Markov chains. The difficulty of finding the entropy rate for these processes was first noted in the late 1950's.[64] In fact, many questions about this class are very difficult. It is only recently, for example, that a procedure for determining the equivalence of two such processes has been given.[65] The awareness that this problem area bears on the complexity of observation and the result that finite complexity processes can appear infinitely complex is also recent.[59]

The new notion of state here that needs to be innovated involves the continuity of real variables. The causal states are no longer discrete. More precisely, an ϵ-machine state for a hidden Markov model is a *distribution* over the hidden states. Since this distribution is a vector of real numbers, the causal states are continuous. In addition, this vector is normalized since the components are state probabilities. The result is that the ϵ-machine states for these processes can be graphically represented as a point in a simplex. The simplex vertices themselves correspond to δ-distributions concentrated on one of the hidden Markov model states; the dimension of the simplex is one less than the number of hidden states. In the simplex representation the ϵ-machine models are stochastic deterministic automata (SDA). Only a subset of the simplex is recurrent. This subset corresponds to an attracting set under the dynamics of the agent's step-by-step prediction of the hidden state distribution based on the current data stream. Figure 17 shows the causal state simplices for three example processes.

The infinite, but countable state machine of Figure 15(a) is shown in this simplex representation in Figure 17(a) as a discrete point set in the 3-simplex. Each point, somewhat enlarged there, corresponds to one of the states in Figure 15. Two more examples of ϵ-machines are given in Figures 17(b) and 17(c). They suggest some of the richness of stochastic finite processes. In Figure 17(b), for example, the ϵ-machine has a partial continuum of states; the causal states lie on a "fractal". Figure 17(c) shows an ϵ-machine whose causal states limit on a full continuum of states.

The different classes shown in Figure 17 are distinguished by a new complexity measure of the ϵ-machines's state structure — the ϵ-machine dimension $d_{\epsilon M}$. $d_{\epsilon M}$ is the information dimension of the state distribution on the simplex. In the case of the countable state ϵ-machine of Figure 17(a) C_μ is finite due to the strong localization of the state distribution over the earlier states. But, since the states are a discrete point set, $d_{\epsilon M} = 0$. For the fractal and continuum SDAs the statistical complexity diverges, but $d_{\epsilon M} > 0$. $d_{\epsilon M}$ is noninteger in the first case and $d_{\epsilon M} = 2$ in the second. Further results will be presented elsewhere.[59]

The finitary stochastic hierarchy

The preceding examples of stochastic finitary processes can be summarized using the computational model hierarchy of Figure 18. This hierarchy borrows the finite memory machines of Figure 2 and indicates their stochastic generalizations. As before each ellipse denotes a model class. As one moves up the diagram classes become more powerful in the sense that they can finitely describe a wider range of processes than lower classes. A class below and connected to a given one can finitely describe only a subset of the processes finitely described by the higher one. Again, this hierarchy is only a partial ordering of descriptive capability. It should be emphasized that "descriptive capability" above the Measure-Support line refers to finitely representing a *distribution* over the sequences; not just the distribution's supporting formal languages.

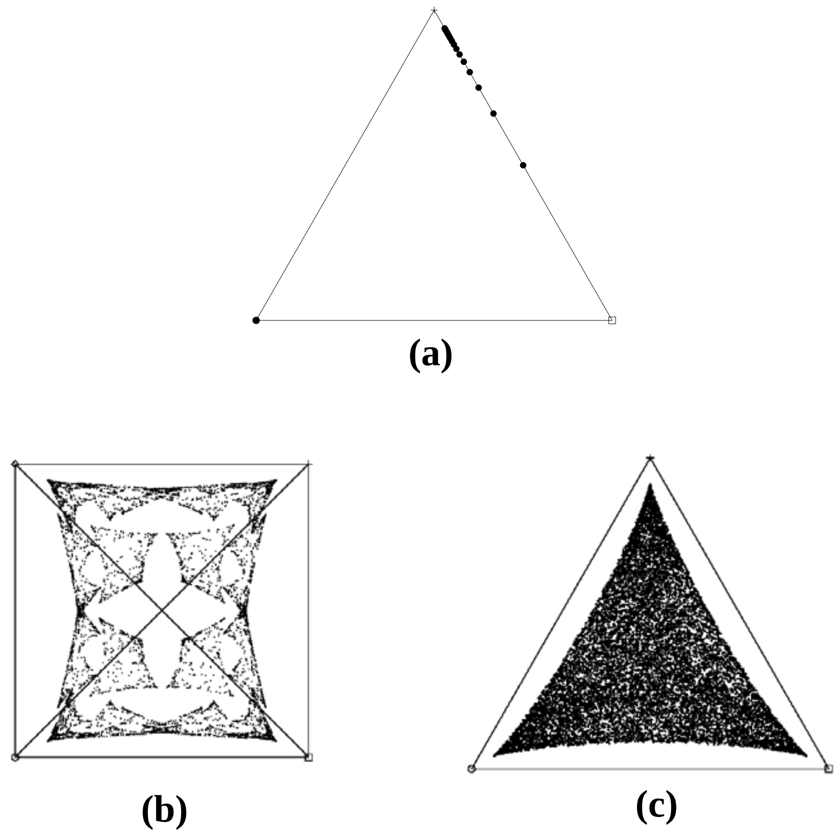

Figure 17 Stochastic deterministic automata (SDA): (a) Denumerable SDA: A denumerable ϵ-machine for the simple nondeterministic source of Figure 11. It is shown here in the (two-dimensional) 3–simplex defining its possible deterministic states (indicated with enlarged dots). Since the state probability decays exponentially, the simulation only shows a very truncated part of the infinite chain of states that, in principle, head off toward the upper vertex. Those dots correspond to the 1s backbone of Figure 15. The state on the lower lefthand vertex corresponds to the "reset" state $1A0B$ in that figure. (b) Fractal SDA: A nondenumerable fractal ϵ-machine shown in the 4–simplex defining the possible deterministic states. (c) Continuum SDA: A nondenumerable continuum ϵ-machine shown in the 3–simplex defining the possible deterministic states.

In formal language theory it is well-known that deterministic finite automata (DFA) are as powerful as nondeterministic finite automata (NFA).[31] This is shown in the hierarchy as both classes being connected at the same height. But the equivalence is just topological; that is, it concerns only the descriptive capability of each class for sets of observed sequences. If one augments these two classes, though, to account for probabilistic structure over the sequences, the equivalence is broken in a dramatic way — as the above example for the "mismeasured" logistic map demonstrated. This is shown in Figure 18 above the Measure-Support line. The class of SNFA is higher than that of the stochastic deterministic finite automata (SDFA). Crudely speaking, if a DFA has transition probabilities added to its edges, one obtains the single class of SDFA. But if transition probabilities are added to an NFA, then the class is qualitatively more powerful and, as it turns out, splits into three distinct classes. Each of these classes is more powerful than the SDFA class. The new causal classes — called stochastic deterministic automata (SDA) — are distinguished by having a countable infinity, a fractional continuum, or a full continuum of causal states: the examples of Figure 17.

Initially, the original logistic map process as represented in Figure 11 was undistinguished as an SNFA. Via the analysis outlined above its causal representation showed that it is equivalent

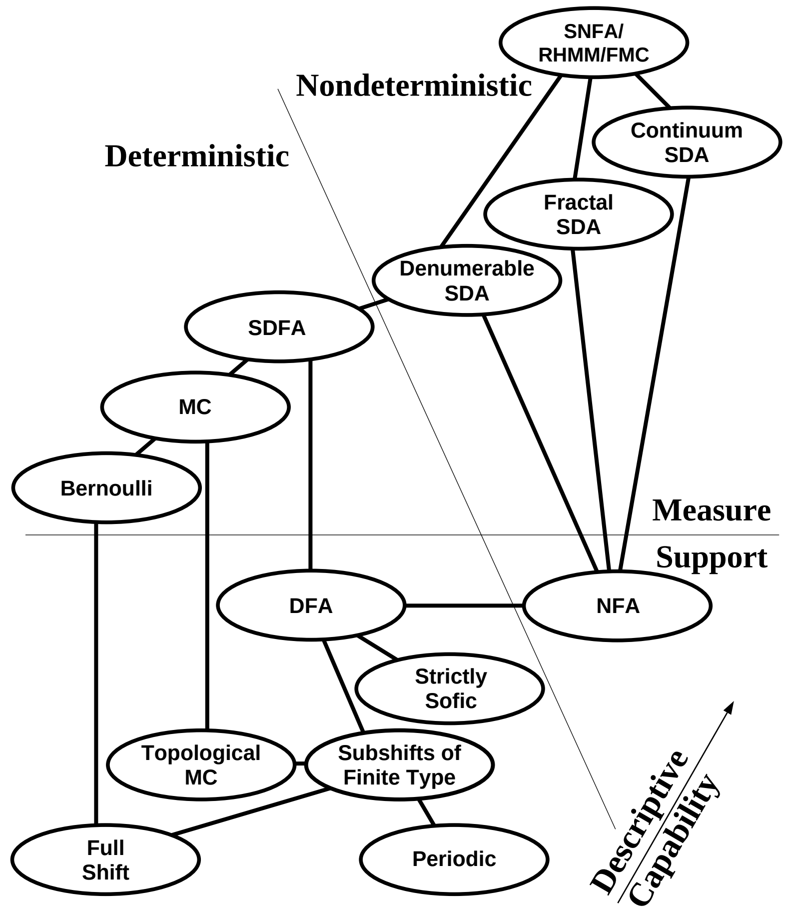

Figure 18 The computational hierarchy for finite-memory nonstochastic (below the Measure-Support line) and stochastic discrete processes (above that line). The nonstochastic classes come from Figure 2, below the Finite-Infinite memory line. Here "Support" refers to the sets of sequences, i.e. formal languages, which the "topological" machines describe; "Measure" refers to sequence probabilities, i.e. what the "stochastic" machines describe. The abbreviations are: A is automaton, F is finite, D is deterministic, N is nondeterministic, S is stochastic, MC is Markov chain, HMM is hidden Markov model, RHMM is recurrent HMM, and FMC is function of MC.

to a denumerable stochastic deterministic automaton (DSDA). Generally, in terms of descriptive power $\mathrm{DSDA} \subset \mathrm{SNFA}$ as Figure 18 emphasizes. Recall that we interpret the SNFA as the environment, which is a Markov chain (MC), viewed through the agent's sensory apparatus. So the computational class interpretation of the complexity divergence is that $\mathrm{MC} \rightarrow \mathrm{DSDA}$ under measurement distortion. That is, MC and DSDA are qualitatively different classes and, in particular, $\mathrm{MC} \subset \mathrm{DSDA}$, as shown in the hierarchy of Figure 18. The representational divergence that separates them is characteristic of the transition from a lower to a higher class.

5.3 The costs of spatial coherence and distortion

As an example of hierarchical structure and innovation for spatial processes this section reviews an analysis of two cellular automata (CA). CA are arguably the simplest dynamical systems with which to study pattern formation and spatio-temporal dynamics. They are discrete in time, in space, and in local state value. The two examples considered have two local states

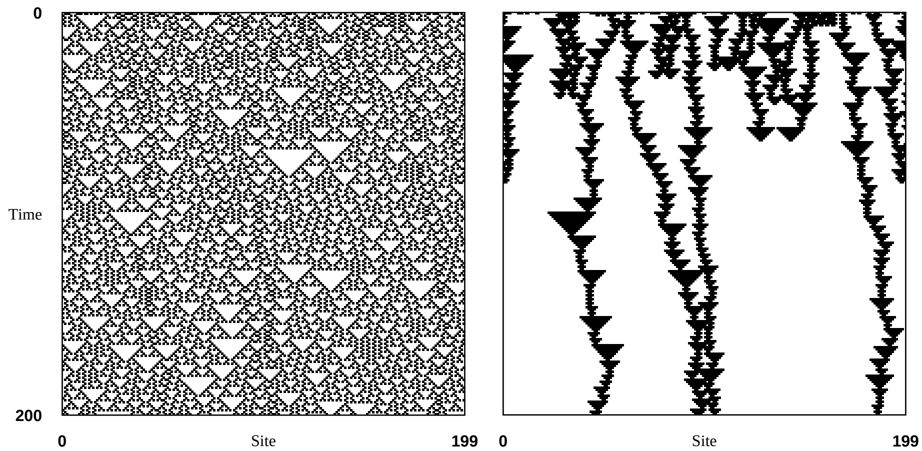

Figure 19 (a) Elementary cellular automaton 18 evolving over 200 time steps from an initial arbitrary pattern on a lattice of 200 sites. (b) The filtered version of the same space-time diagram that reveals the diffusive-annihilating dislocations obscured in the original. (After Ref. [66].)

at each site of a one-dimensional spatial lattice. The local state at a site is updated according to a rule that looks only at itself and its two nearest neighbors.

Figure 19(a) shows the temporal development of an arbitrary initial pattern under the action of elementary CA 18 (ECA 18). This rule maps all of the neighborhood patterns to 0; except 001 and 100, which are mapped to 1. A space-time diagram for elementary CA 54 (ECA 54) is shown in Figure 20(a).

ϵ-machine reconstruction applied to the patterns reveals much of the internal structure of these systems' state spaces and the intrinsic computation in their temporal development of spatial patterns.[60] An important component of CA patterns are domains. CA domains are dynamically homogeneous regions in space-time defined in terms of the same set of pattern regularities, such as "every other site value is a 0". From knowledge of a CA's domains nonlinear filters can be constructed to show the motion and interaction of domains, walls between domains, and particles. The result of this filtering process — shown in Figures 19(b) and 20(b) — is a higher level representation of the original space-time diagrams in Figures 19(a) and 20(a). This new level filters out chaotic (ECA 18) and periodic (ECA 54) backgrounds to highlight the various propagating space-time structures — dislocations, particles, and their interactions that are hidden in the unfiltered space-time diagrams.

In turns out that ECA 18 is described by a single type of chaotic domain — "every other local state is 0, otherwise they are random" — and a single type of domain wall — a dislocation that performs a diffusive annihilating random walk.[66] At the level of particles one discovers an irreducibly stochastic description — the particle motion has vanishing statistical complexity. Because the description of ECA 18 is finite at this level, one can stop looking further for intrinsic computational structure. The correct ϵ-machine level, that providing a finite description, has been found by filtering the chaotic domains.

In ECA 54 there is a single periodic domain with a diversity of particles that move and interact in complicated — aperiodic, but nonstochastic — ways. ECA 54 is more complex that ECA 18, in the sense that even at the level of particles ECA 54's description may not be finite. Thus for ECA 54, one is tempted to look for a higher level representation by performing

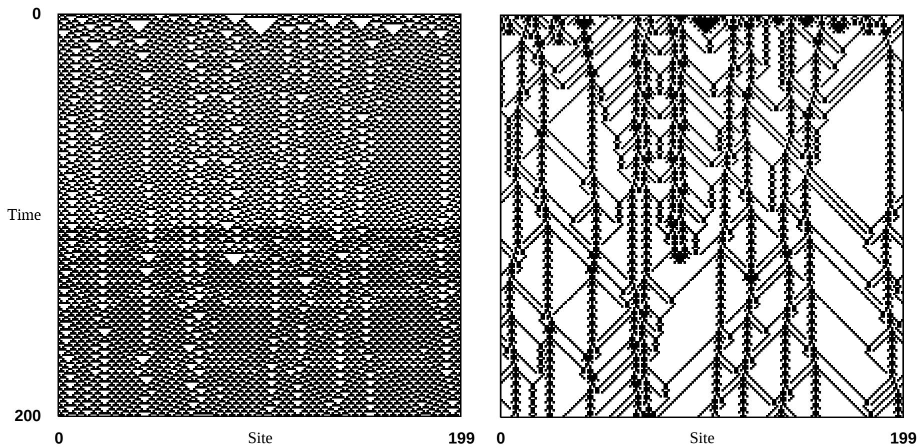

Figure 20 (a) Elementary cellular automaton 54 evolving over 200 time steps from an initial arbitrary pattern on a lattice of 200 sites. (b) The filtered version of the same space-time diagram that reveals a multiplicity of different particle types and interactions. (From Ref. [67]. Reprinted with permission of the author. Cf. [68].)

ϵ-machine reconstruction at this "particle" level to find further regularity or stochasticity and ultimately to obtain a finite description.

These hierarchical space-time structures concern different levels of information processing within a single CA. In ergodic theory terms, the unfiltered spatial patterns are nonstationary — both spatially and temporally. With the innovation of domains and particles, the new level of representation allows for stationary elements, such as the domains, to be separated out from elements which are nonstationary. For ECA 18, discovering the dislocations lead to a stochastic model based on a diffusive annihilating random walk that could be analyzed in some detail and which explained the nonstationary aspects of the dislocation motion. For ECA 54 the description at this level is still not completely understood.

Intrinsic computation in spatio-temporal processes raises a number of interesting problems. Aside from the above, there is an analog for spatio-temporal processes of the apparent complexity explosion in the stochastic nondeterministic processes. This led to the introduction of a new hierarchy for spatial automata, in which CA are the least capable, that accounts for spatial measurement distortion.[29] This spatial discrete computation hierarchy is expressed in terms of automata rather than grammars.[63]

5.4 What to glean from the examples

Understanding the hierarchical structure embedded in the preceding examples required crossing a model class boundary: from determinism to indeterminism, from finitary support to finitary measure, from predictability to chaos, from undifferentiated patterns to domains and particles, and from observed states to hidden internal states. Each of these transitions is a phase transition in the sense that there is a divergence in a macroscopic observable — the representation's complexity — and in that the finitely-describable behaviors on either side of the divergence are qualitatively distinct. Figure 21 attempts to summarize in a very schematic way their relationship.

The point that the examples serve to make is that innovation of new computational structures is required at each step. Just what needs to be innovated can be determined in very large measure by the state-grouping step in hierarchical ϵ-machine reconstruction. As was stated before, beyond the state-grouping it seems that there will always be some undetermined aspect during the innovation step required of the agent. The undetermined aspect has been isolated by the algorithmic portion of hierarchical ϵ-machine reconstruction that groups lower level states into a higher level state. To the extent it is undetermined, though, this step can be the locus of a now highly-constrained search over a space of new mechanisms. In any evolutionary process, such a space would be largely circumscribed by the agents's internal structure and the search, by how the latter can be modified. In this way, hierarchical ϵ-machine reconstruction reduces the search space for innovation by a substantial fraction and indicates where random search could be most informative and effective.

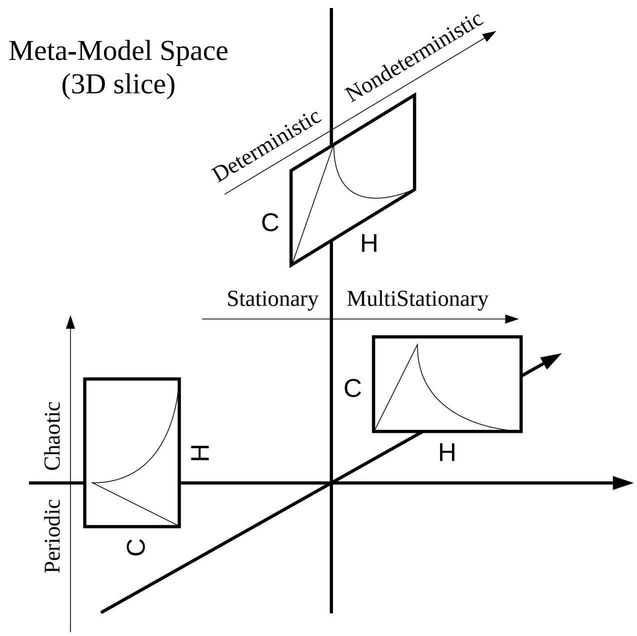

Figure 21 A schematic summary of the three examples of hierarchical learning in metamodel space. Innovation across transitions from periodic to chaotic, from stochastic deterministic to stochastic nondeterministic, and from spatial stationary to spatial multistationary processes were illustrated. The finite-to-infinite memory coordinate from Figure 2 is not shown. The periodic to chaotic and deterministic to nondeterministic transitions were associated with the innovation of infinite models from finite ones. The complexity (C) versus entropy (H) diagrams figuratively indicate the growth of computational resources that occurs when crossing the innovation boundaries.

PART IV
OBSERVATIONS AND HYPOTHESES

1 Complexity as the Interplay of Order and Chaos

Neither order nor randomness are sufficient in and of themselves for the emergence of complexity. Nor is their naive juxtaposition. As Alfred North Whitehead and many before and after him knew, order and randomness are mere components in an ongoing process. That process *is* the subtlety in the problem of emergence. For Heraclitus, the persistence of change was the by-product of an "attunement of opposite tensions".[69] Natural language, to take one ever-present example, has come to be complex because utility required it to be highly informative *and* comprehensibility demanded a high degree of structure. For the process that is human language, emergence is a simple question: How do higher semantic layers appear out of this tension?

As far as the architecture of information processing is concerned, these questions for natural language have direct analogues in adaptation, evolution, and even in the development of scientific theories.[70] While conflating these fields suggests a possible general theory of the emergence of complexity, at the same time it reveals an ignorance of fundamental distinctions. Due to the sheer difficulty of the nonlinear dynamics involved and to the breadth of techniques that must be brought to bear, one must be particularly focused on concrete results and established theory in discussing "emergence".

In spite of receiving much popular attention, the proposals for "computation at the edge of chaos",[71] "adaptation toward the edge of chaos",[72] and "life at the edge of chaos"[73] are recent examples of the problems engendered by this sort of blurring of basic distinctions. Despite the preceding sections' review of the complexity versus entropy spectrum and their analyses of processes in which higher levels of computation arise at the onset of chaos, there is absolutely no general need for high computational capability to be near an "edge of chaos". The infinite-memory counter register functionality embedded in the highly chaotic hidden Markov model of Figure 11 is a clear demonstration of this simple fact. More to the point, that stability and order are necessary for information storage, on the one hand, and that instability is necessary for the production of information and its communication, on the other hand, are basic requirements for any nontrivial information processing. The trade-off between these requirements is much of what computation theory is about. Moreover, natural systems are not constrained in some general way to move toward an "edge of chaos". For example, the very informativeness ($h_\mu > 0$) of natural language means, according to Shannon, that natural language is very far and must stay away from any "edge of chaos" ($h_\mu = 0$).

A critique of the first two "edge of chaos" proposals, which concern cellular automata, can be found elsewhere.[74] Aside from the technical issues discussed there, all three proposals exhibit a fatal conceptual difficulty. To the extent that they presume to address the emergence of complexity, with reference neither to intrinsic computation nor to the innovation of new information-processing architectures, they can provide no grounding for the key concepts — "order", "chaos", "complexity", "computation" — upon which their arguments rely.

The preceding development of statistical complexity, hierarchical ϵ-machine reconstruction, and various extensions to discrete computation theory is an alternative. The presentation gave a broad overview of how three tool sets — the calculi of emergence: computation, dynamics, and innovation — interrelate in three different problem areas: various routes to chaos, measurement distortion and nondeterminism, and cellular automaton pattern formation. It suggested a way to frame the question of how structure appears out of the interplay of order and randomness. It demonstrated methods for detecting and metrics for quantifying that emergence. It showed how the information processing structure of nonlinear processes can be analyzed in terms of computational models and in terms of the effective information processing.

2 Evolutionary Mechanics

The arguments to this point can be recapitulated by an operational definition of emergence. A process undergoes emergence if at some time the architecture of information processing has changed in such a way that a distinct and more powerful level of intrinsic computation has appeared that was not present in earlier conditions.

It seems, upon reflection, that our intuitive notion of emergence is not captured by the "intuitive definition" given in Part I. Nor is it captured by the somewhat refined notion of pattern formation. "Emergence" is a groundless concept unless it is defined within the context of processes themselves; the only well-defined (nonsubjective) notion of emergence would seem to be intrinsic emergence. Why? Simply because emergence defined without this closure leads to an infinite regress of observers detecting patterns of observers detecting patterns .... This is not a satisfactory definition, since it is not finite. The regress must be folded into the system, it must be immanent in the dynamics. When this happens complexity and structure are no longer referred outside. No longer relative and arbitrary, they take on internal meaning and functionality.

Where in science might a theory of intrinsic emergence find application? Are there scientific problems that at least would be clarified by the computational view of nature outlined here?

In several ways the contemporary debate on the dominant mechanisms operating in biological evolution seems ripe. Is anything ever new in the biological realm? The empirical evidence is interpreted as a resounding "yes". It is often heard that organisms today are more complex than in earlier epochs. But how did this emergence of complexity occur? Taking a long view, at present there appear to be three schools of thought on what the guiding mechanisms are in Darwinian evolution that produce the present diversity of biological structure and that are largely responsible for the alteration of those structures.

The Selectionists hold 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.[75] The second, anarchistic camp consists of the Historicists who hold fast to the Darwinian mechanisms of selection and variation, but emphasize the accidental determinants of biological form.[76,77] What distinguishes this position from the Selectionists is the claim that major changes in structure can be and have been nonadaptive. Lastly, there are the Structuralists whose goal is to elucidate the "principles of organization" that guide the appearance of biological structure. They contend that energetic, mechanical, biomolecular, and morphogenetic constraints limit the infinite range of possible biological form.[16,73,78–81] The constraints result in a relatively small set of structural attractors. Darwinian evolution serves, at best, to fill the waiting attractors or not depending on historical happenstance.

What is one to think of these conflicting theories of the emergence of biological structure? The overwhelming impression this debate leaves is that there is a crying need for a theory of biological structure and a qualitative dynamical theory of its emergence.[82] In short, the tensions between the positions are those (i) between the order induced by survival dynamics and the novelty of individual function and (ii) between the disorder of genetic variation and the order of developmental processes. Is it just an historical coincidence that the structuralist-selectionist dichotomy appears analogous to that between order and randomness in the realm of modeling? The main problem, at least to an outsider, does not reduce to showing that one or the other view is correct. Each employs compelling arguments and often empirical data as a starting point. Rather, the task facing us reduces to developing a synthetic theory that balances the tensions between the viewpoints. Ironically, evolutionary processes themselves seem to do just this sort of balancing, dynamically.

The computational mechanics of nonlinear processes can be construed as a theory of structure. Pattern and structure are articulated in terms of various types of machine classes. The overall mandate 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 transformation of structure. It suggests one mechanism with which to study what drives and what constrains the appearance of novelty. The next step, of course, would be to fold hierarchical ϵ-machine reconstruction into the system itself, resulting in a dynamics of innovation, the study of which might be called "evolutionary mechanics".

The prototype universe of Figure 1 is the scaffolding for studying an abstract "evolutionary mechanics", which is distinguished from chemical, prebiotic, and biological evolution. That is, all of the "substrates" indicated in Figure 1 are thrown out, leaving only those defined in terms of information processing. This more or less follows the spirit of computational evolution.[83,84] The intention is to expunge as much disciplinary semantic content as possible so that if novel structure emerges, it does so neither by overt design nor interpretation but (i) via the dynamics of interaction and induction and (ii) according to the basic constraints of information processing. Additionally, at the level of this view mechanisms for genotypic and phenotypic change are not delineated. In direct terms evolutionary mechanics concerns the change in the information-processing architecture of interacting adaptive agents. The basic components that guide and constrain this change are the following.

1. *Modeling*: Driven by the need for an encapsulation of environmental experience, an agent's internal model captures its knowledge, however limited, of the environment's current state and persistent structure.
2. *Computation*: Driven by the need to process sensory information and produce actions, computation is the adaptive agent's main activity. The agent's computational resources delimit its inferential, predictive, and semantic capabilities. They place an upper bound on the maximal level of computational sophistication. Indirectly, they define the language with which the agent expresses its understanding of the environment's structure. Directly, they limit the amount of history the agent can store in its representation of the environment's current state.
3. *Innovation*: Driven by limited observational, computational, and control resources, innovation leads to new model classes that use the available resources more efficiently or more parsimoniously.

The computational mechanics approach to emergence attempted to address each component in turn, though it did not do so in the fully dynamic setting suggested by the prototype universe. But evolutionary mechanics is concerned specifically with this dynamical problem and, as such, it leads to a much wider range of questions.

At the very least, the environment appears to each agent as a hierarchical process. There are subprocesses at several different scales, such as those at the smallest scale where the agents are individual stochastic dynamical systems and those at larger scales at which coordinated global behavior may emerge. Each agent is faced with trying to detect as much structure as possible within this type of environment. How can an evolutionary system adapt when confronted with this kind of hierarchical process?

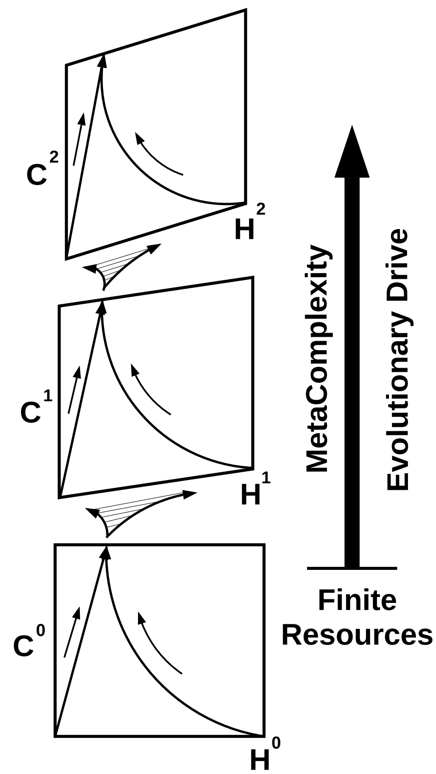

Figure 22 Schematic diagram of an evolutionary hierarchy in terms of the changes in information-processing architecture. An open-ended sequence of successively more sophisticated computational classes are shown. The evolutionary drive up the hierarchy derives from the finiteness of resources to which agents have access. The complexity-entropy diagrams are slightly rotated about the vertical to emphasize the difference in meaning at each level via a different orientation. (Cf. Table 1.)

Figure 22 gives a very schematic summary of one of the consequences following from this framework in terms of successive changes in information-processing architecture. The underlying representation of the entire prototype universe, of course, is the realization space of nonlinear processes — the orbit space of dynamical systems. The figure shows instead a temporal analog of the discrete computation hierarchy of Figure 2 in which the trade-offs within each computational level are seen through the complexity-entropy plane. The overall evolutionary dynamic derives from agent interaction and survival. Species correspond in the realization space to metastable invariant languages: they are temporary (hyperbolic) fixed points of the evolutionary dynamic. Each species is dual to some subset of environmental regularities — a niche — whose defining symmetries some subpopulation of agents has been able to discover and incorporate into their internal models through innovation. Innovation then manifests itself as a kind of speciation. The invariant sets bifurcate into computationally distinct agents with new, possibly-improved capabilities. The statistical mechanical picture of this speciation is that of a phase transition between distinct computational "thermodynamic" phases. Through all of this the macro-evolutionary observables — entropy rate h_μ, statistical complexity C_μ, and innovation $\mathcal{I}$ — monitor changes in the agents' and the universe's information-processing architecture.

Evolutionary mechanics seems to suggest that the emergence of natural complexity is a manifestly open-ended process. One force of evolution appears in this as a movement up the inductive hierarchy through successive innovations. This force of evolution itself is driven at each stage by the limited availability of resources.

Some new principle of refreshment is required. The art of progress is to preserve order amid change, and to preserve change amid order. Life refuses to be embalmed alive.

A. N. Whitehead in **Process and Reality**.[1]

Acknowledgments

Many thanks are due to Lisa Borland, Don Glaser, Jim Hanson, Blake LeBaron, Dan McShea, Melanie Mitchell, Dan Upper, and Karl Young for helpful discussions and comments on early drafts. This work was partially funded under AFOSR 91-0293 and ONR N00014-92-4024. For the opportunity to present this work thanks also are due to the organizers of the Oji International Seminar on *Complex Systems — from Complex Dynamics to Artificial Reality* held 5 - 9 April 1993 Numazu, Japan; to the organizers of the Santa Fe Institute *Integrative Themes of Complex Adaptive Systems* Meeting held on 8 July 1992 in Santa Fe, New Mexico; and to the organizers of the *Artificial Life 2* conference, held during 5 - 9 February 1990.

Bibliography

1. A. N. Whitehead. *Process and Reality*. The Free Press, New York, corrected edition, 1978.
2. C. W. Reynolds. Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics*, 21:25 – 34, 1987.
3. B. Holldobler and E. O. Wilson. *The Ants*. Belknap Press of Harvard University Press, Cambridge, Massachusetts, 1990.
4. E. F. Fama. Efficient capital markets II. *J. Finance*, 46:1575 – 1617, 1991.
5. E. Land. The retinex. *Am. Scientist*, 52:247 – 264, 1964.
6. B. A. Wandell. Color appearance: The effects of illumination and spatial pattern. *Proc. Nat. Acad. Sci.*, 10:2458 – 2470, 1993.
7. I. Kovacs and B. Julesz. A closed curve is much more than an incomplete one - effect of closure in figure ground segmentation. *Proc. Nat. Acad. Sci.*, 90:7495 – 7497, 1993.
8. J. P. Crutchfield, N. H. Packard, J. D. Farmer, and R. S. Shaw. Chaos. *Sci. Am.*, 255:46 – 57, 1986.
9. H. L. Swinney and J. P. Gollub, editors. *Hydrodynamic Instabilities and the Transition to Turbulence*, Berlin, 1981. Springer Verlag.
10. A. M. Turing. The chemical basis of morphogenesis. *Trans. Roy. Soc., Series B*, 237:5, 1952.
11. A. T. Winfree. *The Geometry of Biological Time*. Springer-Verlag, Berlin, 1980.
12. Q. Ouyang and H. L. Swinney. Transition from a uniform state to hexagonal and striped Turing patterns. *Nature*, 352:610 – 612, 1991.
13. H. E. Stanley. *Introduction to Phase Transitions and Critical Phenomena*. Oxford University Press, Oxford, 1971.
14. P. Bak and K. Chen. Self-organized criticality. *Physica A*, 163:403 – 409, 1990.
15. J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman. *The Theory of Critical Phenomena*. Oxford University Press, Oxford, 1992.
16. D. W. Thompson. *On Growth and Form*. Cambridge University Press, Cambridge, 1917.
17. H. Meinhardt. *Models of Biological Pattern Formation*. Academic Press, London, 1982.
18. S. Forrest. Emergent computation: Self-organizing, collective, and cooperative behavior in natural and artificial computing networks: Introduction to the Proceedings of the Ninth Annual CNLS Conference. *Physica D*, 42:1 – 11, 1990.
19. J. G. Kemeny. The use of simplicity in induction. *Phil. Rev.*, 62:391, 1953.
20. C. S. Wallace and D. M. Boulton. An information measure for classification. *Comput. J.*, 11:185, 1968.
21. J. Rissanen. Modeling by shortest data description. *Automatica*, 14:462, 1978.
22. J. P. Crutchfield and B. S. McNamara. Equations of motion from a data series. *Complex Systems*, 1:417 – 452, 1987.
23. J. Rissanen. *Stochastic Complexity in Statistical Inquiry*. World Scientific, Singapore, 1989.
24. J. P. Crutchfield and K. Young. Inferring statistical complexity. *Phys. Rev. Let.*, 63:105 – 108, 1989.

25. S. Wolfram. Computation theory of cellular automata. *Comm. Math. Phys.*, 96:15, 1984.
26. L. Blum, M. Shub, and S. Smale. On a theory of computation over the real numbers. *Bull. AMS*, 21:1, 1989.
27. M. G. Nordahl. Formal languages and finite cellular automata. *Complex Systems*, 3:63, 1989.
28. J. E. Hanson and J. P. Crutchfield. The attractor-basin portrait of a cellular automaton. *J. Stat. Phys.*, 66:1415, 1992.
29. J. P. Crutchfield. Unreconstructible at any radius. *Phys. Lett. A*, 171:52 – 60, 1992.
30. C. Moore. Real-valued, continuous-time computers: A model of analog computation, part I. Technical Report 93-04-018, Santa Fe Institute, 1993.
31. J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Reading, 1979.
32. H. R. Lewis and C. H. Papadimitriou. *Elements of the Theory of Computation*. Prentice-Hall, Englewood Cliffs, N.J., 1981.
33. J. G. Brookshear. *Theory of computation: formal languages, automata, and complexity*. Benjamin/Cummings, Redwood City, California, 1989.
34. J. P. Crutchfield. Semantics and thermodynamics. In M. Casdagli and S. Eubank, editors, *Nonlinear Modeling and Forecasting*, volume XII of *Santa Fe Institute Studies in the Sciences of Complexity*, pages 317 – 359, Reading, Massachusetts, 1992. Addison-Wesley.
35. A. V. Aho. Indexed grammars - an extension of context-free grammars. *J. Assoc. Comp. Mach.*, 15:647, 1968.
36. A. V. Aho. Nested stack automata. *J. Assoc. Comp. Mach.*, 16:383, 1969.
37. B. Weiss. Subshifts of finite type and sofic systems. *Monastsh. Math.*, 77:462, 1973.
38. A. Walker. Adult languages of L Systems and the Chomsky hierarchy. In G. Rozenberg and A. Salomaa, editors, *L Systems*, volume 15, page 201, 1974.
39. A. Lindenmayer and P. Prusinkiewicz. Developmental models of multicellular organisms: A computer graphics perspective. In C. G. Langton, editor, *Artificial Life*, page 221, Redwood City, California, 1989. Addison-Wesley.
40. G. Rozenberg and A. Salomaa, editors. *The Book of L*. Springer-Verlag, Berlin, 1986.
41. E. Allevi, A. Cherubini, and S. Crespi-Reghizzi. Breadth-first context-free grammars and queue automata. *Lect. Notes in Comp. Sci.*, 324:162, 1988.
42. A. Cherubini, C. Citrini, S. Crespi-Reghizzi, and D. Mandrioli. Quasi-real-time fifo automata, breadth first grammars, and their relations. *Theo. Comp. Sci.*, 85:171 – 203, 1991.
43. H. Rogers. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill, New York, 1967.
44. D. B. Searls. The linguistics of DNA. *Am. Sci.*, 80:579 – 591, 1992.
45. G. Chaitin. On the length of programs for computing finite binary sequences. *J. ACM*, 13:145, 1966.
46. A. N. Kolmogorov. Three approaches to the concept of the amount of information. *Prob. Info. Trans.*, 1:1, 1965.
47. R. J. Solomonoff. A formal theory of inductive control. *Info. Control*, 7:224, 1964.

48. C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Champaign-Urbana, 1962.
49. A. A. Brudno. Entropy and the complexity of the trajectories of a dynamical system. *Trans. Moscow Math. Soc.*, 44:127, 1983.
50. J. Ford. How random is a coin toss? *Physics Today*, April 1983.
51. C. H. Bennett. Dissipation, information, computational complexity, and the definition of organization. In D. Pines, editor, *Emerging Syntheses in the Sciences*. Addison-Wesley, Redwood City, 1988.
52. J. P. Crutchfield and N. H. Packard. Symbolic dynamics of noisy chaos. *Physica*, 7D:201 – 223, 1983.
53. R. Shaw. *The Dripping Faucet as a Model Chaotic System*. Aerial Press, Santa Cruz, California, 1984.
54. P. Grassberger. Toward a quantitative theory of self-generated complexity. *Intl. J. Theo. Phys.*, 25:907, 1986.
55. D. Angluin and C. H. Smith. Inductive inference: Theory and methods. *Comp. Surveys*, 15:237, 1983.
56. J. P. Crutchfield and K. Young. Computation at the onset of chaos. In W. Zurek, editor, *Entropy, Complexity, and the Physics of Information*, volume VIII of *SFI Studies in the Sciences of Complexity*, pages 223 – 269, Reading, Massachusetts, 1990. Addison-Wesley.
57. J. P. Crutchfield. Knowledge and meaning ... chaos and complexity. In L. Lam and V. Naroditsky, editors, *Modeling Complex Phenomena*, pages 66 – 101, Berlin, 1992. Springer-Verlag.
58. J. P. Crutchfield. Reconstructing language hierarchies. In H. A. Atmanspracher and H. Scheingraber, editors, *Information Dynamics*, page 45, New York, 1991. Plenum.
59. J. P. Crutchfield and D. R. Upper. In preparation, 1994.
60. J. P. Crutchfield and J. E. Hanson. Turbulent pattern bases for cellular automata. *Physica D*, 69:279 – 301, 1993.
61. M. J. Feigenbaum. Universal behavior in nonlinear systems. *Physica*, 7D:16, 1983.
62. L. P. Kadanoff, M. J. Feigenbaum, and S. J. Shenker. Quasiperiodicity in dissipative systems: a renormalization group analysis. *Physica*, 5D:370, 1982.
63. J. P. Crutchfield. Observing complexity and the complexity of observation. In H. Atmanspacher, editor, *Inside versus Outside*, pages 235 – 272, Berlin, February 1994. Springer-Verlag.
64. D. Blackwell and L. Koopmans. On the identifiability problem for functions of Markov chains. *Ann. Math. Statist.*, 28:1011, 1957.
65. H. Ito, S.-I. Amari, and K. Kobayashi. Identifiability of hidden Markov information sources and their minimum degrees of freedom. *IEEE Info. Th.*, 38:324, 1992.
66. J. P. Crutchfield and J. E. Hanson. Attractor vicinity decay for a cellular automaton. *CHAOS*, 3(2):215–224, 1993.
67. J. E. Hanson. *Computational Mechanics of Cellular Automata*. PhD thesis, University of California, Berkeley, 1993. Published by University Microfilms Intl, Ann Arbor, Michigan.

68. N. Boccara, J. Nasser, and M. Roger. Particle-like structures and their interactions in spatio-temporal patterns generated by one-dimensional deterministic cellular automaton rules. *Phys. Rev. A*, 44:866 – 875, 1991.

69. B. Russell. *A History of Western Philosophy*. Simon and Schuster, New York, 1945.

70. T. S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, 1962.

71. C. G. Langton. Computation at the edge of chaos: Phase transitions and emergent computation. In S. Forrest, editor, *Emergent Computation*, page 12. North-Holland, Amsterdam, 1990.

72. N. H. Packard. Adaptation toward the edge of chaos. In A. J. Mandell J. A. S. Kelso and M. F. Shlesinger, editors, *Dynamic Patterns in Complex Systems*, pages 293 – 301, Singapore, 1988. World Scientific.

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

74. M. Mitchell, P. Hraber, and J. P. Crutchfield. Revisiting the edge of chaos: Evolving cellular automata to perform computations. *Complex Systems*, 7:89 – 130, 1993.

75. J. Maynard-Smith. *Evolutionary Genetics*. Oxford University Press, Oxford, 1989.

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

77. S. J. Gould. *Wonderful Life*. Norton, New York, 1989.

78. C. H. Waddington. *The Strategy of the Genes*. Allen and Unwin, London, 1957.

79. B. Goodwin. Evolution and the generative order. In B. Goodwin and P. Sanders, editors, *Theoretical Biology: Epigenetic and Evolutionary Order from Complex Systems*, pages 89 – 100, Baltimore, Maryland, 1992. Johns Hopkins University Press.

80. W. Fontana and L. Buss. What would be conserved if the tape were played twice? *Proc. Nat. Acad. Sci.*, 91:757 – 761, 1994.

81. W. Fontana and L. Buss. “The Arrival of the Fittest”: Toward a theory of biological organization. *Bull. Math. Bio.*, 56:1 – 64, 1994.

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

83. L. J. Fogel, A. J. Owens, and M. J. Walsh, editors. *Artificial Intelligence through Simulated Evolution*. Wiley, New York, 1966.

84. J. H. Holland. *Adaptation in Natural and Artificial Systems*. MIT Press, Cambridge, MA, 1992. Second edition (First edition, 1975).