To appear in Luciano Floridi, ed.,
Blackwell Guide to the Philosophy of Computing and Information

Artificial Life

Mark Bedau
Department of Philosophy
Reed College
Portland OR 97202
USA
mab@reed.edu
http://www.reed.edu/~mab

Artificial life (also known as "ALife") is a broad, interdisciplinary endeavor that studies life and life-like processes through simulation and synthesis. The goals of this activity include modelling and even creating life and life-like systems, as well as developing practical applications using intuitions and methods taken from living systems. Artificial life both illuminates traditional philosophical questions and raises new philosophical questions. Since both artificial life and philosophy investigate the essential nature of certain fundamental aspects of reality like life and adaptation, artificial life offers philosophy a new perspective on these phenomena. This chapter provides an introduction to current research in artificial life and explains its philosophical implications.

The Roots of Artificial Life

The phrase "artificial life" was coined by Christopher Langton. He envisioned a study of life as it could be in any possible setting, and he organized the first conference that explicitly recognized this field (Langton 1989). There has since been a regular series of conferences on artificial life and a number of academic journals have been launched to publish work in this new field.

Artificial life has broad intellectual roots, and shares many of its central concepts with other, older disciplines: computer science, cybernetics, biology, complex systems theory and artificial intelligence, both symbolic and connectionist (on the topics see chapters 3, 9 and 14).

John von Neumann (von Neumann 1966) implemented the first artificial life model (without referring to it as such), with his famous creation of a self-reproducing, computation-universal entity, using cellular automata (see Glossary). Von Neumann was trying to understand some of the fundamental properties of living systems, such as self-reproduction and the evolution of complex adaptive structures. His approach was to construct simple formal systems that exhibited those properties. This constructive and abstract methodology typifies contemporary artificial life, and cellular automata are still widely used in the field.

At about the same time, cybernetics (Wiener 1948) applied two new tools to the study of living systems: information theory and the analysis of self-regulatory processes (homeostasis). One of the characteristics of living systems is their spontaneous self-regulation: their capacity to maintain an internal equilibrium in the face of changes in the external environment. This capacity is still a subject of investigation in artificial life. Information theory concerns the transmission of signals independently of their physical representation. The abstract and material-independent approach of information theory is characteristic of artificial life.

Biology's contribution to artificial life include a wealth of information about the life forms found on Earth. Artificial life seeks to understand all forms of life that could exist anywhere in the universe, and detailed information about life on Earth is one good clue about this. Biology has also provided artificial life with models that were originally devised to study a specific biological phenomenon. For example, random Boolean networks (discussed below), which were originally devised by Stuart Kauffman as models of gene regulation networks, are now a paradigm of artificial life research.

Physics and mathematics have also had a strong influence on artificial life. One example is the study of cellular automata as exemplars of complex systems (Wolfram 1994). In addition, artificial life's methodology of studying model systems that are simple enough to have broad generality and to permit quantitative analysis was pioneered in statistical mechanics and dynamical systems. For example, the Ising model consists of a lattice of up and down "spins" that have simple local interactions and that are randomly perturbed by "thermal" fluctuations. This model is so abstract that it contains almost none of the detailed internal physical structure of such materials as a cup of water or a bar of iron. Nevertheless, the model provides a precise quantitative description of how liquid water turns into water vapor or a bar of iron loses its magnetization as temperature rises.

Artificial life also has deep roots in artificial intelligence (AI). Living and flourishing in a changing and uncertain environment seems to require at least rudimentary forms of intelligence. Thus, the subject matter of artificial life and AI overlap. Their methodology is also similar, since both study natural phenomena by building computational models. The computational methodology of artificial life is especially close to the connectionist movement that has recently swept through AI and cognitive science.

The Methodology of Artificial Life

The computer model methodology of artificial life has several virtues. The discipline of expressing a model in feasible computer code requires precision and clarity. It also insures that hypothesized mechanisms are feasible. Computer models also facilitate the level of abstraction required of maximally general models of phenomena. The bottom-up architecture of artificial life models creates an additional virtue. Allowing micro-level entities continually to affect the context of their own behavior introduces a realistic complexity that is missing from analytically studied mathematical models. Analytically solvable mathematical models can reveal little about the global effects that emerge from a web of simultaneous nonlinear interactions. The obvious way to study the effects of these interactions is to build bottom-up models and then empirically investigate their emergent global behavior through computer simulations.

There is an important difference between the modeling strategies AI and ALife typically employ. Most traditional AI models are top-down-specified serial systems involving a complicated, centralized controller that makes decisions based on access to all aspects of global state. The controller's decisions have the potential to affect directly any aspect of the whole system. On the other hand, many natural living systems exhibiting complex autonomous behavior are parallel, distributed networks of relatively simple low-level "agents" that simultaneously interact with each other. Each agent's decisions is based on information about only its own local situation, and its decisions directly affect only its own local situation. ALife's models characteristically follow nature's example. The models themselves are bottom-up-specified parallel systems of simple agents interacting locally. The models are repeatedly iterated and the resulting global behavior is observed. Such lower-level models are sometimes said to be "agent-based" or "individual-based." The whole system's behavior is represented only indirectly. It arises out of the interactions of a collection of directly represented parts ("agents" or "individuals"). Two ALife models illustrating this pattern are described in this section below.

The parallel, distributed character of ALife models is similar to the structure of the models studied in the connectionist (parallel distributed processing, neural network) movement. Both involve bottom-up models in which a population of autonomous agents follows simple local rules. In fact, the agents in many artificial life models are themselves controlled by internal connectionist nets. But there are at least three important differences between typical artificial life models and the connectionist models that have attracted the most attention, such as feed-forward networks that learn by the back-propagation algorithm.

- First, artificial life and connectionism depend on different kinds of learning algorithms. Connectionist models often employ supervised learning algorithms like back-propagation. These learning algorithms are typically turned on when the network is learning and then turned off when the acquired information is applied. This distinction between training and application phases is sometimes unnatural. In addition, supervised learning algorithms require an omniscient teacher, which is also often unnatural. By contrast, the learning algorithms employed in artificial life models usually avoid these criticisms. They are typically unsupervised and in continual operation. Often the algorithm is simply natural selection.
- Second, human intervention and interpretation play different roles in artificial life and connectionism. Typical connectionist models passively receive sensory information prepackaged by a human designer and produce output that must be interpreted by a human designer. In artificial life models, on the other hand, a micro-level agent's sensory input comes directly from the environment in which the agent lives. In many cases, this environment is itself part of the computer model. A human designer originally creates the model, of course, but the specific way it impinges on the agents in typically the result of an unpredictable collection of low-level interactions in the model. In ALife models the micro-level agents' output is to perform actions in their environment, and those actions have direct consequences for the agents' well being. Thus their output has an intrinsic meaning regardless of human interpretation.
- Third, artificial life and connectionism typically seek different kinds of dynamical behavior. Much connectionist modeling aims to produce behavior that settles into an equilibrium. This is because both learning and

> applying knowledge are conceived as fixed and determinate goals. By contrast, artificial life views much of the distinctive behavior of living systems as a process of continual creative evolution, so the aim of many ALife models is an open-ended evolutionary dynamic that is forever far from equilibrium.

The biological world is often viewed as a nested hierarchy of levels. These levels include (among other things) chemicals, organelles, cells, organs, organisms, and ecologies. Artificial life models usually explicitly represent one level with the aim of generating the characteristic phenomena of a higher level. One of the ambitious goals of artificial life is the search for a single model that generates the behavior of all these levels from the explicit specification of only the lowest level. So far, the field has had difficulty producing a model that generates even two levels of emergent phenomena.

The most primitive phenomenon explored by some artificial life models is self-organization. Such models study how structure can emerge from unstructured ensembles of initial conditions, such as models of chemical soups in which fundamental structures such as self-maintaining autocatalytic networks might be seen to emerge. A host of models target the organismic level, sometimes with significant interactions between organisms. These models typically allow changes in the organisms as part of the system's dynamics (e.g., through a genetic mechanism). The most common goal of research using these models is to identify and elucidate structure that emerges in the ensuing evolutionary process. Some models fit in between the chemical level and the organismic level, aiming to understand development by modeling interacting cells. Other models are inter-organismic, in the sense that they aim explicitly to model interactions between different types of organisms or agents. These models often contain elements of game theory.

Many artificial life models are designed not to represent known biological systems but to generate wholly new and extremely simple instances of life-like phenomena. The simplest example of such a system is the so-called "Game of Life", devised by the mathematician John Conway in the 1960s (Berlekamp et al. 1982). Conway's Game of Life can be thought of as a model at the physical or chemical level, embodying an extremely simple and unique form of "chemical" interactions. However, the self-organization exhibited in the Game of Life is not a representation of chemical self-organization in the real world but a wholly new instance of this phenomenon. The Game of Life is a two-state two-dimensional cellular automaton with a trivial nearest-neighbor rule. Think of this "game" as taking place on a two-dimensional rectangular grid of cells, analogous to a huge checker-board. Time advances in discrete steps, and a cell's state at a given time is determined by the states of its eight neighboring cells according to the following simple "birth-death" rule: a "dead" cell becomes "alive" if and only if exactly 3 neighbors are just "alive," and a "living" cell "dies" if and only if fewer than 2 or more than 3 neighbors are just "alive." From inspection of the birth-death rule, nothing particular can be discerned regarding how the whole system will behave. But when the system is simulated, a rich variety of complicated dynamics can be observed and a complex zoo of structures can be identified and classified (blinkers, gliders, glider guns, logic switching circuits, etc.). It is even possible to construct a universal Turing machine in the Game of Life, by cunningly positioning the initial configuration of living cells. In such constructions gliders perform a role of passing signals. Analysing the computational potential of cellular automata on the basis of glider interactions has become a major research thrust.

An example of an organismic level artificial life system is Tierra (Ray 1992). This ALife system consists of “organisms” that are actually simple, self-replicating computer programs populating an environment consisting of computer memory and consuming CPU time as a resource. A Tierran genotype consists of a string of machine code, and each Tierran creature is a token of a Tierran genotype. A simulation starts when computer memory is inoculated with a single self-replicating program, the ancestor, which is then left to self-replicate on its own. The ancestor and its descendants repeatedly replicate, until the available memory space is teeming with creatures that all share the same ancestral genotype. To create space in memory for new descendants, older creatures are continually removed from the system. Errors (mutations) sometimes occur when a creature replicates, so the population of Tierra creatures evolves by natural selection. If a mutation allows a creature to replicate faster, that genotype tends to take over the population. Over time, the ecology of Tierran genotypes becomes remarkably diverse. Quickly reproducing parasites that exploit a host’s genetic code evolve, and this prompts the evolution of new creatures that resist the parasites. After millions of CPU cycles, Tierra typically contains many kinds of creatures exhibiting a variety of competitive and cooperative ecological relationships.

Computer simulation is crucial for the study of complex adaptive systems. It plays the role that observation and experiment play in more conventional science. The complex self-organizing behavior of the Game of Life would never have been discovered without simulating thousands of generations for millions of sites. Similarly, it would have been impossible to discover the emergence of complex ecological interactions in Tierra without simulating many millions of generations. Simulation of large-scale complex systems is the single most crucial development that has enabled the field of artificial life to flourish and distinguish itself from precursors such as cybernetics.

Rather than merely producing computer simulations, some artificial life research aims to implement systems in the real world. The products of this activity are physical devices such as robots that exhibit characteristic life-like behavior. Some of these implementations are motivated by the concern to engineer practical devices that have some of the useful features of living systems, such as robustness, flexibility, and autonomy. But some of this activity is primarily theoretical, motivated by the belief that the best way to confront the hard questions about how life occurs in the physical world is to study real physical systems. Again, there is an analogy with biological levels. The “chemical” level is represented by work on evolvable hardware, often using programmable logic arrays, which attempts to use biologically-inspired adaptive processes to shape the configuration of micro-electronic circuitry. The “organismic” level is represented by new directions in biologically-inspired robotics, such as using evolutionary algorithms to automate the design of robotic controllers. A swarm of robots communicating locally to achieve some collective goal is an example at the “population” level. An “ecological” level might be represented by the Internet along with its interactions with all its users on computers distributed around the world.

Emergence

Both living systems and artificial life models are commonly said to exhibit emergent phenomena; indeed, many consider emergence to be a defining feature of life. However, the notion of emergence remains ill defined. In general, emergent phenomena share two broad hallmarks: they are constituted by and generated from underlying phenomena, and yet they are also autonomous from those underlying phenomena. There are abundant examples of apparent emergent phenomena, and most involve life or mind. Yet the two hallmarks of emergence seem inconsistent or metaphysically illegitimate: How can something be autonomous from underlying phenomena if it is constituted by and generated from them? This is the problem of emergence. A solution would both dissolve the appearance of illegitimate metaphysics and enfold emergence in constructive scientific explanations of phenomena involving life and mind.

One can distinguish emergent properties, emergent entities, and emergent phenomena. Being alive, for example, is an emergent property, an organism is an emergent entity, and the life history of an organism is an emergent phenomenon. An entity with an emergent property is an emergent entity, and an emergent phenomenon involves an emergent entity possessing an emergent property. So the first step toward solving the problem of emergence is to explain the notion of an emergent property. There are three main views of what an emergent property is.

According to the first view, emergent properties apply only to "wholes" or "totalities", not to their component "parts" considered in isolation (e.g., Harré 1985, Baas 1994). For example, the constituent molecules in a cup of water, considered individually, do not have properties like fluidity or transparency, though these properties do apply to the whole cup of water. The "wholes" at one level of analysis are sometimes "parts" of a larger "whole" at a higher level of analysis, so a hierarchy can contain successive levels of this sort of emergence. This view easily explains the two hallmarks of emergence. Macro-level emergent phenomena are constituted from and generated by micro-level phenomena in the trivial sense that wholes are constituted and generated by their constituents; and emergent phenomena are autonomous from underlying phenomena in the straightforward sense that emergent properties do not apply to the underlying entities. This notion of emergence is very broad, applies to a large number of intuitive examples of emergent phenomena, and corresponds to the compelling picture of reality consisting of autonomous levels of phenomena. Its breadth is its greatest weakness, however, for it applies to all macro-properties that are not possessed by micro-entities. Macro-properties are usually classified into two kinds: genuine emergent properties and mere "resultant" properties. Resultant properties are those that can be predicted and explained from the properties of the components. For example, a circle consists of a collection of points, and the individual points have no shape. So being a circle is a property of a "whole" but not its constituent "parts." Thus being a circle is an emergent property according to the first view. However, if you know that all the points in a geometrical figure are equidistant from a given point, then you can conclude that the figure is a circle. So being a circle is a resultant property. To distinguish emergent from resultant properties one must turn to other views.

The second main view construes emergent properties as supervenient properties with causal powers that are irreducible to the causal powers of micro-level constituents (e.g., Kim 1999). On this view, supervenience explains the sense in which the underlying processes constitute and generate the emergent phenomena, and irreducible causal powers explain the sense in which they are autonomous from underlying phenomena. These irreducible causal powers give emergent properties a dramatic form of ontological novelty that many people associate with the most puzzling kinds of emergent phenomena, such as consciousness. However, an irreducible but supervenient causal power by definition cannot be explained in terms of the aggregation of the micro-level potentialities. No evident mechanism explains these irreducible supervenient powers, so they must be viewed as primitive or "brute" facts of nature. In addition, this strong form of emergence seems to be scientifically irrelevant. Illustrations of it in recent scientific literature almost universally focus on one isolated example: Sperry's explanation of consciousness from over thirty years ago (Sperry 1969). There is little if any evidence that this form of emergence is empirically relevant in the sciences studying emergent phenomena.

A third view of emergence is poised midway between the first two. It refers to the resultant aggregate global behavior of complex systems. In this sense, a system's macrostate is emergent just in case it can be derived from the system's boundary conditions and its micro-level dynamical process but only through the process of iterating and aggregating all the micro-level effects (e.g., Bedau 1997a). In this case, the micro-level phenomena clearly constitute and generate the macro-level phenomena. At the same time, the macro-level phenomena are autonomous in that the only way to recognize or predict them is by empirically observing the macro-level effect of aggregating all the micro-level phenomena. In effect, this view identifies emergent properties with a special subset of resultant properties: those that cannot be predicted or explained except by empirically aggregating the interactions among micro-level entities. This form of emergence is common in complex systems found in nature. Artificial life's models also exhibit it, since their bottom-up dynamics consists of the continual iteration of micro-level interactions. This view attributes the unpredictability and unexplainability of emergent phenomena to the complex consequences of myriad, non-linear and context-dependent local micro-level interactions. Emergent phenomena can have causal powers on this view, but only by means of aggregating micro-level causal powers. There is nothing inconsistent or metaphysically illegitimate about underlying processes constituting and generating phenomena by iteration and aggregation. Furthermore, this form of emergence is prominent in scientific accounts of exactly the natural phenomena like life and mind that apparently involve emergence. However, this form of emergence sheds no light on those mysterious emergent phenomena, like consciousness, that science still cannot explain. In addition, the autonomy of these kinds of emergent phenomena seems to be merely epistemological rather than ontological. Emergent phenomena are epistemologically autonomous in the sense that knowledge of the underlying phenomena does not provide knowledge about the emergent phenomena. However, metaphysically, the emergent phenomena seem wholly dependent on the constituent phenomena, since emergent causal powers result from micro-level causal powers. This will not satisfy those who think emergent phenomena have a strong form of ontological autonomy.

Artificial life can be expected to play an active role in the future philosophical debate about emergence and related notions like supervenience, reduction, complexity, and hierarchy. Living systems are one of the primary sources of emergent phenomena, and artificial life's bottom-up models generate impressive macro-level phenomena wholly out of micro-level interactions. Exploration and modification of these models is a constructive way to analyze the nature and causes of different kinds of emergent phenomena.

Adaptationism

Adaptive evolutionary explanations are familiar from high school biology. It is a cliché to explain the giraffe's long neck as an adaptation for browsing among the tops of trees, on the grounds that natural selection favored longer-necked giraffes over their shorter-necked cousins. But the scientific legitimacy of these adaptive explanations is controversial, largely because of a classic paper by Stephen Jay Gould and Richard Lewontin (1979). Gould and Lewontin directly challenge *adaptationism*: the thesis that the activity of pursuing adaptive explanations of biological traits is a legitimate part of empirical science. They accept that adaptive explanations are appropriate in some contexts, but they despair of identifying those contexts in any principled and rigorous way. Biology provides many alternatives to adaptive explanations, such as explanations appealing to allometry, genetic drift, developmental constraints, genetic linkage, epistasis, and pleiotropy. But Gould and Lewontin complain that those alternatives receive only lip-service. The presupposition that a trait is an adaptation and so deserves an adaptive explanation is treated as untestable. The fundamental challenge for adaptationism raised by Gould and Lewontin, then, is to find some empirical method for testing when an adaptive explanation is needed. This problem is often especially acute in artificial life. Those studying artificial models have the luxury of being able to collect virtually complete data, but this mass of information only compounds the problem of identifying which evolutionary changes are adaptations.

The canonical response to Gould and Lewontin makes two claims. The first claim is that *specific* adaptive hypotheses, hypotheses about the specific nature of a character's adaptation, are testable. Second, although the *general* hypothesis that a trait is an adaptation might itself not be testable, it is a working hypothesis and empirical science normally treats working hypotheses as untestable. For example, Richard Dawkins claims that "hypotheses about adaptation have shown themselves in practice, over and over again, to be easily testable, by ordinary, mundane methods of science" (Dawkins, 1983, pp. 360f). Dawkins's point is that specific adaptive hypotheses have observable consequences that can be checked. The canonical response reflects and explains evolutionary biology's emphasis on formulating and testing specific adaptive hypotheses. But this response does not address the fundamental challenge to adaptationism, for that challenge is about the testability of *general* adaptive hypotheses, hypotheses to the effect that a trait is an adaptation. Different specific adaptive hypotheses usually have different observable consequences. A general adaptive hypothesis entails that some specific adaptive hypothesis is true, but it gives no indication which one is true. So the general adaptive hypothesis makes no particular empirical prediction. Dawkins admits that general adaptive hypotheses cannot be tested. “It is true that the one hypothesis that we shall never test is the hypothesis of no adaptive function at all, but only because that is the one hypothesis in this whole area that really *is* untestable” (1983, p. 361). Dawkins can defend the appeal to adaptive explanations when a specific adaptive hypothesis has been corroborated. But in the absence this—which is the typical situation—Dawkins must concede Gould’s and Lewontin’s fundamental challenge.

Artificial life has been used to develop and illustrate a new defense of adaptationism. It is argued that it is possible to test general adaptive hypotheses empirically, by recording and analyzing so-called “evolutionary activity” information collected from the evolving system (Bedau 1996, Bedau & Brown 1999). The fundamental intuition behind this method is that we can detect whether an item (gene, gene complex, genotype, etc.) is an adaptation by observing the extent to which it persists in the face of selection pressures. Whenever an item that is subject to heritable variation is “active” or expressed, natural selection has an opportunity to provide feedback about its adaptive value, its costs and benefits. If it persists and spreads through a population when it is repeatedly active, and especially if it exhibits significantly more activity than one would expect to see if it had no adaptive value, then we have positive evidence that the item is persisting *because of* its adaptive value. This means that we have positive evidence that it is an adaptation and deserves an adaptive explanation, even if we have no idea about its specific adaptive function. Since natural selection is not instantaneous, maladaptive items persist for a while before they are driven out by natural selection. Adaptations are distinguished by accruing much more activity than would be expected in a non-adaptive item. A general way to measure the activity expected of non-adaptive items is to construct a “neutral shadow” of the target system—that is, a system that is similar to the target in all relevant respects *except* that none of the items in it have any adaptive significance. The activity in the neutral shadow is a no-adaptation null hypothesis for the target system. If the target system shows significantly more activity than the neutral shadow, this excess activity must be due to natural selection and the target system must contain adaptations. The evolutionary activity method responds directly to Gould and Lewontin. It provides an empirical method for determining when evolution is creating adaptations. Rather than just assuming that traits are adaptations, it puts this assumption to empirical test. Another advantage of the activity method is that statistics based on activity information can be used to measure various aspects of the dynamics of adaptive evolution, thus allowing the process of adaptation in different systems to be classified and quantitatively compared (Bedau et al. 1997, Bedau et al. 1998). One weakness of the evolutionary activity method is that practical problems sometimes make activity data difficult to collect. Another weakness is that genetic hitchhikers—non-adaptive or maladaptive traits that persist because of a genetic connection to an adaptive trait—can accumulate more activity than expected in a neutral shadow. Thus, a trait that is not an adaptation can have significant excess activity if it is connected to a trait that is an adaptation. Significant excess activity in a cluster of traits shows that there are adaptations in the cluster, but it does not separate out the hitchhikers.

The adaptationist perspective on evolution emphasizes natural selection's role in creating the complex adaptive structures found in living systems. Artificial life has been the source of a new and fundamental challenge to this whole perspective. Stuart Kauffman (1993, 1995) has used artificial life models to show that many features of metabolisms, genetic networks, immune systems, and ecological communities should be viewed not as the products of selection but largely as the spontaneous self-organized behaviors of certain abstract complex systems. Kauffman also argues that spontaneous self-organized structures—what he calls "order for free" (Kauffman 1995)—explain both life's origin and its subsequent ability to evolve. Kauffman can make sweeping claims about order for free because the artificial life models he studies are abstract enough to apply to a wide variety of contexts. Random Boolean networks are one such class of models. These consist of a finite collection of binary (ON, OFF) variables with randomly chosen input and output connections. The state of each variable at each step in discrete time is governed by some logical or Boolean function (AND, OR, etc.) of the states of variables that provide input to it. The network is started by randomly assigning states to each variable, and then the connections and functions in the network determine the successive state of each variable. Since the network is finite, it eventually reaches a state it has previously encountered, and from then on the network will forever repeat the same cycle of states. Different network states can end up in the same state cycle, so a state cycle is called an attractor. Kauffman found that the number of variables in the network, the number of connections between the variables, and the character of the Boolean functions determine many biologically crucial properties of the networks. These properties include the number and length of attractors, the stability of attractors to perturbation and mutation, etc. If the variables are highly connected, then the network's attractors contain so many states that the time it takes to traverse the attractor vastly exceeds the lifetime of the entire universe. Furthermore, any perturbation or mutation in the network causes a vast change in its behavior. For all practical purposes, the network behaves chaotically. The network acts differently when each variable takes input from only a biologically plausible number of other variables and when the variables are governed by biologically realistic Boolean functions. In this case, the network has a tiny number of attractors, it maintains homeostatic stability when perturbed, and mutations have limited consequences; in other words it exhibits "order for free." Furthermore, these biologically realistic Boolean networks explain a number of empirically observed features of biological systems, such as how the number of different cells types and cell replication times vary as a function of the number of genes per cell. Kauffman's non-adaptationist explanations of the origins of order are controversial, partly because of the sweeping scope of his analysis. But the suggestion that self-organization rather than natural selection can explain much of the structure in living systems is plausible. The issue is not whether self-organization explains structure, but how much.

The problem of adaptationism is as acute in artificial life as it is in biology. Artificial life can make a distinctive contribution to the debate, for the evolutionary processes studied by artificial life provide many diverse examples of the process of adaptation. Furthermore, the systems can be analyzed with the kind of detail and rigor that is simply impossible to achieve in the biosphere, because the historical data are unavailable or impractical to examine. For analogous reasons, we can expect artificial life to contribute to our understanding of many other fundamental issues in the philosophy of biology, such as the nature of functions, the nature of species, whether and how selection operates at different biological levels, the nature of the niche, and the nature of the relationship between organisms and their environment.

Evolutionary Progress

The evolution of life shows a remarkable growth in complexity. Simple prokaryotic one-celled life lead to more complex eukaryotic single-celled life, which then lead to multicellular life, then to large-bodied vertebrate creatures with complex sensory processing capacities, and ultimately to highly intelligent creatures that use language and develop sophisticated technology. This illustration of evolution's creative potential has led some to propose a ladder of complexity hypothesis according to which open-ended evolutionary processes have an inherent, law-like tendency to create creatures with increasingly complicated adaptive structure. But the evolution of life is equally consistent with the denial of the ladder of complexity. The observed progression could be a contingent result of evolution rather than a reflection of any inherent tendency. The ladder of complexity hypothesis is difficult to test because we do not have a variety of different histories of life to compare. A sample size of one makes it difficult to distinguish inherent trends from artifacts.

Stephen Jay Gould (1989) devised an ideal way to address this issue, namely the thought experiment of replaying the tape of life. Imagine that the process of evolution left a record on a tape. Gould's thought experiment consists in rewinding the evolutionary process backward in time and then replaying it again forward in time but allowing different accidents, different contingencies to reshape the evolution of life. The evolution of life is rife with contingencies. Repeatedly replaying the tape of life with novel contingencies could produce as large a sample of evolutionary histories as desired. It would be relatively straightforward to determine whether a general pattern emerges when all the evolutionary trajectories are compared.

There is substantial controversy about the outcome of Gould's thought experiment. Gould himself suggests that "any replay of the tape would lead evolution down a pathway radically different from the road actually taken" (1989, p. 51). He concludes that the contingency of evolution will debar general laws like the hypothesized ladder of complexity. Daniel Dennett (1995) draws exactly the opposite conclusion. Dennett argues that certain complex features like sophisticated sensory processing provide a distinct adaptive advantage. Thus, natural selection will almost inevitably discover significantly advantageous features that are accessible from multiple evolutionary pathways. Examples of multiple independent evolutionary convergence, such as flight and eyesight, illustrate this argument. Dennett concludes that replaying life's tape will almost inevitably produce highly intelligent creatures that use language and develop sophisticated technology.

Artificial life can make number of contributions to this debate. Experience in artificial life has shown time and again that expectations about the outcome of thought experiments like replaying life's tape are highly fallible. The only sure way to determine what to expect is to create the relevant model and observe the results of repeated simulation. In fact, artificial life is exactly where this sort of modeling activity occurs. A central goal of artificial life is to discover the inherent trends in evolving systems by devising a model of open-ended evolution, repeatedly replaying life's tape with different historical contingencies and searching for patterns that hold across all the results. The best evidence in favor of the ladder of complexity hypothesis would come from showing that a tendency toward increasing adaptive complexity is the norm in such ALife models. However, no one has yet conducted the experiment of replaying life's tape, because no one has yet been able to create a system that exhibits continual open-ended evolution of adaptive complexity. Achieving this goal is one of the key open problems in artificial life (Bedau et al. 2000). All conjectures about the ladder of complexity will remain unsettled until one can actually replay the tape of life.

The Nature of Life

Philosophy traditionally addressed the nature of life but most philosophers ignore the issue today, perhaps because it seems too "scientific." At the same time, most biologists also ignore the issue, perhaps because it seems too "philosophical." The advent of artificial life raises the question anew, for two reasons. Modeling the fundamental features of living systems presupposes an understanding of life, and new artificial life systems push the boundaries of what life could be.

There are three prominent views about the nature of life: life as a cluster of properties, life as metabolization, and life as evolution. The cluster conception takes two forms, depending on whether the properties in the cluster are taken to be individually necessary and jointly sufficient for life. Skeptics argue that life is characterized merely by a loose cluster of properties typically but not necessarily possessed by living entities. This view treats something as alive if it possesses a sufficient number of properties in the cluster, but no precise number of properties is sufficient. On this view, the diversity of living forms have only a family resemblance. Viewing life as a loose cluster of properties provides a natural explanation of why life has vague boundaries and borderline cases. Life is also sometimes characterized by a list of properties intended to provide something much closer to individually necessary and jointly sufficient conditions. Ernst Mayr (1982) produced a comprehensive list of such properties:

1. Living systems have an enormously complex and adaptive organization.
2. Organisms are composed of a chemically unique set of macromolecules.
3. Living phenomena are predominantly qualitative, not quantitative.
4. Living systems consist of highly variable groups of unique individuals.
5. Organisms engage in purposeful activities by means of evolved genetic programs.
6. Classes of organisms have historical connections of common descent.
7. Organisms are the product of natural selection.
8. Biological processes are especially unpredictable.

Cluster conceptions of life account for the characteristic hallmarks of life, although they do this merely by *fiat*. Lists like Mayr's raise rather than answer the question why this striking collection of features is present in an indefinite diversity of natural phenomena. The main drawback of all cluster conceptions is that they inevitably make life seem rather arbitrary or mysterious. A cluster conception cannot explain why any particular cluster of properties is a fundamental and ubiquitous natural phenomenon.

Schrödinger illustrated the second view of life when he proposed persistence in the face of the second law of thermodynamics by means of the process of metabolization as the defining feature of life.

> It is by avoiding the rapid decay into the inert state of "equilibrium" that an organism appears so enigmatic; . . . How does the living organism avoid decay? The obvious answer is: By eating, drinking, breathing and (in the case of plants) assimilating. The technical term is metabolism. (Schrödinger 1969, p. 75)

Living systems need some way to self-maintain their complex internal structure. So metabolization seems to be at least a necessary condition of all physical forms of life. The view that life centrally involves the process of metabolization also nicely explains our intuition that a crystal is not alive. There is a metabolic flux of molecules only at the crystal's edge, not inside it. One drawback of metabolization as an all-encompassing conception of life is that many metabolizing entities seem not to be alive and not to involve life in any way. Standard examples include a candle flame, a vortex, and a convection cell. A second problem is whether metabolization can explain the hallmarks of life (recall Mayr's list). It is doubtful whether metabolization can explain those characteristics on Mayr's list that depend on evolution.

The third main conception of life focuses on the evolutionary process of adaptation. The central idea is that what is distinctive of life is the way in which adaptive evolution automatically fashions new and intelligent strategies for surviving and flourishing as local environments change. As John Maynard Smith explains:

> We shall regard as alive any population of entities which has the properties of multiplication, heredity and variation. The justification for this definition is as follows: any population with these properties will evolve by natural selection so as to become better adapted to its environment. Given time, any degree of adaptive complexity can be generated by natural selection. (Maynard Smith 1975, p. 96f)

The view of life as evolution has two forms. Maynard Smith illustrates one form, according to which living systems are the entities in an evolving population. Recently, Bedau (1996, 1998a) has argued that, in fact, an evolving system itself should be viewed as alive in the primary sense. One virtue of the conception of life as evolution is that it explains why Mayr's hallmarks of life coexist in nature. We would expect life to involve the operation of natural selection producing complex adaptive organization in historically connected organisms with evolved genetic programs. The random variation and historical contingency in the evolutionary process explains why living phenomena are especially qualitative and unpredictable and involve unique and variable individuals with frozen accidents like chemically unique macromolecules. This view can also explain why metabolism is so important in living systems, for a metabolism is a physically necessary prerequisite in any system that can sustain itself long enough to adapt and evolve. There are two main objections to viewing life as evolution. The first is that it seems to be entirely contingent that life forms were produced by an evolutionary process. The Biblical story of Adam and Eve shows that is easy to imagine life forms in the absence of any evolutionary process. A second objection calls attention to evolving systems that seem devoid of life. Viruses and prions evolve but are questionably alive, and cultural evolution provides much starker counterexamples.

The advent of artificial life has revitalized investigation into the nature of life. This is partly because one can simulate or synthesize living systems only if one has some idea what life essentially is. Artificial life's self-conscious aim to discern the general nature of life as it could be encourages liberal experimentation with novel life-like organizations and processes. Thus, artificial life both fosters a broad perspective on life and has the potential to create radically new forms of life. In the final analysis, the nature of life will be settled by whatever provides the best explanation of the rich range of natural phenomena that seem to characterize living systems. Better understanding of how to explain these phenomena will also help resolve a cluster of puzzles about life. These puzzles include whether life admits of degrees, how the notion of life applies at different levels in the biological hierarchy, whether life is essentially connected with mental capacities, and the relationship between the material embodiment of life and the dynamical processes in those materials.

Strong Artificial Life

Artificial life naturally raises the question whether artificial constructions could ever literally be alive. Agreement about the nature of life would make this question easier to answer. For example, if the defining property of living systems were the process of sustaining a complex internal organization through a metabolism, then the issue would be whether an artificially created system could literally exhibit this property (see Boden 1999 for discussion). But the debate over creating real but artificial life currently proceeds in the absence of agreement about what life is.

It is important to distinguish two questions about creating artificial life. The first concerns whether it is possible to create a physical device such as a robot that is literally alive. Aside from controversy about what life is, the challenge here is less philosophical than scientific. It concerns our ability to synthesize the appropriate materials and processes. The philosophically controversial question is whether the processes or entities inside a computer that is running an artificial life model could ever literally be alive. This is the issue of whether so-called "strong" artificial life is possible. Strong ALife is contrasted with "weak" ALife, the uncontroversial thesis that computer models are useful for understanding living systems.

The strong ALife question is sometimes put in terms of computer simulations: can a computer simulation of a living system ever literally be alive? This formulation prompts the response (e.g., Pattee 1989, Harnad 1994) that it is a simple category mistake to confuse a *simulation* of something with a *realization* of it. A flight simulation for an airplane, no matter how detailed and realistic, does not really fly. A simulation of a hurricane does not create real rain driven by real gale-force winds. Similarly, a computer simulation of a living system produces merely a symbolic representation of the living system. The intrinsic ontological status of this symbolic representation is nothing more than certain electronic states inside the computer (e.g., patterns of high and low voltages), and this constellation of electronic states is no more alive than is a series of English sentences describing an organism. It seems alive only when it is given an appropriate interpretation. This interpretation might be fostered if the description dynamically reflects how the living system changes over time and if the simulation produces a vivid life-like visualization, but it is still only an interpretation.

A number of considerations can blunt this charge of category mistake. It is important to recognize that an artificial life model that is actually running on a computer consists of a real physical process occurring in a real physical medium consuming real physical resources. The software specifying the model might be a static abstract entity with the ontological nature of a Platonic universal, but an actual simulation of the model has the ontological status of any physical process. Furthermore, as emphasized earlier, artificial life models are often intended not as simulations or models of some real-world living system but as novel examples of living systems. Conway's Game of Life (Berlekamp et al. 1982), for example, is not a simulation or model of any real biochemical system. Rather, it is simple system that exhibits spontaneous macroscopic self-organization. Similarly, Ray's Tierra (Ray 1992) is not a simulation or model of the ecology and evolution of some real biological system. Instead, it is an instance of ecological and evolutionary dynamics in a digital domain. So, when the Game of Life and Tierra are actually running in computers, they are new physical instances of self-organization and evolution. Processes like self-organization and evolution are multiply realizable and can be embodied in a wide variety of different media, including the physical media of suitably programmed computers. So, to the extent that the essential properties of living systems involve processes like self-organization and evolution, suitably programmed computers will actually be novel realizations of life. Models that merely represent some phenomenon differ from models that actually generate it. For example, a two-dimensional model of a branching process with random pruning can be viewed as a description of the evolution of more or less complex insects, if one dimension is taken to represent time and the other is taken to represent complexity. But exactly the same branching process can equally be viewed as a description of the evolution of more or less tall humans. It can even be viewed as a description of various non-temporal and non-biological processes, such as the pattern of tributaries in a geography. In itself, the model does not intrinsically involve any of these things. By contrast, a glider in Conway's Game of Life is not an electronic pattern that is merely interpretable as a self-sustaining dynamic collective. It really *is* an electronic self-sustaining collective, whether or not anyone notices it and regards it as such. Likewise, the self-replicating machine-language programs in Ray's Tierra genuinely evolve by natural selection and genuinely engage in host/parasite relations. The nature of ALife's key problem of modeling the open-ended evolution of adaptive complexity can be appreciated in this light. It is easy to make a model that can be interpreted as exhibiting this phenomenon; the challenge is to make a model that actually generates it.

The Turing test in artificial intelligence was an attempt to settle whether computing could be indistinguishable from thinking in the absence of any agreement about the nature of thinking itself. Thus the proposal to settle the strong ALife debate with a “Turing test” for life often arises in artificial life. Some (e.g., Sober 1992) warn that the Turing test in AI is an insufficient test for intelligence because it is possible in principle for an unthinking device to pass the test. A typical example of such a hypothetical device is a machine that stores an appropriate output for all the different input that might be encountered. The characteristic drawback of such devices is that, even to exhibit modest capabilities, the number of pieces of information they must store is larger than the number of elementary particles in the entire universe. Though possible in principle, such a device is clearly impossible in practice. Artificial life’s computational methodology demands models that actually produce the phenomenon of interest. In this context, what is possible in principle but impossible in practice is irrelevant. So the experience in ALife prompts one to ignore unfeasible counterexamples to Turing tests. Harnad (1994) has advocated ecological and evolutionary indistinguishability from biological life as a Turing test for life. The motivation for this test for life is that it would be arbitrary to deny life to anything that is indistinguishable ecologically and evolutionarily from biological life. But this test is biased against life forms that are isolated from the biosphere. Systems existing inside computers running artificial life models might exhibit all the ecological and evolutionary richness found in the biosphere. Yet they might not interact with biological life, so they might fail Harnad’s test for life. Thus, Harnad’s test begs the question against some forms of artificial life.

The debate about strong artificial life is intertwined with philosophical questions about functionalism and computation. A significant source of support for strong ALife is the belief that life concerns form more than matter. Although certain carbon-based macromolecules play a crucial role in the vital processes of all known living entities, metabolization creates a continual flux of molecules through living systems. Thus, life seems more like a kind of a process than a kind of material entity. This implies that life could be realized in a variety of media, perhaps including suitably programmed computer hardware. This motivation for strong ALife prompts a functionalist and computationalist view of life, analogous to contemporary functionalism and computationalism with respect to mind. Sober (1992) points out that many essential properties of organisms involve their interaction with the environment. Thus, the computational character of the processes inside organisms would not alone support functionalism and computationalism about life. But since many artificial life models situate artificial organisms in an artificial environment, artificial life still promotes functionalism and computationalism. Bedau (1997b) argues that artificial life’s models generate macro-level dynamics with a suppleness that is distinctive of adaptive intelligence and that cannot be captured by any fixed algorithm. The models are implemented in a computer but adaptive processes like natural selection continually change the micro-level rules that govern the system. Thus, the macro-level processes that emerge are non-computational. This perspective still supports functionalism with respect to life, but a form of functionalism divorced from computationalism.

Artificial life models generate behavior that is characteristic of living systems, so the practice of artificial life will continually raise the question whether a computer model of life could literally be alive. By continually challenging the boundaries between life and non-life, artificial life will also spur novel perspectives on the issue. The debate about strong ALife will also enliven and inform many related issues in the philosophy of mind and artificial intelligence, including functionalism, computationalism, intelligence, intentionality, and representationalism.

Philosophical Methodology

Artificial life also has implications for the methodology of philosophy. Philosophy and artificial life are natural partners. Both seek to understand phenomena at a level of generality that is sufficiently deep to ignore contingencies and reveal essential natures. In addition, artificial life's computational methodology is a direct and natural extension of philosophy's traditional methodology of *a priori* thought experiments. In the attempt to capture the simple essence of vital processes, artificial life models abstract away as many details of natural living as possible. These models are for exploring the consequences of certain simple ideas or premises. They are "thought experiments" explored with the help of a computer. Like the traditional armchair thought experiments employed in philosophy, artificial life simulations attempt to answer "What if X?" questions. Artificial life's thought experiments are distinctive in that they can be explored only by computer simulation; armchair analysis is simply inconclusive. Synthesizing thought experiments on a computer can bring a new clarity and constructive evidence to bear in philosophy (see chapter 26).

References

Baas, N. A. (1994). Emergence, hierarchies, and hyperstructures. In C.G. Langton, ed., *Artificial Life III* (pp. 515-537). Redwood City, CA: Addison-Wesley. A mathematical and technical presentation and illustration of the view of emergent properties as novel macro-properties.

Bedau, M. A. (1996). The nature of life. In M. Boden (Ed). *The Philosophy of Artificial Life* (pp. 332-357). Oxford: Oxford University Press. A defense of the view of life as supple adaptation or open-ended evolution, illustrated in artificial life models. For post-secondary school audiences.

Bedau, M. A. (1997a). Weak emergence. *Philosophical Perspectives* 11, 375-399. A defense of emergence as complicated iteration and aggregation of micro-level interactions, illustrated in artificial life models. For post-secondary school audiences.

Bedau, M. A. (1997b). Emergent models of supple dynamics in life and mind. *Brain and Cognition* 34, 5-27. Describes a characteristic suppleness of the dynamics of mental states, argues that artificial life models capture this kind of dynamics, and draws out the implications for functionalism about life. For post-secondary school audiences.

Bedau, M. A. (1998a). Four puzzles about life. *Artificial Life* 4, 125-140. An explanation of how the view of life as supple adaptation (Bedau 1996) explains four puzzles about life. For post-secondary school audiences.

Bedau, M. A., and Brown, C. T. (1999). Visualizing evolutionary activity of genotypes. *Artificial Life* 5, 17-35. Shows how evolutionary activity graphs reveal the dynamics of adaptive evolution of genotypes in an artificial life model. For post-secondary school audiences.

Bedau, M. A., McCaskill, J. S., Packard, N. H., Rasmussen, S., Adami, C., Green, D. G., Ikegami, T., Kaneko, K., & Ray, T. S. (2000). Open problems in artificial life. *Artificial Life* 6, 363-376. Describes fourteen grand challenges in artificial life, each of which requires a major advance on a fundamental issue for its solution. Intelligible to secondary school audiences.

Bedau, M. A., Snyder, E., Brown, C. T., & Packard, N. H. (1997). A comparison of evolutionary activity in artificial evolving systems and the biosphere. In P. Husbands & I. Harvey (Eds). *Proceedings of the Fourth European Conference on Artificial Life, ECAL97* (pp. 125-134). Cambridge: MIT Press. Comparison of evolutionary activity in two artificial life models and in the fossil record reveals qualitatively differences. For post-graduate audiences.

Bedau, M. A., Snyder, E., & Packard, N. H. (1998). A classification of long-term evolutionary dynamics. In C. Adami, R. Belew, H. Kitano, and C. Taylor, eds., *Artificial Life VI* (pp. 228-237). Cambridge: MIT Press. Evolutionary activity is used to classify qualitatively different kinds of evolving systems. For post-graduate audiences.

Berlekamp, E. R., Conway, J. H., & Guy, R. K. (1982). *Winning ways for your mathematical plays* (Vol. 2): *Games in particular*. New York: Academic Press. Chapter 25 is the authoritative description of the Game of Life, including a description of how to embed a universal Turing machine in the Game of Life. For a general audience.

Boden, M. A. (1999). Is metabolism necessary? *British Journal of the Philosophy of Science* 50, 231-248. Distinguishes three senses of metabolism and examines their implications for strong artificial life. For post-secondary school audiences.

Dawkins, R. D. (1983). Adaptationism was always predictive and needed no defense. *Behavioral and Brain Sciences* 6, 360-361. A defense of adaptationism. For post-secondary school audiences.

Dennett, D. C. (1995). *Darwin's dangerous idea: evolution and the meanings of life*. New York: Simon and Schuster. An extended essay on how natural selection transforms our view of humanity's place in the universe. Accessible to a general audience.

Gould, S. J. (1989). *Wonderful life: the Burgess shale and the nature of history*. New York: Norton. A defense of radical contingency in the evolution of life, in the context of detailed examination of fossils in the Burgess shale. Accessible to a general audience.

Gould, S. J., and Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society B* 205, 581-598. The classic criticism of the use of adaptive explanations in biology. For post-secondary school audiences.

Harnad, S. (1994). Levels of functional equivalence in reverse bioengineering. *Artificial Life* 1, 293-301. Distinguishes synthetic and virtual artificial life, argues that virtual artificial life is impossible, and suggests a Turing test for life to settle whether synthetic artificial life is possible. For post-secondary school audiences.

Harré, Rom. 1985. *The philosophies of science*. Oxford: Oxford University Press. An introduction to the philosophy of science, for secondary and post-secondary school audiences.

Kauffman, S. A. (1993). *The origins of order: self-organization and selection in evolution*. New York: Oxford University Press. A technical presentation of Kauffman's views of the place of spontaneous order in living systems. For post-graduate audiences.

Kauffman, S. A. (1995). *At home in the universe: the search for the laws of self-organization and complexity*. New York: Oxford University Press. Kauffman's views presented to a general audience.

Kim, J. (1999). Making sense of emergence. *Philosophical Studies* 95, 3-36. An investigation of the credibility of emergence with special reference to the philosophy of mind. For post-secondary school audiences.

Langton, C. G. (Ed). (1989). *Artificial life: the proceedings of an interdisciplinary workshop on the synthesis and simulation of living systems*. Redwood City:

Addison-Wesley. Proceedings of the first "artificial life" conference identified as such. The editor's introduction (expanded in Boden 1996) is a classic introductory overview of the field. Includes 25 technical papers and a 40 page annotated bibliography of works relevant to artificial life. For a general audience.

Maynard Smith, J. (1975). *The theory of evolution*, 3rd edition. New York: Penguin. A classic, for a general audience.

Mayr, E. (1982). *The growth of biological thought: diversity, evolution, and inheritance*. Cambridge: Harvard University Press. A history of the ideas in biology, for post-secondary school audiences.

Pattee, H. H. (1989). Simulations, realizations, and theories of life. In C. G. Langton (Ed). *Artificial life: the proceedings of an interdisciplinary workshop on the synthesis and simulation of living systems* (pp. 63-78). Redwood City: Addison-Wesley. A criticism of strong artificial life, reprinted in Boden (1996). For post-secondary school audiences.

Ray, T. S. (1992). An approach to the synthesis of life. In C. Langton, C. Taylor , D. Farmer, & S. Rasmussen (Eds). *Artificial Life II* (pp. 371-408). Redwood City: Addison-Wesley. The classic discussion of one of the best-known artificial life models. For post-secondary school audiences. Reprinted in Boden (1996).

Schrödinger, E. (1969). *What is life?* Cambridge: Cambridge University Press. The classic statement of the metabolic view of life. For a general audience.

Sober, E. (1992). Learning from functionalism—prospects for strong artificial life. In C. Langton, C. Taylor , D. Farmer, & S. Rasmussen (Eds). *Artificial Life II* (pp. 749-765). Redwood City: Addison-Wesley. Explores what might be concluded about strong artificial life from recent related work in the philosophy of mind. For post-secondary school audiences. Reprinted in Boden (1996).

Sperry, R. W. (1969). A modified concept of consciousness. *Psychological Review* 76, 532-536. A classic psychological presentation of consciousness as an emergent phenomenon. For post-secondary audiences.

Von Neumann, J. (1966). *Theory of self-reproducing automata*. Urbana-Champaign: University of Illinois Press. Von Neumann's classic work on self-reproducing automata, completed and edited after his death by Arthur Burks. For post-graduate audiences.

Wiener, N. (1948). *Cybernetics, or control and communication in the animal and the machine*. New York: Wiley. The classic work on cybernetics.

Wolfram, S. (1994). *Cellular automata and complexity*. Reading: Addison-Wesley. A collection of technical papers on cellular automata as complex systems. For post-secondary school audiences.

Suggested Further Reading

Artificial Life. The primary journal for artificial life research, published quarterly since 1994 by MIT Press.

Adami, C. (1998). *Introduction to artificial life*. New York: Springer. An introduction to how a physicist might approach issues in artificial life. Includes a CD containing software for a derivative of Ray's Tierra (Ray 1992). Advanced undergraduate or graduate level.

Bedau, M. A. (1998b). Philosophical content and method of artificial life. In T. W. Bynum and J. H. Moor, (Eds). *The digital phoenix: how computers are changing philosophy* (pp. 135-152). Oxford: Basil Blackwell. An introductory explanation of the impact of artificial life on the content and method of philosophy. Discusses emergence, the supple dynamics of mental states, and evolutionary progress. For post-secondary school audiences.

Bedau, M. A., McCaskill, J. S., Packard, & N. H., Rasmussen, S. (Eds). (2000). *Artificial Life VII*. Cambridge: MIT Press. Proceedings of the seventh international artificial life conference, containing 65 papers spanning the whole field. For post-graduate audiences.

Boden, M. (Ed). (1996). *The philosophy of artificial life*. Oxford: Oxford University Press. A collection of articles on philosophical issues involving artificial life. Level of discussion varies, but generally for post-secondary school audiences.

Clark, Andy. (2001). *Mindware: an introduction to the philosophy of cognitive science*. New York: Oxford University Press. Chapter 6, "Robots and Artificial Life," illustrates robotic work in artificial life and explores the implications of artificial life for representationalism, emergence, and the nature of life and mind. For a general audience.

Emmeche, C. (1994). *The garden in the machine: the emerging science of artificial life*. Princeton: Princeton University Press. An introduction to artificial life and some of its philosophical implications. Accessible to a general audience.

Farmer, J. D., Lapedes, A., Packard, N.H., & Wendroff, B. (Eds). (1986). *Evolution, games, and learning: models for adaptation for machines and nature*. Amsterdam: North Holland. Proceedings of what could be considered the first artificial life conference, but before the field was named. For post-graduate audiences.

Floreano, D., Nicoud, J.-D., & Mondada, F. (Eds). (1999). *Advances in artificial life: 5th European conference, ECAL'99*. Berlin: Springer. Proceedings of the fifth European Conference on Artificial Life (ECAL), containing 90 technical papers. For post-graduate audiences.

Holland, J. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, 2nd edition. Cambridge: MIT Press. An expanded edition of Holland's classic technical book on genetic algorithms. For post-graduate audiences.

Holland, J. H. (1995). Hidden order: how adaptation builds complexity. Reading: Helix Books. A perspective on complex adaptive systems, by one of the founding fathers of artificial life. For a general audience.

Langton, C. G., Taylor, C., Farmer, J. D., & Rasumssen, S. (Eds). (1992). *Artificial Life II*. Redwood City: Addison-Wesley. Proceedings of the second international artificial life conference, containing 29 technical papers. Widely viewed as the best single collection of scientific work in artificial life to date. For post-secondary school audiences.

Langton, C. G. (Ed). (1995) *Artificial life, an overview.* Cambridge: MIT Press. Contains 17 scientific papers surveying many of the areas of artificial life, originally published as the three inaugural issues of the journal *Artificial Life*. For post-secondary school audiences.

Levy, S. (1992). *Artificial life, the quest for a new creation*. New York: Pantheon. A popular but scientifically respectable account of the origin of artificial life.

Varela, F.J., & Bourgine, P. (1992). *Towards a practice of autonomous systems*. Cambridge: MIT Press. Proceedings of the first European conference on artificial life, containing 55 technical papers covering all aspects of the field. For post-secondary school audiences.

Glossary

Autocatalysis: self-catalysis, catalysis of a chemical reaction by one of the products of the reaction.

Boolean network: a network comprised of some number of binary variables. The state of each variable at each step in discrete time is governed by some logical switching or "Boolean" function applied to the states of some specific set of other variables in the network.

Catalysis: a modification (usually an increase) in the rate of a chemical reaction induced by a substance (e.g., a catalyst like an enzyme) that alters the speed of, or makes possible, a biochemical or chemical reaction while itself remaining unchanged at the end of the reaction.

Cellular automaton: A regular spatial lattice of "cells," each of which can be in any one of a finite number of states. The states of all the cells in the lattice are updated simultaneously and the state of the entire lattice advances in discrete time steps. The state of each cell in the lattice is updated according to a local rule that may depend on the state of the cell and its neighbours at the previous time step. Each cell in a cellular automaton could be considered to be a finite state machine which takes its neighbours' states as input and outputs its own state. The best known example is John Conway's Game of Life.

Eukaryote: one of the two major groupings into which all organisms are divided (the other is prokaryote). Included are all organisms, except bacteria and cyanobacteria. The cells of eukaryotes possess a clearly defined nucleus, bounded by a membrane, within which DNA is formed into distinct chromosomes. Eukaryotic cells also contain mitochondria, chloroplasts, and other structures (organelles) that, together with a defined nucleus, are lacking in the cells of prokaryotes.

Homeostasis: The ability or tendency of an organism or a cell to maintain internal equilibrium by adjusting its physiological processes.

Prion: acronym for "proteinaceous infectious particle", it is a infectious micro-organism a hundred times smaller than a virus. It is composed solely of protein, without any detectable amount of nucleic acid (genetic material). How it can operate without nucleic acid is not yet known.

Prokaryote: one of the two major groupings into which all organisms are divided (the other is eukaryote). Prokaryotes are organisms (bacteria and cyanobacteria, i.e., blue-green algae) that do not have a distinct nucleus.

Turing machine: an abstract automaton that can in be any one of a number of states and that is capable of moving back and forth on an infinitely long discrete tape of instructions (customarily zeros and ones), reading and writing instructions on each segment of tape as it moves. A Turing machine's state at a given time is a finite function of both the machines current state and the information on the currently scanned section of tape. A universal Turing machine is a Turing machine capable of executing any algorithm.

Virus: any of a large group of parasitic, acellular entities that are regarded either as the simplest micro-organisms or as extremely complex molecules. A virus typically consists of a protein coat surrounding a core of DNA or RNA. It is capable of growth and reproduction only if it can invade a living cell to use the cell's system to replicate itself. In the process, it may disrupt or alter the host cell' s own DNA and hence cause various common diseases in other organisms.

MARK A. BEDAU
Mark Bedau received his Ph. D. in Philosophy from University of California at Berkeley in 1985. He is currently Professor of Philosophy and Humanities at Reed College in Portland, Oregon, Adjunct Professor of Systems Science at Portland State University, and Editor-in-Chief of the journal *Artificial Life* (MIT Press). He has published extensively on both scientific and philosophical aspects of artificial life.