

# From Matter to Life

## Information and Causality

*Edited by*

SARA IMARI WALKER  
*Arizona State University*

PAUL C. W. DAVIES  
*Arizona State University*

GEORGE F. R. ELLIS  
*University of Cape Town*

 CAMBRIDGE  
UNIVERSITY PRESS

## 12 Life's Information Hierarchy

Jessica Flack

### SUMMARY

I propose that biological systems are information hierarchies organized into multiple functional space and time scales. This multi-scale structure results from the collective effects of components estimating, in evolutionary or ecological time, regularities in their environments by coarse-graining or compressing time-series data and using these perceived regularities to tune strategies. As coarse-grained (slow) variables become for components better predictors than microscopic behavior (which fluctuates), and component estimates of these variables converge, new levels of organization consolidate. This process gives the appearance of downward causation – as components tune to the consolidating level, variance at the component level decreases. Because the formation of new levels results from an interaction between component capacity for regularity extraction, consensus formation, and how structured the environment is, the new levels, and the macroscopic, slow variables describing them, are characterized by intrinsic subjectivity. Hence the process producing these variables is perhaps best viewed as a locally optimized collective computation performed by system components in their search for configurations that reduce environmental uncertainty. If this view is correct, identifying important, functional macroscopic variables in biological systems will require an understanding of biological computation. I will discuss how we can move toward identifying laws in biology by studying the computation inductively. This includes strategy extraction from data, construction of stochastic circuits that map micro to macro, dimension-reduction techniques to move toward an algorithmic theory for the macroscopic output, methods for quantifying circuit collectivity, and macroscopic tuning and control.

## INTRODUCTION

A significant challenge before biology is to determine whether living systems – composed of noisy, adaptive, heterogeneous components with only partly aligned interests – are governed by principles or laws operating on universal quantities that can be derived from microscopic processes or reflect contingent events leading to irreducible complexity (Gell-Mann and Lloyd, 1996; Goldenfeld, 1999; Krakauer and Flack, 2010; Krakauer et al., 2011; Flack et al., 2013). We know the answer to this question for physical systems and it is useful to recall that understanding in physics was achieved only after extensive debate. This debate began with the observation that certain average quantities – temperature, pressure, entropy, volume, and energy – exist at equilibrium in fundamental relationship to each other, as expressed in the ideal gas law. This observation led to thermodynamics, an equilibrium theory treating aggregate variables. When these variables were derived from first principles using statistical mechanics – a dynamical theory treating microscopic variables – the debate about whether regularities at the macroscopic scale were fundamental was partly resolved, by providing the microscopic basis for the macroscopic variables of thermodynamics, statistical mechanics established the conditions under which the equilibrium relations are no longer valid or expected to apply.

This brief summary of the relation between thermodynamics and statistical mechanics in physics is illuminating because it raises the possibility of a potentially deep division between physical and biological systems. So far, and admittedly biology is young, biology has had limited success discovering relationships among macroscopic variables and deriving these variables from first principles rooted in physical laws or deep evolved constraints. Two areas in which there has been success include metabolic scaling (West and Brown, 2005; West et al., 1997) and neural coding (Schneidman et al., 2006; Tkavcik et al., 2013). Both allometric scaling and neural coding theories help to explain how large populations of cells achieve a coordinated maximization of a group-level property. For scaling, this property is the

maximization of efficient metabolic energy use, and for neural coding the most efficient information extraction from environmental inputs. In the scaling case in particular the theory has served to validate that the macroscopic observation that mass scales with metabolic rate to the 3/4 power has a basis in mechanism – in other words the macroscopic variables can be said to be fundamental-obeying laws – rather than nominal (Krakauer and Flack, 2010), and in that sense are getting closer to temperature, pressure, entropy, etc.

Although the idea of optimization subject to simple constraints can explain a surprisingly broad range of quantifiable variation in evolved systems, from physiology to ecology, and from genetics to development (e.g., Beggs, 2008; Couzin, 2009; Frank, 2013; Shriki et al., 2013; West and Brown, 2005), the number of fundamental macroscopic variables known in biological systems remains small and limited to cases like scaling where energy plays a direct role in influencing structure. Whether there is some null expectation for how many we might expect in biological systems given the complexity of the microscopic behavior is a totally open question and perhaps not yet well posed.

## THE IMPORTANCE OF INFORMATION

Beyond the obvious heterogeneity another possible reason for the so-far limited progress identifying biological laws or principles is that biological systems are collective, computational, and involve information processing (Couzin et al., 2005; Flack, 2012, 2014; Flack and Krakauer, 2011; Flack et al., 2013; Hartwell et al., 1999; Krakauer et al., 2011; Levin, 1992; Mitchell, 2011; Vetsigian et al., 2006; von Dassow et al., 2000; Walker and Davies, 2013; Yuh et al., 1998). Examples include swarming behavior in social insects (Conradt and Roper, 2005), feature detection in the peripheral visual pathway (Olshausen and Field, 1996), and voting behavior or consensus formation in primate groups (Flack and Krakauer, 2006; Brush et al., 2013). These examples hint at limitations in the scope of application of traditional physical theories. In each of these systems the basic elements,

i.e., ants, neurons, or primates, often live in relatively small populations, are adaptive with interaction rules that are subject to learning, and have functional properties at the aggregate level that feed back to influence the decision-making rules or strategies guiding behavior.

We have proposed that in systems like these functionally important macroscopic properties arise as heterogeneous, adaptive components extract regularities from their environments to reduce uncertainty. This facilitates adaptation, thereby promoting survival and reproduction. Hence in biological systems functional macroscopic properties are *constructed* over evolutionary, developmental, and ecological time.

In our work, which we review briefly in Examples, we have shown that the collective effects of this regularity extraction can be captured with coarse-grained (or compressed) variables; endogenous coarse-graining represents the average perceived regularity in the environment at the microscopic level. As estimates of the regularities by components (as opposed to the scientist-observers) converge with exposure to larger data samples, the coarse-grained variables consolidate, providing a new effective background against which components tune strategies, and creating new space and time scales. This estimation and consensus assessment process makes such variables subjective, perhaps nonstationary, and inferential in character.

This recasting of the evolutionary process as an inferential one (Bergstrom and Rosvall, 2011; 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 and Feldman, 2001; Krakauer and Zanutto, 2009). This premise, of course, holds only 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 et al., 2013).

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

With these ideas in mind let's return for a moment to the question of a biological laws. Biological systems – from cells to tissues to individuals to societies – have nested organizational levels (e.g., as reviewed in Maynard Smith and Szathmary, 1998). These levels can be quantitatively described by their space and time scales, and each new level has associated with it some new or emergent function – a new feature with positive payoff consequences for the system as a whole or for its components (Flack et al., 2013). This hierarchical organization can be thought of as a nesting of functional encodings of lower-level regularities. As I argue in this chapter and elsewhere (Flack, 2012; Flack et al., 2013), these functional encodings form an *information hierarchy* (see also Walker and Davies, 2013) that results from biological systems manipulating space and time to reduce uncertainty, thereby facilitating efficient extraction of energy, promoting adaptation.

When macroscopic variables describing these levels are not directly tied to energetic constraints, as in the scaling case, but have a profoundly informational character – arising only as component estimates of regularities converge – they may not be obvious a priori from observation at the aggregate level. Discovery of the principles and possibly laws governing biological systems in these cases presumably requires we take information processing and the subjective nature of regularity extraction seriously.

To find the coarse-grainings favored by the system, we need to adopt its perspective. One way to do this is to proceed inductively,

working upward from the data, starting with dynamical many-body formalisms and finding, through empirically grounded simulation and modeling, equilibrium descriptions with a few favored macroscopic degrees of freedom.

Hence I am proposing that to identify fundamental information quantities of biology, we start by identifying provisional macroscopic properties thought to have functional consequences for components, but instead of next looking for equilibrium relationships among these variables, as in physics and in the scaling case, ask instead whether

1. these provisional variables can be derived from microscopic data on strategic interactions known to be important in the system, and
2. they are tunable and 'readable' by components (hence functional) individually or collectively in evolutionary or ecological time.

If we can establish that the provisional macroscopic variables satisfy these criteria, they become good candidate fundamental *biological* variables, and the search for law-like relationships among them may be more straightforward.

#### EXAMPLES

To make these ideas more concrete, let's consider some examples. The first comes from my own work on conflict management in animal societies (for a review of this work, see Flack, 2012) – specifically, third-party policing in primate groups.

Policing, a form of conflict management in which an individual breaks up fights among other individuals, is the new or emergent function. The provisional macroscopic property supporting this new function is the distribution of social power, where power is operationalized as the degree of consensus in the group that an individual can win fights (see Brush et al., 2013, and references therein). When the power structure becomes effectively institutionalized (here meaning associated with a relatively slow time scale and hard to change because in order for an individual's power to change many opinions about fighting ability need change), it becomes a good predictor of the

future cost of social interaction and provides information to the individuals about the kinds of conflict and conflict management behavior they can afford given how power is distributed. When the distribution is heavy tailed, policing, which is an intrinsically costly strategy, becomes affordable, at least to those individuals in the distribution's tail. These are the super-powerful monkeys who are rarely or never challenged when they break up fights (Flack et al., 2006, 2005).

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, 2012; Flack et al., 2013). In the case of the monkeys, a slowly changing status signaling network that sums up the outcomes of fights arises from the conflict interactions and encodes an even more slowly changing power structure. 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. So the monkeys coarse-grain over their fight histories with other individuals to figure out who in a pair will likely win the fight. The collective coarse-grained assessment, which changes yet more slowly, of who can win fights gives the consensus in the group about who has power and provides the basis for the power distribution (Brush et al., 2013).

Proteins offer another example from a very different level of biological organization. Proteins 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 pig-tailed 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 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 *slow variables* (Flack, 2012; Flack and de Waal, 2007; Flack et al., 2013; see also Feret et al., 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, or, minimally, only a temporal component, as in the monkey example.

The basic idea is that as a consequence of integrating overabundant 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 free components to search at low cost a larger space of strategies for extracting resources from the environment (Schuster and Fontana, 1999; Rodriques and Wagner, 2009; Flack and de Waal, 2007; Flack, 2012; Flack et al., 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 (e.g., Schuster and Fontana, 1999).

#### SLOW VARIABLES TO FUNDAMENTAL MACROSCOPIC PROPERTIES

As an interaction or environmental 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 – the slow variables become fundamental macroscopic properties. We speak of a new organizational level when

1. 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*

2. when the coarse-grained estimates made by components are largely in agreement (Flack et al., 2013).

The idea is that convergence on these 'good-enough' estimates underlies nonspurious correlated behavior among the components. This, in turn, leads to an increase in local predictability and drives the construction of the information hierarchy. (Note that increased predictability can give the appearance of downward causation in the absence of careful analysis of the bottom-up mechanisms that actually produced it; see also Walker and Davies, 2013). The probability of estimate convergence should increase as the sample size grows, if the computational capacities of the components are similar, and through a feedback amplification process as new organizational levels consolidate.

#### CHALLENGES

##### *Biology as Collective Computation*

If, as I am arguing, life is an information hierarchy that results from biological components collectively estimating environmental or social regularities by coarse-graining or compressing time-series data, a natural (and complementary) approach is to treat the micro and macro mapping as a computation.

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 and Krakauer, 2011; see also Mitchell, 2011; Valient, 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 time scales, with the slower time-scale processes (the slow variables) providing the 'background' against which the fitness of a component using a given strategy is evaluated (Flack and Krakauer, 2011). The idea is that the system makes a prediction based on its prior experience in this stable,

slowly changing environment about which strategy will increase its fit (measured, for example, in terms of mutual information or fitness) to the environment and tunes its behavior to implement the strategy.

As an example, consider the distribution of social power discussed above. Because the DSP is slowly changing (compared with the interaction and status signaling rates at the individual level), it provides a stable background against which the monkeys can "predict the future cost of interaction" and hence tune their behavioral strategies. It is this ability to predict, derived from the degree of time-scale separation between the power distribution and underlying distribution of fighting abilities, that reveals when the computation is "correct." Here, in contrast to computer science, the output is continuously computed and the notion of correctness comes from the utility of the output for prediction.

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, is the result of a distributed and coordinated sequence of component interactions under the operation of a strategy set, and is a stable output of input values that converges (terminates) in biologically relevant time (Flack and Krakauer, 2011). Examples studied in biology include aspects of vision, such as edge detection (Olshausen and Field, 1996); phenotypic traits, such as the average position of cells in the developing endomesoderm of the sea urchin (e.g., Peter and Davidson, 2011); switching in biomolecular signal-transduction cascades (e.g., Smith et al., 2011); and social structures, such as the distribution of fight sizes (e.g., DeDeo et al., 2010; Flack and Krakauer, 2011) and the distribution of power in monkey societies (e.g., Brush et al., 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 a well-studied phenotypic trait such as the development

of a 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.

In a biological system, the input plus the strategies constitute the system's microscopic behavior. There are many approaches to reconstructing the system's microscopic behavior from raw data. The most powerful is an experiment in which upstream inputs to a target component are clamped off and the output of the target component is held constant. This allows the experimenter to measure the target component's specific contribution to the behavior of a downstream component.

When such experiments are not possible, causal relationships can be identified using time-series analysis in which clamping is approximated statistically. My collaborators and I have developed a novel computational technique, called *inductive game theory* (IGT) (DeDeo et al., 2010; Flack and Krakauer, 2011), that uses a statistical clamping principle to extract strategic decision-making rules, game structure, and (potentially) strategy cost from time-series data. (IGT is one of many approaches being developed in a growing body of literature on causal network reconstruction from time-series and correlation data.)

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 work on echinoderm gene regulatory networks controlling embryonic cell position (the output) in the sea urchin (Peter and Davidson, 2011). In the case of our work on micro to macro mappings in animal societies, we describe the space of

microscopic configurations – fight decision-making rules – using Markovian, probabilistic, social circuits (DeDeo et al., 2010; Flack and Krakauer, 2011).

Nodes in the gene regulatory circuits and social circuits described above are the input to the computation. As discussed above, the input can be genes, neurons, individuals, subgroupings of components, etc. A directed edge between two nodes in the circuit indicates that the “receiving node” has a strategy for the “sending node,” and the edge weight can be interpreted as the 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 metastrategies can be captured in a circuit using different types of gates specifying how a component's myriad strategies combine (DeDeo et al., 2010; Flack and Krakauer, 2011; see also Feret et al., 2009). By varying the types of gates and/or strength of allowed 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. 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. By testing the empirically parameterized circuits against each other in simulation we can determine which best recovers the actual measured macroscopic behavior of the study system and in this way discover if our provisional macroscopic variable may indeed be a candidate fundamental variable.

#### *Circuit Logic*

The circuits describing the microscopic behavior can be complicated, with many ‘small’ causes detailed, as illustrated by the gene regulatory circuits constructed by Eric Davidson and his colleagues.

The challenge once we have rigorous circuits is to figure out the circuit logic.

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 biological systems composed of adaptive components require an additional criterion beyond compactness. As discussed earlier in this chapter, components in these systems are tuning their behaviors based on their own effective theories – coarse-grained (or compressed) rules (see also Feret et al., 2009) that capture the regularities they perceive. 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 average system behavior*, as these models guide component strategy choice. In other words, our effective theory should explain how the system itself is computing (see also Walker and Davies, 2013).

My collaborators and I begin the search for cognitively principled, *algorithmic* effective theories using what we know about component cognition to inform how we coarse-grain and compress the circuits (Daniels et al., 2012). 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 et al., 2010).

#### *Information and Energy in Biology*

As discussed above, once we have a satisfactory family of candidate effective theories for how our system is computing its macroscopic output, we need to choose from among them the one that best recovers our observable and is also the most mechanistically principled. Generally we require two criteria be met to claim a model (or circuit) is mechanistically principled. The model must capture



causal relationships supported by the data – how individual strategies or some simplification of them actually combine to produce the output. The second criterion is that the model must be cognitively or computationally parsimonious. This requires knowing something about the cognitive or computational burden that each of the models assumes of the system and its components (Krakauer et al., 2010). In other words, we need to be able to measure the number of bits required to parameterize each model for a given level of performance, and for most systems we study in biology the number of allowable (given component computational capacity) bits should (probably) be relatively small.

#### The Informational Cost of Biological Computation

Calculating for a reasonably sized data set the number of bits required to perform the computation is generally achievable and allows comparison across the models within the study, but it is not at all clear what it means in an absolute sense to say that a particular theory or compressed representation of the system behavior requires  $x$  number of bits to encode.

For example, in our work on monkey conflict dynamics (Daniels et al., 2012), we found that about 1,000 bits of information are required to encode which individuals and subgroups are regular and predictable participants in fights, assuming a sparse coding algorithm. Our other models performed worse (required more bits). In this sense our bits measurement was a useful bar against which to compare models, but we cannot yet claim to have any idea whether a model requiring 1,000 bits is reasonable given our subjects' cognitive capacities. Establishing this requires an experimental approach.

Another open question includes how the informational cost of computation changes when the computation is collective. Work on robustness (e.g., Ay et al., 2007) and distributed computation suggests that reliability of the output may increase, but this work does not explicitly address how variance in the output at the component level affects the number of bits required to encode the collective

computation. Hence the output may be more reliable, but the total informational cost could be much higher. This may not matter if there are strong, shared constraints on the component computations and/or the tuning is only at the component level.

#### Bits to Joules and the Energetic Cost of Computation

We understand the relationship between energy and information in the limit, as Landauer's principle tells us that there is a minimum amount of energy required to erase one bit of information (Bennett, 2003). However, even if we could in a compelling model-free way quantify the number of bits required to encode a model or make a decision, we have no idea how bits translate into watts in the adaptive, stochastic, information-processing, many-body systems of biology and society, or how this question could be approached empirically.

Yet from the perspective of evolutionary theory it seems likely that information processing is adaptive – meaning it allows biological systems to more efficiently, given constraints, extract energy and do the work required to promote survival and reproduction. Another way of putting this is that even if physical theory says information-processing is energetically costly (Parrondo et al., 2015), evolutionary theory suggests that, in the long run, or given an understanding of the full set of constraints to which a biological system is subject, information-processing *saves* watts. Unpacking this proposition may be the key to go from the information-processing mechanisms producing slow variables and emergent function to the identification of biological laws.

#### COLLECTIVE COMPUTATION TO STRATEGIC STATISTICAL MECHANICS FOR MANY-BODY, ADAPTIVE SYSTEMS

As understanding of the micro-macro mapping is refined through identification of cognitive effective theories that parsimoniously reduce circuit complexity and compactly encode the macroscopic output, we also refine our understanding of the natural scales

of the system. This includes distinguishing strategic microscopic behavior from noise, and hence allows us to extract from (rather than imposing on) the raw data the building blocks of our system. And by investigating whether our best-performing empirically justified circuits can also account for other functionally important macroscopic properties, we can begin to establish which macroscopic properties might be biologically fundamental and whether they stand in law-like relation to one another. One can think of this approach as a *strategic statistical mechanics*, embedding complex decision rules in formalisms for calculating emergent properties and discovering law-like behavior at the aggregate level.

These ideas are very closely related to the pioneering ideas of John von Neumann (D. C. Krakauer, personal communication), who in the 1940s and 1950s began the development of a statistical mechanics of biologically inspired sensing and computing devices (von Neumann, 1987). He writes (von Neumann, 1954):

anything that can be exhaustively and unambiguously described, anything that can be completely and unambiguously put into words, is *ipso facto* realizable by a suitable finite neural network ... we get an image of the strong limitations that our sensations, our intuitions, our logic and our language have to obey. We can put all these things in a more complete statement: The following restrictions are mutually equivalent: to be macroscopic; to be Euclidean (i.e. to adopt the parallel axiom in the way we represent space and spatial relations); to be Galileo-Newtonian in the way we represent motion, time and energy; to capture the surrounding and to act according to our sensorial-intuitive perception of reality; to use and to represent language, in both its natural and artificial variants.

Where von Neumann speaks of neural networks constrained by sensorial perception of classical observations, we consider (D. C. Krakauer, personal communication) how strategic, behavioral rules (de Waal, 1991; Flack et al., 2004), when combined into stochastic

circuits expressed in the language of Markov decision processes, produce and respond to coarse-grained aggregate information through near-critical system states. In less technical language, we explore chains of probabilistic events – decisions or state transitions – that generate and respond efficiently to average features of the world. This enables these systems to tune adaptively to the needs of their social environment. My collaborators and I believe that to develop a statistical mechanics that can accommodate these kinds of adaptive, strategic systems we will need to extend existing physical theories by incorporating ideas from theoretical computer science, information theory, and evolutionary biology.

#### ACKNOWLEDGMENTS

These ideas were developed in collaboration with David Krakauer, who would no doubt state them some other way. Many of the ideas also developed during research projects and conversations with our postdocs and students – Bryan Daniels, Chris Ellison, Eleanor Brush, Eddie Lee, Simon DeDeo, Philip Poon, and Jess Banks – and with SFI colleagues – Eric Smith, Geoffrey West, Jim Crutchfield, and, especially, Walter Fontana. I also thank Sara Walker and Paul Davies for organizing the meeting on information that triggered this chapter. I acknowledge support from two grants to the Santa Fe Institute (SFI) from the Templeton Foundation to study complexity, a grant from the Templeton Foundation to the SFI to study the mind-brain problem, and NSF grant no. 0904863.

#### REFERENCES

- Ay, N., Flack, J. C., and Krakauer, D. C. 2007. Robustness and complexity co-constructed in multi-modal signaling networks. *Philosophical Transactions of the Royal Society, London, Series B*, 362, 441–447.
- Beggs, J. M. 2008. The criticality hypothesis: how local cortical networks might optimize information processing. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 366(1864), 329–343.
- Bennett, C. H. 2003. Notes on Landauer's principle, reversible computation, and Maxwell's demon. *Studies in the History and Philosophy of Modern Physics*, 34, 501–510.

- Bergstrom, C. T., and Rosvall, M. 2011. The transmission sense of information. *Biology and Philosophy*, 26, 159–176.
- Brush, E. R., Krakauer, D. C., and Flack, J. C. 2013. A family of algorithms for computing consensus about node state from network data. *PLOS Computational Biology*, 9(7), e1003109.
- Conradt, L., and Roper, T. J. 2005. Consensus decision making in animals. *Trends in Ecology & Evolution*, 20(8), 449–456.
- Couzin, I. D. 2009. Collective cognition in animal groups. *Trends in Cognitive Sciences*, 13(1), 36–43.
- Couzin, I. D., Krause, J., Franks, N. R., and Levin, S. A. 2005. Effective leadership and decision-making in animal groups on the move. *Nature*, 433(7025), 513–516.
- Crutchfield, J. P., and Feldman, D. P. 2001. Synchronizing to the environment: information-theoretic constraints on agent learning. *Advances in Complex Systems*, 4, 251–264.
- Daniels, B. C., Krakauer, D. C., and Flack, J. C. 2012. Sparse code of conflict in a primate society. *Proceedings of the National Academy of Sciences*, 109(35).
- De Waal, F. B. M. 1991. The chimpanzee's sense of social regularity and its relation to the human sense of justice. *American Behavioral Scientist*, 34, 335–349.
- DeDeo, S., Krakauer, D. C., and Flack, J. C. 2010. Inductive game theory and the dynamics of animal conflict. *PLOS Computational Biology*, 6(5), e1000782.
- Feret, J., Danos, V., Krivine, J., Harmer, R., and Fontana, W. 2009. Internal coarse-graining of molecular systems. *Proceedings of the National Academy of Sciences of the United States of America*, 106(16), 6453–6458.
- Flack, J. C. 2012. Multiple time-scales and the developmental dynamics of social systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1597), 1802–1810.
- Flack, J. C. 2014. Life's information hierarchy. *Santa Fe Institute Bulletin*, 28, 13.
- Flack, J. C., and de Waal, F. 2007. Context modulates signal meaning in primate communication. *Proceedings of the National Academy of Sciences of the United States of America*, 104(5), 1581–1586.
- Flack, J. C., and Krakauer, D. C. 2006. Encoding power in communication networks. *American Naturalist*, 168(3).
- Flack, J. C., and Krakauer, D. C. 2011. Challenges for complexity measures: a perspective from social dynamics and collective social computation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 21(3), 037108.
- Fluck, J. C., Jeannotte, L. A., and de Waal, F. B. M. 2004. Play signaling and the perception of social rules by juvenile chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 118, 149–159.
- Flack, J. C., de Waal, F. B. M., and Krakauer, D. C. 2005. Social structure, robustness, and policing cost in a cognitively sophisticated species. *American Naturalist*, 165(5).
- Flack, J. C., Girvan, M., de Waal, F. B. M., and Krakauer, D. C. 2006. Policing stabilizes construction of social niches in primates. *Nature*, 439(7075), 426–429.
- Flack, J. C., Erwin, D., Elliot, T., and Krakauer, D. C. 2013. Timescales, symmetry, and uncertainty reduction in the origins of hierarchy in biological systems. *Cooperation and Its Evolution*, 45–74.
- Frank, S. A. 2013. Input-output relations in biological systems: measurement, information and the Hill equation. *Biology Direct*, 1–25.
- Gell-Mann, M., and Lloyd, S. 1996. Information measures, effective complexity, and total information. *Complexity*, 2(1), 44–52.
- Goldenfeld, N. 1999. Simple lessons from complexity. *Science*, 284(5411), 87–89.
- Hartwell, L. H., Hopfield, J. J., Leibler, S., and Murray, A. W. 1999. From molecular to modular cell biology. *Nature*, 402, C47–C52.
- Krakauer, D. C. 2011. Darwinian demons, evolutionary complexity, and information maximization. *Chaos*, 21, 037110.
- Krakauer, D. C., and Flack, J. C. 2010. Better living through physics. [Letter to the Editor]. *Nature*, 467(7316), 661.
- Krakauer, D. C., and Zanotto, P. 2009. Viral individuality and the limitations of life concept. Pages 513–536 of Rasmussen, M. A., et al. (eds.), *Protocells: bridging non-living and living matter*. MIT Press.
- Krakauer, D. C., Flack, J. C., Dedeo, S., Farmer, D., and Rockmore, D. 2010. Intelligent data analysis of intelligent systems. Pages 8–17 of Cohen, P. R., Adams, N. M., and Berthold, M. R. (eds), *Advances in intelligent data analysis IX*. Springer.
- Krakauer, D. C., Collins, J. P., Erwin, D., Flack, J. C., Fontana, W., Laubichler, M. D., Prohaska, S. J., West, G. B., and Stadler, P. F. 2011. The challenges and scope of theoretical biology. *Journal of Theoretical Biology*, 276(1), 269–276.
- Levin, S. A. 1992. The problem of pattern and scale in ecology. *Ecology*, 73(6), 1943.
- Maynard Smith, J., and Szathmari, E. 1998. *The major transitions in evolution*. Oxford University Press.
- Mitchell, M. 2011. What is computation? Biological computation. *Ubiquity*, February, 1–7.
- Olshausen, B. A., and Field, D. J. 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381, 607.
- Parrondo, J. M. R., Horowitz, J. M., and Sagawa, T. 2015. Thermodynamics of information. *Nature Physics*, 11, 131–139.
- Peter, I. S., and Davidson, E. H. 2011. A gene regulatory network controlling the embryonic specification of the endoderm. *Nature*, 474, 635–639.
- Rodrigues, J. F. M., and Wagner, A. 2009. Evolutionary plasticity and innovations in complex metabolic reaction networks. *PLOS Computational Biology*, 5, e1000613.
- Schneidman, E., Berry, M. J., Segev, R., and Bialek, W. 2006. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, 440(7087), 1007–1012.

- Schuster, P., and Fontana, W. 1999. Chance and necessity in evolution: lessons from RNA. *Physica D: Nonlinear Phenomena*, 133, 427–452.
- Shriki, O., Alstott, J., Carver, F., Holroyd, T., Henson, R. N. A., Smith, M. L., Coppola, R., Bullmore, E., and Plenz, D. 2013. Neuronal avalanches in the resting MEG of the human brain. *Journal of Neuroscience*, 33(16), 7079–7090.
- Smith, E., Krishnamurthy, S., Fontana, W. D., and Krukauer, D. C. 2011. Nonequilibrium phase transitions in biomolecular signal transduction. *Physical Review E*, E84, 051917.
- Tkavcik, G., Marre, O., Mora, T., Amodei, D., Berry, M. J. II, and Bialek, W. 2013. The simplest maximum entropy model for collective behavior in a neural network. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(03), P03011.
- Valient, L. 2013. *Probably approximately correct*. Basic Books.
- Vetsigian, K., Woese, C., and Goldenfeld, N. 2006. Collective evolution and the genetic code. *Proceedings of the National Academy of Sciences of the United States of America*, 103(28), 10696–10701.
- Von Dassow, G., Meir, E., Munro, E. M., and Odell, G. M. 2000. The segment polarity network is a robust developmental module. *Nature*, 406(6792), 188–192.
- Von Neumann, John. 1954. The general and logical theory of automata. In Taub, A. H. (ed.), *John von Neumann: collected works*. Pergamon Press.
- Von Neumann, John. 1987. *Papers of John von Neumann on computing and computer theory*. MIT Press.
- Walker, S. I., and Davies, P. C. W. 2013. The algorithmic origins of life. *Journal of the Royal Society, Interface*, 10(December 2012), 20120869.
- West, G. B., and Brown, J. H. 2005. The origin of allometric scaling laws in biology from genomes to ecosystems: towards a quantitative unifying theory of biological structure and organization. *Journal of Experimental Biology*, 208(Pt 9), 1575–1592.
- West, G. B., Brown, J. H., and Enquist, B. J. 1997. A general model for the origin of allometric scaling laws in biology. *Science*, 276(5309), 122–126.
- Yuh, C., Bolouri, H., and Davidson, E. H. 1998. Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene. *Science*, 279(5358), 1896–1902.

## 13 Living through Downward Causation

### From Molecules to Ecosystems

Keith D. Farnsworth, George F. R. Ellis, and Luc Jaeger

Downward causation (first defined by Campbell, 1974) is both a philosophical concept and an apparent phenomenon of nature attracting great controversy. Most scientists usually assume that all observable phenomena derive from elemental fundamental physics, so that even human behaviours ultimately result from interactions of subatomic particles, via a unidirectional chain of causes and effects. On closer inspection, the act of living seems able to spontaneously generate events, breaking this chain; it is as though life possessed ‘free will’ by acting without a prior physical cause. In this chapter, we analyse this puzzling behaviour using information and control theory as a general framework, applying it to a range of scales of organisation in biological systems: from the molecular to the ecological. An essential element (and possibly a defining feature) of life emerges from this analysis. It is the presence of downward causation by information selection and control. Through a series of examples, we show how this phenomenon works to produce the appearance of autonomous action from information constructed and maintained by the process of living. After a brief introduction to the concept of downward causation, we set it more firmly within the concepts of biological information processing used within this volume. From this we attempt to derive a general classification of causation across scales of biological organisation. We show how selection from random processes and information embodiment in molecules, organism systems, and ecological systems combine to emerge with the properties of downward causation and the appearance of autonomy. These phenomena seem to be exclusive to life.