The Von Neumann Self-reproducing Architecture, Genetic Relativism and Evolvability

Barry McMullin

Artificial Life Laboratory
Research Institute in Networks and Communications Engineering (RINCE)
Dublin City University, Dublin 9, Ireland
Barry.McMullin@dcu.ie

Abstract

Within any suitable framework of primitive automaton components (CA or otherwise) the von Neumann architecture for self-reproduction can generally give rise to (an infinity of) infinite sets of self-reproducing automata. The automata within any single such set are characterised as sharing a particular "constructing" or "decoding" subsystem (a "general constructive automaton"). This means that all the automata within such a set share the same formal "genetic language"; this, in turn, means that they are connected by a network of potential mutations. The latter was an important and significant finding by von Neumann, as it established that some, at least, of the conditions for the evolutionary growth of automaton complexity can be met in such mechanistic frameworks, within any single one of these von Neumann sets of self-reproducers (McMullin, 2000). This note explores a further elaboration of these results, based on the possibility of mutational pathways *between* (as opposed to *within*) von Neumann sets; i.e., pathways between automata implementing *different* genetic languages. Von Neumann himself considered this issue briefly, but apparently discounted it (von Neumann, 1949, p. 86). By contrast, I argue that it may be deeply significant for both natural and artificial evolutionary systems.

Introduction

In a paper at the main ALife VII conference (McMullin, 2000) I mention in passing that von Neumann's architecture for self-reproducing machines supports the possibility of *genetic relativism* and that it seems to me that this allows for a more profound form or degree of evolvability in von Neumann's system compared to, say, `Tierra` (Ray, 1992). This note elaborates briefly on this point. It is drawn from comments originally presented in my PhD thesis *Genetic Relativism* (McMullin, 1992*a*, Chapter 4).

I will refer several times below to what I call *von Neumann's Problem*. By this I will mean the problem of how machines can manage to construct other machines more "complex" that themselves, in a general and open-ended way—i.e., with the potential for unbounded evolutionary growth of complexity. This is, of course, essentially a problem of *artificial evolvability.*

The von Neumann Architecture for Self-reproduction

The von Neumann architecture for self-reproduction is based on the idea of a 'general constructive automaton". I will denote (an example of) such an automaton by u_0. u_0 is, in effect, a programmable constucting machine. In a manner reminiscent of a turing machine, if a "tape", d, is attached to u_0 then u_0 will interpret this tape as a description of some other machine, which u_0 is intended to construct. We require that u_0 should also *copy* the attached description tape (and attach this copy in turn to the constructed machine). We assume that u_0 is capable of successfully constructing a wide range of such target machines (this is why it is termed a *general* or programmable constuctive automaton). Let $d(m)$ denote a tape describing target machine m (relative to the decoding conventions, or genetic language, realised by u_0). We then write:

$$(u_0 \oplus d(m)) \rightsquigarrow (m \oplus d(m))$$

Assuming that the set M of machines that u_0 can construct includes u_0 itself, it follows that:

$$(u_0 \oplus d(u_0)) \rightsquigarrow (u_0 \oplus d(u_0))$$

This is the essential structure for a single self-reproducing automaton which has been commonly identified as von Neumann's "result". But such an interpretation seriously understates the true depth of von Neumann's work. It completely misses the single most significant aspect of this particular architecture, namely that this core design in turn implies the existence of a whole infinite set of general constructive automata which incorporate u_0 as a *subsystem*:

$$u_i = u_0 \oplus i$$

With some limited assumptions about the interactions of u_0 and the "ancillary" machinery i, it follows that there is an infinite set of self-reproducing automata of the form:

$$(u_i \oplus d(u_i))$$

It turns out that the elements of this set are connected by a network of potential *mutations* (interpreted simply as spontaneous perturbations of the description tape). Which means that there are evolutionary pathways linking the simplest element of this set $(u_0 \oplus d(u_0))$ with arbitrarily complex elements. Exhibiting the detailed design of a particular u_0 in a particular framework (which is what von Neumann actually did) thus leads directly to an effective solution to von Neumann's problem of the evolutionary growth of machine complexity.[1]

On Genetic Absolutism

Now this set of von Neumann self-reproducers anchored on a single u_0 have precisely this in common: they process the same formal "genetic language" for describing machines. In biological terms we may say that this set incorporates a fixed, or *absolute* mapping between genotype (description tape) and phenotype (self-reproducing automaton).

Thus, in committing ourselves (following von Neumann) to solving the problem of the evolutionary growth of complexity purely within the resources of a single such set, we are also committing ourselves to the equivalent of what I have elsewhere called *Genetic Absolutism* (McMullin, 1992*b*, Section 5.3), within the analysis of our formal or artificial system.[2] I should note that, in that paper, I argue at length against the idea of Genetic Absolutism; but not in the sense that it is "bad" in itself—it just is not a tenable theory of biological evolution. Now von Neumann is not yet trying to capture all the complications of biological evolution: he is merely trying to establish that some key features, at least, can be recreated in a formal, or artificial, system. If this can be done within what is, in effect, a framework of Genetic Absolutism, *and if there is some advantage to doing this in that particular way,* then the fact that it is still "unbiological" (in this specific respect) should not be held too severely against it. (Indeed, there are arguably much more severe discrepancies than this in any case.)

[1]This is, of course, an extremely concise summary; see (McMullin, 2000) for much more detail.

[2]Note carefully that this is strictly a limitation of the way *we choose to analyse* the automata framework in question; it need not, and generally will not, reflect an inherent limitation of any such framework *in itself.*

Now, as it happens, adopting Genetic Absolutism *does* have a significant advantage for von Neumann. Working within such a framework it *is* necessary (for the solution of von Neumann's problem) to exhibit one core general constuctive automaton, u_0; and it *is* necessary to establish that this is sufficiently powerful to satisfy the informal requirements of the evolutionary growth of complexity; and it *is* finally necessary to show that, based on the formal genetic language processed by u_0, there is a reasonable likelihood that most, if not all, of the corresponding self-reproducers will be directly or indirectly connected under mutation. But if all this can be done, then the problem immediately at issue for von Neumann can, indeed be solved.

On Genetic Relativism

In any case, what would be the alternative if Genetic Absolutism were not adopted?

Well, the alternative to Genetic Absolutism is *Genetic Relativism* (McMullin, 1992*b*, Section 5.4), which envisages that the mapping between genotype (description tape) and phenotype (self-reproducing automaton) is *not* fixed or absolute but may vary from one organism (automaton) to another.

If we tackle von Neumann's problem in a framework of Genetic Relativism, we do *not* restrict attention to a single u_0, giving rise to an "homogenous" set of self-reproducers, all sharing the same genetic language. Instead we introduce the possibility of having many *different* core automata—u_0^1, u_0^2 etc. Each of these will process a more or less *different* genetic language, and will thus give rise to its own unique set of related self-reproducers. We must still establish that most if not all self-reproducers in each such set are connected under mutation; but, *in addition,* we must try to show that there are at least some mutational connections between the *different* such sets.

The latter is, of course, a much more difficult task, because the mutations in question are now associated with changes in the very languages used to decode the description tapes. But, if such connections could be established, then, for the purposes of solving von Neumann's problem we are no longer restricted to considering the range of complexities of any single von Neumman set of self-reproducers (i.e., anchored on a single u_0, with a common description language), but can instead consider the union of many—indeed a potential infinity—of such sets.

Now clearly, in terms simply of solving von Neumann's problem, Genetic Relativism introduces severe complications which are not necessary, or even strictly useful. For now we have to exhibit not one, but multiple core general constructive automata, processing not one, but multiple genetic languages; and we have to characterise the range of complexity, and mutational connectivity, of not one but multiple sets of self-reproducers; and finally, we still have to establish the existence of mutational links *between* these different sets of self-reproducers. The only benefit in this approach seems to be that maybe—just maybe—the distinct general constructive automata can be, individually, significantly simpler or less powerful than the single one required under Genetic Absolutism; but it seems quite unlikely that this could outweigh the additional complications.

So Now What?

Let me say then that I actually accept all this: that for the solution of von Neumann's problem, as I have stated it, adopting the framework of Genetic Absolutism seems to be quite the simplest and most efficacious approach, and I endorse it as such. Nonetheless, I think it worthwhile to point out the *possibility* of working in the alternative framework of Genetic Relativism for a number of distinct reasons.

Firstly, it would be easy, otherwise, to mistake what is merely a pragmatic preference for using Genetic Absolutism in solving von Neumann's problem with the minimum of effort, for a claim that Genetic Absolutism is, in some sense, *necessary* for the solution of this problem. It is not. More generally, our chosen problem is *only* concerned with what may be possible, or sufficient—not what is necessary.

A second closely related point is this: *prima facie,* our solution based on Genetic Absolutism may seem to imply that a *general* constructive automaton (i.e., capable of constructing a very wide range of target machines) is a pre-requisite to *any* evolutionary growth of complexity. It is not. Indeed, we may say that, if such an implication *were* present, we should probably have to regard our solution as defective, for it would entirely beg the question of how such a relatively complex entity as u_0 (or something fairly close to it) could arise in the first place. Conversely, once we recognise the *possibility* of evolution within the framework of Genetic Relativism, we can at least see how such prior elaboration of the powers of the constructive automata could occur "in principle"; this insight remains valid, at least as a coherent conjecture, even if we have not demonstrated it in operation. This has a possible advantage in relation to the solution of von Neumann's problem in that it may permit us to work, initially at least, with significantly more primitive constructive automata as the bases of our self-reproducers.

Thirdly, Genetic Absolutism views all the self-reproducers under investigation as connected by a *single* "genetic network" of mutational changes. This is sufficient to solve von Neumann's problem, as stated, which called only for exhibiting the *possibility* of mutational growth of complexity. In practice, however, we are interested in this as a basis for a *Darwinian* growth of complexity. Roughly speaking, this can only occur, if at all, along paths in the genetic network which lead "uphill" in terms of "fitness". If the genetic network is fixed then this *may* impose severe limits on the practical paths of Darwinian evolution (and thus on the practical growth of complexity). Again, once we recognise the *possibility* of evolution within a framework of Genetic Relativism—which offers the possibility, in effect, of changing, or jumping between, *different* genetic networks—the *practical* possibilities for the (Darwinian) growth of complexity are evidently greatly increased.

This last point represents a quite different reason for favouring the framework (or perhaps we may now say "research programme") of Genetic Relativism, and it is independent of the "power" of particular core constructive automata. In particular, even if we can exhibit a single full blown general constructive automaton, which yields a mutationally connected set of self-reproducers spanning (virtually) every possible behaviour supported in the system, there could still be advantages, from the point of view of supporting Darwinian evolution, in identifying alternative constructive automata, defining alternative genetic networks (viewed now as evolutionarily accessible pathways through the space of possible automaton behaviours).

Indeed, this need not be all that difficult to do: it provides a particular reason to consider combining a basic constructive automaton with a turing machine (or something of similar computational powers): the latter is arranged so that it "pre-processes" the description tape in some (turing computable) fashion. The program of the turing machine could then effectively encode a space of alternative genetic languages (subject to the primitive constructional abilities of the original constructive automaton); with moderately careful design, it should be possible to open up an essentially infinite set of constructive automata, which are themselves connected under mutation (of the program for the embedded turing machine—another tape of some sort), thus permitting a multitude of *different* genetic networks for potential exploitation by a Darwinian evolutionary process. This should greatly enhance the possibilities for Darwinian evolution of *any* sort, and thus, in turn, for evolution involving the growth of complexity.

This particular idea seems to have been anticipated by Codd:

> A further special case of interest is that in which both a universal computer and a universal constructor [*sic*] exist and the set of all tapes required by the universal constructor is included in the Turing domain T. For in this case it is possible to present in coded form the specifications of configurations to be constructed and have the universal computer decode these specifications ... Then the universal constructor can implement the decoded specifications. Codd (1968, pp. 13–14)

While Codd did not elaborate on *why* such flexibility in "coding" should be of any special interest, it seems plausible that he had in mind precisely the possibility of opening up alternative genetic networks.

Conclusion

I close this brief review with two final remarks relating to Genetic Relativism.

Firstly, von Neumann himself seems to have discounted even the *possibility* of Genetic Relativism being applicable to his models. In his discussion of different kinds of mutations, he stated explicitly that mutations affecting that part of a description tape coding for the core part of the self-reproducer (i.e. coding for u_0 in the terms used above) would result in the production of "sterile" offspring (von Neumann, 1949, p. 86): the implication is that this would *always* be the outcome of such mutations. I suggest that such a claim is too strong, in general. My view is that, on von Neumann's model, it is probably fair to say that such mutations would *almost* always yield sterile offspring—but that, depending on the detailed design of the constructive automaton, and the nature of the particular mutation, there *might* be exceptional cases where the offspring would still be an self-reproducer, but containing an altered core constructive automaton.

Secondly, when tackling von Neumann's problem within the framework of Genetic Absolutism, it was *necessary* to assume a degree of compositionality in the genetic language, to assure that there would exist a range of mutations *not* affecting the core constructive automaton in a self-reproducer; without this assumption it would be difficult, if not impossible, to argue that the set of self-reproducers anchored on this single core general constructive automaton would be connected under mutation. This compositionality assumption is more or less equivalent to the biological hypothesis of *Genetic Atomism*, which holds that genomes may be systematically decomposed into distinct *genes* which, individually, have absolute effects on phenotypic characteristics (see McMullin 1992*b*, p. 11; Dawkins 1989, p. 271). This again represents a divergence between von Neumann's pragmatically convenient solution schema for his particular problem, and the realities of the biological world (where any simple Genetic Atomism is quite untenable). I conjecture therefore that, should we wish to move away from a strict Genetic Absolutism in our formal or artificial systems we might well find it useful, if not essential, to abandon simple compositionality in our genetic language(s) (i.e. Genetic Atomism) also. This, in turn, would ultimately lead away from self-reproducer architectures in which there is any simple or neat division between the core constructive automaton and the rest of the automaton (though there might still be a fairly strict separation of the description tape—i.e. a genotype/phenotype division).

References

Codd, E. F. (1968), *Cellular Automata*, ACM Monograph Series, Academic Press, Inc., New York.

Dawkins, R. (1989), *The Selfish Gene*, new edn, Oxford University Press, Oxford.

McMullin, B. (1992*a*), Artificial Knowledge: An Evolutionary Approach, PhD thesis, Ollscoil na hÉireann, The National University of Ireland, University College Dublin, Department of Computer Science.
`http://www.eeng.dcu.ie/~alife/bmcm_phd/`

McMullin, B. (1992*b*), Essays on Darwinism. 3: Genic and Organismic Selection, Technical Report `bmcm9203`, School of Electronic Engineering, Dublin City University, Dublin 9, Ireland.
`http://www.eeng.dcu.ie/~alife/bmcm9203/`

McMullin, B. (2000), John von Neumann and the Evolutionary Growth of Complexity: Looking Backwards, Looking Forwards..., *in* 'Proceedings of the Conference Artifical Life VII'.

Ray, T. S. (1992), An approach to the synthesis of life, *in* C. G. Langton, C. Taylor, J. D. Farmer and S. Rasmussen, eds, 'Artifical Life II', Vol. X of *Series: Sante Fe Institute Studies in the Sciences of Complexity*, Addison-Wesley Publishing Company, Inc., Redwood City, California, pp. 371–408. Proceedings of the workshop on Artificial Life held February, 1990, in Sante Fe, New Mexico.

von Neumann, J. (1949), Theory and Organization of Complicated Automata, *in* A. W. Burks, ed., 'Theory of Self-Reproducing Automata [by] John von Neumann', University of Illinois Press, Urbana, pp. 29–87 (Part One). Based on transcripts of lectures delivered at the University of Illinois, in December 1949. Edited for publication by A.W. Burks.