Implications of Creation*

David E. Hiebeler
Santa Fe Institute
1399 Hyde Park Road
Santa Fe, NM 87501 USA

December 12, 1992

Abstract

If the field of Artificial Life ("ALife") is successful, we will be forced to confront some difficult moral and philosophical issues which we might otherwise have been able to avoid. The ability to create new life forms as well as destroy existing ones will place a greater responsibility upon us. In addition, the existence of living systems within computer-simulated environments will present some new and unusual moral issues, as a result of the nature of computers and our control over them.

It is the purpose of this paper to stimulate some questions that we may be forced to directly confront in the future; this paper will not attempt to resolve these issues. It is the author's hope to encourage speculation about the moral role of scientists engaging in ALife endeavors, and to remind the ALife scientist that this research does not take place in a moral vacuum.

1 Introduction

The study of systems with lifelike behavior has a long history; it seems very natural for people to wonder what makes themselves different from other creatures, or what makes all creatures different from inanimate matter. Many

*Appeared in *Idealistic Studies*, v. 23 no. 1, Winter 1993

mechanical models of living systems have been constructed over the ages, their complexity reflecting the current state of technology at the time. Finally, during recent decades, the wonderful tools of computer simulation have become available, which have encouraged increasing activity in exploring the properties of life. Computers provide the novel idea of simple, self-contained "artificial universes" where we can create systems containing a large number of simple, interacting components. Each system has its own dynamics or "laws of physics"; after specifying those laws, we set the system into motion and observe its behavior.

In addition to computer simulation, another approach to ALife has been to build robots with sensors such as video cameras and microphones, and computer programs to drive their behavior. If the creatures' programs are capable of learning and interacting with each other, their behavior can often be surprisingly diverse [1, 2, 3, 4]. Still another approach is to try to chemically synthesize new life forms, for example trying to recreate the conditions believed to exist when life first developed on Earth, and accelerating the progress of the system in some manner.

In 1987, Chris Langton organized the first Artificial Life workshop, to try and bring together scientists conducting research in this area. In his contribution to the proceedings of the workshop, he discusses some of the ideas that came out of the conference. Among them, one of the more fundamental tenets is that "Life is a property of *form*, not *matter*, a result of the organization of matter rather than something that inheres in the matter itself." [5] The truth of this statement will be taken for granted throughout this paper.

This paper shall focus primarily on the issues arising from computer-based ALife — that is, living systems existing in simulated ecosystems in digital computers. Some of the issues discussed in this paper, particularly in section 4.6, will probably only arise with the success of ALife in this form, as opposed to other methods, such as chemical ones. This method also seems to hold great promise, especially in light of the continuing rapid increase in available computing power.

2 Two examples from fictional literature

There are many references to some possible effects of successful ALife endeavors in the literature of recent centuries, and especially in this century's science fiction. Two particularly striking examples are mentioned here: one first appeared in 1816 and is now considered a classic; the other first appeared in 1941, and more recently was published in a collection of outstanding science fiction short stories.

2.1 Frankenstein

Mary Shelley's classic novel *Frankenstein* [6] communicates some excellent points related to the creation of life, that are still just as relevant as when the story was written more than 170 years ago.

Probably one of the strongest lessons to be learned from the story is the necessity of maintaining a clear perspective on the directions current research is proceeding in. Victor Frankenstein fails to do this; he works vigorously on his dream of creating life, largely unaware of what he's really doing. Simply put, his passion for learning and achieving his goals blinds him to the implications of his work.

This spell is broken for him immediately when he successfully completes his project. Upon seeing his creation begin to live, he feels disgust and fear as he realizes what he has done. Because of this, he immediately runs away from his creature, and ignores his responsibility for his actions. This is what sets his tragedy into motion, as his creature feels alienated, confused and alone when its creator abandons it.

We can at least be reasonably certain that we will not repeat Frankenstein's grave mistake, simply because of the manner in which success will be achieved. Victor was able to create his creature because he understood some "special characteristic of life". In our endeavors, we will almost certainly have smaller incremental successes, which will give us a chance to react to preliminary, intermediate results before creating something as complex as Victor's creature. Hopefully, these intermediate results will encourage us to use the opportunities to step back and view our accomplishments and goals, to think about what we are doing and where the work is leading.

2.2 Microcosmic God

Theodore Sturgeon's short story "Microcosmic God" [7] is an optimistic look at some of the possible rewards of a very successful ALife experiment. Because the story was written before the introduction of digital computers into our society, the creatures in the story are biological, rather than computational. However, many interesting issues related to the possible results of creating life were indirectly addressed, such as the morality of controlled evolution (see section 4.3).

To summarize the story, a brilliant but reclusive scientist named Kidder lives on his private island, conducting independent research. When he becomes frustrated with his slow pace of learning (in his opinion), he realizes he can't substantially accelerate the learning process for himself or other humans to a degree that would satisfy him. Instead, he learns a great deal of biology (genetics and animal metabolisms), and manages to create some very simple but extremely accelerated organisms. Not only is their metabolism accelerated, but their evolution as well, since Kidder deliberately drives their evolution so that they will develop as quickly as possible.

Kidder's creatures have soon developed opposable thumbs and even language, and their development only continues to accelerate. Kidder then uses these creatures to solve problems for himself, and to develop new technologies. He gets them to cooperate by punishing or killing those who disobey his orders. He communicates with the creatures via a teletype; the creatures perceive this as a "word machine" through which God speaks to them.

In later sections, we will address in more detail some of the issues raised in Sturgeon's story.

3 Benefits

Before asking what are some issues that must be faced, we should first ask what we will get out of ALife. Why do people pursue this field of research at all?

3.1 Establishing a context for biology

ALife originally seemed to be largely a "pure research" discipline — not many of the earlier ALife publications emphasized practical applications, although practical benefits are being discussed more often as time goes on. One important reason for researching ALife, as described by Langton [5], is to develop a true theoretical biology, in order to locate life-as-we-know-it within the larger class of life-as-it-could-be. Currently, biology is restricted to studying carbon-based life as it has evolved on earth. While some biologists do speculate about life based on other elements, ALife encourages people to learn what it is that makes something alive in the first place, how those properties can exist and interact, and helps us to understand emergent phenomena in general.

3.2 Practical applications

ALife has some practical applications as well. We expect that by studying artificial life, we will observe some behaviors corresponding to phenomena in existing carbon-based life. The understanding of our model can then lead us to a deeper understanding of the biological phenomenon, and perhaps lead us to immediately practical results, for example understanding the behavioral and reproductive patterns of mosquitos, so that we may control the spread of malaria (e.g. see references in [8], which is also a good survey of ALife techniques and motivations).

Another application that is gaining widespread interest is the development of new drugs. By using a tight loop of selective reproduction, it is becoming possible to deliberately evolve useful organic molecules [9].

3.3 A far-out goal

A much more optimistic goal, but perhaps still practical in a very long-term sense, is to use ALife to deliberately develop a race of creatures in order to help us solve some of the problems we face (such as pollution, energy crises, etc). While this might sound like science fiction, it may not be as unreachable as it first seems. By developing a race of creatures with the capability to easily modify their own genetic structure, we will be closing a *positive feedback loop*, allowing them to create more advanced creatures, which in turn can create even more advanced creatures, and so on. After several generations, we may have a collection of creatures which are highly adapted to solve the given problem at hand.

Positive feedback between the phenotype and the genotype is a powerful tool. An analogy would be the case where humans learn how to genetically engineer a human baby which is much "smarter" than average humans (I am using the word "smart" very loosely here, roughly meaning "able to learn more quickly"). If that new human were trained in genetics, he or she could go on to create an even smarter human, which could in turn create an even smarter human, and so on. After a few generations, this positive feedback would have substantially changed what we started with into something very different.

Humans are beginning to close this feedback loop themselves through genetic engineering, but a big hindrance to this is the complex mapping of genotype to phenotype. That is, there is no simple “smartness” gene which we can easily adjust. Artificial organisms will most likely also have this problem; but because we may design organisms with this feedback capability built in from the beginning, they may be able to engage in much more genetic experimentation than we are allowed to do by our own morals.

4 Questions and issues

In this, the main section, we will raise some of the difficult issues that may be faced as ALife becomes more successful. Several, if not all, of these issues can apply to artificial intelligence as well as artificial life, since ALife creations may eventually become intelligent as the systems become more complex. However, we have tried to view these issues from the perspective of the ALife researcher, keeping in mind that we will first have to deal with systems that are considered alive but not especially intelligent. Also, it is worth mentioning that anything we consider “alive” will quite likely be considered at least somewhat “intelligent” by the AI community, since such living creatures will be adaptive to changes in their environment and will presumably exhibit complex behavior, either individually or in groups.

The issues are not presented in any particular order.

4.1 Artificial creatures on the loose

One danger of artificial life has already become quite obvious to many people — the damage that an artificial organism may inflict if it manages to “get loose in the wild”.

It may be argued that computer viruses possess many of the qualities that we require of something in order to consider it alive [10]. But whether or not computer viruses are in fact considered alive, the widespread damage they have caused cannot be denied. Such viruses, when released out into the open, may bring entire companies, or wide-area networks spanning the world, to a standstill.

In some cases, these viruses are deliberately released with malicious intent; in other cases, they are unintentionally let loose. The accidental cases are more easily dealt with, using preventive measures; as long as great care is taken to ensure a truly isolated laboratory environment for ALife experimentation, the outside world should not be at risk. The necessary steps for such isolation depend on the specific nature of the experiments. But typical experimentation with ALife in computers should simply be done on a computer or cluster of computers that are completely isolated from any outside networks; before one of the computers is ever reconnected to the outside world, all of its memory and external storage should be purged, and the software reinstalled. Any external storage (such as floppy disks) that came in contact with the laboratory system in any way should also be purged.

Another method used to ensure containment of artificial life forms is to develop them within a "virtual computer" that is being run on some host computer. No matter how complex any of the creatures may become, their basic instructions are incompatible with the instruction set of the physical machine they are running on, and so they may not escape the virtual environment that has been set up for them. This method has been successful so far, although its safety is perhaps not quite as strong as physical isolation, since it may be feasible for an artificial creature to someday exploit an error in the system and generate "native code" for the physical machine it is running on. But such an "escape" seems so extremely unlikely, that most people consider it to be effectively impossible.

As for the deliberate release of artificial creatures into the outside environment, there are essentially two deterrents — punishment and protection. If the laws against such actions are harsh, at least many people otherwise inclined to do so may refrain from such antics. As for protection, there are many counter-methods being used to combat potentially harmful computer viruses. Besides screening any new software for known viruses, a variety of methods are currently in use that are able to watch for typical "virus-like" behavior, and stop it before it spreads. Such protection efforts basically degenerate into an arms-race between those developing viruses and those developing anti-viral software.

4.2 Complexity dependence of responsibility

Just as with biological life ("BLife"), artificial life comes in many degrees of complexity. Currently, even the most advanced ALife forms are much simpler than relatively simple forms of BLife, although that may not be the case for long, as ALife creations continue to become more sophisticated.

An issue which is currently raging among many animal-rights activists is the dependence of a creature's rights on its biological complexity (more specifically, how well-developed its nervous system and mental capacities are). For example, most people do not currently believe an amoeba deserves the same considerations and treatment as a horse; the horse's more highly-developed nervous system and mental capacity makes it more deserving of better treatment. Scientists are free to expose amoeba (and of course fruit flies) to harmful radiation and chemicals, and place them into various harsh environments, without being loudly accused of cruelty to animals. Such treatment toward a horse, dog, or rabbit would draw much unpleasant attention from a large audience.

Unfortunately, much of the current animal-rights literature is more concerned with arguing about specific animals' complexities and abilities to suffer, or specific experiments that are being conducted, than with the general issues of whether it is morally justified to use animals for our own purposes even when it causes them to suffer, and how the complexity of the animals involved affects the issue.

The questions we should really be asking are "is our moral obligation toward a form of life dependent upon the complexity of that life," and if so, "what is the relationship between complexity and responsibility," and "how do we measure that complexity?" Hopefully, the current animal-rights debate will lead to some acceptable resolutions of these questions, and at the very least, increased awareness of the issue.

4.3 Morality of evolution

Assume that the current complex life on our planet evolved from some relatively simple origins. Next, imagine that this evolution was not the result of random mutation and natural selection, as we usually speak of, but rather it was driven by an *external intelligence*, for example very powerful aliens orbiting the earth but hidden from us. These aliens drove the evolution of life on Earth, by deliberately causing mutations (using their advanced technology), and by acting as a force of selection — i.e. preventing certain groups of animals from reproducing, in order to improve the species. We refer to this as *artificial selection*, as opposed to *natural selection*. This is essentially what the scientist did in Sturgeon's short story "Microcosmic God" — he caused his creations to develop by presenting them with new challenges, and punishing (or even killing) those that didn't meet the challenges well.

What would we think of such actions by aliens, driving our evolution? Is it unjustified, regardless of the outcome, or does the end justify the means? That is, if humanity turns out well because of this "experiment" of the aliens, was the experiment justified? If we turn out poorly, either destroying ourselves, or irresponsibly exploiting our planet and perhaps others, was it all a cruel experiment which caused us undue suffering, and damaged a planet as well?

Or was what they did unjust, no matter how things turn out? It was an enormous responsibility, taking things into their own hands and trusting their own judgement of what was right and wrong. After all, what is the difference between their experiment, and someone here on Earth deciding that green-eyed people are superior and others should be sterilized, to "improve the species"? How can someone be sure that his or her own judgement is so correct, and that he or she should control the life and death of another species?

One possible answer to the question is that perhaps artificial selection is fine, until the creatures reach a certain level of complexity. Beyond that, we should let natural selection take over, and those creatures successful in competing for natural resources will be the ones to survive. However, as we'll see in section 4.4, unless we are very careful, such actions may be more cruel than continuing the interference.

These are important questions, as we will be taking on the role of the aliens to an ever-increasing degree as time goes on and progress is made.

4.4 Distribution of resources

What responsibility do we have to care for any creatures we create? For example, in our current world, once we've raised a wolf in captivity, aren't we morally obligated in some way to continue feeding it, because it's our fault that it hasn't learned to survive in the wild?

This question actually comes as a tangent to the question about the morality of controlled evolution. In section 4.3, we proposed that perhaps beyond a certain point, we should let *natural* selection take the place of *artificial* selection, because we should not trust our own judgement about what is "best" for other creatures. However, it may simply be the case that we *cannot* escape the role we have taken on, once we've carried it out for some time. If we have a race of creatures existing within our computer, and we have been supplying resources as needed, to subsets of the population we decide should survive, then *how can we stop*? We would have to make some provisions for resources to continue to be distributed throughout the ecosystem. We could simply distribute the resources "equally" (assuming we know that that means for the particular system at hand), but this might just be another case of releasing a tame wolf into the wild — it may seem like a good deed on the surface, while in fact we may be condemning creatures to their death.

A better understanding of population dynamics and various effects of changes in the ecosystem might help us better decide how to "phase out" our interference in the simulation without disrupting things too much. Ironically, one way to acquire this better understanding is to experiment with artificial ecosystems. Perhaps by the time research has progressed to the point where we must directly confront this issue, we will have a sufficiently strong understanding of the issues involved, from experiments with simpler systems. This would help us to make a well-educated decision about how to relinquish the responsibility of selection to more "natural" forces within the simulation, without being too cruel by doing so. Current efforts at training animals (such as wolves) raised in captivity to survive in the wild may also give us helpful experience with this issue.

4.5 Is the cage too small?

Some people, upon discussing computer ALife simulations, express dissatisfaction with the nature of the artificial ecosystem. In particular, since we will undoubtedly be pushing the limits of whatever computer technology is available, we will want to keep the simulated ecosystem as small as possible, to make the simulation computationally feasible. The problem is that if we do in fact create artificial life within this simulation, it will effectively be in a very small "cage".

The heart of the issue is that if we are to truly say we are not imprisoning our creatures, then to the creatures within the simulation, the containing environment must appear "comfortably large," or even unbound or so large they cannot conceivably reach its limits.

This does not imply that we need to carry on a detailed simulation the size of our universe! With some care, the simulated ecosystem's boundaries may be constructed such that casual observation from within the system would not reveal the boundary. This is similar to techniques used with cellular automata to reduce finite-array effects; either the array is made periodic, or statistical boundary conditions are used. Even though the systems run on a small array, the boundary conditions make it "seem" like the array is just a small patch of a very large system. Similar techniques could be used in our ALife environments. When the creatures are simple, the boundaries don't need to stand up to careful scrutinization; as the creatures become more complex, the boundaries must become more elaborate. Note that the size of the simulation will also have to increase, since the creatures will likely cover a larger area as time goes on. The size and complexity of the *environment* must develop along with the *inhabitants*, if we don't want the cage to become too small.

4.6 The right to CPU time (or life)

If we create life, are we morally obligated to not pull the plug on it? Or if we decide we don't want to continue the simulation, should we make a backup copy of the life forms somewhere, and should we run it occasionally? Is it okay to indefinitely suspend a simulation? We can consider some analogous hypothetical situations in our society.

Imagine we are able to freeze humans, and revive them at a later time with no ill effects on their health. Certainly it would be cruel to freeze someone without their prior consent, and revive them later on, because the rest of society would have changed during that time. Freezing someone for a while would have a potentially enormous impact on her life (she might miss important meetings, her friends might assume she's dead, etc). However, what if we were to freeze someone *along with her entire universe*, for an arbitrary length of time? This is a much different situation, as *no creature* in the universe will even be aware that anything has happened! This could be happening to us regularly, for all we know!

In some sense, it is a meaningless question, then — there are no harmful effects on the creatures when their entire universe is frozen. The only worry might be that if we freeze them for a very long time, we might misplace or damage the copy of their universe, thereby "killing" them. This leads us to an even more outrageous (and more difficult question): what about freezing a universe, with *no* intention of *ever* resuming the simulation? When we suspend a simulation for some period of time, how long is too long? We could freeze a universe for thousands of years, but there would always be the possibility that we would resume the simulation someday, and if we did, the creatures in that universe would never know that they had been frozen at all.

Being morally obligated to preserve any simulations we halt, so that they may potentially be resumed later, would lead to enormous computer storage requirements. We may someday have a warehouse full of computer storage media, holding countless frozen simulations, sitting in permanent storage, that we dare not destroy for fear of "killing" the creatures they contain (if they may be considered alive, in such a frozen state).

In the author's opinion, this is one of the most difficult (and interesting) issues raised in this paper that must someday be confronted, as we will certainly construct simulations which we wish to terminate or indefinitely suspend. This issue is also the most closely tied to ALife implemented in digital computers, since that medium gives us the capability for perfect indefinite storage of creatures and their "universe".

4.7 Humanity's destiny

What will become of humanity if we make something new that may compete with us? Most people have a strong emotional reaction to the notion that we are not the end products of evolution, but merely one of the current survivors. It is entirely possible and natural (and, some would argue, inevitable) that at some time, something will eventually replace us. However, ALife may substantially accelerate that transition.

Eventually, our carbon-based method of survival may no longer be as efficient as other means of passing on the patterns of information that make us up. Computer hardware, or robots if you will, may become the vehicle for passing on the information. Such creatures may be able to couple learning with biological evolution, since they may be able to *deliberately* alter their evolution, by designing their offspring as they wish. Once something like this begins, it could pick up at a fantastic rate, since it is a positive-feedback loop, as mentioned in section 3.3. The new generations that the previous generation designed to be better, will go on to design even better offspring, and so on.

This doesn't mean that humanity is doomed. In fact, humanity will bring these new creatures into the scheme of things, so they will in some sense be our descendants, if evolution does proceed in this manner. But the fact that ALife researchers may be dramatically shaping the future of life on Earth is certainly cause for reflection, to say the least.

5 Conclusions

Many questions were posed here, with very few solutions. Deciding on moral issues is almost never easy, since the line between right and wrong is often blurred when viewed from many different perspectives. Many issues that ALife will bring upon us lie in very grey areas indeed, and debate over such topics will surely become very heated in the future when they become current issues instead of mere speculation.

Even if it is decided that ALife is just too morally dangerous for us to proceed until we are more mature, it will be impossible to stop the individual scientist who has a dream of creating life no matter what the results. Hopefully, by promoting general awareness of the implications of this research, such individuals will be more likely to pause in their work, and consider where they are going.

6 Acknowledgements

I would like to thank Hong Xia Li for the many discussions on these topics; Chris Langton and Gary Doolen for the opportunity to begin working on ALife at Los Alamos and SFI; Henry Hollinger and Sam Kim for all their help and inspiration; Michael Zenzen for encouraging thought on these issues; and Yvette Pirrone for her special help.

References

[1] Valentino Braitenberg, *Vehicles: Experiments in Synthetic Psychology*, MIT Press, 1984.

[2] Mitchel Resnick, "Lego, Logo, and Life", in *Artificial Life*, ed. Chris Langton, Addison-Wesley 1990.

[3] W.G. Walter, "An Imitation of Life", *Scientific American*, **182 (5)** (May 1950).

[4] W.G. Walter, "A Machine That Learns", *Scientific American*, **185 (2)** (August 1951).

[5] Christopher G. Langton, "Artificial Life", in *Artificial Life*, ed. Chris Langton, Addison-Wesley 1990.

[6] Mary Shelley, *Frankenstein: Or, The Modern Prometheus*, New American Library, 1983.

[7] Theodore Sturgeon, "Microcosmic God", reprinted in *Science Fiction Hall of Fame*, Doubleday & Company, 1970.

[8] Charles E. Taylor, "'Fleshing Out' Artificial Life II", in *Artificial Life II*, ed. Langton, Taylor, Farmer and Rasmussen, Addison-Wesley 1992.

[9] Amber A. Beaudry and Gerald F. Joyce, "Directed Evolution of an RNA Enzyme", *Science*, **vol. 257** (July 31, 1992).

[10] Eugene H. Spafford, "Computer Viruses — A Form of Artificial Life?", in *Artificial Life II*, ed. Langton, Taylor, Farmer and Rasmussen, Addison-Wesley 1992.