Book Reviews

Introduction to Artificial Life—C. Adami (Santa Clara, CA: Springer-Verlag, The Electronic Library of Science, 1998, 374 pp., CD-ROM included). *Reviewed by Michael G. Dyer.*

Adami's book, *Introduction to Artificial Life* (abbreviated here as *IAL*), provides a welcome physicist's perspective on the field of artificial life. The emphasis of *IAL* is on developing a solid mathematical and theoretical foundation for the field, through the application of concepts, models, and techniques borrowed mainly from physics, and with some concepts taken from computer science and molecular biology. The book jacket states that the reader is assumed to have already mastered the central concepts and techniques used in the fields of computer architecture, scientific computing, statistical physics, thermodynamics, and biology. This statement is really true. As a nonphysicist I found Adami's book difficult reading in numerous places.

According to Adami, the goal of artificial life (AL) is "...to find a minimum number of characteristics that all those systems classified as living have in common..." (p. 3). Furthermore, any theory must be *universal*, in the sense that it should not need to refer to the materials that make up living systems, but only to the principles involved. Thus, carbon-based life would be just an instance within a more general theory. The most important aspects of living systems for Adami are self-replication, genetic information storage, and transmission in the face of noise, selectional pressure, and low entropy.

Adami considers physiological, metabolic, and biochemical approaches to defining life but finds them overly narrow and constraining. The biochemical approach relies too much on nucleic acids and thus leaves out the possibility of life within other media. The genetic approach is more general because it allows for the modeling of systems that encode information that undergoes mutation and replication, without necessarily specifying the material used to embody such a code. The most general approach according to Adami is the thermodynamic one, in which living systems are defined in terms of their ability to maintain low levels of entropy. Adami's working definition of life is that it is

> ... a property of an ensemble of units that share information in a physical substrate and which, in the presence of noise, manages to keep its entropy significantly lower than the maximum entropy of the ensemble, on timescales exceeding the "natural" timescale of decay of the (information-bearing) substrate by many orders of magnitude. (p. 6)

Thus life is an emergent phenomenon in that it arises as the property of an ensemble of elements and in that it is also defined with respect to its environment.

Before focusing in detail on the thermodynamic definition, Adami provides an overview of other different "flavors" of artificial life. These include a) simulation of artificial agents and the evolution of populations of such individuals, b) carbon-based models, and c) virtual machine models.

For each of these approaches, Adami briefly reviews some well-known models. For instance, in the area of simulation, he describes Sims' [1] model of evolved neural network controllers for creatures composed of multiple attached blocks of different shapes. He also describes a model by Theraulaz and Bonabeau [2] of artificial wasps whose collective local interactions result in the construction of different types of three-dimensional nesting structures.

In the carbon-based area, Adami briefly reviews the work of Wright and Joyce [3] in evolving artificial cultures of ribozymes *in vitro*. This work may ultimately lead to the discovery of an RNA replicase, namely, an RNA molecule capable of self-replication, and believed by many to have been critical in the beginning of cellular life over 3 billion years ago.

In the virtual machine area, Adami reviews the work of both Wolfram and Langton in the classification of cellular automata (CA) models. Wolfram [4] originally classified CA models according to their long-term behavior. Class I models display limit point behavior. That is, from initial states they ultimately arrive at homogeneous terminal states. Class II models develop toward limit cycles, in which they exhibit periodic behavior. Class III models develop aperiodic, chaotic behavior. Finally, Class IV CA's exhibit potentially life-like behavior, in which highly complex patterns appear to self-organize and sustain themselves. Wolfram suggested that it is in this class that there would exist CA's capable of universal computation, which is the ability of a CA to mimic the behavior of other CA's that are encoded as input to it. This is analogous to a genetic code being interpreted by a cell to create new cells.

Langton [5] showed how Wolfram's classes could be reclassified in terms of a single parameter λ that represents the probability that a CA rule will map a cell to a nonquiescent state. Langton examined CA rules for the cellular geometry of a two-dimensional (2-D) regular lattice of cells where each cell contains four neighboring cells (those immediately to its north, south, east, and, west) and whose cells may take on one of eight possible states (0–7, with zero being quiescent). When $\lambda = 0$, all rules map to a quiescent state and so no patterns develop. Langton experimented with many different CA rules with $\lambda > 0$ and discovered that when λ is less than 0.2 then the patterns that form tend ultimately to die out, which corresponds loosely with Wolfram's Class I. When λ is in the general range of $0.2 < \lambda < 0.3$, then patterns tend to exhibit periodicity, which is similar to Wolfram's Class II. When λ is in the range of about $0.3 < \lambda < 0.4$, complex patterns form, which corresponds to Wolfram's Class IV. When λ is greater than 0.5, the patterns become chaotic and any particular complex pattern cannot survive for long, which corresponds with Wolfram's Class III.

Thus λ acts somewhat like the "temperature" of a CA world. If the temperature is too cold (tending toward quiescence) or too hot (tending toward too active) then complex, self-replicating patterns (needed to exhibit life) cannot form. Langton hypothesized that life-like behavior may self-organize most naturally just at "the edge of chaos." Much of *IAL* is devoted to studying this hypothesis, both formally and via simulation.

Probably the most famous virtual machine model is the Tierra system, developed by Ray [6]. Adami only briefly describes this system. Since it was such a landmark I believe it deserves a bit more coverage. Adami, however, concentrates mainly on the avida system, which is an improved version of Tierra, designed by Adami, Brown, Ofria, and other Cal Tech researchers. Avida differs from Tierra in that each virtual machine lives on a 2-D grid, has eight

The reviewer is with the Computer Science Department, University of California, Los Angeles, CA 90095 USA.

Publisher Item Identifier S 1089-778X(98)09427-2.

neighbors, and an orientation. In contrast, Tierra virtual programs interact in a one-dimensional (1-D) world with little sense of locality. In Tierra the instruction pointer (IP) of one creature can continue into the code of a nearby creature. Thus one Tierra creature can execute the code of another. Adami and colleagues felt that having the IP of one creature execute the code of a another (possibly distant) creature is not realistic. Thus, the code of avida creatures is in the form of a loop, so that any IP that continues forward ends up looping back on itself, which is similar to the genomes of many bacteria. Finally, the creation of a daughter cell in Tierra could result in that cell being placed anywhere in Tierra's virtual memory. In contrast, daughter cells in avida are always placed at the end of the mother cell. Thus the avida approach is closer to that of cell division.

The above introductory material appears in Chapters 1 and 2. Chapters 3–8 constitute the core of the book. In these chapters Adami develops the mathematical concepts used later in analyzing the avida experiments which appear in Chapters 9–11. *IAL* comes with a CD containing the avida system code, and in the appendix of the book is an avida user's manual.

In Chapter 3 Adami briefly reviews (without proof) Shannon's fundamental theory of information, which showed that information can be transmitted through a noisy channel with arbitrary accuracy and that the "cost" of doing so merely involves reducing the transmission rate to that of the channel's capacity. A channel's capacity is the maximal mutual entropy between input and output distributions. Adami points out that information

> is *not* the description of an object in terms of bits, or the number of bits necessary to describe an object, but rather the mutual entropy between two ensembles. In other words, information measures the amount of *correlation* between two ensembles, which allows you to make predictions about one ensemble armed only with probabilities garnered from another. (pp. 82–83)

Information theory is of fundamental importance in artificial life because living systems must encode information and also transmit that information through noisy channels. The encoding is that of DNA/RNA; the transmission is that of expressing genes as proteins, and the channel is noisy because errors occur in the DNA/RNA, due to viruses, heat, ultraviolet light, etc. Errors also arise in the processes of replication. Adami shows, mathematically, how uncertainty is related to entropy and information. He calculates the amount of redundancy in the genetic code, consisting of codons (triples of the four nucleic acids C, G, A, and U) that encode 20 amino acids. Adami calculates the information transmission capacity for this system. He shows that "genetic channels are very different from the ordinary ones, because the process that corrupts the messages also controls how much entropy is being sent across the channel" (p. 83).

In Chapter 4, Adami applies concepts from statistical mechanics and thermodynamics. He states that a major goal of artificial life is to "establish a baseline for minimal living systems" (p. 85). To do this, we must be able to describe the collective behavior of a living entity in terms of the forces acting between its constituents. Thus Adami argues that "statistical physics must be the basis of any theory of complexity" (p. 85). In this chapter Adami uses the concepts of phase space (in which every state of a system is represented as a single point), density and trajectories through phase space, and ergodicity (i.e., systems whose statistical distribution function can be traced out by a single trajectory). He argues that genetic space is highly nonergodic because only a tiny fraction of the space will every be traced out in reality.

Adami also introduces the concepts of thermodynamic equilibrium and relaxation time (how long it takes a system to return to equilibrium after perturbations). He argues that large genetic systems tend to be off-equilibrium most of the time. He introduces other concepts, such as Liouville's Theorem, the Second Law of Thermodynamics (how systems approach thermodynamic equilibrium), the mathematical relationships among temperature, energy and entropy, the Gibb's distribution, and first-order phase transitions, which he argues is a useful language for describing major changes that occur during evolution.

In Chapter 5 Adami introduces the concepts of Maxwell's demon and Kolmogorov/Chaitin complexity. Maxwell proposed his demon in 1871. The demon makes measurements of molecules passing between two chambers and allows only fast molecules to pass into one of the chambers. As a result, one chamber becomes hotter, ostensibly without the expenditure of any significant work. In 1961 Landauer [7] resolved the paradox posed by Maxwell's demon by showing that the storage of information requires work, even if the measurement process itself can be achieved without work. Thus information is something physical and must conform to physical laws.

Life exhibits a high level of complexity. Both Kolmogorov and Chaitin [8] independently developed an algorithmically based definition of complexity. The complexity of a string of symbols can be defined in terms of the shortest program that, when fed to a universal Turing machine, generates that string. Random strings cannot be produced by programs that are any shorter in length than the strings themselves. In contrast, strings with much inherent structure can be compressed into very small programs that will regenerate those strings. Adami points out that the *algorithmic* complexity of a string (e.g., a genetic code) might not be the best measure for a string's *physical* complexity because, in the algorithmically based measure, the most complex strings are random strings while, physically, random strings are not very complicated. Adami goes on to calculate the physical complexity of ensembles of strings in terms of the entropy of the ensemble. He points out that evolution can be viewed as Turing machines computing on strings and attempting to extract from random parts of the string those parts that refer to environment in which the Turing machines reside. Thus these machines are behaving somewhat like Maxwell's demon, which is measuring the environment also in an attempt to reduce entropy by selecting the faster moving molecules within the environment. This is similar to information entering a genome but not being allowed to escape it.

Adami argues that the complexity of a string is conditional on its environment because the environment is needed to determine the mutual entropy. Thus the best measure of complexity is that of *mutual* Kolmogorov complexity, which is the length of a string s minus the length of the smallest program that can compute s. As an example Adami then goes on to apply these concepts by calculating the complexity of tRNA, the molecule that translates a DNA sequence into its corresponding protein.

In Chapter 6 Adami introduces the notion of self-organized criticality (SOC), which is applied in areas of physics concerned with phase-transitions in condensed matter systems. The classical model for SOC is the sandpile [9]. Grains of sand are dropped anywhere on a sandpile. As the grains accumulate, sometimes the height of the pile rises while other times the dropping of a single grain will cause a phase-transition (an avalanche) to occur. Small avalanches will occur more commonly than very large avalanches, which tend to be rarer events. Although the size of a particular avalanche may not be predictable, the relative occurrence of different-sized avalanches follows what is known as a power law. Adami goes on to explain the three main types of power laws that arise in physical systems. For instance, one type of power law, such as the Gutenberg–Richter law concerning earthquakes, relates the size of an event (the magnitude of the earthquake) to its frequency.

After examining SOC in sandpile avalanches on a 2-D lattice, Adami proposes a set of requirements that must be met for a system to exhibit SOC. This list is neither necessary nor sufficient because "as of yet, there is no universal theory of SOC" (p. 146). Adami's requirements include that the system be a "dissipative dynamical system with (locally) interacting degrees of freedom." Adami assumes that readers are already familiar with notions such as "dissipative systems" and "dissipative transport equations."

The Tierra system exhibits long periods of stasis followed by rapid evolution of novel organisms (similar to avalanches). Adami discusses whether or not Tierra can be characterized as exhibiting SOC because SOC is apparently associated more "with second-order versus first-order critical phenomena" (p. 152). (Again, Adami assumes that the reader is already familiar with the difference between first-order and second-order phase-transition dynamics.) If Tierra is a self-organized critical system, then all sizes of evolutionary advances in the model can be explained via the concept of SOC, even though each individual major evolutionary event will be unpredictable.

In Chapter 7 Adami introduces the notion of percolation, which seems to be a process by which some property spreads throughout a medium. Percolation has been studied in the abstract, for example, by placing dots randomly in unfilled areas on a 2-D lattice with some probability between zero and one. Dots near each other are then linked. Percolation theory then studies the types of clusters (sets of connected dots) that form on different types of lattices. For example, an *infinite cluster* is one that spans a lattice by connecting one edge to another. "If such a cluster exists, the system is said to *percolate*...The appearance of an infinite cluster changes the properties of the system drastically" (p. 177). For instance, if connected dots represent electrical conductors, then percolation suddenly allows electric current to flow. Thus, percolation represents a geometrical phase transition. Percolation theory is useful in studying the connectedness of any space. In genetics, percolation theory can be used to examine what the probability is that a random genome might support self-replication, or what the probability is that some evolutionary path connects two genomes. Adami goes on to develop mathematics for percolation on 1-D, 2-D, and high-dimensional Euclidean lattices and on also what is called a Bethe lattice, which is geometrically a tree. Adami speculates that "evolution is much like percolation on a Bethe lattice, where the connectivity of the lattice is controlled by the length of the genome" (p. 193).

In Chapter 8, Adami introduces mathematical techniques to characterize fitness landscapes, by extending the theory of percolation. A Lyapunov function is used, which "provides a ranking between sites that determines their occupation probability according to some functional value" (p. 200). Landscapes Adami examines include Derrida (random energy) landscapes, Kauffman's $N-k$ landscapes [10], fractal landscapes, and RNA landscapes.

Chapter 9 concerns experiments with the avida system. Adami states that avida "implements a simple artificial chemistry in which the molecules are computer programs, and the chemical action of these molecules is obtained by executing those programs" (p. 225). The experiments described are restricted to a 2-D regular lattice where the boundaries wrap around. Adami shows how avida's adaptation behavior produces curves that display a *Devil's staircase* type of property, which indicates a fractal landscape.

Chapter 10 involves running percolation/diffusion experiments using sanda, which is "a variant of avida designed to run on arbitrarily many parallel processors" (p. 251). Chapter 11 involves more advanced discussions, including evolution in avida that includes cosmic-ray-based mutations and the mathematics of molecular evolution as an Ising model. Again, the reader is assumed to be familiar with the concepts of Ising crystals, spin chains, and spin glass models.

I believe that nonphysicists with find it difficult reading, but this book clearly makes an important contribution to the field of artificial life by bringing together in one text many of the relevant mathematical concepts developed originally within physics and by showing how these concepts can be applied in analyzing complex systems that exhibit storage and diffusion of information, self-organization, criticality, phase-transitions, and adaptation.

References

[1] K. Sims, "Evolving 3D morphology and behavior by competition," *Artif. Life*, vol. 1,no. 4, pp. 353–372, 1994.

[2] G. Theraulaz and E. W. Bonabeau, "Coordination in distributed building," *Science*, vol. 269, vol. 269, pp. 687–688, 1995.

[3] M. C. Wright and G. F. Joyce, "Continuous in vitro evolution of catalytic function," *Science*, vol. 276, pp. 614–617, 1997.

[4] S. Wolfram, "Universality and complexity in cellular automata," *Physica D*, vol. 10, pp. 1–35, 1984.

[5] C. G. Langton, "Studying artificial life with cellular automata," *Physica D*, vol. 22, pp. 135–149, 1986.

[6] T. Ray, "An appproach to the synthesis of life," in *Artificial Life II*, C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, Eds. Reading, MA: Addison-Wesley, 1992, pp. 391–408.

[7] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM J. Res. Develp.*, vol. 5, pp. 183–191, 1961.

[8] G. J. Chaitin, "On the length of programs for computing finite binary sequences," *J. ACM*, vol. 13, pp. 547–569, 1966.

[9] pp. Bak, *How Nature Works: The Science of Self-Organized Criticality*. New York: Springer-Verlag, 1996.

[10] S. A. Kaufman, *The Origins of Order*. London, U.K.: Oxford Univ. Press, 1993.