EMCSR: pp-pp, year
ISSN xxxx-xxxx
http://conference.emcsr.net

EMCSR
European Meetings on Cybernetics and Systems Research

Von Neumann Redux: Revisiting the Self-referential Logic of Machine Reproduction using the Avida World

Barry McMullin
The Rince Institute, Dublin City University, Ireland. barry.mcmullin@dcu.ie
Tomonori Hasegawa
The Rince Institute, Dublin City University, Ireland. tomonori.hasegawa2@mail.dcu.ie

Abstract: *We introduce the distinctive, self-referential, logic of self-reproduction originally formulated by John von Neumann and present some initial results from a novel implementation of this abstract architecture, embedded within the Avida world. These show that, with this particular implementation, in this particular world, the von Neumann architecture proves to be evolutionarily unstable and degenerates, surprisingly easily, to a primitive, non-self-referential, "copying" or "template replication", mode of reproduction. We briefly discuss some implications, and sketch prospects for further investigation.*

Keywords: Self-reference, self-reproduction, semantic closure, Avida, von Neumann, artificial life

Acknowledgement: This work has been supported by the European Complexity Network (Complexity-NET) through the Irish Research Council for Science and Technology (IRCSET) under the collaborative project EvoSym.

1. Introduction

As early as 1948, John von Neumann had already formulated and essentially resolved a fundamental paradox in the theory of the evolutionary growth of machine complexity: namely, how any (assumed or "divinely" created) seed machine can, directly or indirectly, give rise to machines arbitrarily more complex than itself (McMullin, 2000). Inspired by Turing's general purpose (programmable) computing machines, his resolution relied on a machine architecture comprising a general purpose programmable *constructor* which could act to decode a symbolic description of an arbitrarily (more) complex target machine and thus construct it. As a special case, this also led to a generic architecture for machine *self-reproduction* (where the description is now a *self*-description, and must be *copied* as well as *decoded*). This self-reproduction architecture, formulated very abstractly by von Neumann, was subsequently found to be strikingly reminiscent of the biological role of DNA (as "symbolic description") and of the molecular machinery of the "genetic code" whereby ribosomes (supported by tRNAs and other enzymes) decode or "translate" symbolic descriptions (presented as mRNAs) into arbitrarily complex protein molecules (and protein machinery).[1] More generally, von Neumann's architecture gave a concrete mechanical interpretation and implementation of the traditional biological idea that an organism can be decomposed into a set of tacit hereditary "factors" (genome) and a corresponding, manifest, functional, form (phenome). As subsequently emphasised by Pattee (1982), however, von Neumann's architecture (and its real-world biological counterparts) also carries with it an intriguing example of *self-reference:* the decoding relationship (the "genotype-phenotype mapping", in biological terms) implemented by the programmable constructor is also represented, in encoded form, within the symbolic description (genome) – and this encoding must be precisely according to, or at least consistent with, the very same mapping that the constructor itself (part of the phenome) implements. This most primitive and original form of *self-reference* has been dubbed "semantic closure" by Pattee, and has also been explicitly discussed by Hofstadter (1985); but the full implications of this self-referential closure for understanding, and fabricating, complex self-organising systems are, as yet, poorly understood. The present contribution presents a brief

[1] The prescient nature of von Neumann's contribution is made clear from the fact that the chemical structure of DNA was not elucidated until 1953, and the programmable "decoding" or "translation" function of the ribosome was not fully formulated until 1960 (the code "proper" only later clarified as being implemented by the aminoacyl-tRNA synthetases).

summary of one preliminary attempt to revisit and explore this issue afresh, through building and characterising abstract computational models in the *Avida* world.

2. Self-Referencing Reproduction and Evolution in the Avida World

Avida is an abstract ("simulated" or "virtual") world which has been extensively used to investigate very general properties of spontaneous evolutionary processes (Adami, 1997; Ofria & Wilke, 2004). It is loosely inspired by the structure of a conventional, large scale, cluster computer, with many separate computational nodes, each with one general purpose CPU and a limited local memory. The nodes are sparsely interconnected, typically in a regular two dimensional lattice.[2] The CPU instruction set is configurable on a system wide basis. It is normally reminiscent of a conventional microcontroller, but with some specialised features. A program running on a given node can overwrite the memory of a neighbouring node and in this way replace the program running on that node (effectively re-program the node). Based on this, a suitably designed program may be able to repeatedly reproduce itself into neighbouring nodes. Such a program is regarded as an abstraction of a biological organism. If an Avida world is initialised or seeded with a single instance of some such hand-designed organism, a population of organisms will grow to occupy the entire world roughly in the manner of bacteria in a petri dish. Certain CPU operations in Avida are made unreliable by design. This has the effect that mutant strains of organism can spontaneously arise, multiply, and compete in a darwinian manner for the finite available "space" (nodes) in the system. Unlike the petri dish analog, a culture of Avida organisms can be continuously replenished with "nutrients" (analogous to a continuous flow bioreactor) and thus the ongoing evolutionary process can, in principle, be continued indefinitely.

The "standard" mechanism whereby self-production is achieved in Avida is based on an approach pioneered in several predecessor systems, including *Coreworld* (Rasmussen, Knudsen, Feldberg, & Hindsholm, 1990) and *Tierra* (Ray, 1994). In effect, the parent organism simply inspects and copies its own program directly. We may classify this as reproduction by "self-inspection" or "self-copying". Such a mechanism is not possible for complex organisms in the real world for several practical reasons (McMullin, Taylor, & von Kamp, 2001); but it is closely analogous to the more primitive template replication process underlying *in vitro* RNA evolution, and, indeed, to the DNA replication process that is one component of normal biological reproduction. In particular, it does support inheritable variation and evolutionary exploration of a combinatorially large (for practical purposes, infinite) space of distinct organism strains.

In contrast to this standard reproduction mechanism, used in all previous studies with Avida, we have designed a novel seed program which incorporates the characteristic genotype-phenotype structure and self-referential genotype-phenotype mapping originally described by von Neumann. In the first instance the mapping has been simply modelled on the standard biological genetic code. That is, it is a sequential mapping from discrete "codon" symbols in the genome to functional "instructions" in the phenome, implemented via a lookup table located in the (parental) phenome (directly analogous to the functionality implemented by the aminoacyl-tRNA synthetases in RNA-protein translation). The lookup table is itself, self-referentially, encoded into the genome. While it would be expected that this self-referential mapping would be highly conserved (robust) in evolution, we nonetheless conjecture that some significant long term evolution (either selective or by drift) should be observed.

As yet, only preliminary experiments have been run and analysed with this novel self-referential seed organism. However, the consistent experience to date has been that instead of observing either simple conservation or long term evolution in the genotype-phenotype mapping we see

[2] The Avida world bears some superficial resemblances to von Neumann's own early formulation of an abstract cellular automaton (CA) world, particularly in its 2D network of discrete computational nodes. However, there are also fundamental differences. In the von Neumann CA, each node was a simple finite state automaton with no general purpose memory system (29 states per node, equivalent to less than 5-bits of special purpose memory); whereas each Avida node comprises a general purpose CPU and – by comparison – a substantial general purpose memory system, typically of capacity at least some hundreds or thousands of bits and potentially configurable to be much bigger.

relatively rapid degeneration of the entire reproduction mechanism – i.e., emergence of "conventional", non-self-referential, self-copying organisms, comparable to the standard seed organisms. These organisms lack the decomposition into distinct genome and phenome components, lack any genotype-phenotype mapping process, and therefore also lack the characteristic von Neumann self-reference (or Pattee's semantic closure). Once such organisms emerge they are selectively favoured in this world (as they avoid the computational load of translation/decoding without incurring any immediate offsetting penalty). It follows that this degeneration is essentially irreversible.

3. Discussion and Future Prospects

As noted, it is not surprising that self-copiers should selectively displace self-referencing organisms in Avida; nor is it *very* surprising that there might be some available mutational pathways for such degenerative strains to appear, as the Avida world is specifically designed to make reproduction by self-copying extremely easy (it can be achieved with a program as short as 15 instructions under the default instruction set). However, what was surprising was that this degeneration could occur with just a single point mutation in our newly developed seed organism. Given that several aspects of the reproduction cycle need to be well co-ordinated for reproduction to succeed, we had not thought that a single mutation would be likely to already yield a viable self-copier. Further analysis is ongoing to fully understand the mechanism for this transition but it does reinforce the creative power of evolutionary search. More importantly, our next step will be to introduce specific mechanisms that will selectively favour self-referencing organisms over self-copiers (several potential mechanisms have already been identified). In this way, we should avoid the degenerative takeover by self-copiers and return to the our core research question which is to observe and characterise evolutionary change in the self-referencing genotype-phenotype mapping within a suitably configured Avida world.

References

Adami, C. (1997). *Introduction to Artificial Life*. Springer.

Hofstadter, D. (1985). The genetic code: Arbitrary. *Metamagical Themas: Questing for the Essence of Mind and Pattern* (pp. 671–699). London: Penguin Books.

McMullin, B. (2000). John von Neumann and the Evolutionary Growth of Complexity: Looking Backward, Looking Forward. *Artificial Life*, *6*(4), 347-361.

McMullin, B., Taylor, T., & von Kamp, A. (2001). Who Needs Genomes? Presented at the Atlantic Symposium on Computational Biology and Genome Information Systems & Technology. Retrieved from http://alife.rince.ie/bmcm-cbgi-2001/

Ofria, C., & Wilke, C. O. (2004). Avida: A Software Platform for Research in Computational Evolutionary Biology. *Artificial Life*, *10*(2), 191-229.

Pattee, H. H. (1982). Cell psychology: an evolutionary approach to the symbol-matter problem. *Cognition and Brain Theory*, *5*(4), 325–341.

Rasmussen, S., Knudsen, C., Feldberg, R., & Hindsholm, M. (1990). The coreworld: Emergence and evolution of cooperative structures in a computational chemistry. *Physica D: Nonlinear Phenomena*, *42*(1–3), 111-134.

Ray, T. (1994). An Evolutionary Approach to Synthetic Biology: Zen and the Art of Creating Life. *Artificial Life*, *1*(1/2), 179-209.

About the Author

Barry McMullin

Associate Professor and Director of the Rince Institute at Dublin City University, Ireland. Professor McMullin has focussed particularly on the problem of realising evolutionary growth of complexity in artificial (engineered) systems.

Tomonori Hasegawa

After majoring in Information Systems Science at Soka University, Japan, he has started pursuing research on the EvoSym project as a postgraduate student since 2010 at the Artificial Life Laboratory of the Rince Institute, Dublin City University, Ireland. His primary interest is centred around Artificial Life, specifically, the evolution of symbolisation.