# Information, life, and artificial evolution

## Information and life

The notion that the deep roots of the origin and evolution of life are to be found in information theory seems increasingly common [1, 4, 5]. Information theory [9] specifies that information is about messaging or signaling, and thus requires a receiving system. For example, a computer digital storage device would be devoid of information without a device to interpret it.

We see life everywhere on our planet, and yet never do we observe it to arise spontaneously, likely because such life would be instantly consumed by its ruthlessly efficient predecessors. We also observe life nowhere else in our solar system as yet. The moons Europa and Enceladus beckon us with hints of warm oceans beneath icy sheaths. What is below?

We yearn to know whether life is unique in the cosmos, where millions of earth-like worlds lie scattered among the stars. Would it be more shocking to land on such a planet, watery and warm for billions of years, and find life, or not to find life at all? Fermi's Paradox [8] speaks of alien intelligence, but we members of Homo sapiens, as carriers of the condition on earth have only been about for a few million years, and our claim to immortality is far from settled. Life itself should have a firmer claim to ubiquity. But no one knows the dice that were thrown nor how many throws before life took hold here.

One approach to define and formalize life is to frame it as a process of entropy and thermodynamics. We know that complex systems lie in a middle ground between high and low entropy [2], and that the thermal gradient imbued by our star warmed and stirred the primordial soup suitably for life, the quintessential complex system, to arise.

One position, that taken by Adami [1], is that systems that persistently and maintain an entropy that is less than maximal must be considered in some sense lifelike:

"Living systems can stay away from maximum entropy for much longer, indeed arbitrarily long (the biotic time scale is, for all we know, only limited by the existence of the biosphere). It is then this ability: to persist in a state of reduced entropy for biotic as opposed to abiotic time scales, that defines a set of molecules as living, and this set of molecules must achieve that feat via the self-replication of information."

Stephen Wolfram has in the past conjectured that not only life but all of reality is in a type of cellular automaton [17], drawing illustrations about how simple rule sets can produce dynamic patterns within them. More recently he has argued that graphs, another abstract informational construct, might model the observed universe [18].

Morphogenesis is a process of diversifying organs and tissues within an organism to complete its life cycle. Nature features many mechanisms, chemical, mechanical, electrical, etc. to accomplish this. However, at its core it is information that is being signaled and utilized. To abstract this within a cellular automaton, the *Morphozoic* model [13] has shown that tractable rule sets composed of hierarchically nested cellular neighborhoods are capable of producing both local and global information processing effects that can simulate reaction-diffusion morphogenesis [16], gastrulation, axon pathfinding, and other phenomena.

There is evidence that natural complex systems, rather than becoming more fragile, actually become more robust and able to sustain and mutate. Even artificial systems seem to show this property under certain circumstances. For example, variations of circuits that can implement a logical function become more numerous as the number of components grows [14]. Ball [3] writes:

"These ideas suggest that evolvability and openness to innovation are features not just of life but of *information itself*. That is a view long championed by Schuster's sometime collaborator, Nobel laureate chemist Manfred Eigen, who insists that Darwinian evolution is not merely the organizing principle of biology but a "law of physics," an inevitable result of how information is organized in complex systems. And if that's right, it would seem that the appearance of life was not a fantastic fluke but almost a mathematical inevitability."

Continuing along these lines, studies of genetic variations indicates that "fit" configurations tend to be located on "paths" in the space of all possible configurations. This allow one successful genotype to migrate through mutation into other successful ones that could be more complex. This can be taken as a refutation of the oft-cited implausibility of life arising due to its sheer combinatorial complexity. Instead, we have systems that can evolve from simple origins to assume some of the many potential successful forms that nature permits. Ball summarizes thusly [3]:

"Successful forms of RNA, proteins, etc., although taken in isolation are immensely improbable, are not the product of shaking a bunch of parts in a box and expecting a watch to assemble. Instead, there are pathways in highly dimensional, complex systems that are connected by relatively simple steps that lead to "fit" phenotypes. And strikingly, more complexity leads to more paths to fitness instead of making it harder and harder. So if you start with a minimally fit phenotype in a complex physical/chemical/biological milieu, mutable pathways will exist to take it to other fit phenotypes."

Supporting views about informational complexity arise from the arena of artificial neural networks, where research into why these networks are so successful at classifying an astronomical number of possible inputs has supported a position that the nature of the universe is such that there are relatively few fundamental functions within it [10]. For example, the known laws of physics are relatively compact and of consist of polynomial of low order.

## Artificial evolution

The technique of genetic algorithms was introduced in the 1980s by John Holland [7]. This is a popular and highly successful means of optimizing a process in a high dimensional and rough space that mimics biological evolution. It encompasses such biological counterparts as genes, genotypes, and phenotypes that are evaluated by some fitness function. Mutations and genotype mating with gene cross-over are also used to produce fitter offspring within an evolving population of genotypes. Fitter offspring are selected to populate future generations. Genetic algorithms have been found to be capable of finding solutions within large and rough feature spaces, and thus are often employed as optimizers for problems of this sort.

One striking example of the power of artificial evolution as an optimizer comes from the following [6]:

"A few years ago, Michael Levin faced a conundrum. He and his colleagues at the Tufts Center for Regenerative and Developmental Biology just outside Boston wanted to find a model that would explain why the flatworm—a model organism used throughout biology—looks the way it does. At a fundamental level, they wanted to be able to describe the cascade of events that leads to the growth of a head in one place and a tail in the other."

So a genetic algorithm was developed to try to find commonalities in 1,000 experiments related to the head-trunk-tail pattern in the flatworm. It worked: "In the end, it took 6 billion simulated experiments, 26,727 generations of models, and about 42 hours of processing by the Stampede computer before the computer came up with one result." The model even explained results of papers not supplied to the algorithm that would have possibly affected the viability of solutions. It even predicted unknown results that were subsequently verified.

This is an example of how computational biology, and artificial evolution in particular, can be used to explain and even predict biological phenomena.

Another optimization utilizing genetic algorithms involved the weightings on the synapses that connect neurons in the nematode worm C. elegans. The worm has 302 neurons and a known wiring diagram called a connectome. There are over 3,000 synaptic connections between these neurons. Synapses are weighted such that the activation states of source neurons affect target neurons variably. These weights are difficult to determine. Portegys [12], using a hybrid genetic algorithm, showed how the weights could be optimized to produce arbitrary input-output sequences, suggesting possible weighting schemes for the actual synapses to produce observed behaviors.

The 1990s were a hotbed of artificial life forms that evolved and mutated into emergent and unexpected results. Among these were:

1. Tierra [15]. The goal was for competing chunks of computer code to vie for CPU and memory. It featured evolvability, mutations, replication, recombination, host-parasite co-evolution, and punctuated equilibrium.
2. Avida [2]. Similar in many ways to Tierra, but organisms compete for CPU only as they are isolated within memory bounds.

3. Polyworld [19]. 2D creatures live, forage, prey, reproduce, evolve and mutate. The evolvable genome expresses not only the form of a creature, but also its behavior, which is controlled by an artificial neural network.

Along the lines of engineered artificial evolution, in order to avoid malware detection, some computer viruses feature *metamorphic code*, which is self-editing but preserves the original function. This might entail adding bits that accomplish a sub-function through alternative means, or adding dead-code bits that obfuscate the signature libraries of virus detection programs.

## References


1. Adami, Christoph. (2015). Information-Theoretic Considerations Concerning the Origin of Life. In Origins of Life and Evolution of the Biospheres. Volume 45, Issue 3, pp 309–317.
2. Adami, C. and Brown, C.T. (1994). Evolutionary Learning in the 2D Artificial Life Systems Avida, in: R. Brooks, P. Maes (Eds.), Proc. Artificial Life IV, MIT Press, Cambridge, MA, p. 377-381. arXiv:adap-org/9405003
3. Ball, Philip. (2015). The Strange Inevitability of Evolution, Good solutions to biology's problems are astonishingly plentiful. Nautilus. http://nautil.us/issue/20/creativity/the-strange-inevitability-of-evolution
4. Ben-Naim, Arieh, (2015). Information, Entropy, Life and the Universe, What We Know and What We Do Not Know. World Scientific. ISBN: 978-981-4651-66-0.
5. Carroll, Sean. (2016). The Big Picture: On the Origins of Life, Meaning, and the Universe Itself. Penguin Random House. ISBN 9780525954828.
6. Graber, Cynthia. (2016). Replicating Life in Code. Nova Next. http://www.pbs.org/wgbh/nova/next/evolution/ai-biology/
7. Holland, John (1992). Adaptation in Natural and Artificial Systems. Cambridge, MA: MIT Press. ISBN 978-0262581110.
8. Jones, E. M. (1985). "Where is everybody?" An account of Fermi's question. Los Alamos National Laboratory. OSTI 785733.
9. Küppers, Bernd-Olaf. (1990). Information and the Origin of Life. The MIT Press. ISBN-10: 026211142X.
10. Lin, H.W. and Tegmark, M. (2016). Why does deep and cheap learning work so well? https://arxiv.org/abs/1608.08225
11. Pierce, John R. (1980). An Introduction to Information Theory, Second Revised Edition. Dover Publications.
12. Portegys, T. (2015). Training sensory-motor behavior in the connectome of an artificial C. elegans. Neurocomputing. pp. 128-134. DOI: 10.1016/j.neucom.2015.06.007
13. Portegys, T., Pascualy, G., Gordon, R., McGrew, S., Alicea, B. (2017). Morphozoic: cellular automata with nested neighborhoods as a metamorphic representation of morphogenesis. In Multi-Agent Based Simulations Applied to Biological and Environmental Systems. ISBN: 978-1-5225-1756-6.

14. Raman, K. & Wagner, A. (2011).The evolvability of programmable hardware. Journal of the Royal Society Interface 8, 269-281.
15. Ray, T. S. (1991). Evolution and optimization of digital organisms. In Billingsley K.R. et al. (eds), Scientific Excellence in Supercomputing: The IBM 1990 Contest Prize Papers, Athens, GA, 30602: The Baldwin Press, The University of Georgia. Publication date: December 1991, pp. 489–531.
16. Turing, A.M. (1952). The chemical basis of morphogenesis. Phil. Trans. Roy. Soc. London B237, 37-72.
17. Wolfram, Stephen. (2002). A New Kind of Science. Wolfram Media. ISBN-10: 1579550088.
18. Wolfram, Stephen. (2015). What Is Spacetime, Really? Stephen Wolfram Blog. http://blog.stephenwolfram.com/2015/12/what-is-spacetime-really/
19. Yaeger, L. (2008). Polyworld. https://github.com/polyworld/polyworld/tree/master/docs