When Evolution is Revolution —Origins of Innovation

James P. Crutchfield
Santa Fe Institute
1399 Hyde Park Road
Santa Fe, New Mexico 87501
chaos@santafe.edu

A pictorial tour of the theories of epochal evolution and structural complexity is presented with a view toward the dynamical origins, stabilization, and content of evolutionary innovations. A number of alternative explanations for the occurrence of long periods of stasis that are interrupted by sudden change have been proposed since the first days of mathematical evolutionary theory. Here contrasts are drawn between the mechanisms underlying epochal evolution and those implicated in the classical theory of stochastic intermittency (drift) due to Fisher, Wright's adaptive landscapes, Kimura's neutral evolution, and Gould and Eldredge's notion of punctuated equilibria. The comparisons suggest what a synthetic theory of the evolution of complexity might look like, while at the same time emphasizing that it will remain incomplete without a theory of biological structure. The computational mechanics theory of structural complexity is offered as an approach to the latter.

1 EPOCHAL EVOLUTION AND INNOVATION

The emergence of biological form and function through evolution is often considered to happen by a process of gradual adaptation: through a series

Evolutionary Dynamics—
Exploring the Interplay of Selection, Neutrality, Accident, and Function
edited by J. P. Crutchfield and P. Schuster, 2001

of small changes observable features and improved behaviors appear. When niches alter in character or when a species first moves into a existing niche, the context of prior diversity is changed and relative fitnesses in a population adjust. Selection then acts to reshape the cloud of diverse individuals in directions appropriate to the new environment. The diversity of individuals is the expression of genetic variation—variations whose origins are not correlated with individual fitness. Darwin's analysis of Galapagos finches is the paradigmatic case: The diversity in beak length and shape was seen as reflecting incremental adaptations to small geographic variations in type of food source [14].

This has given rise to a view of evolutionary dynamics as an optimization process: The environment provides constraints and species either go extinct or are able to incrementally change in ways that take advantage of or mitigate the constraints. When there is this kind of tight coupling between organism and environment and when the time scales of change allow adaptation, the form and behavior of the resulting organisms mirror niche structure and environmental constraint. Moreover, individual biological traits take on functional meaning, since they reflect the "solutions" to "problems" imposed by the environment. Finally, in this view the environment is a source of novelty and the instigator of change. Though it can react, Darwinian evolution cannot, in and of itself, produce novel biological structures and functions.

For well over a half century, however, it has been known that gradual adaptation is a substantially incomplete picture of evolutionary dynamics. Early mathematical analyses of stochastic processes, the rise of molecular genetics, investigations of the fossil record, the development of nonlinear population dynamics, and recent laboratory evolutionary experiments reveal that evolution need not be gradual, but can be episodic. Perhaps the most extreme examples are seen in *evolutionary metastability*: Long periods of stasis are interrupted by rapidly emerging innovations. Importantly, evolutionary metastability loosens the coupling between individual diversity and adaptive response to the environment: There can be substantially more individual diversity than adaptation to the environment requires. Another consequence is that this loose coupling opens up the possibility that evolutionary dynamics can produce novel structures on its own, not only in lock-step response to environmental change. Unfortunately, current explanations of metastability do not define what biological structure is and so are not yet complete theories of the innovation of novel form and function.

One of the earliest recognitions of metastability in evolutionary dynamics is Fisher's analysis of *stochastic intermittency* in multi-allele drift processes [19]. He showed that in the absence of selection, when only drift was operating, there could be a transient fixation on one or another allele and that when a shift to a new allele occurred, the transition came quickly—compared to the time scale of allele fixation. (This is illustrated in figure 1.) Being a fundamental property of random finite-sample processes, stochastic intermittency

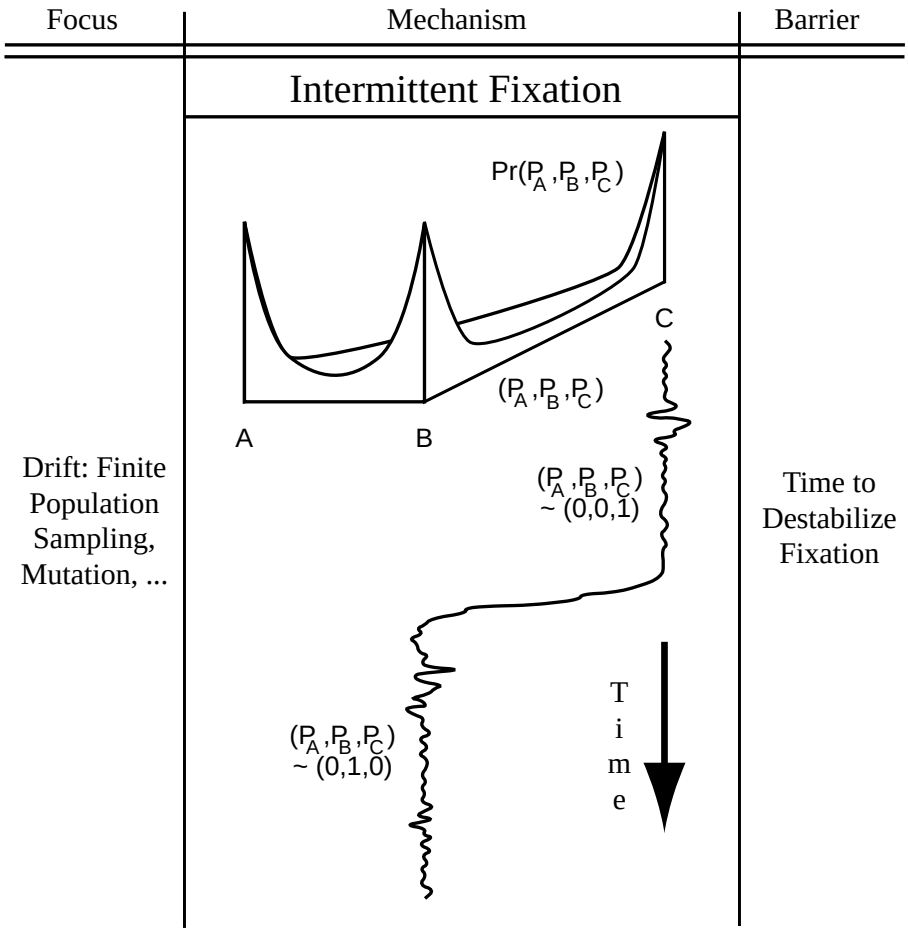

FIGURE 1 Metastability through intermittent fixation: A population is described by the proportion $(P_{\mathrm{A}}, P_{\mathrm{B}}, P_{\mathrm{C}})$ of individuals with one of three alleles A, B, and C. The probability $\Pr(P_{\mathrm{A}}, P_{\mathrm{B}}, P_{\mathrm{C}})$ of the population exhibiting proportion $(P_{\mathrm{A}}, P_{\mathrm{B}}, P_{\mathrm{C}})$ is highly peaked at the pure populations: $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$. Nonetheless, there is some (low) probability of being in intermediate, "mixed" populations and so transitions between the pure populations are possible. Interestingly, the transitions when they occur, occur rapidly.

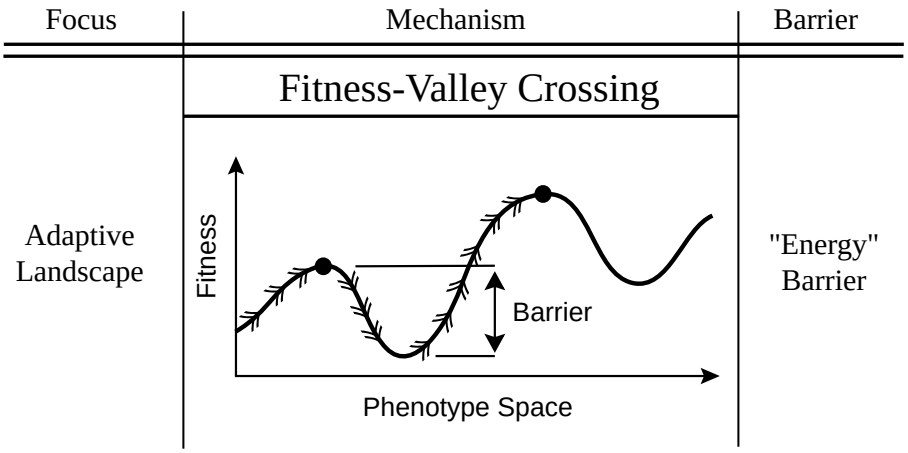

FIGURE 2 Metastability through fitness-valley crossing: A population residing at a local optimum in a fitness landscape must cross through a valley of lower fitness in order to reach another, possibly higher-fitness, peak. The landscape is defined over a space of phenotypes or traits.

can occur at a number of different levels in an evolutionary process: e.g., at both the genotype and phenotype (character) levels [4].

Another description of metastability is found in Wright's early attempt to explain the dynamics of evolutionary change. Wright introduced the notion of *adaptive landscapes* to describe the (local) stochastic adaptation of populations to environmental constraints [70]. This geographical metaphor has had a pervasive influence on theorizing about natural and artificial evolutionary processes. The basic picture is that of a gradient-following dynamics moving over a "landscape" determined by a fitness "gravitational potential". Adaptive landscapes admit two kinds of (related) metastability. First, in *fitness-valley crossing* an evolving population stochastically crawls along a surface determined, perhaps dynamically, by the fitness of individuals, moving to peaks and very occasionally hopping across fitness "valleys" to nearby, and possibly higher fitness, peaks. (See figure 2.) The *barriers* to innovation here are determined by the depth of the valley intervening between two peaks. Due to this, they are sometimes referred to as "energy" barriers, highlighting the physical metaphor. Second, in the *shifting balance* theory periods of stasis correspond to times when populations are isolated at local optima in the landscape, as before. Innovations, however, correspond to populations adapting in response changes in stability of landscape extrema—changes that are initiated by exogenous forces (e.g., environmental) and that alter the locations of peaks and valleys. (See figure 3.) In the shifting balance theory the barriers to innovation are determined by the time scale of behaviors largely external to the evolutionary process.

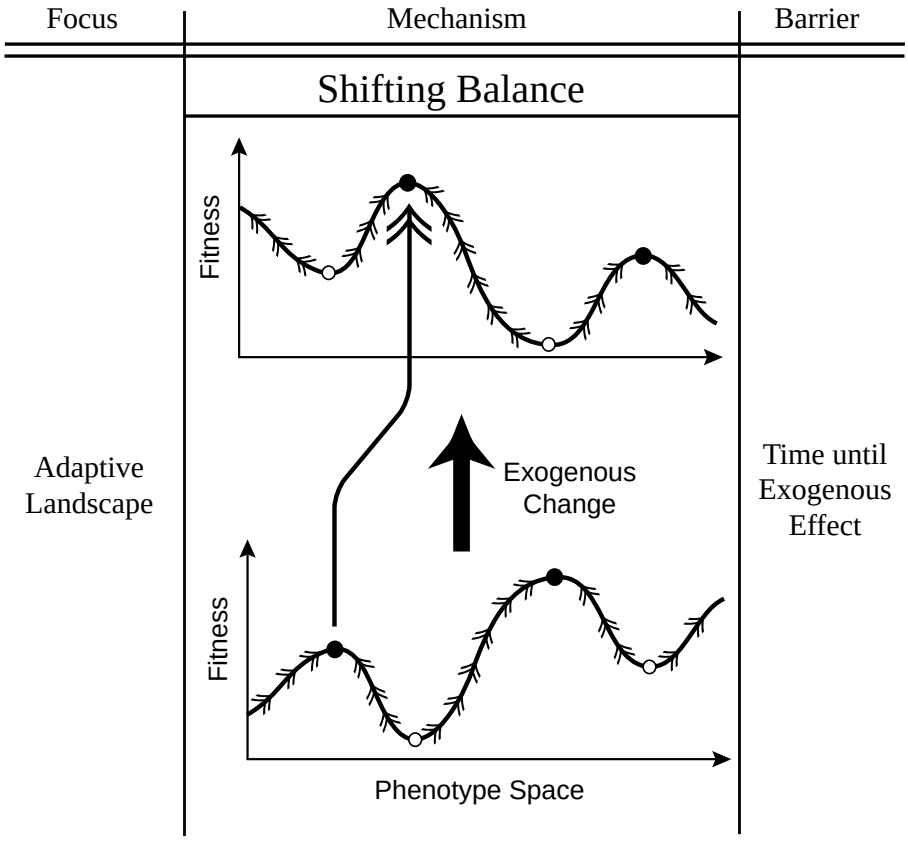

FIGURE 3 Metastability through shifting balance: A population resides at a local optimum in an adaptive landscape. The landscape is defined over a space of phenotypes or traits. An exogenous (e.g., environmental) change occurs that alters the shape of the adaptive landscape in such a way that the population's local optimality disappears and its stability is lost. The population then climbs to a neighboring peak.

More recently, extending Wright's notion of adaptive landscapes, it has been proposed that the processes underlying combinatorial optimization and biological evolution can be modeled as "rugged landscapes" [38, 44]. These are landscapes with wildly fluctuating fitnesses even at the smallest scales of single-point mutations. It is generally assumed that these "landscapes" possess a large number of local optima. With this picture in mind, the common interpretation of stasis and change in evolving populations is that of a population being "stuck" at a local peak, until a rare mutant crosses a valley of relatively low fitness to a higher peak; a picture more or less consistent with Wright's.

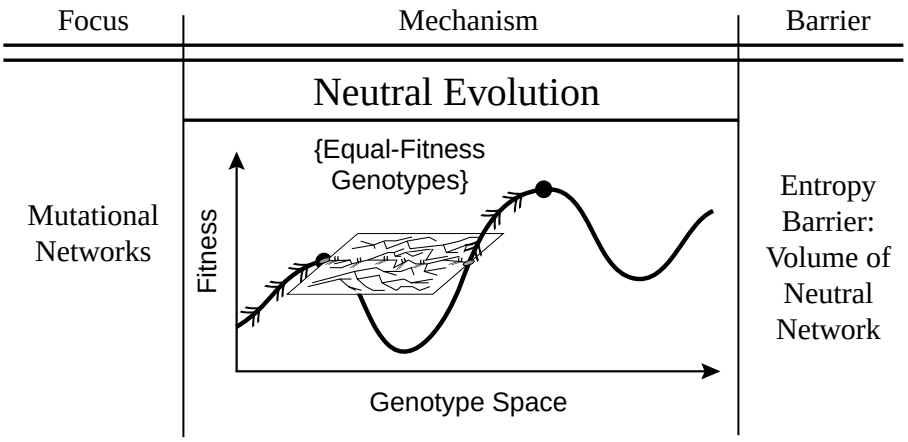

FIGURE 4 Metastability through neutral evolution: A population resides at a local optimum in a fitness landscape—here defined over the space of genotypes. Rather than passing through a valley of lower-fitness genotypes, driven by genetic variation (e.g., mutational) the population diffuses over a network of equal-fitness genotypes, until a higher-fitness genotype is found that leads to a new peak. While diffusing over the neutral network, the population's average fitness does not change. There is phenotypic stasis during a period of relatively rapid genotypic variation.

Metastability also occurs when genetic variations do not produce changes in fitness. Since selection cannot act on those variations, there can be long periods of phenotypic constancy, despite the accumulation of substantial genomic change. The history of this idea—*neutral evolution*—goes back to Kimura [39, 53], who in the 1960s argued that on the genotypic level, most genetic variation occurring in evolution is adaptively neutral with respect to the phenotype. In this situation, many genotypes code for single phenotypes. Additionally, due to intrinsic or even exogenous variations (e.g., environmental fluctuations on relatively fast time scales), there simply may not exist a deterministic "fitness" value for each possible genotype. In this case, fluctuations induce variations in fitness such that genotypes with similar average fitness are not distinct at the level of selection. Differences in fitness are simply washed out and selection cannot act on them. Thus, metastability can be induced either by many genotypes coding for a given phenotype or by "noise" in the fitness evaluation of individuals. (See figure 4.)

Today, the occurrence of neutral evolution is supported by a large and increasing body of evidence that there are substantial degeneracies (many-to-oneness) in genotype-to-phenotype and phenotype-to-fitness mappings. Neutrality has been implicated in the evolutionary optimization methods [9, 65] and the evolution of RNA structure [15, 22, 23, 24, 31, 32, 34, 35], protein structure [2, 36], and ribozymes [41, 69]. When degeneracies in the genotype-to-fitness map are operating, a large number of different genotypes in a pop-

ulation fall into a relatively small number of distinct fitness classes with approximately equal fitness, resulting in metastable evolution.

Probably the best known example of evolutionary metastability, though, is the *punctuated equilibria* behavior attributed to macroevolutionary processes by Gould and Eldredge [29]. They proposed punctuated equilibria to explain the observation in the fossil record of long periods of (morphological) constancy, which are interrupted by relatively short bursts of change, and so argued that gradual adaptation was inadequate. Although exact mechanisms supporting the metastable periods were not analyzed, the causes of punctuations were thought to originate typically in the environment, such as in planetwide climatic change [28].

The fossil record, however, is not amenable to experimental testing. Fortunately, new experimental techniques for the study of bacterial evolution have led to controllable laboratory model systems with sufficiently short replication times that evolution can now be observed in detail over many thousands of generations [42]. These systems promise to yield the detailed and extensive data required for testing theories of evolutionary dynamics. In fact, recently Lenski and collaborators have reported punctuated-equilibrium-like behavior in the evolution of *E. coli* cell size—a proxy for fitness [16]. Even more recently, a genetic analysis of individuals taken from populations during the periods of stasis showed that there was substantial genetic variation—changes that were not phenotypically expressed [54]. Bacterial evolution appears to be a relatively clear and testable case of evolutionary metastability.

Metastability in artificial evolution has been observed in simulation studies of the population dynamics of machine-language programs [56]. In these studies, programs compete for memory and processing resources, replicate by copying themselves, and mutate when errors in copying occur. By directly observing changes in program structure and also by monitoring average replication rate—both of which are straightforward in simulation models, unlike biological experiment and the fossil record—periods of stasis and sudden change were observed over the course of many thousands of generations [1, 56].

There has also been a substantial amount of simulation and theoretical work recently on evolutionary search and optimization processes which exhibit metastability. One thread of this was directed at testing conjectures about evolution's ability to collect together functional "gene" groups by preferentially assembling *building blocks* or partial solutions [50]. In addition to concluding that building-block assembly was not responsible for the evolution of optimal solutions, it was discovered that the evolutionary search dynamics was not a gradual optimization process. Rather, it was dominated by periods of stasis and sudden change [65]. See figure 5 for an example run of a simple evolutionary algorithm that searches a space of binary strings for one with the largest number of functional gene groups.

Similar kinds of evolutionary metastability have been investigated in some detail in alternative models—"rugged landscapes" and others—by using dis-

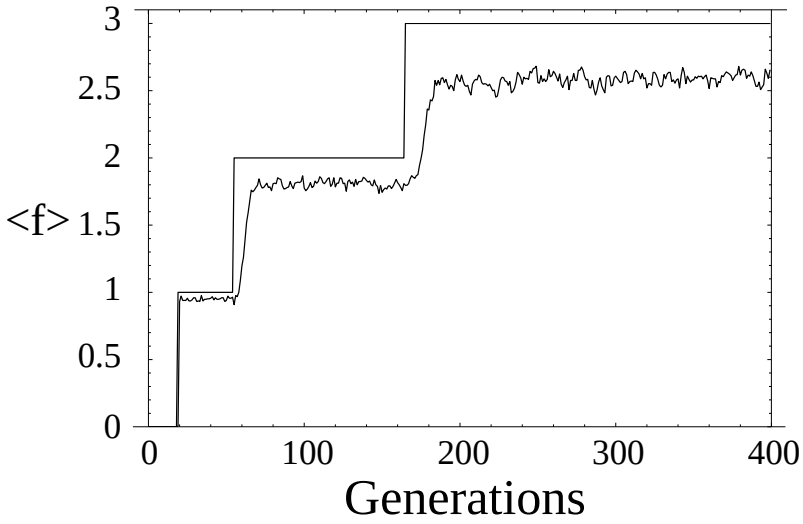

FIGURE 5 Macroscopic view of epochal evolution: survival dynamics, the level on which selection operates. Behavior of the average fitness $\langle f \rangle$ (lower curve) and best fitness (upper curve), for a population of individuals with 30 (binary) genes. They were evolved under a fitness function with three gene constellations, each consisting of 10 (binary) genes. *Constellations* are functional groups of genes that must all be properly set for a genotype to increase in fitness by one unit. The evolving population consisted of 250 individuals that at each generation were selected to replicate in proportion to their fitness and then mutated with probability 0.005 per gene. The fitness starts at 0 (no constellations set properly) and increases in a series of steps to a maximum fitness of 3 (all three constellations properly set). (Reprinted with permission from [66].)

crete, rather than continuous, fitnesses in order to produce fitness plateaus over genotype space [3, 24, 25, 52].

Thus, it appears there is no shortage of examples, from basic theory and simulation to field and laboratory data, of metastable evolution—a behavior quite different from that implied by the view of evolution as gradual adaptation. Aside from the overt behavioral differences in the population dynamics, these cases indicate that evolution, on its own, can generate change and novelty. Except in the cases of shifting balance and punctuated equilibria, there are no appeals to environmental pressures that drive innovations. Moreover, in some of these cases, it has been reported that innovations do not lead to improved structures or functionality, calling into question functional ascriptions for evolutionary innovations.

These case studies of evolutionary metastability do not attempt to explain how novel form and function arise nor do the theories quantify form and functional change. One response to these concerns is found in early work

on the artificial evolution of computation. There the author introduced the phrase *epochal evolution* to describe the stepwise emergence of sophisticated strategies observed when evolving cellular automata to perform spatial computational tasks [49]. Why invent a new descriptor and not call these these stages punctuated equilibria or not label them by one of the other alternatives? First, the theory of punctuated equilibria was introduced to describe the fossil record—a manifestly richer and more complex process than the artificial evolution of computational models. Second, it was clear that the evolutionary stages were due neither to Fisher's intermittent fixation nor to pinning at local optima. Third, more fundamentally it was important to not prejudice the analysis of the mechanisms driving the evolving cellular automata population dynamics. Now, however, based on the analyses of the cellular automata evolutionary dynamics, a theory of epochal evolution has been developed and the generality of the underlying mechanisms is better appreciated. Since there is little chance of confusion, we now refer to the examples of metastability given above as epochal evolution and ask which combinations of its constituent mechanisms produce the observed behaviors of stasis and rapid innovation in various cases.

The goals in the following discussion are two-fold. The first motivation is to provide an accessible tour, augmented by illustrations, of recent theoretical results developed by Erik van Nimwegen and the author on the origins of metastability in evolutionary dynamics. The overview focuses on the central mechanisms underlying epochal evolution, leaving out the mathematical theory [11, 64, 65, 66]. (See van Nimwegen's thesis [62] for a detailed development.) The second motivation is to connect these ideas, which fall largely in the domain of mathematical population dynamics, with a parallel project on quantifying organization and structural complexity in natural systems, which falls largely in the domains of statistical physics and dynamical systems theory [7, 12, 58]. At the end, the discussion returns to compare epochal evolution to the various alternative mechanisms mentioned above and to the current doctrines of evolutionary theory. One conclusion drawn from the comparison is that at present the various evolutionary theories do not offer a mathematical basis on which to analyze the emergence of biological form and function. Thus, the ultimate goal, suggested by juxtaposing epochal evolution and structural complexity, is to knit the two threads of innovation and complexity together to build a predictive theory—an evolutionary mechanics [8]—of the emergence of novel structure.

2 STATISTICAL DYNAMICS OF EPOCHAL EVOLUTION

What, if any, are the common mechanisms that can explain the examples of epochal evolution given above? How are we to begin understanding the general process of epochal evolution? It turns out that answering these questions requires comparing, contrasting, and analyzing three different views of evolu-

tion: its appearance in genotype space, in phenotype space, and in a functional (fitness) space. (We have already seen an example of the latter in figure 5.) Comparing and contrasting these spaces is the burden of the following sections. The mathematical analyses that justify the approach and the results quoted in the following are found in the work just cited.

2.1 SUBBASINS AND PORTALS: MICROSCOPIC EVOLUTION

We think of *genotype space*—the collection of all genotypes—as a network whose nodes are genotypes and whose links connect genotypes that can be transformed into each other by simple genetic modifications, such as single-point mutations. Taking this and the biological facts of neutral evolution and the many-to-one structure of genotype-to-fitness maps into account, we see that genotype space decomposes into a set of neutral networks, or *subbasins* of approximately isofitness genotypes, which are entangled in a complicated fashion; see figure 6. As illustrated there, the space of genotypes is broken into strongly and weakly connected sets with respect to paths generated by genetic modifications. Equal-fitness genotypes form one or several strongly connected neutral subbasins. The volume of each subbasin is determined by the number of genes that can vary without changing fitness: the more *wildcard* genes within a genotype, the larger the volume. In fact, subbasin volume grows exponentially with the number of wildcard genes. Moreover, subbasins of high fitness are generally much smaller than subbasins of low fitness, since higher-fitness genotypes (typically) tend to require more *fixed* genes to maintain their fitness. One consequence is that a subbasin tends to be only weakly connected to subbasins of higher fitness. This is depicted by the tube-like portals in figure 6.

The genotype space for the epochal evolution example of figure 5 consists of all genotypes of 30 (binary) genes; a set of 2^{30} ($\sim 10^9$) binary strings. There are three functional gene-constellations, the 10 genes of which must be properly set to obtain a unit of fitness. Due to this, genotype space contains four subbasins of fitnesses 0, 1, 2, and 3, respectively. There is only one genotype with fitness 3; 3069 genotypes have fitness 2; $\sim 3 \times 10^6$ have fitness 1; and all others ($\sim 10^9$) have fitness 0. Thus, there are large degeneracies in the mapping from genotype to fitness.

Since the different genotypes within a subbasin are not distinguished by fitness selection, neutral evolution—driven by random sampling and genetic variation of individuals—dominates when the population resides in the subbasins. Selection stills acts to stabilize the population, of course, but only by culling low-fitness individuals; e.g., those in low-fitness subbasins. This leads to a rather different interpretation of the processes underlying stasis and change from that suggested by "landscape" models, for example. In landscape models a population stays pinned at a local optimum in genotype space, since all variation leads to decreased fitness. In epochal evolution, however, a population is free to diffuse randomly through subbasins of isofitness genotypes. A balance

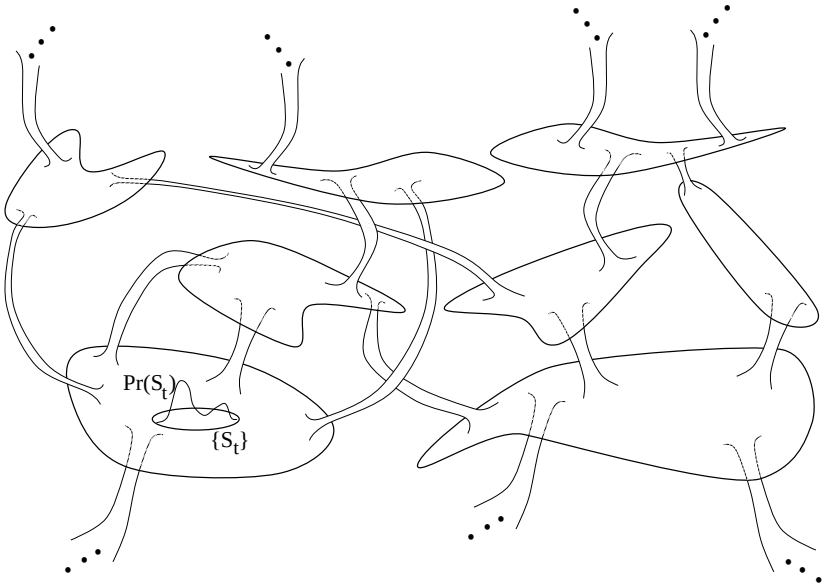

FIGURE 6 Subbasin and portal architecture in genotype space underlying epochal evolutionary dynamics. A population—a collection of individuals $\{S_t\}$ with distribution $Pr(S_t)$—diffuses in the subbasins (large sets) until a portal (tube) to a higher-fitness subbasin is found. (Reprinted with permission from [11].)

between selection and deleterious mutations leads to a (meta-) stable distribution of fitness (or of phenotype), while the population searches through spaces of neutral genotypic variants. During the neutral diffusion process the population of genotypes *accumulates* in the wildcard genes the history of the particular genetic variations that occurred. Even though there is no genotypic stasis during epochs, there is phenotypic stasis. As was first pointed out in the context of molecular evolution in [35], through neutral mutations, the best individuals in the population diffuse over the neutral network of isofitness genotypes until one of them discovers a connection to a neutral network of higher fitness. The fraction of individuals on this network then grows rapidly, reaching a new equilibrium between selection and deleterious mutations, after which the new subset of most-fit individuals diffuses again over the newly discovered neutral network.

Note that in epochal dynamics time scales are a naturally separated. During an epoch, selection acts to establish an equilibrium in the proportions of individuals in the different subbasins, but it does not induce *adaptations* in the population. Adaptation occurs only in a short burst during an innovation (passage through a portal), after which equilibrium on the level of fitness is

reestablished in the population. On a time scale much slower than that of innovations, members of the population diffuse through a subbasin of isofitness genotypes until a (typically rare) higher-fitness (portal) genotype is discovered. Thus, long periods of stasis occur because the population must search most of the subbasin before a portal to a higher-fitness subbasin is discovered. We refer to this as an *entropy barrier* to innovation, since the long duration of epochs is controlled by the *volume* of the most-fit neutral network on which the population resides.

In this way, we shift our view away from the geographic metaphor of evolutionary adaptation "crawling" along a "landscape", impeded by "energy" barriers (fitness-valleys), to the view of a diffusion process constrained by the subbasin-portal architecture. That architecture is induced, in turn, by degeneracies in the genotype-to-phenotype and phenotype-to-fitness mappings. This is not only a shift in architectural view, though, since it places a strong emphasis on the *dynamics* of populations as they move through subbasins, find portals, and so evolve increased fitness. It turns out that, while genotype-space architecture is a key component, it is not the only determinant of evolutionary population dynamics.

2.2 FINITE-POPULATION DYNAMICAL SYSTEMS: MESOSCOPIC EVOLUTION

From a microscopic point of view of genotype space, the exact state of an evolving population is only fully described when a list $\mathcal{S}$ of all genotypes with their frequencies of occurrence in the population is given. On the microscopic level, the evolutionary dynamics is implemented as a *Markov chain* with the conditional transition probabilities $\Pr(\mathcal{S}'|\mathcal{S})$ that the population at the next generation will be the collection $\mathcal{S}'$, given that the current population is $\mathcal{S}$. For any reasonable genetic representation, however, there is an enormous number of these microscopic states $\mathcal{S}$ and so too of their transition probabilities. The large number of parameters, $\mathcal{O}(2^L!)$ for L genes, makes it almost impossible to quantitatively study the dynamics at this microscopic level.

More practically, a full description of the dynamics on the level of microscopic states $\mathcal{S}$ is neither useful nor typically of interest. One is much more likely to be concerned with relatively coarse statistics of the dynamics, such as the evolution of the best and average fitness in the population or the waiting times for evolution to produce a genotype of a certain quality. The result is that quantitative mathematical analysis faces the task of finding a coarser description of the microscopic evolutionary dynamics that is simple enough to be tractable numerically or analytically and that, moreover, facilitates predicting the quantities of interest to an experimentalist. The key, and as yet unspecified, step in developing such a description of evolutionary processes is to find an appropriate set of intermediate-scale *mesoscopic* variables, or *mesostates*, with which to define the dynamics.

Fortunately, the very formulation of Neo-Darwinian evolution suggests a natural decomposition of the microscopic population dynamics into a part that is guided by selection and a part that is driven by genetic diversification. Simply stated, selection is an ordering force that operates on the level of the phenotypic fitness in a population. In contrast, genetic diversification is a disordering and randomizing force that drives a population to an increased diversity of genotypes. Thus, it seems natural to choose as mesostates the proportions of genotypes in different *fitness classes* (subbasins). Additionally, one can assume that, due to random genetic diversification within each subbasin, the distribution of individuals within a subbasin is determined only by these proportions and is, otherwise, as random and unstructured as possible. (This is the *maximum entropy* assumption of statistical physics.)

Following this reasoning, we describe a population in terms of the proportions $P_0, P_1, \ldots, P_N$ of individuals located in each of the subbasins $B_0, B_1, \ldots, B_N$. The maximum entropy assumption entails that within subbasin B_i, individuals are equally likely to be any of the genotypes in B_i. (This is a rather strong assumption that works surprisingly well in predicting observed population dynamics.) In other words, we assume that all wildcard genes are equally likely to be set in any possible way, as long as this does not lead to a portal configuration that changes fitness. Thus, we use the coarser *distribution of fitnesses*, rather than the much more unwieldy genotype distribution, to describe a population. (For simplicity of description, we are assuming that the subbasins have distinct fitnesses.) The immediate benefit is that we work with a space of populations that is vastly smaller—its dimension is the number N of subbasins—than the exponentially large space of genotypes.

Figure 7 illustrates how epochal evolution appears in the intermediate-scale mesoscopic representation afforded by fitness distributions. The figure plots fitness distributions $\vec{P} = (P_0, P_1, P_2, P_3)$ from the run of figure 5. In the figure the P_0 axis indicates the proportion of fitness-0 genotypes in the population, P_1 the proportion of fitness-1 genotypes, and P_2 the proportion of fitness-2 genotypes. Of course, since $\vec{P}$ is a probability distribution, $P_3 = 1 - P_0 - P_1 - P_2$ is completely determined, and the space of possible fitness distributions forms a solid three-dimensional *simplex*.

We see that initially $P_0 = 1$ and the population is located exactly in the lower-left corner of the simplex. Later, between $t = 20$ and $t = 60$, the population is located at a metastable fixed point on the line $P_0 + P_1 = 1$ and is dominated by fitness-1 genotypes ($P_1 \gg P_0$). Some time around generation $t = 60$ a genotype with fitness 2 is discovered, and the population moves into the plane $P_0 + P_1 + P_2 = 1$—the front plane of the simplex. From generation $t = 70$ until generation $t = 170$, the population fluctuates around a metastable fixed point in the upper portion of this plane. Finally, a genotype of fitness 3 is discovered, and the population moves to the asymptotically stable fixed point in the interior of the simplex. It reaches this fixed point around $t = 200$ and remains there fluctuating around it for the rest of the evolution experiment.

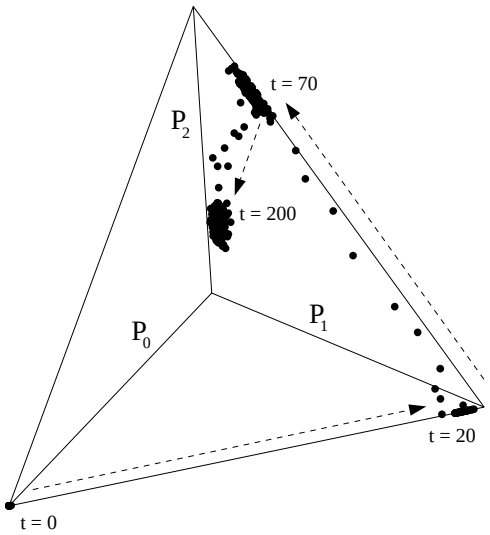

FIGURE 7 Mesoscopic view of epochal evolution—the level of population dynamics: the simplex (a solid tetrahedron) of allowed populations (dots) and the dynamic. The flow dynamics, including the clustering and regions of stability, is induced by selection and genetic variation. The stochasticity seen is the result of finite-population sampling. In this example, the fitness distribution $\vec{P} = (P_0, P_1, P_2, P_3)$ is shown for a population evolving under the fitness function of figure 5 which gives genotypes three levels of fitness: $0, 1, 2, 3$. The location of the fitness distribution at each generation is shown by a dot. The dashed lines indicate the direction in which the fitness distribution moves from metastable to metastable cluster. (The population data used here comes from the same run as in figure 5.) The times at which the different metastable states are first reached are indicated as well. These should be compared to the innovation times of figure 5. (Reprinted with permission of the authors [66].)

2.3 SURVIVAL DYNAMICS: MACROSCOPIC EVOLUTION

Having described epochal evolution at the microscopic level of diffusion through subbasins and portals and the mesoscopic level of the population dynamics, we can return to the highest level of evolution: survival dynamics in the space of functionality—the space on which selection acts. In the simple example already shown in figure 5 we used fitness as a proxy for functionality. Recall that figure 5 showed the fitness dynamics of a population of 30-gene individuals evolving under a three-constellation fitness function.

At time $t = 0$ the population started out with 250 random genotypes. As can be seen from figure 5, during the first few generations all individuals were located in the largest subbasin with fitness 0, since both average and best fitness are 0. The population randomly diffused through this subbasin until, around generation 20, a portal was discovered that led into the subbasin with fitness 1. The population was quickly taken over by genotypes of

fitness 1, until a balance was established between selection and deleterious mutation: selection increasing the fraction of fitness-1 individuals and deleterious mutations (that go from fitness 1 to 0) decreasing their number. The individuals with fitness 1 continued to diffuse through the subbasin with fitness 1, until a portal was discovered connecting to the subbasin with fitness 2. This happened around generation $t = 60$ and by $t = 70$ a new selection-mutation equilibrium was established. Individuals with fitness 2 continued diffusing through their subbasin until the globally optimal genotype with fitness 3 was discovered some time around generation $t = 170$. Descendants of this genotype then spread through the population until around $t = 200$, when a final stable equilibrium was reached.

2.4 PORTRAIT OF AN INNOVATION: UNFOLDING AND STABILIZING NOVELTY

Putting together the views of evolution at the three different levels of genotype subbasins and portals, population dynamics, and survival dynamics, one sees that epochal evolution is a process of state-space unfolding (see figure 8):

1. Initially, the population moves in (say) n mesoscopic dimensions of the population-dynamics space of fitness distributions.
2. It is attracted to a (noisy) fixed point—the metastable collection of populations observed during the epoch.
3. At the same time it diffuses neutrally in the very high dimensional microscopic space of genotypes. During epochs, many genotypic changes occur and accumulate, but do not alter the phenotype. This invariance of the phenotype is a *symmetry* of the fitness distribution with respect to microscopic change.
4. An innovation occurs when, having accumulated a certain combination of changed genes, a portal to increased fitness is discovered in the microscopic space.
5. This breaks the existing epoch symmetry, since genetic changes now affect fitness.
6. A new mesoscopic dimension becomes activated, fitness increases and selection begins to stabilize the innovated feature by removing lower-fitness genotypes (without the feature) and adding higher-fitness ones (with the feature).
7. The mesoscopic population dynamics now moves in an $(n+1)$-dimensional space.

In the unfolding process microscopic variation is amplified through the innovations and becomes locked-in due to the dynamics at the mesoscopic and macroscopic levels. Randomness serves to drive the diffusion in the microscopic dimensions and eventually leads to the discovery of portals to innovation. Selection acts to stabilize the structure of the mesoscopic spaces, once a new

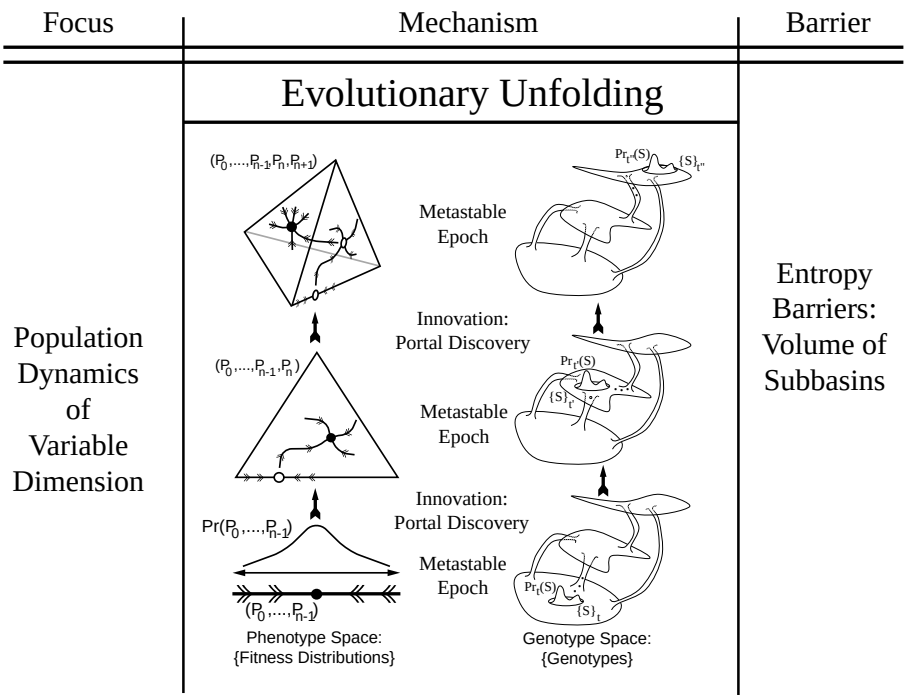

FIGURE 8 Portrait of an innovation: The mesoscopic (left) and microscopic (right) views of two innovations produced by epochal evolution. See text for description.

dimension has been activated. Complementing this flow of information from the microscopic to the macroscopic, there is also feedback from the macroscopic level that determines the constraints on the microscopic dynamics. That is, the macroscopic organization of possible individuals—e.g., attainable fitness levels—is reflected in the subbasin-portal architecture of the microscopic space. The mathematical analysis of these mechanisms and their interaction we call *statistical dynamics*.

It will be helpful at this point, having outlined the statistical dynamics of epochal evolution, on the one hand, and having earlier mentioned several alternative descriptions of the causes of metastability, on the other, to draw the contrasts more sharply between them. First, epochal evolution is not Fisher's intermittent fixation. Though aspects of drift due to sampling and mutation are components of epochal evolution, the epochs are stabilized by selection removing low-fitness genotypes, which is not part of Fisher's model. Second, epochal evolution is not Kimura's purely neutral evolutionary dynamic, since a key part of the former is an explicit mechanism for finding and then locking in structural innovations. In this sense, the theory of epochal evolution proposes an overall architecture for piecewise-neutral evolutionary processes.

Third, epochal evolution is not Gould-Eldredge punctuated equilibria in that it is a predictive quantitative theory with specific mechanisms that, with new automated evolutionary experiments, will be laboratory testable. The theory of punctuated equilibria did not commit to much, if any, underlying mechanism, other than exogenous (environmental) causes of change. (The statistical dynamics of epochal evolution is a theory of endogenous change.) Punctuated equilibria served more as a descriptive summary of phenomenological aspects of the fossil record, as one naturally expects of paleontology. Perhaps at some future date, with more fossil data and a more elaborated theory of epochal evolution, it may be shown that punctuated equilibria in macroevolution is a kind of epochal evolution. At present, all one has is observational consistency, without the ability to positively identify underlying mechanisms from the fossil record.

One notable consequence of the statistical dynamics analysis is that epochal evolution is a kind of open-ended evolution. It is explicitly a dynamics by which a sequence of innovations can be discovered and then become the structural substrate for further evolution. Also, depending on which (randomly chosen) sequence of portals is realized, the course of macroscopic evolution can be very different and so reflect the accumulation of, what some call, *frozen accidents*. In these respects, the statistical dynamics of epochal evolution is a partial response to the criticism of population dynamical systems modeling of evolution as being evolutionarily closed and incapable of intrinsic novelty. The claim is that such models must at the outset build-in the ultimate dimensionality of an evolutionary process which, in turn, caps evolutionary innovations [20]. Epochal evolution shows that this is not an intrinsic failing of population dynamical systems: they can be open-ended in the way epochal evolution unfolds and then stabilizes new state space. The main limit imposed on the continuing emergence of increasingly complex structures—assuming other parameters, such as population size and mutation rate, are compatible—comes from the structure of the space of individual function, not directly from population dynamics.

These observations on innovation processes in evolution lead immediately to questions about what one means by "structure" and "function"—largely open questions, as yet incompletely addressed by evolutionary theory. How can we ever say unambiguously, for example, that an evolutionary system evolved toward complexity, or that it was or was not open-ended, without a theory of structure and function that allows us to quantitatively monitor their change?

At its current stage of development, in the statistical dynamics of epochal evolution, "individuals" are simple and direct genotype-to-fitness maps. Beyond gene constellations that confer fitness when properly set, they are nearly structureless individuals. For example, they have no spatial structure and no temporal behavior. These aspects can play no role in determining individual fitness. Thus, there is no analog of development in the theory, except that which is implicit in the genotype-to-fitness maps. In contrast, the laboratory

experiments and the simulations of evolving dynamical systems mentioned above do have structured individuals, often exhibiting complex structures and rich dynamical behavior. To develop a predictive theory of epochal evolution for these, one needs to be precise about how a given individual is structured, how it functions, and how its functionality confers fitness, in order to track increases (or decreases) in evolutionary and developmental complexity. In short, to quantify changes in structure and function one needs to define "complexity".

3 STRUCTURAL COMPLEXITY AND INNOVATION

Fortunately, recent progress has brought us to a level of understanding complexity that suggests we are close to defining it in ways that are germane to evolution. In particular, results on how complexity emerges in nonbiological systems give important insights into the structures that can emerge in evolutionary processes and also into the constraints that structural innovations must respect when they occur. Before describing these ideas, however, it will be helpful to set the historical context and to make several important distinctions.

As a label for natural systems that are difficult to model and analyze, over the last two decades "complexity" has served a useful role by ambiguously referring both to randomness and to organization. The study of *complex systems* has sometimes focused on simple (albeit, nonlinear) processes that appear random and are difficult to predict—e.g., deterministic chaos and fractal separatrices. The question there, to say it most directly, is, How does disorder emerge from simplicity? At other times, studies of complex systems have focused on large-scale processes consisting of many interconnected components—what one might call *complicated systems*. The question there has been, How is it that order arises despite so much possible disorganization? It is not surprising that two such opposite phenomena—disorder emerging from order and order from disorder—falling under the same rubric of complex systems would lead to confusion. Fortunately, the confusions and resulting debates about what "complexity" is led to a useful clarification. There are two basic and different categories of complexity: complication versus structure. As we now appreciate, although complication emerging from simplicity and organization emerging from disorder appear to be opposite kinds of phenomena, the complexities to which they refer are, in fact, complementary and not opposites.

On the one hand, we have complexity as varying degrees of randomness or complication in a system's behavior or in its architecture. The behavior of a dynamical system ranges, say as we change a control parameter to make it more nonlinear, from being regular, periodic, and predictable to chaos and unpredictability. The organization of social systems ranges from the predictable delivery of vast amounts of food to major cities to the seeming turbulence and uncoordinated deal-making behavior of traders in a stock exchange. Thus, we

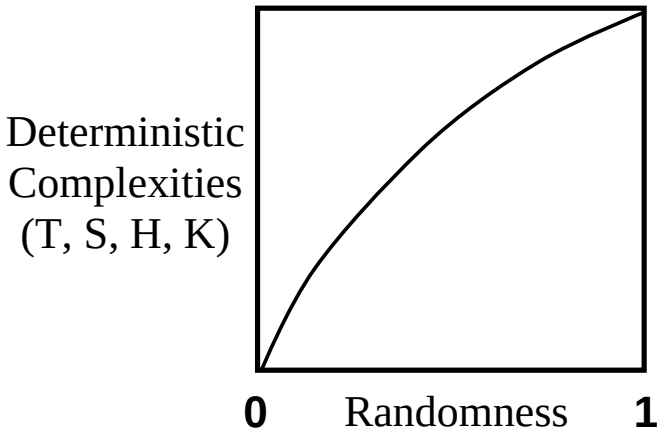

FIGURE 9 Measures of complicatedness—temperature T, thermodynamic entropy S, Shannon information H, and Kolmogorov-Chaitin complexity K—(vertical axis) are all proportional to the amount of randomness in a process (horizontal axis). In one way or another, these quantities assume that all of the randomness in a process must be described using deterministic models—such as, a universal Turning machine on which the Kolmogorov-Chaitin complexity is defined. This is why they are referred to as *deterministic* complexities. (Reprinted with permission from [7].)

think of natural complication as a spectrum of randomness: from pure order to utter disorder.

This spectrum is quite familiar to us. So much so that many fields have developed their own vocabularies for degrees of complicatedness. In physics, for example, one uses temperature (T) and thermodynamic entropy (S) to monitor where a system is in its spectrum of randomness: low temperature or entropy indicate an ordered system, high temperature or entropy a disordered one [55, 57, 71]. In the theory of communication, one uses Shannon's measure (H) of information: predictable messages are uninformative; unpredictable messages are highly informative [6, 59]. In the theory of computation, one uses Kolmogorov-Chaitin complexity (K) as an algorithmic measure of an object's randomness [5, 40, 43, 45], and so on. We can illustrate very simply, as done in figure 9, the relationship between these ways of measuring degrees of complicatedness: T, S, H, and K are all proportional to randomness.

On the other hand, we have complexity as varying degrees of organization—or structure, regularity, symmetry, and intricacy—in a system's behavior or in its architecture. We say that a ferromagnet is more structured at the transition between its low-temperature ordered phase and its high temperature disordered phase, since only there does it exhibit aligned-spin clusters of all sizes. The network of financial, technological, and industrial interdependencies that support the production of modern microprocessors is certainly neither a regular and fixed architecture, in which case it would be too rigid and un-

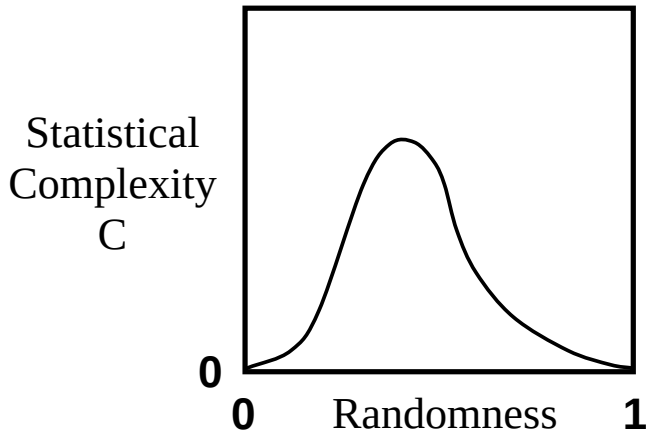

FIGURE 10 Structural complexity C (vertical axis) peaks in the intermediate region between the extremes of randomness (horizontal axis). (Reprinted with permission from [7].)

adaptive, nor one that is entirely unstructured, in which case it would simply be nonfunctional. The required institutional memory, flexibility, competition, and cooperation have led it to some state intermediate between these organizational extremes. In a way analogous to randomness, in the space of organizational architectures, we can think of a spectrum of structure: from simple symmetric architectures to sophisticated and hierarchical ones.

The corresponding measure of structuredness we call *structural complexity* (C). There have been many more or less specific proposals for structural complexity—some, it turned out, actually measure randomness. (See the review and especially the long list of citations in [17, 18, 58].) Nonetheless, we can summarize the basic idea behind structural complexity by contrasting it with the spectrum of randomness. This is done schematically in the *complexity-entropy diagram* shown in figure 10. The peaked curve shows that (i) the extremes of randomness, highly predictable and highly random, are structurally simple (low C) and (ii) structural complexity is largest in the intermediate regime between the extremes.

We now consider structural complexity to be a complementary coordinate to degrees of randomness, as depicted in figure 10. It characterizes a different feature of a system—for temporal processes, the amount of historical *memory*—than randomness—which is the amount of information a system *produces*. That is to say, randomness and structural complexity are both necessary descriptors: the former captures surprise, the latter organization. In analyzing dynamical systems, for example, one uses complexity-entropy diagrams like that of figure 10 to display the spectrum of how a collection of systems generate and store information to varying degrees.

We refer to the theory of structural complexity as *computational mechanics* since it extends statistical mechanics—a theory of randomness—to include definitions of structure that capture computational architectures. The main questions asked when analyzing a system in this framework do not focus on the storage and transduction of energy. They ask instead how a system stores, transmits, and transforms *information*. Briefly, what are a system's *intrinsic computational* properties? Originally introduced over a decade ago [12], the mathematical foundations are now well developed. (See [61] and, again, [17, 58], which also review alternative approaches to structural complexity.) Computational mechanics defines structural complexity (C) in terms of a decomposition of a system's behavior into its minimal causal architecture—a representation called an *ϵ-machine*. C is the amount of information, including spatial correlation and temporal memory, which this minimal architecture stores. In a well defined sense C is the size of the set of minimal causal components embedded in a system. The procedure for identifying a system's minimal causal architecture is called *ϵ-machine reconstruction*.

Aside from providing a first-principles approach to extracting an ϵ-machine for a system and so measuring its structural complexity, one of the main results is that novel structures (forms of intrinsic computational architecture) emerge in pattern-forming systems that are at phase transitions. More generally, it is often observed that structural complexity emerges from the dynamical interplay of ordering and disordering forces—such as, those operating when discovering portals in neutral networks. (These results justify the rather coarse and schematic view captured in the complexity-entropy diagram of figure 10.)

Comparisons of how novel structural complexity emerges at different kinds of transition, such as phase transitions, and over time in cellular automata [33], for example, give some insight into the structural innovations that can emerge in evolutionary processes. (Investigations of evolving cellular automata give many examples of just this kind of structural innovation in an artificial evolutionary process—innovations that can be structurally analyzed in some detail using computational mechanics [9, 10].) First, the central way to detect that some new thing has emerged in an innovation is to monitor the causal architecture—either over time, if analyzing a temporal process such as the evolutionary population dynamics of cellular automata, or over a range of parameters, if it is a controlled process, such as a system undergoing a phase transition. Second, increased structural complexity can appear either smoothly, as shown in figure 10, or abruptly as in a critical phase transition; see, for example, [13]. In the latter case, there is a qualitative change in a system's causal architecture: a divergence in the number of causal components and a shift to a more powerful computational class. Innovations have been analyzed in systems that show a shift from a disorganized initial "heat bath" to patterned levels of coherent domains and particles, from finite-memory to infinite-memory processes, and from finite-state-machine to pushdown-stack architectures. Such innovations have structural signatures and, using computational mechanics, there are now ways to detect them and quantify what

novelty has emerged. Thus, the emergence of structural complexity from the interplay of a system's tendency to order and its tendency to disorder suggests where to look for innovations, how to detect them, and how to describe what has been created.

4 EVOLUTION TO COMPLEXITY

In the computational mechanics of structural complexity one sees the beginnings of a principled approach to form and function in evolutionary processes. First, identifying and then quantifying the kinds and amounts of structure embedded in natural systems are the first steps to making the concept of form precise and testable. To the extent that one considers biological form to include symmetry, regularity, hierarchy, pattern, modularity, and so on, structural complexity, as defined in computational mechanics, is an appropriate operational approach to it. Second, building on an unambiguous concept of form, one can view functionality as arising from the relationships between a system's intrinsic structures—their static architecture and their dynamical interaction—and intrinsic or externally determined evaluation of those structures.

In the computational mechanics view, then, evolutionary innovations are changes in the architecture of information processing. These changes can be reflected at the level of either a population or an individual. The novelty of an innovation is built out of structures on lower levels and occurs in "orthogonal" coordinates when something truly new emerges. Unlike the purely structural emergence observed in pattern formation processes—such as, the appearance of spiral waves in a Belusov-Zhabotinsky chemical reaction-diffusion system—innovations can take on meaning and function *within* in an evolutionary process. Unlike the spiral waves, this *intrinsic emergence* does not require an outside observer to monitor the changes in structural complexity [8]. The meaning and function of intrinsically emerging organization derives from the fitness evaluation of individuals and the persistence of traits over time—features that are part and parcel of an evolutionary process.

In computational mechanics, the process by which open-ended innovation can occur is called *hierarchical ϵ-machine reconstruction*. We think of hierarchical ϵ-machine reconstruction, or some dynamically instantiated version, as specifying the minimal requirements for open-ended evolution: successive innovation of levels of distinct structural classes that build on the lower levels' component structures. Figure 11 illustrates an open-ended series of evolutionary innovations: nested levels of information processing of increasing computational power. At each level, there is a spectrum of structures, some of which are more appropriate (e.g., useful or functional), since they balance both parsimonious resource use against minimal degrees of randomness. To store and use better structures, however, requires increased resources—they are in one or another sense larger than less optimal structures.

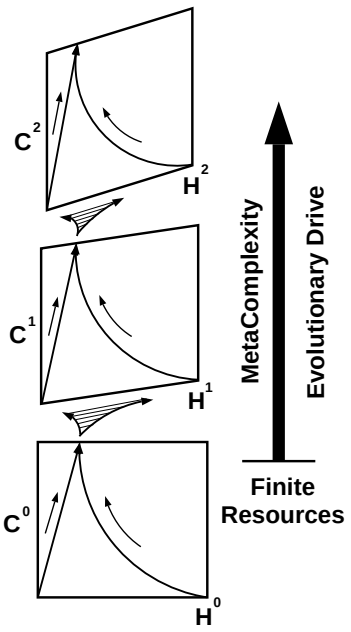

FIGURE 11 Computational mechanics view of open-ended evolution: a sequence of innovations leading up through a hierarchy of increasingly sophisticated classes of information processing. The distinct classes of structure are illustrated with a complexity-entropy diagram (H^i, C^i) that represents the trade-offs between randomness and structural complexity appropriate at each organizational level. The arrows indicate an adaptive dynamics that leads within each level to increasingly sophisticated structures. When adaptation exhausts finite resources, there is an innovation of a new class of structure. (Reprinted with permission from [7].)

It is perhaps not surprising that finite resources drive the process of innovation. Why? If an individual at some level of organization had infinite resources, say to model its environment, then there would no benefit to restructure existing resources or to incorporate new ones. It would gain no predictive advantage, since the current model is as predictive as any alternative. When resources are limited there is an effective pressure to innovate a new class of organization—one that more efficiently uses the available resources while improving efficacy. In this way, computational mechanics describes an open-ended series of innovations as an unfolding hierarchy of recursively embedded structural classes. Though this view of open-ended evolution focuses on innovations in structural complexity, one sees the parallel between this process and how innovations arise during epochal evolution by a process of unfolding and then stablizing new state spaces. In computational mechanics one focuses on nested levels of information processing; in epochal evolution one focuses on activating and then stabilizing new state spaces.

5 ORIGINS OF FORM AND FUNCTION: SELECTION, ACCIDENT, OR MORPHOGENESIS?

It is often said that organisms today are more complex than in earlier times. But what (exactly) is this complexity and how did it emerge? Having reviewed the dynamics of epochal evolution and the theory of structural complexity and the roles they play in innovation, we can now contrast them more directly with views from within evolutionary biology on the emergence of form and function. According to Williams, three doctrinal bases have been used over the last century to address the evolution of complexity [68]:

1. *Natural selection*: "trial and error, as opposed to rational plan";
2. *Historicity*: "the role of historical contingency in determining the Earth's biota"; and
3. *Mechanism*: "only physico-chemical processes are at work in an organism ... as opposed to vitalism".

Natural selection holds that structure in the biological world is due primarily to the fitness-based selection of individuals in populations whose diversity is maintained by genetic variation [46]. That is, genetic variation is a destabilizing force that provides the raw diversity of structure. Natural selection then is a stabilizing dynamic that acts on the expression of that variation, which is structural diversity. It "generates" organization by culling individuals based on their relative fitness, which is determined by their structure. This view identifies a source (genetic variation) of new structures and a mechanism (selection) for altering one form into another. Thus, the adaptiveness accumulated via selection is seen as the dominant mechanism driving the appearance of form and function.

The historicity doctrine acknowledges the Darwinian mechanisms of selection and variation, but emphasizes the accidental determinants of biological form [28, 51]. What distinguishes this position from the emphasis on natural selection is the claim that major changes in structure can be and have been nonadaptive. While these changes have had the largest effect on the forms of present day life, at the time they occurred they conferred no survival advantage. Furthermore, today's existing structures need not be adaptive, reflecting instead a history of frozen accidents. One consequence is that a comparative study of parallel earths would reveal very different collections of life forms on each. Like the doctrine of natural selection, historicity accounts for the emergence of structure by a process of preferentially culling one or several structures within pre-existing structural diversity. But it is a dynamics that is manifestly capricious or, at least, highly stochastic with few or no causal constraints. Due to this, the historicity doctrine is not a theory of the *origins* of diverse form and function.

In the mechanistic view of evolution the goal is to elucidate "principles of organization" that underlie the appearance of biological form. In this,

it focuses directly on the question of what biological complexity might be. The doctrine contends that energetic, mechanical, biomolecular, and morphogenetic properties guide and limit the infinite range of possible biological form [20, 21, 26, 37, 60, 67]. The constraints result in a relatively small set of structure archetypes. In a sense, these play a role in morphology analogous to the Platonic solids in geometry: they pre-exist, before any evolution takes place. In the evolutionary emergence of complexity, then, natural selection chooses between these "structural attractors", possibly fine-tuning their adaptiveness. In this view, Darwinian evolution serves, at best, to fill the waiting attractors or not, depending on historical happenstance. It does not, however, create the structure of those attractors.

What is one to think of these conflicting theories of the evolution of complexity?

First, although natural selection's culling of genetic variation provides a theory of gradual structural transformation, it does not provide a theory of structure itself. For example, what is the average time under an evolutionary dynamic and under the appropriate environmental pressures for a fish fin to be transformed into a leg? If one knew the genetic trajectory—the required sequences of modified and innovated genes—in principle one could use the theory of population genetics to estimate how long the transformation would take. But this assumes and hides too much—How did those genes determine the functionality of fins and legs? To estimate from first principles the time to evolve a leg from a fin one needs a measure of the structures concerned and of the functionality they do or do not confer.

Second, historicity too provides a theory of transformation and not of structure. Moreover, for its highly stochastic transformational dynamics to be successful—or, at least, to not destroy all structures—and for there to be the requisite broad structural diversity on which it acts, historicity requires that the space of possible biological structures be populated with a high fraction that are functional. Whether this is true or not is simply unknown. Additionally, in emphasizing the dominance of historical accident, it advocates an extra-evolutionary theory for the origins of novel organization and form, side-stepping the issue of biology's role in actively producing them. An explanation that appeals to a meteor crashing into the earth simply falls outside the domain of evolutionary theory. Moreover, the occurrence of such events is unlikely ever to be explained by the principles of physical dynamics. The collision just happened: a consequence of particular initial conditions that occurred in a celestial dynamical system which is most likely chaotic and, if so, demonstrably unpredictable. Such accidents impose significant constraints; they do not constitute an explanation of the origins or biological form or function.

Finally, the mechanistic doctrine does not offer a theory of evolutionary transformation, though it focuses on morphogenesis which certainly interacts with evolutionary processes. Although it employs methods from pattern-formation and bifurcation theories, it too falls short in that it does not provide a theory of structure itself nor of the functions of evolved structures. In partic-

ular, the structural attractors are not quantitatively analyzed in terms of their internal architecture nor in terms of system-referred functionality or fitness.

It would appear that the three doctrines rely on undefined concepts of form and function. What about the theory of neutral evolution? What does it say about form and function and their evolution? In natural selection's emphasis on gradual adaptation each and every biological thing, embodying the direct solutions to the survival problems posed by environmental constraints, has a function and so a "story". Neutrality, though, breaks the logic of functional ascription. The direct consequence of neutral evolution is the appearance of nonadaptive, nonfunctional, and nonfitness-conferring genotypes or phenotypes. Previously, Kimura argued that neutrality or near-neutrality is the rule in molecular evolution. One can also develop a different kind of argument, that one should expect neutrality to be common in the evolution of form and function, using the theory of structural complexity.

To say it most simply, whenever there is structure, there will be a many-to-one mapping from genotype to phenotype and to fitness. Why? To say there is "structure" is to say that the range of possible entities is constrained, not random. If the range of possible forms is not fully random (is not structureless), then the mapping of genotypes—an exponentially large number of long strings in a high-dimensional space—to structures is degenerate: many genotypes will code for individual structures, the former substantially out numbering the latter. The many-to-oneness derives most fundamentally, though, from a collapse of dimensionality in going from the microscopic realm of genotype space to the macroscopic realm of form and function. Even if genotypic coding consisted of continuous parameters rather than discrete genes so that the preceding (combinatorial) argument did not apply, any reduction in the dimension (from genotype space to phenotype space to function space) results in neutrality. Thus, the evolution of structurally complex organisms appears to implicate in a fundamental way neutral evolution and so inherently epochal population dynamics. Moreover, when properly calibrated against landscape-optimization processes, evolution along neutral pathways dominates since the time it takes to find innovations is markedly shorter than the time taken by fitness-valley crossing [63].

If neutral evolution is to be expected in the emergence of complex organisms, then there need not always be functional "stories" for each of their component structures, since some structures may have arisen during periods of stasis. Of course, they may become functional later on, say even contributing to an innovation. (Gould and Vrba call this recontextualizing originally nonfunctional traits *exaptation* [30].) Imagine examining a contemporary organism. Which of its structures emerged during periods of stasis and which not? The ambiguity here is only heightened when one realizes that, in many cases for which data is available, epochs of stasis are defined in terms of morphological, and not functional, constancy. Much of what we see in the biological world need not be there because it fulfills a purpose—not even for survival. Functionality, perhaps emerging through adaptive innovations, comes equally

from the context of a given form—something much harder to detect than form itself. Thus, confronted with the possibility of metastable evolution, one comes to appreciate diversity of all kinds, even that which is not functional and which appears to serve no purpose. One may have to adopt a very long view. Present diversity may be highly determinant at some later time and in a different context.

6 NONE OF THE ABOVE

The impression the doctrinal debate leaves, though, is that there is a pressing need for both a qualitative dynamical theory of structural emergence [27] and a theory of biological structure itself [47, 48]. The main problem, at least to an outsider, does not reduce to showing that one or the other existing doctrine is correct. Each employs a compelling argument and often empirical data as a starting point. Rather, as a first step, the task facing us is to develop a synthetic theory that can balance the tensions between selection, accident, neutrality, and mechanism. The analysis of epochal evolution—how it unfolds and stabilizes novel macroscopic spaces, when this occurs and when it is precluded—does suggest what this qualitative dynamical theory might look like. It resolves the tension between the microscopic, mesoscopic, and macroscopic levels on which evolutionary processes act and gives an architectural view of the microscopic and mesoscopic consequences of function.

If we ask about the *origin* of function, though, does it lie in selection, historical accident, morphogenesis, or some combination, the answer here has to be "none of the above". There is some basic thing missing in these three approaches. (Neutral evolution, as just noted, plays no direct role and mostly serves to complicate the question of function.) They do not directly address the question of functionality, nor are they equipped to do so. I argued that, on the way to addressing the origins of function, what is missing is a theory of form based on structural complexity and a theory of its emergence based on epochal evolution.

At this point, however, structural, and functional constraints operating during epochal evolution are only reflected, and indirectly so, in the subbasin-portal architecture of genotype space. Like the three existing doctrines, the theory of epochal evolution is not a theory of structure, nor does it yet incorporate one. Thus, there is a second, much more difficult, step: to develop a quantitative theory of structure and then out of that, a theory of function. Without these, we appear to be in no position to explain the evolution of complexity. It would appear that if one stops here, when evolution is revolution, we simply cannot say what has been innovated.

The computational mechanics of nonlinear processes, however, is a theory of structure. Pattern and structure are articulated in terms of various types of causal architecture—what we called computational classes. The overall mandate there is to provide both a qualitative and a quantitative analysis

of natural information processing architectures. If computational mechanics is a theory of structure, then innovation via hierarchical ϵ-machine reconstruction is a computation-theoretic approach to the open-ended transformation of structure. It suggests one mechanism with which to study what drives (finite resources) and what constrains (intrinsic computation) the appearance of novelty.

The discussion has brought us to a possible next step toward an evolutionary dynamics of structural emergence. This would be to fold hierarchical ϵ-machine reconstruction into an evolutionary process, resulting in an intrinsic dynamics of innovation. In a rough way, something like this is observed in the evolution of cellular automata, mentioned earlier. The theoretical analysis there is incomplete. However, at least evolving cellular automata provides a concrete case, which appears to be tractable and which can be used to ferret out the many taxing definitional problems in the evolutionary dynamics of form and function.

There are two main points to draw from the parallel threads of epochal evolution and structural complexity. First, epochal evolution arises intrinsically: long periods of stasis and sudden change need not be driven by external forces. They are the product of the many-to-one mappings from genotype to phenotype and phenotype to fitness. Epochal evolution is to be expected and it occurs by an open-ended process of discovering and stabilizing novelty—novelty that becomes substrate for further evolution. Second, the emergence of structure can be monitored as an open-ended hierarchy of novel kinds of embedded computation and information processing. When these two threads are knitted together, one hopes that, when evolution is revolution, we will be able to say what novelty has been created.

ACKNOWLEDGMENTS

I thank Victoria Alexander, Dave Feldman, Erik van Nimwegen, and Cosma Shalizi for many stimulating and wide-ranging discussions. I also appreciate discussions with the participants of the workshop **Towards a Comprehensive Dynamics of Evolution—Exploring the Interplay of Selection, Neutrality, Accident, and Function**. Partial support for this work comes from DARPA contract F30602-00-2-0583, NSF grant IRI-9705830, AFOSR via NSF grant PHY-9970158, and the Keck Foundation's Evolutionary Dynamics Program at the Santa Fe Institute.

REFERENCES

[1] Adami, C., "Self-organized criticality in living systems". *Phys. Lett. A* **203** (1995):29–32.

[2] Babajide, A., I. L. Hofacker, M. J. Sippl, and P. F. Stadler, "Neutral networks in protein space: A computational study based on knowledge-based potentials of mean force". *Folding Design* **2** (1997):261–269.

[3] Barnett, L., *Tangled webs: Evolutionary Dynamics on Fitness Landscapes with Neutrality*. Master's thesis, School of Cognitive Sciences, University of East Sussex, Brighton, 1997, http://www.cogs.susx.ac.uk/lab/adapt/nnbib.html.

[4] Bergman, A. and M. W. Feldman, "On the Population Genetics of Punctuation". In *Evolutionary Dynamics: Exploring the Interplay of Selection, Neutrality, Accident, and Function*, edited by J. P. Crutchfield and P. Schuster, 83–102, Santa Fe Institute Studies in the Sciences of Complexity, New York: Oxford University Press, 2001.

[5] Chaitin, G., "On the Length of Programs for Computing Finite Binary Sequences". *J. ACM* **13** (1966):547–569.

[6] Cover, T. M. and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.

[7] Crutchfield, J. P., "The Calculi of Emergence: Computation, Dynamics, and Induction". *Physica D* **75** (1994):11–54.

[8] Crutchfield, J. P., "Is Anything Ever New? Considering Emergence". In *Complexity: Metaphors, Models, and Reality*, edited by G. Cowan, D. Pines, and D. Melzner, vol. XIX of *Santa Fe Institute Studies in the Sciences of Complexity*, 479–497, Reading, MA: Addison-Wesley, 1994.

[9] Crutchfield, J. P. and M. Mitchell, "The Evolution of Emergent Computation". *Proc. Natl. Acad. Sci. U.S.A.* **92** (1995):10742–10746.

[10] Crutchfield, J. P., M. Mitchell, and R. Das, "The Evolutionary Design of Collective Computation in Cellular Automata". In *Evolutionary Dynamics—Exploring the Interplay of Selection, Neutrality, Accident, and Function*, Santa Fe Institute Series in the Sciences of Complexity, Oxford University Press, 2001.

[11] Crutchfield, J. P. and E. van Nimwegen, "The Evolutionary Unfolding of Complexity". In *Evolution as Computation*, edited by L. F. Landweber, E. Winfree, R. Lipton, and S. Freeland, to appear, Lecture Notes in Computer Science, New York: Springer-Verlag, 2000.

[12] Crutchfield, J. P. and K. Young, "Inferring Statistical Complexity". *Phys. Rev. Lett.* **63** (1989):105–108.

[13] Crutchfield, J. P. and K. Young, "Computation at the Onset of Chaos". In *Entropy, Complexity, and the Physics of Information*, edited by W. Zurek, vol. VIII of *SFI Studies in the Sciences of Complexity*, 223–269, Reading, Massachusetts: Addison-Wesley, 1990.

[14] Darwin, C., *On the Origin of Species*. New York: Modern Library, 1993.

[15] Ekland, E. H. and D. P. Bartel, "RNA-catalysed RNA polymerization using nucleoside triphosphates". *Nature* **382** (1996):373–376.

[16] Elena, S. F., V. S. Cooper, and R. E. Lenski, "Punctuated evolution caused by selection of rare beneficial mutations". *Science* **272** (1996):1802–1804.

[17] Feldman, D. P. and J. P. Crutchfield, "Discovering Noncritical Organization: Statistical Mechanical, Information Theoretic, and Computational Views of Patterns in Simple One-Dimensional Spin Systems". *J. Stat. Phys.* (1997):submitted, Santa Fe Institute Working Paper 98-04-026.

[18] Feldman, D. P. and J. P. Crutchfield, "Measures of Statistical Complexity: Why?" *Phys. Lett. A* **238** (1998):244–252.

[19] Fisher, R. A., *The Genetical Theory of Natural Selection*. Oxford, The Clarendon Press, 1930.

[20] Fontana, W. and L. Buss, ""The Arrival of the Fittest": Toward a Theory of Biological O rganization". *Bull. Math. Bio.* **56** (1994):1–64.

[21] Fontana, W. and L. Buss, "What Would be Conserved if the Tape were Played Twice?" *Proc. Nat. Acad. Sci. USA* **91** (1994):757–761.

[22] Fontana, W. and P. Schuster, "Continuity in Evolution: On the Nature of Transitions". *Science* **280** (1998):1451–1455.

[23] Fontana, W., P. F. Stadler, E. G. Bornberg-Bauer, T. Griesmacher, I. L. Hofacker, M. Tacker, P. Tarazona, E. D. Weinberger, and P. Schuster, "RNA folding and combinatory landscapes". *Phys. Rev. E* **47** (1992):2083–2099.

[24] Forst, C. V., C. Reidys, and J. Weber, "Evolutionary dynamics and optimizations: Neutral Networks as model landscapes for RNA secondary-structure folding landscape". In *Advances in Artificial Life*, edited by F. Moran, A. Moreno, J. Merelo, and P. Chacon, vol. 929 of *Lecture Notes in Artificial Intelligence*, 128–147, Springer, 1995.

[25] Gavrilets, S., "Evolution and Speciation in Hyperspace: The Roles of Neutrality, Selection, Mutation, and Random Drift". In *Evolutionary Dynamics: Exploring the Interplay of Selection, Neutrality, Accident, and Function*, edited by J. P. Crutchfield and P. Schuster, 103–130, Santa Fe Institute Studies in the Sciences of Complexity, New York: Oxford University Press, 2001.

[26] Goodwin, B., "Evolution and the Generative Order". In *Theoretical Biology: Epigenetic and Evolutionary Order from Complex Systems*, edited by B. Goodwin and P. Sanders, 89–100, Baltimore, Maryland: Johns Hopkins University Press, 1992.

[27] Goodwin, B. and P. Sanders, editors, *Theoretical Biology: Epigenetic and Evolutionary Order from Complex Systems*, Baltimore, Maryland, 1992, Johns Hopkins University Press.

[28] Gould, S. J., *Wonderful Life*. New York: Norton, 1989.

[29] Gould, S. J. and N. Eldredge, "Punctuated equilibria: The tempo and mode of evolution reconsidered". *Paleobiology* **3** (1977):115–251.

[30] Gould, S. J. and E. S. Vrba, "Exaptation: A missing term in the science of form". *Paleobiology* **8** (1982):4–15.

[31] Grüner, W., R. Giegerich, D. Strothmann, C. M. Reidys, J. Weber, I. L. Hofacker, P. F. Stadler, and P. Schuster, "Analysis of RNA sequence structure maps by exhaustive enumeration: I. Neutral networks; II. Struc-

ture of neutral networks and shape space covering". *Monatsh. Chem.* **127** (1996):355–389.

[32] Gutell, R., A. Power, G. Z. Hertz, E. J. Putz, and G. D. Stormo, "Identifying Constraints on the Higher-Order Structure of RNA: Continued Development and Application of Comparative Sequence Analysis Methods". *Nucleic Acids Res.* **20** (1993):5785–5795.

[33] Hanson, J. E. and J. P. Crutchfield, "Computational Mechanics of Cellular Automata: An Example". *Physica D* **103** (1997):169–189.

[34] Huynen, M., "Exploring Phenotype Space through Neutral Evolution". *J. Mol. Evol.* **43** (1996):165–169.

[35] Huynen, M., P. F. Stadler, and W. Fontana, "Smoothness within ruggedness: The role of neutrality in adaptation". *Proc. Natl. Acad. Sci. USA* **93** (1996):397–401.

[36] Huynen, M. A., T. Doerks, F. Eisenhaber, C. Orengo, S. Sunyaev, Y. Yuan, and P. Bork, "Homology-based fold predictions for *Mycoplasma genitalium* proteins". *J. Mol. Biol.* **280** (1998):323–326.

[37] Kauffman, S. A., *Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press, 1993.

[38] Kauffman, S. A. and S. Levin, "Towards a General Theory of Adaptive Walks in Rugged Fitness Landscapes". *J. Theo. Bio.* **128** (1987):11–45.

[39] Kimura, M., *The Neutral Theory of Molecular Evolution*. Cambridge University Press, 1983.

[40] Kolmogorov, A. N., "Three Approaches to the Concept of the Amount of Information". *Prob. Info. Trans.* **1** (1965):4–7.

[41] Landweber, L. F. and I. D. Pokrovskaya, "Emergence of a dual-catalytic RNA with metal-specific cleavage and ligase activities: The spandrels of RNA evolution". *Proc. Natl. Acad. Sci. USA* **96** (1999):173–178.

[42] Lenski, R. E., "Evolution in Experimental Populations of Bacteria". In *Population Genetics of Bacteria*, edited by S. Baumberg, J. P. W. Young, E. M. H. Wellington, and J. R. Saunders, 193–215, Cambridge, UK: Cambridge University Press, 1974.

[43] Li, M. and P. M. B. Vitanyi, *An Introduction to Kolmogorov Complexity and its Applications*. New York: Springer-Verlag, 1993.

[44] Macken, C. A. and A. S. Perelson, "Protein Evolution in Rugged Fitness Landscapes". *Proc. Nat. Acad. Sci. USA* **86** (1989):6191–6195.

[45] Martin-Lof, P., "The Definition of Random Sequences". *Info. Control* **9** (1966):602–619.

[46] Maynard-Smith, J., *Evolutionary Genetics*. Oxford: Oxford University Press, 1989.

[47] McShea, D., "Complexity and Evolution: What Everybody Knows". *Biology and Philosophy* **6** (1991):303–324.

[48] McShea, D., "Metazoan Complexity and Evolution: Is there a Trend?" *Evolution* **50** (1996):477–492.

[49] Mitchell, M., J. P. Crutchfield, and P. T. Hraber, "Evolving Cellular Automata to Perform Computations: Mechanisms and Impediments". *Physica D* **75** (1994):361–391.

[50] Mitchell, M., J. H. Holland, and S. Forrest, "When Will a Genetic Algorithm Outperform Hillclimbing?" In *Advances in Neural Information Processing Systems*, edited by J. D. Cowan, G. Tesauro, and J. Alspector, vol. 6, 51–56, San Mateo, CA: Morgan Kauffman, 1993.

[51] Monod, J., *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology*. New York: Vintage Books, 1971.

[52] Newman, M. and R. Engelhardt, "Effect of Neutral Selection on the Evolution of Molecular Species". *Proc. R. Soc. London B.* **256** (1998):1333–1338.

[53] Ohta, T. and J. Gillespie, "Development of Neutral and Nearly Neutral Theories". *Theo. Pop. Bio.* **49** (1996):128–142.

[54] Papadopoulos, D., D. Schneider, J. Meier-Eiss, W. Arber, R. E. Lenski, and M. Blot, "Genomic Evolution During a 10,000-Generation Experiment with Bacteria". *Proc. Natl. Acad. Sci. USA* **96** (1999):3807–3812.

[55] Plischke, M. and B. Bergensen, *Equilibrium Statistical Physics*. Englewood Cliffs, New Jersey: Prentice Hall, 1989.

[56] Ray, T. S., "An approach to the synthesis of life". In *Artificial Life II*, edited by C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, 371–408, Reading, MA: Addison-Wesley, 1992.

[57] Robertson, H. S., *Statistical Theromphysics*. Prentice Hall, 1993.

[58] Shalizi, C. R. and J. P. Crutchfield, "Computational Mechanics: Pattern and Prediction, Structure and Simplicity". *J. Stat. Phys.* (2001):in press.

[59] Shannon, C. E. and W. Weaver, *The Mathematical Theory of Communication*. Champaign-Urbana: University of Illinois Press, 1962.

[60] Thompson, D. W., *On Growth and Form*. Cambridge: Cambridge University Press, 1917.

[61] Upper, D. R., *Theory and Algorithms for Hidden Markov Models and Generalized Hidden Markov Models*. Ph.D. thesis, University of California, Berkeley, 1997, Published by University Microfilms Intl, Ann Arbor, Michigan.

[62] van Nimwegen, E., *Statistical Dynamics of Epochal Evolution*. Ph.D. thesis, Universiteit Utrecht, The Netherlands, 1999.

[63] van Nimwegen, E. and J. P. Crutchfield, "Metastable Evolutionary Dynamics: Crossing Fitness Barriers or Escaping via Neutral Paths?" *Bull. Math. Bio.* **62** (2000):799–848.

[64] van Nimwegen, E. and J. P. Crutchfield, "Optimizing Epochal Evolutionary Search: Population-Size Dependent Theory". *Machine Learning* (2001):in press.

[65] van Nimwegen, E., J. P. Crutchfield, and M. Mitchell, "Finite Populations Induce Metastability in Evolutionary Search". *Phys. Lett. A* **229** (1997):144–150.

[66] van Nimwegen, E., J. P. Crutchfield, and M. Mitchell, "Statistical Dynamics of the Royal Road Genetic Algorithm". *Theo. Comp. Sci.* **229** (1999):41–102, special issue on Evolutionary Computation, A. Eiben and G. Rudolph, editors.

[67] Waddington, C. H., *The Strategy of the Genes*. London: Allen and Unwin, 1957.

[68] Williams, G. C., *Natural Selection: Domains, Levels, and Challenges*. New York: Oxford University Press, 1992.

[69] Wright, M. C. and G. F. Joyce, "Continuous *in vitro* evolution of catalytic function". *Science* **276** (1997):614–617.

[70] Wright, S., "Character Change, Speciation, and the Higher Taxa". *Evolution* **36** (1982):427–443.

[71] Yeomans, J. M., *Statistical Mechanics of Phase Transitions*. Clarendon Press, Oxford, 1992.