

Designed and Evolved Blueprints For Physical Self-Replicating Machines

Efstathios Mytilinaios, Mark Desnoyer, David Marcus, and Hod Lipson

Computational Synthesis Lab, School of Mechanical & Aerospace Engineering,
Cornell University, Ithaca NY 14853, USA
hod.lipson@cornell.edu

Abstract

Self-replication is a process critical to natural and artificial life, but has been investigated to date mostly in simulation and in abstract systems. The near absence of physical demonstrations of self-replication is due primarily to the lack of a physical substrate in which self-replication can be implemented. This paper proposes a substrate composed of simple modular units, in which both simple and complex machines can construct and be constructed by other machines in the same substrate. A number of designs, both hand crafted and evolved, are proposed.

Introduction

Self-replication is a process critical to both natural and artificial life, but has been investigated to date mostly in simulation and in abstract systems (Sipper and Reggia, 2001). Even in formal systems, self-replication has been difficult to define consistently (Nehaniv and Dautenhahn, 1998; Sanchez *et al* 1997; McMullin, 2000). Self-replication has been examined empirically numerous times, mostly in the context of replicating programs, cellular automata, and artificial chemistries (e.g. Arbib, 1965; Burks, 1970; Moore, 1970; Langton, 1984; Lohn and Reggia, 1997). In contrast, physical machines capable of self-reproduction have been scarcely discussed.

The practical potential of physical self-reproduction was recognized in the early 80's as a possible method for remote colonization (Freitas and Gilsbreath, 1981) but was abandoned due to many unresolved technical difficulties. More recently, interest is being revived in this area as a fabrication paradigm for both macroscopic space applications and for micro- and nano-scale systems (Freitas, 2003). At small scales, such as at the molecular level, self-replication can occur through *stochastic* self-assembly processes catalyzed by a self-replicating entity. At large scales, such as multicellular organisms and robots, self-replication through stochastic processes is energetically implausible, and explicit reproduction must take place.

Artificial physical self-replication at macro-scale was first demonstrated using stochastic tumbling blocks with special geometries (Penrose, 1959). Deterministic self-reproduction of robotic systems has only recently been demonstrated by Chirikjian (2002), where a Lego™ robot composed of three modules was able to assemble three

other modules into a new identical robot. To do so, the base module of the robot followed a path drawn on the ground, pushing the other modules and joining them into an assembly using magnetic connections or with the assistance of an external passive joining rig.

Chirikjian's work demonstrated a physical, deterministic self-reproducing machine for the first time, but there is more to be desired. If a robot composed of two components, for example, is able to reproduce by actively assembling these components together, then self-replication has indeed occurred, but in a way nowhere as impressive as, say, a machine able to reproduce itself from raw materials or as biological life that is able to reproduce from amino acids.

The apparent existence of different *levels* of self-replication has led us to abandon the view of self-replication as a binary quality, simply existing or not. Instead, we see self-replication (or self-reproduction) as a continuum, quantifiable based on the amount of information being replicated. For example, the amount of information needed to reproduce a two-component robot *given* these two components is less than the amount of information needed to reproduce a robot given many lower-level components. Similarly, the amount of information needed to reproduce an inaccurate copy of a machine is less than the amount of information necessary to produce a more exact duplicate. The amount of information needed to assemble a machine that is independently likely to spontaneously appear in a domain is less than the information needed to assemble a unique machine unlikely to appear by chance. The amount of information involved in the replication is dependent on the definition of the replicating system and the contribution of the environment to the replication process. Based on these assumptions, we have formally defined a domain-independent metric of self-replication (Adams and Lipson, 2003). The metric compares the probability of spontaneous emergence of a system in a given environment, to the probability of emergence of a system given that one instance of the system is already present in the environment. This metric not only avoids some of the difficulties of earlier definitions (how to deal with trivial replicators, and how to apply it to non-formal systems), but also provides a graded value that is more amenable to evolution, if a self-replicating machine is to be evolved directly.

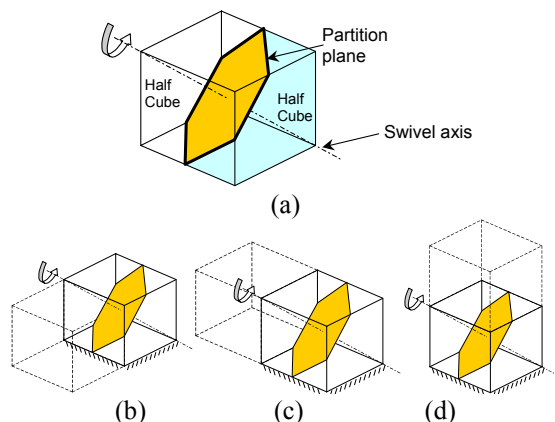


Figure 1: A 3D molecule: (a) Two halves split across the (1,1,1) plane. (b,c,d) Swiveling the top half causes an adjacent block to cycle into new configurations.

Following this formal definition, we have sought a physical substrate in which a variety of self-reproducing machines can systematically be constructed, ranging from simple to complex replicators. We also seek a substrate that is both physically plausible and faithfully simulatable, so that questions about the origin of physical self-replication can be studied.

This paper reports on some of our progress towards this goal. We first describe a physically-plausible substrate, both in its two-dimensional and three-dimensional versions. We then show a number of hand-designed machines and evolved machines, and the way in which they were evolved. Finally, we show a preliminary physical implementation of this substrate.

A space for physical self-replication

The space of machines we put forward is based on a single cubical building block – a *molecube*. Each cube, shown in Figure 1 below, can connect to its adjacent neighbors. Patterns on the surface of the cubes assure that the assemblies are perfectly aligned when joined. Specific bonding and patterning mechanisms depend on the scale of the implementation, and can be magnetic, electrostatic, or hydrophilic/hydrophobic, for example.

A cube is split into two parts along a plane that is perpendicular to its long diagonal (e.g. 1,1,1); the plane is shaded in Figure 1a. One half of the cube can swivel about the long axis in increments of 120 degrees, each time cycling the faces of the cube. If magnets on all three faces on one half have a polarity of north, and magnets on the three faces of the other half have a polarity of south, then polarities are retained after swiveling and then the structure will always have consistent global polarities.

Swiveling a cube while other cubes are attached to it causes a reconfiguration of the structure. For example, the cube shown in Figure 1b has another cube on its left. Swiveling this cube once causes the adjacent cube to move to the back (Figure 1c). Swiveling once more causes the

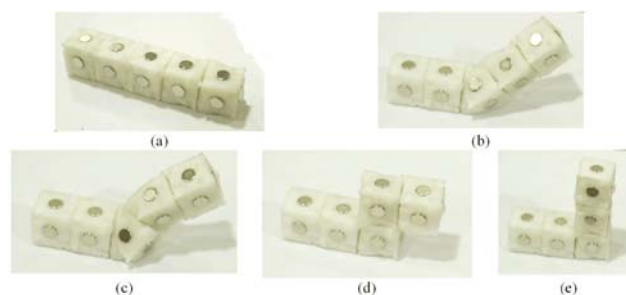


Figure 2: Physical model of a 3D molecule: (a) A single block, with magnets and swivel line. (b-f) a sequence of swivels that transform an object from a line to a Z-shape to an L-shape.

adjacent cube to move to the top (Figure 1d). Swiveling again will restore the original configuration (Figure 1b). Each cube thus has 3 possible swivel states, and 4 possible orientations for the swivel axis.

Figure 2 shows a physical model of a set of five building blocks. Two of these blocks are swiveled 120 degrees in turn, causing the entire structure to reconfigure into a Z-shape and then into an L-shape. Other swivels may yield three-dimensional configurations.

Structure topologies may have loops and branches, and multiple blocks can swivel simultaneously. However, while swiveling, a structure needs to go through intermediate states. Due to collisions some of these configuration transformations may not be possible. Similarly, bonding will occur if two cubes become adjacent during reconfiguration and their magnets are in attractive polarities, and the structure will lock. Other physically-realistic constraints may be placed on the structure depending on its environment, such as collision with the ground, gravitational stability, actuation torque limits and motion dynamics. However, because the end positions of the cubes lie at regular grid locations, actuation sequences can be calculated rapidly and simulated without accumulation of error.

A second form of actuation is the ability to change polarities of the faces, so that adjacent faces attract, repel, or are passive. If magnetic bonding is used, then electromagnets can switch between 'north', 'south' and 'off' states. Each cube thus has $3^6=729$ possible states. Transitioning states allows a block to pick up, hold and drop other blocks or groups of blocks, as well as grip and climb other structures.

Control of the machine is specified using a sequence of swiveling and bond-state switching commands, executed in open loop (without feedback). It is possible to envision more elaborate controllers that incorporate sensing, branching, memory, and stochastic elements, as well as distributed control where cubes and groups of cubes execute programs locally.

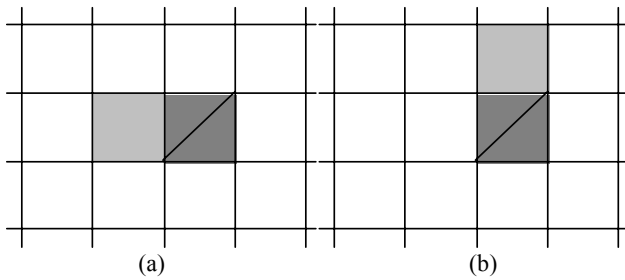
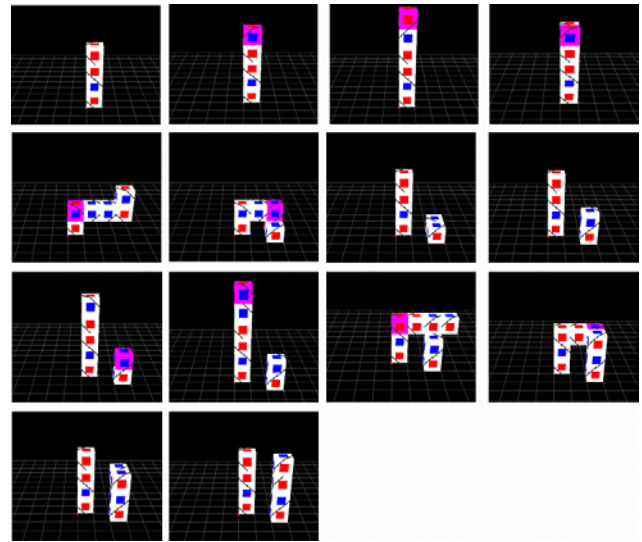


Figure 3 (above): A 2D molecule: (a) Two halves split across the diagonal. (b) Swiveling the top half causes any adjacent blocks to cycle into new configurations.

Figure 4 (right): A self-assisted reproduction sequence (hand designed): Top-left: The original structure contains four cubes. New cubes are dispensed from the top. The machine positions these cubes and the newly formed structure reconfigures during the reproduction to assist in its own construction.



To maintain physical plausibility, power for actuation and bonding is provided through the ground, and transmitted via local connections. Battery operation is possible though inconvenient for physical implementation due to excessive weight and short operation runs.

A two dimensional space

A simpler version of this space exists in two dimensions. In this case each cube is a square; the square is split along its diagonal. A swivel of the square causes two faces to switch as shown in Figure 3. Each square thus has 2 possible swivel states, 2 possible swivel axes and $3^4=81$ possible bonding states. Since the swiveling motion causes the squares to go out of plane during transition, no intermediate collisions exist. Because of this reduced space size and simpler physics, a two-dimensional molecule space is amenable to fast simulation.

The molecube space can also be generalized to other dimensionalities, though it is more difficult to visualize. For example, each four dimensional unit would have 4 possible swivel states, 6 possible swivel axes, and 3^8 possible bonding states.

Possible modes of self-replication

There are a number of ways self-replication may occur in a molecube space. Since any replication process requires material, we assume some grid positions may act as *dispensers*, where new cubes reappear when removed from that location. A machine is considered replicated only when the new copy is *identical* and *detached* from its parent.

- **Direct reproduction:** A machine reconfigures to pick cubes from a dispenser and place them in a new location, gradually building a copy from the ground up.

- **Tandem reproduction:** Multiple machines are required to produce a single copy. One machine may place cubes while the other reorients the constructed machine.
- **Self-assisted reproduction.** The machine being constructed reconfigures during the construction process to facilitate its own construction.
- **Multi-stage reproduction.** Intermediate constructions are required before the target machine can be made. The intermediate machine is then discarded as a waste product, or can be used to catalyze the production of additional machines.

Various combinations and extensions of these modes of operation are possible, and one may imagine an ecology of competing and cooperating machines.

Manually-designed replicators

We initially explored this space manually, looking for possible self-replicating designs. Exploration of the three-dimensional space was carried out using a simulator, able to simulate arbitrary 3D molecube structures and execute sequences of swivel and bonding commands. The simulator accounts for collisions during transformation, loops and locked structures, as well as incompatible bonding polarities and maximum torque loads due to gravity and moment arms.

A number of self-reproducing machines (structure + control) were found. Figure 4 shows one of the simplest machines containing 4 cubes. New cubes are dispensed from the top. The machine positions these cubes, and the newly formed structure reconfigures during the reproduction to assist in its own construction. This is a form of self-assisted reproduction. Larger and more complex machines were found, including a 9-piece machine.

Evolved physical replicators

A more difficult challenge is to *evolve* self-replicators, rather than design them manually. Ideally, such replicators would emerge spontaneously out of a primordial soup of cubes, where – as in nature – self-replication is an *implicit* fitness criterion in itself. However at this initial stage we experimented with direct evolution of replicators where replication is an *explicit* fitness criterion.

Treating self-replication as a binary criterion would not provide any gradient for the evolutionary process to follow, and so would be unlikely to yield any viable solution in this vast space of machines, including both structure and control. Instead we broke down the evolutionary process into two stages, and used the graded definition of self-replication to produce a gradient:

- **Stage 1: Evolve morphologies** of machines that are mechanically capable of reaching an area large enough to contain a detached copy of themselves. The percentage of coverage provides a gradient.
- **Stage 2: Evolve controllers** that would make a given morphology pick cubes from dispensers and place them at the correct position. The number of dispensers needed provides a gradient.

We carried out the initial experiments on the two-dimensional version of molecube as it is faster to simulate and provides a smaller search space. In this particular experiment we also required that each cube be either a swiveling block with permanent magnets or a non-swiveling block with switchable electromagnetic (an 'end effector'), but not both. This restriction was placed both for practical consideration for physical implementation, and also to rule out the trivial solution of a single cube sitting at the dispenser location (this was of course one of the first 'unintended' solutions to be found).

At the initial morphology-search stage, the fitness function first exhaustively mapped out the area that the end-effectors covered, while pruning illegal configurations due to collisions and self-locking. This step can be done in polynomial time using convolution of reconfiguration steps. Once the coverage of the machine was obtained, then a copy of the machine was tried exhaustively to be fitted within that space in any of the four orientations. This step can be done in linear time. The maximum amount of the original structure that would fit in the mapped area provided the fitness for the first stage.

Morphologies were represented as a series of code-pairs. Starting with a cursor at the origin, the first code moves the cursor in one of the four cardinal directions, while the second code defines the type of block to try to place. If there is already a block at that position, the new block is ignored and the cursor continues to move as defined by the next element in the array. Morphology strings may be of variable length, but were limited to under 20 units. Variation was achieved through crossover ($p=0.9$) and three types of mutation: change ($p=0.001$), addition ($p=0.06$), and removal ($p=0.0005$). The population

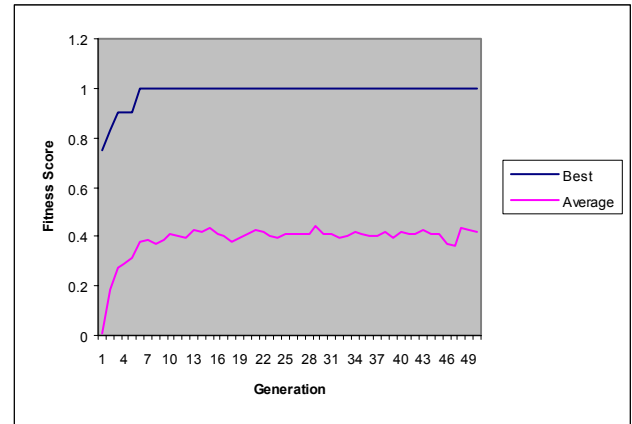
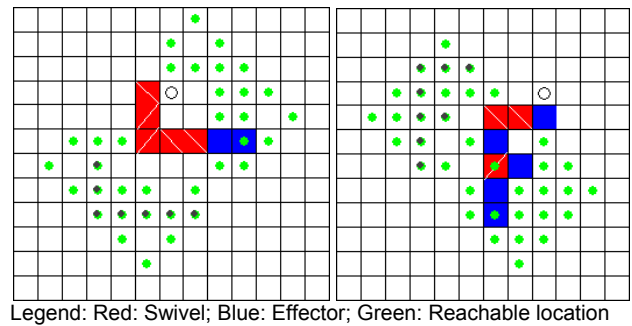


Figure 5: Evolutionary progress for morphology stage (top) Two evolved shapes (morphologies) that can cover an area that contains themselves. (bottom) fitness vs. generations for the run that produced these two results.

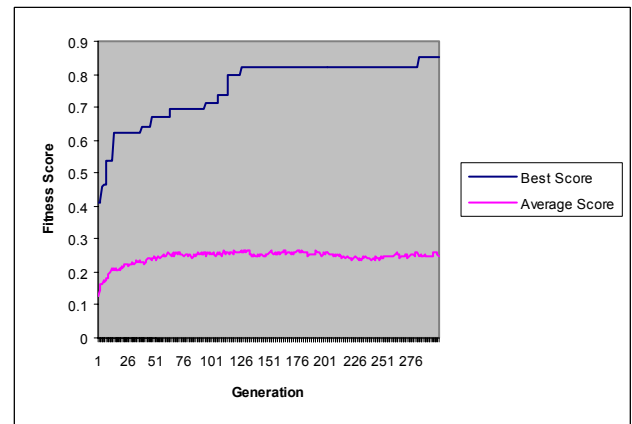


Figure 6: Evolutionary progress for controller stage for the F-Shape.

contained 400 individuals and underwent generational fitness-proportionate selection.

At the second controller-search stage, the fitness function evaluated a control sequence for the given morphology by executing that sequence and measuring the percentage of the potential duplicate that was covered.

Controllers were represented as a series of code-triplets, describing a set of commands to be executed in sequence. Each triplet first described a command ('Swivel', 'Attach', or 'Detach'), and a block number. For the 'Attach'

command, the third parameter also specified which of the four sides a new block should attach to. The attach operation also implicitly defined where dispensers are expected to exist – a factor that influences fitness.

Control strings may be of variable size, but were limited to under 300 commands. Variation was achieved through crossover ($p=0.9$) and three types of mutation: change ($p=0.002$), addition ($p=0.12$), removal ($p=0.001$). Population contained 1000 individuals and underwent generational fitness-proportionate selection for 300 generations. Progress of this stage is shown in Figure 6 for the F-morphology.

The final fitness function used was

$$g(cmds) = 0.8p + 0.2 \frac{c-s+1}{c} - 0.03e$$

where

- p = weighted fraction of the goal location covered
- c = number of blocks inside the model
- s = number of dispensers assumed
- e = number of excess blocks beyond those needed

Of the three morphologies found by the evolutionary process in the first stage, only two were successful at the second stage, yielding a functional, physically-plausible self-replicating machine. The morphologies and sequences of reconfigurations associated with these two machines are shown in Figure 7 and Figure 8.

Physical Realization

Physical realization of the molecule environment involves several challenges, including reliable transfer of power, information and moments among the units. In particular, since each cube involves six sides with four possible orientations for each side, the number of interface contacts can become impractical and unreliable very quickly.

Our design involves two concentric rings on each face of the cube. The inner ring is an electromagnet, and the outer ring is a conductor. Together, they transfer both power and force, and their circularity avoids orientation issues. Control information is passed on top of the power line as a power modulation signal, readable using conventional power-line modem (PLM) components. The power feeds from the base unit into the rest of the machine.

To date we have constructed two physical prototypes, shown in Figure 9. These are 4"x4" ABS cubes, and contain a PIC microcontroller, power modulation and decoding circuitry, electromagnets, power management circuitry, and a servomotor. Though preliminary, they demonstrate that the key challenges in transmission of power, force and data are feasible, and a realization of a simple machine such as the one shown in Figure 4 is within reach.

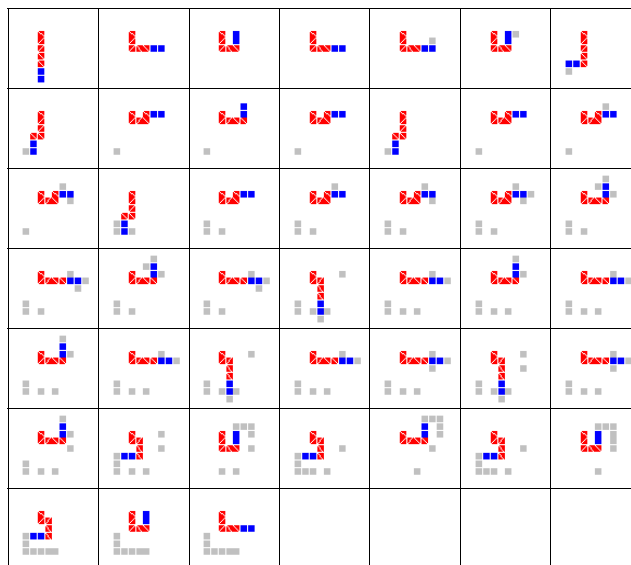


Figure 7: Evolved replication sequence for the L-morphology. Red: Swivel blocks; Blue: Effector blocks; Gray: New material. Note that the sequence contains several redundant cycles.

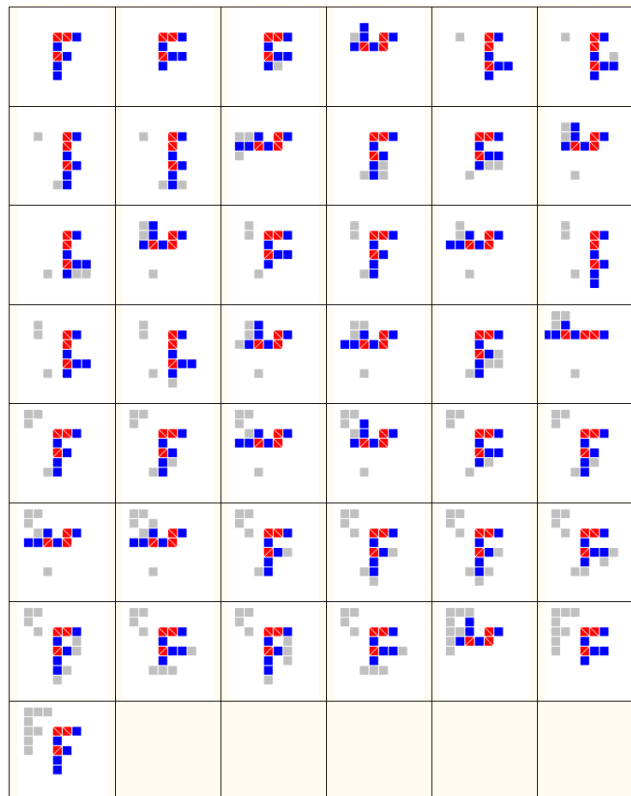


Figure 8: Evolved replication sequence for the F-morphology. Red: Swivel blocks; Blue: Effector blocks; Gray: New material. Note that the sequence contains several redundant cycles.

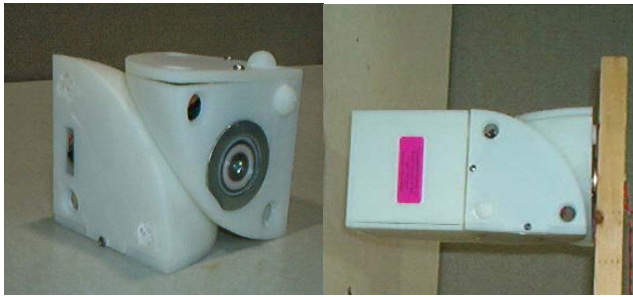


Figure 9: Prototype realization of substrate. (a) The 4"x4" swivel cube, (b) the cube is able to connect, hold, and transmit power and data to adjacent block.

Conclusions

This paper proposes a physical substrate composed of simple modular units, in which both simple and complex machines can construct and be constructed by other machines in the same substrate. A number of designs, both hand crafted and evolved, are proposed, and the key challenges in physical realization of this substrate have been overcome.

Our purpose in this research is to identify a rich substrate in which physically-feasible self-replication can be investigated both in simulation and in reality. The substrate we have identified and propose here is both realizable and easy to simulate quickly and faithfully due to its discrete nature. We have demonstrated preliminary results showing how various self-replicating machines can be found *systematically*. Although our evolutionary processes required user guidance embedded in the fitness function, we expect that future steps along this path will allow us to examine more ways for self-replication to emerge from first principles, thus examining another critical dimension of artificial-life.

Acknowledgments

This paper was funded by US department of energy (DOE), grant #DE-FG02-01ER45902.

References

Adams B., Lipson H., "A metric for self-replication phenomena", In proceedings of the European Conference on Artificial Life, ECAL 2003.

Arbib, M. A., Comments on Self-Reproducing Automata In: Hart, J. F. and Takasu, S, Systems and Computer Science (1965) 42-59

Chirikjian, G.S., Zhou, Y., Suthakorn, J., Self-replicating Robots for Lunar Development, In: IEEE/ASME Trans. on Mechatronics 7. 4 (2002) 462-472

Freitas, R.A., Gilsbreath, W.P., A self-replicating, growing lunar factory In: Proceedings of the Fifth Princeton/AIAA Conference May 18-21 (1981)

Freitas, R.A., Manufacturing Systems for Molecular Nanotechnology: Kine-matic Self-replicating Machines In preparation (2003)

Langton, C. G., Self-Reproduction in Cellular Automata in Physica 10D (1984) 134-144

Lohn J.D. Reggia J.A. (1997). Automatic discovery of self-replicating structures in cellular automata. IEEE Trans. Evolutionary Computation, 1(3):165-178

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

Moore, E. F., Machine Models of Self-Reproduction in Burks, A. W., Essays on Cellular Automata (1970) 187-203

Nehaniv C., Dautenhahn K., Self-Replication and Reproduction: Considerations and Obstacles for Rigorous Definitions. Abstracting and Synthesizing the Principles of Life, Verlag Harri Deutsch, pp. 283-290, 1998.

Penrose, L.S., Self-reproducing machines. In: Scientific American, 200 (6) (1959) 105-114

Sanchez, D., et al., Phylogeny, Ontogeny, and Epigenesis: Three Sources of Biological Inspiration for Softening Hardware. In: T. Higuchi, et al, editors, Proceedings of ICES96, Lecture Notes in Comp. Sci., Vol. 1259, (1997) 35-54.

Sipper, M., Reggia, J.A., Go forth and replicate. In: Scientific American 285/265(2), August 2001, 35-43

Smith, A., Turney, P., and Ewaschuk, R., JohnnyVon: Self-Replicating Automata in Continuous Two-Dimensional Space. (2002) Technical Report ERB-1099, Institute for Information Technology, National Research Council Canada.

Von Neumann, J, completed and edited by Burks, A. W., Von Neumann's Self-Reproducing Automata In: Burks, A. W., Essays on Cellular Automata (1970) 4-65

Hutton T.J., "Evolvable self-replicating molecules in an artificial chemistry", Artificial Life Vol. 8. No. 4, 2002

Bratley P. Millo J. "Computer recreations: Self-reproducing programs." Software Practice and Experience, Vol. 2, pages 397-400, 1972

Ray, T. S. (1992), An Approach to the Synthesis of Life, in C. G. Langton, C. Taylor, J. D. Farmer & S. Rasmussen, eds, 'Artificial Life II', pp. 371-408

Wilke, C. O, J. Wang, C. Ofria, R. E. Lenski, and C. Adami. (2001). Evolution of digital organisms at high mutation rate leads to survival of the flattest. Nature 412:331-333