Prospects for open-ended evolution in artificial life

Russell K. Standish
HPCSU, University of New South Wales
2052, Australia
r.standish@unsw.edu.au
http://parallel.hpc.unsw.edu.au/rks

December 21, 2001

Abstract

Of all the issues discussed at *Alife VII: Looking Forward, Looking Backward*, the issue of whether it was possible to create an artificial life system that exhibits *open-ended evolution* of novelty is by far the biggest. Of the 14 open problems settled on as a result of debate at the conference, some 6 are directly, or indirectly related to this issue.

Most people equate open-ended evolution with complexity growth, although a priori these seem to be different things. In this paper I report on experiments to measure the complexity of Tierran organisms, and show the results for a *size-neutral* run of Tierra. In this run, no increase in organismal complexity was observed, although organism size did increase through the run. This result is discussed, offering some signposts on path to solving the issue of open ended evolution.

1 Introduction

An open discussion forum was conducted at Alife VII last year to draw up a list of open problems that might guide the field of artificial life in the coming years, styled on the famous Hilbert problems that guided much of Mathematics in the 20th century. The resulting list[3] contains 14 open problems, divided into 3 categories: How does life arise from the non-living; What are the potentials and limits of living systems; How is life related to mind, machine and culture.

The issue of *open-ended evolution* can be summed up by asking under what conditions will an evolutionary system continue to produce novel forms. Artificial Life systems such as Tierra and Avida produced a rich diversity of organisms initially, yet ultimately peter out[1]. By contrast, the Earth's biosphere appears

to have continuously generated new and varied forms throughout the 4×10^9 years of the history of life. There is also a clear trend from simple organisms at the time of the first replicators towards immensely complicated organisms such as mammals and birds found on the Earth today. This raises the obvious question of *what is missing* in artificial life systems?

This issue is touched on most directly in problem 6 — "What is inevitable in open-ended evolution of life", but in fact is intimately connected with problems 1, 2, 3, 5 and 7 also.

The issue of complexity growth is related to that of open-endedness. Complexity is related to information in a direct manner[15, 1]. Loosely speaking, available complexity is proportional to the dimension of phenotype space, and an evolutionary process that remained at low levels of complexity will quickly exhaust the possibilities for novel forms. However, intuitively, one would expect the number of novel forms to increase exponentially with available complexity, and so perhaps increasing complexity might cease to be important factor in open-ended evolution beyond a certain point. Of course, it is by far from proven that the number of possible forms increases as rapidly with complexity as that, so it may still be that complexity growth is essential for continual novelty.

For the purposes of this paper, we will examine the possibilities for open-ended growth in complexity in artificial life, since this most closely resembles the evolution of our biosphere. It is worth bearing in mind that the biosphere is subject to the *anthropic principle*[14] while no artificial life model is. Whether this fact is important in the generation of complexity is a question worthy of artificial life studies.

Results will also be reported of an experiment to measure the complexity of Tierran organisms. Whilst it is widely believed that the version of Tierra used here does not exhibit open ended complexity growth, the Tierran team are hoping to produce a version of *network Tierra* that does. The importance of this work is to demonstrate the feasibility of these sorts of experiments on simple ALife systems, before attempting more complex ones.

2   Complexity in Tierra

Complexity is related logarithmically to the volume of genotype space occupied by the *neutral network* of that genotype:

$$C(x) = \lim_{s \to \infty} s \log_2 N - \log_2 \omega(s, x) \tag{1}$$

where $C(x)$ is the complexity (measured in bits), N the size of the alphabet used to encode the description (in the case of Tierra, $N = 32$) and $\omega(s, x)$ the number of organisms phenotypically equivalent to x and of length less than s[15, 16]. Tierra is an artificial life system created by Tom Ray[9] where the artificial organisms are computer programs written in a special assembler-like language.

The equivalence principle for defining the context is phenotypic equivalence, ie when two Tierran organisms behave identically, even if they contain a different sequence of instructions. Adami's group pioneered this technique with *Avida*[1, 2], a similar artificial life system inspired by Tierra. In Avida, the situation is particularly simple, as organisms have only one phenotypic characteristic, namely their reproduction rate. To make matters more interesting, the Avida group set the creatures a range of computational tasks to attempt — success at these tasks is rewarded by extra CPU resources. The success or otherwise of the creatures at these tasks further distinguishes between Avida phenotypes.

Tierra presents considerably greater difficulties than Avida in that Tierran organisms can *interact* with each other via template matches or simply "falling off the end of the code". So, the Tierra phenotype can be categorised by examining the interactions between all possible phenotypes. Since genotype space is so huge, what better way of searching for viable phenotypes than using a genetic algorithm, namely taking the results from a long Tierra run. These are then pitted against each other in pairwise tournaments[13, 11] in a specially crafted simulator called *miniTierra*[12] that can perform the tournament nearly 1000 times faster than Tierra itself. Once all the tournaments have been completed, the organisms are sorted into phenotypically equivalent classes, resulting in a small list of archetypal organisms. These are labeled by the genotype with the earliest creation date. Table 1 lists the archetypes from one such run, along with their creation time, non-volatile site count (see [16]) and measured complexity. The algorithm employed is described in greater detail in [16].

The code and data used for this experiment is available from http://parallel.hpc.unsw.edu.au/rks/software/eco-tierra-1.1.tar.gz. The use of the software is briefly documented in the README file.

There are several key points to note:

- Over this Tierra run of 6×10^9 instructions, there is no sign of complexity increase, although lengths of the organisms do increase from 80 initially to 526 instructions long by 1.4×10^9 instructions executed. CPU resources are divided amongst the organisms in a *size neutral* fashion (ie proportional to organism size) to allow this growth in genome size to happen. This is not an especially lengthy Tierra run, however, and perhaps real complexity takes much longer accumulate in the junk parts of the genome.

- The non-volatile site count C_{NV} is a good proxy measure of the total complexity. The biggest discrepancy in this database occurred for 0194aag, and the real complexity was only 58% higher. This is good news, as it is far more tractable to compute C_{NV}, than the full complexity. The full C_2 computation for this dataset of organisms required nearly 8 CPU years of contemporary processor time, as opposed to about 3 CPU hours for C_{NV}. Fortunately, it is a highly parallel problem, and was computed within about 3 months on a major high performance computing facility.

- The precise definition of the *phenotype* has a big influence over complexity values. As has been argued by Adami[1] and myself[15], this is highly context dependent. No particular definition is wrong per se, however some definitions will be better than others. I hope, dear reader, you consider my definition based on pairwise ecological interactions to be a good definition, and that therefore the complexity measures obtained are useful. It is known that Tierran organisms can exhibit what is called *social hyperparasitism*[9]. This is necessarily a 3-way interaction between Tierran organisms — *hyperparasitism* refers to stealing another organisms CPU to reproduce itself, and *social* refers to the fact that multiple organisms must cooperate to perform this feat. However both of these properties show up as distinct patterns amongst 2 way interactions - the hyperparasite still manages to interfere with its prey's reproductive capability, even if it is unable to gain from the act alone, and mutual interaction with members of its own kind show up also.

3   Discussion

Thus far, the focus of this work has been on presenting a practical method of computing the complexity of Tierran organisms. However, it has been long recognised that neither Tierra nor Avida exhibit open-ended evolution evolution. It is thought that these systems rapidly adapt to and exhaust the possibilities of a fairly simple environment. Adami et al. have approached the problem by artificially increasing the complexity of the environment by specifying a sequence of arithmetical operations that the digital organisms can attempt. Success is rewarded by extra CPU time. In these circumstances, they have shown increasing complexity of the organisms over time[1, 2]. Ray, on the other hand, is attempting to exploit naturally occurring environmental complexity provided by the Internet[8], and hopes to engineer a "Cambrian explosion". To date, his team has had mixed success — whilst they have managed to get organisms to persist in a multicellular state under mutational load, as opposed to reverting to a single celled, they haven't achieved their aim of complexity growth[10].

The obvious next step in research is to attempt to measure the complexities of network Tierra organisms. It is by no means clear that this is tractable. The success of this approach with the original Tierran and Avidan organisms may well be due to the simplicity of the environment.

As I observed in [14], the natural biosphere operates under the constraints of the *anthropic principle*. This means that we must necessarily observe a path to increased complexity in the biosphere. Furthermore, there are some statistical regularities in evolutionary history that indicates the evolution of the biosphere to be an extremely rare process rather than an inevitable one[5]. The universe appears to be performing a quantum search algorithm to find the precise conditions required to generate intelligent life. Probably the way forward is to

4

Organism	Creation	C_{NV}	C_2
0038aep	5104	21	26.2362
0045aaa	2	34	36.8368
0065aac	3172	29	35.5448
0073aaa	0	63	65.2125
0078aal	59	53	63.0305
0078aan	3556	37	43.5753
0080aaa	0	62	69.9581
0080aea	3	57	66.1046
0081aaj	2989	27	32.7449
0128aad	123	21	27.8352
0132abi	289	39	50.0246
0134aae	298	41	51.6064
0138aab	190	18	25.9443
0138aai	179	18	26.1733
0139aaa	258	47	62.8791
0150aag	55	53	62.8791
0155aab	66	4	5
0157aaa	66	53	62.9423
0182aaa	402	38	49.1947
0186aah	2536	28	41.2842
0194aag	2461	27	42.7729
0198aad	2434	29	39.3959
0218aab	323	48	57.3709
0236aaa	359	48	63.836
0260aae	3310	29	35.8535
0397aab	3321	28	34.2954

Table 1: Table of archetypal organisms, with the time of first appearance in genebank record, their non-volatile site count C_{NV} and their computed complexity truncating the neutral network at 2 hops C_2.

perform a similar sort of "scattergun" approach, using the power of quantum computers if and when they become available.

A key step in doing this is to generate a process that adaptively recognises complexity, since it will be impossible to include humans in the loop, even when run on conventional computing platforms. To this end, techniques developed for *data mining*[6] should prove useful. As a curious twist, artificial life techniques are being successfully applied to the domain of data mining[7]. Developing a coevolving system observer is something that can be started now with present technology.

Acknowledgments

References

[1] Chris Adami. *Introduction to Artificial Life*. Springer, 1998.

[2] Chris Adami and N.J.Cerf. Physical complexity of symbolic sequences. *Physica D*, 137:62–69, 2000.

[3] Mark A. Bedau, John S. McCaskill, Norman H. Packard, Steen Rasmussen, Chris Adami, Devid G. Green, Takashi Ikegami, Kinihiko Kaneko, and Thomas S. Ray. Open problems in artificial life. *Artificial Life*, 6:363–376, 2000.

[4] Dario Floreano, Jean-Daniel Nicoud, and Francesco Mondada, editors. *Advances in Artificial Life: 5th European Conference, ECAL 99*, volume 1674 of *Lecture Notes in Computer Science*. Springer, Berlin, 1999.

[5] Robin Hanson. Must early life be easy? the rythm of major evolutionary transitions. *Origins of Life*, 2000. submitted; http://hanson.gmu.edu/hardstep.pdf.

[6] George M. Marakas. *Decision Support Systems in the Twenty-First Century*. Prentice Hall, Upper Saddle River, NJ, 1999.

[7] N. Monmarché, M. Slimane, and G. Venturini. On improving clustering in numerical databases with artificial ants. In Floreano et al. [4], pages 626–635.

[8] Tom Ray. A proposal to create two biodiversity reserves: One digital and one organic. See ftp://tierra.slhs.udel.edu/tierra/doc/reserves.tex, http://www.hip.atr.co.jp/~ray/pubs/reserves/reserves.html. Also see New Scientist, vol 150, no 2034, pp32–35.

[9] Tom Ray. An approach to the synthesis of life. In C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, editors, *Artificial Life II*, page 371. Addison-Wesley, New York, 1991.

[10] Tom Ray and Joseph Hart. Evolution of differentiated multi-threaded digital organisms. In Chris Adami, Richard Belew, Hiroaki Kitano, and Charles Taylor, editors, *Artificial Life VI*, pages 295–304, Cambridge, Mass., 1998. MIT Press.

[11] R. K. Standish. Embryology in Tierra: A study of a genotype to phenotype map. *Complexity International*, 4, 1997.

[12] R. K. Standish. On an efficient implementation of Tierra. *Complexity International*, 4, 1997.

[13] Russell K. Standish. Some techniques for the measurement of complexity in Tierra. In Floreano et al. [4], page 104.

[14] Russell K. Standish. Evolution in the Multiverse. *Complexity International*, 7, 2000.

[15] Russell K. Standish. On complexity and emergence. *Complexity International*, 9, 2001.

[16] Russell K. Standish. Open-ended artificial evolution. In *Proceedings Inaugural Artificial Life Workshop*, December 2001.