

# Toward the Physical Implementation of Artificial Self-Replicating and Evolving Systems

Hiroki Sayama

*The University of Electro-Communications, Tokyo, Japan, sayama@hc.uec.ac.jp*

**Abstract:** Artificial Life is an interdisciplinary research field on the simulation and synthesis of living systems [1,2]. Although it is about fifteen years past since the advent of the field, many important questions are still to be addressed [3]. One of the most challenging questions is how to create an artificial system that demonstrates robust self-replication and evolution in the real world. This question was initiated by von Neumann's famous self-replicating automaton [4-6], which has remained out of physical implementation for more than five decades due to its fragility against perturbations [7]. In this paper, I present a "workplace construction" model [8] as a possible solution for enhancing the robustness of the self-replication processes. Workplaces are assumed to be solid structures that can be easily assembled under perturbations and can rigidly hold other components during the automaton construction. Similar strategies are found to be prevalent in natural organisms, suggesting the effectiveness of the proposed idea for the physical implementation of artificial self-replicating and evolving systems.

**Key words:** artificial life, self-replication, self-replicating automata, robustness, workplace construction.

## 1. Introduction

Artificial Life, advocated by C. Langton in 1987 [1,2], is an emergent interdisciplinary field of research on the simulation and artificial synthesis of living systems that demonstrate various life-like behaviors. The approaches used there range quite widely, from computational modeling of virtual organisms that evolve in an artificial universe created within a computer [9-11], to physical implementation of robots wandering on an assembly table [12,13], to biochemical synthesis of self-replicators using RNAs in a test tube [14].

In Artificial Life, the state of "living" is understood as a kind of dynamic organization of a system which is independent of its physical realization. This claims that, once all the key interactions acting between components of living organisms are identified and understood, similar systems may be made artificially using different materials, and when it is done, the system being made is considered as "living", in the sense that it shares the same characteristic properties of the imitated original ones. This viewpoint is believed to expand the subject of biology from the range of terrestrial life forms we can

observe so far (*life-as-we-know-it*) to the inclusion of any kind of possible life forms that share the same characteristics of life (*life-as-it-could-be*) [1]. Then we will be able to distinguish the essential nature of life in general from material-dependent features of specific terrestrial life. An synthetic approach to artificially create virtual creatures in a virtual world and investigate them as experimental subjects plays a crucial role in the research efforts on *life-as-it-could-be*. For this purpose, computer simulations of artificial organisms have been used quite extensively in the light that the essential aspects of life are the inheritance and variation of genotypic / phenotypic information.

Recently, there have been increasing efforts to evaluate past achievements and lay out future challenges in the field of artificial life [3,15]. One of the most interesting but difficult grand challenges discussed there is to create an artificial system that demonstrates robust self-replication and evolution in the real physical world. This question was initiated by John von Neumann's seminal work on his very complex, sophisticated self-replicating and evolving automaton model that possesses the capability of universal construction [4-6] (it will be briefly summarized in the next section). A number of succeeding studies followed this work [16], implementing much simpler artificial self-replicating systems using theoretical universes, e.g. cellular automata [17-22] or core memory spaces [9,11,23-27]. Unlike his another well known idea of universal computer, however, von Neumann's universal constructor has never been subject to physical implementation. This is mainly due to the fragility of its mechanisms against perturbations that are unavoidable in more realistic kinematic settings, and the same problem has been inherited by all the succeeding models as well.

In this paper, I discuss systematically what are the key issues to be addressed to approach the physical implementation of artificial self-replicating and evolving systems, using the general framework of von Neumann's automaton model [4]. In particular, as a possible solution (at least at conceptual level) to the fragility of the self-replicating mechanism, I present a simple idea of enhancing the robustness of the self-replication processes by introducing an additional subsystem that constructs a "workplace" prior to automaton construction [8]. Workplaces are assumed to be solid structures that can be easily assembled under perturbations and can rigidly

hold other components during the automaton construction processes. Similar kind of strategies are found to be prevalent in biological organisms, which suggests the effectiveness of the presented idea for the realization of kinematic models of self-replication and evolution.

## 2. Von Neumann and Self-Replicating Automata

It is well acknowledged that John von Neumann, a great mathematician / physicist who may be best known as a father of the programmable architecture adopted in today's computers, is also one of the pioneers of artificial life for his seminal work on the self-replicating and evolving automaton [4-6]. In his latest years, von Neumann challenged an empirical rule that was believed in engineering disciplines that the complexity of products is always smaller than that of the manufacturing machines that produce them. To grapple with this problem, he developed a theoretical discussion and concluded that a counterexample to this rule could exist if a system is made of the following [4]:

- A: a universal constructor that creates any arbitrary structure by referring to a static “description tape”.
- B: a tape duplicator that makes a copy of the description tape.
- C: a controller that reads the description tape and passes the written information to the above A and B and appropriately coordinates their behaviors.
- $I_{A+B+C}$ : a description tape that specifies how to construct the entire system itself.

These are symbolically written as

$$A + I_X \quad A + I_X + X, \quad (1)$$

$$B + I_X \quad B + I_X + I_X, \quad (2)$$

$$(A + B + C) + I_X \quad (A + B + C) + I_X + X + I_X, \quad (3)$$

$$(A + B + C) + I_{A+B+C} \\ (A + B + C) + I_{A+B+C} \\ + (A + B + C) + I_{A+B+C}, \quad (4)$$

where the last form represents a self-replicating process. More importantly, this model also captures the capability of the evolutionary growth of complexity [28]: If the description tape happens to contain an additional content that does not interfere the correct functioning of  $A + B + C$ , the system produces a mutated system, possibly of higher complexity than its parent, i.e.,

$$(A + B + C) + I_{A+B+C+F} \\ (A + B + C) + I_{A+B+C+F} \\ + (A + B + C + F) + I_{A+B+C+F}, \quad (5)$$

where  $F$  is an additional component that happened to emerge in the tape by mutation. The existence of and the relationship between phenotypes ( $A + B + C$ ,  $A + B + C + F$ ) and genotypes ( $I_{A+B+C}$ ,  $I_{A+B+C+F}$ ) illustrated in this formulation hold a close resemblance with those found in

reproduction and variation of real (asexually reproducing) organisms.

Later, von Neumann tried to show that such a counterexample *does* exist, at least theoretically, by implementing his very complex universal constructor in a 29-state 5-neighbor 2D cellular automata space—an artificial but totally mechanistic universe. Although he passed away before the completion of this work, it was posthumously completed and published by Burks in 1966 [5]. Von Neumann's idea of self-replicating automata was so profound that it was followed by a number of succeeding studies, which now forms one of the central parts of artificial life [6,16].

## 3. Universal Computation vs. Universal Construction

Many think of von Neumann's work on self-replicating machines as something different, and even weird, compared to his other (probably more famous) achievements in computer science, game theory, quantum physics, nuclear physics, and so on. However, it is worth pointing out that his idea of universal construction with instruction tapes has a fundamental correspondence with his another idea of universal computation with program codes stored in memory. These two are actually the same in that a system contains an arbitrary sequence of information inside itself, and provided it has a universal capability, it can imitate the behavior of any arbitrary system described in the sequence. Presumably von Neumann had no substantial distinction between these two in his mind when he considered these ideas.

There is a practical difference, however, between the universal computer and the universal constructor in terms of the requirement for the robustness of the systems to perturbations. The universal computer does not have to be robust by itself, because the computation theory assumes discrete mathematical entities and their state transition in a noiseless world. The responsibility of realizing such a noiseless condition is primarily on the device manufacturing side and not on the system itself. On the other hand, the universal constructor should be robust *per se*, because its construction capability best makes sense in our real, three-dimensional kinematic universe. Thus the system should be able to deal with more or less continuous physical entities that inevitably involve fluctuations and uncertainty. In such settings, the system itself *is* responsible for the robustness of its workings.

The robust universal construction is apparently a very difficult problem to attain. Von Neumann himself tried to consider a kinematic model at first, but later he abandoned it and switched to cellular automata to avoid this difficulty. Although his work was still monumental and stimulative enough even on cellular automata, the lack of robustness was crucial when considered as a model of real biological and/or engineering systems.

Therefore, in spite that his universal computer has been implemented hundreds of millions of times (i.e., computers we use nowadays), there has been *no* physical implementation of his universal constructor [7]. So the relevant question to ask here is: *How can we enhance the robustness of self-replication processes?*

#### 4. Workplace Construction

In this section I present a simple extension of von Neumann's theoretical model, which I tentatively call a *workplace construction model*, to illustrate a possible solution for the question mentioned above. The key idea is the introduction of an additional subsystem that constructs a “workplace” prior to automaton construction. This model can be described within von Neumann's framework by putting the added subsystem as a variant of automaton  $F$  in his formulation, except for three additional assumptions I am going to make in what follows.

The first assumption is a reasonable statement that we all know empirically:

**Assumption 1:** The construction and duplication processes that are originally sensitive to perturbations can be substantially stabilized by putting them on some solid supporting structure, or *workplace*.

For example, drawing a straight line is rather difficult by a free hand but is quite easy and precise by using a ruler. Using this assumption, I symbolically represent such stabilized processes by

$$A + I_X + S^A_X \quad A + I_X + S^A_X + X, \quad (6)$$

$$B + I_X + S^B_X \quad B + I_X + S^B_X + I_X, \quad (7)$$

$$(A + B + C) + I_X + S^A_X + S^B_X \\ (A + B + C) + I_X + S^A_X + S^B_X + X + I_X, \quad (8)$$

where  $S^A_X$  and  $S^B_X$  represent the workplaces that support automaton construction and tape duplication, respectively. Subscript  $X$  means that the size and/or shape of the workplaces may depend on the product, if not always. The bold right arrow “ $\rightarrow$ ” denotes that the process is significantly robust to perturbations, i.e., the outcome is not affected virtually by small but positive amount of perturbations. Here let us keep ourselves quite loose in evaluating the robustness; I just classify processes as either “robust” or “sensitive” to perturbations. Such a simplification enables us to think about the problem more clearly and concisely. Detailed quantitative analysis of the robustness is far beyond the range of this brief paper.

The above forms (6) (8) look quite similar to the original ones (1) (3), except for the addition of the two workplaces, so one may want to simply apply  $X = (A + B + C) + S^A_X + S^B_X$  to the last form to obtain a robust self-replication process. However, this would result in a non-trivial problem: By definition, workplaces are

assumed to hold and support the construction process of the product, so its size in general should be equal to or greater than that of the product. Therefore, if one lets  $X = (A + B + C) + S^A_X + S^B_X$ , then, for example,  $S^A_X$  must be large enough to hold the entire product  $X$  that includes  $S^A_X$ , resulting in a vicious circle that never closes.

Note that this problem is deeply related to the complexity decreasing rule empirically seen in engineering: To obtain a highly organized product, one needs a stabilized and well controlled manufacturing process, which requires a machine that can control the local environment that should be equal to or larger than the product. Part of the reason why von Neumann's machine can seemingly construct a product more complex (larger) than itself is, in this context, because his machine is implicitly supported by an infinite array of cellular automata that virtually works as a solid workplace. This, however, never applies to our real world where the original complexity decreasing rule came from. In other words, he intended to overcome this empirical rule, but what he actually chose was avoiding it by putting away the crux of the issue out of consideration.

In order to resolve the above problem, here I make the second assumption:

**Assumption 2:** Workplaces are generally made of a simple but extensive repetition of the same kind of components, so they can be generated with enough preciseness by itself (without another workplace) even under perturbations.

Readers may wonder if making this assumption might be just another way of avoiding the problem. Since the main aim of this paper is to show the argument and promote discussions on it, I do not mean this is the one and only right solution. Nonetheless, many empirical observations seem to support this assumption. For example, one can create a straight line in a robust fashion by combining tiny rods into triangular meshes or trusses. As long as the tiny rods are at the same length, the outcome can be precise enough. This is intuitively because the product of this process is not complex; it can be produced by a repetition of the same simple tasks, which is not the case for more sophisticated processes like automaton construction.

With this second assumption, I introduce a set of new robust construction processes, i.e.,

$$R^A + I_X \quad R^A + I_X + S^A_X, \quad (9)$$

$$R^B + I_X \quad R^B + I_X + S^B_X, \quad (10)$$

where  $R^A$  and  $R^B$  are subsystems that estimate the size of needed workplaces for  $A$  and  $B$  from the description tape and construct them in a robust fashion. Note that the exact estimation of the size of needed workplaces might be another complex task that needs another workplace to carry it on. We thus have to make the third assumption:

**Assumption 3:** The size (or the upper bound of the

size) of workplaces for construction and duplication processes can be estimated from the instruction tape in a simple operation, even under perturbations.

This assumption is much less obvious than the previous two, and should be subject to discussion and verification. If it would not be the case in some condition, the arguments developed below would lose its universality, but I believe it would still have sufficient implications for how to create robust systems.

Provided the third assumption applies, the previous form of stabilized automaton construction (8) can be rewritten as

$$\begin{aligned} & (A + B + R^A + R^B + C) + I_X \\ & (A + B + R^A + R^B + C) + I_X + S^A_X + S^B_X \\ & (A + B + R^A + R^B + C) + I_X + S^A_X + S^B_X \\ & + X + I_X, \end{aligned} \quad (11)$$

where the entire system  $(A + B + R^A + R^B + C)$ , which we call  $G$  hereafter, is now capable of both constructing  $X$  and duplicating  $I_X$  in a robust manner, while keeping its own size finite and independent of what is written in  $I_X$ . Note that the function of controller  $C$  is more complex than before; it now has to coordinate the behaviors of four other subsystems:  $A$ ,  $B$ ,  $R^A$ , and  $R^B$ .

Then, finally, we put the system  $G$  itself into  $X$  in the above form, to obtain

$$\begin{aligned} G + I_G & \quad G + I_G + S^A_G + S^B_G \\ G + I_G + S^A_G + S^B_G & + G + I_G. \end{aligned} \quad (12)$$

This represents a robust self-replication process of  $G$ , which also produces some byproducts ( $S^A_G$  and  $S^B_G$ ).

## 5. Discussion

The workplace construction model presented above is by no means theoretically or experimentally proven. It uses three assumptions, each of which must be carefully checked about its validity. Whether such a robust automaton can be actually implemented is another key issue to be investigated, probably with a lot of efforts, just like what von Neumann did with cellular automata to show a concrete example of his self-replicating machine.

Despite these problems all, one can obtain a supportive implication for the model from a variety of phenomena observed in real biological systems at various scales. They imply the applicability and effectiveness of the central idea of the presented model. For example, at the smallest level, the formation of cell membranes is probably the most fundamental instance of workplace construction; it isolates the metabolic process of the cell from environmental perturbations and keeps all the sensitive parts at the same place without diffusion. At a much higher level, niche construction seen in ecology [29] is another clear instance of workplace construction, such as dam building by beavers or development of artificial living environment by humans. These controlled environments are literally “workplaces”

for their activities.

More directly relevant cases can be found in animals that are about to produce their next generation. For example, firm eggshells of birds and reptiles are probably the most direct example of workplace construction. Although the automaton constructors are embedded in the fetuses in these cases, the purpose of eggshells are exactly the same as is discussed in this paper, i.e., to stabilize the process of offspring construction by isolating it from the outside and holding it on a solid structure. More sophisticated instance is the uterus of mammals, where the workplace is included in the parent's body, but can be refurbished and extended as needed to hold its offspring under construction. It is also interesting that the production of  $S^A_G$  and  $S^B_G$  in (12) correctly captures the nature of the process that these workplaces constructed in real organisms (eggshells, endometria, placentas, etc.) are all just for temporary use and they will eventually become garbage after birth of the offspring.

The only point I wish to make, by presenting the workplace construction model, is the following trivial fact: the more complex a system becomes, the better controlled local environment the system needs in order to construct its replica under perturbations. Biological organisms seem to have evolved such sophisticated “workplaces” for their survival and prosperity. However, this point has been long missing in the earlier studies of artificial life, and it must be taken into account to proceed toward the next step of this interesting research field.

## References

- [1] C. G. Langton (1989) “Artificial Life”, in C. G. Langton, ed., *Artificial Life*, Santa Fe Institute Studies in the Sciences of Complexity, vol. VI, 1 47, Addison-Wesley, Redwood City, CA.
- [2] C. G. Langton, ed. (1995) *Artificial Life: An Overview*, MIT Press, Cambridge, MA.
- [3] M. A. Bedau, J. S. McCaskill, N. H. Packard, S. Rasmussen, C. Adami, D. G. Green, T. Ikegami, K. Kaneko, and T. S. Ray (2000) “Open Problems in Artificial Life”, *Artificial Life* 6:363 376.
- [4] J. von Neumann (1951) “The General and Logical Theory of Automata”, in L. A. Jeffress, ed., *Cerebral Mechanisms in Behavior The Hixon Symposium*, 1 41, John Wiley, New York, NY. Originally presented in September, 1948. Also collected in (1987) W. Aspray, and A. W. Burks, eds., *Papers of John von Neumann on Computing and Computer Theory*, 391 431, MIT Press, Cambridge, MA.
- [5] J. von Neumann (1966) *Theory of Self-Reproducing Automata*, University of Illinois Press, Urbana, IL. Edited and completed by A. W. Burks.
- [6] P. Marchal (1998) “John von Neumann: The Founding Father of Artificial Life”, *Artificial Life* 4:229 235.

- [7] G. Friedman (1996) "The Space Studies Institute View on Self-Replication", *The Assembler: Newsletter of the Molecular Manufacturing Shortcut Group of the National Space Society* 4:4 (<http://www.islandone.org/MMSG/9612.html>).
- [8] H. Sayama (2002) "Von Neumann's Machine in the Shell: Enhancing the Robustness of Self-Replication Processes", in *Artificial Life VIII: Proceedings of the Eighth International Conference on Artificial Life*, in press.
- [9] T. S. Ray (1991) "An Approach to the Synthesis of Life", in C. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, eds., *Artificial Life II*, Santa Fe Institute Studies in the Sciences of Complexity, vol. XI, 371-408, Addison-Wesley, Redwood City, CA.
- [10] K. Sims (1994) "Evolving 3D Morphology and Behavior by Competition", in R. Brooks and P. Maes, eds., *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, 28-39, MIT Press, Cambridge, MA.
- [11] R. E. Lenski, C. Ofria, T. C. Collier, and C. Adami (1999) "Genome Complexity, Robustness, and Genetic Interactions in Digital Organisms", *Nature* 400: 661-664.
- [12] P. Husbands and J. A. Meyer, eds. (1999) *Evolutionary Robotics: Proceedings of the First European Workshop on Evolutionary Robotics (EvoRobot'98)*, Lecture Notes in Computer Science 1468, Springer-Verlag, Berlin.
- [13] H. Lipson and J. B. Pollack (2000) "Automatic Design and Manufacture of Robotic Lifeforms", *Nature* 406:974-978.
- [14] G. F. Joyce (1997) "Evolutionary Chemistry: Getting There from Here", *Science* 276:1658-1659.
- [15] H. Sayama, Y. Suzuki, and K. Takadama, eds. (1999) *What's Been Done, What's Not Yet, and What Comes Next in Artificial Life: Position Papers for the 3rd ALIST (Artificial Life Independent Seminar in Tokyo): ALIREN+ALIST Joint Meeting*, March 1999, Tokyo, Japan (presented in Japanese), available at <http://complex.hc.uec.ac.jp/sayama/alist/>
- [16] M. Sipper (1998) "Fifty Years of Research on Self-Replication: An Overview", *Artificial Life* 4:237-257.
- [17] E. F. Codd (1968) *Cellular Automata*, ACM Monograph Series, Academic Press, New York.
- [18] C. G. Langton (1984) "Self-Reproduction in Cellular Automata", *Physica D* 10:135-144.
- [19] J. A. Reggia, S. L. Armentrout, H. H. Chou, and Y. Peng (1993) "Simple Systems That Exhibit Self-Directed Replication", *Science* 259:1282-1287.
- [20] H. H. Chou and J. A. Reggia (1997) "Emergence of Self-Replicating Structures in a Cellular Automata Space", *Physica D* 110:252-276.
- [21] H. Sayama (1999) "A New Structurally Dissolvable Self-Reproducing Loop Evolving in a Simple Cellular Automata Space", *Artificial Life* 5:343-365.
- [22] H. Sayama (2000) "Self-Replicating Worms That Increase Structural Complexity through Gene Transmission", in M. A. Bedau, J. S. McCaskill, N. H. Packard, and S. Rasmussen, eds., *Artificial Life VII: Proceedings of the Seventh International Conference on Artificial Life*, 21-30, MIT Press, Cambridge, MA.
- [23] A. K. Dewdney (1984) "Computer Recreations: In the Game Called Core War Hostile Programs Engage in a Battle of Bits", *Scientific American* 250(5):14-22.
- [24] S. Rasmussen, C. Knudsen, R. Feldberg, and M. Hindsholm (1990) "The Coreworld: Emergence and Evolution of Cooperative Structures in a Computational Chemistry", *Physica D* 42:111-134.
- [25] A. N. Pargellis (1996) "The Evolution of Self-Replicating Computer Organisms", *Physica D* 98:111-127.
- [26] T. S. Ray and J. Hart (1998) "Evolution of Differentiated Multi-Threaded Digital Organisms", in C. Adami, R. K. Belew, H. Kitano, and C. E. Taylor, eds., *Artificial Life VI: Proceedings of the Sixth International Conference on Artificial Life*, 295-304, MIT Press, Cambridge, MA.
- [27] H. Suzuki (2000) "Evolution of Self-reproducing Programs in a Core Propelled by Parallel Protein Execution", *Artificial Life* 6:103-108.
- [28] B. McMullin (2000) "John von Neumann and the Evolutionary Growth of Complexity: Looking Backwards, Looking Forwards...", in M. A. Bedau, J. S. McCaskill, N. H. Packard, and S. Rasmussen, eds., *Artificial Life VII: Proceedings of the Seventh International Conference on Artificial Life*, 467-476, MIT Press, Cambridge, MA.
- [29] K. N. Laland, J. Odling-Smee, and M. W. Feldman (2000) "Niche Construction, Biological Evolution, and Cultural Change", *Behav. Brain Sci.* 23:131-175.