The cover features a sepia-toned portrait of Charles Darwin, an elderly man with a full white beard and a dark bowler hat. The portrait is centered and occupies the middle ground. Overlaid on the entire image is a complex circuit board pattern. The pattern consists of numerous thin, light blue lines that trace intricate paths across the cover, interspersed with small, glowing yellow and blue dots that represent electronic components or data points. The background of the portrait shows some bare, tangled branches on the left side.

# SFI Bulletin

2014, vol. 28

How Life  
Got Complex



## About the Santa Fe Institute

Founded in 1984, the Santa Fe Institute is an independent, nonprofit research and education center that has pioneered the science of complex systems. Its missions are supported by philanthropic individuals and foundations, forward-thinking partner companies, and government science agencies.

The *SFI Bulletin* is published periodically to keep SFI's community informed of its scientific work. The *Bulletin* is available online, either as a web-based (html5) tablet publication or as a pdf, at [www.santafe.edu/bulletin](http://www.santafe.edu/bulletin).

To receive email notifications of future online issues, please subscribe at [www.santafe.edu/bulletin](http://www.santafe.edu/bulletin). To request a subscription to the printed edition, please email the editor at [jdg@santafe.edu](mailto:jdg@santafe.edu).

## Editorial Staff

SFI Chair of the Faculty: **Jennifer Dunne**  
SFI Director of Communications: **John German**  
Art Director: **Paula Eastwood, Eastwood Design**  
Design & Function: **8 Arms Creative**  
Printing: **Starline Printing**

## Published by the Santa Fe Institute

1399 Hyde Park Road  
Santa Fe, New Mexico 87501, USA  
Phone 505.984.8800

## Printed On:

100% recycled fiber  
100% post-consumer waste  
Green-e certified  
Low-VOC vegetable-based ink

[www.santafe.edu](http://www.santafe.edu)

COVER: CHARLES DARWIN, ADOC-PHOTOS/ART RESOURCE, NEW YORK; CIRCUIT, SHUTTERSTOCK.COM; COMPOSITE, PAULA EASTWOOD

## Perspectives

# The Complexity of Life

BY **DAVID KRAKAUER**

DIRECTOR, WISCONSIN INSTITUTE  
FOR DISCOVERY, UNIVERSITY  
OF WISCONSIN-MADISON

CO-DIRECTOR, CENTER FOR COMPLEXITY  
& COLLECTIVE COMPUTATION,  
WISCONSIN INSTITUTE FOR DISCOVERY,  
UNIVERSITY OF WISCONSIN-MADISON

EXTERNAL PROFESSOR, SANTA FE INSTITUTE

*“Verily at the first Chaos came to be, but next wide-bosomed Earth, the ever-sure foundations of all the deathless ones who hold the peaks of snowy Olympus, and dim Tartarus in the depth of the wide-pathed Earth, and Eros, fairest among the deathless gods, who unnerves the limbs and overcomes the mind and wise counsels of all gods and all men within them. From Chaos came forth Erebus and black Night; but of Night were born Aether and Day, whom she conceived and bare from union in love with Erebus. And Earth first bare starry Heaven, equal to herself, to cover her on every side, and to be an ever-sure abiding-place for the blessed gods.”* — Hesiod, from the *Theogony*, Part 2, translated by H.G. Evelyn-White

**T**HE CLASSICAL UNIVERSE is made up brick by brick, starting in the void and culminating with the earth: from the emptiness of night that gave rise to day, to the day that produces the outward order of the heavens, and finally to life upon the ground. The *Theogony* is a story describing the origins of energy and matter and information in the form of life. The *Theogony* exemplifies humanity's great surprise that the universe should have emerged from chaos, that emptiness has not reigned eternal, and that the earth should be hospitable and supportive of multiform sentience.

After almost three millennia our concerns are essentially the same as those of this celebrated Greek farmer and poet. In just less than 14 billion years, the universe has generated, from nothing, more than 100 billion galaxies, each of which contains on average 100 billion stars, and around many of these stars a system of planets. In our own Milky Way, tucked away in a local bubble of the Orion-Cygnus arm of the galaxy, 27,000 light years from the galactic center, spins our solar system, home to eight planets – four small and dense, and four large and gaseous. On one of these planets, the third nearest the sun, we find life. To the best of our knowledge, it is the only planet in our solar system supporting adaptive matter.

From physical law we can derive essential properties of the sun, the elements, and the planets. The incredible machinery of the theories of gravity, quantum mechanics, and the standard model

---

**“Inorganic chemistry is essentially silent on the topic of biology. We do not exist. The theory of everything is a theory of everything except of those things that theorize.”**

---

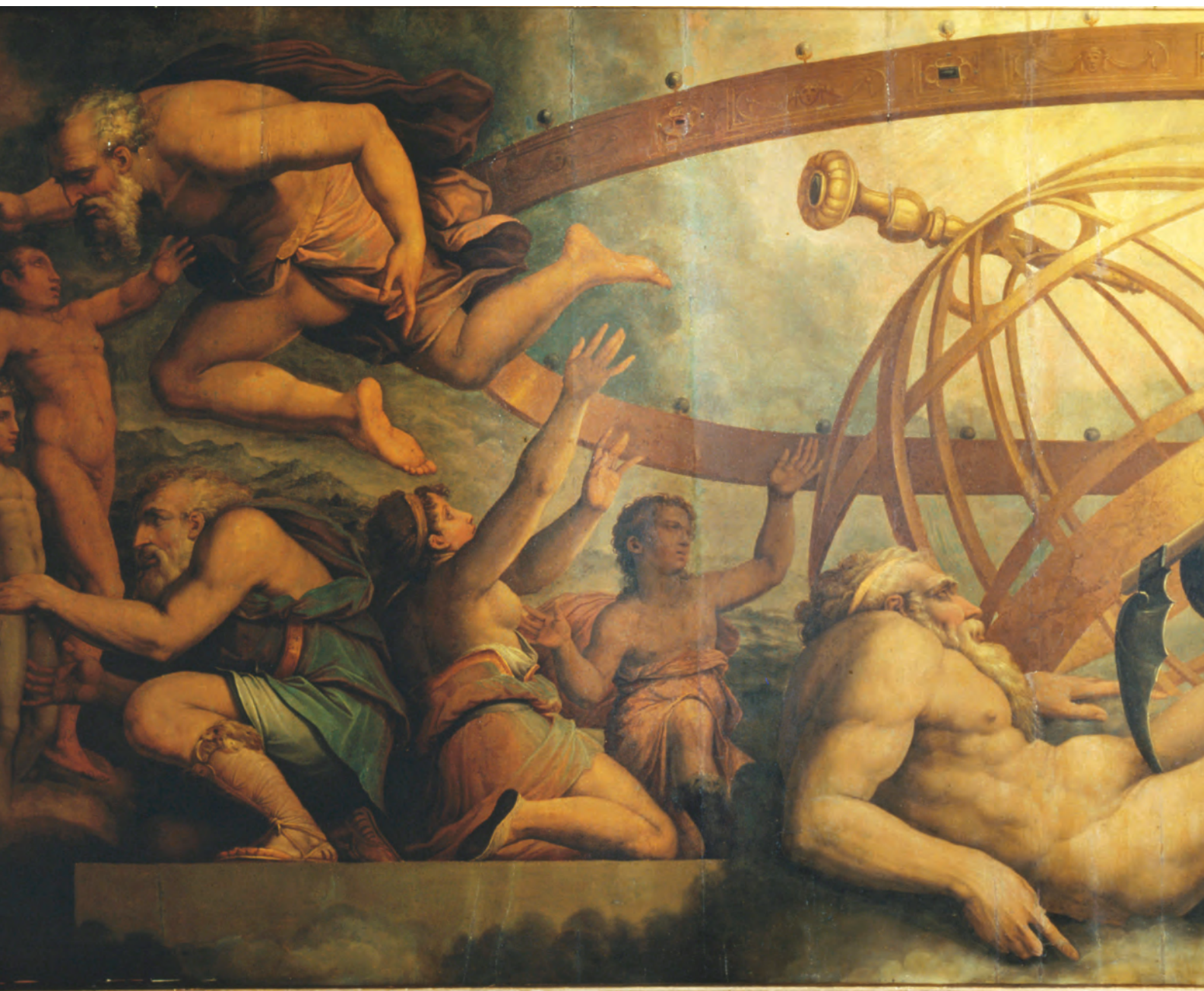


David Krakauer

give us significant insights into the observable structure in the universe. Optimistically, we can even deduce simple molecules from inorganic chemistry. And then the theory machine stops. Physics runs out of gas. Chemistry dries up. From the perspective of physics, our own solar system or galaxy are not in any way different from those anywhere else in the universe. Inorganic chemistry is essentially silent on the topic of biology. We do not exist. The theory of everything is a theory of everything except of those things that theorize.





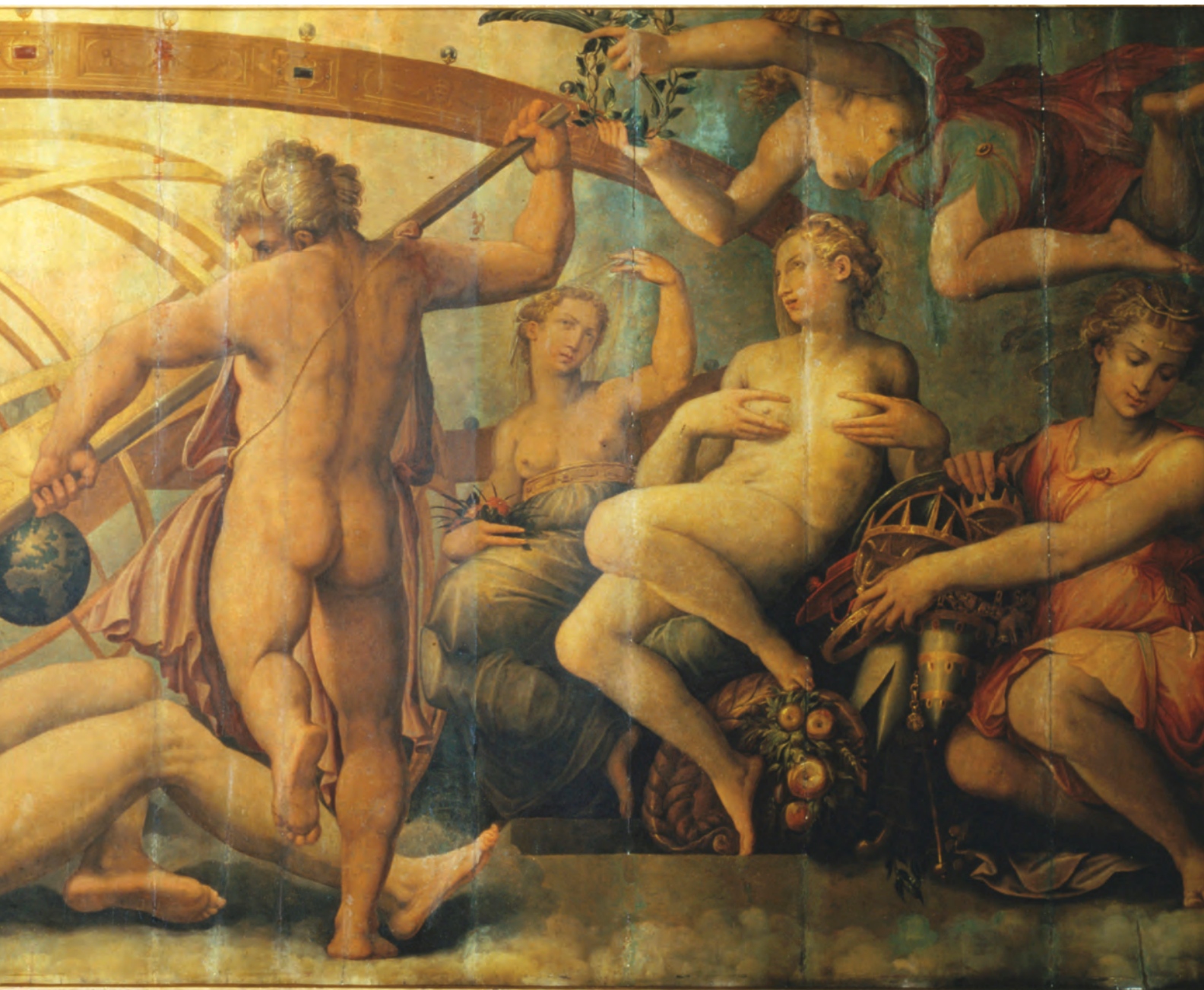


The projects described in this issue of the *Bulletin* are all efforts to grapple with some of the key ideas and concepts required to understand living systems, with an emphasis on evolution, both biological and cultural. We ask how, over long stretches of time, successively more effective mechanisms for storing and processing information have been adaptively engineered, and how these biological computing systems are used to predict, model, and control relevant states of noisy and living environments. We consider complexity through the lens of information processing, we seek to quantify how information is encoded in living systems (genomes and brains), and we suggest estimates for upper bounds in adaptive information

capacity. Some of the concepts relevant to understanding biological complexity include hierarchy, individuality, criticality, information/uncertainty, computation, and sociality.

Stanislaw Lem in his science fiction *chef-d'œuvre* *Solaris* (1961) considers a planet swaddled by an inscrutable ocean capable of astonishing acts of reasoning. So vastly more intelligent and powerful is the Ocean to the human explorers and scientists (Solarists) who dedicate their lives to its analysis and explication that humanity is forced to resign itself to ignorance over its ultimate mechanisms and motives. I have often wondered whether *Solaris* is not a metaphor for life on earth, where the methods of the Solarists are the traditional





**Cronus mutilates his father, Uranus, at the behest of his mother, Gaia, in this scene from Greek mythology featured in Hesiod's *Theogony*.**

methods of science. Solaris is a huge interconnected system of energy and information flows that defy traditional methods of reduction. Perhaps Solaris is waiting on complexity science: information theory, scaling, network theory, evolutionary dynamics, and computation. Perhaps we only stand a chance of understanding complex life when we approach it through the sciences of complexity – in which case, consider this issue of the *Bulletin* a temporary visa granting access to all of those restless explorers intent on the understanding of our terrestrial Solaris. ■



# Biological Systems, Energy, and Information

**R**ESearch on the origins, development, and dynamics of complexity in biological systems has been a core topic of inquiry at the Santa Fe Institute since its founding 30 years ago in 1984. SFI's scientists have worked to develop an understanding of a dizzying array of biological phenomena, from the origins of life, to transitions from single- to multi-cellularity, to evolutionary innovation at different levels of biological organization, to the relationship between ecological complexity and dynamical stability.

The intertwined concepts of energy and information are fundamental to any understanding of biological complexity. Biological systems are far from equilibrium, requiring a constant flow of energy to maintain their organization and functionality. The processing and encoding of information provides a means for life to manage and maintain energy acquisition, use, and dissipation.

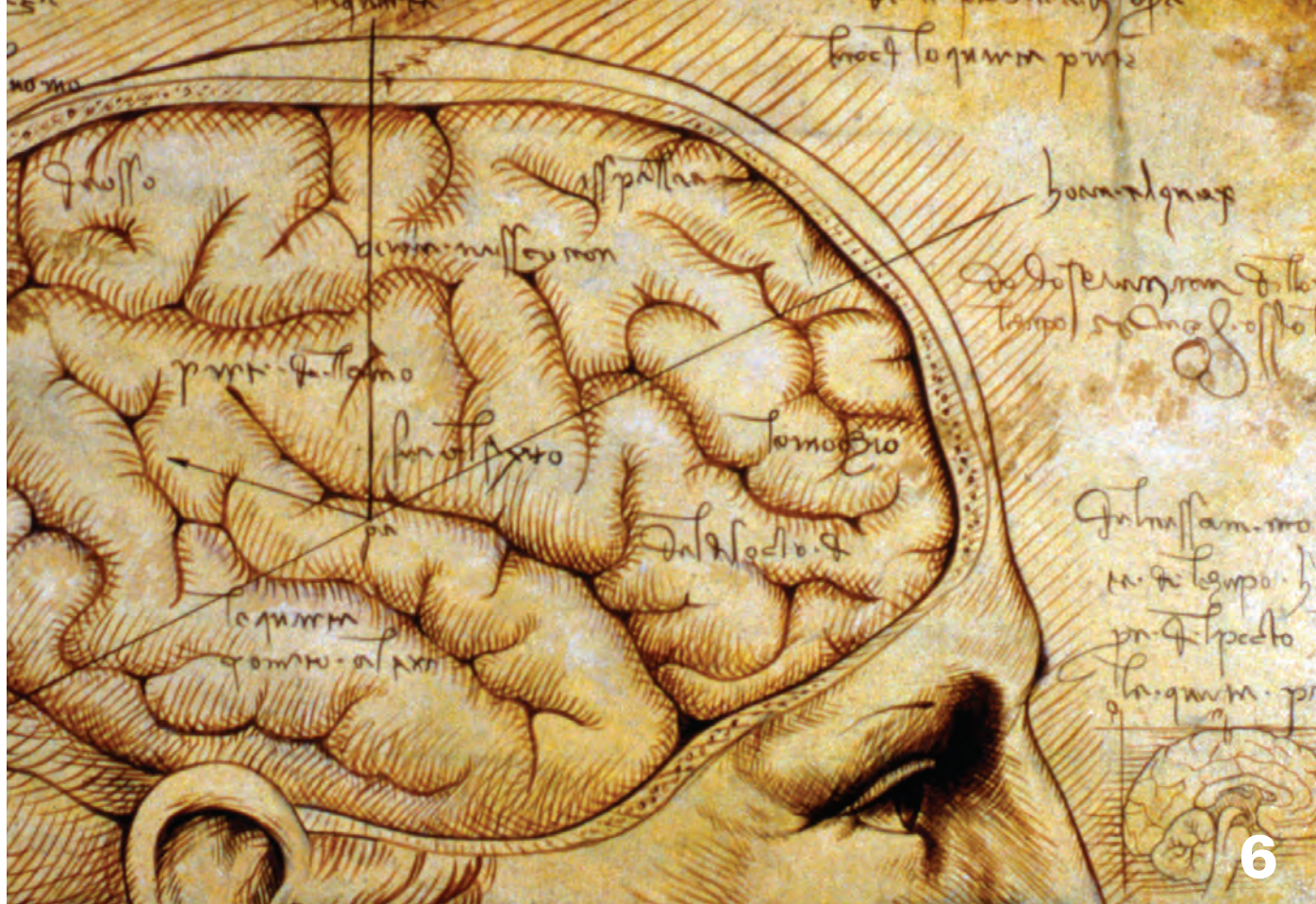
The work of former resident professors and current external faculty members Jessica Flack and David Krakauer (and their many colleagues) highlighted in this issue of the *SFI Bulletin* addresses the intimate dance between energy and information processing in biological systems – a dance they suggest gives rise to the complex, multi-scale structure we observe.

This computational-thermodynamic view of biology provides a powerful, potentially generic framework for understanding the properties of any kind of complex adaptive system, including socio-economic systems.

Sincerely,

Jennifer Dunne, Chair of the Faculty, Santa Fe Institute





# Contents

**SFI BULLETIN**  
VOL. 28 APRIL 2014

“There is nothing that living things do that cannot be understood from the point of view that they are made of atoms acting according to the laws of physics.”

—RICHARD FEYNMAN



## 0. Perspectives: The Complexity of Life

By DAVID KRAKAUER

## 4. Biological Systems, Energy, and Information

By JENNIFER DUNNE

## 6. Why Nature Went to the Trouble of Creating...You

By NATHAN COLLINS

## 12. Life's Information Hierarchy

By JESSICA FLACK

IMAGE CREDITS, THIS PAGE: FROM TOP: BRAIN DRAWING, HESSDESIGNWORKS.COM; EARTH FROM SPACE, NASA; ROUSSEAU, THE DREAM, WIKIMEDIA COMMONS. FACING PAGE: PORTRAIT OF JENNIFER DUNNE, MINESH BACRANIA PHOTOGRAPHY







# Why Nature Went to the Trouble of Creating... You

BY NATHAN COLLINS

In a complex world where plants and animals and everything else are duking it out to survive, an organism stands to gain from becoming more complex.

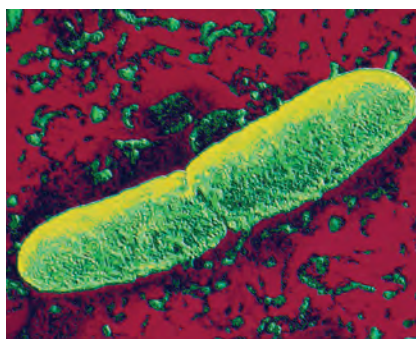
Life on earth began some 3.5 billion years ago, not all that long after the planet itself first formed, and for 1.5 billion years it chugged along, single-celled creatures self-replicating, dividing, diversifying. Remarkably, it took that long – 7,500 times longer than all of human history – for the first multicellular life to emerge, and still longer for it to evolve into life as we know it.

Despite that, the question many researchers ask isn't what took so long, but rather why complex life would have evolved in the first place. Consider this: single-celled organisms make up more than half of the biomass on earth, and even one of the tiniest organisms – *Y. pestis*, better known as bubonic plague – can effortlessly, thoughtlessly kill you.

Nature, it seems, doesn't need you. Indeed, there isn't any obvious reason it would go to the trouble of creating something as complex as a human being, complete with its differentiated organs and top-down control systems. And yet, despite four billion years of Nature's great "meh," here you are, alive, multicellular, complex – even intelligent enough to ponder your own existence.

What was Nature thinking? According to one argument, bacteria first bound together in colonies that enhanced cooperation and hence survival. Eventually, those bacteria bound together physically as well, creating the first multicellular life. And so on.

"You could say that's an answer, but then you could go a bit further and ask what is it exactly that makes you a better competitor," says David Krakauer, who with SFI



*Y. pestis*, the bacterium that causes bubonic plague (a.k.a. The Black Death of the Middle Ages)

External Professor Jessica Flack co-directs the Wisconsin Institute for Discovery's Center for Complexity and Collective Computation, or C4, and SFI's John Templeton Foundation-funded "Evolution of Complexity and Intelligence on Earth" research project.

"Well, you're outsmarting everyone else," Krakauer says.

## Simple versus complex

Despite its seeming indifference, Nature does seem to have thought highly enough

of complex structures to produce a few of them, and to have ratcheted up the complexity further by embedding complex structures within complex structures – animals with hearts and lungs and circulatory systems, or groups of people capable of building their own social institutions.

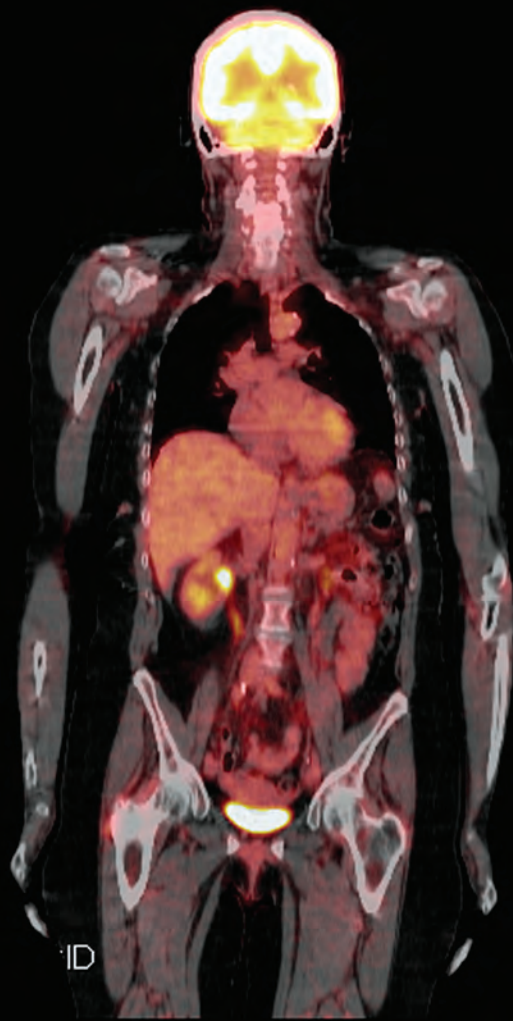
But why? What purpose does it serve? "Why life is hierarchically organized is not at all obvious," Flack says, and how an organism's or a society's complexity relates to the complexity of its environment remains unclear.

Our anthropomorphized Nature might have started with one very simple idea, what Krakauer calls the reflection principle, which presupposes that living things can't be more complex than their environments, an idea rooted in experiments. "If you take organisms and you place them in simpler environments, they just throw everything [superfluous] away. They lose genes," Krakauer says.

At the same time, the world does seem to favor an intelligent creature. Even the tiniest living things need to be able to comprehend, predict, and react to their environments; that's what allows them to outsmart each other, he says. In a complex world where plants and animals and everything else are duking it out to survive, an organism stands to gain from becoming more complex.

That tension between simplicity and complexity is the starting point for C4 postdoctoral fellow Christopher Ellison.





What does Nature care about hearts, brains, and other organs, or, for that matter, political parties – in other words, structures within structures?

“We’d like to understand the implications this has for the environment,” he says. “For example, do simple or complex organisms experience and live in simple or complex environments?”

Working with Flack and Krakauer, Ellison developed “Markov organisms,” computer-simulated creatures that merge

organisms themselves, he focuses on how information flows in the ecological system.

It’s early days, Ellison says, but his simulations suggest that life will often evolve to match its environment’s complexity – findings that are in line with the reflection principle, but with some interesting caveats. For one thing, evolving Markov organisms tend to

overshoot their worlds’ complexity and might take a long time to prune unnecessary complexities.

They’re also susceptible to “basis mismatch,” a problem you know well

if you’ve ever tried to explain to a tourist how to get around in your hometown. To you there are just a few steps, but to the novice it’s a complex process with many twists and

turns, and every intersection represents a possible misdirection; in sprawling cities like Los Angeles, a direction as simple as “take Sunset to Vermont and turn left” becomes infinitely complex. Markov organisms are the same: if their way of solving a problem doesn’t line up with how their environments constructed it, Ellison’s simulations show, Markov organisms’ complexity keeps evolving upward forever.

But Ellison’s information-centric approach has some benefits. One, he says, “is that it attempts to answer the question of how complexity evolves in an organism-independent fashion,” meaning that the ideas apply equally well to anything from bacteria to politics. Similarly, Ellison’s method allows the team to describe both an organism and its environment’s complexity in the same terms, because both derive from the same underlying models of information processing. Surprisingly, that’s something few, if any, other researchers have done.

### Constructing predictability

So it appears Nature might favor multicellular life if it affords a certain computational power not readily available to single-celled organisms. But what does Nature care about hearts, brains, and other organs, or, for that matter, political parties – in other words, structures within structures?

The answer, Flack says, is that living things like their worlds to be predictable, and what makes cells and people more likely to survive, Nature favors.

Much of the structure we observe in the world, Flack says, probably evolved because structure begets stability, hence predictability. Groups of genes, cells, or animals change their collective behaviors slowly compared with individual genes, cells, or animals, giving the faster-moving individual components a chance to anticipate changes more easily.

Biologists call the idea that plants and animals – and genes and organs and so on – structure their environments to be more stable and predictable “niche construction,” and it usually applies to physical structure like ants building nests. But it also can be applied to temporal structure.

Politics offers, perhaps, a simple example. Early U.S. politicians were explicit about designing Congress and the rest of government so that it would change gradually and

Watch David Krakauer’s interview at [www.santafe.edu/life](http://www.santafe.edu/life)

insights from biology with information processing techniques from computer science, to help figure out how life balances these trade-offs. Rather than modeling real





According to the reflection principle, organisms are reflections of their environments. Here the environment is represented as a prose excerpt from Darwin's *Origin of Species*, wherein he contemplates an entangled bank filled with "endless forms most beautiful and most wonderful," each evolved through natural selection. Organism A matches the environment's complexity while Organism B is less complex than the environment and Organism C is more complex than the environment. The research team is asking when evolution satisfies or violates the reflection principle.

be more stable. Even when we complain about our slow-moving government, we are undoubtedly comforted by that very characteristic, because slow is predictable. And when our environment is predictable, we know how to make sound decisions.

So people construct relatively slow-moving, predictable institutions. At the same time, those institutions help shape the behavior of the individuals who created them, points out C4 postdoctoral fellow Philip Poon, who is examining feedback from institutions to explain (in the government case) why, for example, Democrats and Republicans seem to trade off controlling the White House and Congress every few terms. A critical issue is, of course, how that feedback works from a mechanical perspective – that is, how the ways individuals perceive and understand institutions influence their decision-making.

Drawing on the theory of phase transitions, the same one that explains why changing a seemingly minor variable can suddenly shift an entire system from one state to another, Flack, Krakauer, and C4 researcher and former SFI Postdoctoral Fellow Bryan Daniels argue that hierarchical structures bestow another advantage: efficient information flow from the

collective to its individual parts. Systems perched on the edge of a phase transition are exquisitely sensitive, so that a small or localized change "leads to a large change in the global dynamics," Daniels says. Though that might seem unstable or chaotic, systems near the critical point where a phase transi-

**"When monkeys aren't feeling especially aggressive, things are stable, but if the group is sitting at the critical point, one extra monkey picking a fight is enough to kick off a large-scale brawl."**

tion begins are actually quite predictable.

Groups of macaque monkeys – one of Flack and the team's earliest sources of data and inspiration – are one system that appears to be resting near a phase transition. When monkeys aren't feeling especially aggressive, Daniels says, things are stable, and "if I suddenly act out, nothing's going to happen... but if [the group is] sitting at the critical point then [my] contribution is always more important," and one extra monkey picking a fight is enough to kick off a large-scale brawl.

Below the critical point, individual monkeys act fairly independently, but right at the transition, their behavior is tightly

coordinated and individual monkeys act together as one. That, Krakauer emphasizes, eases the flow of information from the system as a whole to its constituent parts, making it all the more predictable. Nature is rife with examples: one bird's sudden course correction changing the

direction of an entire flock, alternating Democrat and Republican control of Congress, and so on.

### Social circuitry

In the abstract, Nature has good reason to favor complexity and hierarchy – each in its own way makes the tasks of comprehension, prediction, and strategizing easier. But it's a third concept, circuits, that grounds Flack and Krakauer's team and lays the practical foundation for much of their work.

Usually the word "circuit" conjures images of the transistors, microprocessors,









People gather in Washington, D.C. before Martin Luther King Jr.'s "I Have a Dream" speech on Aug. 28, 1963, to demand equalities. Individuals tend to move faster and be less predictable than the slow-moving institutions they create.

responsible for power structures, fight-size distributions, or other macroscopic features, using "what we know about individual or component cognition to coarse grain or compress" social circuits, Flack says. Such compression is essential, she says, because living things can't base their decisions on what every other living thing is doing; instead, they're forced to pay attention to just a few patterns or details of what's going on around them.

The key question here, as collaborator and Princeton University graduate student Eleanor Brush puts it, is how little information individuals need to successfully outsmart others.

### Accidental or inevitable

Answering questions like that one – or testing some of the team's more abstract predictions – remains a central challenge. Poon, for example, describes his studies of election cycles and policy change as "toy models" – they capture qualitative features of the data such as party switching, but don't stand up to more precise, quantitative tests.

Meanwhile, Krakauer and others say it's not always clear how to test particular hypotheses, such as the prediction that basis mismatch leads to ever-increasing complexity

**"Living things can't base their decisions on what every other living thing is doing; instead, they're forced to pay attention to just a few patterns."**

in living things. "We're still looking for some compelling example," Ellison says. "Part of the issue is that one can often play the devil's advocate and call into question the example," one reason why his work to construct formal, precise measures of complexity is so important, he says.

New techniques for rapidly analyzing genetic data, Krakauer says, might improve the situation. Combining those techniques with laboratory-based "experimental evolution," in which researchers study the effects of precise environmental changes on small organisms such as bacteria, could help test some of the endeavor's core ideas, such as the reflection principle or the role of phase transitions.

Another potential avenue is to use "digital sources like computer games, where we can control, to a large extent, the form of the data or the conditions under which they were collected," Krakauer says.

Testing their theories is just one part of the team's ambitious aims. They hope, Flack says,

to achieve nothing less than an understanding of why life is organized the way it is, from the smallest bacteria to the largest human institutions. That requires combining real-world observation and abstract mathematical theory in novel and creative ways.

And as if that wasn't enough, Krakauer has one more question in mind: is life an accident, or is it inevitable? And if life is inevitable, well, are we alone?

"If it's not a product of a series of random accidents, but there's an underlying law-like regularity, that would give us confidence in believing in the possibility of life present everywhere in the universe," he says. "So when one asks why does it matter whether it's chance or necessity, it matters if we care whether we're alone are not." ■

*Nathan Collins is a freelance science writer, new father, and film aficionado based in San Francisco.*



**“...in mere Time, all things follow one another, and in mere Space all things are side by side; it is accordingly only by the combination of Time and Space that the representation of coexistence arises.”**

— ARTHUR SCHOPENHAUER, *On the Fourfold Root of the Principle of Sufficient Reason*, 1813



Yves Tanguy, *Indefinite Divisibility*, 1942



# LIFE'S INFORMATION HIERARCHY

The explanation for the complex, multi-scale structure of biological and social systems lies in their manipulation of space and time to reduce uncertainty about the future.

BY JESSICA C. FLACK

CO-DIRECTOR, CENTER FOR COMPLEXITY & COLLECTIVE COMPUTATION,  
WISCONSIN INSTITUTE FOR DISCOVERY, UNIVERSITY OF WISCONSIN-MADISON

EXTERNAL PROFESSOR, SANTA FE INSTITUTE

**B**iological systems – from cells to tissues to individuals to societies – are hierarchically organized (*e.g.* Feldman & Eschel, 1982; Buss 1987; Smith & Szathmáry, 1998; Valentine & May, 1996; Michod, 2000; Frank, 2003). To many, hierarchical organization suggests the nesting of components or individuals into groups, with these groups aggregating into yet larger groups. But this view – at least superficially – privileges space and matter over time and information. Many types of neural coding, for example, require averaging or summing over neural firing rates. The neurons' spatial location – that they are in proximity – is, of course, important, but at least as important to the encoding



Jessica Flack

is their behavior in time. Likewise, in some monkey societies, as I will discuss in detail later in this review, individuals estimate the future cost of social interaction by encoding the average outcome of past interactions in special signals and then summing over these signals.

In both examples, information from events distributed in time as well as space (Figure 1) is captured with encodings that are used to control some behavioral output. My collaborators and I in the Center for Complexity & Collective

Computation are exploring the idea that hierarchical organization at its core is a nesting of these kinds of functional encodings. As I will explain, we think these functional encodings result

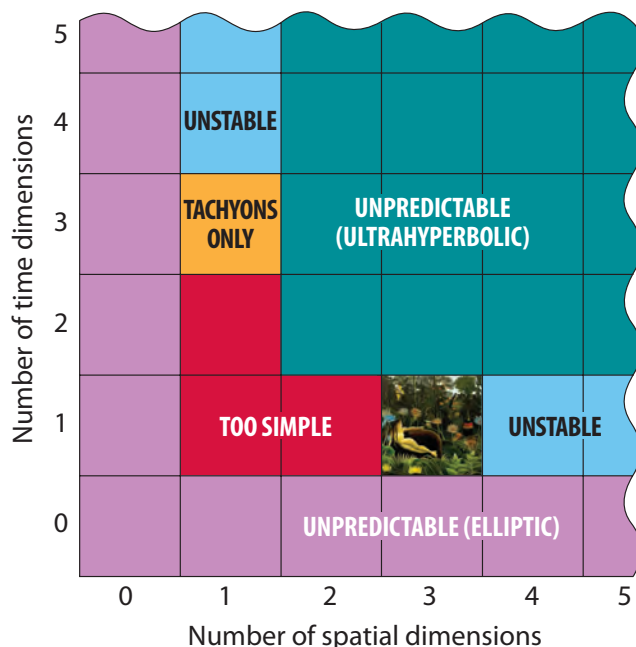
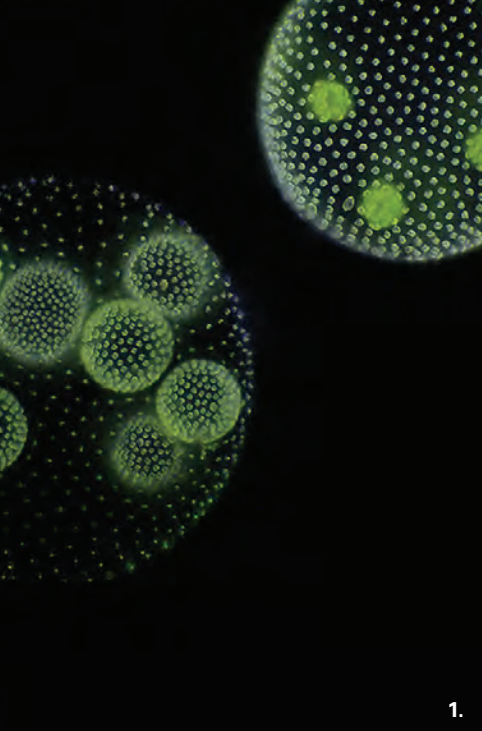


Figure 1. The dimensionality of the time-space continuum, with properties postulated when  $x$  does not equal 3 and  $y$  is larger than 1. Life on earth exists in three spatial dimensions and one temporal dimension. Biological systems effectively “discretize” time and space to reduce environmental uncertainty by coarse-graining and compressing environmental time series to find regularities. Components use the coarse-grained descriptions to predict the future, tuning their behavior to their predictions.





1.



2.



Figure 2. Biological systems – from (left to right) Volvox colonies, to slime molds, to animal societies, to large-scale ecosystems such as reefs, to human cities – are hierarchically organized, with multiple functionally important time and space scales. All feature: 1) components with only partially aligned interests exhibiting coherent behavior at the aggregate level; 2) components that turn over and that co-exist in the system at varying stages of development; 3) social structure that persists but component behavior that fluctuates; and 4) macroscopic variation in temporal and spatial structure and coupling with microscopic behavior, which has functional implications when the components can perceive – in evolutionary, developmental, or ecological time – regularities at the macroscopic scale.

from biological systems manipulating space and time (Figure 2) to facilitate information extraction, which in turn facilitates more efficient extraction of energy.

This *information hierarchy* appears to be a universal property of biological systems and may be the key to one of life's greatest mysteries – the origins of biological complexity. In this essay, I review a body of work by David Krakauer, myself, and our research group that has been inspired by many years of work at

This recasting of the evolutionary process as an inferential one<sup>1</sup> (Bergstrom & Rosvall, 2009; Krakauer, 2011) is based on the premise that organisms and other biological systems can be viewed as hypotheses about the present and future environments they or their offspring will encounter, induced from the history of past environmental states they or their ancestors have experienced (e.g. Crutchfield & Feldman, 2001; Krakauer & Zannotto, 2009; Ellison, Flack, & Krakauer, in prep). This premise, of course, only holds if the past is prologue – that is, has regularities, and the regularities can be estimated and even manipulated (as in niche construction) by biological systems or their components to produce adaptive behavior (Flack, Erwin, Elliot, & Krakauer, 2013; Ellison, Flack, & Krakauer, in prep).

If these premises are correct, life at its core is computational, and a central question becomes: How do systems and their components estimate and control the regularity in their environments and use these estimates to tune their strategies? I suggest that the answer to this question, and the explanation for complexity, is that biological systems manipulate spatial and temporal structure to produce order – low variance – at local scales.

#### UNCERTAINTY REDUCTION

The story I want to tell starts with the observation that with each new level of organization typically comes new functionality – a new feature with positive payoff consequences for the system as a whole, or for its components (Flack, Erwin, Elliot,

the Santa Fe Institute (e.g. Crutchfield, 1994; Gell-Mann, 1996; Gell-Mann & Lloyd, 1996; Fontana & Buss, 1996; West, Brown, & Enquist, 1997; Fontana & Schuster, 1998; Ancel & Fontana, 2000; Stadler, Stadler, Wagner, & Fontana, 2001; Smith, 2003; Crutchfield & Görnerup, 2006; Smith, 2008). Our work suggests that complexity and the multi-scale structure of biological systems are the predictable outcome of evolutionary dynamics driven by uncertainty minimization (Krakauer, 2011; Flack, 2012; Flack, Erwin, Elliot, & Krakauer, 2013).

<sup>1</sup> This idea is related to work on Maxwell's Demon (e.g. Krakauer, 2011; Mandal, Quan, & Jarzynski, 2013) and the Carnot cycle (e.g. Smith, 2003), but we do not yet understand the mapping.



& Krakauer, 2013). Policing in a pigtailed macaque group is an example. Once a heavy tailed distribution of power – defined as the degree of consensus in the group that an individual can win fights (see Flack & Krakauer, 2006; Boehm & Flack, 2010; Brush, Krakauer, & Flack, 2013) – becomes effectively institutionalized (here meaning hard to change) policing (an intrinsically costly strategy) becomes affordable, at least to those animals that sit in the tail of the power distribution: those super powerful monkeys who are rarely or never challenged when they break up fights (Flack, de Waal, & Krakauer, 2005; Flack, Girvan, de Waal, & Krakauer, 2006).

My collaborators and I propose that a primary driver of the emergence of new functionality such as policing is the reduction of environmental uncertainty through the construction of nested dynamical processes with a range of characteristic time constants (Flack, Erwin, Elliot, & Krakauer, 2013). These nested dynamical processes arise as components extract regularities from fast, microscopic behavior by coarse-graining (or compressing) the history of events to which they have been exposed.

Proteins, for example, can have a long half-life relative to RNA transcripts, and can be thought of as the summed output of translation. Cells have a long half-life relative to proteins,

and are a function of the summed output of arrays of spatially structured proteins. Both proteins and cells represent some average measure of the noisier activity of their constituents. Similarly, a pigtailed macaque's estimate of its power is a kind

of average measure of the collective perception in the group that the macaque is capable of winning fights, and this is a better predictor of the cost the macaque will pay during fights than the outcome of any single melee, as

these outcomes can fluctuate for contextual reasons. These coarse-grainings, or averages, are encoded as slow variables (Flack & de Waal, 2007; Flack, 2012; Flack, Erwin, Elliot, & Krakauer, 2013; see also Feret, Danos, Krivine, Harner, & Fontana, 2009, for a similar idea). Slow variables may have a spatial component as well as a temporal component, as in the protein and cell examples (Figure 6), or, minimally, only a temporal component, as in the monkey example.

As a consequence of integrating over abundant microscopic processes, slow variables provide better predictors of the local future configuration of a system than the states of the fluctuating microscopic components. In doing so, they promote accelerated rates of microscopic adaptation. Slow variables facilitate adaptation in two ways: They allow components to fine-tune their behavior, and they free components to search, at low cost, a larger

**“Slow variables provide better predictors of the local future configuration of a system than the states of the fluctuating microscopic components.”**



This model is frequently revised. It is based on the latest laboratory data, some of which is not yet published.

Additional data sources for selected notes: 1: McClay lab; 2: Angerer lab; 3,4: McClay lab; 5: Rogers and Calestani, 2010; 6: Croce and McClay

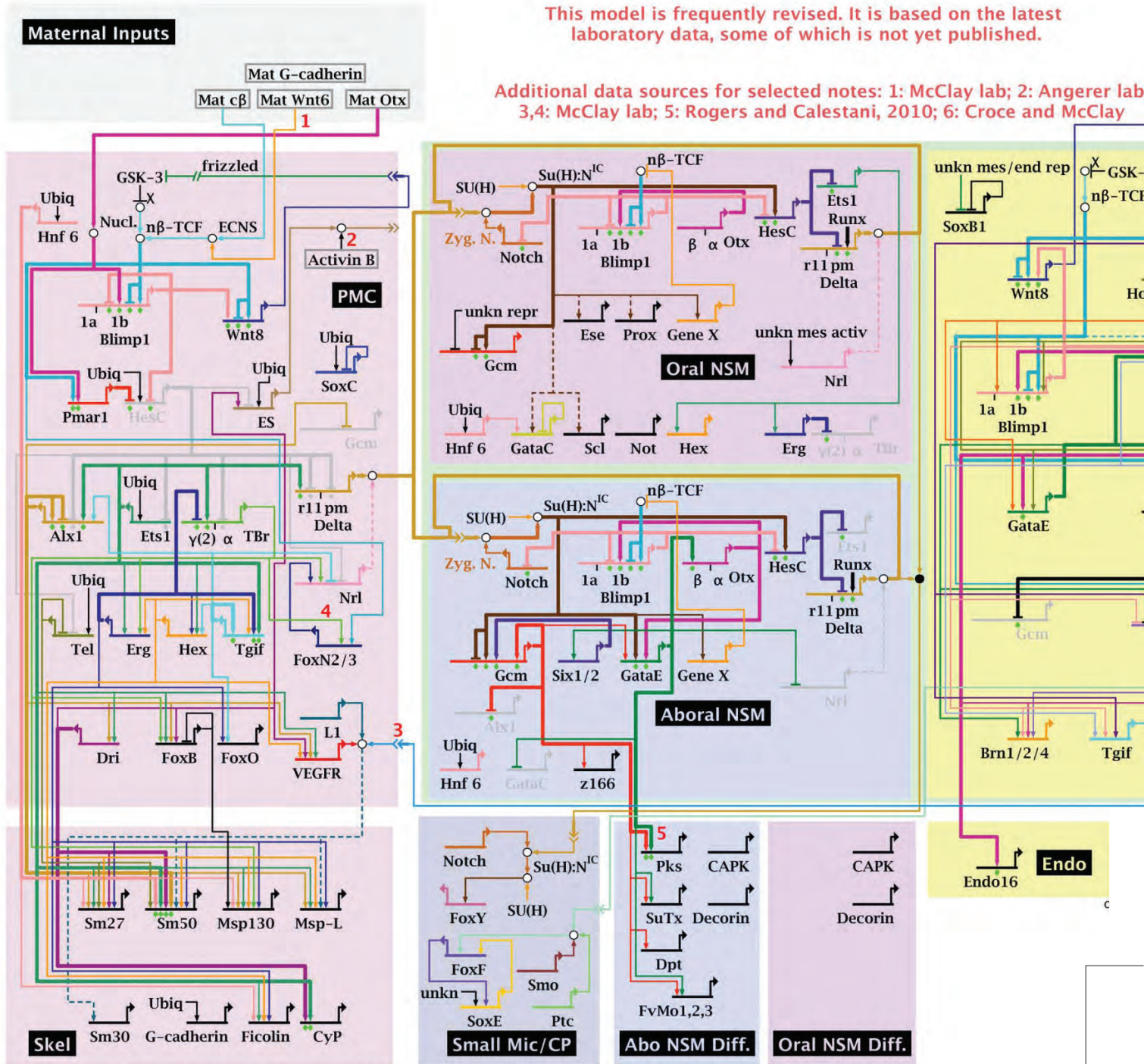
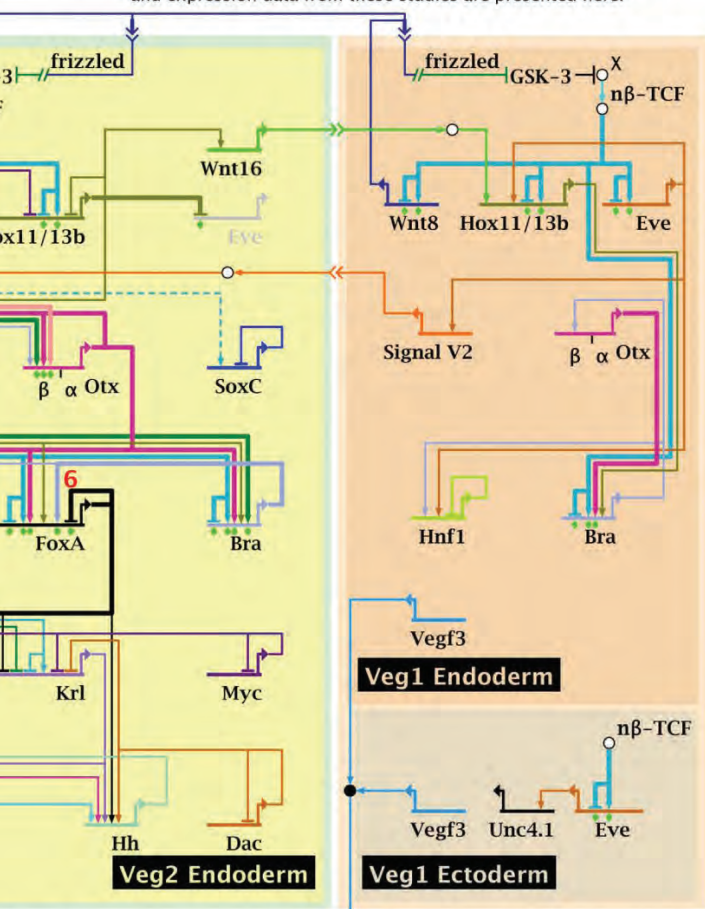


Figure 3. A sea urchin gene regulatory circuit. The empirically derived circuit describes the Boolean rules for coordinating genes and proteins to produce aspects of the sea urchin's phenotype – in this case, the position of cells in the endomesoderm at 30 hours since fertilization. Edges indicate whether a node induces a state change in another node, here genes and proteins. The circuit is a rigorous starting point for addressing questions about the logic of development and its evolution. In computational terms, the input is the set of relevant genes and proteins and the output is the target phenotypic feature.

The current VFA includes not yet published cis-regulatory data of Smadar de Leon, Joel Smith (in press), Andrew Cameron, Qiang Tu, Sagar Damle, Andrew Ransick, Christina Theodoris, and, in addition to published data, is based on recent perturbation and other results of Isabelle Peter (endoderm domains), Stefan Materna (NSM domain), and Joel Smith (CP domains) of the Davidson Lab. Relevant perturbation and expression data from these studies are presented here.



This "Up to 30 Hour Overview" primarily shows the endomesoderm network architecture as it exists after 21 hours, with the addition of all PMC components starting at 6 hours, the inclusion of the Delta-Notch signal from PMC to Veg2, the presence of Wnt8 in Veg2 Endoderm, the nB-TCF and Otx inputs into Blimp1 in NSM, and Gene X in the NSM; the latter four of these features are no longer present by 21 hours. Consult the other models to see all the network elements and interactions in the correct temporal context.

Copyright © 2001–2011 Hamid Bolouri and Eric Davidson

space of strategies for extracting resources from the environment (Flack, 2012; Flack, Erwin, Elliot, & Krakauer, 2013). This phenomenon is illustrated by the power-in-support-of-policing example and also by work on the role of neutral networks in RNA folding. In the RNA case, many different sequences can fold into the same secondary structure. This implies that over evolutionary time, structure changes more slowly than sequence, thereby permitting sequences to explore many configurations under normalizing selection (Fontana & Schuster, 1998; Schuster & Fontana, 1999; Ferrada & Krakauer, in prep).

## NEW LEVELS OF ORGANIZATION

As an interaction history builds up at the microscopic level, the coarse-grained representations of the microscopic behavior consolidate, becoming for the components increasingly robust predictors of the system's future state.

We speak of a new organizational level when the system's components rely to a greater extent on these coarse-grained or compressed descriptions of the system's dynamics for adaptive decision-making than on local fluctuations in the microscopic behavior and when the coarse-grained estimates made by components are largely in agreement (Krakauer, Bertschinger, Ay, Olbrich, & Flack, in review). The idea is that convergence on these "good-enough" estimates underlies non-spurious correlated behavior among the components. This in turn leads to an increase in local predictability (e.g. Flack & de Waal, 2007; Brush, Krakauer, & Flack, 2013) and drives the construction of the information hierarchy. (Note that increased predictability can seem the product of downward causation in the absence of careful analysis of the bottom-up mechanisms that actually produced it.)

## THE STATISTICAL MECHANICS & THERMODYNAMICS OF BIOLOGY

Another way of thinking about slow variables is as a functionally important subset of the system's potentially many macroscopic properties. An advantage of this recasting is that it builds a bridge to physics, which over the course of its maturation as a field grappled with precisely the challenge now before biology: understanding the relationship between behavior at the individual or component level and behavior at the aggregate level.

### In physics

As discussed in Krakauer & Flack (2010), the debate in physics began with thermodynamics – an equilibrium theory treating aggregate variables – and came to a close with the maturation of statistical mechanics – a dynamical theory treating microscopic variables.

Thermodynamics is the study of the macroscopic behavior of systems exchanging work and heat with connected systems or their environment. The four laws of thermodynamics all operate on average quantities defined at equilibrium – temperature, pressure, entropy, volume, and energy. These macroscopic variables exist in fundamental relationships with each other, as expressed, for example, in the ideal gas law. Thermodynamics is an extremely powerful framework as it provides experimentalists with explicit, principled recommendations about what variables should be measured and how they are expected to change relative to each other, but it is not a dynamical theory and offers no explanation for the mechanistic origins of the macroscopic variables it privileges. This is the job of statistical mechanics. By providing the microscopic basis for the macroscopic variables in thermodynamics, statistical mechanics establishes the conditions under which the equilibrium relations are no longer valid or expected to apply. The essential intellectual technologies behind much of statistical mechanics are powerful tools for counting possible microscopic configurations of a system and connecting these to macroscopic averages.

### In biology

This brief summary of the relation between thermodynamics and statistical mechanics in physics is illuminating for two reasons. On the one hand it raises the possibility of a potentially



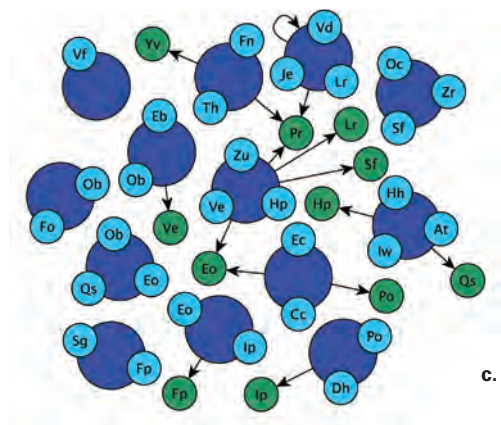
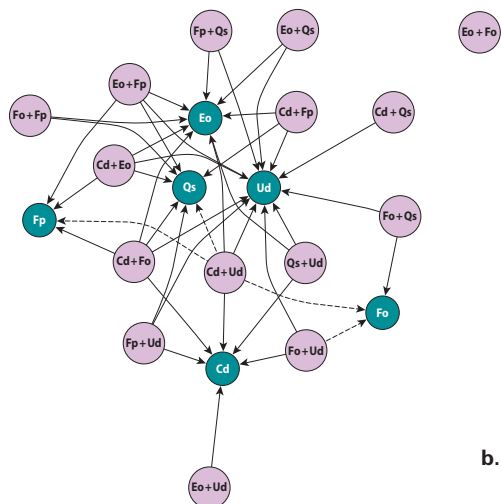
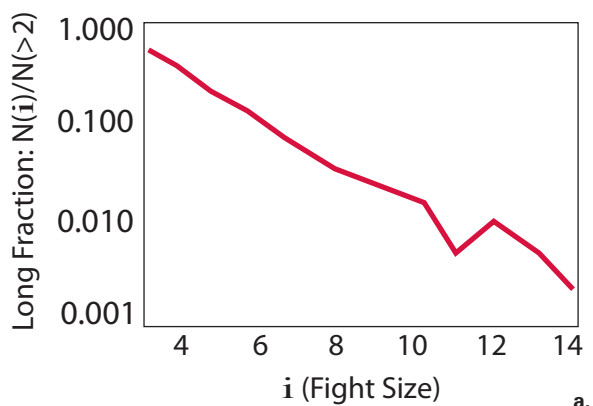


Figure 4. Cognitive effective theories for one macroscopic property of a macaque society: the distribution of fight sizes (a). To reduce circuit complexity we return to the raw time series data and remove as much noise as possible by compressing the data. In the case of our macaque dataset, this reveals which individuals and subgroups are regular and predictable conflict participants. We then search for possible strategies in response to these regular and predictable individuals and groups. This approach returns a family of circuits (b is an example), each of which has fewer nodes and edges than the full circuit (c). These circuits are simpler and more cognitively parsimonious. We then test the reduced circuits against each other in simulation to determine how well they recover the target macroscopic properties.

deep division between physical and biological systems: So far – and admittedly biology is young – biology has had only limited success in empirically identifying important macroscopic properties and deriving these from first principles rooted in physical laws or deep evolved constraints<sup>2</sup>. This may be the case because many of the more interesting macroscopic properties are slow variables that result from the collective behavior of adaptive components, and their functional value comes from how components use them, making them fundamentally subjective (see Gell-Mann & Lloyd, 1996 for more on subjectivity) and perhaps even nonstationary<sup>3</sup>.

On the other hand, the role of statistical mechanics in physics suggests a way forward. If we have intuition about which macroscopic properties are important – that is, which macroscopic properties are slow variables – and we can get good data on the relevant microscopic behavior, we can proceed by working upward from dynamical many-body formalisms to equilibrium descriptions with a few favored macroscopic degrees of freedom (Levin, Grenfell, Hastings, & Perelson, 1997; Krakauer & Flack, 2010; Krakauer et al., 2011; Gintis, Doebeli, & Flack, 2012).

#### A STATISTICAL MECHANICS-COMPUTER SCIENCE- INFORMATION THEORETIC HYBRID APPROACH

The most common approach to studying the relationship between micro and macro in biological systems is perhaps dynamical systems and, more specifically, pattern formation (for examples, see Sumpter, 2006; Ball, 2009; Couzin, 2009; Payne et al., 2013). However, if, as we believe, the information hierarchy results from biological components collectively estimating regularities in their environments by coarse-graining or compressing time series data, a natural (and complementary) approach is to treat the micro and macro mapping explicitly as a computation.

#### Elements of computation in biological systems

Describing a biological process as a computation minimally requires that we are able to specify the output, the input, and the algorithm or circuit connecting the input to the output (Flack & Krakauer, 2011; see also Mitchell, 2010; Valiant, 2013). A secondary concern is how to determine when the desired output has been generated. In computer science this is called the termination criterion or halting problem. In biology it potentially can be achieved by constructing nested dynamical processes with a range of timescales, with the slower timescale processes providing the “background” against which a strategy is evaluated (Flack & Krakauer, 2011), as discussed later in this paper in the section on *Couplings*.

2 The work on scaling in biological systems shows a fundamental relationship between mass and metabolic rate, and this relationship can be derived from the biophysics (e.g. West, Brown, & Enquist, 1997). Bettencourt and West are now investigating whether similar fundamental relationships can be established for macroscopic properties of human social systems, like cities (e.g. Bettencourt, Lobo, Helbing, Kuhnert, & West, 2007; Bettencourt, 2013).

3 With the important caveat that in biology the utility of a macroscopic property as a predictor will likely increase as consensus among the components about the estimate increases, effectively reducing the subjectivity and increasing stationarity (see also Gell-Mann & Lloyd, 1996).

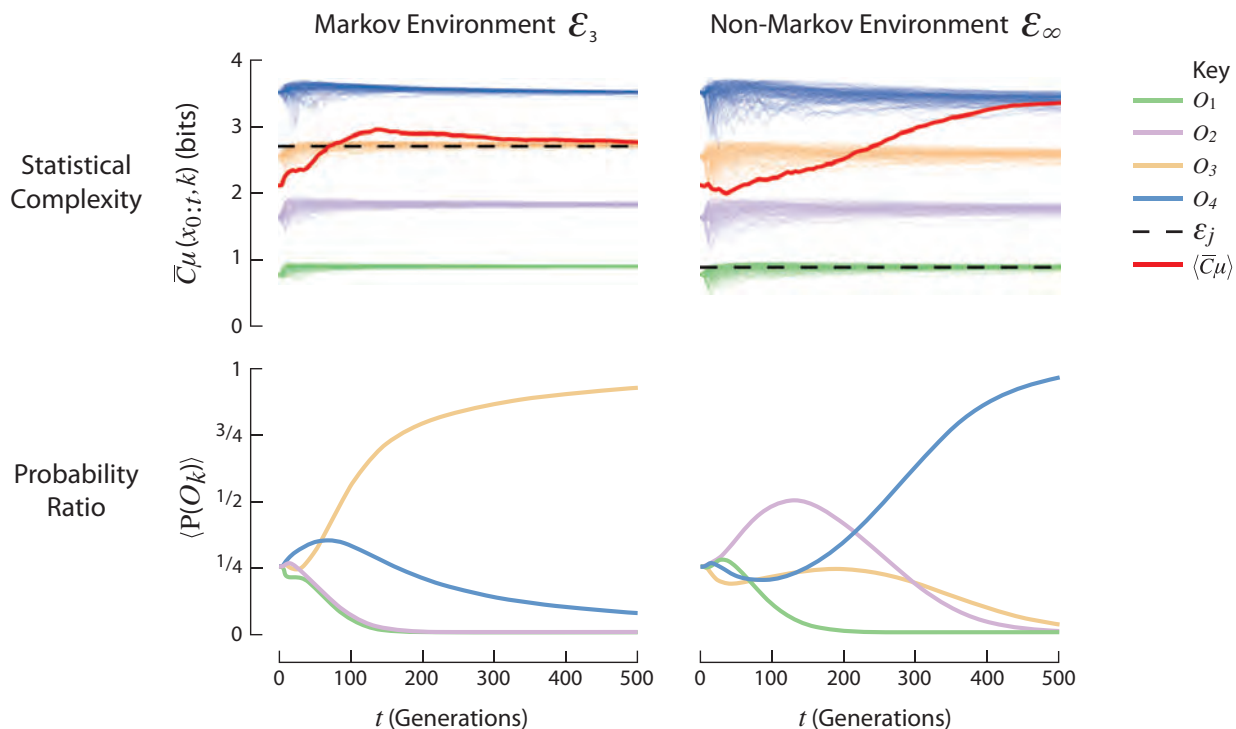


Figure 5. A comparison of Markov organisms in two environments: a Markov environment (left) and a non-Markov environment (right). In the top two plots, organismal complexity is plotted against time for each organism (organisms are represented by varying colors) and for many different sequences of 500 environmental observations; the bold red line shows the average organismal complexity, which in the Markov environment tends toward the environmental complexity and in the non-Markov environment exceeds it. In the bottom plots, the probability that a random organism has order  $k$  is plotted against time.

A macroscopic property can be said to be an output of a computation if it can take on values that have functional consequences at the group or component level, if it is the result of a distributed and coordinated sequence of component interactions under the operation of a strategy set, and if it is a stable output of input values that converges (terminates) in biologically relevant time (Flack & Krakauer, 2011). Examples studied in biology include aspects of vision such as edge detection (e.g. Olshausen & Field, 2004), phenotypic traits such as the average position of cells in the developing endomesoderm of the sea urchin (e.g. Davidson, 2010; Peter & Davidson, 2011), switching in biomolecular signal-transduction cascades (e.g. Smith, Krishnamurthy, Fontana, & Krakauer, 2011), chromatin regulation (e.g. Prohaska, Stadler, & Krakauer, 2010), and social structures such as the distribution of fight sizes (e.g. DeDeo, Krakauer, & Flack, 2010; Flack & Krakauer, 2011; Lee, Daniels, Krakauer, & Flack, in prep) and the distribution

of power in monkey societies (e.g. Flack, 2012; Flack, Erwin, Elliot, & Krakauer, 2013).

The input to the computation is the set of elements implementing the rules or strategies. As with the output, we do not typically know *a priori* which of many possible inputs is relevant,

and so we must make an informed guess based on the properties of the output. In the case of the sea urchin's endomesoderm, we might start with a list of genes that have been implicated in the regulation of

cell position. In the case of the distribution of fight sizes in a monkey group, we might start with a list of individuals participating in fights.

### Reconstructing the microscopic behavior

In biological systems the input plus the strategies constitute the system's microscopic behavior. There are many approaches to reconstructing the system's microscopic behavior. The most powerful is an experiment in which upstream inputs to a

**“In all biological systems there are multiple components interacting and simultaneously coarse graining to make predictions about the future.”**



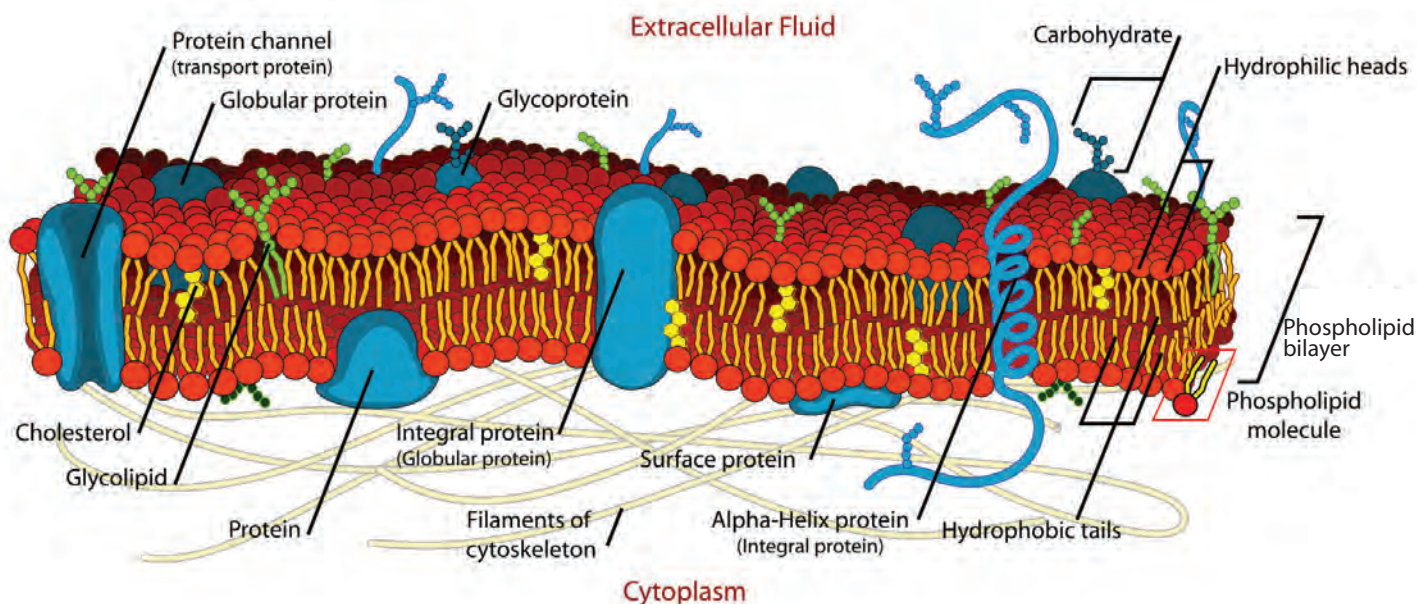


Figure 6. The cell can be thought of as a slow variable to the extent it is a function of the summed output of arrays of spatially structured proteins and has a long half-life compared to its proteins. Features that serve as slow variables provide better predictors of the local future configuration of a system than the states of the fluctuating microscopic components. We propose that when detectable by the system or its components, slow variables can reduce environmental uncertainty and, by increasing predictability, promote accelerated rates of microscopic adaptation.

target component are clamped off and the output of the target component is held constant. This allows the experimentalist to measure the target component’s specific contribution to the behavior of a downstream component (Pearl, 2010). This type of approach is used to construct gene regulatory circuits mapping gene-gene and gene-protein interactions to phenotypic traits (Figure 3).

When such experiments are not possible, causal relationships can be established using time series analysis in which clamping is approximated statistically (Ay, 2009; Pearl, 2010). My collaborators and I have developed a novel computational technique, called Inductive Game Theory (DeDeo, Krakauer, & Flack, 2010; Flack & Krakauer, 2011; Lee, Daniels, Krakauer, & Flack, in prep), that uses a variant of this statistical clamping principle to extract strategic decision-making rules, game structure, and (potentially) strategy cost from correlations observed in the time series data.

### Collective computation through stochastic circuits

In all biological systems, of course, there are multiple components interacting and simultaneously coarse graining to make predictions about the future. Hence the computation is inherently collective. A consequence of this is that it is not sufficient to simply extract from the time series the list of the strategies

in play. We must also examine how different configurations of strategies affect the macroscopic output. One way these configurations can be captured is by constructing Boolean circuits describing activation rules as illustrated by the gene regulatory circuit shown in Figure 3, which controls cell position (the output) at thirty hours from fertilization in the sea urchin (Peter & Davidson, 2011). In the case of our work on micro to macro mappings in animal societies, we describe the space of microscopic configurations using stochastic “social” circuits (Figure 4) (DeDeo, Krakauer, & Flack, 2010; Flack & Krakauer, 2011; Lee, Daniels, Krakauer, & Flack, in prep).

Nodes in these circuits are the input to the computation. As discussed above, the input can be individuals or subgroups, or they can be defined in terms of component properties like age or neurophysiological state. A directed edge between two nodes indicates that the “receiving node” has a strategy for the “sending node” – and the edge weight can be interpreted as the above-null probability that the sending node plays the strategy in response to some behavior by the receiving node in a previous time step. Hence, an edge in these circuits quantifies the strength of a causal relationship between the behaviors of a sending and receiving node.

Sometimes components have multiple strategies in their repertoires. Which strategy is being played at time  $t$  may

vary with context. These meta-strategies can be captured in the circuit using different types of gates specifying how a component's myriad strategies combine (see also Feret, Davis, Krivine, Harmer, & Fontana, 2009). By varying the types of gates and/or the strength of causal relationships, we end up with multiple alternative circuits – a family of circuits – all of which are consistent with the microscopic behavior, albeit with different degrees of precision (Lee, Daniels, Krakauer, & Flack, in prep). Each circuit in the family is essentially a model of the micro-macro relationship and so serves as a hypothesis for how strategies combine over nodes (inputs) to produce the target output (Lee, Daniels, Krakauer, & Flack, in prep). We test the circuits against each other in simulation to determine which can best recover the actual measured macroscopic behavior of our system.

### Cognitive effective theories for collective computation

The circuits describing the microscopic behavior can be complicated, with many “small” causes detailed, as illustrated by the gene regulatory circuit shown in Figure 3. The challenge – once we have rigorous circuits – is to figure out the circuit logic (Flack & Krakauer, 2011; see also Feret, Davis, Krivine, Harmer, & Fontana, 2009).

There are many ways to approach this problem. Our approach is to build what's called in physics an *effective theory*: a compact description

of the causes of a macroscopic property. Effective theories for adaptive systems composed of adaptive components require an additional criterion beyond compactness. As discussed earlier in this essay, components in these systems are tuning their behaviors based on their own effective theories – coarse-grained rules (see also Feret, Davis, Krivine, Harmer, & Fontana, 2009) – that capture the regularities (Daniels, Krakauer, & Flack, 2012). If we are to build an effective theory that explains the origins of functional space and time scales – new levels of organization – and ultimately the information hierarchy, the effective theory must be consistent with component models of macroscopic behavior, as these models, through their effects on strategy choice, drive that process. In other words, our *effective theory should explain how the system itself is computing*.

We begin the search for cognitively principled effective theories using what we know about component cognition to inform how we coarse-grain and compress the circuits (Flack & Krakauer, 2011; Lee, Daniels, Krakauer, & Flack, in prep). This means taking into account, given the available data, the kinds of computations components can perform and the error associated with these computations at the individual and collective levels, given component memory capacity and the quality of the “data sets” components use to estimate

regularities (Krakauer, Flack, DeDeo, & Farmer, 2010; Flack & Krakauer, 2011; Daniels, Krakauer, & Flack, 2012; Ellison, Flack, & Krakauer, in prep; all building on Gell-Mann, 1996).

As we refine our understanding of the micro-macro mapping through construction of cognitive effective theories, we also refine our understanding of what time series data constitute the “right” input – and hence the building blocks of our system. And, by investigating whether our best-performing empirically justified circuits can also account for other potentially important macroscopic properties, we can begin to establish which macroscopic properties might be fundamental and what their relation is to one another – the thermodynamics of biological collectives.

### Couplings, information flow, and macroscopic tuning

Throughout this essay I have stressed the importance of slowness (effective stationarity) for prediction. Slowness also has costs, however. Consider our power example. The power structure must change slowly if individuals are to make worthwhile investments in strategies that work well given the structure, but it cannot change too slowly or it may cease to reflect the underlying distribution of fighting abilities on which it is based

and, hence, cease to be a good predictor of interaction cost (Flack, 2012; Flack, Erwin, Elliot, & Krakauer, 2013). The question we must answer is, what is the optimal coupling between macroscopic and microscopic change,

and can systems, by manipulating how components are organized in space and time, get close to this optimal coupling?

One approach to this problem is to quantify the degeneracy of the target macroscopic property and then perturb the circuits by either removing nodes, up- or down-regulating node behavior, or restructuring higher order relationships (subcircuits) to determine how many changes at the microscopic level need to occur to induce a state change at the macroscopic level.

Another approach is to ask how close the system is to a critical point – that is, how sensitive the target macroscopic property is to small changes in parameters describing the microscopic behavior. Many studies suggest that biological systems of all types sit near the critical point (Mora & Bialek, 2011). A hypothesis we are exploring is that sitting near the critical point means that important changes at the microscopic scale will be visible at the macroscopic scale (Daniels, Krakauer, & Flack, in prep). Of course this also has disadvantages as it means small changes can potentially cause big institutional shifts, undermining the utility of coarse-graining and slow variables for prediction (Flack, Erwin, Elliot, & Krakauer, 2013).

If balancing trade-offs between robustness and prediction on the one hand, and adaptability to changing environments

**A hypothesis we are exploring is that sitting near the critical point means that important changes at the microscopic scale will be visible at the macroscopic scale.”**





on the other, can be achieved by modulating the coupling between scales (Flack, Hammerstein, & Krakauer, 2012; Flack, Erwin, Elliot, & Krakauer, 2013), we should be able to make predictions about whether a system is far from, near, or at the critical point based on whether the data suggest that robustness or adaptability is more important given the environment and its characteristic timescale (Daniels, Krakauer, & Flack, in prep). This presupposes that the system can optimize where it sits with respect to the critical point, implying active mechanisms for modulating the coupling. We are working to identify plausible mechanisms using a series of toy models to study how the type of feedback from the macroscopic or institutional level to the microscopic behavior influences the possibility of rapid institutional switches (Poon, Flack, & Krakauer, in prep; see also Sabloff, in prep for related work on the rise of the state in early human societies).

### COMPLEXITY

This essay covers a lot of work, so allow me to summarize. I suggested that the origins of the information hierarchy lie in the manipulation of space and time to reduce environmental uncertainty. I further suggested that uncertainty reduction is maximized if the coarse-grained representations of the data the components compute are in agreement (because this increases

the probability that everyone is making the same predictions and so tuning the same way). As this happens, the coarse-grained representations consolidate into robust, slow variables at the aggregate level, creating new levels of organization and giving the appearance of downward causation.

I proposed that a central challenge lies in understanding what the mapping is between the microscopic behavior and these new levels of organization. (How exactly do everyone's coarse grainings converge?) I argued that in biology, a hybrid statistical mechanics-computer science-information theoretic approach (see also Krakauer et al., 2011) is required to establish such mappings. Once we have cognitively principled effective theories for mappings, we will have an understanding of how biological systems, by discretizing space and time, produce information hierarchies.

Where are we, though, with respect to explaining the origins of biological complexity?

The answer we are moving toward lies at the intersection of the central concepts in this essay. If evolution is an inferential process with complex life being the result of biological systems extracting regularities from their environments to reduce uncertainty, a natural recasting of evolutionary dynamics is in Bayesian terms (Ellison, Flack, & Krakauer, in prep). Under this view, organism and environment can be interpreted



as k-order Markov processes and modeled using finite-state hidden Markov models (Figure 5). Organisms update prior models of the environment with posterior models of observed regularities. We are exploring how the Markov order (a proxy for memory) of organisms changes as organisms evolve to match their environment, quantifying fit to the environment with model selection. We use information-theoretic measures to quantify structure. Our approach allows us to evaluate the memory requirements of adapting to the environment given its Markov order, quantify the complexity of the models organisms build to represent their environments, and quantitatively compare organismal and environmental complexity as our Markov organisms evolve. We hypothesize that high degrees of complexity result when there is regularity in the environment, but it takes a long history to perceive it and an elaborate model to encode it (Ellison, Flack, & Krakauer, in prep).

### Acknowledgements

This essay summarizes my view of the past, present, and predicted future of the core research program at the Center for Complexity & Collective Computation (C4). In addition to our current collaborators – Nihat Ay, Dani Bassett, Karen Page, Chris Boehm, and Mike Gazzaniga – and the super smart students and postdoctoral fellows Eleanor Brush, Bryan Daniels, Simon DeDeo, Karl Doron, Chris Ellison, the late Tanya Elliot, Evandro Ferrada, Eddie Lee, and Philip Poon, who have carried out much of this work on a daily basis, I am deeply grateful to the Santa Fe Institute for its support over the years and to the Santa Fe folks whose ideas have provided inspiration. First and foremost this includes David Krakauer, my main collaborator. Other significant SFI influences include Jim Crutchfield, Doug Erwin, Walter Fontana, Lauren Ancel Meyers, Geoffrey West, Eric Smith, Murray Gell-Mann, Bill Miller, David Padwa, and Cormac McCarthy. I am indebted to Ellen Goldberg for making possible my first postdoctoral position at SFI. Finally, much of this research would not be possible without the generous financial support provided by the John Templeton Foundation through a grant to SFI to study complexity and a grant to C4 to study the mind-brain problem, a National Science Foundation grant (0904863), and a grant from the U.S. Army Research Laboratory and the U.S. Army Research Office under contract number W911NF-13-1-0340.

### References

Ancel, L. W., & W. Fontana. 2000. *Plasticity, evolvability, and modularity in RNA*. J. Exp. Zoology (Molec. & Devel. Evol.) 288: 242–283.

Ay, N. 2009. *A refinement of the common cause principle*. Discrete Appl. Math. 157: 2439–2457.

Ball, P. 2009. *Nature's patterns: A tapestry in three parts*. Oxford, UK: Oxford University Press.

Bergstrom, C. T., & M. Rosvall. 2009. *The transmission sense of information*. Biol. & Phil. 26: 159–176.

Bettencourt, L. M. A. 2013. *The origins of scaling in cities*. Science 340: 1438–1441.

Bettencourt, L. M. A., J. Lobo, D. Helbing, C. Kuhnert, & G. B. West. 2007. *Growth, innovation, scaling, and the pace of life in cities*. PNAS 104: 7301–7306.

Boehm, C., & J. C. Flack. 2010. *The emergence of simple and complex power structures through niche construction*. In *The Social Psychology of Power*, ed. A. Guinote & T. K. Vescio, 46–86. New York: Guilford Press.

Brush, E. R., D. C. Krakauer, & J. C. Flack. 2013. *A family of algorithms for computing consensus about node state from network data*. PLOS Comp. Biol. 9: e1003109.

Buss, L. W. 1987. *The evolution of individuality*. Princeton, NJ: Princeton University Press, 224 p.

Couzin, I. D. 2009. *Collective cognition in animal groups*. Trends Cog. Sci. 13: 36–43.

Crutchfield, J. P. 1994. *The calculi of emergence: Computation, dynamics, and induction*. Physica D 75: 11–54.

Crutchfield, J. P., & D. P. Feldman. 2001. *Synchronizing to the environment: Information-theoretic constraints on agent learning*. Adv. Compl. Sys. 4: 251–264.

Crutchfield, J. P., & O. Görnerup. 2006. *Objects that make objects: The population dynamics of structural complexity*. J. Roy. Soc. Interface 22: 345–349.

Daniels, B., D. C. Krakauer, & J. C. Flack. 2012. *Sparse code of conflict in a primate society*. PNAS 109: 14259–14264.

Daniels, B., D. C. Krakauer, & J. C. Flack. n.d. *Conflict tuned to maximum information flow*. In preparation.

Davidson, E. H. 2010. *Emerging properties of animal gene regulatory networks*. Nature 468: 911–920.

DeDeo, S., D. C. Krakauer, & J. C. Flack. 2010. *Inductive game theory and the dynamics of animal conflict*. PLOS Comp. Biol. 6: e1000782.

Ellison, C., J. C. Flack, & D. C. Krakauer. n.d. *On inferential evolution and the complexity of life*. In preparation.

Feldman, M., & I. Eschel. 1982. *On the theory of parent-offspring conflict: A two-locus genetic model*. Amer. Natur. 119: 285–292.

Feret, J., V. Danos, J. Krivine, R. Harmer, & W. Fontana. 2009. *Internal coarse-graining of molecular systems*. PNAS 106: 6453–6458.

Ferrada, E., & D. C. Krakauer. n.d. *The Simon modularity principle*. In preparation.

Flack, J. C. 2012. *Multiple time-scales and the developmental dynamics of social systems*. Phil. Trans. Roy. Soc. B: Biol. Sci. 367: 1802–1810.

Flack, J. C., & D. C. Krakauer. 2006. *Encoding power in communication networks*. Amer. Natur. 168: E87–102.

Flack, J. C., & F. B. M. de Waal. 2007. *Context modulates signal meaning in primate communication*. PNAS 104: 1581–1586.

Flack, J. C., & D. C. Krakauer. 2011. *Challenges for complexity measures: A perspective from social dynamics and collective social computation*. Chaos 21: 037108–037108.

Flack, J. C., F. B. M. de Waal, & D. C. Krakauer. 2005. *Social structure, robustness, and policing cost in a cognitively sophisticated species*. Am. Natur. 165: E126–39.



- Flack, J. C., P. Hammerstein, & D. C. Krakauer. 2012. *Robustness in biological and social systems*. In *Evolution and the mechanisms of decision-making*, ed. P. Hammerstein, & J. Stevens, 129–151. Cambridge: MIT Press.
- Flack, J. C., D. Erwin, T. Elliot, & D. C. Krakauer. 2013. *Timescales, symmetry, and uncertainty reduction in the origins of hierarchy in biological systems*. In *Cooperation and its evolution*, ed. K. Sterelny, R. Joyce, B. Calcott, & B. Fraser, 45–74. Cambridge: MIT Press.
- Flack, J. C., M. Girvan, F. B. M. de Waal, & D. C. Krakauer. 2006. *Policing stabilizes construction of social niches in primates*. *Nature* 439: 426–429.
- Fontana, W., & L. W. Buss. 1996. The barrier of objects: *From dynamical systems to bounded organizations*. In *Boundaries and barriers*, ed. J. Casti & A. Karlqvist, 56–116. Reading, MA: Addison-Wesley.
- Fontana, W., & P. Schuster. 1998. *Continuity in evolution: On the nature of transitions*. *Science* 280: 1451–1455.
- Frank, S. A. 2003. *Repression of competition and the evolution of cooperation*. *Evolution* 57: 693–705.
- Gell-Mann, M. 1996. *Fundamental sources of unpredictability*. Talk presented at conference of the same name. <[http://www-physics.mps.ohio-state.edu/~perry/p633\\_sp07/articles/fundamental-sources-of-unpredictability.pdf](http://www-physics.mps.ohio-state.edu/~perry/p633_sp07/articles/fundamental-sources-of-unpredictability.pdf)>. Accessed 01/10/2014.
- Gell-Mann, M., & S. Lloyd. 1996. *Information measures, effective complexity, and total information*. *Complexity* 2: 44–53.
- Gintis, H., M. Doebeli, & J. C. Flack. 2012. *The evolution of human cooperation*. *Cliodynamics: J. Theor. & Math. Hist.* 3.
- Krakauer, D. C. 2011. *Darwinian demons, evolutionary complexity, and information maximization*. *Chaos*, 21: 037110.
- Krakauer, D. C., N. Bertschinger, N. Ay, E. Olbrich, J. C. Flack. *The information theory of individuality*. In *What is an Individual?* eds. L. Nyhart & S. Lidgard. University of Chicago Press. In review.
- Krakauer, D. C., & J. C. Flack. 2010. *Better living through physics*. *Nature* 467: 661.
- Krakauer, D. C., & P. Zanotto. 2009. *Viral individuality & limitations of the life concept*. In *Protocells: Bridging non-living and living matter*, ed. M. A. Rasmussen et al., 513–536. Cambridge: MIT Press.
- Krakauer, D. C., J. C. Flack, S. DeDeo, & D. Farmer. 2010. *Intelligent data analysis of intelligent systems*. *IDA 2010 LNCS* 6065: 8–17.
- Krakauer, D. C., J. P. Collins, D. Erwin, J. C. Flack, W. Fontana, M. D. Laubichler, S. Prohaska, G. B. West, & P. Stadler. 2011. *The challenges and scope of theoretical biology*. *J. Theor. Biol.* 276: 269–276.
- Lee, E., B. Daniels, D. C. Krakauer, & J. C. Flack. n.d. *Cognitive effective theories for probabilistic social circuits mapping strategy to social structure*. In preparation.
- Levin, S. A., B. Grenfell, A. Hastings, & A. S. Perelson. 1997. *Mathematical and computational challenges in population biology and ecosystems science*. *Science* 275: 334–343.
- Mandal, D., H. T. Quan, & C. Jarzynski. 2013. *Maxwell's refrigerator: An exactly solvable model*. *Phys. Rev. Lett.* 111: 030602.
- Michod, R. E. 2000. *Darwinian dynamics: Evolutionary transitions in fitness and individuality*. Princeton, NJ: Princeton University Press.
- Mitchell, M. 2010. *Biological computation*. Working Paper 2010-09-021, Santa Fe Institute, Santa Fe, NM.
- Mora, T., & W. Bialek. 2011. *Are biological systems poised at criticality?* *J. Stat. Phys.* 144: 268–302.
- Olshausen, B., & D. Field. 2004. *Sparse coding of sensory inputs*. *Curr. Opin. Neurobiol.* 14: 481–487.
- Payne, S., L. Bochong, Y. Cao, D. Schaeffer, M. D. Ryser, & L. You. 2013. *Temporal control of self-organized pattern formation without morphogen gradients in bacteria*. *Mol. Sys. Biol.* 9: 697.
- Pearl, J. 2010. *Causality, 2nd ed.* Cambridge, MA: Cambridge University Press.
- Peter, I. S., & E. H. Davidson. 2011. *A gene regulatory network controlling the embryonic specification of endoderm*. *Nature* 474: 635–639.
- Poon, P., J. C. Flack, & D. C. Krakauer. n.d. *Niche construction and institutional switching via adaptive learning rules*. In preparation.
- Prohaska, S. J., P. F. Stadler, & D. C. Krakauer. 2010. *Innovation in gene regulation: The case of chromatin computation*. *J. Theor. Biol.* 265: 27–44.
- Sabloff, J. A. (ed.). n.d. *The rise of archaic states: New perspectives on the development of complex societies*. Santa Fe Institute. In preparation.
- Schuster, P., & W. Fontana. 1999. *Chance and necessity in evolution: Lessons from RNA*. *Physica D: Nonlinear Phenomena* 133: 427–452.
- Smith, E. 2003. *Self-organization from structural refrigeration*. *Phys. Rev. E* 68: 046114.
- Smith, E. 2008. *Thermodynamics of natural selection I: Energy flow and the limits on organization*. *J. Theor. Biol.* <http://dx.doi.org/10.1016/j.jtbi.2008.02.010>, PDF.
- Smith, E., S. Krishnamurthy, W. Fontana, & D. C. Krakauer. 2011. *Nonequilibrium phase transitions in biomolecular signal transduction*. *Phys. Rev. E* 84: 051917.
- Smith, J. M., & E. Szathmáry. 1998. *The major transitions in evolution*. Oxford, UK: Oxford University Press.
- Stadler, B. M. R., P. F. Stadler, G. Wagner, & W. Fontana. 2001. *The topology of the possible: Formal spaces underlying patterns of evolutionary change*. *J. Theor. Biol.* 213: 241–274.
- Sumpter, D. J. T. 2006. *The principles of collective animal behaviour*. *Phil. Trans. Roy. Soc. Lond. B* 361: 5–22.
- Valentine, J., & R. May. 1996. *Hierarchies in biology and paleontology*. *Paleobiology* 22: 23–33.
- Valiant, L. 2013. *Probably approximately correct*. New York: Basic Books.
- West, G. B., J. H. Brown, & B. J. Enquist. 1997. *A general model for the origin of allometric scaling laws in biology*. *Science* 276: 122–126.



# Support Complexity Science

## IT'S SIMPLE

More than 70% of our annual operating budget comes from private sources - forward-thinking individuals, corporations, and foundations. If you believe complex systems science can generate the insights that matter most for science and society, make a gift today. Visit our website at **[www.santafe.edu/support](http://www.santafe.edu/support)** and help us achieve great things.



