Artificial Life as Theoretical Biology:

How to do real science with computer simulation

Geoffrey F. Miller

Cognitive and Computing Sciences

University of Sussex

Falmer, Brighton BN1 9QH, England

geoffm cogs.susx.ac.uk

Abstract

Artificial Life A Life research offers, among other things, a new style of computer simulation for understanding biological systems and processes. But most current A Life work does not sho

This romantic, colonialist view seems implicit in much of Artificial Life A Life research It maintains a kind of Orientalism of the Organic that views complex adaptive systems as newly discovered things and evolution, development, and learning as newly disco

designed through CS research. CS methods are much closer to analytical methods used by engineers (such as finite element analysis, stability analysis, or perturbation analysis) than to empirical methods used by scientists. This basic distinction leads to several other differences between RS and CS, which I will list with brazen simplicity (and tongue slightly in cheek).

RS seeks knowledge of nature, whereas CS seeks profit through technical improvements in hardware and software. RS chooses problems for their theoretical and practical importance, whereas CS chooses problems for their economic relevance. RS analyzes existing natural systems through developing speculative theories subject to hypothesis testing by observation, experimentation, simulation, whereas CS builds new artificial systems through engineering and debugging. RS requires skills in observation, experimentation, comparative analysis, statistics, scholarship, interdisciplinary communication, and bold imagination, whereas CS requires skills in programming and mathematical analysis. RS rewards discoveries for their theoretical generality, experimental clarity, and practical applicability, whereas CS rewards inventions for their patentability, marketability, and complexity. RS advances through public, loosely organized collaborations (called "research areas") by huge numbers of researchers working over many years, whereas CS advances through largely private, more tightly organized collaborations (called "project teams") by smaller numbers of programmers working over seve whereas the phenomena studied by computer science (e.g. computation) depend on the science for their very existence. Computer science is more similar in nature and spirit to architecture and aeronautical engineering than it is to physics or biology.

Granted, computer science has been spectacularly successful as an engineering discipline: computer speeds and computer sales have grown exponentially over several decades, continuing to double every few years. But computer scientists run into trouble when they try to do real science, because they are simply not trained for it. Whenever computer scientists try to do real psychology for example, a boom and bust cycle results, as in artificial intelligence during th[illegible]s, cognitive science during th[illegible]s, and neural networks research during th[illegible]s. With suggestive initial results come wild promises of further progress and massive influxes of research funding; intellectual stagnation sets in as the promised conceptual break-throughs remain elusive, but modestly useful real-world applications keep the field limping along for another few years. I am afraid that A-Life will repeat this same pattern.

Again and again, the same problems arise when computer scientists develop a new field that claims to be real science. [illegible]: the field shows historical amnesia, interdisciplinary blindness, and ignorance of current work in the relevant existing sciences. [illegible]: the field lacks explicit hypothesis testing, systematic observation, controlled experimentation, and statistical analysis. [illegible]: the field avoids recognizing or understanding its failures and over-generalizes its successes in wildly inappropriate ways. [illegible]: the field jinks from one research fad to another, failing to replicate and extend its findings in ways that could lead to a conceptually integrated discipline. Many of these problems result from an "engineering mentality" that seeks to build impressively complicated "masterpiece" systems rather than to develop simple theories that explain complex phenomena. Masterpieces of computer programming (such as those developed by

Ph.D. students in artificial intelligence or cognitive science are almost always useless as scientific models because they usually require too much pre-processing of the input, too generous an interpretation of the output, and too many hidden assumptions and hacks.

These precedents suggest that unless A-Life is very careful, it will become a historical curiosity in short order after being more or less successful in squandering many millions of dollars of public research funding that could have been better spent on a few bright evolutionary biologists and evolutionary psychologists. The remainder of this paper suggests some ways we can learn from the mistakes of the past, by identifying some methodological heuristics for doing A-Life as good theoretical biology, and some areas of biology that may benefit especially from A-Life simulations.

3 Six Methodological Heuristics for A-Life

1 [illegible] know [illegible]

This step is much harder than it sounds. Biology is a mature, successful science that has become quite sophisticated over the last few decades. One cannot just pick up a copy of [illegible] Dawkins, [illegible] or an introductory undergraduate biology textbook, find an interesting sounding issue, and forge ahead with a simulation. A basic maxim of modern science is: [illegible] biological issues much better than A Life researchers do.

So, it is difficult to find a significant unsolved problem that can be addressed through A Life computer simulations, and that has not already been addressed by standard methods of theoretical biology e.g. verbal argument, formal population genetics models, optimality theory, evolutionary game theory

comparative method Harvti

uniform age and sex structure, and no geographic structure. Phenotypes are usually skipped entirely, without attending to life history, learning, or contingent behaviour. The results of such models can be important in understanding simple evolutionary dynamics, and in exploring the implications of hypotheses about those dynamics. But such models are very weak at coping with phenomena such as complex phenotypes, flexible behaviour, co evolution, or evolutionary innovation.

A powerful way of using A Life simulations is to take an existing formal model from theoreti intellectually exciting. It is also more fun to speculate about simulation s metaphysics than its methodology: philosophizing about A Life has its own seductions and snares that distract attention from A Life s real potentials and problems as science.

To ensure A Life s relevance as theoretical biology, we must develop better methods for measurement and experimentation in our simulations. Whereas graphics may briefly catch the eyes of theoretical biologists, only solid experiments will win over their hearts. Our observational and experimental methods will have to approach the sophistication of taxonomy, comparative biology, ethology, psychology, ecology, and evolutionary biology. And we will have to develop new methods, because simulation can yield data that empirical biology cannot.

Early A Life research consisted largely of proof of concept demonstrations: local interactions can lead to certain emergent effects suggestive of biological systems. Such results were important in establishing the concepts of self or infer the dynamics of long-term processes. Such processes could not be studied directly in replicable, controlled experiments that could yield large amounts of relevant data.

Simulation allows access to much richer information as evo specific problems, it is not trivial to know whether the code is working appropriately. Making sure the code compiles and the program doesn't crash is just the first step. With humbling frequency, a simulation that seems to produce reasonable and interesting data under one set of conditions will produce, under slightly different conditions, weeks later (and often immediately before a conference), anomalous output that highlights some hidden, tiny, critical bug that invalidates weeks of results. There are three strategies for a

biologists bother to read it

4 What distinguishes A-Life from other biological simulation?

Simulation is no stranger to biology Theoretical biology papers that include simulations appear regularly in journals such as *An* [illegible], and [illegible] Several journals are heavily biased towards biological simulation, such as [illegible] and [illegible] What does A Life offer that theoretical biology does not already have

Traditional mathematics and simulations in theoretical biology try to capture self organizational or evolutionary dynamics directly in equations or simple procedures that aim straight for the collective, emergent level Such methods are only tractable when they implicitly represent the components of biological systems as simple, stable, homogeneous, and predictable Equations don t generally allow surprising, emergent behaviour and neither do simulations based on equations

The key advance in A Life has been to allow emergence by [illegible] representing individual biological things explicitly as computational procedures, allowing self organizational and evolutionary processes to emerge spontaneously from these things, and making observations and measurements about the resulting patterns and dynamics at the individual and collective levels see Taylor Jefferson [illegible] This advance has allowed A Life to explore the interaction of many biological units molecules, cells, organisms, or populations at several levels of description over different time scales behaviour, development, or evolution

A Life has used emergence in two main ways: as a proof of concept to show that certain biological phenomena can arise from distributed interactions among many local components, or as an extension to make current theoretical biology models more complete and realistic. The first way has led to some messianic predictions that theoretical biology will be revolutionized, perhaps with emergence replacing evolution as the central explanatory principle of life. The second way is more conservative and, I think, more useful: it embraces emergence without getting obsessed with it. Kauffman's 1993 research, despite its Emergentist rhetoric, seems a powerful example of this second strategy.

A Life simulations have other advantages. Conceptually, the requirements of programming force researchers to make assumptions explicit and processes computable; formal equations are actually quite ambiguous compared to computer programs. Computationally, the speed of modern hardware allows biological systems to be simulated at levels of complexity unimaginable only a couple of decades ago. Experimentally, simulation offers a high degree of control, flexibility, and replicability. Analytically, simulation allows accurate measurement of very large amounts of data, and automatic statistical analysis and visualization of that data. Socially, simulation code can be shared over computer networks, promoting easy replication and extension of results.

There is a continuum between A Life as high level theoretical biology, studying general processes and patterns of evolution, and A Life as empirical biology, modelling specific data from certain taxa. Some A Life models the mechanisms or effects of a certain set of behaviours from a single species; but many biologists already develop such models routinely in their empirical work. More usefully, A Life can model more general classes of behaviours shared across many species, such as courtship, cooperation, pursuit and evasion, communication, collective behaviour, or flocking. Models that predict different varieties of the behaviour for different species under different conditions would prove especially valuable to empirical biologists.

5 Open questions in theoretical biology that A-Life might usefully address

Given these strengths of A Life, we can turn now to ask what open questions in theoretical biology might be especially amenable to A Life simulation The following areas seem promising: evolutionary innovations, interactions between different adaptive processes and different selective forces, origins and effects of mental and behavioural adaptations, and life as it could be , logically and extra terrestrially

Theoretical biology cannot yet explain major evolutionary innovations such as the the evolution of life Eigen Schuster sex Williams ; Margulis Sagan ; Maynard Smith, ; Michod Levin , multi cellular bodies Buss , or the human brain Miller ; Ridllain maj e; Mara

A Life as Theoretical Biology

and comparative psychology only reveal the outcomes of neural and cognitive evolution, and fossils do not provide details of neural circuits during evolution. A Life systems that explicitly simulate the evolution of nervous systems interacting with each other and with complex environments may be our only hope for constructing theories of mental and behavioural evolution.

Once behavioural and cognitive adaptations have evolved, they can influence the further course of evolution within and across species. The role of mind in guiding evolution has usually been overlooked entirely, or conflated with a mystical, progressivist, animism as in the work of Herbert Spencer [illegible]; see Godfrey Smith, [illegible] and William McDougall [illegible]; see Boden, [illegible]. Only a few biologists developed a Darwinian view of minds as selective forces Morgan, [illegible]; Thompson, [illegible]; see Richards, [illegible]. But very recently, several theoretical biologists have begun to recognize the importance of perception and cognition as selective forces in the evolution of diverse phenomena such as camouflage, mimicry, warning colouration, sexual ornaments, flowers, fruits,

6 Does Strong A-Life allow stronger theoretical biology research than Weak A-Life?

The debate over "strong A-Life" (computer processes as realizations of living systems) versus "weak A-Life" (computer processes as simulations of living systems) can be extended in a methodological direction by asking: would it make any difference to theoretical biology if an A-Life system were construed as a realization rather than a simulation? Clearly, empirical biology would be affected: we would have to add a sixth kingdom of life to the current five (see Margulis & Schwartz), and I suppose that databases of biological phylogenies would have be be updated every time a new Ph.D. thesis in A-Life was written. Also granted is that acceptance of strong A-Life would imply that "life" (like "mind") is a functional, emergent property of certain systems with certain internal relations, external relations, and evolutionary histories (see Millikan). But my focus is on biological theory: what could we learn about life and evolution from *doing* strong A-Life that we could not learn from doing weak A-Life?

In answering this question, we must appreciate that many theoretical biologists view instances of real terrestrial life as little more than the outcomes of "simulations by Nature" to inform them about how evolution works. That is, empirical biology constrains biological theory in almost the same way that simulation does, so for some theoretical biologists, even real terrestrial biology could be considered "weak A-Life". Life forms that result from artificial selection by human breeders or genetic engineering also blur the distinction between realization and simulation. If one makes a strong division between Nature and Culture, such life forms are experimental simulations of what *would* happen if a lineage were subjected to some selective pressure or mutation in Nature; from a more integrated perspective, such life forms are simply the outcome of a thoroughly Natural process that happens to include humans as selective forces. Likewise, experimental biology research that records animal behaviour in unnatural laboratory conditions could be viewed either as "realizations of behaviour *n*

x , or simulations of natural behaviour . Does it make any difference to theoretical biology either way

Consider an issue in evolutionary theory that might be solved in two ways. Ms. Goodmaths develops a mathematical population genetics model that represents changes in gene frequencies using differential equations, whereas Mr. Badmaths programs an A Life model that represents genes themselves in a genetic algorithm. In the latter case, one might argue that the genes are alive in the limited sense of replicating because they really are copied within computer data structures. Suppose the two models are both good and give the same answer. The aliveness of the genes in the A Life model is simply irrelevant to the theoretical biology. If the models are formally equivalent, one can move smoothly from the differential equations through discrete iterative approximations to genetic algorithms without affecting the results in the slightest. The evolutionary dynamics have multiple realizability in formal equations, computational approximations, and procedural simulations. Only if Mr. Badmaths simulation goes beyond the complexity that Ms. Goodmaths equations can represent, is there any point in doing the simulation.

An analogy to high energy physics may also help here. The events that occur within colliders at CERN have an ambiguous status. Empirical physicists treat them as real physical events that reveal forces operating outside colliders. But for theoretical physicists, colliders may as well be viewed as very special, very expensive computers that simulate the physics of the very early universe, shortly after the Big Bang. It does not really matter for theoretical physics whether collider events are viewed as realizations of current physics or simulations of early universe physics , because in both cases the problem remains of how to generalize to processes outside the collider.

Todd [illegible] Such simulations will probably become the most important scientific tool for understanding evolution since Darwin [illegible] first de

Acknowledgments

This research was supported by NSF NATO Post Doctoral Research Fellowship RCD [illegible] and NSF Research Grant INT [illegible] Thanks to Maggie Boden, Dave Cliff, Inman Harvey, Peter Todd, and Michael Wheeler for useful discussions

References

Ackley, D., Littman, M [illegible] Interactions between learning and evolution In C Langton, C Taylor, J D Farmer, S Rasmussen Eds , *A [illegible] II* pp [illegible] New York: Addison Wesley

Ackley, D., Littman, M [illegible] A case for Lamarckian evolution In C Langton Ed , *A [illegible] III*, pp [illegible] New York: Addison Wesley

Alcock, J [illegible] *An [illegible] o An o on [illegible]* 4th Ed Sunderland, MA: Sinauer Associates

Andersson, M B [illegible] *[illegible] x [illegible] on* Princeton U Press

Barkow, J H, Cosmides, L, Tooby, J Eds [illegible] *[illegible] n o on [illegible] o o [illegible] n n [illegible] on o* Oxford U Press

Barth, F G [illegible] *In [illegible] n o [illegible] o o o [illegible] n* Princeton U Press

Bateson, P [illegible] The active role of behavior in evolution In M W Ho S W Fox Eds , *o on [illegible] o [illegible] n [illegible] o* pp [illegible] New York: John Wiley

Bedau, M, Packard, N [illegible] Measurement of evolutionary activity, teleology, and life In C Langton, C Taylor, J D Farmer, S Rasmussen Eds , *A [illegible] II* pp [illegible] New York: Addison Wesley

Boden, M [illegible] *o x [illegible] on n o o* Harvard U Press

Boekhorst, I. J. A. te, Hogeweg, P. [illegible]. Effects of tree size on travelband formation in Oran Utans: Data analysis suggested by a model study. In R. A. Brooks & P. Maes (Ed.), *A I*, pp. [illegible]. MIT Press/Bradford Books.

Boyd, R., & Richerson, P. J. [illegible]. *n o on o*. U. Chicago Press.

Buss, D. M. [illegible]. *o on o n n*. New York: Basic Books.

Buss, L. W. [illegible]. *o on o n*. Princeton U. Press.

Byrne, R., & Whiten, A. (Eds.). [illegible]. *n n n o x n o on o n n onk n n*. Oxford U. Press.

Cariani, P. [illegible]. Emergence and artificial life. In C. Langton, C. Taylor, J. D. Farmer, & S. Rasmussen (Eds.), *A II* pp. [illegible]. New York: Addison-Wesley.

Cavalli-Sforza, L. L., & Feldman, M. W. [illegible]. *n on n o on*. Princeton U. Press.

Cliff, D. [illegible]. Computational neuroethology: A provisional manifesto. In J. A. Meyer & S. W. Wilson (Eds.), *o n o n o n o In on on n on on o A o*, pp. [illegible]. MIT Press.

Cliff, D., & Miller, G. F. (in press). Co-Evolution of pursuit and evasion II: Simulation methods and results. For *A o*.

Clutton-Brock, T. H. [illegible]. *o on o n*. Princeton U. Press.

Collins, R. J., & Jefferson, D. R. [illegible]. The evolution of sexual selection and female choice. In F. J. Varela and P. Bourgine (Eds.), *o o ono o o n o o n on n on A*, pp. [illegible]. MIT Press.

Cronin, H. [illegible]. *n n o k A n x on o n o o*. Cambridge U. Press.

Daly, M., & Wilson, M. [illegible]. *x o on n o* (2nd Ed.). Boston: Willard Grant Press.

Daly

Eldredge, N. [illegible]. [illegible]. Oxford U. Press.

Eldredge, N. [illegible]. [illegible]. New York: McGraw Hill.

Endler, J. A. [illegible]. Signals, signal conditions, and the direction of evolution. [illegible].

Ewald, P. [illegible]. [illegible]. Oxford U. Press.

Fagen, R. [illegible]. [illegible]. Oxford U. Press.

Fisher, R. A. [illegible]. [illegible]. Clarendon Press.

French, R. M., Messinger, A. [illegible]. Genes, phenes, and the Baldwin Effect: Learning and evolution in a simulated population. In R. A. Brooks, P. Maes (Ed.), [illegible], pp. [illegible]. MIT Press/Bradford Books.

Foley, R. [illegible]. [illegible]. Harlow, Essex, England: Longman Scientific Technical.

Futuyama, D. J. [illegible]. [illegible] nd Ed. Sunderland, MA: Sinauer.

Futuyama, D. J., Slatkin, M. (Eds.). [illegible]. Sunderland, MA: Sinauer.

Gale, J. S. [illegible]. [illegible]. London: Unwin Hyman.

Gallistel, C. R. [illegible]. [illegible]. MIT Press.

Godfrey-Smith, P. [illegible]. Spencer and Dewey on life and mind. In R. A. Brooks, P. Maes (Ed.), [illegible], pp. [illegible]. MIT Press/Bradford Books.

Guilford, T., Dawkins, M. S. [illegible]. Receiver psychology and the evolution of animal signals. [illegible].

Harvey, P. H., Pagel, M. D. [illegible]. [illegible]. Oxford U. Press.

Hirschfeld, L. A., Gelman, S. A. Eds. *n n o n n o n on n*. Cambridge U. Press.

Kauffman, S. *o n o o o n on n on n o on*. Oxford U. Press.

Kohn, D. Ed. *n n*. Princeton U. Press.

Krebs, J. R., N. B. Davies Eds. *o o o An o on o* nd Ed. Oxford: Blackwell Scientific.

Langton, C. Artificial life. In Langton, C. Ed. *A* pp. New York: Addison Wesley.

Lewin, R. *n o on* rd Ed. Boston: Blackwell Scientific.

Mangel, M., Clark, C. W.

Miller, G. F. [illegible]. Two dynamic criteria for validating claims of optimality. [illegible]

Miller, G. F. [illegible]. [illegible]. Ph.D. thesis, Stanford University Psychology Department. To be published as a book by MIT Press Bradford Books in [illegible].

Miller, G. F. [illegible]. Exploiting mate choice in evolutionary computation: Sexual selection as a process of search, optimization, and diversification. In T. C. Fogarty (Ed.), [illegible], pp. [illegible]. Springer Verlag.

Miller, G. F., Cliff, D. [illegible]. Protean behavior in dynamic games: Arguments for the co evolution of pursuit evasion tactics in simulated robots. In D. Cliff, P. Husbands, J. A. Meyer, S. Wilson, [illegible], pp. [illegible]. MIT Press.

Miller, G. F., Freyd, J. J. [illegible]. [illegible]. Cognitive Science Research Paper CSRP [illegible], University of Sussex.

Miller, G. F., Todd, P. M. [illegible]. Exploring adaptive agency I: Theory and methods for simulating the evolution of learning. In D. S. Touretsky, J. L. Elman, T. J. Sejnowski, G. E. Hinton (Eds.), [illegible], pp. [illegible]. San Mateo, CA: Morgan Kaufmann.

Miller, G. F., Todd, P. M. [illegible]. Let evolution take care of its own. [illegible]

Miller, G. F., Todd, P. M. [illegible]. Evolutionary wanderlust: Sexual selection with directional mate preferences. In J. A. Meyer, H. L. Roitblat, S. W. Wilson (Eds.), [illegible] pp. MIT Press.

Miller, G. F., Todd, P. M. (accepted). The role of mate choice in biocomputation: Sexual selection as a process of search, optimization, and diversification. For W. Banzaf, F. Eeckman (Eds.),

Millikan, R. MIT Press.

Morgan, C. L. Natural selection and elimination. Aug.

Nitecki, M. H. (Ed.). U. Chicago Press.

Pattee, H. H. Simulations, realizations, and theories of life. In Langton, C. (Ed.), *A* pp. New York: Addison Wesley.

Plotkin, C. G. (Ed.). MIT Press.

Raff, R. A., Raff, E. C. (Eds.). New York: Alan R. Liss.

Richards, R. J. U. Chicago Press.

Ridley, Mark. Boston: Blackwell Scientific.

Ridley, Matt. London: Viking.

Roff, D. A. New York: Chapman Hall.

Romanes, G. J. 2nd Ed. Chicago: Open Court Publishing.

Ryan, M. J. Sexual selection, sensory systems, and sensory exploitation.

Said, E. New York: Pantheon Books.

Schoemaker, P. J. H. The quest for optimality: A positive heuristic for science.

[illegible]

Sigmund, K. [illegible] [illegible] Oxford U. Press.

Simpson, G. [illegible] [illegible] Columbia U. Press.

Sober, E. [illegible] Learning from functionalism Prospects for strong artificial life. In C. Langton, C. Taylor, J. D. Farmer, S. R

Webb, B. [illegible]. Robotic experiments in cricket phonotaxis. In D. Cliff, P. Husbands, J. A. Meyer, S. Wilson, *o An o An o n o In n on on n on [illegible] ono A o*, pp. [illegible] MIT Press.

Weismann, A. [illegible]. The selection theory. In *o on n o n o [illegible] k o on n nn n* [illegible], pp. New York: Boni and Liveright.

Williams, G. C. [illegible]. *A on n n on*. Princeton U. Press.

Williams, G. C. [illegible]. *[illegible] x n o on*. Princeton U. Press.

Wilson, E. O. [illegible]. *[illegible]o [illegible] o o n n*. Harvard U. Press.

Wright, S. [illegible]. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *o* . *[illegible]x In on n*, [illegible].