Biol Philos
DOI 10.1007/s10539-007-9066-x

Propagating organization: an enquiry

**Stuart Kauffman · Robert K. Logan · Robert Este
Randy Goebel · David Hobill · Ilya Shmulevich**

Received: 2 January 2007 / Accepted: 19 February 2007

Abstract Our aim in this article is to attempt to discuss propagating organization of process, a poorly articulated union of matter, energy, work, constraints and that vexed concept, ''information'', which unite in far from equilibrium living physical systems. Our hope is to stimulate discussions by philosophers of biology and biologists to further clarify the concepts we discuss here. We place our discussion in the broad context of a ''general biology'', properties that might well be found in life anywhere in the cosmos, freed from the specific examples of terrestrial life after 3.8 billion years of evolution. By placing the discussion in this wider, if still hypothetical, context, we also try to place in context some of the extant discussion of information as intimately related to DNA, RNA and protein transcription and translation processes. While characteristic of current terrestrial life, there are no compelling grounds to suppose the same mechanisms would be involved in any life form able to evolve by heritable variation and natural selection. In turn, this allows us to discuss at least briefly, the focus of much of the philosophy of biology on population genetics, which, of course, assumes DNA, RNA, proteins, and other features of terrestrial life. Presumably, evolution by natural selection—and perhaps self-organization—could occur on many worlds via different causal mechanisms. Here we seek a non-reductionist explanation for the synthesis, accumulation,

S. Kauffman · R. K. Logan · R. Este
Institute for Biocomplexity and Informatics, University of Calgary, Calgary, AB, Canada

R. K. Logan (✉)
Department of Physics, University of Toronto, Toronto, ON, Canada
e-mail: logan@physics.utoronto.ca

R. Goebel
Department of Computer Science, University of Alberta, Edmonton, AB, Canada

D. Hobill
Department of Physics and Astronomy, University of Calgary, Calgary, AB, Canada

I. Shmulevich
Institute for Systems Biology, Seattle, WA, USA

and propagation of information, work, and constraint, which we hope will provide some insight into both the biotic and abiotic universe, in terms of both molecular self reproduction and the basic work energy cycle where work is the constrained release of energy into a few degrees of freedom. The typical requirement for work itself is to construct those very constraints on the release of energy that then constitute further work. Information creation, we argue, arises in two ways: first information as natural selection assembling the very constraints on the release of energy that then constitutes work and the propagation of organization. Second, information in a more extended sense is ''semiotic'', that is *about* the world or internal state of the organism and requires appropriate response. The idea is to combine ideas from biology, physics, and computer science, to formulate explanatory hypotheses on how information can be captured and rendered in the expected physical manifestation, which can then participate in the propagation of the organization of process in the expected biological work cycles to create the diversity in our observable biosphere. Our conclusions, to date, of this enquiry suggest a foundation which views information as the construction of constraints, which, in their physical manifestation, partially underlie the processes of evolution to dynamically determine the fitness of organisms within the context of a biotic universe.

Keywords Propagating organization · Constraints · Information · Shannon · Work · Semiosis · Adjacent possible

An organized being is then not a mere machine, for that has merely moving power, but it possesses in itself formative power of a self-propagating kind which it communicates to its materials though they have it not of themselves; it organizes them, in fact, and this cannot explained by the mere mechanical faculty of motion.

Immanuel Kant – Critique of Judgement

Introduction

Our broad aim is to understand propagating organization as exemplified by the vast organization of the co-evolving biosphere. Our effort is a rather mysterious undertaking, for we entirely lack a theory of organization of process, yet the biosphere, from the inception of life to today manifestly propagates organization of process. Indeed, we believe that the evolving universe as a whole also manifests the propagation of organization. We shall focus most of our efforts on the biotic case, but undertake an initial extension of our analysis to the abiotic case as well.

The role of information in biology, what it ''is,'' how it accumulates, and how it is ''used,'' has been directly addressed by mainstream biologists and philosophers of biology. By and large, the biological concept of information derives from the DNA, RNA, protein processes of ''coding'', transcription, and translation. Yet in the broader sense that we seek to articulate, information in terrestrial life is likely to be one of the key unifying concepts in the emerging field of systems biology. As part of the propagating organization within living cells, the cell operates as an information-processing unit, receiving information from its environment, propagating that information through complex molecular networks, and using the information stored in its DNA and cell-molecular systems to mount the appropriate response. Indeed, biology is acquiring many characteristics of an information science (Hood and Galas 2003).

It is sometimes the case that science progresses by finding the concepts and language to "see that which is directly in front of us." Such is the case with the present enquiry. We are persuaded that we are not wholly successful, but hope that we shall have at least started a far broader discussion.

Two predecessors to this article can be found in *Investigations* (Kauffman 2000), and "Emergence, Autonomous Agents, and Organization" (Kauffman and Clayton 2006). At its core, *Investigations* seeks to understand the physical nature of agency itself, and proposes that a molecular autonomous agent, able to act on its own behalf in an environment, is an autocatalytic system carrying out at least one thermodynamic work cycle. Much follows from this tentative definition, which implies that an autonomous agent is an open non-equilibrium chemical system, and finds general biotic importance in the fact that work cycles link spontaneous and non-spontaneous (exergonic and endergonic) processes. This linkage has built up the enormous complexity of the biosphere.

Further analysis reveals this work to be the constrained release of energy into a few degrees of freedom. But if one asks where the constraints themselves come from—as in the example of a cylinder and piston that confine the expansion of the working gas in the head of the cylinder to yield the translational motion of the piston, hence the release of energy into a few degrees of freedom—one finds that it typically takes work to construct the constraints[1]

Thus we arrive at the first surprise—it takes constraints on the release of energy for work to happen, but work for the constraints themselves to come into existence. This circle of work and constraint shall turn out to be part of the theory of propagating organization that we shall discuss.

Most importantly, contemporary cells are both collectively autocatalytic and do work cycles, in part to construct constraints on the release of energy. When released, this energy constitutes further work that drives non-spontaneous processes, builds structures, drives processes, and also builds further constraints on the release of energy, which when released can build still more such constraints. In short, cells carry out propagating work linking spontaneous processes, constraints, work, and non-spontaneous processes, and more broadly as we shall see, the propagating organization of process. In doing so, the cell carries out a set of interlocked tasks that achieve a closure of tasks whereby the cell literally builds a rough copy of itself. We know this, yet we have no clear way to say what we know. This closure of work, constraints, tasks, and information, as we shall see below, is a new state of matter, energy, information, and organization that constitutes the living state.

The new insight that we explore in this article is that the constraints that allow autonomous agents to channel free energy into work are connected to information: in fact, simply put, the constraints *are* the information, are partially causal in the diversity of what occurs in cells, and are part of the organization that is propagated.

In "Emergence, Autonomous Agents, and Organization" (Kauffman and Clayton 2006), the tentative definition of autonomous agent is extended to include construction of boundaries enclosing the agent, discrimination of "yuck" (meaning

[1] Here we use the word "constraint" in a very general sense that includes "global constraints" (e.g. conservation of energy, symmetry conditions etc.,) and "local constraints" or boundary conditions (e.g. initial conditions, reflection or absorption at a spatial location).

poison) or ''yum'' (meaning food), and at least one choice of action: flee (or not), approach (or not). Our language is teleological. We believe that autonomous agents constitute the minimal physical system to which teleological language rightly applies.

It is important that our definition of a molecular autonomous agent applies to terrestrial life, but is, in principle far broader. The concepts identify a new class of far from equilibrium chemical thermodynamic systems, and we suspect, could form the basis of life in a variety of molecular instantiations. For example, already Sievers and von Kiedrowski (1994) and Lee et al. (1997a, b), have made collectively self-reproducing DNA and peptide systems. Quite directly, Lee et al. (ibid.) have shown that self-reproduction does not depend upon the double helix structure of DNA and RNA. Thus self-reproduction on a basis other than template replication, transcription and translation has been achieved. Further, work on the origin of life based on self reproducing liposomes (Mavelli and Luisi 1996), the theory of the probable emergence of collectively autocatalytic sets of molecules (Kauffman 1993), and autocatalysis in organic reaction mixtures (Smith and Morowitz 2004), begins to suggest a broad physical basis for life in the cosmos. Molecular autonomous agents have yet to be created, but work cycles and molecular motors are accomplished experimental facts.

If we succeed in creating, or finding life which is radically different from contemporary earth life, the way will open up for a general biology, and a new union of physics, chemistry, biology, and the information sciences. Core to this, we feel, will be an understanding of propagating organization of process.

We comment that it is unlikely that very early life was based on DNA, RNA and proteins via transcription and translation, given the huge complexity of the molecular apparatus to achieve these events, including encoded enzymes that charge transfer RNA with the correct amino acids to achieve translation. We emphasize this, because we wish to place a discussion of information, and its relation to work, constraint, and propagating organization in a wider context that the contemporary debate among philosophers of biology and biologists about the information status of the DNA $\rightarrow$ RNA $\rightarrow$ protein chain.

In turn, a general biology will necessarily confront us with a discussion of evolution by heritable variation and natural selection, perhaps typically without the familiar concepts of DNA, gene frequencies, alterations of gene frequencies as the microevents of microevolution and the diverse philosophic opinions that have ranged over these issues. We will have to explore the general conditions allowing for evolution and the emergence of biospheres.

This article is organized as follows:

In the ''Darwinian adaptations and preadaptations'' section we discuss Darwinian adaptations and preadaptations, argue that the first implies that biology cannot be reduced to physics, while the second, stunningly, implies that the future evolution of the biosphere cannot be finitely prestated. Much follows from these surprising conclusions.

In the ''Shannon information'' section we discuss Shannon information and argue that it does not apply to the evolution of the biosphere. One reason is that due to Darwinian preadaptations, the ensemble of possibilities and their entropy cannot be calculated.

In the ''Schrödinger's aperiodic crystal...'' section we begin with Schrödinger's famous statement that a periodic crystal cannot ''say'' a lot, while an aperiodic crystal can say a lot, and will contain a microcode. We shall argue that the proper and deep understanding of Schrödinger's intuition is that an aperiodic crystal contains a very large number of diverse constraints that are partially causal in guiding the huge diversity of specific events and processes which occur physically in cells. From this we shall arrive at a new formulation: constraints are information and information is constraints. The first part of this twosome, constraints are information is, we believe, secure. The second part, information is constraints, may be more problematic.

In "The relativity of information" section we discuss the relativity of the concept of information.

In the "Semiosis as a special case..." section we shall place our definition of biotic information in the larger context in which information is "about" something, arguing that when an autonomous agent discriminates yuck or yum, the molecular signatures of yuck or yum are about yuck or yum, hence the rudiment of semiotics. We shall locate biotic (but not linguistic) semiosis, as a subcase of information as constraints.

In the "Heritable variation and natural..." section we shall stress that constraints as information, and, derivatively, semiotic information, must have causal consequences for the autonomous agent. These consequences increase its fitness such that the information is assembled by natural selection into the ongoing evolution of the biosphere. Without this coupling to fitness, the information and its effects would not come to exist in the universe. Therefore we shall argue that natural selection constitutes the assembly machinery, when coupled with heritable variation, that literally assembles the propagating organization of matter, energy, constraint, work, and information. This constitutes the propagating organization in autonomous agents, whose co-evolution drives the biosphere's progressive exploration of what we call the Adjacent Possible. This discussion is reminiscent of some aspects of Maynard Smith's argument (2000a, b) that selection confers on genes a specific informational character, and Sterelny and Griffiths (1999) broadened concept that selection confers on many features of a cell or organism the features of information.

In "The evolution of the abiotic..." section we attempt to extend our analysis to the abiotic universe. We find that our analysis that considers information as constraints is equivalent to the statement that information consists in boundary conditions and in global constraints. But, in classical and quantum physics, boundary conditions—like the cylinder and piston—are only partially causal for what occurs. Physicists often "put in by hand" the boundary conditions of a problem, such as the behavior of the cylinder, piston, and working gas system. But in the unfolding of the biosphere or universe since the Big Bang, the very coming into existence of new boundary conditions—information we argue—is itself part of the full dynamics of the total system. We thus assume a context with information understood as boundary conditions on the release of energy that makes diverse processes happen. So we argue that in the proper union of matter, energy and information it is precisely the union of matter, energy, and boundary conditions that, in an expanding and cooling universe, progressively break symmetries, invade the Adjacent Possible, and cause an increasing diversity of events, processes and structures to come into existence. The evolution of the biosphere is but one case of this general process.

In the "Population genetics and evolution..." section we briefly discuss the general context of successful evolution by heritable variation and natural selection in a general biology. Here the ''neighborhood'' relations between different autonomous agents is an issue. More essentially, without propagating organization of process there would be nothing upon which selection could act. Thus, we suggest, the habit of population genetics of ignoring the root physical basis of life may first of all constrain our understanding of evolution unnecessarily to contemporary earth life, and misses entirely what we shall describe as the evolution of perhaps any biosphere into its ''Adjacent Possible'', a fundamental feature of life that underlies the specifics of evolution by altering gene frequencies.

Darwinian adaptations and preadaptations

Were one to have asked Darwin what the function of the heart is, he would presumably have responded that the function of the heart is to pump blood. But the heart has a wealth of other causal consequences, such as heart sounds. Heart sounds are not the function of the heart. That is, the causal consequence of the heart that matters, the virtue for which it was selected, was the pumping of blood. So the function of a part (or organ) of an organism is typically, if not always, a subset of its causal consequences. This has major implications. Among these, the function of a part (or organ) of an organism cannot be analyzed except in the context of the whole organism in its selective environment. But further, this fact is just one of the reasons that biology cannot be reduced to physics. In Kauffman and Clayton (2006), it is argued that, if we grant the physicist a theory of everything, say string theory to cite one example, and the capacity to deduce upwards to all that occurs in the universe—an impossibility given throws of the quantum dice—the physicist could deduce all the causal features of the heart, *but would have no way to pick out the pumping of blood as the relevant causal property which is the function of the heart* and which is the property that gave rise to the evolutionary emergence of this organ.

To do so, the physicist would have to discuss whole organisms as causal agents in their own right, evolving under natural selection in changing environments. That is, the physicist would have to become a biologist and talk biology talk. Thus, biology cannot be reduced to physics, rather physics has to be lifted up to biology.

A second reason we feel biology is emergent with respect to physics is that Darwin's natural selection is utterly neutral with respect to the physical basis of heritable variation and hence natural selection. Life might be based on DNA, RNA and proteins, or might be based on autocatalytic organic chemical reactions systems, and/or polymer systems that create a bounding membrane. This implies that a physicist armed with a theory of everything might, (actually could not) deduce that a specific molecular autonomous agent would have offspring of differential reproductive success, the physicist cannot deduce Darwin's natural selection itself which transcends any specific realization of it. Indeed, for small changes in the constants of nature, life might still be possible, hence Darwin's ''higher order'' or emergent law cannot even be reduced to the physics of this universe.

In short, for these and other reasons, we wish to join forces with those who argue for a limitation of reductionism, and the reality of emergence with respect to the furniture of the universe (Silberstein 2003).

Darwin had many brilliant insights. Among these is what is now called a Darwinian preadaptation. Here the central concept is that a causal property of a part of an organism that is not of selective significance in the normal environment might become useful in a different environment, and hence become subject to selection. It is critical to point out, first that Darwinian preadaptations have occurred repeatedly in evolution, and second, that such an evolutionary step results in the emergence in the biosphere of a novel function. For example, lungs evolved from the swim bladders of certain early fish. The swim bladders, partially filled with water and partially with air, adjusted the height in the water column to establish neutral buoyancy of the fish. But the swim bladder, with air in it, was preadapted for use as a lung, and air breathing was a novel functionality with its own causal consequences that allowed life to conquer land thereby changing the universe.

We now raise a central question discussed in *Investigations*. Is it possible to say ahead of time what all possible Darwinian preadaptations are for human beings, or for the whole biota of the contemporary biosphere for that matter? The answer appears to be "No." We cannot finitely prestate all possible Darwinian preadaptations. Part of the difficulty, or impossibility, in doing so is that we cannot even begin the task of prestating what all possible selective environments will be. That is, there appears to be no finitely stateable procedure which would allow us to enumerate all possible selective environments.

Part of the challenge is that the concept of such environments is systematically vague. It is not even clear how to begin on the project of listing all possible environments for all actual, let alone possible, organisms. While we do not know how to prove our claim, we believe it to be true and shall assume that it is.

We point out that the property or causal consequence which becomes the subject of a Darwinian preadaptation need not be a mutant property. It might be a normal feature of the organism, but normally of no selective significance until the new environment is encountered. Therefore, an attempt to enumerate the possible preadaptations by trying to count the number of mutations possible to a genome is irrelevant. Darwinian preadaptations cannot, in general, be prestated.

Much follows from the claim that we cannot finitely prestate all possible Darwinian preadaptations of all contemporary organisms. First, it means in a radical sense that we cannot predict the future evolution of the biosphere. We literally have no idea of what such preadaptations may be. Second, it means that a frequency interpretation of probability statements does not apply to possible probability statements about the evolution of the biosphere. In the normal frequency interpretation of probability, say that a fair coin will be heads about 5,000 times out of 10,000 coin flips, one can finitely prestate all possible outcomes. This is not possible for the evolving biosphere. Third, and dramatically, the incapacity to say ahead of time what the relevant preadaptations will be means that we cannot write down a stateable set of variables in equations whose dynamics captures the evolution of the biosphere. But all our mathematical techniques in physics begin with a prestatement of the full set of variables and the configuration space of the system. This is true in Newtonian dynamics, statistical mechanics, general relativity and in quantum mechanics if one does not believe in hidden variables. If one believes in hidden variables then because they are hidden they cannot be prestated hence the caveat for quantum mechanics.

But we cannot prestate the configuration space of the biosphere. Now a classical physicist might argue that, if we take the solar system, it is just a large classical $3N$

dimensional system where N is the number of particles in the solar system and the current biosphere within the rest of the solar system is a point in that space. Let us grant the move. Then, we rejoin, the physicist has no way to pick out the collective variables, the lungs and hearts and wings, and features of the environment that are the relevant causal variables for the ongoing evolution of the biosphere. Thus, again we see that we cannot write down causal laws with a prestated set of (collective) variables for the evolution of the biosphere.

We shall not discuss it further here, but the same incapacity to prestate the evolution of the economy and its technology also arises, as does the incapacity to prestate the evolution of human culture. But all this has the deepest implications. Reductionist science is powerful, but is limited. This sets us free in astonishing ways, for organisms live their lives forward, they do not deduce them. We appear to live in a universe in which our reductionistic world view is inadequate: there is the emergence of life, and value as we discuss below. Human language and culture also represent propagating organization (Logan 2006, 2007). Moreover we live in and partially co-create a ceaselessly ''creative'' biosphere, economy, and human culture. This glimmers a new scientific world view, beyond reductionism with broad potential societal ramifications (Kauffman 2006).

Shannon information

Shannon (1948) information theory has been a brilliant mathematical construct. At its core, Shannon envisioned a Source with a set of messages, symbol strings, over which a well defined probability distribution might be attributed. Then he envisioned a (perhaps noisy) channel over which information is transmitted. He then envisioned a receiver and, importantly, a decoder. Shannon's move was to calculate the entropy of the set of messages at the Source. The information that propagated down the channel and was received at the receiver removed uncertainty with respect to the entropy of the Source. This reduction of uncertainty, hence the lowering of the entropy of the Source, constitutes the amount of information transmitted. One interpretation, not given by Shannon himself who abjured to say what information ''is,'' is that information is just the reduction in uncertainty at the receiver. This definition leaves open exactly what the claim might mean. It might be the reduction of uncertainty in a human receiver's mind, for example.

Importantly, and widely recognized, is the fact that Shannon information considers the amount of information, nominally in bits, but is devoid of semantics. There is no sense of what information is ''about'' in Shannon information.

Now we ask whether Shannon information applies to the evolution of the biosphere. We answer that it does not. In particular, Shannon information requires that a prestated probability distribution (frequency interpreted) be well stated concerning the message ensemble, from which its entropy can be computed. But if Darwinian preadaptations cannot be prestated, then the entropy calculation cannot be carried out ahead of time with respect to the distribution of features of organisms in the biosphere. This, we believe, is a sufficient condition to state that Shannon information does not describe the information content in the evolution of the biosphere.

There are further difficulties with Shannon information and the evolving biosphere. What might constitute the ‘‘Source’’? Start at the origin of life, or the last common ancestor. What is the source of something like ‘‘messages’’ that are being transmitted in the process of evolution from that Source? The answer is entirely unclear. Further, what is the transmission channel? Contemporary terrestrial life is based on DNA, RNA, and proteins via the genetic code. It is insufficient to state that the channel is the transmission of DNA from one generation to the next. Instead, one would have to say that the actual ‘‘channel’’ involves successive life cycles of whole organisms. For sexual organisms this involves the generation of the zygote, the development of the adult from that zygote, the pairing of that adult with a mate, and a further life cycle. Hence, part of one answer to what the ‘‘channel’’ might be is that the fertilized egg is a channel with the Shannon information to yield the subsequent adult. But it has turned out that even if all orientations of all molecules in the zygote were utilized, there is not enough information capacity to store the information to yield the adult. This move was countered by noting that, if anything, development is rather more like an algorithm than an information channel (Apter and Wolpert 1965). In short, a channel to transmit Shannon information along life cycles does not exist, so again, Shannon information does not seem to apply to the biosphere.

It seems central to point out that the evolution of the biosphere is not the transmission of information down some channel from some source, but rather the persistent, ongoing, co-construction, via propagating organization, heritable variation, and natural selection, of the collective biosphere. Propagating organization requires work. It is important to note that Shannon ignored the work requirements to transmit ‘‘abstract’’ information, although it might be argued that the concept of constraints is implicit in the restrictions on the messages at the Source. While we mention this, we have no clear understanding physically of what such constraints are.

One might be tempted to argue that a Shannon-like information theory could be applied to the vast set of selective events that have led to the specific DNA sequences that are in contemporary organisms. But does this move work? Can we specify a finite ensemble of possible DNA sequences out of which the present DNA sequences have been derived? If we consider all DNA sequences longer than, say 1,000 nucleotides, it would take vastly large repetitions of the history of the universe for the universe to construct one copy of each possibility. This cannot physically constitute the ensemble. Is the ensemble the set of DNA sequences that have been explored in the actual evolution of the biosphere, some accepted, most rejected? This approach initially seems promising, but has the obvious difficulty that we cannot specify the ensemble explored in 3.8 billion years, hence do not and cannot know the Shannon information content of the biosphere. A further difficulty with this approach is that it measures the information content of the biosphere as a function of the number of DNA sequences ‘‘tried’’ in evolution. But very different numbers of attempted mutations might have led to the same biosphere, hence quantitating the information of the biosphere by the number of attempted DNA mutations is not in direct correspondence to any specific biosphere.

We conclude that a Shannon Information content analysis of the information content of the evolving biosphere is not legitimate.

Schrödinger's aperiodic crystal: "instructional" information as constraint or boundary condition

In *What is Life*, Schrödinger (1992) is concerned with the order in organisms and hence the physical basis of the gene. He argues, based on X-ray mutation induction frequency, that the gene must have a few hundred to a few thousand atoms, and points out that statistical mechanical equilibria cannot account for the stability of the organism over generations. He then posits that quantum mechanics in the form of chemical bonds is the answer. Then he brilliantly points out that the order of life cannot be based on a periodic crystal, for such a crystal cannot say a lot, or carry much information. He places his bet on aperiodic crystals which can, in strong contrast, say a lot, or carry much information, even a microcode which will somehow specify the adult.

He was brilliantly right, and presaged DNA and the genetic code. Now we come to the critical issue. In just what sense can an aperiodic crystal "say a lot?" Schrödinger does not himself say more than suggesting that the aperiodic crystal can contain a microcode.

We believe Schrödinger was deeply correct, and that the proper and deep understanding of his intuition is precisely that an aperiodic solid crystal can contain a wide variety of microconstraints, or micro boundary conditions, that help cause a wide variety of different specific events to happen in the cell or organism. Therefore we starkly identify information, which we here call "instructional information" or "biotic information," not with Shannon, but with constraints or boundary conditions. The amount of information will be related to the diversity of constraints and the diversity of processes that they can partially cause to occur. By taking this step, we embed the concept of information in the ongoing processes of the biosphere, for they are causally relevant to that which happens in the unfolding of the biosphere.

We therefore conclude that constraints are information and, as we argue below, information is constraints which we term as instructional or biotic information to distinguish it from Shannon information. We use the term "instructional information" because of the instructional function this information performs and we sometimes call it "biotic information" because this is the domain it acts in, as opposed to human telecommunication or computer information systems where Shannon information operates. This step, identifying information as constraint or boundary condition, is perhaps the central step in our analysis. We believe it applies in the unfolding biosphere and the evolving universe, expanding and cooling and breaking symmetries, that we will discuss below.

Is this interpretation right? It certainly seems right. Precisely what the DNA molecule, an aperiodic solid, does, is to "specify" via the heterogeneity of its structural constraints on the behavior of RNA polymerase, the transcription of DNA into messenger RNA. Importantly, this constitutes the copying or propagating of information. Also, importantly, typically, the information contained in aperiodic solids requires complex solids, i.e., molecules, whose construction requires the linking of spontaneous and non-spontaneous, exergonic and endergonic, processes. These linkages are part of the work cycles that cells carry out as they propagate organization.

It is essential to note that the set of constraints in a contemporary cell is not merely the DNA and RNA, but lies also in the specific stereochemistry of a vast horde of specific molecular species. So, when an enzyme binds two substrates and holds them in proximity, lowering the potential energy barrier to their joining, the enzyme is acting as a constraint on the motion of the two substrates, hence as a catalyst. The working of a cell is, in part, a complex web of constraints, or boundary conditions, which partially direct or cause the events which happen. Importantly, the propagating organization in the cell is the structural union of constraints as instructional information, the constrained release of energy as work, the use of work in the construction of copies of information, the use of work in the construction of other structures, and the construction of further constraints as instructional information. This instructional information further constrains the further release of energy *in diverse specific ways*, all of which propagates organization of process that completes a closure of tasks whereby the cell reproduces.

Our discussion here has some of the flavor of Sterelny and Griffiths (1999) in their discussion of an extended concept of information beyond DNA, RNA and protein sequences. On the other hand, none of those who have written on the concept of information in biology have taken up the struggle to relate it to constraints, work, and propagating organization of process such as that in reproducing cells.

The relativity of information

In the ‘‘Shannon information’’ section we have argued that the Shannon conception of information are not directly suited to describe the information of autonomous agents that propagate their organization. In the ‘‘Schrödinger's aperiodic crystal...’’ section we have defined a new form of information, instructional or biotic information as the constraints that direct the flow of free energy to do work.

The reader may legitimately ask the question ‘‘isn't information just information?’’, i.e., an invariant like the speed of light. Our response to this question is *no*, and to then clarify what seems arbitrary about the definition of information. Instructional or biotic information is a useful definition for biotic systems just as Shannon information was useful for telecommunication channel engineering, and Kolmogorov (Shiryayev 1993) information was useful for the study of information compression with respect to Turing machines.

The definition of information is relative and depends on the context in which it is to be considered. There appears to be no such thing as absolute information that is an invariant that applies to all circumstances. Just as Shannon defined information in such a way as to understand the engineering of telecommunication channels, our definition of instructional or biotic information best describes the interaction and evolution of biological systems and the propagation of organization. Information is a tool and as such it comes in different forms. We therefore would like to suggest that information is not an invariant but rather a quantity that is relative to the environment in which it operates. It is also the case that the information in a system or structure is not an intrinsic property of that system or structure; rather it is sensitive to history and environment. To drive home this point we will now examine the historic context in which Shannon (1948) information emerged.

Before delving into the origin of Shannon information we will first examine the relationship of information and materiality. Information is about material things and furthermore is instantiated in material things but is not material itself. Information is an abstraction we use to describe the behavior of material things and often is thought as something that controls, in the cybernetic sense, material things. So what do we mean when we say the constraints are information and information is constraints as we did in the ''Schrödinger's aperiodic crystal...'' section.

''The constraints are information'' is a way to describe the limits on the behavior of an autonomous agent who acts on its own behalf but is nevertheless constrained by the internal logic that allows it to propagate its organization. This is consistent with Hayle's (1999, p. 72) description of the way information is regarded by information science: ''It constructs information as the site of mastery and control over the material world.'' She claims, and we concur, that information science treats information as separate from the material base in which it is instantiated. This suggests that there is nothing intrinsic about information but rather it is merely a description of or a metaphor for the complex patterns of behavior of material things. In fact, the key is to what degree information is a completely vivid description of the objects in question.

This understanding of the nature of information arises from Shannon's (1948) original formulation of information, dating back to his original paper:

> The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one selected from a set of possible messages. The system must be designed to operate for each possible selection, not just the one that will actually be chosen since this is unknown at the time of design. If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely.

A number of problems for biology emerge from this view of information. The first is that the number of possible messages is not finite because we are not able to prestate all possible preadaptations from which a particular message can be selected and therefore the Shannon measure breaks down. Another problem is that for Shannon the semantics or meaning of the message does not matter, whereas in biology the opposite is true. Biotic agents have purpose and hence meaning.

The third problem is that Shannon information is defined independent of the medium of its instantiation. This independence of the medium is at the heart of a strong AI approach in which it is claimed that human intelligence does not require a wet computer, the brain, to operate but can be instantiated onto a silicon-based computer. In the biosphere, however, one cannot separate the information from the material in which it is instantiated. The DNA is not a sign for something else it is the actual thing in itself, which regulates other genes, generates messenger RNA, which in turn control the production of proteins. Information on a computer or a telecommunication device can slide from one computer or device to another and then via a printer to paper and not really change, McLuhan's ''the medium is the message'' aside. This is not true of living things. The same genotype does not always produce the same phenotype.

According to the Shannon definition of information, a structured set of numbers like the set of even numbers has less information than a set of random numbers because one can predict the sequence of even numbers. By this argument, a random soup of organic chemicals has more information that a structured biotic agent. The biotic agent has more meaning than the soup, however. The living organism with more structure and more organization has less Shannon information. This is counterintuitive to a biologist's understanding of a living organism. We therefore conclude that the use of Shannon information to describe a biotic system would not be valid. Shannon information for a biotic system is simply a category error.

A living organism has meaning because it is an autonomous agent acting on its own behalf. A random soup of organic chemicals has no meaning and no organization. We may therefore conclude that a central feature of life is organization—organization that propagates.

Semiosis as a special case of constraint as information

We wish next to consider the minimal physical conditions for semiosis. We shall not concern ourselves with fully human linguistic symbols, but with the semiosis of our minimal molecular autonomous agent. Consider an agent that is confronted by molecules in its environment, which constitute "yuck" or "yum." To respond to these environmental features, the agent, assumed to be bounded (Kauffman and Clayton 2006), must also have yuck and yum receptors, capable in the simplest case of "recognizing" molecules of yuck or yum, and responding appropriately by avoiding yuck and eating yum. Assume such molecular machinery exists in the agent. They of course exist in prokaryotic and eukaryotic cells. We wish to say that the agent confronting yuck or yum receives information "about" yuck or yum. This appears to constitute the minimal physical system to which semiotic information might apply. And it is worth noting that the "meaning," or semiotic content of the yuck and yum molecules is built into the propagating organization of the cell. The cell, we want to say, has embodied knowledge and know-how with respect to the proper responses to yuck and yum, which was assembled for the agent and its descendants by heritable variation and natural selection.

The existence of yuck and yum as semiotic signs is a subcase of constraint as information. How does the agent detect yuck? A concrete case would be that a yuck molecule binds a yuck receptor, constraining the receptor's motions, which in turn acts as a constraint in unleashing a cell signaling cascade leading to motion away from yuck. Further, if yuck is present below a detection threshold, it will not be detected by the agent. Hence that threshold, and the receptor itself, act as a constraints partially determining the behavior of the agent in fleeing or not fleeing.

One can construct an underlying set theoretical interpretation for yuck and yum semantics in two equivalent ways: The first posits a set of instances, and a set of properties to which each instance is assigned. The second posits a set of instances and detectors, or classifying operators, that classify "properties" of instances. Note that in the second case, those properties need not themselves be discussed because the detectors do the job. If the second stance is taken, then detectors, "yuck" and "not yuck," suffice and no extension beyond instructional information is required. If the second stance suffices, we want to say not only that constraints are information but also that information is constraints. We recognize that this second step is arguable and do not analyze this issue further here.

Semiotic information can not itself embody "agentness," for it has no agency; but identified agents can be observed to respect the semiotic interpretation like yuck and yum. This inspectable behavior provides the opportunity to attribute constraint-directed behavior to the agent organism.

Another important point in this attempt to understand propagating organization is that the semiotic behavior can identify a source of free energy, yum in this case, from which work can be extracted and propagate in the cell. This behavior is part of a theory that unifies matter, energy, information and propagating organization.

We end this section with the description of a final interesting feature of the yum receptor. A wide variety of molecules might bind to the yum receptor with modest affinity, hence mimic true yum molecules. So the yum receptor can be "fooled." This might allow another agent to emit a poison that mimics the yum molecule, fools the receptor, and leads to the death of the agent. So evolves the biosphere. Now ask, can a Shannon channel be "fooled?" Clearly noise can be present in the channel. Due to noise a 1 value can replace a 0 value in the constrained sense of 1 and 0 as subsets of the physical carriers of 1 and 0. But the Shannon channel cannot be fooled: "fooling" is a semantic property of detectors, hence not present in a Shannon channel. Therefore, while one might be tempted to measure the amount of semiotic information using a Shannon-like approach, the fact that semiosis in an organism can be fooled suggests that a symbol based Shannon move is inappropriate.

We conclude that semiotic information in molecular agents such as organisms is a special case of information as constraint. For semiotic information to be "about" something, and to be extracted, it appears that a constraint must be present in one or more variables that are themselves causally derived from that which the information is about.

Like the threshold level of yum needed for detection, to use the information, the extracted semoitic information must do work on some system. That work might copy the information, for example into a record, or might construct constraints on the release of energy which is further work. Here, semiotic information becomes part of propagating organization.

We comment that in standard semiotic analyses with human agents and language, there are three elements to semiotic information, namely,

1. The subject of the information or the agent being informed;
2. The object of the information or what the information is about; and
3. The possibly arbitrary, sign or symbol referring to the object.
4. With Monod (1971) in Chance and Necessity we add that allosteric chemistry allows arbitrary molecules to cause events. If we wish to call such molecules "symbols" that "refer to" "yum," the standard semiotic analysis just noted applies to molecular autonomous agents. Note that Monod's example is broader than DNA, RNA and proteins. It is the general arbitrariness of allosteric chemistry that allows arbitrary molecules to cause events. Information is thus broader than coding.

Heritable variation and natural selection as assembly processes

We have now grounded biotic information as "instructional information" or constraint, or boundary condition, that partially causes subsequent events in the unfolding of the biosphere. In this view information is not an abstraction, but is causally efficacious in the biosphere and we argue below in the unfolding of the abiotic universe. And we have grounded semiotic information as information detected about external (to the agent) features of the environment about which it learns. These semiotic cases are also cases of constraints, or boundary conditions, detecting and categorizing inputs and partially causing subsequent events. We note again that we remain neutral for the moment about whether information needs to be extended beyond instructional information for a set theory analysis of the categorization of objects.

At the level of complex molecules, as noted above, the universe has not had time to create all possible versions. For example, the universe has not had time to create all proteins to length 200, by about 10 to the 67th power repetitions of the history of the universe.

Consider a simple set of organic molecules and all the reactions they can collectively undergo. Call the initial set of molecules the Actual. Now among the reactions that might happen, some may lead to molecular species that are not present in the initial Actual. Call these new molecular species the Adjacent Possible. They are the molecular species that are reachable in a single reaction step from the current actual. It is of fundamental importance that the biosphere has been evolving into the Adjacent Possible for 3.8 billion years, from an initial diversity of perhaps 1,000 organic molecules to trillions. The biotic world advances into the adjacent possible in terms of molecules, morphologies, species, behaviors, and technologically from pressure flaked stones; it lurks in everything from the global economy to the computer, and the millions of products in the current global economy.

Once at a level of complexity sufficiently above the atom, the universe, the biosphere, and the technosphere can never exhaust the diversity of things and events that can happen. The evolving universe and biosphere advance persistently into the adjacent possible. This means that what comes to exist at these levels of complexity is typically unique in the universe.

Now consider a heritable variation which gives rise to a new constraint, physical biotic information, that helps cause a sequence of events in a molecular agent. If that heritable variation is to the selective benefit of the agent, the new constraint, the new biotic information, will be grafted into the organism, its progeny, and the ongoing evolution of the biosphere.

It is essential to note that in the absence of heritable variation, an increase in fitness, and natural selection, this new functionality would not come to exist in the universe: but lungs and flight have come to exist. The mechanisms of heritable variation and natural selection comprise an assembly process by which propagating organization is modified in normal Darwinian adaptations and preadaptations where new functionalities arise, and these modifications are built into the ongoing evolution of the biosphere.

It is clear then, that heritable variation and natural selection are sufficient mechanisms in the biosphere to build an expanding mesh of functionalities as the biosphere invades the adjacent possible. We will ask next whether similar processes can happen in the abiotic universe.

The evolution of the abiotic expanding universe: propagating organization diversifying sources of constraint, free energy, and coupling of spontaneous and non-spontaneous processes

We here ask whether we can find generalizations of the above analysis of information, matter, energy, constraint, work, in the biosphere, in the abiotic expanding universe.

For some time, scholars have struggled to find the union of matter, energy, and information. Cases such as Maxwell's demon, the Bekenstein bound on the entropy of a black hole, and the holographic principle, all seem to be places in physics where matter, energy, and information come together. These cases merit attention, but we leave them unanalyzed, except for this comment.

For information to be united with matter and energy, information must be part of the physical unfolding of the universe. Thus, consider Maxwell's demon. It has been shown that the demon cannot "win" with respect to the Second Law of Thermodynamics for a closed equilibrium system (Kauffman 2000). However, in a non-equilibrium setting, the demon can win by making measurements that reduce the entropy of the measured system, with respect to the demon, *faster* than the most compressed record of the measured system grows, on average, in length. Now physicists usually end their argument with a claim rather like, "Then, in principle, work could be extracted." Such a statement is inadequate for a theory that unites matter, energy, and information. What is required is that, in the non-equilibrium setting, a displacement from equilibrium that is a source of free energy must be detected by at least one measurement; a physical system able to couple to that source of free energy must have come to exist and must actually extract free energy, and must release that energy in a constrained way to carry out actual work. Thereafter, this work may propagate.

If we conceive of an abiotic physical system able to carry out these processes of measurement and work extraction in the abiotic universe, it will have to be an abiotically derived system able to perform such measurements, recording the results, and employ the record of the measurements to extract actual work. Such a system will be a case of propagating organization with boundary conditions as constraints, including measurements in the record as constraints on the behavior of the system conditional on the recorded measurements, and the constrained release of energy in work. Whether the coming into existence in the universe of such a system is plausible abiotically is certainly open to question but may be worthy of consideration. Biotically, of course, such systems abound: sources of free energy from sunlight to prey are detected and coupled to work extraction. Records of sources of free energy in the form of food are seen in ant phermone trails. The measurement of a source of free energy and extracting that free energy typically involves thresholds and other constraints or boundary conditions. For example, ants will not follow a phermone trail if it is below a detection threshold, and the boundaries of the trail are boundary conditions on the ants' motions.

These considerations suggest that we take information to be constraint or its physical equivalent, boundary conditions that partially cause events, where the coming into existence of the constraint is itself part of propagating organization. If we do so, the issue starts to clarify in a simple way. It is fully familiar in physics that one must specify the laws, particles, the initial and boundary conditions, then calculate the behavior of the system in a defined state space. Now it is common, as noted, in physics, to ‘‘put in by hand’’ the boundary conditions, as in the cylinder and piston case. But in the evolving biosphere, itself part of the evolving universe, and in the evolving universe as a whole, new boundary conditions come into existence and partially determine the future unfolding of the biosphere or the universe. These evolving boundary conditions and constraints are part of the propagating organization of the universe.

We consider a single, but complex case in cosmic evolution. It is well known that molecular grains are found in interstellar space. These grains aggregate up to the scale of planetessimals. Now it is also well known that the grains have surfaces with complex molecular features on which complex chemistry appears to be occurring. The grains themselves act as constraints, or boundary conditions, that confine reacting substrates, hence may catalyze reactions, some of which may be endergonic, requiring, for example, photons. In some cases, the product molecules presumably are bound to the growing grain, thereby modifying the boundary conditions afforded by the grain, which in turn modifies the chemical reactions that can occur. Furthermore, the product molecules can be novel substrates—hence novel sources of free energy—which again allow novel chemical reactions to occur. In short, the grains appear to behave as constraints that can partially guide spontaneous or non-spontaneous processes, can, in addition, link spontaneous and non-spontaneous processes, can create new constraints enabling such processes and linked processes, and can create novel sources of free energy in the form of novel substrates able to enter into new chemical reactions.

Assume the above account is roughly correct. Then the growing grains appear to be cases in which matter, energy, and continuously evolving boundary conditions and novel sources of free energy *emerge*, and condition the future evolution of the grains. The grains are at levels of complexity sufficiently above atoms so that what occurs is typically unique in the universe. It seems virtually sure that no two modest size grains are molecularly identical. Here we confront a union of matter, energy, and evolving and diversifying boundary conditions linking, for example, spontaneous and non-spontaneous processes, and providing diversifying sources of free energy, which alter the ever diversifying structures that come to exist in the evolving expanding universe.

If this approach has merit, it appears to afford a direct union of matter, energy, and information as constraint or boundary condition.

Population genetics and evolution in any biosphere

Philosophy of biology has largely grown up in the constrained environment of current terrestrial life. Its analysis of heritable variation and selection has largely ignored the physical basis of the propagating organization and closure of tasks that achieve the living state and underlie heritable variation. Moreover, we have discussed above the fact that at levels of complexity above atoms, the universe is on a unique trajectory into the Adjacent Possible. These physical facts are utterly requisite to descent with heritable variation and natural selection. But these aspects are simply assumed, without deeper analysis, as available to evolution. Life would have a hard time evolving at the level of complexity of quarks, gluons, and atoms.

The diversity is insufficient at least. While we do not now know the implications of the broadened view of a general biology and the evolution of biospheres in a general biology, we suspect that these issues are worth careful consideration. We will make or find life anew in the next century almost certainly. Adaptation, preadaptation, the relation between the specific physical basis of each form of life and the capacity for heritable variations will become the subjects of intense study. And meanwhile, the possibility of general laws remains open to investigation. For example, it has long been hypothesized that cells are dynamically critical, poised between order and chaos. Recent evidence begins to support this possibility (Ramo et al. 2006; Serra et al. 2004; Shmulevich et al. 2005; M. Nykter et al. in preparation). Since critical networks are rare in the space of dynamical systems, if cells are critical it is precisely a marriage of self organization affording such critical behavior, and the selective usefulness of criticality, that would account for the putative results noted above. Perhaps molecular autonomous agents in any biosphere are dynamically critical. Perhaps the hinted fourth law of thermodynamics discussed in Investigations is true of all biospheres. We simply do not know. But that does not imply that we should not search for such laws—laws that are emergent with respect to physics and part of the emergent, endlessly ‘‘creative’’ universe in which we appear to live.

We would end by inviting philosophers of biology, physics, and others, to help think through the potential implications of a new scientific world view that goes beyond the reductionism of the past three and a half centuries to emergence and a creative evolution in biology and the human economic and cultural realms that cannot be prestated. We believe that such a change in scientific worldview, if merited, will bring with it large societal changes.

Summary

We have traveled along a new path in which we have discussed Darwinian adaptations and the non-reducibility of biology to physics, the mysterious Darwinian preadaptations which seem to preclude finite prestatement and lead to evolution where the state space cannot be prestated. This brings us to serious doubts about whether Shannon information directly applies to the evolution of the biosphere, and leads to Schrödinger’s aperiodic crystal and the hypothesis that information is constraints and boundary conditions, to semiotic information and records, and to the realization that, in the biosphere, it is heritable variation and natural selection that build the intricate web of propagating organization. This provides the basis for considering a new union of matter, energy, information-constraint, and work in cells. This leads to questions about the abiotic universe, where information as boundary conditions affords a simple means to unite matter energy and information.

We have been led to doubt that Shannon information is physically instantiated, whereas the evolving universe and biosphere are.

We seek a new theory of propagating organization, the unfolding of Kant’s statement at the outset of this article. We further seek a theory of the diversifying sources of free energy and constraints that are used to couple spontaneous and non-spontaneous processes into an ever expanding diversity of processes in the biosphere and universe. We do not believe our analysis is fully adequate, but believe it is a start.

References

Apter MJ, Wolpert L (1965) Cybernetics and development. I. Information theory. J Theor Biol 8:244–257
Hayle K (1999) The condition of virtuality. In: Peter L (ed) The digital dialectic. MIT Press, Cambridge, MA
Hood L, Galas D (2003) The digital code of DNA. Nature 421:444–448
Kauffman S (1993) The origins of order: self-organization and selection in evolution. Oxford University Press, New York
Kauffman S (2000) Investigations. Oxford University Press, Oxford
Kauffman S (2006) The third culture beyond reductionism: reinventing the sacred. Edge.org, November 20, 2006
Kauffman S, Clayton P (2006) On emergence, agency, and organization. Biol Philos 21:501–521
Lee DH, Severin K, Reza Ghadiri M (1997a) Autocatalytic networks: the transition from molecular self-replication to molecular ecosystems. Curr Opin Chem Biol 1:491–496
Lee DH, Severin K, Yokobayashi Y, Reza Ghadiri M (1997b) Emergence of symbiosis in peptide self-replication through a hypercyclic network. Nature 390:591–594
Logan RK (2006) The extended mind model of the origin of language and culture. In Nathalie G, Jean P, Van B, Diederik A (eds) Evolutionary epistemology, language and culture. Springer, Dordrecht
Logan RK (2007) The extended mind: the origin of language and culture. University of Toronto Press, Toronto
Mavelli F, Luisi PL (1996) Autopoietic self-reproducing vesicles: a simplified kinetic model. J Phys Chem 100:16600–16607
Maynard SJ (2000a) The concept of information in biology. Philos Sci 67:177–194
Maynard SJ (2000b) Reply to commentaries. Philos Sci 67:214–218
Monod J (1971) Chance and necessity. Knopf, New York
Ramo P, Kessel J, Yli O (2006) Perturbation avalanches and criticality in gene regulatory networks. J Theor Biol 242(1):164–170
Schrödinger E (1992) What is life? Cambridge University Press, Cambridge
Serra R, Kessell J, Semeria A (2004) Genetic network models and statistical properties of gene expression data in knock-out experiments. J Theor Biol 227(1):149–157
Shannon CE (1948) A mathematical theory of communication. Bell Syst Technical J 27:379–423, 623–656
Shiryayev AN (ed) (1993) Selected works of A.N. Kolmogorov, vol III, Information theory and the theory of algorithms (mathematics and its applications). Kluwer Academic Publishing, New York
Shmulevich I, Kauffman SA, Aldana M (2005) Eukaryotic cells are dynamically ordered or critical but not chaotic. PNAS 102: 13439–13444
Sievers D, von Kiedrowski G (1994) Self-replication of complementary nucleotide-based oligomers. Nature 369:221–224
Silberstein M (2003) Reduction, emergence and explanation. In: Peter M, Michael S (eds) The Blackwell guide of the philosophy of science. Blackwell Publishing, Williston, VT
Smith E, Harold M (2004) Universality in intermediary metabolism. Santa Fe Institute Working Paper. http://www.santafe.edu/research/publications/workingpapers/04-07-024.pdf
Sterelny K, Griffiths PE (1999) Sex and death. University of Chicago, Chicago