Artificial Life as Philosophy

Daniel Dennett
Cognitive Studies
Tufts University
Medford, MA 02115
ddennett@pearl.tufts.edu

There are two likely paths for philosophers to follow in their encounters with Artificial Life: They can see it as a new way of doing philosophy, or simply as a new object worthy of philosophical attention using traditional methods. Is Artificial Life best seen as a new philosophical method or a new phenomenon? There is a case to be made for each alternative, but I urge philosophers to take the leap and consider the first to be more important and promising.

Philosophers have always trafficked in thought experiments, putatively conclusive arguments about what is possible, necessary, and impossible under various assumptions. The cases that philosophers have been able to make using these methods are notoriously inconclusive. What "stands to reason" or is "obvious" in various complex scenarios is quite often more an artifact of the bias and limitations of the philosopher's imagination than the dictate of genuine logical insight. Artificial Life, like its parent (aunt?) discipline, Artificial Intelligence, can be conceived as a *sort* of philosophy—the creation *and testing* of elaborate thought experiments, kept honest by requirements that could never be imposed on the naked mind of a human thinker acting alone. In short, Artificial Life research is the creation of prosthetically controlled thought experiments of indefinite complexity. This is a great way of confirming or disconfirming many of the intuitions or hunches that otherwise have to pass as data for the sorts of conceptual investigations that define the subject matter of philosophy. Philosophers who see this opportunity will want to leap into the field, at whatever level of abstraction suits their interests, and gird their conceptual loins with the simulational virtuosity of computers.

But perhaps some philosophers won't see the field this way. They will disagree with this assessment of mine or will worry about some of its presuppositions and implications, and, for them, Artificial Life will appear to be just one more controversial object in the world in need of philosophical analysis, criticism, defense, categorization. What are the n defining doctrines of the Artificial Life creed, and what can be said in defense or criticism of them? Already the stirrings of discussion about whether one wants to distinguish "strong AL" from one or another variety of "weak AL" can be heard in the corridors of philosophy. No doubt there is some useful work to be done identifying the popular misconceptions of the field and exposing them, scolding the overambitious partisans on both sides, and clarifying the actual products, as well as the prospects, of work in the field. It would be a shame, however, if this conceptual policeman role were to be the dominant contribution philosophers make to the field.

If we draw the boundaries of Artificial Life rather broadly, there are many quite traditional philosophical issues in the philosophy of biology, of science, of mind, and even metaphysics and ethics on which AL explorations have already begun to yield important insights. Even such a relatively simple ancestor as Conway's Life game provides a host of insights into traditional questions about causation, levels of explanation, identity over time, ceteris paribus reasoning and other topics [1]. Are Hobbesian just so stories about the possibility of the evolution of cooperation defensible? Certainly Axelrod's pioneering competitions point the way to a rich future of exploration. Under what conditions does (could, would, must, might) communication arise as a feature of interaction between individuals in groups? Can we build a gradualist bridge from simple amoeba-like automata to highly purposive intentional systems, with identifiable

goals, beliefs, and so forth? These questions of manifest philosophical interest merge seamlessly with the delicious conceptual questions of biology: Why is there sex? Are there fixable scales or measures of complexity or designedness or adaptativeness that we can use to formulate hypotheses about evolutionary trends? Under what conditions does the fate of groups as opposed to individuals play a decisive role in evolution? What *is* an individual? The list goes on and on.

Artificial Life has already provided philosophers with a tidy batch of examples that challenge or illustrate points that have figured prominently in contemporary philosophy. I anticipate that as philosophers acquaint themselves with the field and actively enter into its explorations, the philosophical progeny of the early work will multiply like fruitflies. After all, the field could hardly be better designed to appeal to a philosopher's habits: You get to *make up* most of the facts! This, as any philosopher knows, is perfectly kosher in a conceptual investigation.

References

1. Dennett, D. C. (1991). Real patterns. *Journal of Philosophy, 88*, 27–51.