Life origin:

the role of

Complexity

The Edge of Chaos

David L. Abel, Program Director

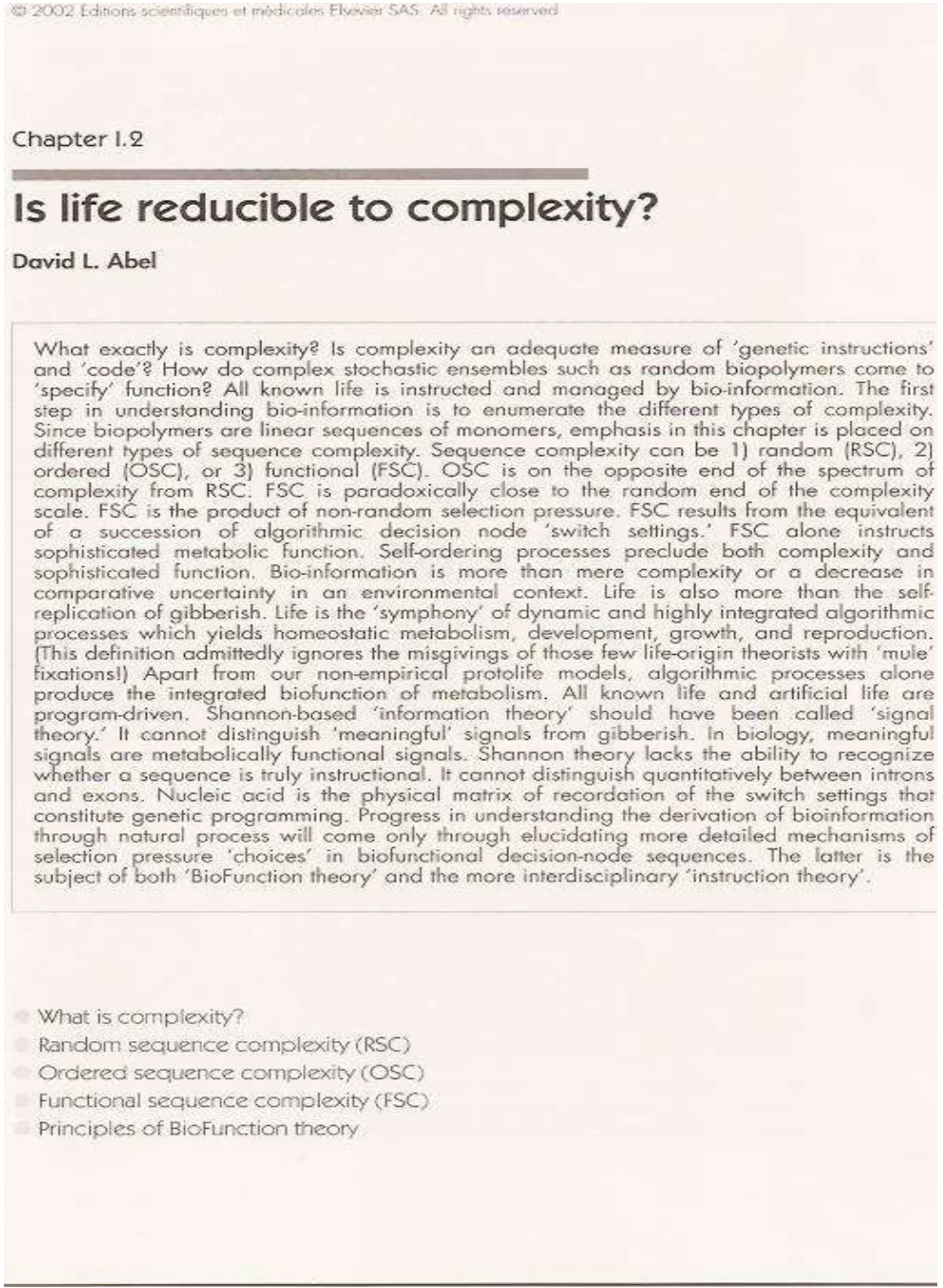
The Gene Emergence Project

The Origin of Life Foundation, Inc.

(501-c-3 U.S. science foundation)

The history of the Gene Emergence Project has always been the pursuit of a natural-process mechanism for the derivation of initial linear digital genetic programming.

The purpose of The Origin of Life Prize is to stimulate research in the area of gene emergence



Chapter I.2

Is life reducible to complexity?

David L. Abel

What exactly is complexity? Is complexity an adequate measure of 'genetic instructions' and 'code'? How do complex stochastic ensembles such as random biopolymers come to 'specify' function? All known life is instructed and managed by bio-information. The first step in understanding bio-information is to enumerate the different types of complexity. Since biopolymers are linear sequences of monomers, emphasis in this chapter is placed on different types of sequence complexity. Sequence complexity can be 1) random (RSC), 2) ordered (OSC), or 3) functional (FSC). OSC is on the opposite end of the spectrum of complexity from RSC. FSC is paradoxically close to the random end of the complexity scale. FSC is the product of non-random selection pressure. FSC results from the equivalent of a succession of algorithmic decision node 'switch settings.' FSC alone instructs sophisticated metabolic function. Self-ordering processes preclude both complexity and sophisticated function. Bio-information is more than mere complexity or a decrease in comparative uncertainty in an environmental context. Life is also more than the self-replication of gibberish. Life is the 'symphony' of dynamic and highly integrated algorithmic processes which yields homeostatic metabolism, development, growth, and reproduction. (This definition admittedly ignores the misgivings of those few life-origin theorists with 'mule' fixations!) Apart from our non-empirical protolife models, algorithmic processes alone produce the integrated biofunction of metabolism. All known life and artificial life are program-driven. Shannon-based 'information theory' should have been called 'signal theory.' It cannot distinguish 'meaningful' signals from gibberish. In biology, meaningful signals are metabolically functional signals. Shannon theory lacks the ability to recognize whether a sequence is truly instructional. It cannot distinguish quantitatively between introns and exons. Nucleic acid is the physical matrix of recordation of the switch settings that constitute genetic programming. Progress in understanding the derivation of bioinformation through natural process will come only through elucidating more detailed mechanisms of selection pressure 'choices' in biofunctional decision-node sequences. The latter is the subject of both 'BioFunction theory' and the more interdisciplinary 'instruction theory'.

- What is complexity?
- Random sequence complexity (RSC)
- Ordered sequence complexity (OSC)
- Functional sequence complexity (FSC)
- Principles of BioFunction theory

Points of confusion in scientific literature

"Complexity" is a garbage-can catch-all term we use to explain everything we don't understand and can't reduce.

What exactly is "Complexity"”?

An unequivocal, pristine, mathematical definition of

“Complexity”

already exists in scientific literature:

What exactly is "Complexity"?

Maximum complexity is randomness

Maximum complexity
cannot be compressed because
it lacks patterns and order.

Kolmogorov/Solomonoff/Chaitin

What exactly is "Complexity"?

As the probability of an event approaches 1.0,

its Shannon uncertainty approaches 0 bits

What exactly is "Complexity"?

High probability is high order

A polyadenosine has maximum order,
no uncertainty,
and therefore no complexity

What exactly is "Complexity"?

"Give me an adenosine.

Repeat 200 times."

What exactly is "Complexity"?

The compression algorithm for a
polyadenosine
contains very little information

A law of physics
contains very little information because
the data, also, is so highly ordered

Order

Ordered Sequence Complexity (OSC)

Polyadenosines on a clay surface

Randomness

Random Sequence Complexity (RSC)

Stochastic ensembles

Increasing complexity →

Minimal Uncertainty (P = 1.0)

Low Shannon bit content

Maximum compressibility

Most patterned

Maximum Uncertainty

High Shannon bit content

Minimum compressibility

Least patterned

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. Modified from Yockey HP: In *Fundamentals of Life*. Edited by Palyi G, Zucchi C, Caglioti L. Paris: Elsevier; 2002: 335-348

Random Sequence Complexity (RSC)

Ordered Sequence Complexity (OSC)

Functional Sequence Complexity (FSC)

Theoretical Biology and Medical Modelling

BioMed Central

Review

Open Access

Three subsets of sequence complexity and their relevance to biopolymeric information

David L Abel[1] and Jack T Trevors*[2]

Address: [1]Director, The Gene Emergence Project, The Origin-of-Life Foundation, Inc., 113 Hedgewood Dr., Greenbelt, MD 20770-1610 USA and [2]Professor, Department of Environmental Biology, University of Guelph, Rm 3220 Bovey Building, Guelph, Ontario, N1G 2W1, Canada

Email: David L Abel - life@us.net; Jack T Trevors* - jtrevors@uoguelph.ca

* Corresponding author

Published: 11 August 2005

Received: 23 May 2005

Accepted: 11 August 2005

Theoretical Biology and Medical Modelling 2005, **2**:29 doi:10.1186/1742-4682-2-29

This article is available from: http://www.tbiomed.com/content/2/1/29

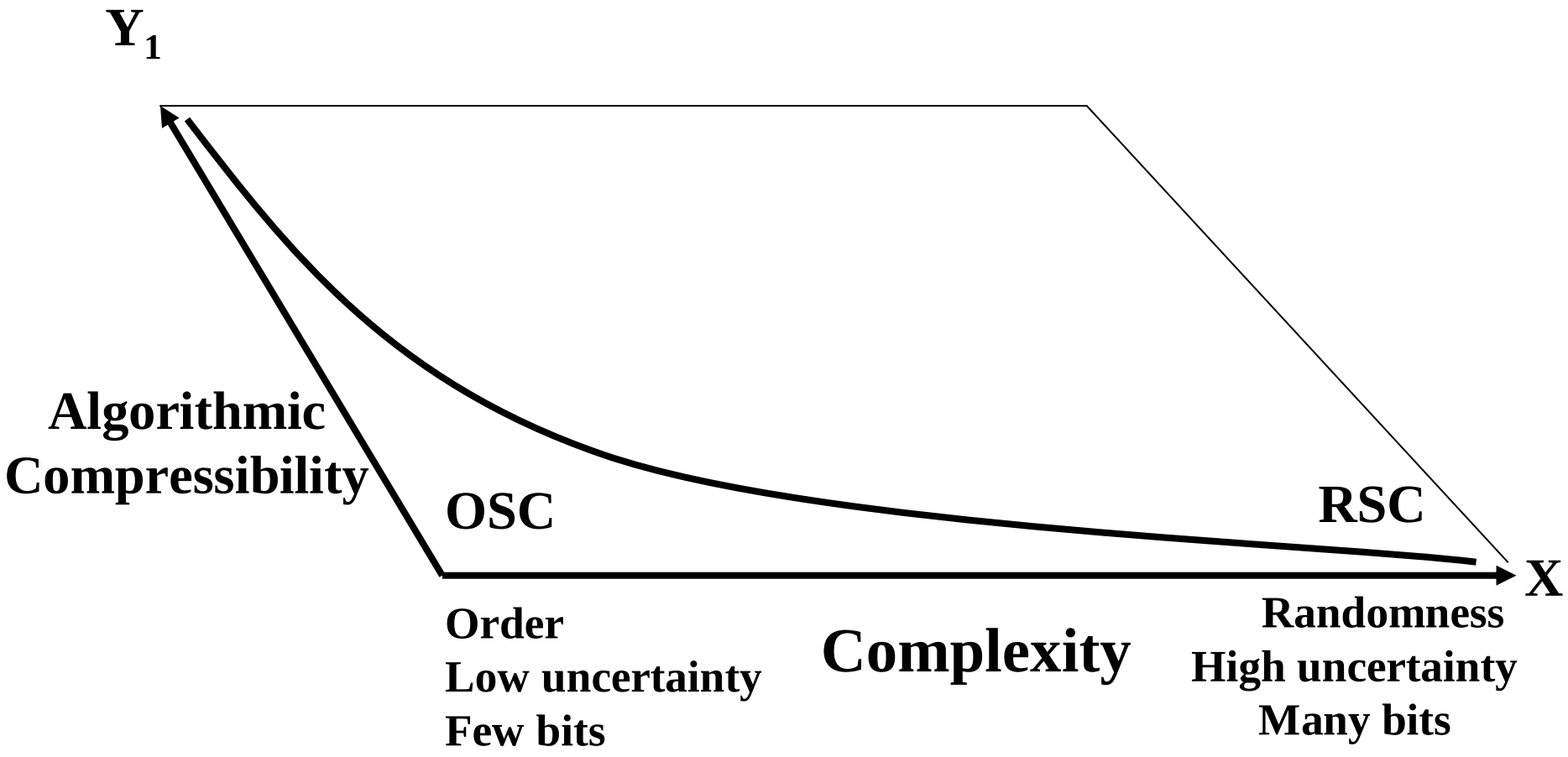
Abstract

Genetic algorithms instruct sophisticated biological organization. Three qualitative kinds of sequence complexity exist: random (RSC), ordered (OSC), and functional (FSC). FSC alone provides algorithmic instruction. Random and Ordered Sequence Complexities lie at opposite ends of the same bi-directional sequence complexity vector. Randomness in sequence space is defined by a lack of Kolmogorov algorithmic compressibility. A sequence is compressible because it contains redundant order and patterns. Law-like cause-and-effect determinism produces highly compressible order. Such forced ordering precludes both information retention and freedom of selection so critical to algorithmic programming and control. Functional Sequence Complexity requires this added programming dimension of uncoerced selection at successive decision nodes in the string. Shannon information theory measures the relative degrees of RSC and OSC. Shannon information theory cannot measure FSC. FSC is invariably associated with all forms of complex biofunction, including biochemical pathways, cycles, positive and negative feedback regulation, and homeostatic metabolism. The algorithmic programming of FSC, not merely its aperiodicity, accounts for biological organization. No empirical evidence exists of either RSC of OSC ever having produced a single instance of sophisticated biological organization. Organization invariably manifests FSC rather than successive random events (RSC) or low-informational self-ordering phenomena (OSC).

Background

"Linear complexity" has received extensive study in many areas relating to Shannon's syntactic transmission theory [1-3]. This theory pertains only to engineering. Linear complexity was further investigated by Kolmogorov, Solomonoff, and Chaitin [4-8]. Compressibility became the measure of linear complexity in this school of thought. Hamming pursued Shannon's goal of noise-pollution reduction in the engineering communication channel through redundancy coding [9].

Little progress has been made, however, in measuring and explaining *intuitive information*. This is especially true regarding the derivation through natural process of


Y_1
Algorithmic
Compressibility
OSC
RSC
X
Order
Low uncertainty
Few bits
Complexity
Randomness
High uncertainty
Many bits

	Probability of Availability				
Adenine	0.25	*	(- log2 0.25)	=	0.500
Uracil	0.25	*	(- log2 0.25)	=	0.500
Guanine	0.25	*	(- log2 0.25)	=	0.500
Cytosine	0.25	*	(- log2 0.25)	=	0.500
	1.00				2.000 bits

Adenine	0.46	*	(- log2 0.46)	=	0.515
Uracil	0.40	*	(- log2 0.40)	=	0.529
Guanine	0.12	*	(- log2 0.12)	=	0.367
Cytosine	0.02	*	(- log2 0.02)	=	0.113
	1.00				1.524 bits

From: Abel, David L. (2002), "Is Life Reducible to Complexity?" in Gyula Palyi, Claudia Zucchi and Luciano Caglioti (eds.), *Fundamentals of Life*, Paris: Elsevier, 57-72.

Fig 1: The difference between sequence “order” and “complexity”

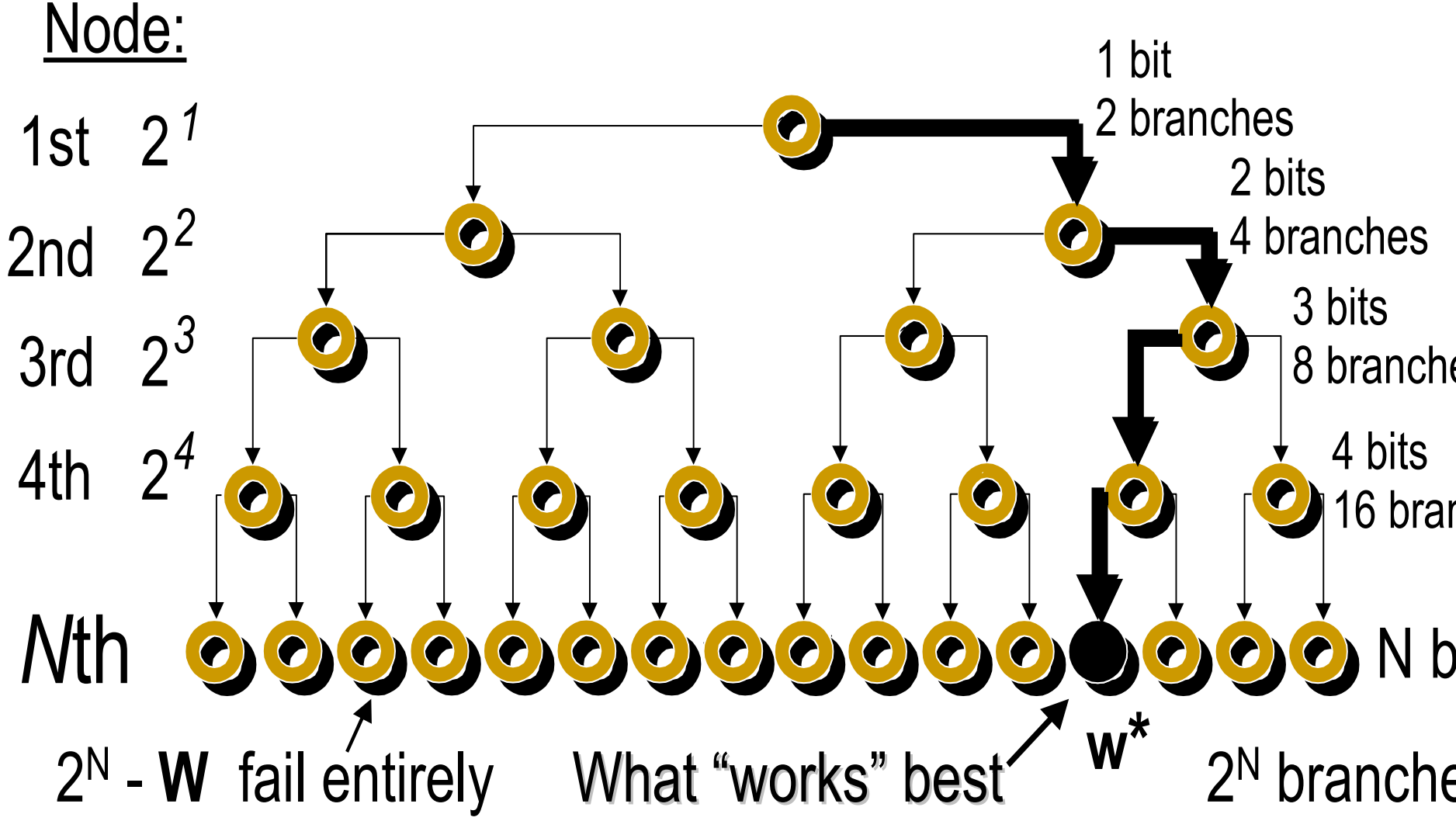
ORDER	COMPLEXITY
regular	irregular
repeating	nonrepeating
redundant	nonredundant
predictable	nonpredictable
symmetrical	asymmetrical
periodic	aperiodic
monotonous	variable
crystal-like patterning	linguistic-like patterning°
reducible	largely irreducible
compressible	noncompressible*

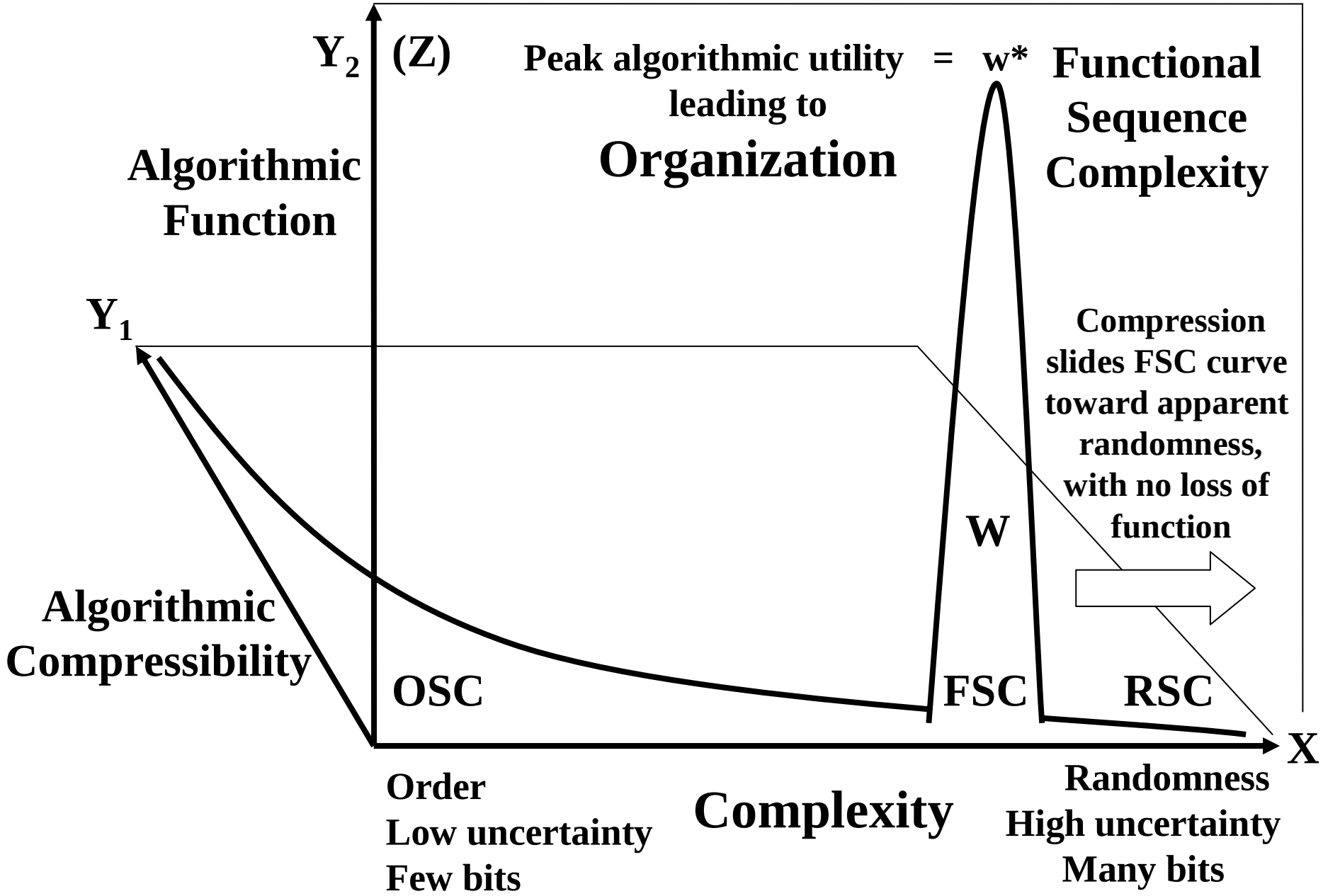
°Maximum complexity (randomness) lacks patterns and order

*Linguistic-like patterning permits some degree of compressibility.
Random complexity does not.

From: Abel, David L. (2002), "Is Life Reducible to Complexity?" in Gyula Palyi, Claudia Zucchi and Luciano Caglioti (eds.), *Fundamentals of Life*, Paris: Elsevier, 57-72.

One of very few paths leading to function out of 2^N branches




Y2 (Z)
Peak algorithmic utility leading to
Organization
=
w*
Functional Sequence Complexity
Algorithmic Function
Y1
Compression slides FSC curve toward apparent randomness, with no loss of function
W
Algorithmic Compressibility
OSC
FSC
RSC
X
Order
Low uncertainty
Few bits
Complexity
Randomness
High uncertainty
Many bits

What provides the third dimension is when each token in the linear string is *selected* for function.

The string becomes a cybernetic program capable of computation only when signs/symbols/tokens are *chosen* to *represent* utilitarian configurable switch settings.

No empirical evidence exists,
not even an anecdotal account,
of maximum complexity
ever having produced
sophisticated algorithmic function or
cybernetic organization
of any kind.

No empirical evidence exists,
not even an anecdotal account,
of highly ordered states
ever having produced
sophisticated algorithmic function or
cybernetic organization
of any kind.

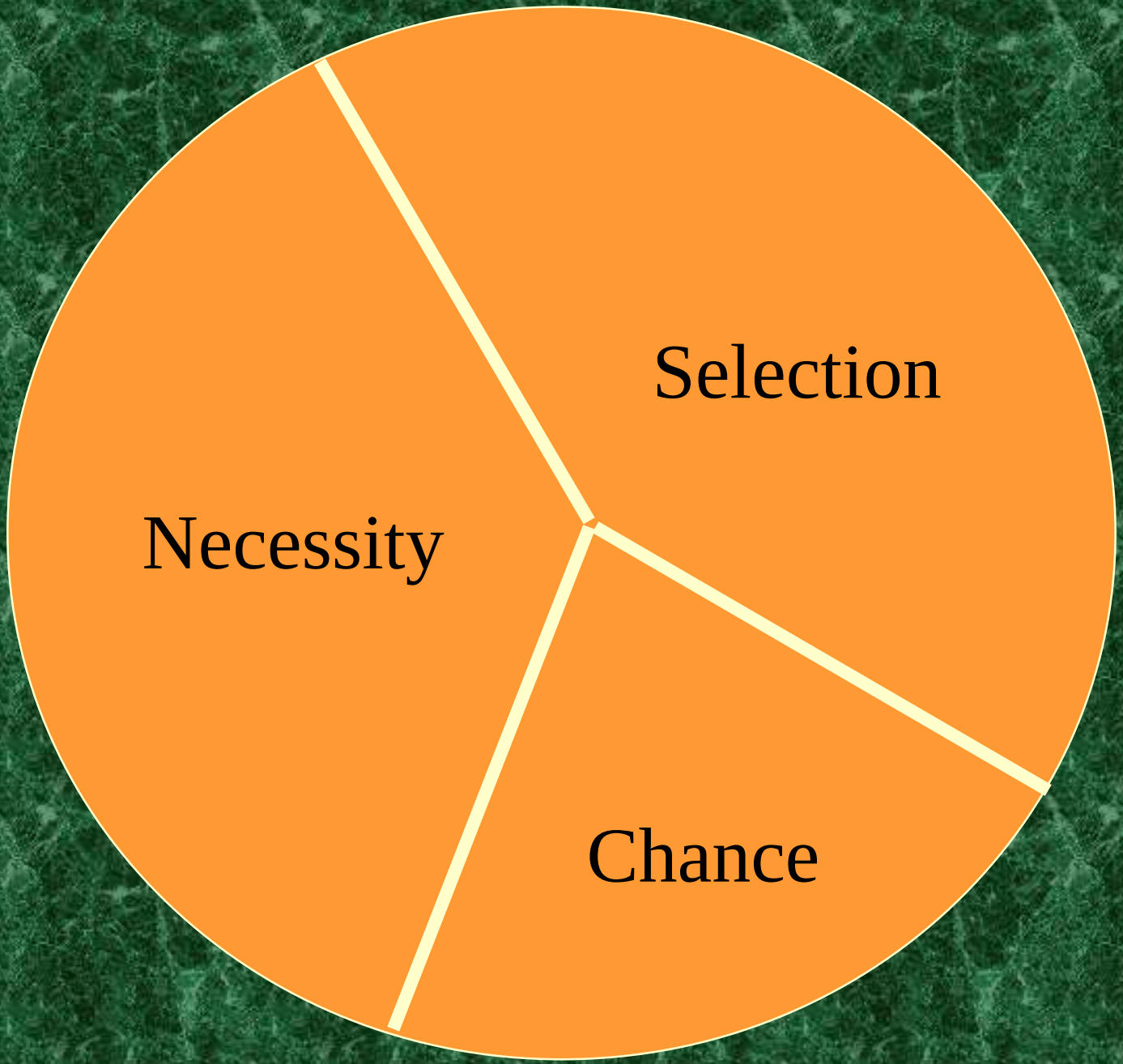
There is no basis in science for attributing linear digital messages and instructions to “order” *or* “complexity”

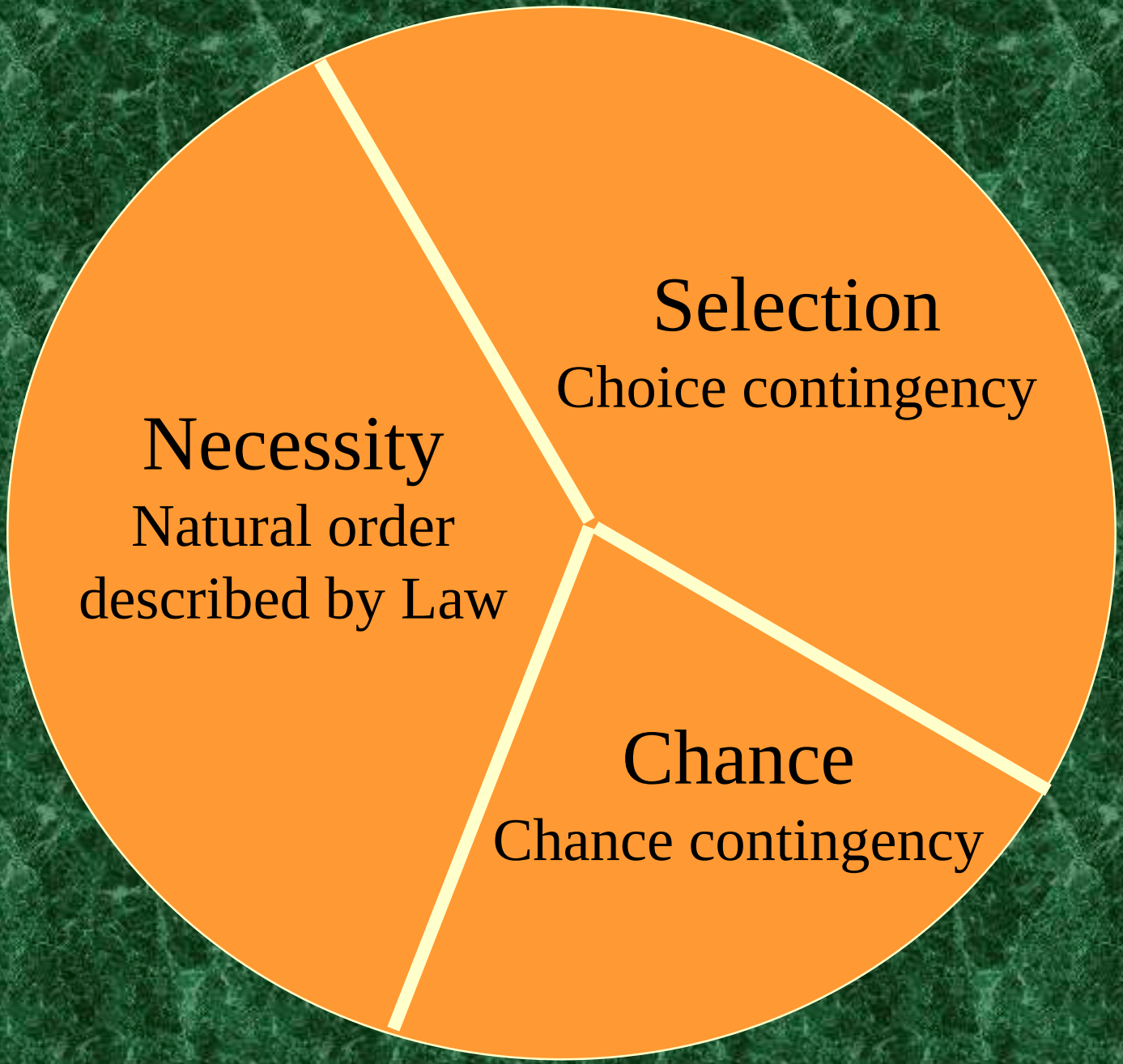
No rational or empirical justification
exists for attributing
linear, digital, encrypted, genetic recipes
to stochastic ensembles
in *any* amount of time.

Linear digital genetic instructions represent *selection-based* cybernetic programming.

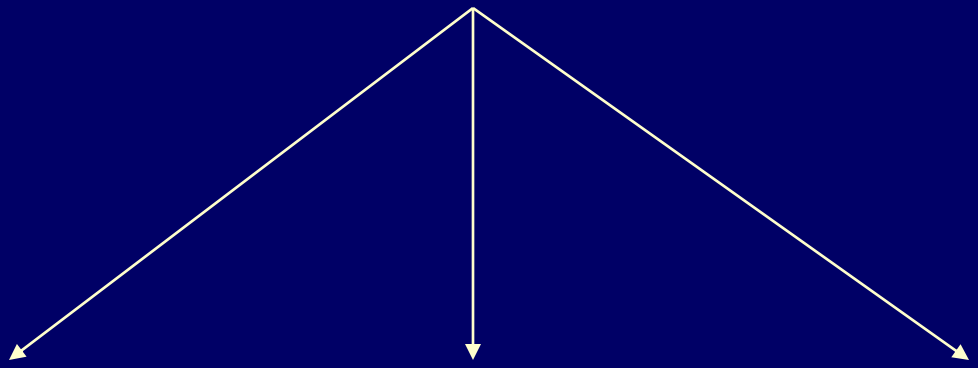
Without *selection*, evolution is impossible.

But if complexity
has nothing to do with *function*,
why do we persist in pointing to complexity
as a scientific explanation
of function?


Selection
Necessity
Chance

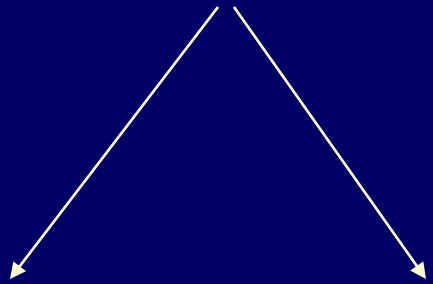

Selection
Choice contingency
Necessity
Natural order
described by Law
Chance
Chance contingency

Scientifically Addressable Presupposed Objective Reality



Chance
(Quantifiable using formal combinatorial probabilism and statistical mechanics)

Selection



Selection *of* Existing Fitness (Natural Selection)
(Selection Pressure)
(Survival of the fittest already-computed phenotypes)

Selection *for* Potential Fitness (Artificial Selection)
(Logic, Math & Computation)
(Algorithmic programming)
(Engineering function)
(Decision nodes; Sign systems)
(Choice with intent: Formality)
(Linear digital genetic instructions)

Necessity
(Formally modelable, low-informational, highly compressible, natural order)

What exactly is "Complexity"?

To appeal to complexity
as an explanation for
selection-based cybernetic programming
at bona fide decision nodes
is a "category error,"
a fallacious logical inference

What exactly is

"The Edge of Chaos"?

Stuart Kauffmann

Does an "edge" of chaos really exist?

What exactly do chaos and complexity interface with
on their "edges"?

(maximum complexity corresponds to randomness and noise)

Order

Ordered Sequence Complexity (OSC)

Polyadenosines on a clay surface

Randomness

Random Sequence Complexity (RSC)

Stochastic ensembles

Increasing complexity →

Minimal Uncertainty (P = 1.0)

Low Shannon bit content

Maximum compressibility

Most patterned

Maximum Uncertainty

High Shannon bit content

Minimum compressibility

Least patterned

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. Modified from Yockey HP: In *Fundamentals of Life*. Edited by Palyi G, Zucchi C, Caglioti L. Paris: Elsevier; 2002: 335-348

What is the scientific basis for believing that anything at the edge of chaos can generate cybernetic programs?

- **The self-ordering phenomena**
- **of chaos theory**
- **have nothing to do with**
- **cybernetic organization.**

What is the scientific basis for believing that anything at the edge of chaos can generate cybernetic programs?

The dissipative structures of Prigogine arise out of high order "necessity." Dissipative structures self-order. They do NOT self-organize.

Dissipative structures are

1) highly ordered,

2) monotonous,

3) predictable,

4) regular (vortices, sand piles)

5) low informational

6) momentary

- **Dissipative structures are**
  - **usually destructive,**
- **not cybernetically constructive**
  - **(e.g., tornadoes, hurricanes)**

If this mystical edge of chaos actually existed, what could chaos interface with other than natural order?

Can “Order” program configurable switches?

Organization requires:

Decision nodes

Algorithmic steering

Algorithmic optimization

Selective controls

Choice with intent

If "Order" programmed configurable switches, they would all be programmed the same way.

They would all be set to "0's"

Or they would all be set to "1's"

The dissipative structures of chaos theory are unimaginative.

- Highly ordered structures contain very little information.
- Information retention in any physical medium requires freedom of selection of configurable switch settings.
- Switches must be “dynamically inert” with respect to their function as decision nodes.

What about fractals? Can't fractals generate genetic algorithms?

Fractals are also

- **1) highly ordered,**
- **2) redundant,**
- **3) predictable,**
- **4) unimaginative**
- **5) low informational**
- **6) non cybernetic**

The Cantor Dust Fractal representative of the high order content of all fractals.

The low complexity is demonstrated by the simple Kolmogorov compression algorithm:

“Take a line segment, remove the center third. Repeat N times.”

If necessity
can't possibly program
configurable switches,
what about chance?

Given enough time, couldn't a
cybernetic program
arise spontaneously?

Has anyone ever seen a random number generator produce a halting computational program?

If not, why do we cling to
blind belief in primordial soup,
stochastic ensembles, and
random mutations
so religiously,
so dogmatically?

Can we name a single cybernetic program that has ever spontaneously arisen from chaos in the history of human observation?

If not, why are we
SO IN LOVE
with this notion?

In an infinite amount of time Maximum complexity would:

- Set all of the configurable switches randomly.
- Not be distinguishable from noise pollution.
- Fail to compute anything; it would fail to halt.
- Consist of nothing but “bugs”
- “Blue screen” every time .
- Be the perfect example of “Garbage in, Garbage out!”

But what about selection pressure?

Can't randomness
plus environmental selection
write genetic programs
given long enough
periods of time?

Environmental selection
works only on
already-computed phenotypes

Selection pressure
cannot choose nucleotide tokens
with programming intent
at the genetic level.

Genomes are linear digital programs.

Genomes employ nucleotide selections as physical symbol vehicles in a sign/symbol/token system.

Each nucleotide selection represents a dynamically-inert configurable switch-setting.

No living organism has ever been found that did not depend upon linear digital cybernetic instructions and management.

Genotype prescribes phenotype

No reason, experience, or fulfilled predictions support the notion that either

self-ordering out of chaos
or
maximum complexity (randomness)
could
cybernetically program formal computations

We are committed to a Kuhnian paradigm rut far worse than Ptolemaic astronomy

Science is about:

1) Repeated observation

2) *Rational* inference

3) Falsifiability

4) Predictability

Believing in spontaneously occurring *formal* instructions and redundancy block coding is nothing less than superstition

There is no basis in science
for attributing
cybernetic programs
to random number generators
(the equivalent of stochastic ensembles in primordial soup)

Computation,
cybernetic programming, and
algorithmic optimization
are all nonphysical
FORMAL functions,
not inanimate, “natural”
physical events

In NO amount of time
can physicodynamic
self-ordering phenomena
set configurable switches
to achieve

FORMAL COMPUTATIONAL HALTING

If complexity has nothing to do with function, why do we persist in pointing to complexity as a supposed explanation of function?

Natural selection favors
the fittest
already-living organisms.

Natural selection cannot
and does not
select configurable switch-settings
at the genetic programming level.

Primary structure
(linear digital sequencing)
determines the range of
secondary and tertiary structures
(three-dimensional protein conformations)
that catalyze metabolism.

What about three-dimensional complexity?

- Primary structure = the sequence of monomers
- Every nucleotide addition to the string constitutes a four-way configurable switch setting
- Dynamic folding of proteins results from minimum-free-energy folding space *controlled by primary structure programming*.

What about three-dimensional complexity?

- The sequence is rigidly bound by covalent bonds before weak-bonded folding ever begins
- Sequencing *determines* the range of folding options the strand can assume.
- The fact that many segments are not critical does not change the primacy of sequencing.

Sequencing is covalently bound, rigidly bound, before folding ever begins.

Programming selections not only precede protein conformation, they determine average protein conformation.

Primary structure
consists of a string of pre-set
configurable switches.

Three-dimensional shapes are computed.

Protein conformation is
cybernetically programmed

Hamming redundancy block-coding is employed in all known organisms.

Triplet codons *represent* each amino acid.

The genetic CODE is a noise-reducing, abstract, conceptual

encryption/decryption system

Fully falsifiable null hypotheses:

Only one observed exception would be required to falsify the following null hypothesis:

"*Stochastic* ensembles of physical units cannot program algorithmic/cybernetic/computational halting function."

Only one observed exception would be required to falsify the following null hypothesis:

"*Dynamically-ordered* sequences
of individual physical units
(physically patterned by natural law causation)
cannot program
algorithmic/cybernetic/computational
formal halting function."

Only one observed exception would be required to falsify the following null hypothesis:

"*Statistically weighted means* (e.g., increased availability of certain units in the polymerization environment of sequence space) giving rise to patterned (compressible) sequences of units cannot program algorithmic/cybernetic/computational formal halting function."

Only one observed exception would be required to falsify the following null hypothesis:

"Computationally halting
configurable switch settings
cannot be set by chance, necessity,
or any combination of the two."

Have we really explained anything by attributing cybernetic function to a mysterious, ill-defined "edge of chaos"?

"The edge of chaos"

(the imagined interface of randomness and

self-ordering phenomena)

cannot program or instruct.

Linear digital symbol systems are formal, not physical

(though they can be instantiated into a physical medium).

- Symbols
  1) have *arbitrarily* assigned meaning
  2) are *selected* specifically *for function*.
- Both operating system and application software require choice contingency to program
- Rules must be followed voluntarily (rules can be broken)

Chance and necessity cannot program an algorithm

A linear digital sign system can only instruct if *Choice Contingency* is wisely exercised at successive decision nodes.

Chance contingency and law cannot generate programming choices, instructions, organizational controls, and computational halting.

Self-ordering events are physicodynamic.

Organization is formal.

Physicality has no ability to perform formal functions

Physicality possesses no capability of self-organizing itself.

The difference between spontaneous “self-ordering” and “self-organizing” systems in nature.

SELF-ORDERING	SELF-ORGANIZING
Increases redundancy	Decreases redundancy
Increases predictability	Decreases predictability
Increases symmetry	Decreases symmetry
Increases periodicity	Decreases periodicity
Increases monotony	Decreases monotony
Produces crystal-like patterns	Produces linguistic-like patterns

Abel, David L. (2002), "Is Life Reducible to Complexity?" in Gyula Palyi, Claudia Zucchi and Luciano Caglioti (eds.), *Fundamentals of Life*, Paris: Elsevier, 57-72.

The difference between spontaneous "self-ordering" and "self-organizing" systems in nature.

SELF-ORDERING	SELF-ORGANIZING
Decreases complexity	Increases complexity
Short-lived (highly dissipative)	Long-lasting (minimal dissipation)
Produced by cause-and-effect	Lacks natural process mechanism
Observed	Has never been observed
Consistent with 2nd Law	Seems inconsistent with the 2nd Law
Non integrative	Integrative
Non conceptual	Abstract, Conceptual, Cybernetic
Not particularly functional	Produces extraordinary function

Abel, David L. (2002), "Is Life Reducible to Complexity?" in Gyula Palyi, Claudia Zucchi and Luciano Caglioti (eds.), *Fundamentals of Life*, Paris: Elsevier, 57-72.

Inanimate nature possesses no ability
to exercise the foresight,
choice contingency,
integrative engineering intent,
or metabolic motivation
necessary for
programming a triplet codon
sign system.

What is the role of "Complexity" In life origin?

Neither linear digital genetic symbol systems
nor noise-reducing
redundancy block-coding
(the triplet codon system)
can be attributed to "complexity"
as scientifically defined.

No amount of time

can generate

Choice Contingency

out of

Chance and Necessity

Life origin:

the role of

Complexity

The Edge of Chaos

David L. Abel, Program Director

The Gene Emergence Project

The Origin of Life Foundation, Inc.

(501-c-3 U.S. science foundation)