What would be conserved if “the tape were played twice”?  
(evolution/organization/self-maintenance/hierarchy/\(\lambda\)-calculus)  

WALTER FONTANA* AND LEO W. BUSS††  

*Santa Fe Institute, 1660 Old Pecos Trail, Santa Fe, NM 87501; and Departments of 7Biology and of 9Geology and Geophysics, Yale University, New Haven, CT 06511  

Communicated by Murray Gell-Mann, September 20, 1993  

ABSTRACT  
We develop an abstract chemistry, implemented in a \(\lambda\)-calculus-based modeling platform, and argue that the following features are generic to this particular abstraction of chemistry; hence, they would be expected to reappear if “the tape were run twice”: (i) hypercycles of self-reproducing objects arise; (ii) if self-replication is inhibited, self-maintaining organizations arise, and (iii) self-maintaining organizations, once established, can combine into higher-order self-maintaining organizations.  

Gould (1) has asked the question whether the biological diversity that now surrounds us would be different if “the tape were played twice.” If we had the option of observing a control earth, would we observe, say, the evolution of \(Homo sapiens\) or the evolution of something unambiguously identifiable as a metazoan or even something akin to a eukaryote? The question is important in that it focuses attention on the fact that historical progressions, such as the history of life, are the product of both contingency and necessity. While Gould’s (1) emphasis on the contingent is well taken, one nevertheless has the sense that certain features would recur. What are those features and how might we discover them?  

The fundamental difficulty with analysis of the questions of contingency and necessity in the distant past is the very fact that they occurred in the distant past. Experiments today cannot be performed with systems as they might have existed billions of years ago. The only alternative is to establish a model universe in which such an exploration is possible. In such a universe, one may unambiguously demonstrate whether the appearance of a given result is necessary or contingent. The question of the validity of such a claim may then be rigorously challenged by questioning the abstractions upon which the model is based or by introducing increasingly realistic elaborations of the model universe.  

A model universe designed to explore what is contingent in the history of life cannot assume the prior existence of organisms. The approach must seek to establish how biological organizations are generated. In this communication, we sketch a framework, developed in greater detail elsewhere (2), that holds promise for such an undertaking. We introduce an abstract chemistry implemented in a modeling platform that permits the study of the origins of self-maintaining organizations in a minimally constrained fashion. In several specific instances, this system spontaneously and robustly generates a number of features that occurred in the history of life. The minimality of our model, then, suggests that these features arise generically and, hence, might be expected to reappear if “the tape were played twice.”  

Theoretical Framework and Modeling Platform  

We seek to develop a model of biological organization that is grounded in a particular abstraction of chemistry. Chemistry is characterized by a combinatorial variety of stable objects—molecules—capable, upon combination, of interacting with each other to generate new stable objects. When two molecules interact, the product is determined by their structure—i.e., the components of which they are built and the manner in which these components are arranged. Thus, a molecule is an object with both a syntactic structure and an associated function. Syntactically, it is built up from component objects according to well-defined rules. Its function, coded by its structure, is revealed by the chemical reactions in which it partakes. Chemical reactions generate a stable product through a series of structural rearrangements driven by thermodynamics. We abstract from chemistry both (i) the interaction between molecules to generate new molecules and (ii) the driving of a reaction to a stable form by structural rearrangement.  

The mathematical machinery that provides us with an implementation of such a situation is known as the \(\lambda\)-calculus (3). In \(\lambda\)-calculus, syntactic structures—that is, objects—are defined inductively in terms of nonlinear combinations of other objects, starting from primitives. This definition implies that each object is a function. The function represented by object \(A\) is the mapping that assigns to any object \(B\) a new object expressed syntactically as \((AB)\), referred to as the action of \(A\) on \(B\). To execute this action, \(\lambda\)-calculus defines axiom schemes for rearranging the structure of objects. Let \((AB)\), say, be restructured by applying the schemes of rearrangement one at a time until no further modification is possible. Such a process generates a series of intermediate objects, \((AB) \rightarrow C_1 \rightarrow C_2 \rightarrow \ldots \rightarrow C\), and is termed reduction. The unique final product thereby reached is called a normal form. The schemes of rearrangement are such that functional equality ensues—i.e., we can replace \((AB)\) by \(C\) since \((AB) = C\). Thus, in \(\lambda\)-calculus, (i) objects combine with other objects to produce new objects, which (ii) are transformed to achieve a stable form.  

\(\lambda\)-calculus, while capturing certain key abstractions from chemistry, is not a theory of actual chemistry or theoretical biophysics. For example, this level of description intentionally lacks any explicit reference to thermodynamic notions. Thermodynamic driving is abstracted solely by requiring that every object in our system be in normal form—i.e., schemes of rearrangement are applied to obtain a stable (normal form) object. From a logical point of view thermodynamics essentially implements a consistency requirement by preventing arbitrary rearrangements in arbitrary reactions from occurring. The reduction process as defined in \(\lambda\)-calculus guarantees such a consistency. Thus, \(\lambda\)-calculus captures what is inherent in such consistency requirements but not necessarily what is inherent in thermodynamics. In addition, the present system does not consider spatial constraints, conservation laws, or unequal reaction rates. Our intention is not to emulate actual chemistry but rather to explore the consequences of those minimal features we abstract from chemistry.
This theoretical framework is instantiated in a model with the following components:

(i) Universe. A universe is specified by the axioms of $\lambda$-calculus, which define the nature of objects and the manner in which objects are transformed syntactically. We use the axioms of $\lambda$-calculus as defined in ref. 4. In our universe, all objects are required to be in normal form, and reduction to normal form is required to be completed within some maximum number of steps; otherwise, the object is not allowed in the universe.

(ii) Collision rule. The basic event in our model universe is the interaction among two objects, $A$ and $B$, upon collision. In the simplest case, the interaction between $A$ and $B$ invokes application in $\lambda$-calculus: $(A)B$. The new object created by the interaction is the normal form of $(A)B$. The collision rule may itself be expressed as an object in $\Lambda$, which provides a powerful generalization.

(iii) Interaction scheme. Let $[A,B]$ denote a collision event between the ordered pair $A$ and $B$. The interaction scheme used here is $(a)$ $[A,B] \rightarrow A + B + \text{normal form of } (A)B$; $(b)$ $A$ must be of the form $\xi f \cdot Q$, with $Q$ arbitrary; and $(c)$ computation of the normal form of $(A)B$ must be completed within 10,000 steps and must not exceed 4000 characters. If any of the requirements are violated, no reaction occurs: $[A,B] \rightarrow A + B$. These limits, imposed for practical computation reasons, were not usually exceeded in our computer experiments.

(iv) System. The system is a well-stirred flow reactor that is initialized with 1000 randomly generated (and reduced) objects unless otherwise noted. A pair of objects, $A$ and $B$, is chosen at random for collision, $[A,B]$, according to the above interaction scheme. The object chosen first is hereafter referred to as the operator. Note that $(A)B \neq (B)A$. On average, however, half of the collisions between $A$ and $B$ will be of the form $[A,B]$; the other half will be of the form $[B,A]$. The newly created collision product is checked against predefined syntactical and/or functional constraints (i.e., boundary conditions). If the object passes the filters, it is added to the system. We keep a constant number of objects at any one time. To do so, one object chosen randomly from the system is eliminated. This gives each object a finite lifetime. The whole procedure is reiterated.

We do not describe the model in further detail here. Rather, we summarize the results of computer experiments, which we hope will be sufficient to introduce the behavior of the system and serve as an inducement to readers to explore the primary literature describing the approach (see refs. 2, 5, and 6; in particular, ref. 2).

**Computer Experiments**

We describe three series of experiments.

**Level 0 Experiments.** Experimental protocol and summary of results. The system was initialized with a series of 1000 randomly generated functions, each initially present in one copy. Such experiments always become dominated by either single self-copying functions or ensembles of hypercyclically (7) coupled copying functions [i.e., functions $f$ with $(f)g = g$ or $f$; for all $g$ in the system]. Under perturbation—i.e., the introduction of random objects—level 0 ensembles reduce to single self-copying functions [i.e., a function $f$ with $(f)f = f$].

**Relationship to the work of others.** These results represent an independent rediscovery of the work of Eigen and Schuster (7) on hypercycles. Specifically, our level 0 experiments generate autocatalytic ensembles of copy reactions of which the hypercycle is one example. Like the original hypercycle, our ensembles are not stable entities upon perturbation (but see ref. 8 for spatial systems).

**Level 1 Experiments.** Experimental protocol and summary of results. Level 1 experiments are identical to level 0 experiments except that copying functions (i.e., level 0 entities) are barred from action. Such experiments generate organizations of considerable complexity. Each such organization is a set of objects, distinct from the initial ones, that maintains itself without any single member engaging in a copying action.

**Example of a level 1 organization.** A particularly simple level 1 organization is illustrated in Fig. 1. All objects maintained in the system are characterized by a particular syntactical architecture. All objects are made of two building blocks ($A$, $B$) such that $i$ contiguous $A$s are followed by $j$ contiguous $B$s with $i \leq j$. The organization can, therefore, be visualized in the $i$, $j$ plane (Fig. 1). All actions that occur within this organization can be described by two invariant laws—i.e., emergent regularities in the behavior of the system. The first law states that an object acting on another object will produce the object immediately below the operator along its diagonal, as illustrated by the boldface arrow in Fig. 1b. This law applies to all objects except those at the end of the diagonals. The second law governing the system states that these objects acting upon any others produce specific objects somewhere up the argument's diagonal, as illustrated by the dashed line in Fig. 1b. Together these two laws ensure that the system is syntactically closed and self-maintaining.

**Features of level 1 organizations.** A zoo of different organizations can be generated by specifying syntactical

---

**FIG. 1.** An example of a level 1 organization. (a) All objects in the organization share a syntactical regularity. Specifically, each object is a string of two elements, $A$ and $B$, so that every object contains a sequence of $A$ elements followed by a sequence of $B$ elements, with the number, $i$, of $A$ elements equaling or exceeding the number, $j$, of $B$ elements. Within the formalism of $\lambda$-calculus, $A$ is encoded as $\lambda x_i$ and a sequence of $j$ $B$s is $\xi_j$. (b) The functional relationships ensuring self-maintenance can be succinctly stated as a consequence of two laws. Specifically, in the $i$, $j$ plane, the boldface arrow illustrates the action of law 1 $((Q)\xi)O_{k,l} = O_{k,l+1}$ for all $j > 1$, $i$, $k$, $l$ and the broken arrow illustrates that of law 2 $((Q)\xi)O_{k,l} = O_{k+1,l+1}$ for all $i$, $k$, $l$. Note that the emergent laws do not make reference to the underlying $\lambda$-calculus but nonetheless are sufficient to describe the product of any object in the system interacting with any other object. For the sake of a less-congested figure, actions are drawn schematically and every other diagonal is missing.
filters that prohibit particular organizations from emerging. Detailed descriptions of level 1 organizations will be published elsewhere (2). All level 1 organizations, however, share certain common features:

(i) Self-maintenance versus reproduction: In level 0 experiments, objects were copied by other objects or were self-copying. In such a case, it is appropriate to speak of reproduction of objects. In level 1 organizations, however, neither single objects nor the organization itself is copied. Level 1 organizations are self-maintaining but not reproducing.

(ii) Emergence of laws: Each organization may be completely specified by a listing of all objects in the system and a corresponding listing of all actions of each object on all other objects. This description, however, is hardly concise—it is rather like describing a rat liver by handing a student a biochemistry textbook. We find that all organizations generated by this system display two classes of emergent patterns. One class refers to patterns at the syntactical level of the objects and the other refers to patterns at the functional level. In the example above, what is emergent at the syntactical level is the restriction of all objects to a particular form of combination of the building blocks A and B. What is emergent at the functional level is that two laws govern all transformations among objects of this type.

Laws are emergent in the sense that global regularities result from the collective behavior of locally interacting objects without those regularities being imposed on the objects initially. The laws represent a distinct level of description in that they do not refer to the detailed micromechanics of the system (i.e., the underlying λ-calculus operations).

Level 1 organizations may be considerably more complex than the example outlined above (2). From a formal point of view, laws describing the syntactical constitution of objects are known as a grammar and the (finite) set of objects confirming with it is a formal language. The laws characterizing the relations among objects specify an algebraic structure that is supported by the language. Both grammar and algebraic structure are invariant with respect to the ongoing interactions among objects. Grammar and algebraic structure constitute a compressed description that completely characterizes the organized state of the system. This leads to a minimal notion of "organization" as a dynamically maintained algebraic structure.

(iii) Synthetic pathways and self-repair: Level 1 organizations are remarkably robust toward deletion of particular objects. Elimination of most objects results in their recreation by the remaining interactions. For example, if one eliminates all end objects of the diagonals in Fig. 1, these objects reappear in the system shortly after their elimination (by virtue of the first law). The ability of these organizations to repair themselves is a direct consequence of self-maintenance—i.e., that all objects survive within an organization only by virtue of being the product of some production pathway.

(iv) Seeding sets: An organization contains a number of different smallest sets of objects that are sufficient to recreate that organization if an experiment were begun only with members of such a set. We refer to such sets as seeding sets. All organizations we have observed so far have a unique self-maintaining seeding set.

Effects of perturbations. The effect of perturbing level 1 organizations was explored by periodically injecting random objects into an existing organization. Four different perturbation schedules were explored: (i) introduction of one random object in 10 copies every 30,000 collisions; (ii) as in i, but with perturbation occurring every 50,000 collisions; (iii) as in i, but with perturbations injecting 50 copies; and (iv) as in i, but with three random objects injected simultaneously. With one exception, the results were identical. The organization persists without any change in the emergent laws that characterize it. Level 1 organizations are very robust to small perturbations.

In the exceptional case, a perturbation resulted in the appearance of a new emergent law. This new law did not displace any of the existing laws but rather represented an addition to the existing laws. The new system proved robust to continued perturbations. The exceptional case is relevant in that it illustrates that level 1 organizations can be altered by perturbations.

We have also explored the consequences of relaxing the restriction on copying. Recall that level 1 organizations were obtained by preventing identity functions from acting. If identity functions are either present upon initialization of the system or allowed to arise early during the course of an experiment, the system typically does not reach level 1. Rather, the system becomes dominated by copying level 0 objects. We have found a variety of conditions under which this restriction may be relaxed. If (i) a restriction is placed on the efficacy of copying or if (ii) objects that copy also support constructive interactions, level 1 organizations are generated in the presence of level 0. In addition, we have found that if (iii) identity functions are allowed to act or are introduced into the system after a level 1 organization has been constructed, the organization remains stable. Specifically, the same laws that characterized the organization prior to perturbation characterize the system after the perturbation. Removing the no-copy constraint results in their kinetically stable integration into the existing organization.

Relationship to the work of others. Our level 1 organizations recall three different lines of research. Our level 1 organizations share with the hypercycle model of Eigen and Schuster (7) a limitation on the advantage to self-copiers but differ fundamentally in that they are founded on constructive interactions and the ensuing network of transformations. A copy action is precisely the negation of a transformation. From a functional point of view, a copy action does not force a new object (i.e., function) into the system. This simple difference accounts for the remarkable stability toward functional perturbations as well as for the specificity by which new functional objects can be stably integrated into an established level 1 organization.

The second and third research traditions are work on autocatalytic sets (9–13) and on autopoietic systems (14, 15), respectively. Our results share with these models the phenomenon of self-maintenance but differ markedly from these efforts in defining a formal framework that allows systematic exploration of the conditions permitting its emergence and characterization.

Level 2 Experiments. Experimental protocol and summary of results. Level 2 experiments are initiated with the products of two different level 1 experiments. The procedure is otherwise identical to the level 1 protocol, except that the system is increased to a constant size of 3000 objects. Such experiments have one of two outcomes: either a single level 1 organization comes to dominate the system or a new self-maintaining metaorganization arises (hereafter referred to as level 2). Such metaorganizations have the two self-maintaining level 1 organizations as components. In addition, the metaorganizations contain a set that is not self-maintaining but that acts to knit the self-maintaining level 1 sets into a higher-order self-maintaining entity. This set contains objects that result from the communication (cross-interaction) between the level 1 organizations and that do not belong to either organization. The grammars and the algebraic laws characterizing level 2 organizations will be described elsewhere (2).

Features of level 2 organizations. As in the case of level 1 organizations, a zoo of different level 2 organizations can be generated by varying boundary conditions such as syntactical
filters or the constituent level 1 organizations. In addition to the features of level 1 organizations enumerated above, level 2 organizations share certain common features unique to this level of organization.

(i) Emergence of level 2 laws. The laws characterizing the level 1 organizations, which are contained within a given level 2 organization, remain unchanged by their inclusion in the higher-order entity. In addition, a new set of laws—i.e., level 2 laws—is found that defines the structure and the actions of the metabolism that glues the two level 1 organizations into a self-maintaining higher-order entity.

(ii) Seeding set. Under the boundary conditions imposed, the seeding set of a level 2 organization is nothing more than the seeding sets of the level 1 organizations from which it is constructed. However, the organizational description, as manifest in the laws, is not a superposition of the descriptions of the level 1 organizations.

(iii) Metabolic flows. In a level 2 organization, the set that links the two self-maintaining level 1 organizations is composed of products derived from syntactical elements of both level 1 organizations. These hybrid products define production pathways from each of the level 1 organizations into the shared metabolism and from there back into the two level 1 organizations. The magnitude of these metabolic flows changes the diversity of products present in each level 1 component as compared to its stationary diversity in isolation. Metabolic flows, and the differences in metabolic diversity they generate, are maintained despite the fact that both the seeding sets and the laws of each component level 1 organization remain unchanged.

Effects of perturbations. Three different perturbation schedules were explored: (i) introduction in 10 copies of three random objects every 30,000 collisions into the set that links the two self-maintaining level 1 organizations, (ii) simultaneous introduction of three random objects in 10 copies every 30,000 collisions into each of the level 1 organizations and into the set that links them, and (iii) sequential introduction of three random objects in 10 copies into each set, so that each set is perturbed every 60,000 collisions. The first perturbation schedule resulted in no changes in the seeding sets of either component level 1 organizations and no change in the law of the level 2 organization. The second and third perturbation schedules resulted in a simplification of the seeding sets of one of the component level 1 organizations. In one case (perturbation schedule ii), this simplification had no effects on the level 2 laws (since the lost part was redundant); in another case (perturbation schedule iii), the simplification had the effect of stabilizing the level 2 laws. In all cases in which simplification occurred, the laws were not fundamentally restructured upon perturbation. These results indicate that level 2 organizations are resistant to small perturbations.

What Would Be Conserved If “The Tape Were Played Twice”?

The results of these experiments can be briefly summarized to reflect three general findings: (i) Hypercycles of self-reproducing objects arise. (ii) When replication is inhibited, self-maintaining organizations of considerable complexity emerge. (iii) Organizations can be hierarchically combined to produce new self-maintaining organizations that contain the lower-level organizations as self-maintaining components.

To assess the consequences of these results for the issue of contingency and necessity, it is important to keep in mind the level of description that our model induces. The model pictures a particular abstraction of chemistry in terms of a calculus and endows it with a simple dynamics. The system spontaneously constructs networks of functional relationships that constitute a formalization of an interesting notion of “phenotype” —i.e., organizational structure. The model allows a natural definition of organizational grades and the exploration of the conditions under which they arise.

Our results invite analogy to the organizational grades as they arose in the early history of life: self-replication, self-maintaining prokaryotic organizations, and self-maintaining eukaryotic organizations. In particular, level 0 hypercycles are equivalent to the RNA hypercycles proposed by Eigen and his colleagues (7, 9). Level 1 organizations are systems of transformations in our abstract chemistry, analogous to the metabolism of prokaryotes (see below). Finally, level 2 organizations are metaorganizations constructed from interactions between level 1 organizations, as they occurred in generation of the eukaryotic cell from prokaryotic precursors.

The correspondence between the results of our computer experiments and real life is intriguing. Yet it is a strong claim indeed to contend that these results are sufficient to inform us literally about what would emerge if “the tape were played twice.” The claim that our results represent the generic behavior of self-maintaining organizations is based on two propositions:

(i) A-calculus is an appropriate set of constructs to encompass the origins of metabolism. A fundamental property of all extant forms of life is the presence of a metabolic organization. A metabolism is to A-calculus as that portion of metabolic machinery, which catalyzed reactions, characterized by the transforming action of catalysts on substrates. We have shown here that two abstractions from chemistry togetherness with a simple dynamics are sufficient to generate self-maintaining organizations. Thus, if it is true that (a) metabolism is an invariant property of living systems and (b) the minimal notion of metabolic organization is captured by the study of the origins of self-maintaining transformation systems, then A-calculus is a natural approach for analyzing the origin of biological organizations. The question of the validity of such a claim may then be rigorously challenged by increasingly sophisticated elaborations of our model universe.

(ii) The assumptions made are biologically meaningful. Our level 1 and level 2 results are dependent on two assumptions with regard to our interaction scheme. We show that these assumptions are equivalent to neglecting “food” and “waste” in a sense that is appropriate at our functional level of description.

(a) Food: We assume that when two functions interact to produce a third, the functions that participate in the interaction are not consumed in the process. If this assumption is relaxed, self-maintaining organizations fail to emerge. This assumption represents a controlled input into the system from the environment—i.e., food. Note, however, that this assumption is not a buffering since it does not guarantee the persistence of an object in an organization. Any object has a finite lifetime induced by the dilution flow. This provides for a sorting mechanism that biases toward those objects that have production pathways involving other objects in the same system. It follows that as soon as a stably self-maintaining seeder set of an organization has formed, we should be able to relax the interaction scheme to a catalytic transformation in which one interaction partner is used up. We have found this to be the case without influencing either the syntactical or algebraic regularities that characterize an organization.

(b) Waste: Interactions that yield products that in combination with existing objects generate a normal form without reduction are prohibited in our system (see model platform, item iiib). If this assumption is relaxed, self-maintaining organizations fail to emerge. Such products are incapable of participating in closed transformation networks—i.e., the product will fail to metabolize. The biological interpretation
of this assumption is straightforward. Metabolic processes are disrupted by the accumulation of waste, and this assumption guarantees that a particular form of waste does not accumulate (see ref. 2 for details).

Selection, Self-Maintenance, and the Emergence of Biological Order

Darwinian selection, as opposed to the mere differential sorting of an arbitrary collection of objects, presupposes the existence of self-reproducing entities (16). In this study, the only reproducing entities are the self-copying functions arising in level 0; level 1 and level 2 organizations are self-maintaining but not self-reproducing. Indeed, we find that organizations (levels 1 and 2) can arise in the absence of self-reproducing entities (level 0).

Self-reproduction and self-maintenance are shared features of all extant organisms, barring viruses. It is not surprising, then, that there has been little attention paid to the generation of one feature independently of the other. While selection was surely ongoing when transitions in organizational grade occurred in the history of life, our model universe provides us the unique opportunity to ask whether selection played a necessary role. Our findings clearly indicate that it need not.

Moreover, separating the problem of the emergence of self-maintenance from the problem of self-reproduction leads to the realization that there exist routes to the generation of biological order other than that of natural selection (17). Indeed, this is apparent upon inspection of the formal structure of the theory. Neo-Darwinism is about the dynamics of alleles within populations, as determined by mutation, selection, and drift. A theory based on the dynamics of alleles, individuals, and populations must necessarily assume the prior existence of these entities. Selection cannot set in until there are entities to select. Our exploration of an abstract chemistry not only provides a route to generate such entities but illustrates that different organization grades can arise in the absence of selection. This raises the problem of determining which features of biological organization are attributable to the emergence of the organization and which features are attributable to its subsequent modification by selection. At issue is the primacy of natural selection in shaping the major features of biological organization.

This paper is communication number 2 from the Center for Computational Ecology of the Yale Institute for Biosopheries Studies. We thank John McCaskill, Günter Wagner, and John Reinitz for discussions that proved central to the development of these ideas and M. Dick, M. Eigen, D. Lane, H. Morowitz, S. Rasmussen, J. Reinitz, P. Schuster, B. Singer, P. Stadler, A. Wagner, and, particularly, J. Griesemer for comments on the manuscript.