CHAPTER 4

THE NECESSITY OF BIOSEMIOTICS: MATTER-SYMBOL COMPLEMENTARITY

H. H. PATTEE*

Department of Systems Science and Industrial Engineering, Thomas J. Watson School of Engineering and Applied Science, State University of New York at Binghamton, Binghamton, NY 13902–6000

Abstract: Biosemiotics distinguishes life from inanimate matter by its dependence on material construction controlled by coded symbolic information. This irreducible primitive distinction between matter and symbol is necessary for open-ended evolvability and the origin of life as we know it. This type of subject/object distinction is reestablished at many levels throughout all of evolution. In physics this becomes the distinction between material laws and symbolic measurements and models; in philosophy this is the distinction between brain and mind. These are all emergent epistemic distinctions, not ontological dualisms. The origin of life requires understanding the origin of this symbolic control and how inanimate molecules become functional messages. I discuss the necessary physical conditions that would allow such evolvable symbolic control of matter to arise

Keywords: self-replication, von Neumann, symbolic control, semantic information, measurement, constraints, epistemic cut, protein folding

LIFE DEPENDS ON SEMIOTIC CONTROLS

We easily agree with Einstein that a Beethoven symphony cannot be appreciated as only “a graph of air pressures,” although in principle it has such a physical description. In the same way we understand Bohr that, “You don’t explain a tea party by quantum mechanics.” On the other hand, it is not so easy to understand why you cannot adequately explain genetics with biochemistry or enzyme catalysis with quantum mechanics. Because we believe no events at tea parties, in genes, or in enzymes violate any physical laws we might assume that their descriptions differ

---

[*] Present address: 1611 Cold Spring Road, Williamstown, MA 01267. hpattee@adelphia.net

M. Barbieri (ed.), Introduction to Biosemiotics, 115–132.

only in their degrees of complexity. What biosemiotics illustrates is that symbolic controls are categorically different from laws and that they are irreducible to physical laws even though their material vehicles obey the laws and have a correct physical description.

What we need to understand is that physical laws are universal and must apply to all conceivable systems. Therefore laws are empirically moot with respect to any particular system until its particular initial conditions are specified. This requires information, and physical laws cannot specify this information. In physics jargon symbol systems are special types of initial condition called boundary conditions or constraints (Polanyi, 1968; Pattee, 1972). Consequently an adequate explanation of any living organism requires more than a detailed lawful physical description or merely the confirmation that the laws of nature are always inerrantly followed. One must explain how informational constraint structures locally control the universal physical laws so as to propagate and evolve.

All living organisms exist by virtue of hierarchies of control by informational constraints. This is the case at all levels, from the genes, to development, to sensorimotor controls, to abstract thinking, and to our technical artifacts. Symbol systems are rate-independent informational constraints that control rate-dependent dynamics by means of coding systems.

To understand what this implies one must first recognize that physical laws are universal and objective. This means that the fundamental principled requirement for a law of nature is that it is as independent as possible of all conceivable individual organisms and observers. Consequently, physical laws are based on invariance and symmetry principles that guarantee the irrelevance and impotence of any observer, organism, or mechanism to affect the laws. In other words, *physics focuses on all those universal regularities of nature over which organisms and observers have no control*. Physical laws are universal and inexorable. By contrast, the study of *biology focuses on those specific events over which the organisms and observers have local control*. Beginning with the organism's specific catalytic rate control by enzymes, evolution progresses by elaborating and testing many types of controls at many hierarchical levels. Over the course of evolution organisms have gradually increased their ability to control their internal and external environments on which their survival depends.

Survival is the ultimate function of controls, but unfortunately controls do not assure survival. In the case of humans the brain through the freedom of language and the prostheses of technology has developed controls with such Promethean powers that the art of artificial control may turn out to have exceeded what is adaptive as a survival strategy for our species.

THE RELATION BETWEEN PHYSICAL LAWS AND CONTROL CONSTRAINTS

There is a clearly problem of language here that creates confusion. What does it mean to say that universal inexorable physical laws over which organisms can have no control are in fact controlled by individual organisms? The answer requires understanding a distinction fundamental to all physical theory, the primitive separation of the laws themselves that are universal and inexorable, and initial conditions and constraints that are local and controllable, and that must be measured to have any effect. Eugene Wigner calls this principled distinction between laws and initial conditions "Newton's greatest discovery." That is because it is an epistemic necessity that is essential for all conceivable physical laws, not just Newton's laws.

Briefly, the idea is this. The universe and all systems within it are assumed to run according to universal laws whether or not observers or life exist. The mathematical descriptions of these laws are interpreted by ontological concepts of space, time, matter and energy but the laws themselves do not include the epistemological concepts of measurement and control events. However, measurement is essential if we want to predict any consequence of laws on a specific observable system. There must be measurement of initial conditions and the measurement process requires local control constraints of a measuring device or instrument.

Physical laws and initial conditions are therefore necessary irreducibly complementary categories. That is, neither can be reduced to, or derived from, the other. Measurement instruments and control constraints are special, usually complex, structures with initial conditions that are largely arbitrary. In practice measurements and controls are distinguished from the action of physical laws by how time and energy enter their descriptions. Fundamental physical laws are time and energy dependent in a mathematically rigorous sense. That is, the equations describing these laws require the concept of time-derivatives or rates of change of the states and energies of the system[1]. Also, the fundamental microscopic laws are time reversible. This physical time, sometimes called "real time," and the rates described by time derivatives are intrinsic to natural laws and are not controllable, although they may be different when measured by different observers in relative motion.

The concept of *control* of rates does not apply to universal laws but only to local structural constraints. The classical example of both rate control and time measurement is a clock. By contrast to the real-time of laws, clock-time depends on some form of local structure or constraint. We speak of clocks *measuring* time intervals but, unlike laws, clocks do not have an intrinsic rate independent of how we measure it. Also, unlike microscopic laws, measurement and control are irreversible concepts. Clocks function only by measuring local periodic structures such as a pendulum with an escapement or counter. Of course the pendulum swings according to laws, but its period depends on its length, and that is entirely arbitrary boundary condition. Escapements, whether mechanical, electronic, or chemical, can be said to control the rate at which energy "escapes" or is dissipated from the driving source, and these constraints are also arbitrary[2]. Some form of measurement is a necessary component of any functional control process.

The point I want to emphasize here is that we say a clock is a control constraint only by virtue of its locally "escaping" the inexorable time, rate, and energy dynamics of physical laws. In other words, the laws exist in time but cannot make measurements of time. Within wide limits imposed by natural laws, a clock keeps its own arbitrary time and runs at its own arbitrary rate. *This concept of local* *"escape" is important because life depends on it.* Enzymes control the rates of specific chemical dynamics in all of life allowing local organisms to locally escape the universal rates we associate with unconstrained physical laws. The existence of an isolated catalyst that controls rates of reaction is not considered as functional. Function, as I use the term, applies only coordinated controls initiated by organisms or executed by their artifacts.

RATE-INDEPENDENT CONSTRAINTS; SYMBOL SYSTEMS

Biosemiotics recognizes many levels of control. Rate control, as in enzyme catalysis, is only the first level of control constraints. Symbolic constraints are a type of record that requires higher levels of organization. Ernst Mayr has often pointed out that biology is not explained by natural laws because life requires the concept of an adaptive evolutionary *history*, an actual record in the organism that while obeying natural laws cannot be derived from these laws. Records are a special form of constraint that can "instruct" controls. Polanyi (1968) classifies these structures in physical terms as special types of boundary conditions that he aptly describes as "harnessing" the laws.

The word "history" has two profoundly different usages. The looser meaning is simply an implied ontological sequence of events, as in the history of the universe, or geological history, where there is no explicit record other than the actual events or structures themselves. The more specific meaning of history as used by Mayr requires a separate record of events. This latter specific meaning that is essential for evolution implies an epistemic record that is a representation or description distinct from the events that it records. In all known living systems, the genes are such historical records of innumerable adaptive natural selection processes. The relative simplicity of the record itself (the DNA) is deceptive. What is important is that for a record to have any function or meaning requires complex coding, reading and interpreting mechanisms. Along with measurement and control the concepts of biological information and instruction are not a part of physical theory in so far as they are functional concepts. That is, we cannot identify a molecule as informational unless we can identify how it is interpreted by the organism and how it functions in the organism[3]. The question remains, how does symbolic information actually get control of physical systems when it appears to be a separate category?

PHYSICAL LAWS CANNOT ADDRESS THIS QUESTION

This matter-symbol separation has been called the epistemic cut (e.g., Pauli, 1994). This is simply another statement of Newton's categorical separation of laws and initial conditions. Why is this fundamental in physics? As I stated earlier, the laws are universal and do not depend on the state of the observer (symmetry principles) while the initial conditions apply to the state of a particular system and the state of the observer that measures them. What does calling the matter-symbol problem "epistemological" do for us? Epistemology by its very meaning presupposes a separation of the world into the knower and the known or the controller and the controlled. That is, if we can speak of knowledge *about* something, then the knowledge representation, the knowledge vehicle, cannot be in the same category of what it is about.

The dynamics of physical laws do not allow alternatives paths between states and therefore the concept of information, which is defined by the number of alternative states, does not apply to the laws themselves. A measurement, in contrast, is an act of acquiring information about the state of a specific system. Two other explicit distinctions are that the microscopic laws are universal and reversible (time-symmetric) while measurement is local and irreversible. There is still no question that the measuring device must obey the laws. Nevertheless, *the results* of measurement, the timeless semantic information, cannot be usefully described by these time-dependent reversible laws (e.g., von Neumann, 1955).

THE EPISTEMIC PROCESS IN BIOLOGY

The problem is that physics greatly simplifies the matter/symbol relation by more or less arbitrarily making an epistemic cut. This avoids as far as possible the influence of measurement on the state of the measured system. Whenever an attempt is made to include the measuring device in the system it becomes the notorious "measurement problem" on which there is still no consensus.

The origin of life must address the question: How did this separation, this epistemic cut, originate? As Hoffmeyer (2000) has pointed out, the assumed sharp conceptual epistemic cut between these highly evolved categories of knower and known makes it difficult to imagine how life began and how these two categories separated at primitive levels. The epistemic cut is often treated as a conceptual discontinuity. Indeed it is difficult to imagine a "gradual cut." How does a reversible dynamics gradually become an irreversible thermodynamics? How does the concept of counting discrete units transform into the concept of a continuum (Zeno's paradox of motion)? How does a paradigm shift from classical determinism to quantum indeterminism occur gradually?

The problem arises acutely with the genetic code. A partial code does not work, and a simple code that continuously works as it evolves is hard to imagine. In fact, this is a universal problem in evolution and even in creative thought. How does a complex functioning set of constraints originate when no subset of the constraints appears to maintain the function? At least in the case of thought we can trace some of the history, but in the origin of life we have no adequate history. Even in the case of creative thought, so much goes on in the subconscious mind that the historical trace has large gaps.

The problem is that conceptually the epistemic cut divides the world in two, and the central problem is how the two worlds are connected. As C. S. Peirce has emphasized, all symbol systems are necessarily triadic systems, and the epistemic cut itself is actually a complex process. It corresponds to the *interpretation* that relates the symbol to its referent. In the cell this is an enormously complex process of transcription, translation, synthesis, folding, distribution, and selective control of many proteins. How this coordinated interpreting system originated is the central problem of the origin of life.

SYMBOLIC CONTROL IS NECESSARY FOR EVOLVABILITY

The categorical distinctions between matter described by physics and matter functioning as symbol are different at each level of biological organization. The distinction needs to be made clearly at each hierarchical level or conceptual and terminological confusion will result. It is easy to distinguish symbols at highly evolved levels like symphonies and tea parties. The words on this page are clearly symbols. Their material embodiment is arbitrary. The font is not relevant, nor do we pay attention to their material embodiment, whether they are displayed on a liquid crystal screen, a cathode ray tube, or ink on paper. Even the language we are using is arbitrary.

It is not so easy to see that the DNA of genes is also an arbitrary embodiment of a record because it happens to be the only one we know from life on earth. However, within the fields of exobiology and artificial life studies the arbitrariness of DNA is generally assumed. Many other copolymer strings or even bit strings in a computer could be interpreted or translated by a suitable coding mechanism to synthesize the same proteins as a DNA sequence.

Why is this arbitrariness of symbols essential for open-ended evolution? The most obvious property of highly evolved symbol systems such as natural language and mathematics is their enormous open-ended variety that is not limited in any significant way by physical laws. This independence is also illustrated by the fact that, unlike physical laws, the function and meaning of symbols is not dependent on the rate at which they are written or read. A mathematical proof does not depend on how long it took to produce or to read. The same it true of a work of literature. In other words, the basic observables of physical laws, space, time, matter, energy, and rates of change, have no significance for the semantic information of symbol systems. The symbolic expressions of physical laws are “about the laws” but the mathematical symbols that describe the laws do not appear to be restricted by the laws. It is just this arbitrariness that allows organisms freedom to harness laws. The necessity of symbols for open-ended evolution was first discussed by von Neumann (1966) in his lecture on the logic of self-replication.

VON NEUMANN’S DESCRIPTION AND CONSTRUCTION

Von Neumann was the first to argue that the two categories, *symbolic description* and *material construction*, are essential for self-replication that is capable of open-ended evolution. His argument was entirely abstract and by no means logically complete. It explicitly abstracted away matter, energy and all physical laws. I will first elaborate on von Neumann’s logic and then I will take up the necessary *physical* conditions to realize this logic, or what he thought, "may be the more important half" of the problem.

Von Neumann's logical argument for the necessity of symbols as distinct from dynamics in self-replication was informal and largely intuitive. Nevertheless, if you understand his argument you will find it hard to think how evolvable self-replicating units could work any other way. The motivation for his argument was to understand the "threshold of complication" that would allow systems to evolve increasing complexity rather than wearing out or decaying. His logic is all the more remarkable because it correctly predicted how cells actually replicate before the discovery of the mechanisms of genetic description, coding and protein synthesis. Von Neumann began by observing that the medium of communication that feeds a material automaton is completely different than the automaton itself or its output. This was his recognition that symbols are a different category than matter. He also recognized that this was important for general-purpose computers, what is called the software-hardware distinction.

Von Neumann emphasized the "completely decisive property of complexity, that there exists a critical size below which the process of synthesis is degenerative, but above which the phenomenon of synthesis, if properly arranged, can become explosive." He was thinking of biological evolution and its open-ended variety. The essential condition here is that the individual self that is being replicated must be only one of an indefinite number of different potential selves all of which can be replicated by the same process. This raised two questions: (1) what defines the set of all possible individual selves that potentially can be replicated? And (2) how do you describe or represent the individual *self* that is being replicated?

Logic will get you only so far with these questions. For example, the concept of replication implies assembling or constructing a new individual that is like another. Von Neumann realized that how this construction can be done will depend on the nature of the available parts and on how the parts are to be assembled. He saw that if the parts were too elementary, like atoms, then both the description and construction would be a long and complicated process, while if the parts were too complex, like robots or rabbits, then there would be no real problem. He called this the "parts problem" and abstracted away the matter and energy of real construction by defining some functional operations on parts, like recognizing, moving, cutting, joining, etc., that are to be symbolically represented. There is a great amount of arbitrariness in these choices of parts and operations, but as we shall see, the basic logical separation of symbolic description and material construction does not depend on these choices.

The more fundamental question is how you make sure the replicated individual is like the original. How do you construct a copy of an organized structure made up of parts from a reservoir of these parts? There are two approaches. One is to identify the original parts themselves by *inspection* and then assemble the corresponding parts to form the copy. The other approach is to use a *description* of the original that when interpreted amounts to instructions enabling the assembly of the parts in the copy. Note that the concepts of inspection and description require an epistemic cut that separates the object being inspected or described and the record of the inspection or description. Both of these methods have advantages and disadvantages that go beyond logic and depend on the physical nature of space, time, and the nature of the parts. Von Neumann using heuristic reasoning found that taking advantage of both approaches gives the most promising results, and in fact we now know that both approaches are used in all living systems in the way that von Neumann proposed.

VON NEUMANN'S LOGIC OF SELF-REPLICATION

Following these intuitions, von Neumann began simply by postulating the existence of both symbolic and material components in the forms of a *description* and a *constructor*. The constructor would both interpret and construct what was described using parts from a reservoir. The constructor was universal with respect to an open-ended set of descriptions one of which he assumed could be the description of the constructor itself. In his notation, A was the material constructor and $\varphi(\mathrm{A})$ was the description of the constructor. If the description $\varphi(\mathrm{A})$ was fed to the constructor A, then A would construct a copy of itself, A′. We can symbolize this as $\varphi(\mathrm{A}) \rightarrow \mathrm{A} = \mathrm{A}'$. This is not self-replication because the description $\varphi(\mathrm{A})$ has not been replicated. One might at first think that to copy the description we would simply feed the constructor a description of the description, $\varphi(\varphi(\mathrm{A}))$, but this leads to an infinite regress since that description must also be copied, and so on.

This leads to the crucial recognition that *a symbolic description, whatever form it may take, has a physical structure that is independent of its interpretation*. In other words, to *read* the description means to *interpret* the description. To *copy* the description means *not to interpret* the description but to copy only its physical structure. Since the description is quiescent, copying can be done by inspection or by some template process. The constructor is defined to only interpret the description, so it is necessary to add another component, B, called the copier and its description $\varphi(B)$. We then can write $\varphi(\mathrm{A}+\mathrm{B}) \rightarrow (\mathrm{A}+\mathrm{B}) = \varphi(\mathrm{A}'+\mathrm{B}') \rightarrow (\mathrm{A}'+\mathrm{B}')$. This is almost self-replication except it is ambiguous. What is missing is how the new descriptions and constructions are related. Von Neumann "solved" this logically by creating a new control component, C, that takes care of housekeeping details such as inserting the new description into the new hardware constructor and separating the offspring from the parent. This component, C, amounts to what is called the operating system of a computer that takes care of the software-hardware relationship.

Von Neumann's logic and computer analogies are by no means a clear solution to the material semantics of cells. In the cell we know that the control required for cell division is a very complex process that is not yet fully understood. But the essential evolutionary consequence of von Neumann's logic is that now any additional description, D, of some new structure or function when added to this basic description will be constructed and incorporated into all future generations:

$$\begin{aligned} &\varphi\,(\mathrm{A}+\mathrm{B}+\mathrm{C}+\mathrm{D}) \rightarrow (\mathrm{A}+\mathrm{B}+\mathrm{C}) \\ &\quad = \varphi(\mathrm{A}'+\mathrm{B}'+\mathrm{C}'+\mathrm{D}') \rightarrow (\mathrm{A}'+\mathrm{B}'+\mathrm{C}'+\mathrm{D}') \end{aligned}$$

This is as far as von Neumann's logic takes us. The main point of his logic is that open-ended evolution requires more than a complex time-dependent dynamics and complex chemical reactions. There must be a time-independent passive memory that by means of a coded description controls the dynamical rates of specific constructions or chemical syntheses. What I will now take up are the physical requirements that would allow such a complicated symbol-matter logical scheme to actually work in a reasonably effective way. I repeat that I am not solving the origin problem. Von Neumann himself had no clue. He thought, "That such complex aggregations should occur in the world at all is a mystery of the first magnitude." In my view, the place to look for clues is in the actual physical requirements of symbol systems where we may imagine simpler systems than we find in today's highly evolved organisms that satisfy these requirements.

VON NEUMANN'S "MORE IMPORTANT" QUESTION

Von Neumann was fully aware that logic alone was not adequate to explain cells. He warned: "By axiomatizing automata in this manner one has thrown half the problem out the window and it may be the more important half. One does not ask the most intriguing, exciting and important questions of why the molecules or aggregates that in nature really occur…are the sorts of thing they are, why they are essentially very large molecules in some cases, but large aggregations in other cases."

Von Neumann's use of *inspection* and *description* are really generalizations of highly evolved cognitive activities that need to be more precisely defined in the context of the simplest replicating unit. Copying by inspection means using physical interaction with the object directly without the use of symbols, codes, translation, or interpretation. Casting from a mold and template matching are such direct processes, as in base pairing in copying nucleic acids and the binding of a substrate by an enzyme. I should emphasize here that the physical interaction of base pairing and substrate binding are not in themselves functional or semiotic processes. It is only by virtue of their roles in the overall process of self-replication that they are interpreted as functional. Such material matchings might be interpreted in Peirce's terms as iconic signs.

A description, on the other hand, requires more complicated physical interactions that couple the description to what it stands for, its referent. This interaction in the context of self-replication can be called a code or an interpretation, and because the code constraints are themselves constructed from a description they are not determined by physical necessity. It is implicit in the concept of a code that it must apply to more than one description. In fact, to allow evolution the code must apply to an open set of potential descriptions. Again I emphasize that only by virtue of its potential function for an individual's survival can this be distinguished as a semiotic process. This chemical arbitrariness in the coding enzymes Jaques Monod (1971) calls the "principle of gratuity." It is also this construction from a description that Barbieri (2004) calls "artifact-making," a distinguishing characteristic of life. It is because of this freedom or lack of physical necessity that genetic symbol systems and the novelties of evolution have no adequate physical explanation even though they can in principle be correctly described by physical laws in every detail.

It is not clear that von Neumann saw this point since he was concerned with the logic, not the physics. However, he did argue that a description had the advantage of being quiescent, relatively time-independent, and free of the dynamics of the system it describes. It could then be copied by direct inspection. On the other hand, copying a dynamic system by direct inspection in real time would run into a problem with the parts continually changing in time. How would the system choose what state should be copied in that case? He also suggested that a complete and detailed inspection, including inspecting the inspection components themselves, would probably lead to logical antinomies of self-reference. He did not elaborate on this, but he may have been thinking of the measurement process in physics where he showed elsewhere that measuring the initial conditions of the measuring device itself leads to an infinite regress. Only by choosing at some point to make the distinction between the system being measured and the measuring device, i.e., an epistemic cut, can this regress be terminated (von Neumann, 1955).

PHYSICAL REQUIREMENTS FOR EFFICIENT MEMORY

The physical conditions necessary for memory storage are relatively simple to state as contrasted to the conditions for writing and reading of memory. The first condition is that there exist many inherently equiprobable constraint structures with adequate stability. Equiprobable means that the structures are energy degenerate or the energy of each state is the same. These states need not be exactly the same energy as long as the energy differences do not significantly affect the setting of the state by writing or the communication of the state by reading. One-dimensional copolymers and linear symbol strings are the simplest common physical structures satisfying these conditions. Such relatively time-independent memory structures function as long-term, high capacity storage.

Memory structures can also exist physically in one, two, three dimensions, or in n-dimensional networks but explicit syntax for access must be supplied. The advantages of the linear sequence memory, like nucleic acids and Turing machine tapes, and language text are (1) open-endedness or extendable capacity, (2) uniformity and simplicity of writing and reading, including ease of random access, (3) universal coding for all sequences, (4) relative isolation from the dynamics that it controls because of coding or the interpretation process. In the context of the origin of life, copolymer chains are the simplest abiogenic structures that have the necessary stability and potential memory capacity. The disadvantages of linearity are (1) lack of parallel processing or associative access, (2) low density of information storage, and (3) the necessity for an explicit code to couple one-dimensional energy degenerate sequences to the energy-dependent three-dimensional dynamics.

One can also define analog memory and codes as in analog computation. Analogs need not involve discrete symbols. This has been suggested by Hoffmeyer and Emmeche (1991), Juarrero (1998), Hoffmeyer (1998) and Barbieri (2003) in contrast to discrete or digital memories and codes. The problem with analogs is that they are all special purpose structures like individual molecular messengers that have limited informational capacity and that have no common code or interpreting process, as do genetic sequences. An autocatalytic or metabolic network may be interpreted as containing an implicit informational dynamics, but lacking an explicit passive memory structure and code it is difficult to imagine any open-ended evolvability. On the other hand, as Hoffmeyer (2000) suggests, some form of implicit analog codes may have existed as precursors of the explicit discrete codes of present life.

PHYSICAL REQUIREMENTS FOR CODING AND CONSTRUCTION

In even the simplest existing cells the steps from the symbolic base sequence in DNA to a functioning enzyme are too complex to have originated without simpler intermediate stages. However, to control construction or synthesis, even the simplest one-dimensional discrete-state memory storage that exists by virtue degenerate energy states, must somehow control the rates of specific dynamical interactions. This means that the linear degeneracy must be broken. This must be done by new interactions between the linear storage elements. In present cells this is a complex process that requires several steps. First, the DNA sequence is transcribed to messenger RNA by template copying. Next the coding enzymes and transfer RNAs translate the base triplet code to the corresponding amino acids that are then joined in sequence by the messenger RNA and ribosome machinery. Finally, the one-dimensional sequence folds into a functioning enzyme. In this process there are cases of descriptions and constructions by both template inspection and coded descriptive translations.

The discovery of enzymatic RNA made it possible to imagine a much simpler translation process in which RNA can function both as a constructing enzyme and as a symbolic description of an enzyme. By description I mean a passive structure that can be copied by template inspection, and by construction I mean a dynamic catalytic process that joins molecules by strong, covalent bonds. The main point is that this double function is only possible by virtue of the two configurations of RNA, the passive one-dimensional sequence memory and the folded three-dimensional active ribozyme.

THE PHYSICAL REQUIREMENTS FOR FOLDING AND FUNCTION

Folding transformations are the most fundamental semiotic processes in all living systems. Folding is fundamental because it is the process that transforms the passive symbolic gene sequences into the dynamic rate-control of enzymes. Folding transforms what are essentially rate-independent syntactically coded sequences into rate-dependent functional controls. Protein folding is a highly parallel process with so many degrees of freedom that is difficult to model even on supercomputers. Physically to describe folding in any structure requires two types of bonds, strong bonds that preserve the passive topological structure of what is folded, and weaker bonds that acting together hold the active folded structure in place.

This physical requirement follows from the logical definition of "folding." For example, to fold a sheet of paper means forming a three-dimensional shape without changing the two-dimensional topology of the sheet by tearing or gluing. As long as the strong-bond topological sequence structure is energy degenerate it can serves as an informational constraint or a passive memory. Folding removes this degeneracy by allowing new weak bond interactions between the elements resulting in an active enzyme. A *physical description* of protein folding is an energy minimization process or a relaxation of many weak bond interactions under the constraints of the strong bonds holding the linear sequence together (e.g., Frauenfelder and Wolynes, 1994).

How should we describe the semiotics of this process? I want to distinguish the physics and the semiotics. First, I defined a condition for symbolic information storage as a physically indeterminate (energy degenerate) structure. I assumed that all symbol vehicles obey physical laws and have, in principle, a physical description, but as I explained, that does not imply that symbol structures are physically determined. Quite the contrary is the case. Such a degenerate sequence structure can have an immense number of physically indeterminate sequences. Therefore the interpretation or function of any such semiotic or informational sequence is literally metaphysical (beyond physics).

The actual folding process, on the other hand, is an entirely physical process of minimizing the energy under the semiotic constraints of the sequence. In other words, the strong-bonded sequence can be called informational because it is one of many physically equivalent alternative sequences, while the folding dynamics itself is not informational because no new information is added in the process of minimizing the energy. (There are special cases where folding information may be added from scaffolding molecules.)

THE SEMIOTIC CLOSURE REQUIREMENT FOR "SELF"

How do we define the individual system that is interpreting the information? We need an objective criterion for what "self" is doing the interpreting and replicating, because there are innumerable energy degenerate structures that are not descriptions and many catalytic events that are not functional. What additional conditions are required to satisfy a *physical* implementation of the *logical* "self" that reads and interprets descriptions and constructs and assembles parts in von Neumann's formal self-replication.

The essential logical requirement for self-replication that von Neumann described is that all the components that implement description, translation, and construction are themselves described, translated and constructed within the "self" that is being replicated. This amounts to a *logical closure* that defines a "self." Physically this requires elaboration. There is more to the strong and weak bond requirement than the ability of the weak bonds to cause the strong bonds to fold into a functioning enzyme. The strong bonds also stabilize the passive memory and the integrity of the primary structure of enzymes. The weak bonds bind the enzyme to its substrate and control the rate of catalyzed strong bond formation. In effect, the strong bonds form the skeleton for both descriptive and constructive molecules while the coordinated organization of weak bonds define the shapes necessary to control the strong bonds, both the strong bond folding and individual strong bond formation or breaking.

These are the *physical* conditions required to implement von Neumann's logical closure. I have called this *semantic closure*, but Luis Rocha (2001) has more accurately called it *semiotic closure* because its realization also includes the syntax and pragmatic physical control processes. This complex interrelationship of strong and weak bonds is the minimum physical requirement that allows the realization of von Neumann's quiescent symbolic description and dynamic material construction. Of course the actual physical forces come in more than two strengths and evolution has refined structures at many hierarchical levels using different types of forces. Many types of strong and weak bonds enter into the complex process of folding (e.g., Wolynes, et al., 1995).

EVOLUTION REQUIRES POPULATION DISTRIBUTIONS

Based on the concept of semiotic closure, I would define an interpreter as a semiotically closed localized (bounded) system that survives or self-reproduces in an open environment by virtue of its memory-stored constructions and controls. That distinguishes interpreters from inanimate physical systems that evolve dynamically simply because they follow the memoryless state-determined laws of nature. I believe that this minimal concept of interpreter is consistent with Ghiselin's (1997) more elaborate definition of an "individual" that also applies to higher levels, like species. However, just as there are no single symbols that have meaning, so there are no single interpreters capable of efficient evolution.

Symbols exist only in the context of codes and interpreters. Symbols are recognized in an individual interpreting system just because they function in propagating the system. But we cannot stop there. We immediately see that "propagating a system" is ambiguous. The individual interpreter is not enough. The whole idea of evolution by variation and natural selection depends on a *population* of individuals that can differ in their heritable memories. This leads directly to the central issue of evolution: what kinds of symbolic descriptions, control constraints and material constructions promote survival of *populations*? Of course there is no predictable answer to this question except the course of evolution itself. All we can do is look carefully at what is actually going on in existing organisms, and see if we can discover some answers to von Neumann's question of why the molecules are the sorts of thing they are. I will mention some properties of memory, codes, symbolic control, and material construction that studies suggest promote efficient evolutionary search and natural selection.

REQUIREMENTS FOR EFFICIENT SEARCH AND SELECTION

After asking this question von Neumann remarked that it was “a very peculiar range” for the parts since they were many orders of magnitude larger than the physically elementary particles. He did not discuss this except to suggest that the size had to do with the reliability of control since in automata there is a direct correlation between number and size of parts and reliability. A certain level of reliability is certainly one requirement in order to prevent error catastrophe, but another way to look at the question is in terms of function. How small could an enzyme be and accurately bind a substrate and catalyze a specific single bond. It would have to be a large enough structure to establish a shape with the necessary specificity to recognize a substrate by folding up a linear chain. Simple models suggest that of the order of 100 amino acids is necessary.

This size creates two fundamental problems. The first problem is that the number of copolymer sequences of such lengths is immense, well beyond actual enumeration. One of the oldest, non-religious arguments against Darwinian evolution is the apparent improbability of chance mutations producing any successful protein, let alone a species. This is still an argument used by “intelligent design” advocates. This argument is based on the assumption of the sparseness of functional sequences and the immensity of the search space. The weakness of this argument is that the actual probabilities of the events in question are largely unknown.

Formulated in biosemiotic terms, to address this problem we need to know what fraction of the innumerable potential symbol strings in a genetic memory has some meaning or function when expressed by a population of individual interpreters. We need to know how the enormous space of sequences maps into the space of biological functions. The second classical problem is that functions appear to be discretely separated. That is, one function does not smoothly transform into another function. This results in the so-called trapping problem on a function or fitness landscape.

Both these problems have been studied extensively, greatly assisted by the use of computational models. Of course, there are no pure theoretical answers. Some basic empirical knowledge is required of the actual polymers that form the memory sequence space, the nature of codes that map to protein sequences, the nature of folding, and the nature of the constructive or controlling enzymes. The auspicious discovery of molecular genetics was that many mutations are neutral with respect to function and fitness (Kimura, 1983). Along with the redundancy in the genetic code, this neutrality permits searches over a wide region near a function optimum or a local fitness peak thereby alleviating the trapping problem. Trapping is also greatly reduced by the large number of saddle regions that increases with the dimensionality of the memory sequence space (e.g., Kanerva, 1988). This lends weight to the concept of quasispecies and the advantages of mutation rates near the error threshold (Eigen, 1971; Eigen and Schuster, 1979).

This search problem has been studied extensively for the simple RNA worlds of sequences and their folding (e.g., Schuster, et al., 1994; Schuster, 1998; Crutchfield and Schuster, 2003). Again the mapping of passive memory sequences to shapes that could function as enzymes appears to be highly redundant with many sequences resulting the same three-dimensional shape. Furthermore, these sequences are distributed more or less uniformly over the entire sequence space. This means that a random search need not find just one needle in a haystack, but only one of many needles uniformly distributed over the whole haystack. That is, wherever a random search begins in sequence space, it appears likely that a description of a useful molecule will be found nearby.

ANALOGIES AND DISANALOGIES OF GENETICS WITH NATURAL LANGUAGE

Biosemiotics is the study of all forms of signification and communication. It recognizes that life is distinguished from the nonliving world by its dependence on signs and symbols. However, of the innumerable examples of pattern recognition, recording, signaling, and communication throughout all levels of living organizations only two clear examples of open-ended, creative language systems exist, the genetic language and natural languages. The similarities of genetic sequences and natural language have struck linguists as well as biologists and physicists (e.g., Jakobson, 1970). These two languages can be characterized by (1) a small, fixed alphabet, (2) one-dimensional expressions in discrete sequences, (3) an immense sequences space with no significant restriction or bias from physical laws, (4) expressions not limited in what they can potentially describe by what currently exists, (5) the interpretation of sequences, their function or meaning is complex requiring highly parallel processing. In the case of genetic sequences, the essential step is folding in which many strong constraints and weaker forces act in parallel. In the case of the brain, millions of neurons are involved in interpreting even the simplest expressions (Pattee, 1980).

Natural language structure also illustrates the strong and weak bond principle, not with a hierarchy of physical forces but with a hierarchy of rules. The lexical rules are the most rigid beginning with the alphabet and the words in the lexicon. The grammar rules are weaker than the lexical rules in the sense that syntax cannot control or modify the alphabet or the dictionary. The semantics of the text does not generally alter syntax. We usually assume our writing will not change the basic meanings of words or the grammar rules depending on what we write. Similarly the sequence or meaning of the code's base triplets is not changed by the functions of enzymes they describe. Notably however, both languages have evolved exceptions to these rules, the genetic system with special editing enzymes, reverse transcription and cutting and splicing, and natural language with freedom to invent metaphors, add new words, and to violate grammar rules with figures of speech.

Of course there are enormous differences between these languages both in their embodiments, their stability, and in their range of meanings which one would certainly expect considering they originated only at the very beginning and the end of the evolutionary time scale. The genetic language began with the origin of life, and it took 4 billion years of evolution to create brains with the capability to create natural languages. The genetic language can be called highly successful in creating adaptive functions that have kept life going over this enormous time span.

As I suggested at the beginning of this paper, whether language will turn out to be a long-term evolutionary success is not at all obvious. We often refer to natural language as the defining characteristic of human intelligence. The power of language has dominated history and shaped all our cultures. Human language has not existed for more than 100,000 years and it is quite possible that it has become too persuasive for generating myths and wishful thinking that avoid basic survival necessities for the species. Also, the technology that depends on language now allows us to design genetic messages that satisfy immediate human desires rather than long-term survival of the species. Assuming humans survive the dangers of natural language and technology, one wonders what higher level of languages might evolve in 100,000 years. If humans do not survive natural language and technology, one wonders what alternative biosemiotic structures might evolve in its place.

NOTES

Sections of this paper are edited and updated selections from H. H. Pattee, The Physics and Metaphysics of Biosemiotics, *Journal of Biosemiotics* 1(1), 223–238 (2005).

[1] This statement applies to the relatively narrow range of time and energy domains within which living organisms have been found to exist on earth. Fundamental particle and cosmological theories are far outside these domains, although the possible relevance of these theories to other conceivable forms of life is an open question.

[2] Natural periodic motions like the rotation of the earth and the emission frequencies of atoms also serve as a reference for clocks, but without arbitrary and often elaborate dissipative constraints the function of any clock, that is, the measurement of time, does not occur. The word control is also sometimes used in a more general sense to describe parameters in physical systems where no function or measurement is involved.

[3] Physicists and engineers often use information in a structural rather than functional sense because of its formal relation to the entropy of a system. *Structural information* is defined in communication theory (e.g., Shannon and Weaver, 1949). Also in quantum processes one may think of structural information being transferred from the quantum system to the observing system (e.g., Zurek, 1990). I am restricting my usage to *semantic information* that functions in the survival of biological organisms and populations.

REFERENCES

Barbieri, M. (2003). *The Organic codes. An Introduction to Semantic Biology.* Cambridge University Press, Cambridge, UK.

Barbieri, M. (2004). Life is “artifact-making,”*J. Biosemiotics*, 1(1), 81–101.

Crutchfield, J. P. and Schuster, P., Eds. (2003). *Evolutionary Dynamics: Exploring the Interplay of Selection, Accident, Neutrality, and Function.* Oxford University Press, New York.

Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules, *Naturwissenschaften,* 58, 465–523.

Eigen, M. and Schuster, P. (1979). *The Hypercycle – A Principle of Natural Self-organization.* Springer-Verlag, Berlin.

Frauenfelder H. and Wolynes, P. G. (1994). Biomolecules: Where the physics of complexity and simplicity meet, *Physics Today* February 1994, pp. 58–64.

Ghiselin, M. T. (1997). *Metaphysics and the Origin of Species*. University of New York Press, Albany, New York.

Hoffmeyer, J. and Emmeche, C. (1991). Code duality and the semiotics of nature. In M. Anderson and F. Merrell, eds., *On Semiotic Modeling.* Mouton de Gruyter, New York.

Hoffmeyer, J. (2000). "Code-duality and the epistemic cut". In Jerry Chandler and Gertrudis Van de Vijver, Eds., *Closure. Emergent Organizations and their Dynamics.* Annals of the New York Academy of Sciences, Vol. 901. pp. 175–186.

Hoffmeyer, J. (1998). The unfolding semiosphere. In G. V. de Vijver, S. Salthe, and M. Delpos, Eds., *Biological and Epistemological Perspectives on Selection and Self-Organization*, Kluwer, Dordrecht.

Jakobson, R. (1970). Linguistics. *In Main Trends of Research in the Social and Human Sciences* I, Paris-The Hague: Mouton-UNESCO, p. 438.

Juarrero, Alicia (1998). Causality as constraint. In *Evolutionary Systems. Biological and Epistemological Perspectives on Selection and Self-Organization*, G. V. de Vijver, S. Salthe, and M. Delpos, Eds., Kluwer, Dordrecht.

Kanerva, P. (1988). *Sparse Distributed Memory*, MIT Press, Cambridge, MA.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge, UK.

Mayr, Ernst (1982). *The Growth of Biological Thought*. The Belknap Press of Harvard University Press, Cambridge, MA.

Monod, Jacques (1971).*Chance and Necessity*, Knopf, New York.

Pattee, H. H. (1969). How does a molecule become a message? *Developmental Biology Supplement*, 3, 1–16.

Pattee, H. H. (1972). Laws and constraints, symbols and languages. In Waddington, C. H. Ed., Toward a Theoretical Biology 4. Essays, Edinburgh University Press, Edinburgh, pp. 248–258.

Pattee, H. H. (1973). The physical basis and origin of hierarchical control, in *Hierarchy Theory: The Challenge of Complex Systems*, H. Pattee, Ed., George Braziller, New York, pp. 73–108.

Pattee, H. H. (1980). Clues from molecular symbol systems. In *Signed and Spoken Languages: Biological Constraints on Linguistic Forms*, Bellugi, U. and Studdert-Kennedy, M. Eds. Dahlem Konferenzen, Verlag Chemie GmbH, pp. 261–273.

Pattee, H. H. (2001a). The physics of symbols: bridging the epistemic cut, *BioSystems*, 60, 5–21.

Pattee, H. H. (2001b). Irreducible and complementary semiotic forms, *Semiotica*, 134(1–4), 341–358.

Polanyi, M. 1968. Life's irreducible structure. *Science* 160, 1308–1312.

Planck, Max (1960). *Survey of Physical theory.* Dover, New York, p. 64 and p. 66.

Pauli, W. (1994) The philosophical significance of the idea of complementarity, in *Writings on Physics and Philosophy*, C. P. Enz and K. von Meyenn, eds., Springer-Verlag, Berlin, 1994, pp. 35–48 (see p. 41). First published under the title "Die philosophische Bedeutung der Idee der Komplementarität" in *Experientia* 6(Heft 2), pp. 72–75, 1950.

Rocha, Luis (2001). Evolution with material symbol systems, *BioSystems*, 60(1–3), 95–121.

Schuster, P, Fontana, W., Stadler, P. F., and Hofacker, I. L. (1994). From sequences to shapes and back: a case study in RNA secondary structures, *Proc. Roy. Soc. Lond.* B 255, 279–284.

Schuster,P. (1998). Evolution in an RNA world. In M. G. Ord and L. A. Stocken, Eds., *Foundations of Modern Biochemistry, Vol. IV: More Landmarks in Biochemistry*. JAI Press, Stamford, CT, pp. 159–198.

Shannon, C. and Weaver, W. (1949). *The Mathematical Theory of Communication*. Univ. of Illinois Press, Urbana, IL.

Simon, H. A., (1962). The architecture of complexity: hierarchic systems, *Proc. Am. Philosophical Soc.* 106 (Dec. 1962), 467–482. Reprinted in Simon, H. A., *The Sciences of the Artificial*, MIT Press, Cambridge, MA, 1996, pp. 183–216.

Von Neumann, J. (1966). *Theory of Self-reproducing Automata*. Edited and completed by A. W. Burks, University of Illinois Press, Urbana and London, pp. 74–87 and pp. 121–123.

Von Neumann, J. (1955). *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, NJ, pp. 418–421.

Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995). Navigating the folding routes, *Science*, 267, 1.

Zurek, W. H. 1990. *Complexity, Entropy, and the Physics of Information*. Addison-Wesley, Redwood City, CA.