COMPLEXITY, SELF-ORGANIZATION, AND EMERGENCE AT THE EDGE OF CHAOS IN LIFE ORIGIN MODELS

David L. Abel[*]

The Origin of Life Foundation, Inc.

Abstract

"Complexity," "self-organization," and "emergence" are terms used extensively in life-origin literature. Yet precise and quantitative definitions of these terms are sorely lacking. "Emergence at the edge of chaos" invites vivid imagination of spontaneous creativity. Unfortunately, the phrase lacks scientific substance and explanatory mechanism. We explore the meaning, role, and relationship of complexity at the edge of chaos along with self-organization. We examine their relevance to life-origin processes. The high degree of order and pattern found in "necessity" (the regularities of nature described by the "laws" of physics) greatly reduce the uncertainty and information retaining potential of spontaneously-ordered physical matrices. No as-of-yet undiscovered law, therefore, will be able to explain the high information content of even the simplest prescriptive genome. Maximum complexity corresponds to randomness when defined from a Kolmogorov perspective. No empirical evidence exists of randomness (maximum complexity) generating a halting computational program. Neither order nor complexity is the key to function. Complexity demonstrates no ability to compute. Genetic cybernetics inspired Turing's, von Neumann's, and Wiener's development of computer science. Genetic cybernetics cannot be explained by the chance and necessity of physicodynamics. Genetic algorithmic control is fundamentally formal, not physical. But like other expressions of formality, it can be instantiated into a physical matrix of retention and channel transmission using dynamically-inert configurable switches. Neither parsimonious law nor complexity can program the efficacious decision-node logic-gate settings of algorithmic organization observed in all known living organisms.

[*] Dr. David L. Abel is a theoretical biologist focusing on primordial biocybernetics. He is the Program Director of The Gene Emergence Project, an international consortium of scientists pursuing the natural-process derivation of initial biocybernetic/biosemiotic programming and control.

By what natural process did inanimate nature generate:

1. A genetic representational sign/symbol/token system?
2. Decision nodes and logic gates?
3. Dynamically-inert (dynamically incoherent) (Rocha, 2001) configurable switch settings that instantiate functional "choices" into physicality?
4. A formal operating system, software, and the hardware on which to run it?
5. An abstract encryption/decryption system jointly intelligible to both source and destination?
6. Many-to-one Hamming "block codes" (triplet-nucleotide codons prescribing each single amino acid) used to reduce the noise pollution of genetic messages?
7. The ability to achieve computational halting in the form of homeostatic metabolism?

The heuristic/operational value of using computational and linguistic analogies to describe genetic programming is widely accepted. It is nevertheless common to dismiss many of the above-listed parallels with cybernetics as being merely metaphor. Multiple investigators have taken a close look at possible limits to this metaphor (Emmeche and Hoffmeyer, 1991, Fiumara, 1995, Konopka, 2002, Lackoff and Johnson, 1980, Lackoff, 1993, Rosen, 1993, Sarkar, 2003, Torgny, 1997). Others have discussed whether semantic information about phenotypic traits actually exists (Atlan and Koppel, 1990, Godfrey-Smith, 2003, Griffiths, 2001, Maynard Smith, 2000, Moss, 2003, Stegmann, 2005, Sterelny *et al.*, 1996, Wheeler, 2003). Lwoff warned against taking the genetic information and linguistic metaphors too far (Lwoff, 1962). Some claim that the metaphor is misleading (Godfrey-Smith, 2003, Griffiths, 2001, Kauffman, 1993, Kay, 2000, Keller, 2000, Noble, 2002, Stent, 1981). Rocha (Rocha, 2001, Rocha and Hordijk, 2005) fully appreciates Pattee's epistemic cut (Pattee, 1995b) and the need for semantic closure (Pattee, 1995a), but seeks to explain formal self-organization and sign systems physicodynamically. Others view genetic information as real (Jacob, 1974), (Alberts *et al*, 2002), (Davidson *et al.*, 2002), (Wolpert, 2002), (Stegmann, 2005), (Barbieri, 2004) and (Abel, 2002, Abel and Trevors, 2005, 2006a,b, 2007, Abel, 2009).

None of these bioinformation literalists, however, views genetic information as being "everything." Anti-informationists (e.g.,

None of these bioinformation literalists, however, views genetic information as being "everything." Anti-informationists (e.g., infodynamicists) often create this straw-man argument as justification for denial that any bioinformation exists. Such factors as non-genetic inheritance of cytoplasm and membrane, the role of environment in gene expression, epigenetic factors, prions, post-transcriptional and post-translational editing, do not undo the reality of objective prescriptive information instantiated into linear digital genetic code. They only compound the sophistication of life's control mechanisms.

The first problem with trying to reduce the cybernetic nature of molecular biology to mere metaphor is that biological programming predates the very existence of metaphors. Molecular biology provided the model for the entire field of cybernetics. Genetic cybernetics inspired Turing's (Turing, 1936), von Neumann's (von Neumann, 1950), and Wiener's (Wiener, 1948) development of computer science. Had it not been for their observation of linear digital genetic control, computers might never have been invented. The argument is therefore untenable, if not amusing, that computer science generated only an analogy applied to molecular biology in the minds of humans. If anything, computer science is analogous to the formal logic of a molecular biology that not only preceded, but produced *Homo sapiens* brains and minds.

What exactly is Complexity?

Use of the term "complexity" is extensive in life-origin scientific literature. Unfortunately, complexity is a garbage-can catch-all term we use to explain everything we don't understand and can't reduce. Surprisingly, an unequivocal, pristine, mathematical definition of "complexity" does exist in scientific literature (Kolmogorov, 1965, Li and Vitanyi, 1997). We achieve quantification of complexity through measuring algorithmic compressibility. When a sequence cannot be compressed, it is maximally complex. Random sequences are maximally complex. Maximum complexity cannot be compressed because it lacks patterns and order (Chaitin, 1990, 2001).

Charles Bennett's "logical depth" is also worthy of mention here (Bennett, 1989). Logical depth measures the time required for computational halting. But logical depth presupposes many computer science design concepts not relevant to prebiotic molecular evolution questions. We will not be able to elucidate the derivation through natural process of initial genetic algorithmic control through a discussion of logical depth.

The paradox of Kolmogorov/Solomonoff/Chaitin/Yockey algorithmic information theory is that orderliness lies at the opposite end of the complexity scale from uncertainty and potential information. Even more paradoxical is that randomness (maximum complexity) contains the maximum number of bits of non-compressible "information." The reason this seems so confusing is that Shannon equations do not really quantify "information." They quantify uncertainty and reduced uncertainty (before and after acquired knowledge). The real purpose of Shannon theory is to *compare* sequences: the one sent by the transmitter vs. the one received at the receiver. It's also important for us to remember that Shannon quantifications have nothing to do with meaning or function

(Shannon, 1948). Referring to Shannon quantifications using the term "information" leads to much confusion. It displeased Shannon himself. Shannon opposed calling his theory of communication engineering, "information theory" (Shannon, 1951).

As the probability of an event approaches 1.0, its order increases, and its Shannon uncertainty approaches 0 bits. A law of physics is a compression algorithm for reams of data. At first glance, the data seem almost random. The discovery of a law of physics corresponds to the discovery of order and patterns of relationship hidden in that data.

High probability is high order. A polyadenosine theoretically has maximum order, no uncertainty, and therefore no complexity. Uncertainty and Shannon Information are inversely related to order. The reason laws are so parsimonious is that they describe a highly patterned, highly ordered dynamic. Because law-like behavior is so regular, very little information is required to describe the order of inanimate nature. Very little information can potentially be retained in any structure produced by natural force relationships.

The relationship between order and complexity

Hubert Yockey has graphically clarified the relationship between order and complexity (Yockey, 2002).   The inverse relationship between order and complexity is demonstrated on a linear vector progression from high order on the left toward greater complexity on the right (Figure 1).

The arrow point represents theoretical absolute randomness.   How can "maximum complexity" possibly equal "randomness"? The answer is that randomness cannot be algorithmically compressed to any degree. It is therefore maximally complex. Maximum complexity is a low-end probability bound approaching 0. The probability of 0 is a wall rather than an edge.   No probability can go below 0. No event of probability lower than 0 interfaces with events with 0 probability.

The relationship between order and Kolmogorov algorithmic complexity is shown graphically in Figure 2. By adding a second dimension (Axis Y1) to the uni-dimensional linear vector graph of Yockey, we can visualize the high degree of compressibility for a highly ordered sequence like polyadenosine. Note the low degree of compressibility for a random sequence. Ordered Sequence Complexity (OSC) is on the left. Maximum order means maximum compressibility.   Random Sequence Complexity (RSC) is on the right. Random sequences have no compressibility. No compressibility is maximum complexity. The more highly ordered (patterned) a sequence, the more highly compressible that sequence becomes. The less compressible a sequence, the more complex is that sequence. A random sequence manifests no Kolmogorov compressibility. This reality serves as the very definition of a random, highly complex string. Algorithmic compressibility provides a reliable mathematical definition of

"complexity." The shortest statement of a random sequence is the enumeration of every character of the sequence (Chaitin, 1990, 2001, 2002).

Complexity cannot compute

Although a robust mathematical definition of complexity exists, much to our chagrin, complexity has absolutely nothing to do with function. Yet more often than not, we appeal to complexity as an explanation of computational life-origin processes. Algorithmic function, primordial biocybernetics, and initial organization are what we are hoping to explain. Mere complexity provides no mechanism for any of these three.

Computation is formal, not physical. Both computation and any form of algorithmic optimization require efficacious decision-node selections. When these selections are made randomly, computational halting has never been observed to arise. Sophisticated algorithmic optimization has never been achieved by chance. Function must be "selected for" at the logic-gate programming level prior to the realization of that function. Selection of fittest function is always *after the fact* of any computational success. This is called *The GS Principle* (Genetic Selection Principle) (Abel and Trevors, 2005, 2006a, b, 2007, Abel, 2009). Natural selection favors only the fittest already-computed phenotypes. Yet selection must occur at the logic-gate level of genetic programming. Configurable switches are "set in stone" with rigid covalent bonds before folding begins.
Three-dimensional conformation of molecular machines is largely *determined* by the minimum-free-energy sinks of *primary structure* folding. The primary structure of any protein or sRNA is the already-covalently-bound *sequence* of particular monomers that serve as configurable switch-settings.

Maximum complexity is randomness because randomness offers the highest degree of combinatorial uncertainty. But randomness is the equivalent of pure noise. Noise has never been observed to program any algorithm. Adding long periods of time provides no mechanism of selection at the decision node level where programming is accomplished. Although a random sequence could happen to match a program sequence, outside of a specifically chosen operational context and set of rules, such a matching sequence would remain random and nonfunctional. Thus not only would the random sequence itself have to match the program sequence, but the operating system at both ends of the channel would also have to match by chance in order for function to arise.

Order cannot compute

Much life-origin literature appeals to "yet-to-be discovered laws of self-organization". Laws, however, describe highly ordered/patterned behavior. Because they are parsimonious compression algorithms of data, they contain very little information. Given the high information content of life, expecting a new law to explain sophisticated genetic algorithmic programming is ill-founded. Considerable peer-reviewed published literature is erroneous because of failure to appreciate that the 'complexity of life" could never arise from such highly "ordered," low informational

physicodynamic patterning. Tremendous combinatorial uncertainty is required. The complexity of life will never be explained by the highly ordered behavior that is reducible to the low-informational laws of physics and chemistry.

A crystal is highly ordered. Its description can be easily algorithmically compressed. A crystal is about as far from being "alive" as any physical state we could suggest. Every member of a 200-monomer string of adenosines can be specifically enumerated by stating two short clauses. "Give me an adenosine. Repeat 200 times." This is called a compression algorithm. The simplicity and shortness of this compression algorithm is a measure of the extremely low complexity of this polymer. Such a parsimonious statement of the full sequence is only possible because that sequence is so highly patterned. Such a highly ordered sequence lacks uncertainty, complexity, and the ability to instantiate prescriptive information. Such a parsimonious compression algorithm can enumerate each and every member of the 200-mer string with only seven words. This reality defines high order or pattern along with low information retaining potential.

We value Ocham's razor in laws because we realize that physicality is so highly ordered. We consider a law to be elegant and beautiful because of its ability to compress reams of data down to one little parsimonious equation. When we look for new laws of physics, we look for new compression algorithms for reams of data.

When we come to biology, however, we encounter not only the highest degree of complexity known, we encounter linear, digital, cybernetic, prescriptive information of the most sophisticated, abstract, and conceptual nature. The world's finest main frame parallel computer system cannot hold a candle to the central nervous system of any mammal.

If all four RNA bases were equally available in a theoretical primordial soup, each nucleotide selection would represent 2 bits of Shannon uncertainty. If, on the other hand, some bases were more available than others in primordial soup, the uncertainty of each nucleotide selection drops to much less than 2 bits. Unequal availability of bases results in more ordering of the sequence. More ordering = less complexity, and therefore less information retention potential. The particular oligoribonucleotide strand would have mostly one or two bases with less uncertainty, fewer bits, and therefore less complexity than if all four bases were equally available.

All too many life-origin specialists still operate under the mistaken premise that greater complexity contains more order. In reality, order and complexity are antithetical. In addition, neither order nor complexity is the key to function. Neither order nor complexity alone can generate algorithmic organization. Bona fide organization results from algorithmic optimization. The best solutions to any problem must be *selected* to achieve optimization. Apart from selection, noise will increase within any system. A tendency toward randomization and loss of function unfolds from noise. Complexity increases while algorithmic optimization decreases.

This latter point exposes the second common illusion, that increasing complexity produces increasing algorithmic utility. In reality, complexity has nothing to do with integration, organization, or utility. Programming requires formal decision-node choice commitments made with intent. Any attempt to disallow choice or intent from the mix results in the deterioration of programming function, computational halting, integration, and organization.

In addition to showing the Kolmogorov compression in the second dimension (Y axis), Figure 2 also shows the superimposition of a third cybernetic dimension (Z axis), Functional Sequence Complexity (FSC). The Y axis plane plots the decreasing degree of algorithmic compressibility as complexity increases from order towards randomness. The (Z) axis plane shows where along the same complexity gradient (X-axis) that highly instructional sequences and algorithmic programs are generally found.

The Functional Sequence Complexity (FSC) curve includes all algorithmic sequences that work at all (W). The peak of this curve (w*) represents "what works best." The FSC curve is usually quite narrow and is located closer to the random end than to the ordered end of the complexity scale.

The third dimension of utility and organization is when each alphabetical token in the linear string is selected for meaning or function.   The string becomes a cybernetic program capable of computation only when signs/symbols/tokens are chosen to represent utilitarian configurable switch settings. What is the common denominator to all aspects of design and engineering function?   Choice contingency; not chance contingency, not law, not physicodynamics, but *choice* contingency.   The FSC curve is usually quite narrow and is located closer to the random end than to the ordered end of the complexity scale. Compression of an instructive sequence slides the FSC curve towards the right (away from order, towards maximum complexity, maximum Shannon uncertainty, and seeming randomness) *with no loss of function*. This further demonstrates that neither order nor complexity is the determinant of algorithmic function. Functionality arises in *a third dimension of selection* that is unknown to the second dimension of compressibility. This is one of most poorly understood realities in life-origin science. Selection alone produces functionality. Without selection, evolution would be impossible.

Figure 3 is a dendrogram showing all possible sequences (branches or paths) of decision node options. W paths may show some function, but w* represents the best algorithmic path to achieve maximum function. Notice that each path contains equal (N) bits of Shannon uncertainty regardless of whether the path leads to anything useful. The measurement of bits tells us nothing about whether the string *does* anything useful. Only certain strings of specific choice commitments lead to function and organization. Neither $-\log_2 P$ nor the formula for Shannon mutual entropy $[\, I(A:B) = H(x) - H(x\,|\,y)\,]$ measures *prescriptive information* (Abel and Trevors, 2005, 2006a, b, 2007, Abel, 2009, Trevors and Abel, 2004). Prescriptive information either instructs or directly produces sophisticated algorithmic utility. In addition, no reason exists to think that maximum complexity (randomness; noise; maximum uncertainty; maximum bits) has any

functional capability in and of itself.   If anything, we expect no function at all out of maximum complexity.

All known life is cybernetic

Any one of four different nucleotides can be added next to a forming nucleic acid strand in aqueous solution. No physicochemical bias exists (Judson, 1993, Monod, 1972, Polanyi, 1968) for which nucleotide polymerizes apart from base-pairing of an already existing strand, or clay-surface templating. The latter tends to produce polyadenosines, a non-informational sequence because of its extremely high order and extremely low uncertainty. Physicodynamics, therefore, does not explain functional *sequencing*. The effort that has been invested into genome projects affirms the prescriptive nature of nucleotide sequencing.   While not everything, no one can deny that amino acid sequencing is determined by triplet codon sequencing.

Every nucleotide added to an oligoribonucleotide in the pre RNA World represents the specific setting of an additional discrete 4-way configurable switch. The appropriate setting of a string of programmable switches alone accounts for computational success. Computational "halting" in the pre RNA World is defined in terms of catalytic binding success of three dimensional small RNA's. But binding success depends upon secondary and tertiary structure. Secondary and tertiary structure in turn depends upon the thermodynamic minimum-free-energy folding sinks of each primary structure (Rhoades, et al., 2003). Primary structure is the *sequence* of nucleotides. This linear digital sequence of nucleotides is held together by rigid covalent bonds.   Covalent bonds are "written in stone" compared to the weak hydrogen bonds, van der Waals forces, electrostatic attractions and repulsions, and hydrophobicities that contribute to secondary folding.

Atlan et al attempt to elucidate a mechanism for self-classification and self-organization in automata networks (Atlan, et al., 1986). They also explore the notion of self-creation of meaning (Atlan, 1987). Finally they suggest that DNA is data rather than program (Atlan and Koppel, 1990). As with Shannon (Shannon, 1948), Kolmogorov(Kolmogorov, 1965), Chaitin (Chaitin, 1987), and Yockey (Yockey, 2005), Atlan et al's concept of information fails to measure up to what we actually observe in molecular cybernetics.   The reason is a failure to acknowledge and incorporate the literal instructive and controlling role of genetic information. Linear digital genetic information specifically *prescribes* functional sRNA and protein sequences. Post transcriptional and post translational editing does not undo this reality. They only add to the sophistication of the entire system. No progress will be made in quantifying semantic information until we pursue the unique properties of what Abel has termed *prescriptive information* (Abel and Trevors, 2006a, b, 2007, Abel, 2009) Abel, 2005 #5001; Trevors, 2004 #5544}. Prescriptive information is more than just semantic.   It is cybernetic. Prescriptive information alone generates computational halting in the form of homeostatic metabolism. No theory of combinatorial probabilism or compression can explain or measure *computational success*. Charles Bennett's logical depth (Bennett, 1988) comes

the closest, but presupposes human-designed computer science in a fashion inappropriate for prebiotic molecular evolution theory.

The key to everything that Turing, von Neumann, and Weiner did in inventing computers was to recognize that life is made possible because molecular biology uses *dynamically incoherent, dynamically inert, freely-configurable switches* (Rocha, 2000). Any of the four ribonucleotides can polymerize next in aqueous solution. This fact is what makes information recordation into the physical matrix of nucleic acid possible. If selection of the next nucleotide were determined by physicodynamic factors, the sequence would be too highly ordered and redundant for "messenger molecules" to be possible. Using only four alphabetical characters (four different nucleotides), any instructions can be written into DNA. What makes programming possible is that the switch is designed to be freely "configurable." Any of the four letters can be chosen without physicochemical prejudice. This means that no law determines which way the four-way switch knob is pushed.

In computer science, only the programmer's mind determines which way the switch knob is pushed. In evolution science we say that environmental selection "favors" the fittest small groups. But selection is still the key factor, not chance and necessity. If physicodynamics set the switches, the switches would either be set randomly by heat agitation, or they would be set by force relationships and constants. Neither chance nor necessity, nor any combination of the two, can program. Chance produces only noise and junk code. Law would set all of the switches the same way. Configurable switches must be set using "choice with intent" if "computational halting" is expected.

Nucleic acid can spontaneously form without purpose, such as a polyadenosine forming (by physicochemical law) on a montmorillonite clay template surface. But the latter is a classic example of all the switches being set the same when law is involved. A polyadenosine is nucleic acid, but it can't program anything. It can't relay any information, because all of the four-way switches have been set the same way (all adenosines) "by law." What so many fail to realize is that RNA and DNA are nothing but ordinary physical molecules that have *the potential* of being used for information retention only through selection of each nucleotide. It is the sequencing of particular nitrogen base selections that accounts for any information retention in a nucleic acid molecule, not the largely inert DNA itself. Prescriptive information is not physicodynamic. It is formal, though it can be instantiated into a physical medium using dynamically inert configurable switches.

RNA (oligoribonucleotides up to eight monomers, at least) can form spontaneously in aqueous solution. But such strings are "stochastic ensembles" (random strings of nucleotides). As such, they too contain no prescriptive information. They are like linguistic gibberish. They are "garbage in, garbage out" computer code, pure "software bugs."

A "cybernetic program" presupposes a cybernetic context in which it operates. One has to have an operating system of "rules" before one can have an application

software. And of course one must have a hardware system too. All of these components only come into existence through "choice contingency," not through "chance contingency" or law. One of many problems with metaphysical materialism is that it acknowledges only two subsets of reality: chance and necessity. Neither can write operating system rules or application software. Neither can generate hardware or any other kind of sophisticated machinery, including molecular machines (the most sophisticated machinery known).

We see in Figure 2 that complexity as mathematically and scientifically defined is blind to function. Mere complexity cannot generate algorithmic optimization. *Selection for fitness* is required. Complexity cannot do this. Complexity knows nothing of selection, fitness, or meaning. Without selection, evolution is impossible.

The Edge of Chaos

Physical events "at the edge of chaos" have never been observed to select for fitness or binding success. No mechanism has been demonstrated empirically whereby physicodynamics spontaneously generates sophisticated algorithmic optimization or bona fide organization.   Switches must be set a certain way to achieve integrated circuits. Order can spontaneously emerge from chaos. But if chaos sets configurable switches, the result will predictably "blue screen." Without steering towards sophisticated function at each decision node, sophisticated function has never been observed to arise spontaneously. Only disorganization accumulates. No prediction fulfillments have been realized of cooperative integration of biofunction arising spontaneously in nature.

"Emergence at the edge of chaos" is poetic, if not mesmerizing. The phrase invites vivid imagination of mystical powers and ingenious spontaneous creativity. Unfortunately, this notion has not provided detailed scientific mechanism to explain the efficacious selection of pragmatic configurable switch-settings. Organization requires algorithmic optimization. The latter requires expedient decision-node commitments that are instantiated into specific physical configurable switch-settings. To explain life origin requires elucidating how these particular logic gates were selected at the genetic level. Phenotypes must first be computed before the fittest phenotype can be selected. No plausible theoretical mechanism and no empirical evidence for emergence exist in the literature.   No prediction fulfillment of spontaneous emergence exists. In every case that provides the illusion of spontaneous emergence, investigator involvement can be demonstrated in the Materials and Methods section of so-called "evolutionary algorithm" papers. The experimenter's goal and steering are apparent in faulty experimental designs. This is usually evident in the choice of each successive iteration to pursue. Real evolution has no goal. Iterations cannot be steered toward experimenters' goals (e.g., a desired ribozyme using SELEX (Ellington and Szostak, 1990, Robertson and Joyce, 1990, Tuerk and Gold, 1990)). Quality science requires brutal self-honesty. We must be open-minded enough to consider the possibility that emergence and self-organization are closer to metaphysical presuppositions than observed scientific facts.
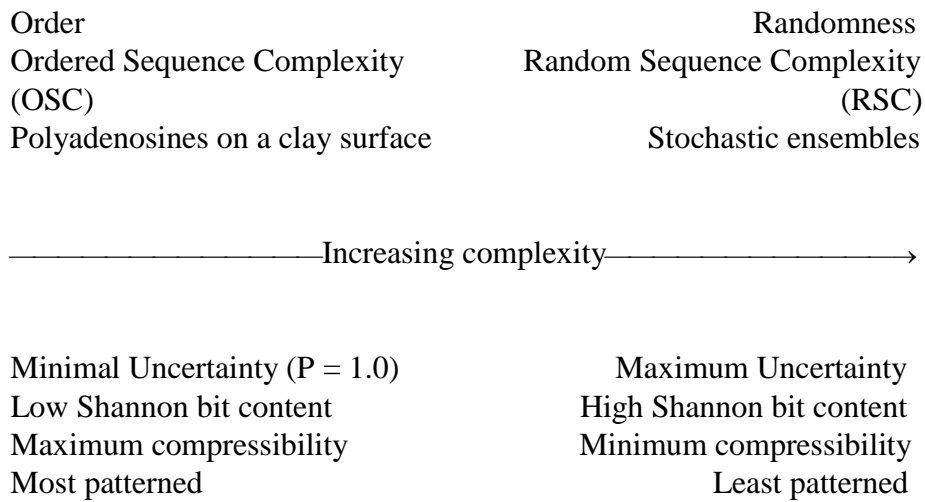
EndNotes:

Abel, D.L., 2002, Is life reducible to complexity? In Palyi, G., Zucchi, C. and Caglioti, L., *Fundamentals of life.* Elsevier, Paris.

Abel, D.L. and Trevors, J.T., 2005, Three subsets of sequence complexity and their relevance to biopolymeric information. Theoretical Biology and Medical Modeling **2**: 29, open access at http://www.tbiomed.com/content/2/1/29

Abel, D.L. and Trevors, J.T., 2006a, More than metaphor: Genomes are objective sign systems. Journal of BioSemiotics **1**: 253-267

Abel, D.L. and Trevors, J.T., 2006b, Self-organization vs. Self-ordering events in life-origin models. Physics of Life Reviews **3**: 211-228

Abel, D.L. and Trevors, J.T., 2007, More than metaphor: Genomes are objective sign systems. In Barbieri, M., *Biosemiotic research trends.* Nova Science Publishers, Inc., New York.

Abel, D.L., 2009, The biosemiosis of prescriptive information. Semiotica In Press

Alberts, B. and et al, 2002, *Molecular biology of the cell*. Garland Science, New York.

Atlan, H., E., B.-E., Fogelman-Soulie, F., Pellegrin, D. and Weisbuch, G., 1986, Emergence of classification procedures in automata networks as a model for functional self-organization. J.Theor. Biol. **120**: 371-380

Atlan, H., 1987, Self-creation of meaning. Physica Scripta **36**: 563-576

Atlan, H. and Koppel, M., 1990, The cellular computer DNA: Program or data. Bulletin of Mathematical Biology **52**: 335-348

Barbieri, M., 2004, Biology with information and meaning. History & Philosophy of the Life Sciences **25**: 243-254

Bennett, C.H., 1989, On the logical "Depth" Of sequences and their reducibilities to incompressible sequences

Bennett, D.H., 1988, Logical depth and physical complexity. In Herken, R., *The universal turing machine: A half-century survey.* Oxford University Pres, Oxford.

Chaitin, G.J., 1987, *Information, randomness & incompleteness : Papers on algorithmic information theory*. World Scientific, Singapore ; Teaneck, NJ, USA.

Chaitin, G.J., 1990, *Information, randomness and incompleteness : Papers on algorithmic information theory*. World Scientific, Singapore ; New Jersey.

Chaitin, G.J., 2001, *Exploring randomness*. Springer, London ; New York.

Chaitin, G.J., 2002, *Conversations with a mathematician : Math, art, science, and the limits of reason : A collection of his most wide-ranging and non-technical lectures and interviews*. Springer, London ; New York.

Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otim, O., Brown, C.T., Livi, C.B., Lee, P.Y., Revilla, R., Rust, A.G., Pan, Z., Schilstra, M.J., Clarke, P.J., Arnone, M.I., Rowen, L., Cameron, R.A., McClay, D.R., Hood, L. and Bolouri, H., 2002, A genomic regulatory network for development. Science **295**: 1669-78

Ellington, A.D. and Szostak, J.W., 1990, In vitro selection of rna molecules that bind specific ligands. Nature **346**: 818-822.

Emmeche, C. and Hoffmeyer, J., 1991, From language to nature: The semiotic metaphor in biology. Semiotica **84**: 1-42

Fiumara, G.C., 1995, *The metaphoric process: Connections between language and life*. Rutledge, London.

Godfrey-Smith, P., 2003, Genes do not encode information for phenotypic traits. In Hitchcock, C., *Contemporary debates in philosophy of science*   Blackwell, London.

Griffiths, P.E., 2001, Genetic information: A metaphor in search of a theory. Philosophy of Science **68**: 394-412

Jacob, Francois, 1974, *The logic of living systems---a history of heredity*. Allen Lane, London.

Judson, H.F., 1993, Frederick sanger, erwin chargaff, and the metamorphosis of specifity. Gene **135**: 19-23

Kauffman, S.A., 1993, *The origins of order: Self-organization and selection in evolution*. Oxford University Press, Oxford.

Kay, L., 2000, *Who wrote the book of life? A history of the genetic code*. Stanford University Press, Stanford, CA.

Keller, E.F., 2000, Decoding the genetic program. In Beurton, P., Falk, R. and Rheinberger, H.-J., *The concept of the gene in development and evolution.* Cambridge University Press, Cambridge.

Kolmogorov, A.N., 1965, Three approaches to the quantitative definition of the concept "Quantity of information". Problems Inform. Transmission **1**: 1-7

Konopka, A.K., 2002, Grand metaphors of biology in the genome era. Computers & Chemistry **26**: 397-401

Lackoff, G. and Johnson, M., 1980, *Metaphors we live by*. University of Chicago Press, Chicago, IL.

Lackoff, G., 1993, The contemporary theory of metaphor. In Ortony, A., *Mataphor and thought, 2nd edition.* Cambridge University Press, Cambridge.

Li, M. and Vitanyi, P., 1997, *An introduction to kolmogorov complexity and its applications*. Springer-Verlag, New York.

Lwoff, A., 1962, *Biological order*. MIT Press, Cambridge, MA.

Maynard Smith, J., 2000, The concept of information in biology. Philosophy of Science **67**: 177-194 (entire issue is an excellent discussion)

Monod, J., 1972, *Chance and necessity*. Knopf, New York.

Moss, L., 2003, *What genes can't do*. MIT Press, Cambridge, MA.

Noble, D., 2002, Modeling the heart--from genes to cells to the whole organ. Science **295**: 1678-82

Pattee, H.H., 1995a, Evolving self-reference: Matter, symbols, and semantic closure. Communication and Cognition-Artificial Intelligence **12**: 9-28

Pattee, H.H., 1995b, Artificial life needs a real epistemology. 3rd European Conference on Artificial Life: Session 1: Foundations and Epistemology, Granada, Spain.

Polanyi, M., 1968, Life's irreducible structure. Science **160**: 1308-12

Rhoades, E., Gussakovsky, E. and Haran, G., 2003, Watching proteins fold one molecule at a time. PNAS **100**: 3197-3202

Robertson, D.L. and Joyce, G.F., 1990, Selection in virtro of an rna enzyme that specifically cleaves single-stranded DNA. Nature **344**: 467-468

Rocha, L.M., 2000, Syntactic autonomy: Or why there is no autonomy without symbols and how self-organizing systems might evolve them. In Chandler, J.L.R. and van

deVijver, G., *Closure: Emergent organizations and their dynamics. Annals of the new york academy of sciences.* Vol 901, 207-223.

Rocha, L.M., 2001, Evolution with material symbol systems. Biosystems **60**: 95-121

Rocha, L.M. and Hordijk, W., 2005, Material representations: From the genetic code to the evolution of cellular automata. Artif Life **11**: 189-214

Rosen, R., 1993, Bionics revisited. In Haken, H., Karlqvist, A. and Svedin, U., *The machine as a metaphor and tool.* Springer-Verlag, Berlin.

Sarkar, S., 2003, Genes encode information for phenotypic traits. In Hitchcock, C., *Comtemporary debates in philosophy of science.* Blackwell, London.

Shannon, C., 1948, Part i and ii: A mathematical theory of communication. The Bell System Technical Journal **XXVII**: 379-423

Shannon, E.C., 1951, Intersymbol influence reduces information content. Bell. Sys. Tech. J. **30**: 50

Stegmann, U.E., 2005, Genetic information as instructional content. Phil of Sci **72**: 425-443

Stent, G.S., 1981, Strength and weakness of the genetic approach to the development of the nervous system. Annu Rev Neurosci **4**: 163-94

Sterelny, K., Smith, K. and Dickison, M., 1996, The extended replicator. Biology and Philosophy **11**: 377-403

Torgny, O., 1997, *Metaphor--a working concept. Kth, royal institute of technology.* CID, Stockholm, Sweden.

Trevors, J.T. and Abel, D.L., 2004, Chance and necessity do not explain the origin of life. Cell Biology International **28**: 729-739

Tuerk, C. and Gold, L., 1990, Systematic evolution of ligands by exponential enrichment -- rna ligands to bacteriophage - t4 DNA-polymerase. Science **249**: 505-510

Turing, A.M., 1936, On computable numbers, with an application to the *entscheidungs problem.* Proc. Roy. Soc. London Mathematical Society **42**: 230-265 [correction in 43, 544-546]

Wheeler, M., 2003, Do genes code for traits? In Rojszczak, A., Cachro, J. and Kurczewski, G., *Philosophic dimensions of logic and science: Selected contributed papers from the 11th international congress of logic, methodology, and philosophy of science.* Kluwer, Dordrecht.

Wiener, N., 1948, *Cybernetics.* Wiley, New York.

Wolpert, L., 2002, *Principles of development.* Oxford University Press, Oxford.

Yockey, H.P., 2002, Information theory, evolution and the origin of life. Information Sciences **141**: 219-225

Yockey, H.P., 2005, *Information theory, evolution, and the origin of life.* Cambridge University Press, Cambridge.

Figure 1: The inverse relationship between order and complexity is demonstrated on a linear vector progression from high order on the left toward greater complexity on the right (modified from Hubert Yockey, *Fundamentals of Life.* Edited by Palyi G, Zucchi C, Caglioti L. Paris: Elsevier; 2002: 335-348.) Used with permission from: Abel, David L., and Jack. T. Trevors (2005), "Three subsets of sequence complexity and their relevance to biopolymeric information." *Theoretical Biology and Medical Modeling* 2:29, open access at http://www.tbiomed.com/content/22/21/29.

Order   Randomness
Ordered Sequence Complexity   Random Sequence Complexity
(OSC)   (RSC)
Polyadenosines on a clay surface   Stochastic ensembles

———————————————Increasing complexity——————————————→

Minimal Uncertainty (P = 1.0)   Maximum Uncertainty
Low Shannon bit content   High Shannon bit content
Maximum compressibility   Minimum compressibility
Most patterned   Least patterned

Figure 2:

Superimposition of Kolmogorov compression (2nd dimension: Y1 axis) and Functional Sequence Complexity (FSC) (3rd dimension: Z axis) onto the single dimension of Figure 1's linear vector graph. The Y1 axis plane plots the decreasing degree of algorithmic compressibility as complexity increases from order towards randomness. The Y2 (Z) axis plane shows where along the same order-complexity gradient (X-axis) that highly instructional and prescriptive sequences are generally found. The Functional Sequence Complexity (FSC) curve includes all algorithmic sequences that work at all (W). The peak of this curve (w*) represents "what works *best*." Used with permission from: Abel, David L., and Jack. T. Trevors (2005), "Three subsets of sequence complexity and their relevance to biopolymeric information." *Theoretical Biology and Medical Modeling* 2:29, open access at http://www.tbiomed.com/content/22/21/29.
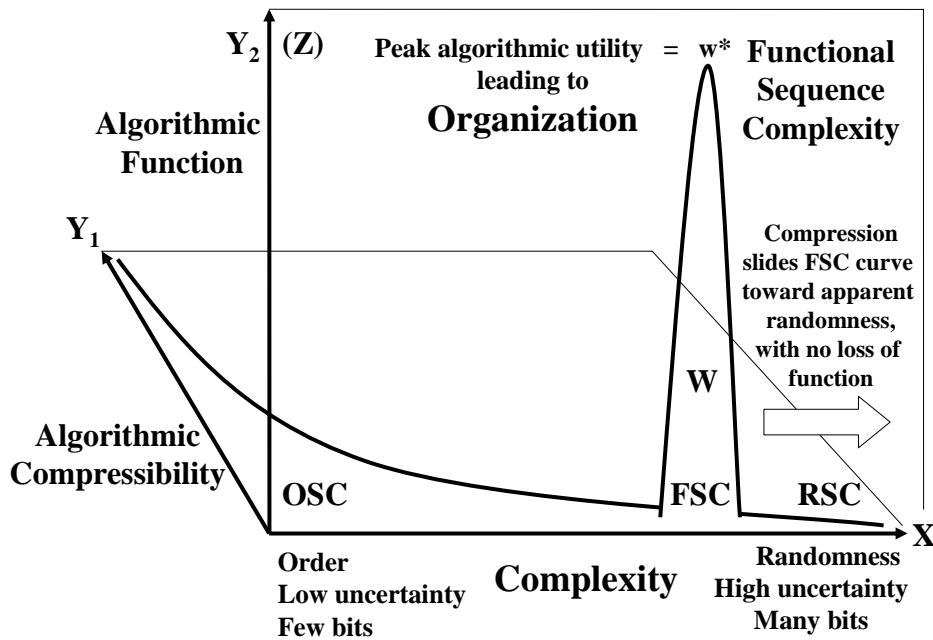
Figure 3: A dendrogram showing all possible sequences (branches or paths) of decision node options. "w*" represents the best algorithmic path to achieve maximum function. "W" represents all paths that produce any degree of algorithmic utility. Notice that *all* paths contain equal (n) bits of Shannon so-called "information" regardless of whether the sequence of specific choice commitments accomplishes anything useful.

One of very few algorithmic processes that function out of 2^n branches

Node:

1st	2^1	1 bit 2 branches
2nd	2^2	2 bits 4 branches
3rd	2^3	3 bits 8 branches
4th	2^4	4 bits 16 branches
nth		n "bits"

w*

2^n - W fail **What "works" best** 2^n branches