**Penultimate version of review which appears in**
***Philosophical Psychology*, V. 11, No. 2 (1998), 251-260. (Please cite that version.)**

Updated: Thu Jan 22 1998
Word Count: 4,246

# Artificial Life for Philosophers

**Brian L. Keeley**
(keeley@twinearth.wustl.edu)
Philosophy/Neuroscience/Psychology Program,
Washington University in St. Louis

**Books reviewed:**

Margaret A. Boden, ed., *The Philosophy of Artificial Life*. (Series: Oxford Readings in Philosophy.) Oxford, Oxford University Press, 1996.

Christopher G. Langton, ed., *Artificial Life: An Overview*. Cambridge, MA: MIT Press, 1995.

**Abstract**

Artificial life (ALife) is the attempt to create artificial instances of life in a variety of media, but primarily within the digital computer. As such, the field brings together computationally-minded biologists and biologically-minded computer scientists. I argue that this new field is filled with interesting philosophical issues. However, there is a dearth of philosophers actively conducting research in this area. I discuss two books on the new field: Margaret A. Boden's *The Philosophy of Artificial Life* and Christopher G. Langton's *Artificial Life: An Overview.* They cover three areas of philosophical interest: the definition of life, the relationship between life and mind, and the possibility of creating life within a computational environment. This discussion allows me to critique past work in the philosophy of ALife that tends to see the field as a proving ground for traditional arguments from the philosophy of artificial intelligence. Instead, I suggest, what is interesting about ALife is how it *differs* from artificial intelligence and that the most interesting philosophical issues in the area are those derived from biology, not psychology. I recommend that these two books taken together constitute an interesting introduction to ALife and the wealth of philosophical issues found therein.

One of the most interesting and exciting intellectual developments of the past decade or so has been the emergence of *artificial life* (ALife). The study of ALife unites biologists interested in modelling and computer scientists interested in biology. ALife stretches across a wide range of issues in these two fields. Although ripe with potentially interesting philosophical issues, the academic microniche "philosophy of ALife" has been incredibly underexploited, with only a handful of active participants (e.g., Emmeche 1994; Bedau 1997a, b, c, forthcoming). No doubt, much of the reason is due to the difficulty obtaining accessible and engaging texts on the subject. To date, most ALife literature has been restricted to two classes: popular books for the "lay science" audience (Levy 1992), and detailed, scientific papers collected together in numerous conference proceedings (Brooks & Maes 1994; Langton 1989, 1994; Langton *et al* 1991; Langton & Shimohara 1997; Varela & Bourgine 1992). However, two new books help rectify this situation: philosopher Margaret A. Boden's *The Philosophy of Artificial Life* and field founder Christopher G. Langton's *Artificial Life: An Overview*.

The Boden collection is a notably long entry in the Oxford University Press *Readings in Philosophy* series, weighing in at 404 pages and containing fifteen selections (plus an introductory essay and a bibliography for further reading). Twelve of the essays have appeared in some form elsewhere, and three are unique to this collection. The book is broken up thematically into five sections. The book opens with introductory essays by Boden and Langton. Since many readers will not be familiar with the field, the second section includes essays that survey some representative examples of Alife research: Tom Ray's *Tierra* system; Stuart Kauffman's work on the self-organizing dynamics of biological systems; John Maynard Smith on artificial evolution; and David McFarland's formal approach to animal (and robot) behavior. The third and fourth sections focus on two issues: (a) architectural and explanatory requirements for robotic and animal systems, and (b) the definition of life. The collection's final section explores the issue of whether computer based entities can ever be deemed "alive" in an interesting sense of the term.

Langton's collection, *Artificial Life: An Overview*, reprints the first three issues of his new journal, *Artificial Life*, now in its third year from MIT Press. For these inaugural issues, he solicited, "a special set of overview articles contributed by members of [the journal's] Editorial Board. In these articles, each editor has attempted to review his or her own thread of special interest within the broad and diverse tapestry of research efforts that have come to be associated with the term 'artificial life'" (*ix*). The result is eighteen comprehensive articles covering all the major research areas within the field—computer viruses, computational models of evolution & genetic algorithms, morphogenesis, molecular evolution, robotics, philosophy, and ALife-in-the-classroom. There is even an extended book review of eleven recent texts related to ALife.

These two collections, coming as they do near the end of the first decade of ALife, offer an opportunity to review what the field has accomplished and to prognosticate as to its future. These books are not without their weaknesses, but luckily the two books work well together, in that the weakness of one are generally the strength of the other. These two texts also offer the opportunity to reflect on the viability of this new field. Is there enough material within ALife to keep the philosopher interested? I believe that there is, but the future of philosophical speculation about ALife lies largely along different lines than has previously been explored.

---

ALife was christened in September 1987, when then-doctoral student, Chris Langton organized a workshop on the "synthesis and simulation of biological systems" in Los Alamos, New Mexico (Langton 1989). Langton soon graduated and moved down the road to the Santa Fe Institute, with which ALife has been associated (for better or worse) ever since. In simple terms, ALife is the study of biology through the attempt to create it. It is the belief of many researchers in this field that our artificial creations—our synthetic chemicals, our robots, and most of all our computers—have developed to a level of complexity such that the quest to achieve human-made life is now worth serious effort.

ALife seems worth the effort because our artificial creations already exhibit, singly,

many attributes once restricted to natural biological entities; humans can create things which *flock*, which *explore their environments,* which can *parasitize a host*, which *compete for limited resources*, which *reproduce*, which *evolve by mutation and selection,* etc. Of course, any one of these attributions can be challenged as being too liberal. For instance, some might argue that the "flocking" of robots is too different from the natural behavior of birds or fish to be called *real* flocking. The same type of argument can be made for any of the other putative biological properties. Further, few would want to argue that the possession of any *one* of these attributes is sufficient to judge an artificial system "living." But these objections notwithstanding, the time has come to take on the (perhaps Herculean) task of attempting to create life artificially.

As Langton discusses in his essay for the Boden collection, ALife should be thought of as complementing traditional biology—particularly theoretical biology—in two interrelated ways: First, whereas traditional biology is primarily concerned with the *analysis* of living organisms and biological phenomena, ALife is concerned with the *synthesis* of life-like behavior and entities in various media. Second, whereas traditional biology studies *life-as-we-know-it* (the carbon-chain-based entities we find on Earth), ALife seeks to explore the possibilities of *life-as-it-could-be*. As science identifies characteristics which purport to be general characteristics of (or phenomena related to) living things, ALife explores the full range of mechanisms which can give rise to such phenomena, regardless of whether those mechanisms are like those that we find in our own particular biosphere.

---

A variety of philosophical issues are raised in the context of ALife. Bedau (1992) identifies fourteen questions the field raises for such diverse sub-disciplines as metaphysics, philosophy of mind, epistemology, philosophy of biology, and ethics. I will restrict my discussion here to the three issues that best allow a discussion of the books at hand.

First, ALife brings back on stage an issue that was once central to philosophy, but which has disappeared from mainstream twentieth century philosophy: the issue of the

proper *definition of life*. If the Langton collection is any indication, ALife researchers are happy to let biology define biological phenomena, while they themselves attempt to model and capture them in artificial media. This is unfortunate and it runs counter to one of Langton's central goals for ALife: to explore life as it *could* be (Keeley 1997). Before seriously exploring possible life, it seems necessary that we give some thought to what is the possible space of living systems.

Contemporary biologist Ernst Mayr (1982: 51-59) discusses a variety of "special characteristics" of life—complexity and organization, chemical uniqueness, possession of a genetic program, etc. Gareth Matthews, in his paper for the Boden collection, shows that this is a venerable approach to the problem. Aristotle, in the *De anima*, sets out to define life by explicating the "life-functions" associated with living things. Matthews provides a cogent analysis of this Aristotlean project. In particular, he notes the need to be clear about the hierarchical or "nested" nature of the properties of living things. Some properties that all (or most) living systems share, say self-nutrition, might simply be in the service of another life-function, say reproduction.

This focus on reproduction is echoed by Mark Bedau's contribution to the Boden volume, entitled "The Nature of Life". Bedau boldly takes a stand and seeks to defend a particular definition of life, namely that life is marked by the property of *supple adaptation*: "The essential principle that explains the unified diversity of life seems to be this *suppleness* of the adaptive process—its unending capacity to produce novel solutions to unanticipated changes in the problems of surviving, reproducing, or, more generally, flourishing" (338, emphasis in original). According to Bedau, the essence of life is its tenacity in the face of opposition; the way evolution by natural selection allows living things to change form in order to survive as a species. A standard counter-example involves sterile creatures, such as mules and postmenopausal placental mammals. Bedau counters this move by noting that all *individual* living creatures (sterile or otherwise) are only "living" in a derivative sense of the term. The primary sense of "alive" applies properly to systems or populations of individuals, not the individuals themselves.

While Bedau's proposal is provocative, and the discussion of Aristotle's approach is

enlightening, one is left wanting more. The definition of life is surely central to ALife, yet it is the primary focus in only three, unrelated papers in Boden's collection. In particular, one wants to see defended a more pragmatic approach to the definition of life. It is often non-productive, during the initial stages of a science, to spend a lot of time attempting to develop highly detailed definitions. Such is the case, I suggest, with ALife. When dealing with the question of the proper domain of psychology, William James (1890/1981) suggests that,

> It is better not to be pedantic, but to let the science be as vague as its subject, and include [vaguely defined] phenomena if by doing so we can throw any light on the main business in hand. It will ere long be seen, I trust, that we can; and that we gain much more by a broad than by a narrow conception of our subject. At a certain stage in the development of every science a degree of vagueness is what best consists with fertility. (19)

A similar pragmatism should be adopted here. ALife is every bit as vague in its proper constitution as psychology was in James' day. Perhaps even more so, for ALife is even younger as an established discipline than psychology was for James. This might go some way in explaining the relative silence on the part of ALife practitioners with respect to defining the concept that is central to their discipline.

The second issued raised by ALife is that of the relationship between life and mind; between ALife and its close intellectual cousin, artificial intelligence (AI). One potential complaint about Boden's collection is the large number of papers from a single institution. Five of fifteen chapters (or six of sixteen, if you count the editor's introduction) are written by individuals associated with the University of Sussex. It would have been nice to have more papers from outside the cognitive science oriented Sussex circle.

In Boden's defense, the School of Cognitive and Computing Sciences (a.k.a. COGS)

at Sussex represents one of the most active and exciting hubs of ALife research. (Indeed, I myself first developed an interest in ALife during a year spent at COGS in 1989-90.) While the issue of institutional diversity is ultimately a rather minor quibble, the correlated issue of intellectual bias is more worrisome. In the Boden collection, which concentrates on the philosopher's take on ALife, it is psychological issues—not the biological ones such as the definition of life—that take up the most space. In the Langton collection, which concentrates on scientific work in the field itself, it is biology and not psychology that receives the most ink.

ALife shares many superficial similarities and some philosophical methodologies with AI. Elliot Sober, in his paper for the Boden collection, suggests that this affinity can be characterized in the form of an analogy: artificial life is to biology as artificial intelligence is to psychology (see also, Keeley (1994)). Both fields depend heavily on computational models of their phenomena of interest. Both attempt, through *synthesis*, to expand domains of study traditionally *analyzed* by biology and psychology. ALife has as a goal the creation of human-made biological entities, while artificial intelligence has as a goal the creation of human-made psychological entities. Both fields place a lot of weight on the assumption that, at the very least, through the attempt to create artificial examples of otherwise natural phenomena we will come to a better understanding of those natural phenomena.

At the same time, it is important to keep in mind the significant differences between the two fields. First, while AI has only a single significant natural instance to model, namely normal, adult humans, ALife can draw on the whole biological world as source and inspiration. Interested in modelling reproduction? Well, there are a variety of natural systems to consider. Reproduction in nature is carried out in a variety of ways: mitosis, budding, sex, etc. There are even a variety of strategies embedded in each of these large categories. Among the sexual strategies, some animals fiercely protect their eggs and young while others leave their offspring to their own devices. Some animals produce only a few offspring, while others produce millions. And so on. Although it is not obvious that it is an *advantage*, it is certainly a difference between AI and ALife that

ALife has available a much larger set of phenomena to investigate and explain.

A second, more striking, difference between psychology and biology is that psychological explanation has *more* to explain than just the behavior of psychological systems. One of the things that makes psychology such a difficult endeavor is that in addition to the straightforward *behavioral*, third-person phenomena which stand in need of explanation, in the case of humans at least, there seem to be additional *experiential*, first-person phenomena. Part of the burden of psychology is to explain (or explain away) phenomena related to the *prima facie* claim that psychological systems exhibit attention, intentionality, consciousness, self-consciousness, a "point-of-view," the property of there being "something-it-is-like-to-be" that entity, qualia, or any other of the constellation of concepts relating to the subjective nature of the psychological. Indeed, it seems plausible that it is this element of the psychological which makes it so apparently resistant to mechanistic or reductionistic explanation. It is the difficulty of even conceiving a *conscious mechanism* which hampers the would-be psychological mechanist. Whatever consciousness is, it seems the sort of thing about which no collection of third-person facts would ever be complete; that after science has done its best, there will still remain first-person facts inaccessible to the traditional scientific method.

Thankfully, neither of the books reviewed here address the issue of consciousness. However, much of the Boden collection focuses on that other warhorse of contemporary philosophy of mind, the nature of *representation*. The question addressed is: What sense of internal representation is needed to build robots with interesting behavior and explain the interesting behavior of animals? Both sides of the debate are well-represented—McFarland and Michael Wheeler are against the invocation of representation (or perhaps more fairly, in favor of a *very minimal* sense of representation); David Kirsh and Andy Clark are generally in favor of representation-talk.

Such a debate is interesting, but I fear its conclusions are more relevant to cognitive science and AI than they are to ALife. For example, the central claim of Kirsh's contribution to the Boden collection is that the animal systems studied by ALife scientists are too different from humans to shed much light on the goals of cognitive science and

AI. The conceptual and representational requirements of an artificial insect are very different from those of an artificial human, he argues:

> If we are ever to build the much awaited household robot, it will have to be designed [to cope with increasingly complex desire systems and to recover resourcefully from failure]. I think designers, however, will have an impossibly difficult time building in such abilities without using conceptual representations. …Any household robot worth its salt must be able to make us a midnight snack. (257)

All this may very well be true, but it has little bearing on the goals of ALife. ALife does not seek to build the robotic man-servant of 1950s science fiction. Define life however you may, the facts of the matter are (a) there are an enormous number of creatures living who are living quite successfully, but (b) we would not want the vast majority of them anywhere near the insides of our houses for fear of what they would do to the carpet.

The representational requirements of a cockroach need not tell us much at all about the representational requirements of humans (although it might), yet that does not make it any less interesting, from a *biological* perspective, to try and build an artificial cockroach. The goals and intellectual backgrounds of AI and ALife are not the same, and it does ALife a disservice to simply translate four decade's worth of work in the philosophy of AI and cognitive science into a new philosophy of ALife. And, I would like to suggest, the possibility of success in ALife might well lie in the nature of these differences from related fields. If ALife can succeed in creating artificial insects or artificially evolved molecules or whatever, where AI has as yet failed to create artificial humans, so much the better for ALife.

This brings us to the third philosophical issue raised by ALife: the prospects for "strong ALife". Just as in artificial intelligence (Searle 1980), it seems reasonable to

distinguish between "weak" and "strong" versions of the goals of ALife. The *Weak* version is the claim that computers, robots, and other physical experiments, if appropriately designed, provide science with powerful tools with which to formulate and test biological theories. According to the weak thesis, ALife would be on par with statistics, electron microscopy, or any other field which develops new techniques and formalisms useful for advancing biological science.

The *Strong* version, on the other hand, would add to the weak version the claim that computers, robots, and other physical experiments, if appropriately designed, would actually *be* biological (or living) systems. According to this stronger thesis, ALife has as its goal the extension of the number of entities in the world which can be considered truly biological (or alive). This is the implication of the notion discussed above that ALife studies *life-as-it-could-be*. The greatest goal of ALife is the creation of new, human-made life.

Steven Harnad, in his contribution to the Langton volume, directly addresses the possibility of strong ALife. In this paper, Harnad does exactly what I complain about above: he translates his previous work applying Searle's "Chinese room" thought experiment to AI into a discussion of ALife. He argues that, "Computational modelling (virtual life) can capture the formal principles of life, perhaps predict and explain it completely, but it can no more *be* alive than a virtual forest fire can be hot" (293, emphasis in original). However, he goes on to argue that *synthetic* life—artificial systems which are properly grounded in the world—could be candidate biological (or living) systems.

Harnad's discussion is interesting, but the translation is woefully incomplete. Left unsaid is how this philosophical argumentation applies to the kind of work ALife actually carries out. Harnad explicitly mentions robots, so we can be pretty certain that they qualify as synthetic, and not virtual, life. But most of ALife's paradigm creations are computer based models. Are we to conclude that because these systems are found on computers that they are therefore ungrounded symbolic systems; that they are mere virtual life? Harnad does not offer a lot of guidance, although it would seem that he say

that yes, they are virtual life.

This would be to miss the point, and—by looking at ALife through AI-colored glasses—to once again miss what is interesting and different about ALife. Take, for example, Ray's *Tierra* system (introduced in the Boden collection and expanded upon in an essay in Langton's volume. I will give a quick description of the system here. For more detail, see Ray's papers and Keeley 1994). Ray was trained as a rain forest ecologist and *Tierra* (spanish for Earth) is his attempt to create an artificial ecosystem using a computer. It contains tiny strings of code (let's call them "codelets") capable of self-reproduction. All these strings of code do is make copies of themselves until they fill up the entire section of memory allotted to them by the computer's operating system, which regularly culls a portion of the codelets, thereby freeing up more room. Replication by these codelets is imperfect and this brings about variation in the population of codelets. Some codelets are better at replicating than others and the result is a open-ended evolutionary ecosystem of replicating strings of code. Ray's papers discuss how the evolutionary dynamics seen in *Tierra* reflect those seen in natural systems, e.g., there are long periods of relative stasis followed by massive population crashes as a new variant enters the ecosystem, etc.

But this is all just *virtual*, right? The "codelets" in *Tierra* are just code and we all know that code requires an interpretation. There is no "real" replication going on, just simulations of replication, correct? Not so fast. We need to remember that, unlike the case in typical AI systems, the fine structure of computation must always be kept in mind. The "codelets" in *Tierra* represent actual, *physical* patterns of high and low voltages within the computer. When a codelet successfully replicates itself, not only does the computational world of symbols and codes change, so does the physical world. Where there once was *this* pattern of high and low voltages residing at this position in space, there are now two identical patterns of high and low voltages.

If speaking of codelets in *Tierra* leaves you unconvinced, think of that related entity, the computer virus. When a computer virus infects a disc placed in a computer, it not only makes a change at the level of computation, it does so by changing the physical

medium that is carried in the disc. Does a computer virus only simulate replication? To the contrary, if replication involves bringing about multiple instances of a spatiotemporal pattern, computer viruses are *bona fide* replicators. And while virtual forest fires do not produce heat, this is of little consolation when a computer virus erases your hard drive.

I do not want to suggest that the idea that the codelets in *Tierra* or computer viruses are legitimately biological is in any way well-established. Quite the contrary, I think the issue, as are most issues in ALife, is very much up for grabs. There is a lot of open territory here for the biologically-minded philosopher to explore. However, I *do* want to argue that things are a great deal more complicated than ALife being simply AI and cognitive science in a new guise. ALife is indeed a philosophically interesting endeavor, but that interest is more than just a new proving ground for the old positions found in the philosophy of AI.

---

Although one cannot expect any finite collection of papers to cover *every* topic of interest, I feel compelled to mention one oversight here. Neither collection contains much, if anything, on the social and ethical implications of ALife. This is striking because these topics have typically been of central concern to practitioners in the field (in person and in talks, at least; far less so in print). ALife scientists are well aware of the spector of Shelley's *Frankenstein*—in which it is Dr. Frankenstein, and not his creation, who is the monster. However, this awareness is not reflected in these two collections. It would be useful to round out the conceptual issues in biology and cognitive science with some discussion of whether one, say, ought to develop new paradigms for evolving computer viruses.

My only other complaint about these two books is actually mute, if one takes the two books together. In the Boden collection, one longs for more discussion of what ALife actually consists in. Tom Ray's paper on *Tierra* is substantial, and Burian & Richardson tell a detailed story of Kauffman's work, but Maynard-Smith explicitly avoids discussing current work in genetic algorithms and computational modelling of evolution, choosing instead to give us his personal opinion on the current state of evolutionary theory.

However, if Langton's collection is anything, it is a detailed and comprehensive review of a lot of interesting scientific work. If somebody is curious to know whether there's any *there* there, Langton's collection is the place to start.

However, for the philosopher *qua* philosopher, the Langton collection is pretty disappointing. The only two explicitly philosophical pieces—one by Dan Dennett, the other by Steven Harnad—leave a lot to be desired. Dennett's paper is less than two pages and does little more than encourage philosophers to design ALife computer simulations rather than standard philosophical "thought experiments." I discussed my problems with Harnad's paper above. However, this failing on the part of the Langton collection is more than made up by the extensive philosophical discussion found in Boden's book.

Taken together, these two collections represent a very worthwhile introduction to what is going on in ALife and why philosophers (and, perhaps, cognitive scientists) should take notice. I imagine that together they would make a nice set of texts for a graduate or upper-division course on the philosophical issues raised by artificial life. Furthermore, I predict that the Boden collection will become *the* standard text in the philosophy of ALife for some time to come. If I am correct about the nature of ALife, there will be future standard texts and these will involve more and more philosophical issues unique to ALife and primarily drawn from biology (and not psychology). But we must start somewhere, and these two volumes represent a sturdy foundation on which to build.

## References:

Bedau, Mark A., "Philosophical Aspects of Artificial Life." In Varela & Bourgine (1992: 494-503).

Bedau, Mark A., "Emergent Models of Supple Dynamics in Life and Mind." *Brain and Cognition 34* (1997a): 5-27.

Bedau, Mark A., "Philosophical Content and Method of Artificial Life." In T. W. Bynam and J. H. Moor, eds., *The Digital Phoenix: How Computers are Changing Philosophy*, Basil Blackwell, 1997b.

Bedau, Mark A., "Weak Emergence." In James Tomberlin, ed., *Philosophical Perspectives: Mind, Causation, and World*, vol. 11 (Blackwell Publishers), 1997c.

Bedau, Mark A., "Four Puzzles about Life." *Artificial Life*, forthcoming.

Brooks, Rodney A. and Maes, Pattie (eds.) *Artificial Life IV*. MIT Press (A Bradford Book), 1994.

Emmeche, Claus, *The Garden in the Machine: The Emerging Science of Artificial Life*. Princeton U Press, 1994.

James, William, *The Principles of Psychology*. Cambridge, MA: Harvard Univ. Press, 1981, p. 19.

Keeley, Brian L., "Against the global replacement: on the application of the philosophy of artificial intelligence to artificial life." In Langton (1994: 569-587).

Keeley, Brian L., "Evaluating Artificial Life and Artificial Organisms", in Langton & Shimohara (1997: 264-271).

Lange, Marc, "Life, 'Artificial Life,' and Scientific Explanation". *Philosophy of Science, 63* (June 1996): 225-244.

Langton, C. G., ed. *Artificial Life: The Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems.* SFI Studies in the Sciences of Complexity, Proc. Vol. VI. Redwood City, CA: Addison-Wesley, 1989.

Langton, C. G. (ed), *Artificial Life III.* SFI Studies in the Sciences of Complexity, Proc. Vol. XVII. Redwood City, CA: Addison-Wesley, 1994.

C. G. Langton & K. Shimohara (eds.), *Artificial Life V: Proceedings of the Fifth International Workshop on the Synthesis and Simulation of Living Systems*. Cambridge, MA: MIT Press, 1997.

Langton, C. G., Taylor, C., Farmer, J. D., and Rasmussen, S., eds. *Artificial Life II*. SFI Studies in the Sciences of Complexity, Proc. Vol. X. Redwood City, CA: Addison-Wesley, 1991.

Levy, Steven, *Artificial Life: The Quest for a New Creation*, New York, Pantheon Books, 1992.

Mayr, E. *The Growth of Biological Thought*. Cambridge: Harvard University Press, 1982.

Searle, John R. "Minds, brains and programs." *Behavioral and Brain Sciences, 3* (1980): 417-424.

Varela, F. J., and Bourgine, P., eds. *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life.* Cambridge, MA: MIT Press, 1992.