I-485/H-400 *Biologically-inspired computing*, Lecture Notes
Spring 2015, Indiana University, By Luis M. Rocha

1. What is Life?

> "*What was life? No one knew. It was undoubtedly aware of itself, so soon as it was life; but it did not know what it was"*. Thomas Mann [1924]

Threshold of Complexity

> "*Seeking a connecting link, they had condescended to the preposterous assumption of structureless living matter, unorganized organisms, which darted together of themselves in the albumen solution, like crystals in their mother-liquor; yet organic differentiation still remained at once condition and expression of all life. One could point to no form of life that did not owe its existence to procreation by parents*". Thomas Mann [1924].

> "*Nothing in biology makes sense without evolution*". Theodosius Dobzhansky [1973]

Biologically-inspired computing is an interdisciplinary field that formalizes processes observed in living systems to design computational methods for solving complex problems, or simply to endow artificial systems with more natural traits. But to draw more than superficial inspiration from Biology we need to understand and discuss the concept of *life*. It should be noted that for the most part of the history of humanity, the question of what *life* is was not an important issue. Before the study of mechanics became important, everything was thought to be alive: the stars, the skies, the rivers and mountains, etc. There was no non-life, so the concept was of no importance. It was only when people started to see the World as determined by the laws of mechanics that the question arose. If all matter follows simple physical laws, then what is indeed the difference between life and non-life, between biology and physics? Let us then start with a current dictionary definition:

> "life adj.— n.1. the general condition that distinguishes organisms from inorganic objects and dead organisms, being manifested by growth through metabolism, a means of reproduction, and internal regulation in response to the environment. 2. the animate existence or period of animate existence of an individual. 3. a corresponding state, existence, or principle of existence conceived of as belonging to the soul. 4. the general or universal condition of human existence. 5. any specified period of animate existence. 6. the period of existence, activity, or effectiveness of something inanimate, as a machine, lease, or play. 7. animation; liveliness; spirit: The party was full of life. 8. the force that makes or keeps something alive; the vivifying or quickening principle." [Random House Webster's Dictionary]

The definitions above fall into three main categories: (1) life as an organization distinct from inorganic matter (with an associated list of properties), (2) life as a certain kind of *animated behavior*, and (3) life as a special, incommensurable, quality—*vitalism*. Throughout this course we will see that all principles, and indeed all controversies, associated with the study of life fall into one of these categories or the differences among them. The third category has been discarded as a viable scientific explanation, because for science nothing is in principle incommensurable. The question of whether life is organized according to a special design, intelligent or mysterious, pertains to metaphysics. If the agent of design cannot be observed with physical means, then it is by definition beyond the scope of science as it cannot be measured, and any theories derived from such a concept cannot tested.

While metaphysical dispositions do not pertain to science, many scientists have observed that a naive mechanistic decomposition of life may also fail to explain it. The traditional scientific approach has lead the study of living systems into a reductionist search for answers in the nitty-gritty of the biochemistry of living organisms. This alternative sees life as nothing more than the complicated physics of a collection of moving

bodies. However, the question remains unanswered since there are many ways to obtain some complicated dynamics, but of all of these, which ones can be classified as *alive*? What *kind of complexity* are we looking for? No one disputes that life is some sort of complex material arrangement, but when do we reach a necessary *threshold of complexity* after which matter is said to be living? Is it a discrete step, or is life a fuzzy concept? To understand it without meaningless reduction, must we synthesize organizations with the same threshold of complexity (first category above), or is it enough to simulate its animated behavior (second category above)?

Information Organizes and Breeds Life

> "*Life is a dynamic state of matter organized by information*". Manfred Eigen [1992]

> "*Life is a complex system for information storage and processing*". Minoru Kanehisa [2000]

Traditionally life has been identified with material organizations which observe certain lists of properties, e.g. *metabolism, adaptability, self-maintenance* (autonomy), *self-repair, growth, replication, evolution,* etc. Most living organisms follow these lists, however, there are other material systems which obey only a subset of these rules, e.g. viruses, candle flames, the Earth, certain robots, etc. This often leads to the view that life is at best a fuzzy concept and at worst something we are, subjectively, trained to recognize—life is what we can eat—and is thus not an objective distinction. The modern-day molecular biology view of life, on the other hand, tends to see life as a material organization that if not completely defined by *genomic information*, is at least fully controlled by it. Thus, when Craig Venter's team [Gibson et al, 2010] recently produced a bacteria with a "*prosthetic genome*" [a termed coined by Mark Bedau, see *Nature* | Opinion, 2010] copied from another bacteria but synthesized in the lab, the momentous synthetic biology feat was announced as the creation of the first synthetic or artificial life form.

The artificial life field, whose members tend to follow the fuzzy list of properties conception of life, does not typically recognize Venter's bacteria with a prosthetic genome as a bona fide synthesis of artificial life, since it relies on the pre-existence of a working, naturally-obtained cell to implant a prosthetic genome into. Even most molecular biologists will agree that we are nowhere near <u>understanding</u>, let alone synthesizing an artificial cell from scratch [e.g. George Church, see *Nature* | Opinion, 2010]. Nonetheless, Venter's achievement begs at least the question of what is it about life's *design principle* that makes it easier to synthesize a working prosthetic genome than a working "prosthetic proteome or metabolome"? It also makes us think about what does "understanding life" mean for biology, biomedical technology, artificial life, and informatics? Why is *genetic information* so important and how does it relate to *information technology*?

Life requires the ability to both categorize and control events in its environment in order to survive. In other words, organisms pursue (or even decide upon) different actions according to *information* they perceive in an environment. Furthermore, living organisms reproduce and develop from genetic *information*. More specifically, genetic information is <u>*transmitted*</u> "vertically" (inherited) in phylogeny and cell reproduction, and <u>*expressed*</u> "horizontally" within a cell in ontogeny for the functioning of living organisms as they interact and react with their environments—we are now sure that genetic information can also be transmitted horizontally between organisms and play an important role in evolution [Goldenfeld & Woese 2007; Riley, 2013]. Indeed, the difference between living and non-living organizations seems to stand on the ability of the former to use *relevant* information for their own functioning. It is this "relevant" which gives life an extra attribute to simple mechanistic interactions. When an organization is able to recognize and act on aspects of its environment which are important *to its own* survival, we say that the mechanisms by which the organization recognizes and acts are *functional* in reference to the organization itself (*self-reference*). Physics is not concerned with function. A physical or chemical description of DNA is certainly possible, but will tell

us nothing as to the function of a DNA molecule as a gene containing relevant information for a particular organism. Only in *reference* to an organism does a piece of DNA *function* as a gene (e.g. an enzyme with some effect in an environment).

Thus it is remarkable that in Venter's experiment, a cell with a synthesized prosthetic genome from a similar but distinct organism, was able to reproduce over and over resulting in a cell with a different phenotype from the original, implanted cell—in effect, a cell *re-programmed* by a synthesized genome. Is life then a type of computer that can be reprogrammed? This also leads us to question how general-purpose can such genomic re-programming be? Will it be restricted to very narrow classes of similar organisms, or will it ever be possible to re-program any prokaryotic or eukaryotic cell ?

Emergence and Explanation

> "*First, nothing in biology contradicts the laws of physics and chemistry; any adequate biology must be consonant with the 'basic' sciences. Second, the principles of physics and chemistry are not sufficient to explain complex biological objects because new properties emerge as a result of organization and interaction. These properties can only be understood by the direct study of the whole, living systems in their normal state. Third, the insufficiency of physics and chemistry to encompass life records no mystical addition, no contradiction to the basic sciences, but only reflects the hierarchy of natural objects and the principle of emergent properties at higher levels of organization*". Stephen Jay Gould [1984].

This issue could be rephrased in terms of the notion of *emergence*. Whatever (macro-level) organization exists after the complexity threshold for life is passed, we may say that it is emergent because its attributes cannot be completely explained by the (micro-) physical level. In particular, function, control, and categorization cannot be explained by the mechanics and dynamics of the components of life alone. Notice, however, that emergence does not imply vitalism or dualism. When we say that certain characteristics of life cannot be explained by physics alone, we mean that they must be explained by different, additional models—namely, *informational*, *historical* and *functional* descriptions. In other words, though biological function, control, and categorization cannot be explained by physics alone, organisms, like anything else, must nonetheless follow physical laws. But information is contextual, and therefore requires more than universal models: it requires contingent, context-specific descriptions. In particular, the *origin of life*, is a problem of emergence of information from a physical milieu under specific constraints [Eigen, 1992]. This is the crux of complex systems: the interplay between micro- and macro-level descriptions determines their behavior, and both levels (emergence) are required to understand complexity.

The definition of emergence as an epistemological, explanatory requirement, is related to the notion of *emergence-relative-to-a-model* [Rosen, 1985; Cariani, 1989] or *intensional emergence* [Salthe 1991]. It refers to the impossibility of epistemological reduction of the properties of a system to its components [Clark, 1996]. As an example, we can think of phase transitions such as that of water in its transition from liquid to gas. Water and its properties cannot be rephrased it terms of the properties of hydrogen and oxygen, it needs a qualitatively different model. Another example of complementary models of the same material systems is the wave-particle duality of light.

Physicists *understand* the laws of nature (as best they can), but it takes engineers to *control* nature. The very best physicists are the very best engineers, but those are exceedingly rare (e.g. Von Neumann). The goal of complex systems is to understand *organized complexity* (life, society, cognition) in the same way physicists understand nature [Weaver, 1948]. Biology, as a discipline, has not entirely "made up its mind" if it wants to understand life as a physicist or control it as an engineer. Due to its focus on the micro-level of life, its biochemistry, molecular biology follows essentially a (reverse-) engineering, black-box methodology

(knockouts, controls, etc.). This leads to a bit of a schizophrenic agenda: focusing exclusively on micro-level experiments in order to *suggest* macro-level understandings. If the goal is control of biology, say for biomedical advances, then we really need to focus on biotechnology engineering. If the goal is understanding, then we need to focus more on macro-level organized complexity. Ideally, a healthy life sciences program would tie the need to understand with the need to control better—like physicists and engineers do.

This is where complex systems, artificial life, and bio-inspired computing can contribute to a wider arena of the life sciences; they can be used as laboratories for experimenting with theories of organized complexity, and thus enrich our understanding of life. Artificial life concerns both the simulation and realization of life in some artificial environment, usually the computer. At least regarding the second of its goals, artificial life aims to understand the fundamental micro/macro-level interaction that leads to organized complexity. Bio-inspired computing, as a more pragmatic endeavor, does not need to concern itself with synthesizing actual life, but only with drawing analogies from life (real and artificial). Nonetheless, if the main motivation of bio-inspired computing is that life with its designs has already solved versions of many complex engineering problems we are interested in, then a thorough and accurate understanding of the essential characteristics of life is inescapable. Moreover, by abstracting context-specific principles of life to make them relevant in other settings, provides a useful laboratory to experiment with theoretical biology.

Further Readings and References:

Cariani, Peter [1989].*On the Design of Devices with Emergent Semantic Functions*. PhD.Dissertation. SUNY Binghamton.

Clark, Andy [1996]. “Happy couplings: emergence and explanatory interlock.” In: *The Philosophy of Artificial Life*. M. Boden (ed.). Oxford University Press, pp. 262-281.

Dobzhansky, T. [1973]. “Nothing in Biology Makes Sense Except in the Light of Evolution”. *The American Biology Teacher*, March 1973 (35:125-129)

Eigen, M. [1992]. *Steps Towards Life*. Oxford University Press.

D. G. Gibson et al [2010]. “Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome”. *Science*. **329** (5987): 52-56

Goldenfeld, Nigel, and Carl Woese [2007]. “Biology’s Next Revolution.” *Nature* **445** (7126): 369.

Gould, Stephen Jay [1984]. *Natural History*; Jan84, Vol. 93 Issue 1, p24.

Mann, T. [1924]. *The Magic Mountain*. As quoted by Eigen [1990].

Nature | Opinion [2010] “Life after the synthetic cell”. *Nature* **465**: 422–424

Pattee, Howard H. [1978]."The complementarity principle in biological and social structures." In: *Journal of Social and Biological Structures* Vol. 1, pp. 191-200.

Polanyi, M. [1968]. “Life’s irreducible structure”. *Science*, **160** (3834), 1308-1312.

Riley, D. R., K.B. Sieber, K. M. Robinson, J. R. White, A. Ganesan, S. Nourbakhsh, and J. C. Dunning Hotopp [2013]. “Bacteria-Human Somatic Cell Lateral Gene Transfer Is Enriched in Cancer Samples.” *PLoS Computational Biology* **9** (6): e1003107.

Salthe, Stanley N. [1991], “Varieties of Emergence”. *World Futures* Vol. 32, pp.69-83

Schrödinger, Erwin [1944]. *What is Life?*. Cambridge University Press.

Weaver, W. [1948]. "Science and Complexity". *American Scientist*, **36**(4): 536-44.

For next lectures read:

Dennet, D.C. [2005]. "Show me the Science". *New York Times*, August 28, 2005.

Gleick, J. [2011]. *The Information: A History, a Theory, a Flood*. Random House. Chapter 8.

Kanehisa, M. [200]. *Post-genome Informatics*. Oxford University Press. **Chapter 1, Blueprint of life**, pp. 1-23.

Langton, C. [1989], “Artificial Life” In *Artificial Life*. C. Langton (Ed.). Addison-Wesley. pp. 1-47.

Nunes de Castro, Leandro [2006]. *Fundamentals of Natural Computing: Basic Concepts, Algorithms, and Applications*. Chapman & Hall. **Chapter 1**, pp. 1-23.

Polt, R. [2012]. "Anything but Human". *New York Times*, August 5, 2012