TR CS-93-11

Simple Systems Exhibiting Self-Directed Replication*

by

James A. Reggia[1]
Steven L. Armentrout[1]
Hui-Hsien Chou[1]
Yun Peng[2]

1. Dept. of Computer Science, University of Maryland, College Park.
2. Dept. of Computer Science, University of Maryland Baltimore County

May 1993

* This article has appeared in *Science*, 26 February 1993, Volume 259, pp. 1282-1286.
Reprint with permission from *Science*.

Simple Systems Exhibiting Self-Directed Replication

James A. Reggia*
Steven L. Armentrout
Hui-Hsien Chou
Yun Peng

Abstract: Biological experience and intuition suggest that self-replication is an inherently complex phenomenon, and early cellular automata models supported that conception. More recently dramatically simpler computational models of self-directed replication called sheathed loops have been developed. It is shown here that "unsheathing" these structures and altering certain assumptions about the symmetry of their components leads to a family of non-trivial self-replicating structures, some substantially smaller and simpler than those previously reported. Dependence of replication time and transition function complexity on initial structure size, cell state symmetry, and neighborhood are examined. These results support the view that self-replication is not an inherently complex phenomenon, but rather an emergent property arising from local interactions in systems that can be much simpler than is generally believed.

J. A. Reggia, S. L. Armentrout, and H. Chou are in the Dept. of Computer Science, A.V. Williams Building, University of Maryland, College Park, MD 20742 . Y. Peng is in the Dept. of Computer Science, University of Maryland at Baltimore County, Catonsville, MD 21228.

*To whom correspondence should be sent.

Mathematicians, computer scientists, and others have been studying artificial self-replicating structures or "machines" for over thirty years (1). Much of this work has been motivated by the desire to understand the fundamental information processing principles and algorithms involved in self-replication, independent of how they might be physically realized (2). A better theoretical understanding of these principles could be useful in a number of ways. For example, understanding these principles may advance our knowledge of the biomolecular mechanisms of reproduction by clarifying conditions that any self-replicating system must satisfy and by providing alternative explanations for empirically observed phenomena (3). Self-replicating systems have thus become a major area of research activity in the field of artificial life (4). Work in this area could also shed light on those contemporary theories of the origins of life that postulate a prebiotic period of molecular replication before the emergence of living cells (5). Additionally, it has been suggested that creating and using self-replicating devices will be important for atomic-scale manufacturing technology, or nanotechnology (6). Unfortunately, much past work on artificial self-replicating structures has been expressed in the technical framework of formal automata theory and has therefore been relatively inaccessible to individuals in other fields.

While the earliest work on artificial self-replicating structures/machines often used mechanical devices (1), subsequent work has been based largely upon computational modelling, especially with cellular automata. The mathematician von Neumann first conceived of using cellular automata to study the logical organization of self-replicating structures (2). In his and subsequent two-dimensional cellular automata models space is divided into cells, each of which can be in one of n possible states. At any moment most cells are in a distinguished "quiescent" or inactive state (designated by a period or blank space in this article) whereas the other cells are said to be in an active state (7). A self-replicating structure is represented as a configuration of contiguous active cells, each of which represents a component of the machine. At each instance of simulated time, each cell or component uses a set of rules called the transition function to determine its next state as a function of its current state and the state of immediate neighbor cells. Thus, any process of self-replication captured in a model like this must be an emergent behavior arising from the strictly local interactions that occur. Based solely on these concurrent local interactions an initially-specified self-replicating structure goes through a sequence of steps to construct a duplicate copy of itself (the replica being displaced and perhaps rotated relative to the original).

Von Neumann's original self-replicating structure was a complex "universal constructor-computer" embedded in a two-dimensional cellular automata space that consisted of 29-state cells. It was literally a simulated digital computer that used a "construction arm" in a step-by-step fashion to construct a copy of itself from instructions on a "tape". Von Neumann's work provided an early demonstration that an artificial information-carrying system capable of self-replication was theoretically possible. It established a logical organization that is sufficient for self-replication, but left open the question of the minimal logical organization necessary for such behavior (2). Subsequent analysis led to several other results: it showed that some simplification of von Neumann's configuration was possible by redesigning specific components (8) or by increasing cell state complexity (9), demonstrated that variations of the configuration could be used to simulate sexual reproduction (10), generalized von Neumann's basic result to other configurations and higher-dimensional cellular spaces (11), established theoretical upper bounds on how rapid a population of self-replicating configurations could grow (12), and examined fundamental definitional issues (13) that continue to generate theoretical interest today (14). Most influential among this early work has been Codd's demonstration that if the components or cell states meet certain symmetry requirements, then von Neumann's configuration could be done in a simpler fashion using cells having only eight states (15). However, although these early studies describe structures that self-replicate, these structures generally consist of tens of thousands of components or active cells, and have thus never actually been simulated computationally because of their tremendous size and complexity.

The complexity of these early cellular automata models seemed consistent with the remarkable complexity of biological self-replicating systems: they appeared to suggest that self-replication is, from an information processing perspective, an inherently complex phenomenon. However, more recently a much simpler self-replicating structure based on 8-state cells, the sheathed loop, was developed (see Fig. 1b) (16). The term "sheathed" here indicates that this structure is surrounded by a covering or sheath (X's in Fig. 1a-c). For clarity, in Fig. 1 and the remainder of this article self-replicating structures are labelled by their type (SL = sheathed loop, UL = unsheathed loop) followed by the number of components, the rotational symmetry of the individual cell states (S=strong, W=weak; explained in the next section), the number of possible states in which a cell may be, and the type of neighborhood (V=von Neumann, M=Moore). For example, the sheathed loop in Fig. 1b is labelled SL86S8V because it spans 86 active cells, has strongly-symmetric cell states with each cell assuming one of 8 possible states, and its transition function is based on the von Neumann neighborhood (17).

In creating a sheathed loop the biologically-implausible requirement of universal computability used in earlier models, in other words, requiring the ability to function as a general purpose computer (18), was abandoned. To avoid certain trivial cases, sheathed loops are required to have a readily-identifiable stored "instruction sequence" that is used by the underlying transition function in two ways: as instructions that are interpreted to direct the construction of a replica, and as uninterpreted data that is copied onto the replica (16). Thus, sheathed loops are truely "information replicating systems" in the sense that this term is used by organic chemists (19).

The original sheathed loop was a modified version of a periodic emitter, a storage element and timing device in Codd's model (15). Codd introduced the concept of a sheathed data path, a series of adjacent cells in state O called the core covered on both sides by a layer of cells in state X called the sheath (Fig. 1a). The sheathed path served as a means for signal propagation. Signals or instruction sequences, represented by cells in other states embedded in the core of a data path, propagate along the path. Codd's periodic emitter was a non-replicating loop similar to that in Fig. 1a except that it contained a sequence of signals that continuously circulated around the loop. Each time the signal sequence passed the origin of the arm (lower right of loop in Fig. 1a) a copy of the signal would propagate out the arm and, among other things, could cause the arm to lengthen or turn. Langton showed that a loop like this could be made self-replicating by storing in it a set of instructions that direct the replication process (16). These instructions cause the arm to extend and turn until a second loop is formed, detaches, and also begins to replicate, so that eventually a growing "colony" of self-replicating loops has appeared. This replicating sheathed loop consists of 86 active cells as pictured in Fig. 1b, and its transition function has 219 rules based on the von Neumann neighborhood (17). Subsequently, two smaller self-replicating sheathed loops containing as few as 12 active cells in one case have been described (Fig. 1c) (20).

Unsheathed Loops

After studying the simplified cellular automata models of self-replicating structures developed in Codd's eight-state framework (15,16,20) we hypothesized that a number of alterations could be made that would result in simpler and smaller self-replicating structures. Such simplification is important for understanding the minimal information processing requirements of self-replication, for relating these formal models to theories of the origins of life, and for identifying configurations so simple that they might actually be synthesized or fabricated. One potential simplifying alteration is removal of the sheath surrounding data paths. It was not obvious in advance that complete removal of the sheath would be possible. The sheath was introduced by Codd and retained in developing sheathed loops because it was believed to be essential for indicating growth direction and for discriminating right from left in a strongly rotation-symmetric space (15, p.40; 20, p. 296). In fact, we have discovered that having a sheath is not essential for these tasks, and its removal leads to smaller self-replicating structures that also have simpler transition functions.

To understand how the sheath (surrounding covering of X's) can be discarded, consider the unsheathed version UL32S8V (shown in Fig. 1e) of the original 86-component sheathed loop (shown in Fig. 1b). The cell states and transition rules of this unsheathed loop obey the same symmetry requirements as those of the sheathed loop, and the signal sequence +−+−+−+−+−+−L−L− directing self-replication is similar (read off of the loop clockwise starting at the lower right corner and omitting the "core" cells in state O). As illustrated in Fig. 2, the instruction sequence circulates counterclockwise around the loop, with a copy passing onto the construction arm. As the elements of the instruction sequence reach the tip of the construction arm, they cause it to extend and turn periodically until a new loop is formed. A "growth cap" of X's at the tip of the construction arm enables directional growth and right-left discrimination at the growth site (seen in Fig. 2b-d). It is this growth cap that makes elimination of the sheath possible. As shown in Fig. 2e, after 150 iterations or units of time the original structure (on the left, its construction arm having moved to the top) has created a duplicate of itself (on the right).

The unsheathed loop UL32S8V not only self-replicates but it also exhibits all of the other behaviors of the sheathed loop: it and its descendents continue to replicate, and when they run out of room for new replicas, they retract their construction arm and erase their coded information. After several generations a single unsheathed loop has formed an expanding "colony" where actively replicating structures are found only around the periphery. Unsheathed loop UL32S8V has the same number of cell states, neighborhood relationship, instruction sequence length, rotational symmetry requirements, and so on, as the original sheathed loop and it replicates in the same amount of time. However, it has only 177 rules compared to 207 for the sheathed loop (21), and is less than 40% of the size of the original sheathed loop (32 active cells versus 86 active cells, respectively). The rules forming the transition function for UL32S8V are given in (22).

Successful removal of the sheath makes it possible to create a whole family of self-replicating unsheathed loops using 8-state cells. Examples of these new self-replicating structures are shown ordered in terms of progressively decreasing size in Fig. 1d-g and are summarized in the first four rows of Table 1. Each of these structures is implemented under exactly the same assumptions about number of cell states available (eight), rotational symmetry of cell states, neighborhood, isotropic and homogeneous cellular space, and so forth, as sheathed loops, in other words, within Codd's framework (15). Given the initial states shown here, it is a straightforward but tedious and time-consuming task to create the transition rules needed for replication of each of these structures using software we developed for this purpose (22). The smallest unsheathed loop in this specific group using 8-state cells, UL06S8V in Fig. 1g, is listed in line 4 of Table 1; it is more than an order of magnitude smaller than the original sheathed loop (SL86S8V; line 9 of Table 1). Consisting of only six components and using the instruction sequence +L, it replicates in 14 units of time (column Replication Time in Table 1). Replication time is defined as the number of iterations it takes for both the replica to appear and for the original structure to revert to its initial state. This very small structure uses a total of 174 rules (Total Rules in Table 1) of which only 83 are needed to produce replication (Replication Rules); the remaining rules are used to detect and handle "collisions" between different growing loops in a colony, and to erase the construction arm and instruction sequence on loops during the formation of a colony. If one counts only those rules which cause a change in state of the cell to which they are applied, this structure uses a total of 91 rules (State Change Rules) of which only 49 are used to produce replication (State Change Replication Rules). This latter measure is taken here to be the preferred measure of the information processing complexity of a transition function because it includes only rules needed for replication and only rules that cause a state change.

The smallest previously described structure that persistently self-replicates (20), designated SL12S6V here, uses 6-state cells, has 12 components (Fig. 1c), and as indicated in Table 1, requires 60 state change replication rules (23). We have created unsheathed loops, designated UL06S6V and UL05S6V, using 6-state cells with half as many components and requiring only 46 or 35 state change replication rules, respectively (last two rows of Table 1). The initial state of UL06S6V is shown in Fig. 1g, and that of UL05S6V is identical except it has one less component in its arm; the complete transition functions are given in (22). To our knowledge, UL05S6V is the smallest and simplest self-replicating structure created under exactly the same assumptions about cell neighborhood, symmetry, and so forth, as sheathed loops.

Varying Rotational Symmetry

Cellular automata models of self-replicating structures have always assumed that the underlying two dimensional space is homogeneous (every cell is identical except for its state) and isotropic (the four directions NESW are indistinguishable). However, there has been disagreement about the desirable rotational symmetry requirements for individual cell states as represented in the transition function. The earliest cellular automata models had transition functions satisfying weak rotational symmetry: some cell states were directionally oriented (2,8). These oriented cell states were such that they permuted among one another consistently under successive 90° rotations of the underlying two-dimensional coordinate system (24). For example, the cell state designated $\uparrow$ in von-Neumann's work is oriented and thus permutes to different cell states $\rightarrow$, $\downarrow$, and $\leftarrow$ under successive 90° rotations; it represents one oriented component that can exist in four different states or orientations. However, Codd's simplified version of von Neumann's self-replicating universal constructor-computer (15) and the simpler sheathed loops (16,20) are all based upon more stringent criteria called strong rotational symmetry (24). With strong rotational symmetry all cell states are viewed as being unoriented or rotationally symmetric. The transition functions for the unsheathed loops shown in Fig. 1d-g also all use this strong rotational symmetry requirement (indicated by S in their labels). Their eight cell states are designated

$$.O\#L - *X+ \tag{1}$$

where the period designates the quiescent state. All of these states are treated as being unoriented or rotationally symmetric by the transition function (25).

The fact that the simplest self-replicating structures developed so far have all been based on strong rotational symmetry raises the question of whether the use of unoriented cell states intrinsically leads to simpler algorithms for self-replication. Such a result would be surprising as the components of self-replicating molecules generally have distinct orientations. To examine this issue we developed a second family of self-replicating unsheathed loops, shown in Fig. 1h-k, whose initial state and instruction sequence are similar to those already described in Fig. 1d-g. However, for the structures in Fig. 1h-k weak symmetry is assumed, and the last four of the eight possible cell states

$$.O\#L\wedge > \vee < \tag{2}$$

are treated as oriented according to the permutation $(.)(O)(\#)(L)(\wedge > \vee <)$. In other words, the cell state $\wedge$ is considered to represent a single component that has an orientation and is thus permuted to $>$, $\vee$ and $<$ by successive 90° rotations of the coordinate system, while the remaining four cell states do not change. For example, in Fig. 1i the states $>$, $\vee$, and $<$ appear on the lower, left and upper loop segments, respectively, to represent the instruction sequence $<<<<<< LL$. While cells in such a model have 8 possible states and are thus comparable in this sense with the above work on sheathed and unsheathed loops (Fig. 1a-g), they also can be viewed as simpler in that they have only 5 distinct possible components. As can be seen in Table 1 (lines 5-8), where the presence of oriented cell states or weak symmetry is indicated by W in the structure labels, relaxing the strong rotational symmetry requirement like this consistently led to transition functions requiring fewer rules than the corresponding strong symmetry version; this is true by any of the measures in Table 1, and is illustrated in Fig. 4. This decrease in complexity occurred in part because the directionality of the oriented cell states intrinsically permits directional growth and right-left discrimination, making even a growth cap unnecessary.

This simplicity and speed of replication made possible by weak rotational symmetry are illustrated in Fig. 3 (top half) where the complete first replication cycle of UL06W8V is shown. Only 31 rules are needed to direct replication of this small structure which makes use of only 5 possible components (26). After several generations the older, inactive structures are surrounded by persistently active, replicating progeny (bottom half of Fig. 3), and this colony formation continues indefinitely. The small but complete set of transition function rules needed for one replication of UL06W8V are listed in the top half of Table 2. Each rule here is of the form CNESW $\rightarrow$ C$'$ where C is the state of the center cell of the neighborhood, the next four characters are the states of the four noncenter neighbors taken clockwise (north, east, south, west), and C$'$ designates the new state of the center cell. Each rule is interpreted assuming weak rotational symmetry as described above.

The results summarized in Table 1 lead to additional observations about unsheathed loops. For systems with either weak or strong symmetry requirements, the number of rules in the transition function required for replication increases as structure size increases but then levels off to a value characteristic of which of the symmetry requirements are in effect (see Fig. 4). Replication time is essentially independent of the type of rotational symmetry used (strong versus weak) but is proportional to the size of the self-replicating loop (see Fig. 5). This proportionality is effectively linear (slope 5.27, y-intercept -19.04 by least squares fit). To assess the effect of neighborhood, we implemented versions of the two arbitrarily-selected unsheathed loops shown in Fig. 1e and 1j using the Moore neighborhood (27). The resultant systems, designated UL32S6M and UL10W8M in Table 1, had the same replication time as identically-structured UL32S8V and UL10W8V but required dramatically more rules in their transition function for replication.

Reduced Rule Sets

As noted earlier, the complete transition function includes a number of rules that are extraneous to the actual self-replication process (such as instruction sequence erasure) and many rules which simply specify that a cell state should not change. The state change rules alone are completely adequate to encode the replication process. For this reason, we believe that the number of state change rules used for one replication is the most meaningful measure of complexity of transition functions supporting self-replication. As shown in the sixth column of Table 1, this measure indicates that, from an information processing perspective, algorithms for self-directed replication can be relatively simple compared to what has been recognized in the past, especially when oriented components are present.

The simplicity of unsheathed loop transition functions when oriented components are used is even more striking if one permits the use of unrestricted placeholder positions in encoding their rules. We implemented a search program that takes as input a set of rules representing a transition function, such as those forming the top part of Table 2, and produces as output a smaller set of reduced rules containing "don't care" or "wildcard" positions (bottom part of Table 2) (22). This program systematically combines the original rules, replacing multiple rules when possible with a single rule containing positions where any cell state is permissible (designated by the underline character _). Introduction of such wildcard positions is done carefully so that the new reduced rules do not contradict any of the original rules, including those that do not change a cell's state. The size of the reduced rule sets that result from applying this program to the complete original set of rules and to only the replication rules of each of the cellular automata models described above is shown in the rightmost two columns of Table 1. For example, with UL06W8V the single new rule L____ → O that means "state L always changes to state O" replaces seven original replication rules, while the single rule >__ . L → L indicating that L follows > around a loop replaces three original replication rules. With UL06W8V this procedure reduces the complete rule set from 101 to 33 rules, and the set of rules needed for one replication from 58 to 20. Thus, by capturing regularities in rules through wildcard positions, it is possible to encode the replication process for unsheathed loop UL06W8V in only 20 rules (Table 2, bottom half). Computer simulations verified that these 20 rules can guide the replication of UL06W8V in exactly the same way as do the original rules. As shown in Table 1, similar reductions occur with other self-replicating structures (verified by additional computer simulations), and we anticipate that additional modifications to rule format might provide even simpler encodings.

The physical and chemical processes underlying self-replication in contemporary biological systems appear to be quite complex. In fact, at present it has not been possible to actually realize any "informational replicating systems" in the laboratory (19), although recent results in experimental chemistry suggest this may someday be possible (28). However, we have shown here that there exist remarkably simple information replicating systems, at least when they are viewed from an information processing perspective (29). Analogous conclusions about unexpectedly simple information processing requirements have been reached regarding other complex physical/chemical processes after cellular automata models of them were developed, such as the appearance of stably rotating spiral forms in the Belousov-Zhabotinskii autocatalytic reaction (30). Further, it seems probable that the self-replicating structures described here are not the simplest possible: variations, refinements, and different structures such as linear sequences might exist that are even simpler. At present the unsheathed loops described here, which are not intended as realistic models of known biochemical processes, have only a vague correspondence to real molecular structures. Their information-carrying loop might be loosely correlated, for example, with a circular oligonucleotide, and the construction arm with a protein that reads the encoded replication algorithm to create the replica. Still, the existence of these systems raises the question of whether contemporary techniques being developed by organic chemists studying autocatalytic systems (19,28) or the many innovative manufacturing techniques currently being developed in the field of nanotechnology (31) could be used to realize self-replicating molecular structures patterned after the information processing occurring in unsheathed loops.

References and Notes

1. L. Penrose and R. Penrose, *Nature*, **179**, 1183 (1957); H. Jacobson, *Amer. Sci.*, **46**, 255 (1958); H. Morowitz, *Amer. Sci.*, **47**, 261 (1959).

2. J. von Neumann, *The Theory of Self-Reproducing Automata* (University of Illinois Press, Urbana, 1966); A. Burks, in *Essays on Cellular Automata*, A. Burks, Ed. (University of Illinois Press, Urbana, 1970), chap. 1.

3. R. Laing, *J. Theor. Biol.*, **54**, 63 (1975).

4. C. Langton, Ed., *Artificial Life* (Addison-Wesley, New York, 1989); C. Langton, C. Taylor, J. Farmer and S. Rasmussen, Eds., *Artificial Life II* (Addison-Wesley, New York, 1992).

5. J. Oró, S. Miller and A. Lazcano. *Ann. Rev. Earth Planet. Sci.*, **18**, 317 (1990); C. Ponnamperuma, Y. Honda, and R. Navarro-González, *J. Brit. Interplanet. Soc.*, **45**, 241 (1992).

6. K. Drexler, in *Artificial Life*, C. Langton, Ed. (Addison-Wesley, New York, 1989), 501.

7. The Game of Life provides a well-known example of this computational framework having only two states: each cell is either in the active state 1 ("alive") or in the quiescent state 0 ("dead"). See E. Berlekamp, J. Conway and R. Guy, *Winning Ways for Your Mathematical Plays*, (Academic Press, New York, 1982), vol. 2, chap. 25.

8. J. Thatcher, in *Essays on Cellular Automata*, A. Burks, Ed., (Univ. of Illinois Press, Urbana, 1970), chap. 5.

9. M. Arbib, *Inform. and Control*, **9**, 177 (1966).

10. P. Vitányi, *Math. Biosci.*, **18**, 23 (1973).

11. J. Myhill, in *Views on General Systems Theory*, M. Mesarović, Ed. (Krieger, Huntington NY, 1964), chap. 7; T. Ostrand, *J. Comp. and System Sci.*, **5**, 623 (1971).

12. E. Moore, *Proc. Fourteenth Symp. Appl. Math*, April 5-8 (American Mathematical Society, 1962), p. 17.

13. R. Rosen, *Bull. Math. Biophys.*, **21**, 387 (1959); E. Banks, *Eleventh Ann. Symp. on Switching and Automata Theory*, Oct. 28-30 (IEEE, 1970), p. 194; G. Herman, *Information Processing Letters*, **2**, 61 (1973); L. Priese, *J. Cybernet*, **6**, 101 (1976).

14. A. Smith, in *Artificial Life II*, C. Langton, C. Taylor, J. Farmer and S. Rasmusen, Eds., (Addison-Wesley, 1992), 709.

15. E. Codd, *Cellular Automata* (Academic Press, 1968).

16. C. Langton, *Physica D*, **10**, 135 (1984); C. Langton, *Physica D*, **22**, 120 (1986).

17. The von Neumann neighborhood of a cell consists of itself and the four contiguous vertical and horizontal cells (designated NESW).

18. More formally, a Universal Turing Machine. See J. Hopcroft and J. Ullman, *Introduction to Automata Theory, Languages and Computation*, (Addison-Wesley, Reading, 1979), chap. 7.

19. L. Orgel, *Nature*, **358**, 203 (1992).

20. J. Byl, *Physica D*, **34**, 295 (1989).

21. This is a slightly smaller number of rules for the sheathed loop SL86S8V than given in (16). Upon re-implementing the sheathed loop we discovered that 12 of the original rules are never used.

22. J. Reggia, H. Chou, S. Armentrout and Y. Peng, CS-TR-2965, available from: Technical Report Librarian, Dept. of Computer Science, University of Maryland, College Park, MD 20742 USA (1992).

23. SL12S6V exactly recreates the sequence given in Fig. 3 of (20). To allow fair comparisons, the number of transition function rules for SL12S6V is measured in the same fashion as that for all of the other self-replicating structures described in this article. Thus, the values listed for SL12S6V in Table 1 differ somewhat from those given in (20) where the number of rules was counted differently. Structures SL12S6V, UL06S6V, and UL05S6V differ from the others in Table 1 in that they do not erase their instruction sequence when replication is no longer possible.

24. We use the formal definition of rotational symmetry in cellular automata given in (15). Let $x = (i, j)$ be the integer coordinates of an arbitrary cell, $v^t(x)$ be the state of cell x at time t, and f be the transition function. Then for the von Neumann neighborhood the transition function f has *weak rotational symmetry* if there exists a permutation p acting on the set of possible cell states such that

1. $p(v_o) = v_o$, where v_o is the distinguished quiescent state; and
2. if $v^{t+1}(x) = f(v^t(x), v^t(x+(0,1)), v^t(x+(1,0)), v^t(x+(0,-1)), v^t(x+(-1,0)))$ then
$p(v^{t+1}(x))=$
$f(p(v^t(x)), p(v^t(x+(-1,0))), p(v^t(x+(0,1))), p(v^t(x+(1,0))), p(v^t(x+(0,-1)))).$

Function f has *strong rotational symmetry* if p is the identity function.

25. Care should be taken not to confuse the rotational symmetry of a cell state as interpreted by the transition function with the rotational symmetry of the character used to represent that state. Here the character L is not rotationally symmetric, for example, but the cell state it represents is treated as such.

26. The permutation among cell states upon rotation might be considered to be a single additional "rule" but this is not included in any of the rule counts in this article.

27. The Moore neighborhood of a cell consists of itself, its four contiguous vertical/horizontal cells (NESW), and its four adjacent diagonal cells (NE,NW,SE,SW).

28. J. Hong, Q. Feng, V. Rotello and J. Rebek, *Science*, **255**, 848 (1992); Q. Feng, T. Park, and J. Rebek, *Science*, **256**, 1179 (1992); G. von Kiedrowski, *Angew. Chem. Int. Ed. Engl.*, **25**, 932 (1986); I. Amato, *Science*, **255**, 800 (1992).

29. All of the new self-replicating structures described in this article are non-trivial, having a readily identifiable instruction sequence that both directs construction of a replica and is also copied onto the replica (16,22). This can be seen, for example, for even the shortest instruction sequences such as $< L$ in Fig. 3.

30. B. Madore and W. Freedman, *Science*, **222**, 437 (1983); M. Gerhardt, H. Schuster and J. Tyson, *Science*, **247**, 1563 (1990).

31. H. Ahmed, *Proc. IEEE*, **79**, 1140 (1991); P. Ball and L. Garwin, *Nature*, **355**, 761 (1992); S. Chu, *Science*, **253**, 861 (1991); Y. Kim and C. Lieber, *Science*, **257**, 375 (1992); C. Schneiker, in *Artificial Life*, C. Langton, Ed., (Addison-Wesley, 1989), 443; see also special issue of *Science*, **254**, Nov. 29, 1991.

32. Supported by NASA Award NAGW-2805 and by a Faculty Research Fellowship Award to Dr. Reggia from the University of Maryland. Dr. Reggia is also with the Institute of Advanced Computer Studies and the Dept. of Neurology.

Figure Captions

Fig. 1. Self-replicating structures in two dimensional cellular automata. Cells in the quiescent state are indicated by blank spaces. (**a**). Non-replicating loop plus arm (lower right) consisting of a core of cells in state O and a sheath of cells in state X. This is essentially a fixed point of the transition function of the sheathed loop, SL86S8V. (**b**). Initial state of the sheathed loop SL86S8V (16); non-numeric characters are used to represent cell states. The instruction sequence + + + + + + L L is embedded in the core of O's of the non-replicating loop shown in (**a**) (reading clockwise around the loop starting at the lower right corner). The full set of cell states used is given by expression (1) in the text. (**c**). Initial state of a smaller self-replicating sheathed structure SL12S6V (20). (**d-g**). Unsheathed self-replicating loops we have discovered that also use strongly rotation-symmetric cell states. Labeled UL48S8V, UL32S8V, UL10S8V, and UL06S8V, respectively. (**h-k**). Analogous to d-g except some cell states are oriented (weak rotational symmetry). Labeled UL48W8V, UL32W8V, UL10W8V, and UL06W8V, respectively. (**l**). A small self-replicating loop UX10W8V where the construction arm extends from the "side" rather than "corner". Each loop shown here is a perfect square which is distorted by a character font that is taller than it is wide.

Fig. 2. Successive states of unsheathed loop UL32S8V pictured in Fig. 1e starting at time t=0. The instruction sequence repeatedly circulates counterclockwise around the loop with a copy periodically passing onto the construction arm. At t=3 (**a**) the sequence of instructions has circulated 3 positions counterclockwise with a copy also entering the construction arm. At t=6 (**b**) the arrival of the first + state at the end of the construction arm produces a growth cap of X's. This growth cap, which is carried forward as the arm subsequently extends to produce the replica, is what makes a sheath unnecessary by enabling directional growth and right-left discrimination even though strong rotational symmetry is assumed. Successive arrival at the growth tip of +'s extends the emerging structure and arrival of L's cause left turns, resulting in eventual formation of a new loop. Intermediate states are shown at t=80 (**c**) and t=115 (**d**). By t=150 (**e**) a duplicate of the initial loop has formed and separated (on the right; compare to Fig. 1e); the original loop (on the left, construction arm having moved to the top) is beginning another cycle of self-directed replication.

Fig. 3. Structure UL06W8V uses only five unique components and its replication can be governed by either of the two relatively small sets of rules in Table 2. (*Top*) Starting at $t = 0$, the initial state shown at the upper left passes through a sequence of steps until at $t = 10$ an identical but

rotated replica has been created. (*Bottom*) After several generations, a colony has formed. Structures around the periphery are still actively replicating; those in the center have retracted their arms and erased the instruction sequence that directs their self-replication. Growth of this colony continues indefinitely (this was verified by computer simulations out to 11 generations for all of the unsheathed loops described in this article).

Fig. 4. Rules required for one replication versus initial size. Regardless of whether one considers all rules used to replicate (filled symbols) or only replication rules involving state changes (open symbols), comparable self-replicating loops using weak rotational symmetry (◇ and ◇) had substantially fewer rules in their transition functions than those using strong rotational symmetry (■ and □). This and the following figure are based on the first eight rows of Table 1.

Fig. 5. Time to replicate versus initial size. Plots for strong (■) and weak (□) rotational symmetry are largely indistinguishable and effectively linear.

```
a.    XXXXXXXX           b.    XXXXXXXX
     XOOOOOOOOX              XO+ OL OLX
     XOXXXXXXOX              X XXXXXX X          c.    XX
     XOX    XOX              X+X    XOX                XLOX
     XOX    XOX              XOX    XOX                XL+X
     XOX    XOX              X X    XOX                 X*
     XOX    XOX              X+X    XOX
     XOXXXXXXOXXXXX          XOXXXXXXOXXXXX
     XOOOOOOOOOOOOOX         X +O +O +OOOOOX
      XXXXXXXXXXXXX           XXXXXXXXXXXXX


d.   -0+-0+-0L-0L
     +          -         e.
     0          0            0+-0L-0L
     -          0            -      -        f.          g.
     +          0            +      0           000         00
     0          0            0      0           0 0         L+00
     -          0            -      0           L++00
     +          0            +      0
     0          0            0      0
     -          0            -+0-+0-+0000
     +          0
     0-+0-+0-+0-+0000


 h.  00<00<LL0000
     v          0         i.
     0          0           00<LL000
     0          0           v      0        j.          k.
     v          0           0      0           000         00
     0          0           0      0           0 0         L>00
     0          0           v      0           L>>00
     v          0           0      0
     0          0           0      0                    l.
     0          0           >00>00>00000                 000
     v          0                                        0 000
     00>00>00>00^0000                                    L>^
```

Figure 1

```
a.                   b.                    c.                 X
    0L-0L-00             0L-00000              -0+-0+-0      X0+-
    -      0             -      +              +      L       X 0
    +      0             L      -              0      -         +
    0      0             0      0              -      0         -
    -      +             -      +              +      L         0
    +      -             +      -              0      -         +
    0      0             0      0   X          -      0         -
    -+0-+0-+0-+0         -+0-+0-+0-+0X         +0-+00000-L0-L0-+0
                                    X

                                       0
                                       +
d.                          e.         -
                                       0
    0L-0L-00  -0+-0+-0          0000+-0+  0+-0L-0L
    -      0  +      +          0      -  -      -
    +      0 X0X     -          -      0  +      0
    0      0  X      0          L      +  0      0
    -      +         L          0      -  -      0
    +      -         -          -      0  +      0
    0      0         0          L      +  0      0
    -+0-+0-+0-+00000-L          0-+0-+0-  -+0-+0-+0000
```

Figure 2

```
                                                                <
OO        O<           vL           LO O         OO^O<        O< vL
L>OO      OL>O         OOL>         >OOL^        L>OOL        OLvOO


              O            O           O            O
 #<           O            O           O            ^
vL LO        LO OO        OO O<       O^ vL        vL LO
OO >O        >O L^        L^ OL       OL OO        OO vO
  >            >#            O           O            O
                             O           O            O
```

```
             ^
             L
            LO OO
            >O OO

         OO OO OO
       ^ OO OO OO  ^
       L           L
      LO OO OO OO LO OO
      >O OO OO OO >O OO

   OO OO OO OO OO OO OO
 ^ OO OO OO OO OO OO OO  ^
 L                       L
LO vL OO OO OO OO OO OO LO OO
>O OO OO OO OO OO OO OO >O L^
   O                       v
   LO OO OO OO OO OO OO OO O
   >  OO OO OO OO OO OO OO

      OO vL OO OO OO OO
      OO OO OO OO OO L^
         O           v
         LO OO OO OO O
         >  OO OO OO

            OO OO
            OO L^
               v
               O
```

Figure 3

Table 1: Replication Time and Number of Rules

Label	Replication Time	Total Rules	Replication Rules	State Change Rules	State Change Replication Rules	Reduced Total Rules	Reduced Replication Rules
UL48S8V	234	177	167	109	104	75	72
UL32S8V	150	177	166	109	104	74	71
UL10S8V	34	163	117	74	54	50	40
UL06S8V	14	174	83	91	49	66	32
UL48W8V	234	142	98	80	52	68	42
UL32W8V	151	134	98	77	52	66	42
UL10W8V	34	114	82	43	35	31	24
UL06W8V	10	101	58	44	31	33	20
SL86S8V	151	207	181	118	101	90	77
UL32S6M	150	-	305	-	129	-	-
UL10W8M	34	-	221	-	56	-	-
UX10W8V	44	173	103	70	36	57	25
SL12S6V	26	145	140	61	60	46	45
UL06S6V	18	115	83	64	46	30	30
UL05S6V	17	65	58	35	35	23	23

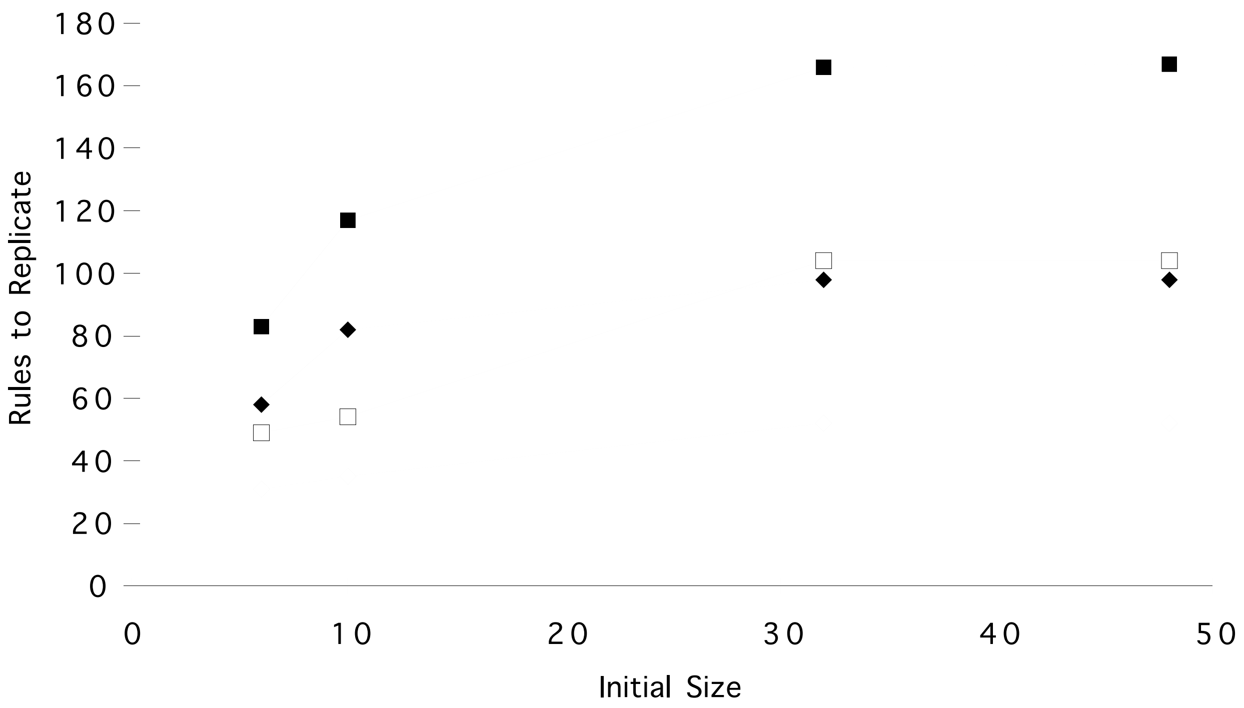
180
160
140
120
100
80
60
40
20
0
Rules to Replicate
0
10
20
30
40
50
Initial Size

Figure 4:

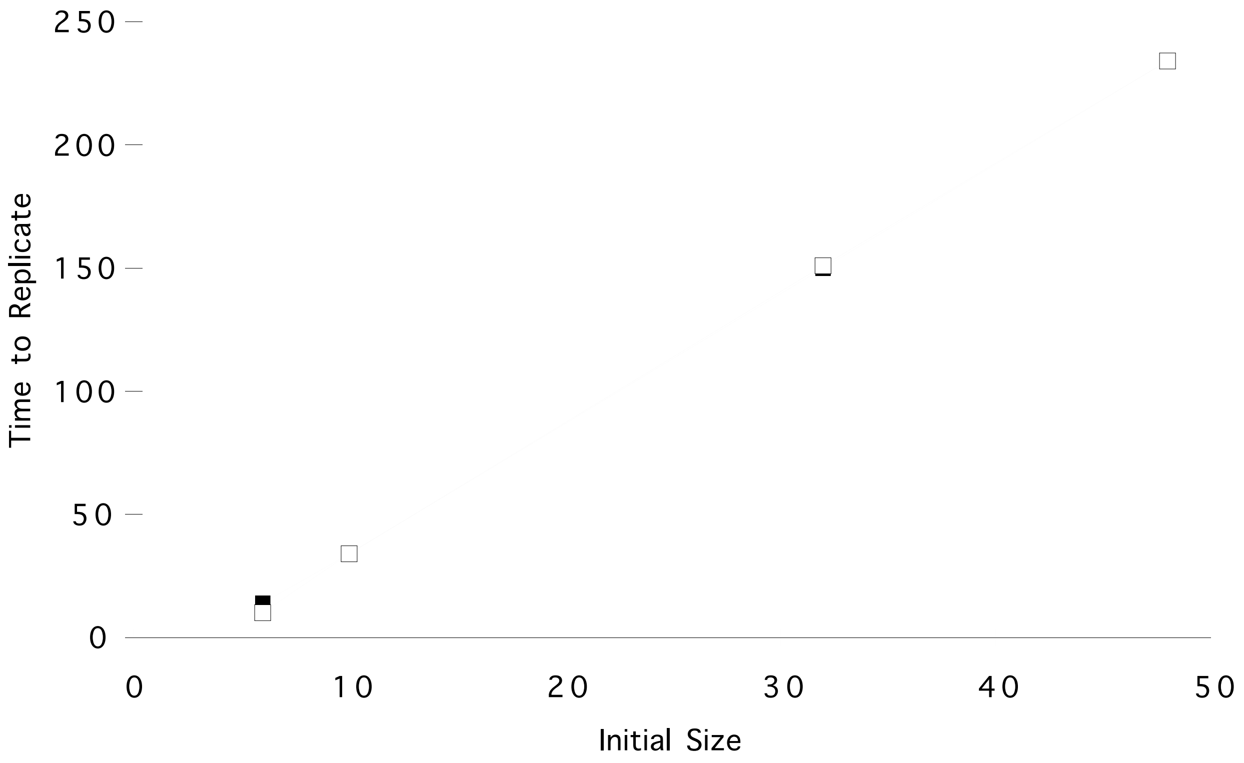
Time to Replicate
250
200
150
100
50
0
0
10
20
30
40
50
Initial Size

Figure 5:

Table 2: Transition Function for UL06W8V

```
CNESW -> C     CNESW -> C     CNESW -> C     CNESW -> C
----------     ----------     ----------     ----------
Replication rules:
..... -> .     >.0.L -> L     0..L. -> 0     .0>.. -> .
...0. -> .     0...> -> >     ..000 -> ^     ..>>0 -> .
..L.. -> .     ...v. -> .     0.^>0 -> <     .<L.L -> .
.>... -> .     L..0v -> 0     0^0.> -> v     L#.0v -> 0
...00 -> .     0LL.0 -> 0     00L.0 -> 0     ...L< -> .
0.0L. -> 0     ...LL -> .     0.<0^ -> v     <...# -> .
L0>.. -> 0     ....> -> ^     ^.000 -> .     ..#v. -> .
>00.L -> L     ...>. -> 0     ...v< -> .     #.<L. -> 0
0..>0 -> <     >...L -> L     L<v.0 -> 0     ...#. -> 0
0.0.> -> >     L.>.0 -> 0     v.0.L -> .     ..L>0 -> .
0...0 -> 0     000.> -> >     0v0.v -> >     ..0.0 -> .
0.<0. -> v     0.L.0 -> 0     .<vv< -> .     00.0L -> 0
<..L0 -> L     L0^.0 -> 0     <.... -> <     ..0L. -> .
L<>.0 -> 0     ^...L -> L     ..<<. -> #     00.0. -> 0
...>< -> .     ...^0 -> <

Reduced set of replication rules:
^.0__ -> .     0._.> -> >     0^___ -> v     L____ -> 0
v.0__ -> .     0v0._ -> >     ...^0 -> <     #____ -> 0
<...# -> .     000._ -> >     0._>_ -> <     >__.L -> L
....> -> ^     0.<0. -> v     ...>. -> 0     <..L_ -> L
..000 -> ^     0.<0^ -> v     ...#. -> 0     ..<<_ -> #
-------------------------------------------------------
Rule neighborhoods are as follows (no rotations listed):
                   N
                 W C E    ->    C
                   S
```