post-traumatic stress disorder (PTSD). This is a challenge for both the SOF and their families.

Giving special operators a significant degree of autonomy presents challenges and risks that are distinct from those of conventional command and control systems. Care must be taken to ensure social tissue is not damaged by unintentional friendly fire, collateral damage, or intentional "rogue operators." The potential for dysfunction is not unlike autoimmune disorders in complex organisms, and the immune system has developed mechanisms for prevention, though no mechanism is failure proof. Local feedback systems including multiple specialized roles rather than centralized control ones must be in place that put checks on the actions of operators. The structure of these feedback mechanisms must be the subject of intensive study.

Rapid growth in recent years has led to institutionalization of SOF using concepts that may be incorrectly adopted from command control military traditions. Bureaucratization runs counter to the ability of SOF for performing the functions we have identified. Rather than enabling SOF function as it grows, institutionalization may result in undermining the effectiveness of SOF as it becomes more like conventional forces. Alternative structures must be developed. They may be inferred from fundamental complex systems analyses, including correspondence with immune system functions or well-designed experimentation.

Institutional structures and relationships between SOCOM and other enabling agencies, including those within DoD, and other departments of the executive branch such as the State Department, need to be carefully considered. For example, how the agenda of an ambassador of a given region and local SOF should interrelate is an open question. If command and control structures do not appropriately interface with SOF, their unique capabilities will not be utilized effectively. This includes knowing when and, crucially, when *not* to utilize SOF to achieve a desired effect. This article is intended to contribute to this clarification.

7 SOF in the Twenty-First Century

There is no doubt that as a global civilization, we will continue to face fine-grained, high-complexity disruptors that have the potential to grow into larger-scale malignancies. The only way to combat this is to promote and enable the flourishing of healthy social tissues. Multiscale control systems theory makes clear the need for an immune-like system embedded within human social systems. It must be sensitive to and embedded within high-complexity psychosocial and sociocultural environments to make decisions locally based on understanding of a given social system, its nuances, and distinctive qualities.

Like the various tissues arranged into functional organs throughout the body, cultures and social systems do not all look, behave, or function alike. Part of a global strategy for the twenty-first century must be the recognition that cultures cannot simply be “exported” or “projected” onto others without pushback and that behavioral diversity at the collective scale is a natural and healthy part of our human civilization. SOF possess the unique organizational capabilities to be sensitive to the healthy behavior of these diverse “social tissues” while providing the direct and indirect action capabilities to neutralize malignant forces when identified.

Moving forward, a major part of the SOF repertoire must include relationship building. Interpersonal relationships with local individuals form the basis of understanding necessary to discern between harmful and beneficial (or neutral) forces to social health. The ability to perceive and understand local tensions, grievances, typical and atypical interactions, customs, and other nuanced features can serve to generate solutions before the normal functioning of healthy social tissue is threatened. The highly complex and fine-grained nature of this endeavor makes it an unsuitable role for conventional forces—they cannot sense nor act on such a fine scale. A focus on direct action is important when specific disruptors have been identified and not otherwise.

SOF is uniquely positioned to serve as a global immune system, keeping the diverse set of social tissues healthy and reserving large-scale intervention for when it is necessary.

We thank Charles Flournoy and Philip Kapusta for helpful comments and discussion.

References

J. D. Kiras, A theory of special operations: These ideas are dangerous, *Special Operations Journal*, **1**, 75–88 (2014).

H. R. Yarger, 21st century SOF: Toward an american theory of special operations, *Tech. rep.*, DTIC Document (2013).

R. G. Spulak Jr, A theory of special operations: The origin, qualities, and use of SOF, *Tech. rep.*, DTIC Document (2007).

W. R. Ashby, *An Introduction to Cybernetics.* (Chapman & Hail Ltd., London, 1956).

Y. Bar-Yam, Multiscale variety in complex systems, *Complexity*, **9**, 37–45 (2004).

B. Allen, B. C. Stacey, Y. Bar-Yam, An information-theoretic formalism for multiscale structure in complex systems, *arXiv preprint arXiv:1409.4708* (2014).

Y. Bar-Yam, D. Harmon, Y. Bar-Yam, Computationally tractable pairwise complexity profile, *Complexity*, **18**, 20–27 (2013).

Y. Bar-Yam, J. R. Cares, J. Q. Dickmann, W. G. Glenney IV, Multiscale representation phase I, *Final Report to Chief of Naval Operations Strategic Studies Group* (2001).

Y. Bar-Yam, *Making Things Work* (Knowledge Press, 2004).

Y. Bar-Yam, *Dynamics of Complex Systems* (Addison-Wesley Reading, MA, 1997).

Y. Bar-Yam, Complexity rising: From human beings to human civilization, a complexity profile, *Encyclopedia of Life Support Systems (EOLSS), UNESCO, EOLSS Publishers, Oxford, UK* (2002).

M. Lagi, Y. Bar-Yam, K. Z. Bertrand, Y. Bar-Yam, Accurate market price formation model with both supply-demand and trend-following for global food prices providing policy recommendations, *Proceedings of the National Academy of Sciences*, **112**, E6119–E6128 (2015).

USSOCOM, SOF truths. [Online; accessed 8-February-2016].

P. Matzinger, The danger model: a renewed sense of self, *Science*, **296**, 301–305 (2002).

D. Trulio, You can't surge trust—Insights from the opening of the aspen security forum (2012). [Online; posted 27-July-2012].

Applying Evolutionary Meta-Strategies to Human Problems

Valerie Gremillion

1 Introduction

Humans are facing increasingly intricate, entwined, and encompassing problems. Climate change (Raftery et al. 2017; Mora et al. 2017), deepening planetary toxicity (Cribb 2017) affecting our intelligence (Bratsberg and Rogeberg 2018) and genetic expression (Baccarelli and Bollati 2009), and the anthropogenic Sixth Extinction (Ceballos et al. 2017) are only some of the obvious, existential threats *Homo sapiens* faces as a civilization and a species.

While engineers plan Mars colonies unlikely to be successful due to our unacknowledged ignorance of how to successfully construct ecosystems, it is time complexity scientists face a difficult truth: because the problems confronting humanity are multiscale, multidisciplinary, nonlinear, and interdependent, they fall in the wheelhouse of complexity science. Complexity science is necessary, and perhaps obligated, to address humanity's complex future.

Complexity scientists in concert with specialists provide the best chance of addressing our entangled global problems. They are experienced in high-dimensional systems and are armed with concepts that apply across scales and disciplines. They have the theoretical, mathematical, and computational skills to map real world problems onto realistic models. Complexity science offers perspectives, approaches, and models to offset the common domain- and field-specific bias of "If all I have is a hammer, everything looks like a nail."

Support: No institutional support or grant was used for this work

---

V. Gremillion (✉)
Neurocomplexity Consulting, Santa Fe, NM, USA
e-mail: valerie@valeriegremillion.com

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_17

Except for climate models, there have been insufficient attempts by scientists to model real-world high-dimensional systems in all their parameters. Complexity scientists, with their roots in physics and mathematics, prefer to model the least complex or realistic version possible when faced with the real-world intricacy of biology, neuroscience, and ecology, often for reasons of mathematical tractability. Accurate analysis is then hampered by these facile models because difficult or unfamiliar structures, functions, and dynamics have been left out. Due to the tendency to study models based on over-simplistic derivations of real-world systems, even experts are often seeking answers in the wrong part of the state space. Questionable conclusions are then drawn from these low-dimensional creations as nonexperts assume they map accurately to reality.

The increasing hyper-specialization of scientific disciplines into largely inaccessible "silos" also causes a lack of integration into Complex Systems (CS) science of the multitude of emergent principles seen in actual complex systems. Brains, organisms, ecosystems, biochemical networks, etc. offer their own lessons, concepts, and mathematics – but CS approaches rarely utilize them, except in a highly abstracted form such as "deep machine learning." Yet uniting the critical lessons of true complex systems with formalizations from theoretical CS and other fields is most likely to yield principles for creating sustainable societal structures.

One major blind spot blocks success in this arena: few models of our planetary reality explicitly include *humans*, the species with the largest disruptive impact on our natural complex systems. Without understanding the roles or the impacts humanity is having, scientists (and nations) remain unaware of the effects and the problems caused by human presence and their cascaded and convergent impacts.

Currently, we have neither clear, standardized methods to delineate ecological and abiotic interactions with humans, nor standard protocols to marry theory and data with policy. In fact, scientists are often unable to even format some problems across the many disciplines involved, due primarily to hyper-specialization, preventing both insight and solutions. The resulting lack of cross-disciplinary integrators/interpreters hinders overall progress tremendously. More generally it is problematic because no one – neither science, industry, nor government – rewards truly "big picture" or "Seven Generations" (Graham 2008) thinking. Quite simply, *it is no one's job to care about our whole world system.*

And yet is the entire planet that is in disequilibrium, and likely to destabilize most local equilibria.

What scientists *can* do is explore how evolutionarily sustainable, working complex systems (like ecologies, biochemical networks, or real neurons) map onto our human-created ones and vice-versa. Using real systems and the principles we derive from them as guides, we can more clearly grasp what we as a species have done, and learn what is necessary to construct, rebuild, and sustain functioning planetary systems.

2 A Preliminary Framework for Encapsulating Human Problems

Solving humanity's entangled multiscale problems require understanding both their genesis and their context. For that, we need clear and relatively comprehensive delineations of our systems, both natural and human-created. To correctly describe these relationships, we need tools – methods, frameworks, mathematics. Most importantly, we require the conceptual tools and cognitive scope necessary to acquire new comprehension of our planet's dynamic interdependencies, and their network solutions.

Ecological network representations are a good starting place for a particularly useful dynamic and functional description of our systems, even if they lack critical components. From early work on community ecology (Paine 1966, 1980; Cohen 1978; Ives et al. 1999, Williams and Martinez 2000) to path analysis (Shipley 1997), ecosystem dynamics (Ulanowicz 1997), and other formulations (Levins 1975; Patten 1995; Shachak and Jones 1995), ecological network models have embraced approaches from the abstract (e.g., McRae et al. 2008) to highly detailed models focused on specific relationships and aspects of functional ecological roles such as predator, prey, parasite, and pollinator (e.g., Proulx et al. 2005). The latter address functional subsets of ecosystem components and/or dynamics, for instance, food webs (Dunne et al. 2002) and pollination networks (Jordano 1987; Kaiser-Bunbury et al. 2010). While these models often tell us enough to model species and sometimes systems behavior, as each address only a subset of flow types or classes of interaction, many aspects of dynamics and informatics are unknown.

A minority of ecological network models (e.g., Odum 1983; Andrewartha and Birch 1984) have attempted to encompass whole systems, but for purposes of general prediction, action, and policy, few if any have addressed Earth's most pressing problem: humans, and the impacts of their behaviors and institutions on its species and systems.

Given ecology's nonmathematical roots, high-dimensionality, and generally empirical approach, theoretical ecology has had a difficult time parsing large-scale models, at least for our purposes. We can use knowledge from other systems to escape this limitation.

3 Mapping from Neurons to Species

Many biological networks have core similarities, although this fact is largely unexplored. Neurons have been quite well-studied, in comparison to global (or even local) eco-dynamics, though we cannot yet model more than a fraction of their behaviors. But as individual entities, the basic dynamics of neurons are well-described by the Hodgkin-Huxley equations (Hodgkin and Huxley 1952). Together with a huge amount of data, this enables us to note core structure and function (see Fig. 1a, b). Most neurons, for instance, maintain a narrow resting membrane potential range of only 15–30 mV of their total 150 mV range of

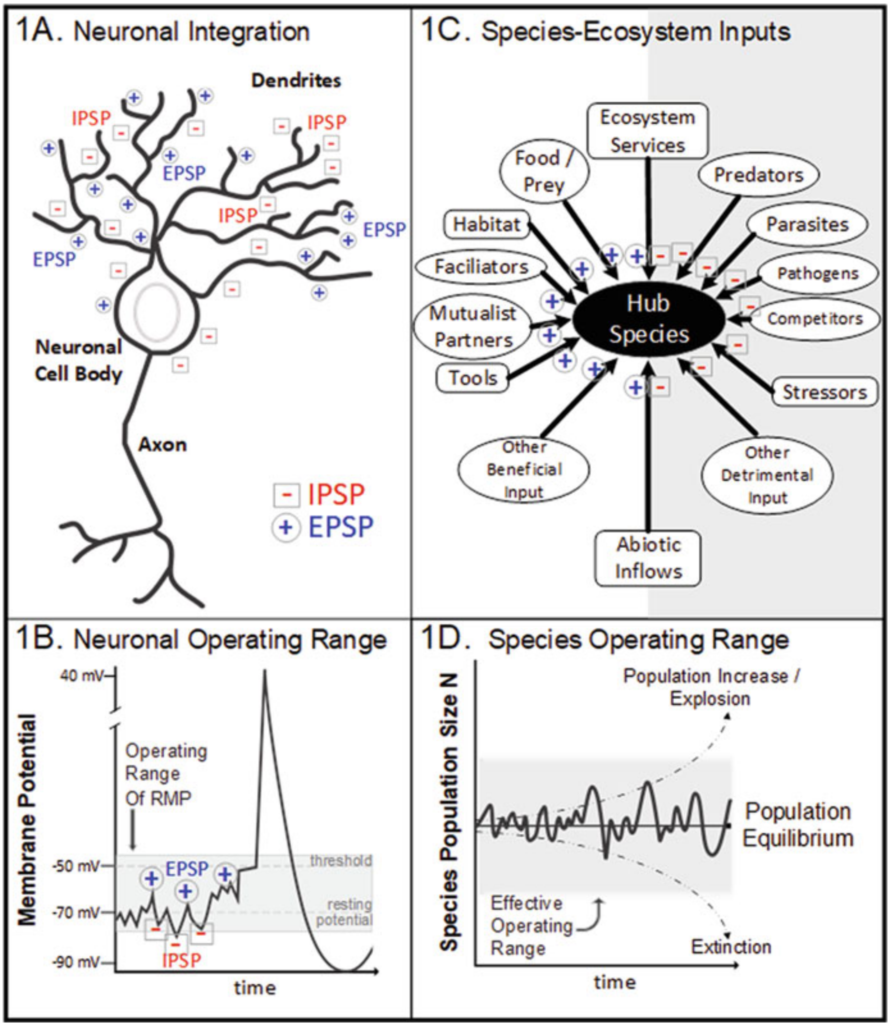


Fig. 1 Deriving concepts from one complex system for application to another. Neurons and species both have their behavior constrained within a narrow range when embedded within a network which provides them with hundreds or more inputs of many classes and dynamics. (**a**) Neuronal integration. Physiological dynamics of a typical neuron mean it has a resting potential (equilibrium state) of between ~−70 mV and −55 mV. This resting potential is displaced by positive inputs in the form of EPSPs (excitatory postsynaptic potentials) which move the neuron closer to its threshold for triggering an action potential or spike or by negative inputs in the form of IPSPs (inhibitory postsynaptic potentials) which move the neuron further away from its threshold for triggering a spike. (**b**) Neuronal operating range.

action. Only in this narrow range does the physiology allow the integration of hundreds/thousands of positive and negative inputs to generate an action potential (spike) capable of signaling to its recipients (ibid). We can infer that as a norm, a neuron's positive (excitatory – moving the membrane potential closer to a spike) and negative (inhibitory – moving the membrane potential further from a spike) inputs are roughly balanced, to keep it within operating range.

Comparisons between real neurons and species (or other entities) in ecological networks are not spurious, as they share many key attributes:

1. Multiple kinds of input types to both neurons and species. Many different classes, pattern, and dynamical response types mean complex structure and complex dynamics are likely.
2. Highly detailed circuitry (in neuroscience, "microcircuitry") may best describe both neural systems, comprised by circuits of classes of neurons and their dynamical relations, and ecosystems, described by circuits of classes of species and their dynamical relations.
3. Opposing positive and negative inputs to each individual entity/node/species constrain dynamical behavior to a subset of the space necessary to maintain a specific operating range. This balancing act at multiple scales is a function of sustainable network structure.
4. Individual entities, whether neurons or species, have a wider behavioral landscape outside of the network than within it; that is, network inputs act to constrain the behavior of these complex entities when they exist within their networks.

Fig. 1 (continued) Summation of both excitatory and inhibitory inputs yield constant variation around the resting membrane potential. Much lower or higher than this, physiological dynamics prevent the possibility of an action potential. In other words, to signal other cells with an action potential, a neuron must stay within a small operating range. What maintains the cell within this operating range is an ongoing relative balance of excitatory (+) and inhibitory (−) postsynaptic potentials which drive the neuron's voltage up and down within the range until it hits threshold and signals with a spike. (**c**) Simple species-ecosystem inputs. A simple visual description of an ecological network that attempts to incorporate all classes of interaction and impact that would alter the population (or other metric) of the entity-in-question or hub species, over time. Functionally beneficial (+) and detrimental (−) inputs roughly balance to create an operating range for a given "species-in-question"/entity, referred to as the hub species as we temporarily treat it as the network's central node. Note that the beneficial or detrimental nature of an input is not intrinsic, but can change if thresholds, magnitude of the flow, or other state changes occur. Water, for example, is necessary and beneficial – but too much is negative and detrimental. Specific functions govern most interactions. (**d**) Integration of the ecocircuitry at any given "hub" organism, species, or entity includes all types of biotic and abiotic interactions. Positive interactions (inputs of energy, materials, and services and information which increase the hub species' population) and negative interactions that decrease the hub species' population, must balance to produce a stable population. If negative inputs predominate over the long term, extinction results; if positive inputs predominate, the species population increases, perhaps exponentially. But species populations also must have an operating range to exist successfully in the long term – dip too low in population, and vulnerability to extinction from a freak virus or storm increases; explosive population growth, and feedback mechanisms (such as depletion of prey) endangers the population

As inputs/constraints are lost, neuron/species/entities may express previously unseen and unexpected behaviors.

While the mapping between neurons and species is not exact, these four analogous aspects of system structure and dynamics form a quite powerful mapping, one whose insights may be drawn on in constructing more realistic ecological and ecocircuitry models.

3.1 A First, Oversimplified Model of "Ecocircuitry"

We propose a theoretical framework for modeling all species, ecosystems, and connected planetary systems. The idea of *ecocircuitry* we define as a realistic ecological network that includes all classes of interaction and behavior, and specific dynamics where possible, but is explicitly extended to include:

- *Homo sapiens*, as a species, and its behaviors, actions, and institutions. This includes all human-related network presences that impact flows of energy, materials, services, and information. We include not merely the physical, such as infrastructure, government, roads, art, corporations, farming, and so on, but beliefs, knowledge, and culture.
- Abiotic input flows. All nonliving things are classed as abiotic – solar heat, rivers, rocks, clouds, weather, and geochemical cycles. Organisms need abiotic inputs (like heat, water), but those inputs have different – yet incorporable – dynamics than species do under evolutionary selection. Many human institutions can usefully be treated as abiotic agents/flows according to their desired optimization and dynamics. We outline this simplest basis for an ecocircuitry model (see Fig. 1c, d) for its illustrative principles.

Ecocircuitry network models include flows of energy, material, services, and information through many classes of inputs/outputs both biotic and abiotic. We posit that the population size of any species varies with the convergence of its network interactions. Since ecologists most often study pairwise interactions, it is difficult to grasp the large number and diversity of inputs to any species-in-question, but inferring the need to maintain a particular operating range, Eq. 1 shows that

$$f\left(\sum_{i=1}^{n} b_i\right) = f\left(\sum_{j=1}^{m} d_j\right) \tag{1}$$

the functional weighting of the sum of all beneficial (positive) inputs, b, must be approximately equal to the functional weighting of the sum of all detrimental (negative) inputs, d. The subscript i enumerates each individual input of the set of n positive inputs, while j does the same for the set of m negative inputs. All functions f

indicate generic, not specific, functions. Applications of functional weightings will be discussed in an upcoming paper.

Equation 2 thus describes the simplest approach to population change over time for the species-in-question/hub of population size N:

$$dN/dt = f\left(\left(\sum_{i=1}^{n} b_i\right) - \left(\sum_{j=1}^{m} d_j\right)\right) \tag{2}$$

where the change in the species population, *dN/dt*, is a function of the sum of the beneficial, or positive inputs, *b*, minus the sum of the detrimental, or negative inputs, *d*.

The hub species is simply the species that is considered the 0th order of the network for a given analysis, with its inputs being 1st order, the inputs of its 1st order inputs being 2nd order, etc. – all within an N-dimensional ecocircuitry network. Parsing the network through the perspective of a single defined “species” as the central hub allows us to grasp the details of ecological and human connectivity patterns beyond food webs, and enables examination of specific evolutionary strategies arising from either particular species interactions or more general rules that govern such interactions.

Examining ecocircuitry through multiple points of view from many individual species supports a form of parallax, necessary due to the size and complexity of these networks. Not only does this provide multiple viewpoints on the system, but it suggests that complete summations over all species’ perspectives may be necessary to accurately assess both unintended consequences and true costs and benefits.

For ecocircuits the interrelated flows of the system (energy/material/information/services) are the drivers of key metrics of “success” – whether increase in population for a species, decrease in infant mortality for a nation, increase in profits for a corporation, and so on – and are addressed in a forthcoming paper.

3.2 A Realistic Basis of Ecocircuitry: Modeling Entities from Organisms to Institutions with a Double Binary Model

Some may say that the simple ecocircuitry model is elegant – but it is too simple. It conveys well the idea of integration of inputs of varying sign or impact, but it is not correct for determining changes in species population, as it portrays only half of the flows. The reason for that is our mapping from neurons to species is useful but incomplete.

Neurobiological network models are primarily focused only on information (spike) output, so flows of materials, energy, and services are ignored in neural equations. This is a primary difference and misses not only classes of input critical to species and ecosystem flows, but their outputs, which are losses of both positive and negative flows. These losses are not a separate and unchanging output like a spike, but rather reflect the loss of both *desired* things – like nutrients or sensors – as well as the loss of *undesired* things, like waste or toxins. An accurate network description of species population thus changes over time as flows wax and wane, requiring a more sophisticated approach that reflects dynamic structure-function relationships.

We call the incorporation of this four-part model of ecocircuitry, the double binary model (see Fig. 2). The double binary model incorporates all aspects of the simple ecocircuitry model, but the dynamics governing it reflect the reality of flows through entities more accurately. We can thus improve the simple binary model above – a balance of beneficial and detrimental inputs *only* – by adding outputs. This produces the "double binary" of both:

- Beneficial and detrimental inputs (positive and negative "gains")
- The output of benefits and detriments (positive and negative "losses")

In this case, it is the balance of the net positive effect (beneficial inputs + detrimental outputs) minus the net negative effect (detrimental inputs + beneficial outputs) that helps all individual entities, thus their species, and thus the system, maintain dynamic stability.

Equation 3 shows that net positive effects must roughly equal net negative effects:

$$f\left(\left(\sum_{i=1}^{n} b_i\right)+\left(\sum_{h=1}^{q} d_h\right)\right)=f\left(\left(\sum_{j=1}^{m} d_j\right)+\left(\sum_{k=1}^{p} b_k\right)\right) \tag{3}$$

The functional weighting of the sum of all beneficial (positive) inputs, b_i, plus the sum of all detrimental (negative) outputs, d_h, must be approximately equal to the functional weighting of the sum of all detrimental (negative) inputs, d_j, plus the sum of all beneficial (positive) outputs, b_k. The subscript i enumerates each individual input of the set of n beneficial inputs, with subscript k enumerates each individual output of the set of p beneficial outputs; the subscript j enumerates each individual input of the set of m detrimental inputs, with subscript h enumerates each individual output of the set of q detrimental outputs.

Equation 4 now describes the simplest realistic approach to population change over time for any "hub" species of population size N:

$$dN/dt=f\left(\left(\sum_{i=1}^{n} b_i\right)+\left(\sum_{h=1}^{q} d_h\right)\right)-f\left(\left(\sum_{j=1}^{m} d_j\right)+\left(\sum_{k=1}^{p} b_k\right)\right) \tag{4}$$

where the change in the species population, dN/dt, is a function of the sum of the net positive effect (beneficial inputs + detrimental outputs) minus the sum of the net negative effect (detrimental inputs + beneficial outputs).

The power of this approach is that it offers a proxy for the stability of an ecological network dependent on each individual species' capacity to optimize

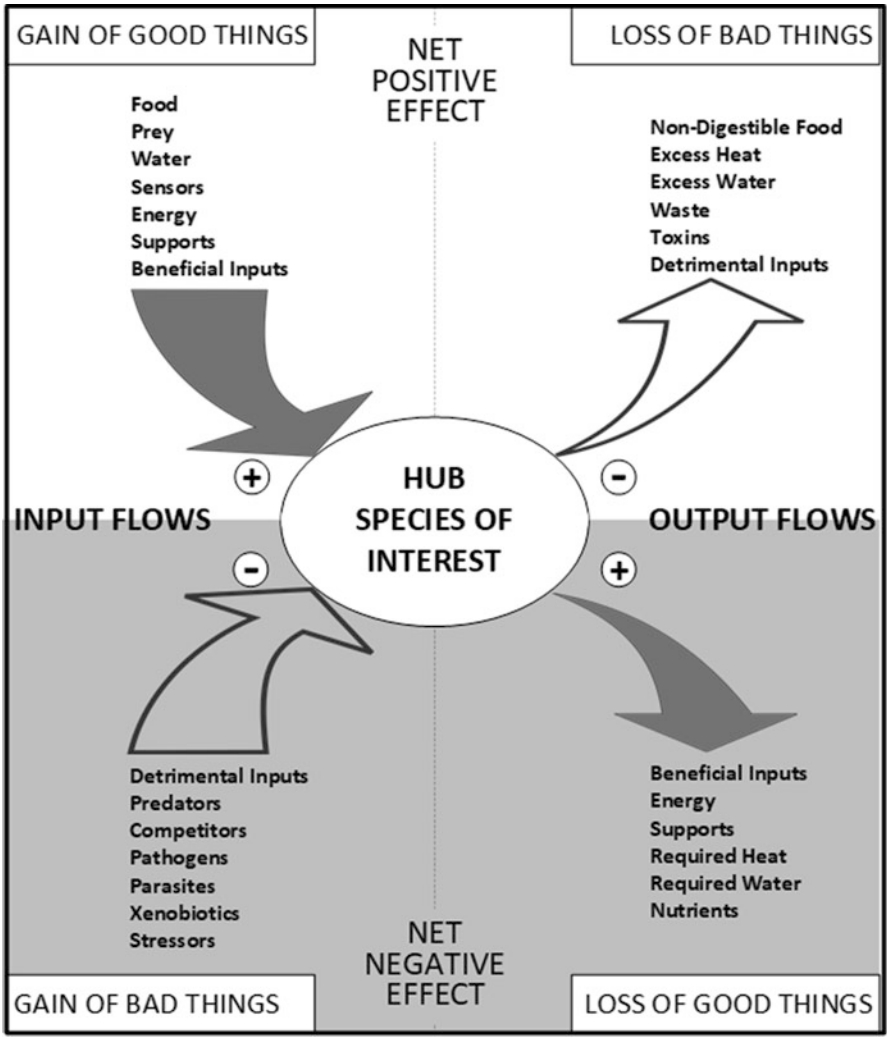


Fig. 2 Ecocircuitry: the double binary model. This ecocircuitry model consists of four quadrants with differing dynamics: two are sets of input flows (intrinsically, added to the hub entity/species-in-question), one beneficial, generating a net positive effect, and one detrimental, producing a net negative effect. The other two are sets of output flows (output, intrinsically subtracted from the hub), one being detrimental things lost, for a net positive effect, and the other beneficial things lost, yielding a net negative effect

its own network inputs by maximizing positive inflows and negative outflows while minimizing negative inflows and positive outflows. Its predators and other interaction partners simultaneously seek to optimize their own network flows. Species networked together in both support and opposition can effectively balance the positive and negative inputs to each species over the long term. Over a densely linked network of species, these coupled dynamics result in a system of adaptive "checks and balances" that tends to homeostatically maintain ecosystem structure and function, often despite frequent and substantial perturbations.

Therefore, neither ecological networks nor the evolutionary strategies they produce is a matter of simple accounting. Rather than merely a "pollinator," a hummingbird is a multifaceted entity that is predator, prey, mutualist, competitor, and many more things – depending on its interactor. What is valuable to one organism – e.g., dung to the dung beetle – may be useless or even negative to the organism that released it. Thus, the network should be viewed through "the eyes" of many species to begin to grasp the multifaceted ecosystem we live in. An insistence on using a single currency in ecocircuitry (as with "money") means a myriad of weighting functions and their interactive impacts will likely be missed as we look through any single lens. An insistence on viewing our system only through the eyes of *Homo sapiens* is equally short-sighted. "Getting rid of all vermin" means one thing to a human with a house, and something different to an ecosystem for which that entity serves multiple purposes.

3.3 Core Evolutionary Strategies and Meta-Strategies: Manipulating the Ecocircuitry

General evolutionary strategies emerge from this model, and intrinsically map to the handling of beneficial and detrimental inputs and outputs. Van Valen (1973) and successors have famously described the "Red Queen" evolutionary dynamic, in which any given species must constantly adapt that they may persist against other, coexisting species that are engaged in similar optimization processes. The connectivity of species in ecological networks and the resulting convergent input to each species implies that all species also continually evolve strategies for manipulating their ecological network inputs. Such strategies may also illuminate evolutionary vs. developmental patterns of gene flow, particularly environmental-epigenetic interactions and resulting natural selection.

Gremillion and Brown (2001) outline core strategies used by any species to optimize its flows:

(i) Altering the number of flows or connections
(ii) Altering the magnitude of flows
(iii) Decreasing the variance of all flows

Thus the core evolutionary strategies as shown in the simple model are quite straightforward *optimizations* in many dimensions:

1. Increase the number and magnitude of beneficial connections. For example, increase nourishment through foraging, locate assistive medicines, and plant nonlocal crops.

2. Decrease the magnitude and/or number of detrimental/connections. For example, kill predators, fight pathogens, and decrease stressors.
3. Decrease the variance of all flows. As variability itself induces instability, it can lead to temporary or permanent vulnerability to extinction. The best example is that of *storage* – which reduces vulnerability to changing conditions; e.g., hibernation, squirrel nut storage, and from grain silos and cisterns, to libraries as stable, external cognitive storage.

3.4 Meta-Strategies

Evolutionary meta-strategies are *strategies for managing the core evolutionary strategies* – managing the number, magnitude, variance, and type of inputs and outputs to the organism, entity, or species in question. A human in ancient Sumeria, for instance, would insure the number of his beneficial inputs is increasing, by using meta-strategies of language and mathematics to explicitly track them. To speed up the output of negative pathogens, human meta-strategies include growing herbal medicines, using chemistry to produce pharmaceuticals, building hospitals. Many, many more such things are meta-strategies. It is on the shoulders of these inventions, insights, and constructions that humanity has built civilization.

There is a vast and uncounted array of such meta-strategies – in part because new ones are constantly being invented. Money/currency, for instance, was invented a long time ago. Banking *systems* took much longer; credit cards another 500 years after that; Bitcoin, the first cryptocurrency, was only invented in 2009. And that is just skimming one class of meta-strategy. A brief and quite incomplete list of some types of meta-strategies may be illustrative:

- Defense: from sharpened sticks and stone axes to predator drones and missiles
- Food/prey: fire, cooking, beneficial spices, stoves and kitchen gadgets, recipes, and gear
- Energy: use of wood, passive solar, fossil fuels, hydro, geothermal, active solar, wind, tidal
- Teaching: stories, apprenticeships, guilds, formal education, hyperspecialized experts
- Water: from baskets to vases to irrigation to plumbing, wells, and vast infrastructure
- Cognitive meta-strategies: concepts, language, money, printing, science, and computation

Corporations and businesses, governments, schools, bureaucracies, and newspapers – even social manners, cultural norms, and languages, in their cognitive-shaping capacities – all impact input/output flows through their cognitive effect on decisions and actions.

But what clearly differentiates a strategy from a meta-strategy?

Say you are a foraging creature. You find a new berry in the forest; you eat it – that is a core strategy: increase your number of positive inputs. But your reaction tells you it is poisonous. If you are a regular forest critter, you never look at that berry again; it is useless for purposes of your core evolutionary strategies. If you are a human, however, you gather that berry, cook it down, and coat your arrows with it. You have now effectively laid the "poison" meta-strategy onto the "throwing spear" meta-strategy, for amplified success in both hunting and protection.

Indirect approaches like this often have multiple benefits: they extend distance from threat, physical or otherwise (many fighting implements do this); they provide leverage through distance; and they are amplifiers of power and capacity to enable needed adaptation, in the same way the Colt 0.45 pistol in the Old West was called "the Equalizer" (Moss 2016). Similarly, the poisoned berry amplifies the lethality, and therefore the operative range of the arrow, as well as broadening the effective target area.

But *Homo sapiens* is hardly the only species to implement innovative meta-strategies, just the most obvious one. We now know a great number of species are tool-users – and tools are the basic meta-strategy. Beyond this, ants keep other ants as slaves (Wilson 1975) and engage in agriculture (Mueller et al. 1998). The externalization of labor and assurance of food sources through cultivation are two meta-strategies used often by *Homo sapiens.* Eco-engineering species like beavers and termites as well as builders of nests and structures are using meta-strategies in combining biotic and abiotic materials to enhance their survival and reproduction. Even coyotes following humans into populated areas, where a fertile foraging ground is available, is a meta-strategy, for it does not use a direct signal to find food, but an indirect one: where humans move is indicative of an abundance of largely uncontested food, in centralized garbage cans.

Once some portion of ecocircuitry is formulated in the double binary model, specific useful evolutionary strategies become clear. Here we outline four examples of meta-strategies that can help generate understanding and solutions, to illustrate potential classes of application.

3.5 Applying Evolutionary Meta-Strategies

Meta-strategies are the water we swim in unknowingly, built on the insights and integration of past and present. Everyone uses meta-strategies many times a day, whether a refrigerator or a meditation technique, a TV show or formal manners, for these and many more form the network of strategies in which humans have become ecological specialists at being *generalists* – of learning, generalizing, and specializing – though the creation of tools and infrastructure are not just for action but for observation, analysis, or playing. *Meta* here can imply meta-meta-meta-meta: layers and levels of utility, benefit, advantage, making the world one in which complexity can be handled more easily if we peel back levels rather than ignore them. A few examples below help illuminate the scope and application of the meta-strategy concept.

4 Meta-Strategy: Perceiving Higher-Order Impacts Through Darwinian Algebra

One of the major problems in ecology and other real-world systems is the complexity of the network itself, and our inability to see, with straightforward logic, far out into the network. This means that humans seeking to understand ecological networks (including *Homo sapiens*) are not easily able to assess the consequences of their actions on the systems in which they play such a powerful role, save for their closest connections and obvious impact.

Thus, humans are rarely able to determine the net effects, beneficial or detrimental, of our manipulations of the ecocircuitry, whether in draining a swamp or in large-scale spraying of mosquito pesticides. Nor do we really try: the combinatorics and higher-order interactions are too complex without a relatively intuitive and logical framework, method, or training.

An ecocircuitry formulation allows us to use the properties of the network, creating a kind of commutative "Darwinian algebra." For instance, we all can grasp the meaning of "the enemy of my enemy is my friend" – either because of personal example or because it is a logical axiom of $[-]\,[-] = [+]$. Using such notation, we can look at interactions at increasing distance and still derive the sign of the impact on the hub species as shown in Fig. 3. While the actual strength or weighting of an input is clearly important, a first approximation of the *sign* of higher-order impacts – negative or positive – is most critical, as it enables us to track the unexpected relationship of cats to clover (see Fig. 3).

This small slice of Darwin's local ecocircuitry shows how a more functionally inclusive and realistic incorporation of multiple kinds of inputs/outputs can illuminate higher-order interactions – generally termed "indirect" in ecology – of facilitation and interference. Figure 3 does not, of course, show the summation of all inputs/outputs at each stage in isolating this part of the red clover ecocircuitry; but it does give us a means to look far out into the network for cascade effects, unintended consequences, costs or benefits, or even interactions between multiple distant indirect inputs or outputs. We can then generate methods by which we may better sense and measure these interactions, invaluable as typically, these are quite difficult as both a network problem and as a practical ecosystem issue.

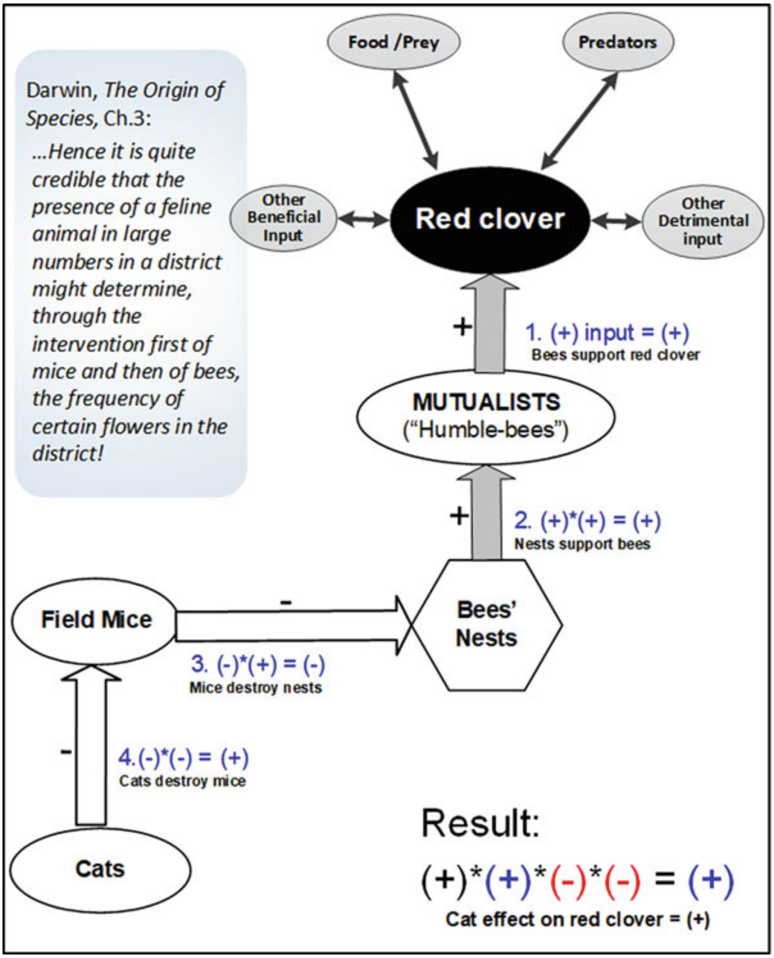


Fig. 3 The enemy of my friend's enemy is my friend: Darwinian "algebra" explains how cats are the friends of red clover. Starting from the species-in-question (hub), Red clover. Its necessary pollination by humblebees (1). The supportive structure of humblebees' nests – which thereby support clover as well (2). Field mice destroy the nests of the bees which pollinate the clover and are therefore its enemy (3). Cats, in this ecosystem, prey on mice and are thereby functional facilitators of clover (4). (Quote: p. 72 of the 2nd edition of Darwin's *Origin of Species* [1859])

5 Meta-Strategy: Language Shapes Acceptance of Climate Change

"Evolutionary meta-strategy," as we coin it, is meant to be fairly encompassing. It includes not only physical expressions of action like trade, art, engineering, agriculture, transport, etc., but also psychological and cognitive meta-strategies such as education, belief structures, and language. The latter leads directly to a multitude of additional strategies such as writing, printing, universities, and science, but also socio-political manipulation and propaganda.

Arguably the most successful capture of real-world complexity is climate change modeling. Brown and Caldeira (2017) found in that arena that the more real system structure/dynamics are included, the more predictive the models. But more facts have not changed all minds. The issue of climate change, intersecting as it does with human nature and cognition, clarifies how correct use of cognitive-*linguistic* meta-strategies could change perception, belief, decision-making, and thus policies. Possible solutions to the dangers of climate change in the United States are currently stymied by "climate change deniers." Examination of this phenomena shows it may be purely a cognitive/linguistic effect, underscoring the need for *strategic* use of meta-strategies.

The initial proclamation of "global warming" led to assumptions that such warming could be ameliorated by turning up the air-conditioning, while record low winter temperatures were interpreted to contradict the entire idea. When terminology switched to "climate change," deniers comforted themselves with the truths that "change happens" and "climate's always changing," irrelevant as those facts are to anthropogenic climate change. In both cases the "languaging" used conveyed the *cause* of many problems to professionals, but lack of context meant no alarm for deniers. Indeed, both original terms sound, to those without context, as though the problem is overstated.

Analysis of this situation through the eyes of meta-strategy can be summed up by the Einstein quote: "Language is an implement of reasoning, in the tool sense of the word." For "climate change," it shows us that the terminology used must instead be indicative of the *effect and its impact* – as impacts are what evolutionary strategies optimize.

We propose substituting the term "extreme climate instability" for its cognitive and emotional connotations as it is far more difficult to shrug off or deny. "Extreme climate instability" indicates both a severity not indicated by previous terms and introduces the concept of climate "instability," a far more alarming concept than change and highly indicative of a far larger scale of weather and even social uncertainty. Moreover, "instability" actually matches the weather (climate) being seen by the average person, so that personal experience and scientific labeling are coherent, not at odds. This is using a meta-strategy to convey contextualized information on which individuals are more motivated to act and thus forestall disaster.

6 Meta-Strategy: Honeybees and the Cause of Colony Collapse Disorder (CCD)

Honeybees (*Apis mellifera)* must apply the core evolutionary strategies to numerous input flows, for their set of potential and realized interactions with other species includes mutualistic pollination relationships with dozens of species of flowering plants (Wilson 1971; Seeley 1997); many nectivorous competitors (Nabhan and Buchmann 1997); and diverse predators, parasites, and diseases (Gould and Gould 1988; Schmid-Hempel 1998).

A large number of causes have been proposed to cause CCD (Steinhauer et al. 2018). Figure 4 shows that massive changes in the ecocircuitry of honeybees have recently occurred in both the number and magnitude of all types of flow, virtually all of them antithetical to the honeybee's self-optimization (Stone et al. 2017). Largely due to *Homo sapiens*, these changes occur in virtually every arena of the circuit. Beneficial inputs (e.g., habitat and plant species for nectar and pollen) are diminished by human activity, while the loss of detrimental things (e.g., toxins) is diminished; detrimental inputs (e.g., pollutions) are increased, while the loss of positive things (e.g., energy, nutrients) is increased. These changes yield an increase in net negative effects and a decrease in net positive effects. Taken together, this degree of negative *convergent causation,* wherein a number of interacting components together produce nonlinear (and often unpredictable) effects, could likely result in colony collapse.

Yet it is possible that the search for a single cause has led to our inability to effectively act. Due to its multicausal nature, we are far less likely to identify *the* cause of CCD. But by believing there to be a single cause, we continue to seek one, and this has prevented us from recognizing that it is the combined negative effects of anthropogenic action, including, as just one factor, the creation of honeybee monocultures brittle to environmental variance, that has enabled this crisis.

While we seek a solitary cause to CCD, when it is far more likely to be the result of multiple convergent causes, we will not find "the answer" nor stop the unnecessary demise of one of our most important pollinators. To halt CCD, many of humanity's net negative effects on *Apis mellifera* may have to be stopped or remedied, at many points of the honeybee's ecocircuitry.

7 Meta-Strategy: *Homo sapiens* Shapes Entire Networks to Attack Its Enemies

Many of humanity's meta-strategies revolve around *defense* – against predators, parasites, pathogens, and competitors. Unlike the teeth and claws of predators, mankind's weapons have been consciously constructed. Over time *Homo sapiens*' meta-strategies around defense have shown numerous themes and patterns. For example:

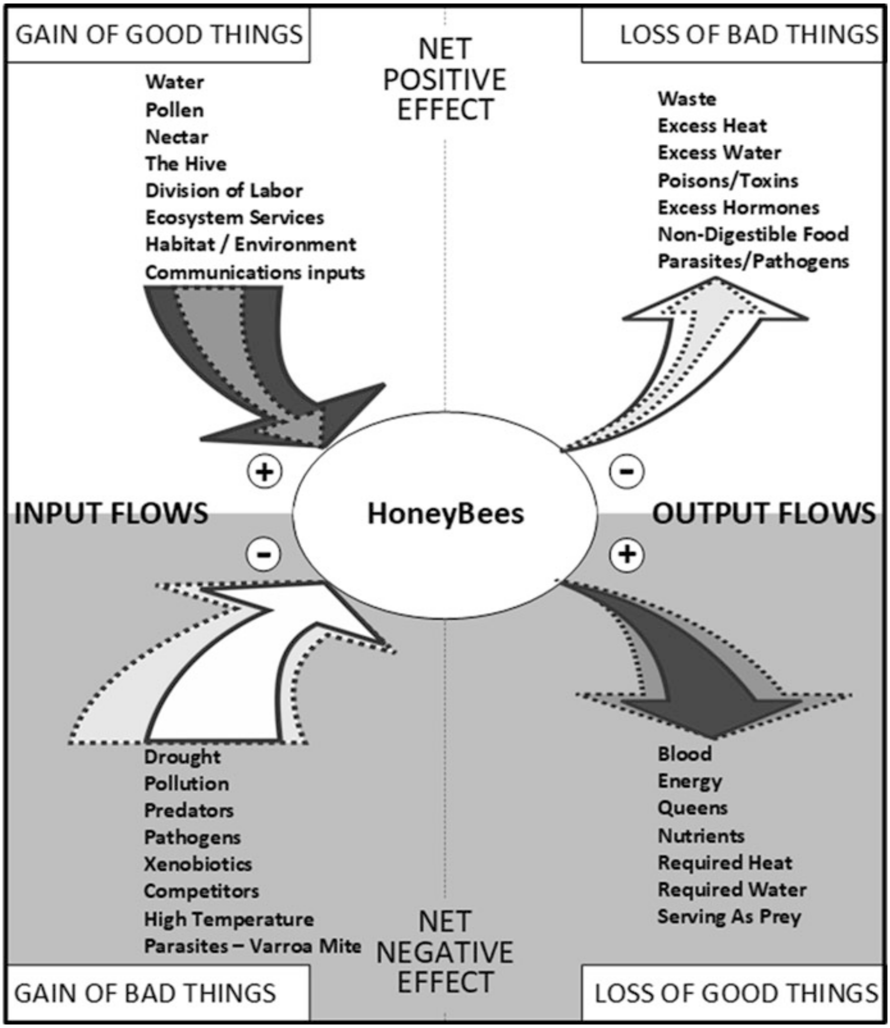


Fig. 4 The ecocircuitry of honeybees in colony collapse disorder. Solid-lined arrows show the normal, balanced flow of inputs and outputs to the honeybee, *Apis mellifera*, while dotted lines show the impact of *Homo sapiens*. Like all species, honeybees try to optimize their own flows, but the status of their inputs and outputs shows that humans have had impacts on most of these flows that are the opposite of that desired for optimization: decreased positive inputs like habitat, nectar and pollen plant species, coupled to increasing negative inputs like xenobiotic pesticides and noise stressors. On the output side, the rejection/loss of negative things like the Varroa mite is decreased due to inhibited immune system function (Evans et al. 2006), while the loss of positive things – nutrients stored in their honey, communication signals interfered with by EMF – creates a net negative effect

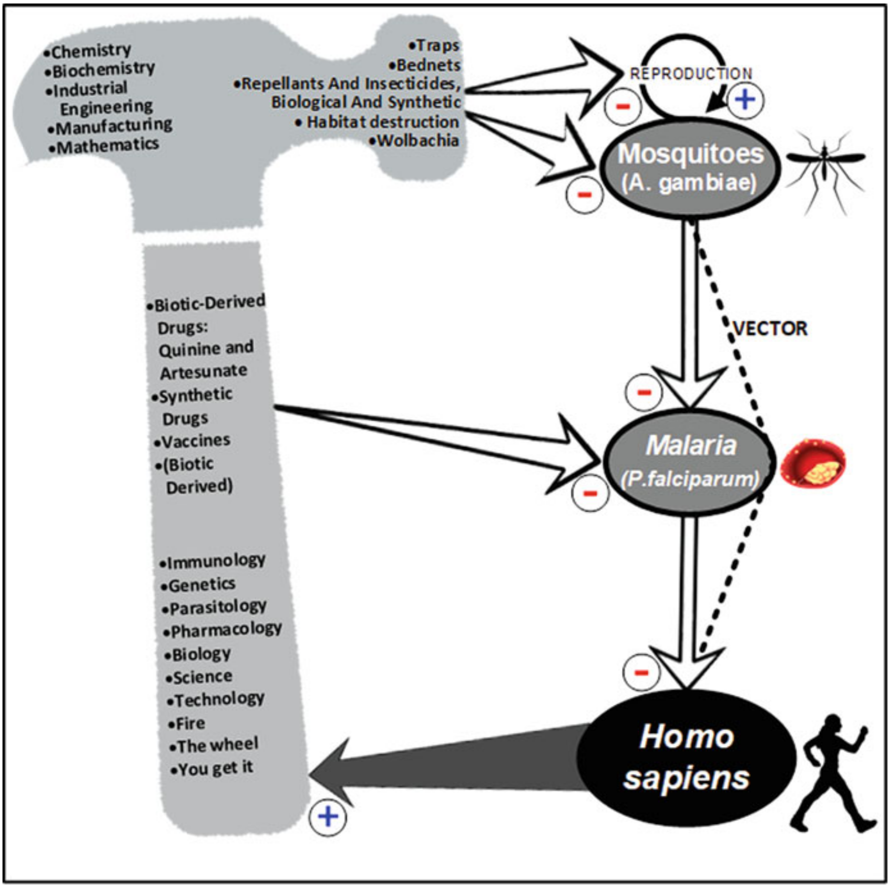


Fig. 5 *Homo sapiens* uses foundational meta-strategies to attack malaria and its vector. Illustration of how humanity has built specific strategies for our defense against known enemies as well as invisible pathogens up from foundational meta-strategies like chemistry and manufacturing. We coordinate ever higher and higher-order meta-strategies in an attempt to completely destroy these enemies

- Introduced distance – spears, catapults, guns, missiles, drones – up to and including physical removal from the site of danger.
- Increased weapon power/effectiveness/leverage – poisons, explosives, projectile rifling.
- Mobility – horses, ships, combat vehicles, airplanes.
- Infrastructure – fortifications, roads, armies/navies, weapons factories, supply chains.
- "Defense" to the point of extinction of the threatening species. While humans have not yet completely extinguished mosquitoes or malaria, Fig. 5 shows many of the ways that we are trying to do so – and the degree to which we will use

huge swathes of our ecocircuitry to try effectively fashioning global-scale threats to other species. Humans have made a great number of species, from wolves to smallpox, go extinct. Other species do not preemptively remove all members of a species it fears – but the efficiency of this defensive strategy has made this not just possible but standard for humanity.

8 Conclusions

Our planet is at an inflection point – a crossroad of direction, potential, and choice. Such transitions are vulnerable: once the giant boulder is rocking, it may be persuaded along many different paths. Ecocircuitry models may provide that direction via a system analysis that incorporates human behavior and institutions into ecological networks encompassing all aspects of the physical world.

Given that humans, their institutions, and their created entities are now essentially the primary "selective agent" for virtually every species on planet Earth, human connectivity and impact on our global ecological networks can no longer be ignored. Indeed, *Homo sapiens* must be explicitly incorporated, for any solutions based on modeling without them are addressing a system that does not currently exist.

Our aim is to actualize the "web of life" – Darwin's "entangled bank" – as a useful framework, analytic tool, and dynamical model. A description simple enough to be easily comprehensible, yet sufficient to capture the real complexity of the structure/dynamics of human-connected ecologies, from local ecosystems to the planet. A tool that provides the ability to identify patterns, dynamics, and blind spots, and allows researchers to conceive and visualize new *kinds* of solutions.

The massive amount of data emerging on ecosystems and their component entities, flows and dynamics, coupled to new "big data" mining techniques, highlights two specific needs. First, the necessity to organize this data into coherent structures that allow human understanding of their complexity; and second, to create generalizable models that can integrate relevant but currently siloed knowledge across all of our planet and species, including *Homo sapiens*.

Our model posits an overarching extension of ecological models to (1) include multivalent sets of dynamical flows (energy, materials, services and information); and (2) delineate the interactions between species/ecosystems, abiota (all nonliving things, from weather to water) and the human sphere, including its entities, systems, and impacts, from cars to constructs.

Taken together, we define these things as Earth's ecocircuitry. Optimization imperatives emerge when viewed through this lens: species have generated a vast host of complex evolutionary strategies to manage, manipulate, safeguard, and signal their I/O flows to maximize the beneficial, minimize the detrimental, and stabilize what exists. With all metrics (whether Darwinian, money, status, etc.), these optimization imperatives point to classes of evolutionary strategy, and further to network-manipulating meta-strategies whose principles can be consciously applied to our planet's problems.

The ecocircuitry model itself *is* such a meta-strategy – a cognitive tool. Humanity and presumably successfully sustained life is supported by huge numbers of meta-strategies of all kinds. With the development of certain "portal" meta-strategies such as language, teaching, and other information sharing, one species, *Homo sapiens*, has illustrated their iterative potential.

Yet a crucial point is that not only have humans used these evolutionary meta-strategies to become who we are, to develop our civilizations and our understanding of reality, but that *we do not recognize them as such*. Critical tools we take for granted – reading and writing, mass printing, libraries and universities, even education, science, and complex systems science itself – are all functional meta-strategies. Yet we do not recognize or understand these strategies as agents of active global change, even as they reshape the world around us.

Thus we can no longer silo concepts and useful models within specialized disciplines, for we run the risk of science – and humanity – using a pixel-by-pixel process on problems requiring *big picture* breadth and depth. Let us gain parallax through multidisciplinary perspectives to generate more accurate, insightful models for solving real-world complex problems, rather than low-dimensional models that will not map to reality sufficiently to be either correct or useful.

We thus urge complexity scientists to go beyond their primary training in nonlinear dynamics, physics, or mathematics, to study natural (neural, biochemical, ecological, physiological) complex systems. The complexity of the latter outpaces simplistic nonlinear-dynamical approaches and places those systems in the wrong part of the state space to find real-world solutions. Instead, complex systems science can and should extract core principles from natural complex systems, and use them to build accurate, integrative models more suitable for both understanding and decision-making.

More specifically, complexity scientists should be challenged to take on cross-cutting problems in collaboration with policy experts, thinkers, and the public, for our current political approach appears incapable of solving the very real and devastating threats which face us. We can – and should – train specialists to be generalists and remind complexity scientists that true complexity does not collapse to $F = ma$, nor is the best system description the most mathematically tractable one. We can – and should – learn to extract principles from specific real systems and apply them generally. And perhaps most importantly, we could – and should – reward big picture, long-plan, interdisciplinary problem solving rather than persist in the short-term approach of "publish-or-perish."

Perhaps most importantly, let us cease using poorly parsed, low-dimensional models too simplistic to apply to the hard problems, and embrace the high-dimensional interdisciplinary modeling that may be less easily done, but is most urgently needed.

When humans accept their effects on – and responsibility to – our global network, we can choose to use evolutionary meta-strategies such as a nuanced description of Earth's ecocircuitry, to understand and illuminate *Homo sapiens*' roles in the network of Earth's global ecology. We can then recognize our place and "ecological function" in the planetary system and the impacts our population, behavior, and institutions have within and upon it. Only then can we be empowered to rebuild our vulnerable and damaged networks into more resilient sustainable ones.

It is up to us to solve the problems of our planet with all possible useful strategies, as conscious beings capable of both analyzing the complex networks we exist within, and of taking action to mitigate destructive practices. Responsibility for the future of Earth lies firmly in the hands of *Homo sapiens* – and especially in the hands of those scientists who have the capacity to understand, model, and illustrate both problems *and* solutions. It is time for a conformation change in science that supports our informing the lay public in a clear and unbiased way of the perils and possibilities before us, and together determining how humanity, its tools and its capacities can become the network solution. Complexity science wedded to principles of biology/ecology, economics, evolution, and behavior may offer one of the tools we need to change our present negative trajectory.

Acknowledgments The author would like to thank E. Todd Hochman for his interdisciplinary brainstorming and assistance in graphics design; Professor James H. Brown for his in-depth collaboration on the initial model; Professor Astrid Kodric-Brown for her valuable collaboration on information flows and honeybee systems; John Smart, for his beneficial comments; the Hitchings-Elion Fellowship that funded some of the initial neuroscience thinking that led to this model; and the Department of Biology, UNM, ABQ for their essential support.

References

Andrewartha, H.G. and L.C. Birch. (1984) *The Ecological Web*. University of Chicago Press, Chicago.

Baccarelli & V. Bollati. (2009) Epigenetics and environmental chemicals. Curr Opin Pediatr. 21: 243–251.

Bratsberg, B. & O. Rogeberg. (2018) Flynn effect and its reversal are both environmentally caused. PNAS 201718793; published ahead of print June 11, 2018. https://doi.org/10.1073/pnas.1718793115

Brown, P. and K. Caldeira. (2017) Greater future global warming inferred from Earth's recent energy budget. *Nature* 552:45–50. doi:https://doi.org/10.1038/nature2467

Ceballos, G., P. Ehrlich and R. Dirzo. (2017) Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *PNAS* 114:E6089–E6096.

Cohen, J. E. (1978) Food webs and niche space. *Monographs in Population Biology*. Princeton University Press, Princeton.

Cribb, J. (2017) *Surviving the 21st Century: Humanity's Ten Great Challenges and How We Can Overcome Them.* NY: Springer International Publishing.

Darwin, C. (1859) *The Origin of Species; And, the Descent of Man*. Modern Library.

Dunne, J.A., Williams, R.J. and Martinez, N.D. (2002) Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology letters*, 5:.558–567.

Evans, J.D., Aronstein, K., Chen, Y.P., Hetru, C., Imler, J.L., Jiang, H., Kanost, M., Thompson, G.J., Zou, Z. and Hultmark, D. (2006) Immune pathways and defense mechanisms in honey bees *Apis mellifera*. *Insect molecular biology* 15:645–656.

Gould, J. L., & Gould, C. G. (1988) The honey bee. Scientific American Library.

Graham, L. (2008) Reparations, Self-Determination, and the Seventh Generation. *21 Harv. Hum. Rts. J. 47*.

Gremillion, MAV. and J. Brown. (2001) An ecosystem network model for human ecological interactions. *Ecological Society of America Abstracts, 86th Meeting*: ID=28381.

Hodgkin, A. and Huxley, A. (1952) A quantitative description of membrane current and its Application to conduction and excitation in nerve. *J Physiol*. 117: 500–544.

Ives, A.R., S.R. Carpenter, and B. Dennis. (1999) Community interaction webs and zooplankton responses to planktivory manipulations. *Ecology* 80:1405–1421.

Jordano, P. (1987) Patterns of mutualistic interactions in pollination and seed dispersal: connectance, dependence asymmetries, and coevolution. The American Naturalist, 129: 657–677.

Kaiser-Bunbury, C., S. Muff, J. Memmot, C. Muller, A. Caflisch. (2010) The robustness of pollination networks to the loss of species and interactions: a quantitative approach incorporating pollinator behavior. *Ecology Letters* 13: 442–452

Levins, R. (1975) Quantitave analysis – loop analysis. In M.L. Cody ad J.L. Diamond, eds. Ecology and Evolution of Communities. Cambridge, Mass. Belknap Press.

McRae, B., B. Dickson, T. Keitt, V. Shah. (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology* 89:2712–2724. https://doi.org/10.1890/07-1861.1

Mora, C. et al. (2017) Global risk of deadly heat. *Nature Climate Change* 7:501–506.

Moss, M. (2016) How the Colt Single Action Army Revolver Won the West. Popular Mechanics retrieved 5-31-18. https://www.popularmechanics.com/military/weapons/a23685/colt-single-action/

Mueller, U., S. Rehner, T. Schultz. (1998) Evolution of Agriculture in Ants. *Science* 281:2034.

Nabhan, G. P., & Buchmann, S. L. (1997) Services provided by pollinators. Nature's Services: societal dependence on natural ecosystems, 133–150.

Odum, H. (1983) *Systems Ecology: An Introduction*, Wiley-Interscience.

Paine, R.T. (1966) Food web complexity and species diversity. American Naturalist 100:65–75.

Paine, R.T. (1980) Food webs: linkage, interaction strength and community infrastructure. *Journal of Animal Ecology* 49:667–685.

Proulx, S. D. Promislow, and P. Phillips. (2005) Network thinking in ecology and evolution. *TREE*, v20:345–353. https://doi.org/10.1016/j.tree.2005.04.004

Schmid-Hempel, P. (1998) Parasites in social insects. Princeton University Press.

Seeley, T. D. (1997) Honey bee colonies are group-level adaptive units. The American Naturalist, 150(S1), s22–S41.

Raftery, A., A. Zimmer, D. Frierson, R. Startz, & P. Liu. (2017) Less than 2 °C warming by 2100 unlikely. *Nature Climate Change* 7:637–641.

Shachak, M. and C.G. Jones. (1995) Ecological flow chains and ecological systems: concepts for linking species and ecosystem perspectives. In *Linking Species and Ecosystems*, C.G. Jones and J.H. Lawton, eds. Chapman and Hall, New York, New York.

Shipley, B. (1997) Exploratory path analysis with applications in ecology and evolution. *American Naturalist* 149:1113–1138.

Steinhauer, N. K. Kulhanek, K. Antunez, H. Human, P. Chantawannakul, M-P Chauzat, D. VanEngelsdorp. (2018) Drivers of colony losses. Curr Op in Insect Sci 26:142–148.

Stone, G., B. Gyawali, J. Sandifer. (2017). Honeybee Colony Collapse Disorder in the USA. http://digitalcommons.murraystate.edu/postersatthecapitol/2018/KSU/6/

Ulanowicz, R.E. (1997) *Ecology: The Ascendant Perspective.* Columbia, New York, New York.

Van Valen, L. (1973) A new evolutionary law. *Evolutionary Theory* 1:1–30.

Williams, R.J. and N.D. Martinez. (2000) Simple rules yield complex food webs. *Nature* 404:180–183.

Wilson, E. O. (1971) *The insect societies*. Cambridge, MA: Harvard University Press.

Wilson, E. O. (1975) Slavery in ants. Scientific American, 232:32–40.

Complex Dynamics in Small Groups

Holly Arrow and Kelly Bouas Henry

1 Introduction

A largely unsolved problem in the field of group dynamics is how and why groups whose members have apparently equivalent skills and potential develop so differently, with some achieving surprisingly good results while others dramatically underperforming their apparent potential. In this chapter, we consider this puzzle by taking a dynamic systems perspective on emergent structures and processes in groups. Our analysis draws on the metaphor of group "thermodynamics" (Henry and Arrow 2010) as a guide to understand how structure and energy in groups reciprocally influence each other as groups change over time.

We propose that small groups in which complex structure facilitates flexible behavioral coordination tend to perform better and adapt more effectively to evolving demands from the environment. Following Levine and Moreland (1995), we define "small groups" as groups with three members up to size 25–30, although the upper boundary is fuzzy. A group is not just a collection of people, however. From a complex systems perspective, a group is "a self-organizing, open system of dynamic connections among interacting members, embedded within and interacting with its context" (Arrow 2005, p. 203).

In the group dynamics literature, work on *group development* addresses the question of how groups change systematically over time. Some models of group development emphasize internally driven change; others focus more on externally triggered change (Arrow et al. 2005). We propose an integrated account that attends

---

H. Arrow (✉)
Psychology Department, University of Oregon, Eugene, OR, USA
e-mail: harrow@uoregon.edu

K. B. Henry
Missouri Western State University, Saint Joseph, MO, USA

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_18

to the interplay between these forces and suggest that differences in developmental trajectories help explain why some groups become more resilient and effective than others. To motivate our examination of the interplay between structure and dynamics, between endogenous development and externally responsive adaptation, we use examples of groups from three different contexts: the classroom, team sports, and military aviation. After each example, we unpack some of our ideas about structure, dynamics, and development, using the concepts of heat and energy transfer as underlying drivers of performance. We believe a better understanding of how emergent structures shape group processes (and vice versa) can help coaches, leaders, and group members guide their groups to dynamic states associated with higher levels of performance.

2 A Class Becomes A Group: Building Structure by Managing Energy Flow

Effective teaching at any level involves creating a climate that engages students. Demonstrating enthusiasm for a topic (Marsh 1987; Marsh and Bailey 1993; Zhang 2014) can spark shared interest that may help warm up a class full of unmotivated, disengaged students. However, enthusiasm is not always contagious. From a structural perspective, students may feel insufficiently connected to the instructor to start caring about a topic just because the person at the front of the classroom does. Instead, many may just sit there, sneaking peaks at their phones, daydreaming, and occasionally jotting down some notes. A collection of students is not a learning group until the structures that support effective information sharing and meaning making emerge.

One strategy for educators to break the ice is the “pair-share” strategy (O’Connor 2013). Rather than punctuating lecture segments with open-ended pauses to solicit “any questions?”, asking students to talk to a neighbor about a topic creates new learning relevant links among the students. After neighbors discuss the topic, the instructor can either have students switch to a new conversation partner or start scaling up the discussion to small groups by having pairs share with other pairs. Physically circulating people around the room to mix and match pairs and build small groups can contribute to the warming trend. As more people talk to each other, the level of engagement rises and the classroom feels more active. That first interpersonal contact is the building block that breaks the ice and contributes to the warmer classroom climate that results. This strategy initiates the “forming” stage of development by seeding a network of links connecting students to one another and facilitating information and energy flow. An unresponsive set of silent students absorbs energy; an engaged, interacting set of students generates energy.

That increased energy may not be directed at learning, however. As the ice breaks and the classroom warms up, discussions can quickly wander off track. Rather than productively channeling shared energy toward the chosen topic, the “everyone talk” dynamic can veer into chaos, with some students on topic, others off, others arguing or flirting, and the teacher simply standing by. The room may warm up, and the students engage, but not much is getting done. Now the instructor needs to redirect the energy. While it may seem natural to regain control of the class by raising one's voice, an experienced teacher knows that yelling is generally ineffective. It only adds energy to a situation with too much energy and not enough structure to channel it. Instead, the energy needs to be dialed back down to a productive level with structures that contain it.

One strategy is to deactivate the dyadic links with new instructions. If circulating students to form temporary and shifting dyads warms the class up, changing the physical dynamic to put students back in fixed seats will cool it back down. Rather than shout to get the students' attention, writing new instructions on the board – for example, "Return to your seat and write one thought that came up in discussion" – will gradually attract attention, and as students comply, the whole class will calm down. The new instruction helps the students bring the energy of discussion back to their own sense of connection to the topic – a link that has ideally been strengthened by discussing the topic. The final stage may be to go around the room and solicit some turn-taking sharing with the full classroom. Instead of isolated responses to "any questions?" this becomes, instead, a selective debrief of a fully engaged class discussion.

3 Structure: An Emerging Network of Links and Attractors

Although group structure is one of the most pervasive aspects of a group (e.g., Bales 1950; Stein and Heller 1979), it has also been identified as one of the most neglected topics of study (Moreland et al. 1994). A common definition of structure is a pattern of relationships among members (Forsyth 1990; Levine and Moreland 1990; Wilke 1996). Structure includes cohesion, which originally referred to the collective attachment of members to the group as a whole (Festinger 1950); roles, which help differentiate members (Cummings 2004); and norms, which generate predictable patterns of interaction that integrate the group (Cummings and Ancona 2005). These traditional elements can be integrated into a network conception of structure (Arrow et al. 2000, pp. 50–53) with interaction patterns understood via the dynamic systems concept of attractors.

Because people are the most obvious elements of a group, it's natural to think of groups as comprised of individuals. However, a collection of people (e.g., a set of people who all show up at the same place at the same time) is not a functioning group. It is the connections among the people, the "pattern of relationships" that turns these individuals into group members. Over time, a class full of students whose primary activity is listening to the teacher and occasionally asking clarification questions will become more group-like. Routines and norms established by the teacher will guide behavior. However, the group structure is quite simple, a set of asymmetrical teacher-student links with mostly one-way communication. Links among students (who may, e.g., sit next to friends and occasionally consult together when they are confused) will be mostly imported from outside connections rather than emerging as a core group structure.

The concept of “attractors” – characteristic patterns of behavior in which a dynamic system tends to settle – is a way of understanding both the emergence of routines and the changes in behavioral states that the teacher helps orchestrate.

Routines are easy-to-execute patterns of action that become attractors for the group, because following established routines takes less effort than doing something new. At the beginning of the classroom example, the “chilly” classroom is in a *fixed state attractor* characterized by negative feedback cycles that drain energy from teacher and students alike and hamper productivity (Henry and Arrow 2010). The pattern will quickly become routine unless something actively disrupts the dynamic and shifts the group into another behavioral regime. Directing students to “pair and share” shifts them into the well-practiced back and forth of dyadic conversation instead. This is an easy-to-access *periodic attractor* that allows for higher energy flow as each conversation partner shares thoughts and then waits to hear what the other person has to say. Instead of the teacher being the only source of energy in the room, the students’ conversations tap energy from everyone and fuel the discussion. In this early phase of group formation, however, it’s easy for the energy to leak out of the group container and wander off into outside conversations or dyadic exchanges that are no longer group-relevant. The resulting mildly *chaotic state* is characterized by divergent energy flows and minimal integration (Henry and Arrow 2010). The developmental trajectory of a group can, in part, be tracked by looking at the emergence of different attractors, which will be strengthened if a group visits the attractor frequently and stays in that state longer.

The strategy described in the classroom case above, in which students are explicitly guided to connect to one another on class topics and circulate to create new ties, builds a more complex structure by proliferating links and evoking a new dynamic state – a periodic attractor of back-and-forth conversation. Ashby’s Law of Requisite Variety (Ashby 1956) states that to achieve consistent outcomes, the variety in a system must match or exceed the complexity and variety of its inputs. An impoverished classroom system has quite limited ability to process inputs, and much of the material presented by the teacher may never take hold. By helping the group form more fully, the teacher enables the group to handle more complex topics and engage students more fully in processing whatever material the teacher provides. As the group develops over time, students may bring more of their own material into the classroom as well, further enriching the available inputs into the learning group.

This boosts the social energy of the group, defined by Canavan (1996, p. 3; cited in Betancur 2005) as “the generalized motivational, engaged state generated by doing something one likes with others whom one likes and who like the same activity.” As individuals share information and enthusiasm over a common interest, they are energized. Fans who compare the experience of watching a game alone or in a stadium with like-minded groups of friends easily recognize the difference. The greater complexity of multiple links (person to person to shared focus of enthusiasm) supports stronger energy flows than a single link (one fan focused on the game). In groups, the coordination network of links provides the scaffold that supports social energy flowing into group work, which develops and tunes the role structure linking members to tasks. When groups have too little energy, structural changes can help ramp energy levels up and also keep it from scattering beyond the group boundaries. As energy increases, new structures need to form or new patterns of activation emerge to accommodate the changing flow. Four "motors" of change help drive this dynamic interplay.

4 Systematic Change: The Four Motors and Progressive Development

Poole et al. (2000) describe four possible motors for generating change in groups. The *life cycle* "motor" drives the sequenced unfolding of development across different stages, analogous to the development of individuals. As new capacities come on line, they create a scaffold for the next focus of group activity. The *teleological* motor drives change as groups work intentionally toward one or more goals. In this case, the demands of the task and the context shift the group from one pattern of activity to another. The *dialectical* motor drives change via tension between opposing entities. This may be tension between a group's leader and followers, among subgroups within the group, conflict with outsiders, or friction between the group agenda and individual member goals. When tension gets resolved via synthesis, this sets the stage for a new conflict to emerge. The *evolutionary* motor drives change through repeated cycles of variation, selection, and retention. Consistent with the operation of evolution at the species level, this motor has been applied mostly to populations (i.e., Hannan and Freeman 1984, who focus on organizations), rather than particular groups.

The most well-known models of group development are progressive stage models (i.e., Tuckman and Jensen 1977; Wheelan 1994) that identify a set number of stages that groups move through to reach productive maturity. These models rely on the endogenous life cycle motor to explain change, but also draw on the teleological motor with maturity as the goal of development. The stages can be understood as a series of attractors that capture a group for a while, stabilizing it in whatever activity dominates for that stage, until new attractors develop and the group moves on. Within each stage, the network of links connecting members and tasks is further elaborated and tuned, and the characteristic patterns of activation as energy flows among these links shift accordingly. The classroom example describes a group in the first stage, *forming* (Tuckman and Jensen 1977) or "dependency and inclusion" (Wheelan 1994). As relationships build and interaction increases across more varied pairings or larger sets of members, the network of links is fleshed out. Social energy increases as group members begin to identify with the group and coordinate their attention to focus on shared topics and tasks.

Once members feel more certain about their inclusion, they start testing the emergent structure by dealing with conflicts around power, authority, and competition: the *storming* stage (Tuckman and Jensen 1977), which is driven by the dialectical motor. As the group resolves the tension, it moves into the *norming* stage, in which patterns of influence among group members are often reconfigured, and agreement emerges on adjusted goals, norms, and role expectations. The energy released by friction becomes less important and social energy increases again as group members coordinate their efforts and attention on achieving shared goals and the teleological motor takes over again as the primary driver. Ideally, the result is the high task productivity of the *performing* stage. According to the model, groups remain in this phase until they terminate, the *adjourning* phase.

Multiple studies support this sequence and content of change in groups (Wheelan et al. 1994; Wheelan and Kaeser 1997), although groups can get stuck and fail to progress, either failing to form (a common occurrence in larger lecture classes) or remaining mired in conflict without achieving the integrative structures that allow them to achieve high productivity. Inadequate elaboration of group structure (e.g., too much homogeneity; insufficient role differentiation or failure to enact roles) can prevent a group from moving on to the next stage (Bennis and Shepard 1956; Kuypers et al. 1986). Using the thermodynamic metaphor, the classroom had to be "warmed up" using energy transfer mechanisms of *contact* (e.g., pair-share interactions) and *circulation* (mixing the dyads and moving them around the room) that also helped elaborate the structure of the group. Our classroom example showed the initial phase of forming. In the next section, we briefly summarize the full lifespan of a sports team whose ultimate performance was legendary.

5 Forging a Miracle on Ice

In the 1980 Winter Olympics, the US Hockey team, comprised entirely of amateur players with an average age of 21, went up against a heavily favored Soviet team of mostly professional players. The US team entered the Olympics seeded seventh, and in an exhibition game 3 days before they would face Sweden for their first Olympic match, they had lost to the Soviet team 3–10. The Soviets had won 5 of the 6 previous gold medals in Olympic Hockey, and most teams considered them unbeatable. After outscoring second seed Czechoslovakia and qualifying for the finals, the coach of the third-seeded Sweden team was asked about his team's chances for gold. Ohlssson replied "We are headed for the silver . . . The gold is out of the question. The Russians are the best . . ." (Gilbert 2008, p. 437).

Like Swedish coach Ohlsson, Team USA Coach Herb Brooks had no illusions about the strength of the team his players would face. But he did not envision defeat. Instead, he infused his team with confidence with an extraordinary pregame speech (recreated based on the players' accounts) before they took to the ice: "If we played them ten times," he told his players, "they might win nine. But not this game, not tonight. . . .Tonight, we are the greatest hockey team in the world."

(Ciardi et al. 2004; Littlefield 2015; Miracle on ice n.d.). He was right. They won the game 4–3 in one of the most extraordinary upsets in hockey.

Brooks built the team capable of this feat in only 7 months. It had taken him 7 years to build an outstanding team at the University of Minnesota. The compressed schedule required him to shepherd his team to the mature state of performing (Tuckman and Jensen 1977), also known as "work and productivity" (Wheelan 1994), quickly. The challenges of the initial forming stage were clear. Among his recruits were a core of "Gophers" from the Minnesota team who were familiar with him and his system. The other players, however, were not new recruits joining the Gophers, but players from teams in Massachusetts and elsewhere with whom the Gophers had a fierce rivalry. These pre-existing rivalries created a huge source of friction and negative energy that would make it difficult to form the players into a single team. Brooks accomplished this not by trying to suppress or block the energy, but by redirecting it to a new target: himself.

As he told Craig Patrick, whom he had just recruited as Assistant Coach, "I'm going to be a real prick, and you will have to keep the guys together. The guys from the East and West hate each other, and the only way to keep them from hating each other is to make them all hate me" (Gilbert 2008, p. 301). This strategy could only work with players able to take the heat that Brooks would train on them over and over in a tempering process designed to forge the hostile factions into a unified whole. As former Gopher Mike Polich described, "When it came to the Olympic team, he was going to select guys who could take it and respond. If you throw gasoline on a flower, it might wilt. Herbie didn't want that. He wanted to know that when he threw gasoline, there'd be an explosion" (Gilbert 2008, p. 296).

To learn to play the fluid, creative version of hockey that he envisioned could beat the best European teams, each player would have to come to absolutely trust one another, a hallmark of the norming stage (Tuckman and Jensen 1977), also known as "trust and structure" (Wheelan 1994). Instead of stamping out the diversity among players with different backgrounds, however, Brooks saw it as a resource. As he told journalist Gilbert when describing his strategy for recruiting: "We can't all have the same styles. We've got to build a house out of all different kinds of bricks" (Gilbert 2008, p. 307).

Although he was famous for the intricacy of the systems that he devised for the players, the structures Brooks built were "aimed at leveraging a disciplined sense of defensive play with unfettered creativity." This is what would enable the team to excel in the pressure of Olympic competition. "Brooks called it 'sophisticated pond hockey,' because he knew that the freedom to experiment . . . is a staple of unstructured pickup games on outdoor ice. . . . [he] selected players for their perceived hockey sense in addition to their skill, and then tried to 'unstructure' them into a pond-hockey mentality, where they might function well in the face of extreme pressure, or make a move they had never attempted before" (p. 343).

The success of the coach's ability to both create effective structures and evoke and channel the energy of the players into the team was key to beating the objectively superior Soviet team, which was unable to match their energy and creativity. As Swiss reporter Andreas Wyden commented in his analysis of the game: “When the U.S. was behind, they had to invent something to get ahead, and they did . . . Now that the Russians are in trouble, they are poised. But they don’t need poise, they need emotion” (461). By the time the players of Team USA reached this critical game, their explosive energy was fused and directed toward a singular purpose. As player Strobel said later, Coach Brooks “knew you can’t coach spontaneity, and that emotion has to come from within” (450). The 7 months of tempering had built a strong and flexible thermodynamic engine, and Brooks’ inspirational pregame speech released the emotion and creativity that carried the team to their win.

6 Strong Synergy Emerges from Repeated Cycles of Tempering

The elusive holy grail of group synergy (Larson 2007, 2010), also known as the assembly bonus effect (Collins and Guetzkow 1964; Shaw and Ashton 1976), has fired the imagination of many group researchers interested in productivity. Synergy is when group productivity exceeds what would be expected based on the abilities of individual members, and Team USA clearly achieved it. Recent work by Larson (2010) has distinguished between “weak” and “strong” synergy. The former is when group performance exceeds what would be expected based on the performance of a typical group member; the latter is when the group exceeds what would be expected based on the individual performance of the group’s best member (pp. 6–7). As Larson points out (p. 15), Steiner’s (1966, 1972) influential definition of group productivity as “potential productivity minus process loss” has been a substantial conceptual and methodological stumbling block in studying synergy, as it has directed researchers to study productivity losses rather than productivity gains. Most research comparing individual and group performance has also ignored the impact of group development by studying newly formed groups. Larson identifies synergy as “an emergent phenomenon rooted in group interaction” (p. 4) but notes that research on synergy has almost completely ignored the time and developmental processes required for synergy to emerge (pp. 365–367).

Ample research using progressive stage models has demonstrated that groups take time to achieve full productivity, with the time to reach the fourth “performing” stage ranging between 4 and 7 months (Wheelan et al. 2003). Many groups never progress to that stage. In a study of 329 work groups that had been together with relatively stable membership for at least 6 months (and some for many years), Wheelan (2009) found that less than a third had made it to Stage 4, and larger groups of more than six members were developing more slowly than groups of 3–6. About half of the larger groups had not progressed beyond Stage 2 (storming). With a final roster of 20 players, a deeply rooted source of intragroup conflict, and 7 months to reach world-class level performance, Team USA was facing a particularly stiff developmental challenge. To understand how Brooks guided the team to power through this challenge, we turn to an alternate group development model, *punctuated equilibrium*.

According to the logic of the life stage motor that drives progressive stage models, it's simply not feasible for groups to skip the first several stages and perform effectively in short order. Yet some groups violate this logic. The punctuated equilibrium model emerged as an alternative to stage models based on a detailed examination of project groups whose life span ranged from 7 days to 6 months. In these groups, the evolution of group structure and the quality of performance shifted between periods of stability maintained by strong inertial forces and sudden, rapid change, similar to the punctuated equilibrium pattern in biological evolutionary theory (e.g., Eldredge and Gould 1972). It provides a different window into how Team USA achieved an optimal balance of stability and flexibility.

Punctuated equilibrium theories initially focused on midpoint transitions in project groups (Gersick 1988), with later iterations of the theory focusing on transitions in general (Gersick 1991; Gersick and Hackman 1990). In the initial studies, Gersick (1988, 1989) documented that task groups jumped right into task-focused work with whatever structure formed in the first set of interactions. This structure persisted until the midpoint of their time together, when heightened concern about the approaching deadline triggered a flurry of activity that restructured the group and jumpstarted a new phase of heightened productivity. Elaborations of the theory for ongoing groups (Gersick and Hackman 1990) proposed other triggers for change, including external interventions, membership change, or other novel developments that either signal a mismatch between the group's current operation and the environmental demands or renders the current structure inoperable in other ways. From a dynamic systems point of view, stability corresponds to a strong attractor, which Gersick (1991) calls "deep structure." When a control parameter that governs which attractors are available to the group shifts, that attractor vanishes, leaving the group scrambling to find a new way of working together. Heightened energy and activity contribute to a more chaotic dynamic that aids in rapid restructuring by scrambling the existing coordination network. This is similar to sudden turbulence in a river when a steep drop in altitude speeds up the flow of water, creating high-energy rapids.

For a true Team USA to emerge from the players he had recruited, Coach Brooks needed to rescramble pre-existing structures of group identification that bonded members to their college teams and regions. He needed to transform the dialectical energy of intergroup hostility into the social energy of shared commitment to Team USA and create a container capable of withstanding both centrifugal forces and the external pressures of competition. He also needed to break down standard North American approaches to hockey into a more fluid mode of play that would allow the team to compete successfully on the world stage. This would require a more complex structure supporting the capability of the team to enter the *complex dynamic state*, a highly creative and variable mode of operation characterized by positive feedback cycles that boost member energy and enable the group to adapt quickly to external demands (Henry and Arrow 2010).

Tempering and annealing are forms of *heat treatment* used to alter the characteristics of metal, and fit with our conception of the underlying group thermodynamics of change in the punctuated equilibrium pattern. Carley and Svoboda (1996) have used "simulated annealing" to explain how organizations, viewed as complex adaptive systems, adapt their design to their environment, focusing on individual learning and executive decisions about restructuring as the drivers. In applying this metaphor, we pay more attention to the thermodynamic element and note that heat treatments of physical materials and the fundamental restructurings characteristic of punctuated equilibrium result in nonreversible changes in structure that increase both toughness and plasticity. The inertia of deep structure resists change, so the group needs a jolt of increased energy to catapult it out of the comfortable gravity well of a familiar attractor.

Coach Brooks got the players riled up and hating him (distracting them from hating each other) and then subjected them to a punishing schedule of conditioning, travel, and competitive matchups, throwing metaphoric gasoline on the emotionally aroused players to create explosions. After the extreme heating, the links among the players and between the players and the emergent Team USA reformed into a strong, resilient, but flexible structure that could handle a lot of pressure without breaking. The performance of the restructured team exceeded what could plausibly be expected by reviewing either the average hockey ability of the individual players or even the talent of the best members. This is the strong synergy of nonadditive, nonlinear performance, when $2 + 2 = 5$... or 7... or 10. The result is nonadditive, because the group performance of a strongly interdependent team is based on the coordinated expenditure of energy that is regulated through interaction as it flows through a network structure of connections.

Coordinated activation of subsets of links in the network orchestrates group performance. Like conductors and resistors in electrical circuits, any link can facilitate or damp down energy flow. If resistance dominates, the energy members bring to the group is frustrated and diminished. If too many links are activated simultaneously, group activity can become chaotic and unfocused. In the complex dynamic state, the shared focus of group members on a motivating task keeps the energy flowing, while openness to the evolving context allows them to adjust their patterns of activity quickly, matching the complexity of inputs in a way that sustains high performance. We turn now to a case that examines performance in a setting that is as dynamic as a hockey game but far more complex and distributed in the demands it makes on the group.

7 Adaptive and Nonadaptive Response in an Aircraft Carrier Case

An aircraft carrier is an enormously complex system in which any error can lead to catastrophic loss of equipment and/or life. Rochlin et al. (1987) describe it as operating at the "edge of the envelope" (p. 76) under extreme circumstances in an unstable environment with enormous tension between needs for safety and reliability versus operational efficiency. Planes land and take off simultaneously, on a single narrow runway awash in seawater that is rocking from side to side (p. 78). When a plane lands or is "recovered," the pilot acts in close coordination with "a set of interrelated activities among air traffic controllers, landing signal officers, the control tower, navigators, deck hands, the helmsman driving the ship, etc." (Weick and Roberts, p. 363). Up to 25 people have the capacity to wave the aircraft off if there is a problem. In the incident described below, a 38-million-dollar aircraft that was waved off was lost at sea. Here's what happened (summary based on pp. 371–372, Weick and Roberts 1993).

The incident started in a night launch with several planes waiting on deck to be launched, while five other planes in the air were in line for landing (recovery). Problems among the planes in the air began to perturb the system, requiring rapid adjustment. The plane fourth in line to land had an apparent hydraulic failure that would prevent its wings from folding up after landing, so it would take up extra space. While that pilot sought help via radio, the plane third in line had a compound hydraulic failure in which no back-up systems seemed to work. The plane fifth in line then developed a control problem. The air boss needed to land these planes interspersed with launching the waiting aircraft on deck. The catapult for launch had a blast-deflector panel that extended part way into the area for landing. So the coordination needed to get the problem planes recovered was highly demanding.

Plane 3, the one with the worst problem, was moved up in the landing pattern, but then had to be waved off, because the person operating the deflector panel for launches (who didn't realize an emergency recovery was in progress) lowered the panel a second too late. Plane 4 landed, and a second pass for Plane 3 was delayed so that Plane 5 (whose problem had been resolved, unbeknownst to the air boss) could be recovered. In the meantime, Plane 3's increasingly stressed pilot reported further hydraulic failures after trying various fixes suggested by the squadron representative. Eventually, the pilot was redirected to a land base, but ran out of fuel. He and the radio intercept officer (RIO) had to eject. They were rescued. The plane was lost.

8 System Breakdown in a High-Performance Crew

To offer a window into this case, we need to first identify some critical features that differentiate this group from the classroom and the hockey team cases. First is the distinct nature of a *crew*, in which the primary set of links that organizes behavior is not a social network of relationships among a relatively stable group of people, but the job network that specifies the sorts of actions needed for the technology and tasks to fit together to coordinate specialized sorts of projects (Arrow et al. 2000, pp. 83–84). Unlike the project groups (Arrow et al. 2000 call them *task forces*) studied by Gersick, crews handle a series of projects that are similar in type (managing flight operations) but differ in details and demands from one project or mission to the next. The people who staff the system come and go as shifts rotate and new members are swapped in and out. To perform effectively, crew members need to adjust their actions based on external inputs that vary across jobs and projects.

Larson (2010) emphasizes how the definition of weak and strong synergy (and how difficult it is to achieve them) depends a lot on the nature of the task. Although the fast-paced adjustment of hockey players shows a fluid response to evolving events, flight operations on an aircraft are far more complex and distributed. Rochlin et al. (1987) note repeatedly that an air craft carrier's complexity of operations is so great that no one person could know the content or sequence of all that is needed to make sure aircraft can fly safely, reliably, and on schedule. That knowledge is distributed across persons and elaborated and adjusted in real time through the network of interaction that Weick and Roberts (1993) call heedful interrelating. So it becomes impossible to compare the performance of the system with what the "average" or "best" member can do. The breakdowns that contributed to the loss of the aircraft as the system became overcome by events (OBE, Weick and Roberts 1993, p. 373) were not individual failures but a breakdown of circuits that failed to regulate energy and the flow of information effectively. In this situation, the sources of energy are the arousal of high stress, potentially dangerous situations and the high motivation to prevent errors and ensure safe launch and recovery. Effective flow of information through the communication network helps regulate that energy.

Had the deflector operator realized that recovering the next plane in line was an emergency situation, he would likely have been more aroused, more attentive, and more successful in the timing of his actions. He didn't receive that information. Plane 5 was no longer in trouble when it landed. But the right people didn't receive that information in time. And neither the RIO in Plane 3 nor the squadron representative attempted to calm the frightened pilot, whose escalating level of arousal made it harder and harder for him to think through the technical problems. Somewhat like a teacher shouting in a chaotic classroom, the growing number of attempted solutions prompted by the squadron representative actually made things worse, triggering further hydraulic failures (Weick and Roberts 1993, p. 373). The plane's back-up systems, like the pilot's emotional state, got overloaded and plane and pilot both became less and less able to fly.

In the classroom, a brief foray into the chaotic state is unproblematic. In Team USA, chaotic release of explosive tensions helped erase prior structures and allow for a new, more flexible and effective structure to emerge. In the midst of a highly dynamic complex system reaching overload, any uncontrolled increase of energy can tip the system past the edge of chaos with catastrophic results. As the catastrophe began to unfold, interaction became more isolated, creating unhelpful local feedback loops, while connections to the wider system were neglected. The structure of the system began to pull apart. Communication circuits that failed to amplify and transmit critical information and regulatory circuits that failed to dampen escalating local inputs that were fouling both the plane and its pilot led to a partial disintegration of a system that demands rapid and effective adaptation to evolving events.

Although we have focused here on a system breakdown, what is most surprising about flight operations on aircraft carriers is that they regularly maintain such a remarkable level of effective performance under astonishingly demanding conditions. Weick and Roberts (1993, p. 357) quote Wilson's (1986, p. 1) characterization of an aircraft carrier as "a million accidents waiting to happen." Yet almost none of these accidents occur. What developmental processes build and maintain these extraordinary groups?

9 Group Development On the Fly

For crews, the progressive stage model of group development makes little sense, for the system is already in full productive operation as new members arrive and fit into their roles. Crews are a type of *concocted* group (Arrow et al. 2000, p. 65) that are designed and formed in advance, before they are staffed. The internal logic of unfolding is also a poor fit for the external orientation of crews. The punctuated equilibrium model, which we applied to the Miracle on Ice hockey team, is also a poor fit. In that model, periods of stability are anchored internally by deep structure and inertia, and during these periods, groups pay less attention to external cues. The high-energy restructuring periods of revolutionary change typically reorient the group to align more closely with external demands – in the hockey case, to prepare for successful competition against top European teams in the Olympics.

Given the high turnover often experienced on air craft carriers (Rochlin et al. 1987) and the regularly changing demands of different missions, stable routines are neither feasible nor effective. Neglect of external cues immediately compromises performance and can quickly lead to a crisis. Flight operations require a highly energetic group that have a shared overall understanding that is enacted and adapted as they respond to each other and changing conditions. Nothing can be routinized, and constant attention to each other, task, and context is necessary to support and adjust the flow of information and energy in real time. Rather than fixed routines, the crew needs to generate and adapt dynamic patterns of responding on the fly as conditions for a given day unfold. In this way, flight operations on an aircraft carrier are a bit like Brooks' fluid "pond hockey" writ large, but with new skaters arriving and departing in quick succession in a fog dense enough that no skater can see the whole field of play.

A third type of group development model, *adaptive response,* is the best fit for describing this fluidity. This pattern applies to groups that actively generate and adjust their coordination networks in response to both internal and external contingencies (Arrow 1997). Rather than the deep structure of overarching routines and fixed patterns, such groups draw on a repertoire of smaller-scale action patterns that link people, tools, and tasks for a particular local purpose. Global structure emerges from the enactment of these local action patterns, integrated by a collective understanding of the larger purpose into which the local activities fit.

The speed of change in groups that fit the adaptive response pattern will depend both on how attentive members are to changes in the group's operating context and how quickly that context changes. The Law of Requisite Variety mentioned earlier (Ashby 1956) states that to achieve consistent outcomes, the variety in a system must match or exceed the complexity and variety of its inputs. For groups in highly dynamic and demanding contexts, the quality of performance will depend on whether the group's potential speed of adjustment matches or exceeds the pace of unfolding events. Hence, such groups need to be especially skilled in tapping reservoirs of energy needed for sudden coordinated shifts, and regulating that flow so that it doesn't overwhelm the capacity of the group to channel it.

10 Lessons Learned

The cases and literature cited in this chapter are couched in language that often references thermodynamic principles. The classroom needs to be "warmed up" or "cooled down," depending on how the class interaction is going. The 1980 gold-medal US hockey team developed a flexible and resilient structure through a "tempering" heat treatment that built structures to channel the emotional energies of the team into fluid, creative play. The aircraft carrier crew's successes and failures demonstrate how complex dynamics of high energy can either be channeled effectively or tip into chaos that dissolves the integrating structures that tie the distributed actions together. No one developmental model fits all three. Instead we drew on a different type of model – progressive stage, punctuated equilibrium, and adaptive response – for each.

Thinking about the thermodynamics of energy flow in groups and how that helps shape and is shaped by the evolving group structure helps illuminate the underlying set of fundamentals that explains change in groups over time. The group development literature offers a catalog of options to identify various patterns of change. Along with the three used in our case analyses, other patterns of stability and change not mentioned earlier include robust equilibrium (e.g., Bales 1955; Gersick and Hackman 1990) and repeating cycle models (e.g., Worchel et al. 1992; Worchel 1994). The specification of mechanisms and drivers in these largely descriptive models shows little consistency across models, and is often underdeveloped.

Which pattern is most useful for understanding what's going on with a particular group? This is a challenging question for leaders and group members unfamiliar with the literature, which lacks integration and an overarching guide to systematically match group characteristics to change patterns. We have pointed to some differences between crews and project groups that may predispose them to adaptive response or punctuated equilibrium patterns. More fundamentally, however, what group members and leaders generally want to know is how to use their understanding to help their groups achieve synergy through positive change. Even for those well versed in group development theory, the patterns described in the literature do not offer much guidance about how to *change* the course of a given group. Rather, they simply describe the course on which a given group *might* be at a given moment in time. Applying these theories to promote useful change is problematic.

Anderson et al. (2001) describe a way of understanding this researcher-practitioner gap by identifying four types of science that vary along two dimensions: practical relevance and methodological rigor. They label science high in rigor but low in relevance *pedantic science*. Science high in relevance and low in rigor they call *populist*. *Pragmatic science* is high in both rigor and relevance, while *puerile science* is low in both. To the extent that extant theories of group development are descriptions grounded in rigorous scientific method, they lean more toward the pedantic quadrant of Anderson et al.'s framework. If leaders of real groups find the research on group development useful for retrospective analysis of a group's trajectory, but difficult to enact for change in the present, it highlights the need for a more pragmatic science of group dynamics.

We see the thermodynamic perspective on change in groups as a step in that direction. It provides an intuitive mechanism for leaders of groups to understand patterns of group interaction and change without having to sift through numerous group development theories on the fly, and then apply those models correctly to achieve a more adaptive pattern of interaction. Thermodynamic thinking helps us understand how, in groups of all types, the flow of energy creates and transforms group structure, and how that structure constrains or supports future interaction patterns and performance. Leaders and group members can use observations of their current patterns of interaction to assess the thermodynamic "state" of the group. Is it too cold, draining energy from group members? Does it need contact to create a warming pattern? Is it too chaotic, scattering member efforts without clear coordination? If so, the group needs to either lower the group "temperature" or redirect the energy into a more complex structure that transforms heat into productive work. Scanning groups to assess the level of energy, how that energy is flowing, and what links are guiding that flow allows leaders, coaches, or group members themselves to make adjustments in groups without necessarily knowing which overall pattern of group development is presently unfolding.

Through our case analysis, we have shown how our thermodynamic perspective connects with extant research on group development. Scientific rigor is critical to understanding any phenomenon, and research on group development has a long history in that regard. We have no desire to either supplant existing theories or to add yet another variety of group development model to the catalog. Instead our goal is to draw attention to the underlying dynamics from which all the patterns arise. Adjusting these dynamics can help stabilize helpful structures, break down problematic patterns, and both stimulate and regulate the flow of energy that makes both stability and change possible. We hope we have demonstrated that the thermodynamic perspective holds promise as a way to integrate across many existing theories of group development and change, rather than simply add to it. A more parsimonious understanding of change over time in groups that increases practical relevance and maintains methodological rigor could benefit scientists and practitioners alike.

References

Anderson, N., Herriot, P., & Hodgkinson, G. (2001). The practitioner-research divide in industrial work and organizational (IWO) psychology: Where are we now and where do we go from here? *Journal of Occupational and Organizational Psychology, 74,* 391–411.

Arrow, H. (1997). Stability, bistability, and instability in small group influence patterns. *Journal of Personality and Social Psychology, 72*, 75–86.

Arrow, H. (2005). Chaos, complexity, and catastrophe: The nonlinear dynamics perspective. In S. A. Wheelan (Ed.), *The handbook of group research and practice* (pp. 201–219). Thousand Oaks, CA: Sage.

Arrow, H., McGrath, J. E., & Berdahl, J. L. (2000). *Small groups as complex systems: Formation, coordination, development, and adaptation.* Thousand Oaks, CA: Sage.

Arrow, H., Henry, K.B., Poole, M. S., Wheelan, S., & Moreland, R. (2005). Traces, timing, and trajectories: The temporal perspective on groups. In M. S. Poole & A. B. Hollingshead (Eds.), *Theories of Small Groups: Interdisciplinary Perspectives* (pp. 313–367). Thousand Oaks, CA: Sage.

Ashby, W.R. (1956). *An introduction to cybernetics*. London: Chapman & Hall.

Bales, R. F. (1950). *Interaction process analysis.* Reading, MA: Addison Wesley.

Bales, R. F. (1955). Adaptive and integrative changes as sources of strain in social systems. In A. P. Hare, E. F. Borgatta, & R. F. Bales (Eds.), *Small groups: Studies in social interaction* (pp. 127–131). New York: Knopf.

Bennis, W. & Shepard, H. (1956). A theory of group development. *Human Relations, 9,* 415–437.

Betancur, A. (2005). Energy made visible: Behavioral effects of social energy. Honors thesis, Boston College. http://hdl.handle.net/2345/401

Canavan, D. (1996). A mind of one's own: The suchness of things. *Proceedings of the 21st Annual International Conference on Improving University Teaching,* Nottingham, UK.

Carley, K. M., & Svoboda, D. M. (1996). Modeling organizational adaptation as a simulated annealing process. *Sociological methods & research*, *25*(1), 138–168.

Ciardi, M. & Gray, G., et al. (Producers), & O'Connor, G. (Director). (2004). *Miracle* [motion picture]. United States: Walt Disney Studios.

Collins, B. E. & Guetzkow, H. (1964). *A social psychology of group processes for decision-making.* New York: Wiley.

Cummings, J. (2004). Work groups, structural diversity, and knowledge sharing in a global organization. *Management Science, 50*(3), 352–364.

Cummings, J.N. & Ancona, D.G. (2005). The functional perspective. In S. Wheelan (Ed.), *The Handbook of Group Research and Practice* (pp. 107–117).

Eldredge, N. & Gould, S. J. (1972). Punctuated equilibria: An alternative to phyletic gradualism. In T. J. M. Schopf (Ed.), *Models in paleobiology* (pp. 82–115). San Francisco, CA: Freeman, Cooper, & Co.

Festinger, L. (1950). Informal social communication. *Psychological Review, 57,* 271–282.

Forsyth, D. R. (1990). The pecking order. In R. Brown (Ed.), *Human behavior: How groups work* (Vol. 15, 1820–1827). New York: Marshall Cavendish.

Gersick, C. J. G. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal, 31,* 9–41.

Gersick, C. J. G. (1989). Marking time: Predictable transitions in task groups. *Academy of Management Journal, 32,* 274–309.

Gersick, C.J.G. (1991). Revolutionary change theories: A multilevel exploration of the punctuated equilibrium paradigm. *Academy of Management Review*, *16*(1), 10–36.

Gersick, C. J. G. & Hackman, J. R. (1990). Habitual routines in task-performing groups. *Organizational Behavior and Human Decision Processes, 47,* 65–97.

Gilbert, J. (2008). *Herb Brooks: The Inside Story of a Hockey Mastermind*. Minneapolis, MN: Voyageur Press.

Hannan, M. T. & Freeman, J. (1984). Structural inertia and organizational change. *American Sociological Review, 49*(2), 149–164.

Henry, K. B., & Arrow, H. (2010). *Making heat work: The thermodynamics of groups.* (Technical Report No. 10–01). Eugene, OR: Institute for Cognitive and Decision Sciences, University of Oregon.

Kuypers, B., Davies, D. & Glaser, K. (1986). Developmental arrestations in self-analytic groups. *Small Group Behavior, 17,* 269–302.

Larson, J. R., Jr. (2007). Deep diversity and strong synergy: Modeling the impact of variability in member's problem solving strategies on group problem solving. *Small Group Research, 38*(3), 413–436.

Larson, J. R., Jr. (2010). *In search of synergy in small group performance.* New York: Psychology Press.

Levine, J. M. & Moreland, R. L. (1990). Progress in small groups. *Annual Review of Psychology, 41,* 585–634.

Levine, J. M., & Moreland, R. L. (1995). Group processes. In A. Tesser (Ed.). *Advanced Social Psychology* (pp. 419–466). New York: McGraw-Hill.

Littlefield, B. (2015, June 6) Hollywood Scores A 'Miracle' With Locker Room Speech. Retrieved 21 February 2018 from www.wbur.org, Only A Game website. http://www.wbur.org/onlyagame/2015/06/06/us-miracle-olympics-herb-brooks

Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research, 11*(3), whole issue.

Marsh, H. W. & Bailey, M. (1993). Multidimensionality of students' evaluations of teaching effectiveness: A profile analysis. *Journal of Higher Education, 64,* 1–18.

Miracle on ice pre-game speech (text of speech by Kurt Russell playing Herb Brooks). (n.d.) Retrieved 21 February 2018 from Genius.com website. https://genius.com/Kurt-russell-miracle-on-ice-pre-game-speech-annotated

Moreland, R. L., Hogg, M. A., & Hains, S. C. (1994). Back to the future: Social psychological research on groups. *Journal of Experimental Social Psychology, 30,* 527–555.

O'Connor, K. (2013). Class participation: Promoting in-class student engagement. *Education*, *3*(5), 340–344.

Poole, M. S., Van de Ven, A. H., Dooley, K., & Holmes, M. E. (2000). *Organizational change and innovation processes: Theory and methods for research.* New York: Oxford University Press.

Rochlin, G.I., LaPorte, T. R., & Roberts, K. H. (1987). The self-designing high-reliability organization: Aircraft carrier flight operations at sea. *Naval War College Review, 40*(4), 76–90.

Shaw, M. E. & Ashton, N. (1976). Do assembly bonus effects occur on disjunctive tasks? A test of Steiner's theory. *Bulletin of the Psychonomic Society, 8*(6), 469–471.

Stein, R. T. & Heller, T. (1979). An empirical analysis of the correlations between leadership status and participation rates reported in the literature. *Journal of Personality and Social Psychology, 37,* 1993–2002.

Steiner, I. D. (1966). Models for inferring relationships between group size and potential group productivity. *Systems Research and Behavioral Science, 11*(4), 273–283.

Steiner, I. D. (1972). *Group processes and group productivity.* New York: Academic Press.

Tuckman, B. W. & Jensen, M. A. (1977). Stages in small group development. *Group and Organizational Studies, 2,* 419–427.

Weick, K. E. & Roberts, K. H. (1993). Collective mind in organizations: Heedful interrelating on flight decks. *Administrative Science Quarterly, 38*(3), 357–381.

Wheelan, S. A. (1994). *Group processes: A developmental perspective*. Boston: Allyn & Bacon.

Wheelan, S. A. (2009). Group size, group development, and group productivity. *Small Group Research, 40*(2), 247–262.

Wheelan, S. A., Davidson, B., & Tilin, F. (2003). Group development across time: Reality or illusion? *Small Group Research, 34*(2), 223–245.

Wheelan, S. A. & Kaeser, R. M. (1997). The influence of task type and designated leaders on developmental patterns in small groups. *Small Group Research, 28*(1), 94–121.

Wheelan, S. A., Verdi, A. F., & McKeage, R. (1994). *The group development observation system: Origins and applications*. Philadelphia: GDQ Associates.

Wilke, H. A. M. (1996). Status congruence in small groups. In E. Witte & J. H. Davis (Eds.), *Understanding group behavior, Vol. 2* (pp. 67–91). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Wilson, G. C. (1986). *Supercarrier.* New York: Macmillan.

Worchel, S. (1994). You can go home again: Returning research to the group context with an eye on developmental issues. *Small Group Research*, *25,* 205–223.

Worchel, S. Coutant-Sassic, D. & Grossman, M. (1992). A developmental approach to group dynamics: A model and illustrative research. In S. Worchel, W. Wood, & J. A. Simpson (Eds.), *Group process and productivity.* Newbury Park, CA: Sage.

Zhang, Q. (2014). Assessing the effects of instructor enthusiasm on classroom engagement, learning goal orientation, and academic self-efficacy. *Communication Teacher, 28*(1), 44–56.

The Building Blocks of Religious Systems: Approaching Religion as a Complex Adaptive System

Richard Sosis

1 Introduction

Over a recent Thanksgiving holiday I visited my parents, who still reside in the house in which I was raised. During the holiday weekend, they invited some longtime family friends over for a meal. I was talking with one of the guests and I asked him how the local synagogue was faring, as I recalled from my youth that he was an active member. About a decade ago, the synagogue in question had moved from its original downtown location to the suburbs to be in closer proximity to the majority of its members, such as my parents. The new building is gorgeous, modern, and considerably larger than the downtown structure, which after more than 50 years as a house of God, was transformed into a hospital parking lot. Our family friend sighed in response to my question and then answered that membership in the synagogue had plummeted in recent years. He noted that it had been assumed that building an elegant facility closer to its constituency would attract more members, but in fact, the move seemed to have had the opposite effect. Surprisingly, membership was about half of what it was in the downtown location. And this was not a result of an exodus of Jews from the area; in fact, the Jewish population in the suburban residential area surrounding the synagogue has likely increased in recent years. It seems, he explained, that Jews who had in previous generations felt an obligation to join the synagogue to support the institution, even if they did not frequently worship there, no longer felt the communal pull to contribute to an institution that they had no intention of using regularly.

Our family friend was not optimistic about the long-term survival of the synagogue. If current trends continued, membership dues would be insufficient to

---

R. Sosis (✉)
Department of Anthropology, University of Connecticut, Mansfield, CT, USA
e-mail: richard.sosis@uconn.edu

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2_19

support such a large building. One synagogue tradition that I fondly remembered was the full bagels and lox breakfast spread that followed the daily morning prayers. I knew our family friend had been a regular attendee at the morning minyan, as it is known, so I asked whether he still attends. He responded that despite a membership of several hundred families, they are rarely able to secure a minyan, that is, the minimum of ten people required to say certain prayers. The breakfast tradition had passed into history, although a group of four retirees who are committed to the minyan regularly convene at a local diner for breakfast following prayers. Indeed, our friend confided that he often just skips the minyan and meets the retirees at the diner for breakfast.

I share this anecdote because it captures four, or at least four, important aspects about religion. First, religions respond to local conditions; in other words, they seek to adapt. Sometimes, of course, they fail; history is littered with many more extinct religions than ones that have endured the test of time. This leads to the second observation: religions are difficult to predict, even for those who are actively engaged in the religion. The leaders of the synagogue I grew up in quite reasonably thought that moving the synagogue would benefit the community, not contribute to its demise. Third, and at the moment probably less apparent than the first two points, ritual is the lifeblood of religions. As I explained to our friend, I did not think the minyan was failing because membership rates were falling (with several hundred families, they could have easily supported dozens of minyanim) but rather that membership was falling because the minyan itself was failing. The fourth point is implied by the previous three aspects: religions are systems.

These insights stem from an approach to religion that I have advocated over the past decade (Sosis 2009, 2016). Specifically, I have proposed with my colleagues Candace Alcorta (Alcorta and Sosis 2005, 2006; Sosis and Alcorta 2003, 2004), Benjamin Purzycki (Purzycki et al. 2014; Purzycki and Sosis 2009, 2010, 2011, 2013), John Shaver (Shaver et al. 2016; Sosis and Shaver 2015), and Jordan Kiper (Kiper and Sosis 2014, 2016; Sosis and Kiper 2014a, b, 2017) that religion may best be understood as an adaptive complex of traits incorporating cognitive, neurological, affective, behavioral, and developmental elements. We argued that these traits derive from prehuman ritual systems and were selected for in early hominin populations because they contributed to the ability of individuals to overcome ever-present ecological challenges. By fostering cooperation and extending the communication and coordination of social relations across time and space, these traits served to maximize the potential resource base for early human populations, thereby increasing individual fitness. The religious system is an exquisite, complex adaptation that serves to support extensive human cooperation and coordination, and social life as we know it.

The goal of this chapter is to describe the religious system, or specifically, how religion can be understood as a complex adaptive system. I will begin by describing the building blocks of religious systems. Then I will outline the feedback processes that constitute the religious system. This will be followed by a discussion of some implications of this systemic approach and why the religious system is best understood as a complex adaptive system.

2 Religions as Systems

I have asserted that religions are systems, but I am not alone in making such a claim. Two of the most influential anthropologists to study religion in the last 50 years, Clifford Geertz and Roy Rappaport, both approached religion as a system, although they did so in very different ways. One of Geertz's most celebrated articles, "Religion as a Cultural System" (1967/1973), embeds the study of religion in the nexus of a hermeneutic approach to culture, whereas Rappaport's magnum opus, *Ritual and Religion in the Making of Humanity* (1999), describes religion as a cybernetic system.

Despite the unabated influence of both Geertz and Rappaport's work, the academic study of religion has not taken a systemic turn. There are a few exceptions, however. For example, Czachesz (2014) develops a formal network model to explore the evolutionary dynamics of religious systems. His model relies on the interrelationship between religious beliefs, such as toward a god or spirit, and religious artifacts, including texts and ritual objects. Cho and Squier (2013) cautiously explore the merits of recognizing religions as complex systems, most notably the facilitation of cross-cultural comparisons. They are at least partially motivated by concerns of reductionism in the scientific study of religion (Cho and Squier 2008).

Notwithstanding these efforts to advance a theoretical foundation for a systemic approach, most religious studies scholars remain highly suspicious of attempts to generalize "religion" as a system, as Cho and Squier, as well as Czachesz, recognize. Nonetheless, advances in the study of complexity over the past several decades suggest that religion not only possesses systemic features but also can be characterized as a complex adaptive system. While the term complex adaptive system is defined elsewhere in this volume, it is worth clarifying how I have understood this term in my studies of religion, and how the term has entered anthropological discourse (Lansing 2003), since transdisciplinary terms can suffer shifts in meaning as they cross disciplinary boundaries. This is particularly true in the case of complexity studies; one of the few unanimities in the field appears to be that it is not (yet) a unified field (Mitchell 2009).

Systems, whether economic, political, or digestive, can be described as a set of interacting or interdependent elements that form an integrated whole (Von Bertalanffy 1972). "Complex" is also a term that carries colloquial meaning, but it holds a specific meaning in the context of complex adaptive systems. Miller and Page (2007: 9) explain that "[c]omplexity arises when the dependencies among the elements become important. In such a system, removing one such element destroys system behavior to an extent that goes well beyond what is embodied by the particular element that is removed... Complicated worlds are reducible, whereas complex ones are not." Lastly, complex adaptive systems are "adaptive" in the sense that they are flexible and they respond successfully – in terms of the system's survival – to local social and ecological conditions.

Yet trying to explain what complex adaptive systems are by defining the words that constitute its taxonomic label is not sufficient and, in fact, highlights a key characteristic of actual complex adaptive systems: they are more than the sum of their parts (Holland 1998; Kauffman 1995). Complex adaptive systems possess a number of defining features; I describe these below and discuss how they are manifest within religious systems. First, however, I explore how religious systems are built.

2.1 The Building Blocks of Religious Systems

Holland (1995) lists "building blocks" as one of the seven basic characteristics that are common to all complex adaptive systems. As Holland explains, the component parts of any complex adaptive system are not arbitrary. Rather, complex adaptive systems consist of categories of elements that combine to create the system. The building block mechanism, that is, the combining of elements to create a system, inevitably generates astonishing variation. Holland illustrates this variation by considering the construction of faces. The building blocks of faces are features such as noses, ears, mouths, eyes, hair, cheeks, foreheads, and chins. Within each of these building block categories, there is variation, such as size, shape, and color: large oval green eyes, long wavy blonde hair, broad angular brown chin, and so forth. Using this variation, an almost limitless number of unique faces can be built simply by combining facial building blocks.

Religious systems also exemplify remarkable diversity, yet underlying all of this diversity is a set of recurring core features. Religious systems typically maintain eight building blocks: authority, meaning, moral obligation, myth, ritual, sacred, supernatural agents, and taboo. Each of these building blocks is most usefully conceived of as a unique category that may have an independent phylogenetic history, but within religious systems, they are inherently interconnected to the other building blocks within the system. Other features of religion are common, such as music, spirit possession, afterlife beliefs, prophecy, superstition, and pilgrimage, but they are not *essential* to the working of the religious system; rather, they are better understood as secondary forms of one of the essential building blocks identified above.

Here I describe the religious system's building block categories.

Ritual Rappaport defines ritual as "the performance of more or less invariant sequences of formal acts and utterances not entirely encoded by the performers" (1999: 24). One of the most obvious aspects of ritual is that it requires human bodies; it is a physical action (Rappaport 1979). This banal observation turns out to be critically important, as will be evident below. Without performance, there is no ritual; unlike myth, for example, rituals are dead if they only exist in books or memories. Rappaport's definition, like most definitions of ritual (e.g., Smith 1979; Turner 1969), recognizes the formality of ritual in the sense that rituals are typically stylized, repetitive, and stereotyped. This formality distinguishes ritual from ordinary behaviors. Norms surrounding ritual generally define the appropriate times and places of performance. And rituals, of course, vary considerably in their intensity and pageantry (Whitehouse 2004); singing from the psalter on a Sunday morning and diving into a partially frozen lake to retrieve a cross are both religious rituals. As I will discuss below, rituals are the central building block of religious systems.

Taboo Taboos are often conceived of as anti-rituals. That is, whereas rituals must be performed, taboos restrict behaviors. Religions maintain taboos on countless activities, including the consumption of food and drink, social relationships, sex, smoking, gambling, wearing jewelry, exposing certain body parts, types of work, and so on. Some taboos are always in effect, such as Mormon prohibitions on smoking or Muslim bans on pork, but many taboos are temporally regulated, such as Catholic meat consumption during Lent. Taboos are effective at creating boundaries between populations, often limiting the types of social engagements that are possible (Douglas 1966).

Authority Religious systems generally have leaders or authority figures who possess particular power and influence within a community. These authorities include priests, prophets, gurus, magicians, shamans, imams, rabbis, ministers, seers, sorcerers, and witches. It is worth emphasizing that each of these authorities have distinct roles and functions within their respective communities. Yet such individuals similarly gain and maintain their authority through access to specialized knowledge and/or specialized access to supernatural beings and the worlds these beings inhabit. Often, individuals rise to prominence as a religious authority because of their charisma and oratory skills (Weber 1947). While some religious leaders, such as the pope, command considerable authority, others such as a synagogue president lead primarily by following the will of the community.

Myth Myths within religious systems serve to provide a contextual narrative for many of the other building blocks of religious systems. For example, myths often describe the origins and deeds of the gods, the reasons for certain rituals and taboos, and how religious leaders have been granted their authority. Myths also serve as explanations for phenomena, such as the sanctity of hills, rivers, and cities, the creation and history of a people, and the justifications for specific moral obligations. However, myths do not offer logical, or what might be described as scientific, explanations (Lévi-Strauss 1964/1994). Myths are lived explanations. They are better understood as one of the fundamental forms of religious discourse, if not the fundamental form. Indeed, communities often define themselves by the myths they share, and divergences in narratives often demarcate the fissure point between religious groups, such as the differing narratives of Jesus among the Abrahamic religions. Myths not only inform group boundaries and identity, they are of course entertaining, both frightening and alarming, and even humorous, such as trickster tales (Radin 1956/1972). For many, religious myths provide a window into what it means to be human.

Sacred Durkheim (1912/1995) famously distinguished between the sacred, that which is set apart, and the profane, that which is mundane. As many anthropologists have remarked (Alcorta and Sosis 2005; Rappaport 1999), sanctity is not discovered through encounter but rather created through ritual. Since ritual can sanctify just about anything, the range of sacred things is quite varied, including food items, books, land, clothing, weapons, animals, plants, people, ideas, symbols, words, and discourse. In the context of religious systems, to say that something is sacred is to suggest that it has particular emotional valence for individuals living within that system.

Supernatural Agent Supernatural agents are beings that exist and operate outside of physical reality, although they typically have impacts on the physical world. They are agents in the sense that they are ascribed actions and motives for those actions (Purzycki et al. 2012). The term "supernatural" itself is contested in the religious studies literature because many cultures do not perceive their gods, spirits, ghosts, demons, angels, and assorted beings as supernatural at all. Rather, these beings are perceived as a natural part of their social and physical landscape (Klass 1995). The concept of the supernatural, these scholars correctly argue, is a western concept, and when we impose it on non-Western cultures, we fail to understand how non-Westerners experience their religious worlds. Westerners often relate to their supernatural agents through what is described as "belief" or "faith." Such a relationship to supernatural agents, though, is quite foreign to many of the small indigenous populations that anthropologists have studied, as well as many historical populations (Kugel 2017; Smith 1998). Nonetheless, while I appreciate that individuals within such populations might not conceive of particular beings as supernatural, I employ the term supernatural agents to describe a building block category because it offers a useful analytic position – an etic perspective – that distinguishes these agents from living beings (humans and animals), as well as fictional characters that populate contemporary entertainment genres.

Moral Obligation Anthropologists have long pointed out that religious systems and moral systems that are concerned with social ethics are generally distinct in traditional societies. It is only after the rise of domestication and intensive agriculture, and the associated sedentary lifestyle, that an indelible link between religious systems and social ethics emerges. Nonetheless, even in traditional foraging societies in which moral prohibitions such as sexual infidelity, theft, and murder are not supported through religious sanctions (i.e., they are immoral behaviors, but there are no priests or holy books that give such prohibitions authority), moral obligation is often established through religious systems, particularly through ritual (Kiper and Sosis 2014). Rappaport (1999: 132) suggests that breach of obligation may be "one of the few, if not, indeed, the only act that is always and everywhere held to be immoral." In short, Rappaport argues that ritual performances establish obligations to behave according to the moral values explicitly or implicitly encoded in the rituals.

Meaning Religious systems not only offer explanations for the existence of humanity and the existence of particular communities, they offer a purpose for this existence. The comprehensiveness of religious systems – their ability to inform every aspect of an individual's life – enables them to generate meaning for individuals. Religious systems work to keep nihilism at bay, and they enable adherents to make sense of their lives and give it purpose. Religious meanings range from fixing society through social justice to seeking individual salvation to preparing the world for the messiah. When religious systems are functioning optimally, religious meanings powerfully organize lives and establish order within communities (e.g., Levine 2003).

Why have these eight features served as building blocks for religious systems? To adequately answer this question, we would need a phylogenetic account of the emergence of the religious system, and unfortunately, our understanding of religion as a complex adaptive system is not yet developed enough to pursue such an analysis. Nonetheless, we do know enough to recognize that like all complex adaptive systems (Miller and Page 2007), the elimination of one of the religious system's building blocks would result in either a collapse or transformation of the system into something else. Remove one of the building blocks, and the adaptive functionality of the religious system will be compromised. A religious system that lacks organization (authority), fails to impart significance (meaning), forsakes ceremonial activity (ritual), maintains no limits on activity (taboo), is unable to sanctify anything (sacred), offers no beings capable of transcending the natural world (supernatural agents), does not establish social commitments (moral obligation), or provides narratives that can link all these elements together into an explanatory framework (myth) will ultimately falter or transform into another social institution. Notably, religious systems are impressively resistant to the elimination of one of their core building blocks, as the persistence of supernatural elements within Buddhist cultures and Jewish Reconstructionist congregations attest.

Holland (2012) distinguishes between two types of building blocks: generators and conglomerates. The imperviousness and essentialness of the religious system's building blocks suggest that they are generators. Generators do not change over their life course, and they behave consistently in similar contexts. Over their existence, they fit together with other generators in the same way and follow the same set of fixed rules in their reaction with other generators. Conglomerates, on the other hand, are building blocks that do change over their lifespan, sometimes dividing to produce other building blocks or adjusting to compensate for the loss of another building block. For example, the brain's ability to self-repair through rewiring or taking over the functions of damaged areas suggests that the brain is built on conglomerate building blocks (Holland 2012: 112). It is possible that some of the religious system's building blocks should be conceived as conglomerates, but this is an area that needs further investigation. In social environments where governments seek to eliminate specific building blocks of religious systems, the conditions might provide a natural experiment in which this possibility could be explored.

A few further comments concerning the religious system's building blocks are necessary. First, the building blocks identified above are likely universal across religious systems, but they are not building blocks because of their universality; there are other universal features of religions that are not core building blocks (e.g., symbolization, the creation of alternative worlds). Rather, these features are building blocks because they each appear to play a distinct and integrative role within religious systems. Second, and related to the first point, the identification of these eight core features is based on our understanding of how religious systems work. In other words, religious systems appear to function in particular ways and exhibit specific structural features, as I will describe below. Third and relatedly, while some common features of religion, such as pilgrimage, altered states of consciousness, spirit possession and so forth, are best depicted as falling within one of eight building block categories (e.g., altered states of consciousness are achieved through ritual), some religious systems may develop such that secondary features become building blocks. For example, the centrality of the peyote hunt among the Huichol Indians of northern Mexico (Myerhoff 1974) may suggest that pilgrimage is a building block of this particular religious system; without the pilgrimage for peyote, the system would likely collapse. The characterization of religious systems that I describe below depicts how most religious systems generally function, but as just noted, there are undoubtedly exceptions to these generalizations.

Our understanding of how the religious system's building blocks are put together is rudimentary, but we can make a few observations with some confidence. The adhesive holding the building blocks together is language. Indeed, the religious system is inconceivable without language; discourse serves to indicate the moral obligations conveyed in ritual performance, describe unseen supernatural agents, articulate prohibited behaviors, reveal myths, and so on.

While language undoubtedly serves to connect the religious system's building blocks, there appear to be inherent patterns through which these building blocks interrelate. These patterns likely constitute a grammar (Bulbulia 2012), and it is the ongoing task of evolutionary and cognitive researchers studying religion to uncover these grammatical rules. This task, however, is genuinely challenging because of the nonlinear nature of the interactions between the building blocks. As Holland (2012) points out, standard statistical observations will not suffice to uncover the rules interconnecting building blocks for any complex adaptive system.

Nonetheless, evolutionary and cognitive researchers have offered various theories that provide insights into the possible rules regulating relations between religion's building blocks. For example, the Modes Theory of Religion provides a useful framework for understanding the mechanisms that enable rituals to create meaning, and significantly, how variation in the frequency of ritual performance is related to variation in the formation of meaning (Whitehouse 2004). Likewise, Cronk's (1994) theory of signal manipulation provides a powerful explanation of how religious authorities employ ritual for means of exploitation (also see Watts et al. 2017). Sacred values research (Ginges et al. 2007; Tetlock 2003), offering another example, highlights the relationship between the sacred-profane distinction and taboo. Other theories, such as supernatural punishment theory (Johnson 2016;

Schloss and Murray 2011), signaling theory (Irons 2001; Bulbulia and Sosis 2011), MCI theory (Purzycki and Willard 2016), hazard precaution system theory (Lienard and Boyer 2006), and ritual form theory (McCauley and Lawson 2002), offer further insights about how all the building blocks within religious systems interrelate.

We are now in a position to provide a general outline of the structure of religious systems and how they work, even if our understanding of the how religion's building blocks interact remains limited. We presently turn to this outline.

2.2 *The Structure of Religious Systems*

Figure 1 depicts a general illustration of the structure of religious systems. Religious systems begin with a group of socially engaged individuals. Individuals are agents of the model, and they can enter as well as depart from the group. Like all communities, the group is influenced by external factors including the social, political, economic, ecological, and religious environment in which the group is situated. Notably, however, religious groups are not simply influenced by their external conditions, they actively shape them (Bulbulia 2012; Purzycki and Sosis 2013). These external factors, as well as the internal social dynamics of the group, motivate human action

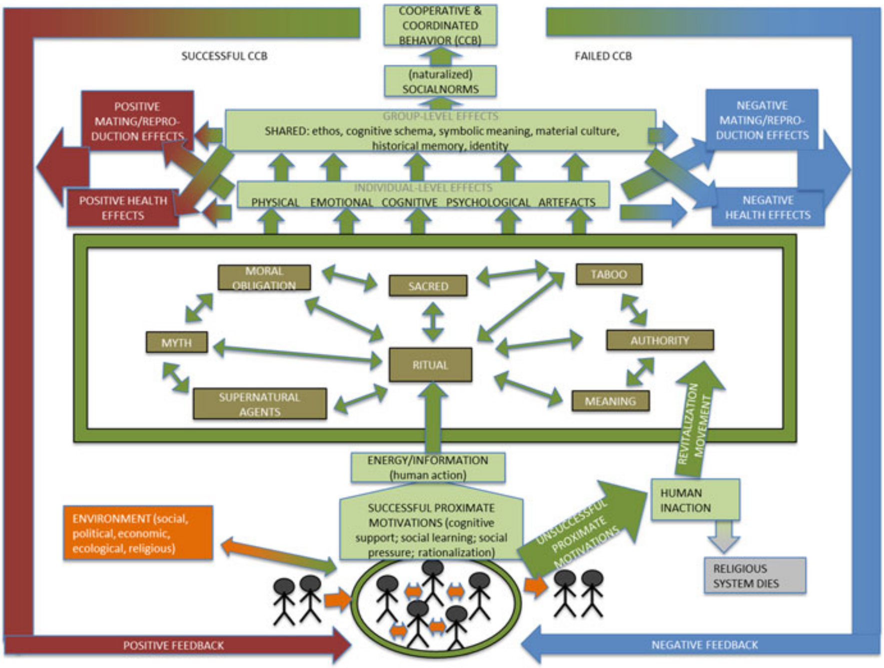


Fig. 1 The religious system

in the form of ritual behavior. Like all systems, religious systems require energy to function. Energy is introduced into the religious system through human action in the form of ritual. Ritual behavior contains energy, as well as social information, that enters the religious system. All systems transform energy; likewise, the religious system transforms the energy and information of human ritual behaviors into human cooperative and coordinated behaviors.

Since ritual is a physical performance, it may be self-evident that it carries energy; however, the information carried by ritual is less obvious. Ritual, as Rappaport (1999) explains, carries two types of information: indexical and canonical. Indexical information refers to messages that reveal the current state of the performer. For example, consider prayer. How loudly one prays or how vigorously one gesticulates can indicate the enthusiasm of the reciter, whereas one who prays with a scowl or teenage eye-roll on their face, or mumbles through their prayers, suggests a less enthusiastic endorsement of Sunday morning in the pews. Rituals also contain information about the past and future; indeed, rituals often appear enduring or even eternal to performers. This information is known as canonical, and it is often (but not always) contained in the verbal part of the ritual (Rappaport's "utterances" from the definition above). Continuing the example of prayer, the actual words being recited in prayer are canonical information. Moral codes are typically embedded, implicitly or explicitly, in ritual's canonical information.

How does human action in the form of ritual behaviors emerge from social groups? The proximate motivations are likely to be diverse and socioecologically vary. Fortunately, to appreciate how religious systems operate, we do not have to fully apprehend this process, although this is an important area for future work. We do know that rituals spontaneously emerge when communities are under threat. The ethnographic literature on cargo cults (e.g., Whitehouse 1995; Worsley 1957) offers abundant examples, and ethnographers have also detailed ritual practices that have emerged during times of war, such as psalm recitation (Sosis 2007; Sosis and Handwerker 2011). But it is clear that "community threat" is just one avenue through which ritual behaviors arise; new technologies, social movements, ecological changes, and demographic factors are among the many potential avenues that need to be further explored.

While we await future work that examines these external forces, it is important to recognize that the structures of rituals themselves play a role in their emergence. Specifically, successful ritual behaviors, that is, those that are performed and passed on to future generations, require cognitive support. Humans have implicit understandings of how rituals are supposed to work (Barrett and Lawson 2001; McCauley and Lawson 2002), and it is likely that rituals that are successfully motivated (i.e., brought to life) are those that are consistent with these implicit understandings. Put simply, some rituals are more compelling than others; those rituals that take a form that is congruent with cognitive expectations are likely to be more compelling than rituals that are more difficult to mentally process.

It is useful to distinguish between proximate motivations that can explain the emergence of ritual behaviors within a group from proximate motivations for continuing ritual behaviors that have previously stabilized within a community.

Mechanisms for the continuance of ritual behaviors include social learning (Henrich 2009), social pressure (Sosis 2003), and rationalizations about the efficacy of the ritual behaviors (Sosis and Handwerker 2011). Significantly, there appear to be critical developmental windows that facilitate the generational transmission of ritual knowledge (Alcorta and Sosis 2005; Finkel et al. 2010).

It is somewhat misleading to describe energy entering religious systems in the form of ritual behavior, because in fact ritual only manifests its full character through its interactions with other elements within the system. More accurately, human intentions and motivations enter the system and are transformed into ritual action through their interactions with other elements of the system. Subsequent inputs into the system take the form of the behavior transformed into religious ritual. Once energy enters the religious system through ritual behaviors, the elements that constitute the system interact with ritual behavior in feedback loops. Within religious systems, for example, ritual behaviors become associated with supernatural agents. Supernatural agents can take on various roles in ritual performance, such as the recipient of sacrificed food or the target of petitionary prayers. But whether supernatural agents are seen as receivers, creators, or enforcers of a ritual performance, once such agents become linked to a ritual, desires to please or appease the agents can proximally motivate the ritual performance. Indeed, the human action that emerges from the social group that provides the seeds of the system will be transformed into what we recognize as religious ritual once it interacts and incorporates the elements of the religious system.

The interaction of the religious system's core building blocks results in five primary individual-level effects: physiological, emotional, cognitive, neurological, and technological effects. The first four are internal responses of ritual performers. These responses span the entire gamut of human experience and are likely to vary significantly depending on whether the base rituals are dysphoric, such as Hindu fire walking (Power 2017), or euphoric, such as Sufi dancing (Trimingham 1971). In addition to these internal effects on individuals, the religious system also produces a primary external effect, or what can be considered an extended phenotype (see Purzycki and Sosis 2013), in the form of ritual objects. These artifacts include masks, mats, beads, pipes, attire, and countless other items fashioned by individuals immersed in their religious system.

These primary individual-level effects yield various group-level effects. Specifically, group-level effects include shared cognitive schema, ethos, symbolic meaning, material culture, historical memory, and group identity. Group-level effects are an emergent property of the religious system, and they can powerfully shape individual lives. Indeed, group-level effects produce societal order by creating structured and stable social worlds – often fantastically imaginative – that individuals inhabit and navigate. They are also generally the most salient features of religious systems and why religions are typically characterized as a collective phenomenon. These effects also give form to the initial population, providing the basis for sustained communal engagement.

2.3 The Emergence of Social Norms in Religious Systems

What emerges from these group-level effects are social norms; specifically, expectations and patterns of behavior that characterize communities. This is no small matter. Humans are able to conceive of alternative ways of engaging, understanding, and organizing life. Consequently, our social norms – that is, the way we pattern our lives – are always at risk of modification (Seligman and Weller 2012). Rappaport (1999) argues that this potential instability is minimized because our social norms become internalized and naturalized. One of the extraordinary features of human experience is that individuals view the norms in which they are entwined as a natural part of their existence (Berger 1967). Indeed, it is only when one is in tension with societal norms, such as facing away from an elevator door, that norms feel unusual or extraordinary.

Rappaport (1999) suggests that ritual plays a key role in the emergence of social norms from religious systems. Let us consider his argument. Rappaport observes that ritual's inherent structure is binary; one either performs a ritual or not. He maintains that while ritual behaviors appear to be shrouded in mystery, they are deliberate and their message to others is clear: participation in a ritual performance indexically signals acceptance of (and not necessarily belief in) the moral values encoded in the ritual. Participation, therefore, always carries obligations, and participants can be held accountable if these obligations are compromised. Lovers can be unfaithful, but adultery can be committed only after the marriage ceremony. Notably, flipping the correspondence theory of truth on its head, when the world does not conform to the moral order encoded in rituals, it is not the ritual that is wrong, but rather the world that needs to be adjusted.

As many anthropologists have emphasized, including Rappaport, the sacred is not discovered, but rather it is created through ritual. Sanctity emerges from the structure of ritual itself via at least two pathways. First, some rituals can evoke numinous experiences. The power of these experiences, which are undeniable to those who experience them (e.g., D'Aquili et al. 1979), make the discursive aspects of ritual unquestionable, that is sacred. Second, rituals consist of behavioral and discursive components. Rappaport (1999) describes the latter as invariant. Again, consider prayer, which consists of bodily movements as well as words. While the intensity of bodily movements, including how loudly prayers are recited, can vary across performers, what is recited (i.e., the canonical messages) is largely invariant. Canonical messages are not encoded by the performer (those sitting in the pews did not write the Psalms), and all performers utter the same words; therefore, this verbal aspect of ritual lacks information (e.g., Bloch 1974). But, it is argued that the meaning of the "informationlessness" that emerges from canonical invariance is certainty, which is understood to be unquestionable and true (Wallace 1966). Consequently, the moral messages carried by ritual seem correct, the arbitrariness of norms is transformed to necessity, and these norms seem natural and continuous with the physical world (Rappaport 1999).

While rituals create and support religious systems, rituals also produce other social constructions such as governments, kin networks, sports teams, and libraries. In other words, rituals spawn many systems; our focus here specifically on religious systems should not obscure the role that these other systems play in the manifestation of social life.

Ritual, sometimes within religious systems and sometimes in other systems, provides the stable grounding of social institutions. While not all institutions or their products, such as collective norms, are directly associated with ritual performance, social institutions are interlocking (e.g., consider how libraries systems are connected to monetary, educational, governmental, and other systems), and it appears that all foundational institutions (e.g., governments, legal systems) engender and sustain rituals. Indeed, this may be one of ritual's defining features. It is plausible that rituals inevitably emerge in such institutions because these physical acts provide the sanctity and grounding for our abstract construction of social reality. Rappaport (1999), consequently, sees ritual as the basic social act; ritual provides the source for the naturalization of social norms, and thus, he maintains that social life would not be possible without ritual performance. Ritual, as Seligman and Weller (2012) note, does not eliminate the ambiguities of social life, but it does enable us to live with this ambiguity.

2.4 Generating Cooperation and Coordination

While religious systems generate diverse social norms through ritual, the norms that sustain religious systems involve community-level cooperation and coordination. The cooperative and coordinated behaviors that are produced via these norms are, in evolutionary terms, the ultimate goals of religious systems. The success of religious systems in motivating cooperative and coordinated behaviors goes a long way toward explaining the emergence and perdurance of this complicated social institution.

The energetic output of religious systems, therefore, is cooperative and coordinated behavior. It is worth bearing in mind that the religious system is a stunningly convoluted way to produce such behavioral responses. Other social organisms have devised ways of achieving collective goals that are less complicated and mysterious. Selection, however, operates on available traits, and the religious system was built on the existing cognitive and behavioral foundation. Also, human language has necessitated complex solutions for sustaining cooperation and coordination. As Rappaport (1999) observes, the symbolic nature of language means there is always the possibility of deceit and lying since the relationships between signs and their significata are arbitrary. Thus, ultimately, actions (i.e., rituals) speak louder than words.

The religious system is cybernetic in the sense that feedback is inherent to its structure. Successful cooperation and coordination supports the group through the successful acquisition of energy, which feeds back into the system. Unsuccessful cooperation and coordination also feed back into the group, and this lack of energetic input informs the group of failure and warns them about impending resource challenges. In addition to these energetic feedbacks, information about health, mating, and reproductive effects also feed back into the group, informing them about group vitality and offering proximate cues about the value of engaging in ritual behavior.

This information about health, mating, and reproduction emerges from individual- and group-level effects. The individual- and group-level effects that impact health can potentially be positive, as numerous studies on the health benefits of religion attest (Koenig et al. 2012). But they can also be negative as many rituals are dangerous, including subincision and scarification ceremonies in unhygienic environments (Hogbin 1970), ritual club fights (Hill and Hurtado 1996), and fire walking (Power 2017). And some religious beliefs can lead to mental instability (e.g., Luhrmann 2012). Moreover, both individual- and group-level effects can impact mating and reproduction, again both positively and negatively. Religions are associated with some of the highest fertility rates in the world (Kaufmann 2010; Shaver 2017), as well as the lowest (e.g., celibate monks, Shakers, etc.).

Within religious systems, when the balance of feedback is positive, individuals perform ritual behaviors, which feed the system with the necessary energy it needs to be sustained. However, when the balance of feedback is negative, proximate factors will not motivate ritual behavior, which will drive the system down one of two possible pathways. If conditions warrant, the group will undergo a religious revitalization (Heimola 2012; Wallace 1966). This will generally require an individual (or group of individuals) who emerges as an inspirational authority that can reinvigorate the group and motivate ritual action. Without the emergence of such a figure, the religious system is likely to die, which has been the fate of the majority of religious systems that have existed in human history. Obviously most religious systems spend much of their existence fluctuating between periods of success, stasis, failure, and revitalization. But ultimately, religious systems either die or transform beyond the recognition of the old system.

2.5 Adaptability of Religious Systems

Religion's ability to change, that is, its ability to adapt, turns out to be one of its most extraordinary, and most misunderstood, features. Religious claims are rarely stagnant or offer permanent truths about the world; they are flexible and respond effectively to changing socioeconomic and ecological conditions (Alcorta and Sosis 2005; Purzycki and Sosis 2009; Sosis 2009). Religions are adaptive systems that are not only responsive to changing conditions, but they are often instrumental in facilitating social change, such as the Ghost Dances among the Sioux (Mooney 1965) and Pawnee (Lesser 1978) and Black Churches in the Civil Rights Movement (Billingsley 1999).

If religions are responsive to changing circumstances, why do religions often appear to be so resistant to change? Why is religion often viewed as a conservative social force? One of the remarkable features of religion is its ability to adapt to local environmental conditions while adherents experience partaking in an eternally consistent and changeless tradition. Rappaport (1999) argues that religion achieves this through a hierarchy of religious discourse, for there is an inverse relationship between the material specificity of a religious claim and the durability of the claim. Religious ideas are hierarchically organized within communities, and at the apex of a community's conceptual hierarchy is what Rappaport refers to as ultimate sacred postulates, such as the *Shahada*, *Shema*, or *Vandana Ti-sarana* for Muslim, Jewish, and Buddhist communities, respectively. These ultimate sacred postulates lack material specificity and are highly resistant to change. However, below ultimate sacred postulates in the religious hierarchy are various cosmological axioms, ritual proscriptions, commandments, directives, social rules, and other religious assertions that do experience varying levels of change, depending on their material specificity.

While the rules of religions change throughout time, those who experience such adjustments consider them as an intensification of their own religious acceptance (Rappaport 1999). Religions rarely invalidate the old completely: change occurs by adding to previous practices and beliefs and also by elaborating upon them, while other beliefs and practices slip away unnoticed. Once sacralization is internalized, it is indeed very difficult to convince adherents that something consecrated is no longer holy. Hence, when undergoing change, religions often retain the most sacralized elements and augment them. For example, Jewish prayers are part of the Catholic Mass and when proselytizing to indigenous populations, missionaries often retain the dates of indigenous ritual celebrations and tolerate the continued commitment to indigenous ancestral spirits (e.g., Shaver 2015). Change for adherents therefore is not experienced as something radically new. It is rather experienced as an increased acceptance of eternal and personally relevant truths that, for the practitioner, have always been part of their religious tradition.

It is important to appreciate that sacred texts such as the Bible, Koran, or Bhagavad-Gita do not impede the ability of religions to adapt. Intuitively, it may seem that once sacred texts become an essential part of a religious system, as they are in contemporary world religions, that the permanence of these texts would make religions more inflexible. In fact, it is a testament (forgive the pun) to the adaptability of religious systems that textual resources often facilitate change.

Religious texts that endure do so because they are open to multiple literary interpretations. They tend to make use of metaphor and poetry that engage subconscious processes of personal significance and create contextual meaning. As a result, each new generation reinterprets religious texts in relation to their own meaningful experiences, thereby keeping them living, relevant, and fresh. Past interpretations are not necessarily rejected per se, but are instead transformed or ignored by the community. They nonetheless remain available should cultural change make their message relevant again. Indeed, the sacred writings of contemporary religious traditions are vast repositories that leaders draw upon, emphasizing aspects that are socially and politically expedient, and disregarding those that are not. Though religious radicals often revive past interpretations to justify their radicalization and violence (Sosis et al. 2012), use of these latent textual resources is not always so contrived and manipulative. For example, the writings of twelfth-century condemned heretic, Peter Abelard, were largely forgotten until his ecumenical voice was "rediscovered" in the nineteenth century, when his writings received a more welcome reception than they did during his lifetime (Armstrong 1993; Carroll 2001).

While religious texts do not inherently impede the ability of religion to adapt, religions are at risk of over-sanctifying texts and other discourse. When religions sanctify – that is, make unquestionable – discourse that is materially specific, it renders the religious system maladaptive. If low-level materially specific discourse, such as social rules about homosexual marriage or the driving of cars by women, is highly sanctified, it limits the religious system's ability to respond adaptively to changing socioecological conditions. These points articulate well with the observations of complexity theorists that complex adaptive systems exist on the edge of chaos (Kauffman 1995). Complex adaptive systems that are too ordered become inflexible, like fundamentalist religions. On the other hand, complex adaptive systems that are too chaotic are unable to gain traction within an environment. Likewise, religious systems without rules and expectations will not endure because the boundaries of such communities will remain undefined.

One last point is worth emphasizing about the adaptability of religious systems: adaptation is local. This is a source of confusion when discussing religions because we use labels such as Hinduism, Islam, Catholicism, and so forth to describe what we imagine are particular religions. But religious systems are local affairs, and therefore, the taxonomic labels that we ascribe to religions actually consist of multiple and often diverse religious systems. Religious systems that fall under the same taxonomic label are linked in the minds and even actions of those who identify with the label. These linked systems, in other words, are part of the environmental input that impacts individuals within a system. Interestingly, complexity scientists recognize that taxonomic labels can influence system dynamics (Holland 1992, 2012). Religious taxonomic labels are often contested with significant implications for human welfare and lives, especially since the rise of nation states (Seeman 2010). The important point here is that the broad taxonomic labels, particularly of major world religions, should not be confused with religious systems; religious systems are locally defined and they locally adapt. World religions consist of many varied localized religious systems. This hierarchical structure is what complexity theorists would anticipate. Holland (2012: 110), for instance, comments that in "most complex adaptive systems, building blocks at one level of complexity are combined to get building blocks for structures at a higher level of complexity."

3 Primary Features of Religions as Complex Adaptive Systems

In addition to building block mechanisms, complexity scholars have delineated many features that are deemed essential to all complex adaptive systems. Here I list some of these features and briefly discuss how they are manifest in religious systems.

Emergence Emergence is a difficult and sometimes heated topic among complexity scientists (see Corning 2002; Deacon 2010; Holland 1998). As is often the case, competing definitions have inflamed debate and complicated rather than clarified matters. Nonetheless, most scholars would agree that emergence is exhibited where properties result from the interactions of a system's components that are not in evidence among the components themselves. Emergence implies that phenomena cannot be explained from linear interactions among the system's components.

Religious systems exhibit emergent properties, including group identities, shared symbolic meanings, and other group-level features discussed above. When the core features of religious systems coalesce, social phenomena are created, specifically group-level properties, which are not in evidence independently among religion's core elements. For example, a child's belief that ghosts reside in her closet does not produce shared symbolic meanings, but belief in ghosts embedded within an animist religious system, such as I encountered during ethnographic fieldwork in Micronesia (Sosis 2005; Spiro 1952), will generate many group-level effects.

The emergent nature of religious systems has significant implications for how we understand religious beliefs. The complex systems approach to understanding religion emphasizes that religious beliefs are not independent propositional claims about the world (Sosis and Kiper 2014a). Religious beliefs emerge from within a cultural system, and they must be understood within that system. In other words, religious belief, as an element of a larger religious system, cannot be analyzed independently of the system in which it is embedded. To do so is like evaluating a symphony when you can hear only one instrument. Moreover, similar to a symphony, religious systems have emergent properties, and thus religion cannot be reduced to independent propositional claims.

Interestingly, the emergent nature of religious beliefs, especially in relation to myth, is a point of potential agreement between some atheists and theologians. Atheist philosopher Daniel Dennett, for instance, argues that telling stories is fundamental to humanity. Notably, he writes "Our tales are spun, but for the most part we don't spin them; they spin us. Our human consciousness, and our narrative selfhood, is their product, not their source" (Dennett 1991: 418). Protestant theologian, Paul Tillich, would likely agree. For him religions employ myth "because symbolic expression alone is able to express the ultimate" (1957: 41). Myths are not history – and Tillich warns that mistaking myth for history is idolatrous – but myths remain powerful because they are able to transcend themselves and express group values and identity (Mecklenburger 2012), a point in which Dennett would likely be in agreement.

One final point concerning emergence: because religions exhibit emergent properties, studying their building blocks independently as though they are not embedded within a religious system can result in misleading conclusions about religions. Since religions are complex and the result of nonlinear interactions of its building blocks, they simply cannot be broken down and easily reassembled. It is the interaction between components of the system that must be understood (Sosis 2009). We will return to this issue below.

Self-Organization Complex adaptive systems spontaneously self-organize, that is, they do not require top-down or bottom-up orchestration to develop; they emerge inherently from the interactions of the system's components. One of the fascinating characteristics of religious systems is that they seem to arise naturally wherever humans live as a community. And as countless commentators of communist China and Russia have remarked, they even arise in the face of social and political forces that are designed to prevent their emergence. Equally intriguing, secular groups who distance themselves from religious dogma nonetheless often adopt the features of religious systems, and some might be considered quasi-religious. Greek fraternities at US universities, for instance, often generate unverifiable mythic narratives, intense ritual routines, and unfalsifiable ideologies concerning "brotherhood" (Shaver et al. 2018; Sosis and Bressler 2003). Likewise, successful secular terrorists, similar to their religious counterparts, employ features of religion, such as emotionally evocative symbols, rituals, and myths (Sosis et al. 2012).

Self-organization does not imply a lack of hierarchical structure. While some complex adaptive systems, such as bird flocks, lack central organization, others, such as the nervous system, develop central controls. Religious systems appear to require hierarchical organization – notably, religious leaders in the form of priests, shamans, healers, gurus, prophets, and so forth emerge – despite countless attempts to build perfectly egalitarian religious communities.

Unconsciousness Constituent entities within complex adaptive systems are unconscious of the process of self-organization. The design of complex adaptive systems can appear to observers to be ingenious, such as the functioning of biological cells, but no intelligence is required among constituent entities, in this case, ribosomes, lysosomes, Golgi apparatus, and so forth. Moreover, constituent entities are often completely unaware of the entirety of the system; their information is limited to their local environment.

While religious systems self-organize without top-down or bottom-up orchestration, like most complex adaptive systems, the agents – i.e., humans – lack complete information about the workings of the systems they enliven. Their interactions, which result in the formation of religious systems, are generally not consciously aimed at creating religious systems. Indeed, one of the successful regulatory mechanisms of religious systems appears to be its ability to shelter agents (i.e., adherents) from the functioning and goals of the system itself. Accurate insight into the workings of religious systems might actually be a destabilizing force; mythic narratives and supernatural rewards and punishments, reinforced through ritual routines, tend to be better motivators than mundane incentives.

Decentralization Complex adaptive systems typically lack central control. However, this does not mean that all constituent entities play an equal role in the emergence and functioning of a complex adaptive system. Indeed, as noted above, complex adaptive systems are often hierarchically organized and that seems to be the case for most religious systems. Importantly, complex adaptive systems that exhibit hierarchical organization maintain their structure despite the constant flow of agents through the system. Cells within the human body are ephemeral, firms succeed and fail within economic markets, and popes are chosen, and ultimately either resign or pass away, while the hierarchical structure of these respective systems endures.

Regulatory Mechanisms Complex adaptive systems require regulatory mechanisms that enable them to respond adaptively to changing environmental conditions. Our discussion of the hierarchy of religious discourse, above, describes one of the regulatory mechanisms that enable religious systems to adapt to changing environmental systems. Moreover, the nature of the feedback loops in religious systems is such that when a system is not producing cooperative and coordinated behavior, or group members experience negative health or reproductive impacts, this information effects the potential performance of ritual behaviors. Accordingly, the system itself will adjust by possibly revising rules, weaving new motivational myths, or even imagining novel supernatural worlds; otherwise the system will die.

Open Complex adaptive systems are open systems; that is, they have fuzzy (i.e., not well-defined) and porous (i.e., easily crossed by agents) boundaries. Religious groups, of course, vary in how open or closed they are (Wilson et al. 2017). Some religions proselytize and seek outside members, such as most forms of Christianity, whereas other religions, such as Judaism, discourage new members from joining. Further, individual mobility varies significantly across religious landscapes. For example, in the USA, where religions are experienced as free market commodities, there are much higher levels of denominational switching than in European countries where state-supported religious monopolies exist (Putnam and Campbell 2010). There are also considerable differences between world and indigenous religions in their openness; major world religions tend to be much more open. In many indigenous communities, the only way to join the religious community is to participate in the local initiation ritual; there are no anonymous members in distant lands (Whitehouse 2004).

Amplification of Random Fluctuations Positive feedback loops direct complex adaptive systems toward divergent evolutionary pathways; in other words, small random changes that are not necessarily adaptive responses to environmental conditions can result in substantial differences across systems. Because interactions among agents are nonlinear, as discussed above, small fluctuations in input can have considerable impacts on output. This of course has significant consequences for understanding religious systems. For example, Jewish communities that I have visited in North America, South America, Europe, Africa, Middle East, and India exhibit extraordinary religious diversity. Their differences in food preferences, dress, greetings, language, styles of prayer, and so forth are a product of the local sociocultural environment. But how and why communities initially settled in a particular area is often the result of a fortuitous decision of a religious leader, a ship simply landing where the winds took it, myths of golden streets, and countless other random factors that are not adaptive responses to environmental conditions. Yet these factors result in astonishing differences between communities over time. Or consider religious holidays that mark a historical event. Jews, for instance, observe four minor fast days to commemorate tragedies in their history, or in one case a tragedy narrowly averted, but the dates themselves are the incidental consequence of history. Yet, once the fast days were canonized as part of the Jewish calendar, the shape and rhythm of the Jewish year and the lived experiences of Jews were altered. The amplification of random fluctuations suggests that researchers must be cautious when analyzing religious systems not to over-interpret their adaptive nature; adaptive responses are often built on random fluctuations.

History One implication of the amplification of random fluctuations for complex adaptive systems is that historical contingency is always partially responsible for present behavior. This means that historical analyses will play an important role in understanding any complex adaptive system, and religious systems are no exception. Indeed, it would be impossible to understand why any religion takes the form it does without understanding the historical factors that shaped the religion accordingly.

Unpredictability of Agents It is generally difficult to predict the behavior of specific agents in a complex adaptive system. Even when group-level behavioral patterns appear to stabilize, it is very difficult, if not impossible, to predict individuals' lives. For example, we can confidently anticipate a gathering of parishioners at a local church on Sunday morning, but even regular worshippers will fail to attend occasionally due to illness, travel, or mood, factors we are unlikely to foresee.

Disequilibrium One of the most interesting and significant features of complex adaptive systems is that they do not operate at equilibria conditions. This does not mean that they do not experience moments of stasis, but their responsiveness to changing environments keeps complex adaptive systems nearly in constant flux. The reason that this is so important is because it implies that religious systems are continuously evolving. This will strike some as extraordinary and even outlandish. But as discussed above, while religious systems give the impression to adherents and outsiders that they are eternal and stable, religious systems are in fact constantly adjusting to local socioecological conditions. As Bulbulia (2009) has astutely noted, a historical glance at any religion will reveal both surprising stability (e.g., the Lord's Prayer) and extensive change (e.g., the language in which it is recited).

3.1 Four Additional Features of Complex Adaptive Systems

One of the world's leading complexity theorists, computer scientist John Holland (1992), offers four additional features of complex adaptive systems. First, agents interact by sending signals and they interact simultaneously, which he terms parallelism. Second, agents are characterized by conditional action in that they respond to signals as if-then statements. Third, the regulatory mechanisms of complex adaptive systems are modular in the sense that groups of rules combine to form what we might think of as subroutines. These subroutines enable complex adaptive systems to deal with novel conditions. Fourth, not only do complex adaptive systems themselves adapt and evolve, but agents and constituent elements within these systems adapt and evolve as well.

There are several points to emphasize when considering how these features relate to religious systems. First, signaling theory offers a robust collection of models that have been rigorously applied to religion by economists, sociologists, biologists, psychologists, and anthropologists (see Bulbulia and Sosis 2011). However, this work has not yet recognized that religious signals lie at the core of the communication structure that enables religious systems to operate. Second, what Holland in the language of computer science refers to as subroutines is particularly important for understanding religious systems. Much of what cognitive and evolutionary researchers have been uncovering concerning the relationships between the religious system's core building blocks, such as Modes Theory and Supernatural Punishment Theory, are subroutines in Holland's terms. Moreover, the linkage of rules also occurs at the level of agent-agent interactions because religious rules are regularly tied together. In religious communities, there are expectations that if an individual follows one rule, there are other rules that are assumed to be obeyed as well. The Muslim who habitually prays five times each day but enjoys a ham sandwich after evening prayers will be eyed with suspicion, and notably, the prayers themselves, even if heartfelt, will be unlikely to serve as effective signals of one's commitment to the community. Third, not only do religious systems adapt, but the individuals who enliven such systems change and the core elements of religious systems change as well. It is indeed likely that the building blocks of religious systems have been transformed over time, lending support to the warning that anthropologists have long voiced about not generalizing across categories of religions, such as tribal, chiefdom, and contemporary world religions (Evans-Pritchard 1965). For instance, religious meaning differs considerably across these categories, and as Whitehouse (2004) observes, it is likely a function of variation in the frequency of ritual performance.

4 Discussion

Religious systems are clearly organic. The coalescence of various cognitive, neurological, behavioral, affective, and developmental aspects of humanity has resulted in self-sustaining and cross-culturally recurring systems in which individuals, interacting through signaling mechanisms, bring life to these systems. One of the most important lessons of the complex adaptive approach to religion, and vital messages of this chapter, is that altering one part of a complex adaptive system, such as a religious systems, has significant effects on other parts. Those effects, notably, are difficult to predict even for those most familiar with religious systems, including scholars and religious leaders. For example, sociologists Rodney Stark and Roger Finke (2000) have argued that when the Second Vatican Council in 1962 repealed many of the Catholic Church's prohibitions and reduced the level of strictness in the church, it had unforeseeable consequences. The Vatican Council was an attempt to regain the commitments of wavering Catholics, but it inadvertently initiated a decline in church attendance among American Catholics and reduced the overall enrollments in seminaries. In the late 1950s almost 75% of American Catholics were attending Mass weekly, but since the Vatican's actions, there has been a steady decline to the current rate below 35% (D'Antonio et al. 2007; Hadaway and Marler 2005). A similar reduction in commitment is associated with the purging of ritual obligations in Reform Judaism as well (Iannaccone 1994; Lazerwitz and Harrison 1979). Though many other instances could be referenced, what these two examples illustrate is that religions grow organically, and thus naïve tampering with them can result in unexpected changes, even stunted growth or collapse.

Most contemporary discussions of religion involve concerns about the rising, and what is perceived to be uncontrollable, geopolitical influence of religion today. The merit of any theory lies in its ability to explain current trends and by that measure the complex adaptive systems approach fares comparatively well. Understanding religions as complex adaptive systems underscores how external pressures that aim to change religions can sometimes result in dangerous consequences. For instance, religious radicalization, such as the emergence of the Muslim Brotherhood in Egypt, Turkey, and elsewhere, appears to have been a response to aggressive secular campaigns (Armstrong 2000; Ruthven 2004). Minimizing religious extremism in the future thus may require secularists to countenance religious traditions and design policies accordingly. This will not be easy, for even externally imposed changes that are intended to benefit religious communities can have long-term negative consequences. For example, on March 3, 1948, during a period of civil war prior to the Israeli War of Independence, Ben Gurion established a military exemption for yeshiva students. He presumably felt he was saving a cultural remnant of European Jewry that was otherwise headed toward extinction with the birth of the secular Israeli state (Efron 2003). As the yeshiva population has grown exponentially because of the extraordinary birth rates of Israeli Ultra-Orthodox Jews, not serving in the military has emerged as a costly signal of one's commitment to the community. For Jewish Israelis, not serving in the military is a stigma with consequences in the labor market. But this stigma serves as a gatekeeper within the religious community: one way of demonstrating one's commitment is staying in yeshiva not just until the possibility of being drafted has passed due to age, but even several years after one is eligible for the draft (Berman 2000). As a result, Yeshiva students and their families are exceedingly poor because under the terms of their military exemption they are permitted only minimal employment. Due to their failure to recognize that the military exemption has been transformed into a religious commitment signal, the government has attempted to alleviate the financial plight of these yeshiva students by increasing their subsidies, but this has only exacerbated the problem (Berman 2009). By increasing payments to yeshiva students, the government has increased the amount of time yeshiva students must remain in the yeshiva to serve as an effective signal of commitment. In short, the government subsidies have effectively decreased the costs of the signal.

Complexity theorists acknowledge that because of the nonlinear dynamics of complex adaptive systems, it will be a challenge to develop a general theory of complex adaptive systems that will generate reliable predictions. However, these theorists also argue that such a pursuit is not hopeless. They emphasize, for example, that all complex adaptive systems exhibit lever points (Holland 1995). These are points in the system where small changes can have predictable system-wide effects. One obvious lever point in religious systems is ritual performance. As Atkinson and Whitehouse (2011) have shown, changes in a ritual's frequency and intensity have predictable consequences for the hierarchical structure of the ritual community as well as cognitive impacts on individual perceptions of local versus universal communities. Of course, likewise, eliminating a ritual entirely can have profound effects on a community. Tuzin (1997), for example, ethnographically documents how social dynamics among the Ilahita Arapesh were transformed following the demise of the central ritual cult known as the Tambaran. The secret male cult was eliminated when a revivalist movement overtook the Ilahita villages where Tuzin was conducting fieldwork. Fathers lost the respect and control of their sons, which had previously been secured by the brutal initiation rites of the Tambaran. Wives, who were left out of the Tambaran, now had social power that enabled them to dominate their husbands. Beyond social relations, work itself also completely changed. Garden work, especially yam cultivation, lost its meaning for Ilahita men; what was once an activity motivated by the possibility of pleasing the spirits and contributing to Tambaran feasts, now became drudgery for these men.

While ritual performance is clearly a lever point in religious systems, the examples discussed throughout this chapter suggest that our ability to anticipate changes at this lever point is limited. Religions are complex adaptive systems, and our understanding of them is incomplete. Like other systems in nature, religions are dynamic, emergent, and difficult to predict. By recognizing religions as complex adaptive systems, it is hoped that students, scholars, and policy-makers appreciate that if there are compelling reasons to control or reform religions, we currently have limited understanding of how to do so, and naïve policies seeking change are likely to have unintended consequences. This does not mean that societies are simply at the mercy of religious exigencies, but it does suggest that effective engagements with religions will require an appreciation that they are dynamic complex systems, and all that this implies.

5 Conclusion

I began this chapter with a story, and I will end it with a story. While many people are aware of Yom Kippur, the Jewish day of atonement on which Jews fast (no food or water) for 25 hours, less well-known are the four minor fasts in the Jewish ritual cycle mentioned above: Tzom Gedaliah, Asarah b'Tevet, Shiva Asar b'Tammuz, and Ta'anit Esther. While conducting ethnographic fieldwork in Israel among Orthodox Jews, I asked a friend whether he observes these minor fasts, as I was aware that not all Orthodox Jews observe them. He responded that although these fasts did not hold much meaning for him (they commemorate various tragedies, or tragedies narrowly averted, in Jewish history), he did indeed observe the fasts. I asked him why and he responded that if he failed to do so the whole "thing" would unravel. I was slow in comprehending what he was saying so I asked him what he meant; what "thing" would unravel? His response: Judaism. He explained that Judaism was a complete package and that if you stopped doing one thing, then soon enough other religious obligations would be neglected and pretty soon, in his own words, everything would begin to unravel. I share this anecdote for my anthropological colleagues who would rightly be concerned that the complex adaptive system approach that I have outlined here fails to capture the experience of lived religion. That is a fair critique, although my goal here was to offer an explanatory model of religion rather than a phenomenological account. Nevertheless, as my friend described, many people see their religion as a system, and a complex one at that, even if they would not use the language of complexity studies to describe their religion. Although the complex adaptive system approach to religion is just in its infancy, one day it may offer an avenue for both etic and emic analyses of religion, fulfilling the elusive promise of a truly holistic anthropology.

Acknowledgments I thank the James Barnett Endowment for support and Michael Price, John Smart, and Connor Wood for helpful comments on an earlier draft of this chapter.

References

Alcorta, C.S. and Sosis, R. 2005. Ritual, emotion, and sacred symbols: The evolution of religion as an adaptive complex. *Human Nature* 16:323–359.

Alcorta, C.S. and Sosis, R. 2006. Why Ritual Works: A Rejection of the By-Product Hypothesis. *Behavioral and Brain Sciences* 29:613–614.

Armstrong, K. 1993. A History of God. New York: Ballantine Books.

Armstrong, K. 2000. *The Battle for God*. New York: Random House.

Atkinson, Q.D., & Whitehouse, H. 2011. The cultural morphospace of ritual form: Examining modes of religiosity cross-culturally. *Evolution and Human Behavior*, 32:50–62.

Barrett, J.L. and Lawson, E.T. 2001. Ritual intuitions: Cognitive contributions to judgments of ritual efficacy. *Journal of Cognition and Culture*, *1*(2), 183–201.

Berger, P. 1967. *The Sacred Canopy*. New York: Anchor Books.

Berman, E. 2000. Sect, Subsidy and Sacrifice: An Economist's View of Ultra-Orthodox Jews. *Quarterly Journal of Economics* 115:905–953.

Berman, E. 2009. *Radical, Religious and Violent*. Cambridge: MIT Press.

Billingsley, A. 1999. *Mighty like a river: The Black church and social reform*. Oxford: Oxford University Press.

Bloch, M. 1974. Symbols, song, dance and features of articulation: Is religion an extreme form of traditional authority? *European Journal of Sociology*, *15*, 54–81.

Bulbulia, J. 2009. Charismatic signalling. *Journal for the Study of Religion, Nature, Culture*, 3(4), 518–551.

Bulbulia, J. 2012. Spreading order: religion, cooperative niche construction, and risky coordination problems. *Biology and Philosophy*, *27*(1), 1–27.

Bulbulia, J. and R. Sosis 2011. Signaling theory and the evolution of religions. *Religion* 413:363–388.

Carroll, J. 2001. *Constantine's Sword*. Boston: Houghton Mifflin Co.

Cho, F., & Squier, R.K. 2008. "He Blinded Me with Science": Science Chauvinism in the Study of Religion. *Journal of the American Academy of Religion*, 76(2):420–448.

Cho, F., & Squier, R.K. 2013. Religion as a complex and dynamic system. *Journal of the American Academy of Religion*, 81:357–398.

Corning, P. 2002. The re-emergence of "emergence": a venerable concept in search of a theory. *Complexity* 7(6):18–30.

Cronk, L. 1994. Evolutionary Theories of Morality and the Manipulative Use of Signals. *Zygon* 29:81–101.

Czachesz, I. 2014. The Evolutionary Dynamics of Religious Systems Laying the Foundations of a Network Model. *Origin of Religion, Cognition and Culture, edited by A. Geertz and J. Sinding Jensen*, pp. 98–120.

D'Antonio, W.V., J.D. Davidson, D.R. Hoge, and M.L. Gautier 2007. *American Catholics Today: New Realities of their Faith and their Church*. New York: Rowman and Littlefield Publishers.

D'Aquili, E.G., Laughlin, C.D. and McManus, J. 1979. *The spectrum of ritual*. New York: Columbia University Press.

Deacon, T. 2010. *Incomplete Nature*. New York: Norton.

Dennett, D. 1991. *Consciousness Explained*. Boston: Little, Brown & Co.

Douglas, M. 1966. *Purity and danger: An analysis of concepts of pollution and taboo*. New York: Routledge.

Durkheim, E. 1912/1995. *The Elementary Forms of Religious Life*. New York: Free Press.

Efron, N.J. 2003. *Real Jews: Secular vs. Ultra-Orthodox and the Struggle for Jewish Identity in Israel.* New York: Basic Books.

Evans-Pritchard, E. 1965. *Theories of primitive religion*. London: Oxford University Press.

Finkel, D., P. Swartwout, and R. Sosis. 2010. The Socio-Religious Brain: A Developmental Model. Eds. R. Dunbar, C. Gamble, J. Gowlett. *Proceedings of the British Academy* 158: 283–307.

Geertz, C. 1967/1973. Religion as a cultural system. In *The Interpretation of Cultures,* pp. 87–125. New York: Basic Books.

Ginges, J., Atran, S., Medin, D., and Shikaki, D. 2007. Sacred bounds on rational resolution of violent conflict. *Proceedings of the National Academy of Sciences*, *104*, 7357–7360.

Hadaway, C.K. and P.L. Marler 2005. How Many Americans Attend Worship Each Week? An Alternative Approach to Measurement. *Journal for the Scientific Study of Religion*. 44:307–322.

Heimola, M. 2012. *Religious Rituals and Norms in the Making of Adaptive Systems*. Doctoral thesis, University of Helsinki.

Henrich, J. 2009. The evolution of costly displays, cooperation and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior*, *30*, 244–260.

Hill, K., and Hurtado, A.M. 1996. *Ache life history: The ecology and demography of a foraging people*. New York: Aldine de Gruyter.

Hogbin, I. 1970. *The island of menstruating men: Religion in Wogeo, New Guinea*. London: Chandler Publishing.

Holland, J.H. 1992. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. Cambridge: MIT University Press.

Holland, J.H. 1995. *Hidden Order: How Adaptation Builds Complexity.* New York: Addison-Wesley.

Holland, J.H. 1998. *Emergence: From Chaos to Order*. Oxford: Oxford University Press.

Holland, J.H. 2012. *Signals and boundaries: Building blocks for complex adaptive systems*. Cambridge: MIT Press.

Iannaccone, L.R. 1994. Why strict churches are strong. *American Journal of Sociology* 99:1180–1211.

Irons, W. 2001. Religion as a hard-to-fake sign of commitment. In R. Nesse (ed.), *Evolution and the capacity for commitment*, pp. 292–309. New York: Russell Sage Foundation.

Johnson, D.D.P. 2016. *God is Watching You: How the Fear of God Makes Us Human.* Oxford: Oxford University Press.

Kauffman, S. 1995. *At home in the universe: The search for the laws of self-organization and complexity*. Oxford: Oxford University Press.

Kaufmann, E. 2010. *Shall the religious inherit the earth? Demography and politics in the twenty-first century*. London: Profile Books.

Kiper, J. and Sosis, R. 2014. Moral intuitions and the religious system: an adaptationist account. *Philosophy, Theology, and Science* 1:172–199.

Kiper, J. and Sosis, R. 2016. The Roots of Intergroup Conflict and the Co-option of the Religious System: An Evolutionary Perspective on Religious Terrorism. In *Oxford Handbook of Evolutionary Perspectives on Religion*, eds. J. Liddle and T. Shackelford. Oxford: Oxford University Press.

Klass, M. 1995. *Ordered Universes: Approaches to the Anthropology of Religion*. Boulder: Westview Press.

Koenig, H., King, D., and Carson, V.B. 2012. *Handbook of religion and health*. Oxford: Oxford University Press.

Kugel, J. 2017. *The Great Shift*. New York: Houghton Mifflin Harcourt.

Lansing, J.S. 2003. Complex adaptive systems. *Annual review of anthropology*, *32*(1), 183–204.

Lazerwitz, B. and M. Harrison 1979. American Jewish Denominations: a social and religious profile. *American Sociological Review* 44:656–66.

Lesser, A. 1978. *The Pawnee ghost dance hand game: ghost dance revival and ethnic identity*. Madison: University of Wisconsin Press.

Levine, S.W. 2003. *Mystics, mavericks, and merrymakers: An intimate journey among Hasidic girls*. New York: NYU Press.

Lévi-Strauss, C. 1964/1994. *The raw and the cooked: Introduction to a science of mythology*. New York: Pimlico.

Lienard, P. and Boyer, P. 2006. Whence collective rituals? A cultural selection model of ritualized behavior. *American Anthropologist*, *108*(4), 814–827.

Luhrmann, T.M. 2012. *When God talks back: Understanding the American evangelical relationship with God*. New York: Vintage.

McCauley, R.N. and Lawson, E.T. 2002. *Bringing ritual to mind: Psychological foundations of cultural forms*. Cambridge: Cambridge University Press.

Mecklenburger, R. 2012. *Our Religious Brains*. Woodstock, VT: Jewish Lights Publishing.

Miller, J.H. and Page, S.E. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton: Princeton University Press.

Mitchell, M. 2009. *Complexity: A guided tour*. Oxford: Oxford University Press.

Mooney, J. 1965. *The ghost-dance religion and the Sioux outbreak of 1890*. Chicago: University of Chicago Press.

Myerhoff, B.G. 1974. *Peyote hunt: The sacred journey of the Huichol Indians*. Ithaca: Cornell Univ. Press.

Power, E.A. 2017. Discerning devotion: Testing the signaling theory of religion. *Evolution and Human Behavior*, 38:82–91.

Purzycki, B.G., Finkel, D.N., Shaver, J., Wales, N., Cohen, A.B., & Sosis, R. 2012. What does God know? Supernatural agents' access to socially strategic and non-strategic information. *Cognitive Science*, 36:846–869.

Purzycki, B.G., Haque, O., and Sosis, R. 2014. Extending Evolutionary Accounts of Religion beyond the Mind: Religions as Adaptive Systems. In *Evolution, religion, and cognitive science: Critical and constructive essays*, eds. F. Watts and L. Turner, pp. 74–91. New York: Oxford University Press.

Purzycki, B.G. and Sosis, R. 2009. The Religious System as Adaptive: Cognitive Flexibility, Public Displays, and Acceptance, in E. Voland and W. Schiefenhövel (eds), *The Biological Evolution of Religious Mind and Behavior*, pp. 243–56. New York: Springer Publishers.

Purzycki, B.G. and Sosis, R. 2010. Religious Concepts as Necessary Components of the Adaptive Religious System, in U. Frey (ed), *Interdisciplinary Perspectives on Philosophy: Evolution and Religion*, pp. 37–59. Marburg, German: Tectum Verlag.

Purzycki, B.G. and Sosis, R. 2011. Our Gods: Variation in Supernatural Minds, in *Essential Building Blocks of Human Nature*, ed. Ulrich Frey, C. Stormer, K.P. Willfuhr, pp. 77–93, New York: Springer.

Purzycki, B.G. and Sosis, R. 2013. The extended religious phenotype and the adaptive coupling of ritual and belief. *Israel Journal of Ecology & Evolution*, *59*(2): 99–108.

Purzycki, B.G. and Willard, A. 2016. MCI Theory: A Critical Discussion. *Religion, Brain & Behavior,* 6(3): 207–248.

Putnam, R.D. and Campbell, D.E. 2010. *American grace: How religion divides and unites us*. New York: Simon and Schuster.

Radin, P. 1956/1972. *The trickster: a study in American Indian mythology*. New York: Schocken.

Rappaport, R.A. 1979. The obvious aspects of ritual. In *Ecology, Meaning, and Religion*, pp. 173–221. Richmond, CA: North Atlantic Books.

Rappaport, R.A. 1999. *Ritual and Religion in the Making of Humanity*. London: Cambridge Univ. Press.

Ruthven, M. 2004. *Fundamentalism: The Search for Meaning*. Oxford: Oxford University Press.

Schloss, J.P. and Murray, M.J. 2011. Evolutionary accounts of belief in supernatural punishment: A critical review. *Religion, Brain & Behavior*, *1*(1):46–99.

Seeman, D. 2010. *One people, one blood: Ethiopian-Israelis and the return to Judaism*. New Brunswick: Rutgers University Press.

Seligman, A.B. and Weller, R.P. 2012. *Rethinking pluralism: Ritual, experience, and ambiguity*. Oxford University Press.

Shaver, J. 2015. The evolution of stratification in Fijian ritual participation. *Religion, Brain & Behavior* 5(2):101–117.

Shaver, J. 2017. Why and how do religious individuals, and some religious groups, achieve higher relative fertility? *Religion, Brain & Behavior*, *7*(4), 324–327.

Shaver, J., Divietro, S., Lang, M., and Sosis, R. 2018. *Costs do not explain trust among secular groups. Journal of Cognition and Culture* 18: 180–204.

Shaver, J., B. Purzycki, and R. Sosis 2016. Evolutionary theory and the study of religion. In *The Oxford Handbook of the Study of Religion*, eds. M. Stausberg and S. Engler, pp. 124–136. Oxford: Oxford University Press.

Smith, W.C. 1998. *Believing: An historical perspective*. Oxford: Oneworld.

Smith, W.J. 1979. Ritual and the Ethology of Communicating. *The Spectrum of Ritual*, New York: Columbia University Press, pp., 51–79.

Sosis, R. 2003. Why aren't we all Hutterites? Costly signaling theory and religion. *Human Nature* 14:91–127.

Sosis, R. 2005. *Ifaluk Atoll: An Ethnographic Account,* in eHRAF World Cultures, ed. Carol Ember. New Haven, CT: HRAF.

Sosis, R. 2007. Psalms for safety: Magico-religious responses to threats of terror. *Current Anthropology*, 48, 903–911.

Sosis, R. 2009. The adaptationist-byproduct debate on the evolution of religion: Five misunderstandings of the adaptationist program. *Journal of Cognition and Culture.* 9:315–332.

Sosis, R. 2016. Religions as complex adaptive systems. In *Mental Religion: The Brain, Cognition, and Culture*, ed. N. Clements, pp. 219–236. Farmington Hills, MI: Macmillan.

Sosis, R., and Alcorta, C.S. 2003. Signaling, solidarity and the sacred: The evolution of religious behavior. *Evolutionary Anthropology* 12:264–274.

Sosis, R. and Alcorta, C.S. 2004. Is Religion Adaptive? *Behavioral and Brain Sciences* 27:749–750.

Sosis, R. and Bressler, E. 2003. Cooperation and commune longevity: A test of the costly signaling theory of religion. *Cross-Cultural Research* 37:211–239.

Sosis, R. and Handwerker, P. 2011. Psalms and coping with uncertainty: Israeli women's responses to the 2006 Lebanon War. *American Anthropologist*, 113, 40–55.

Sosis, R. and Kiper, J. 2014a. Religion is more than belief: What evolutionary theories of religion tell us about religious commitment. *Challenges to religion and morality: Disagreements and evolution*, eds., M. Bergman & P. Kain, pp. 256–276. New York: Oxford University Press.

Sosis, R. and Kiper, J. 2014b. Why religion is better conceived as a complex system than a norm-enforcing institution. *Behavioral and Brain Sciences* 37: 275–276.

Sosis, R. and Kiper, J. 2017. Sacred versus Secular Values: Cognitive and Evolutionary Sciences of Religion and Religious Freedom. In *Homo Religiosus? Exploring the Roots of Religion and Religious Freedom in Human Experience*, eds. T.S. Shah and J. Friedman, pp. 89–119. Cambridge: Cambridge University Press.

Sosis, R., Phillips, E., and Alcorta, C.S. 2012. Sacrifice and Sacred Values: Evolutionary Perspectives On Religious Terrorism, in T. Shackelford and V. Weekes-Shackelford (eds), *The Oxford Handbook of Evolutionary Perspectives on Violence, Homicide, and War*, pp. 233–53. NY: Oxford University Press.

Sosis, R. and J. Shaver 2015. How rituals elicit shared sacred values. *Interdisziplinäre Anthropologie* 3:75–81.

Spiro, M.E. 1952. Ghosts, Ifaluk, and teleological functionalism. *American Anthropologist*, 54:497–503.

Stark, R. and R. Finke 2000. *Acts of Faith: Explaining the Human Side of Religion*. Berkeley: University of California Press.

Tetlock, P.E. 2003. Thinking the unthinkable: Sacred values and taboo cognitions. *Trends in Cognitive Sciences, 7,* 320–324.

Tillich, P. 1957. *Dynamics of Faith*. New York: Harper and Row.

Trimingham, J.S. 1971. *The Sufi Orders in Islam*. London: Oxford University Press.

Turner, V. 1969. *The Ritual Process: Structure and Anti-Structure*. New York: Routledge.

Tuzin, D. 1997. *The cassowary's revenge: the life and death of masculinity in a New Guinea society*. Chicago: University of Chicago Press.

Von Bertalanffy, L. 1972. The History and Status of General Systems Theory. *The Academy of Management Journal*, 15: 407–426

Wallace, A.F.C. 1966. *Religion: An Anthropological View*. New York: Random House.

Watts, J., Sheehan, O., Atkinson, Q.D., Bulbulia, J., & Gray, R.D. 2017. Ritual human sacrifice promoted and sustained the evolution of stratified societies. *Nature*, *532*(7598), 228–231.

Weber, M. 1947. *The Theory of Social and Economic Organization*. Oxford: Oxford University Press.

Whitehouse, H. 1995. *Inside the cult: Religious innovation and transmission in Papua New Guinea*. Oxford University Press.

Whitehouse, H. 2004. *Modes of Religiosity: A Cognitive Theory of Religious Transmission*. New York: Alta Mira Press.

Wilson, D.S., Hartberg, Y., MacDonald, I., Lanman, J.A., & Whitehouse, H. 2017. The nature of religious diversity: a cultural ecosystem approach. *Religion, Brain & Behavior*, *7*(2), 134–153.

Worsley, P. 1957. *The trumpet shall sound*. London: MacGibbon & Kee.

Index

G. Y. Georgiev et al. (eds.), *Evolution, Development and Complexity*,
Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-00075-2