1A) is attributable to aliasing of tidal variability of that sort. There has also been observed a weekly time scale variation of the outflow current in the western Gulf of Cadiz, which may be attributable to the meandering of the current and eddy formation [M. L. Grundlingh, *"Meteor" Forschungsergeb. Reihe A* **21**, 15 (1981)]. To monitor this fast time scale variability, we obtained CTD profiles at several fixed sites west of section A that were revisited up to six times during the course of our week-long experiment. The resulting time series data showed little temporal variation of the outflow thickness or salinity in the eastern Gulf of Cadiz (less than 30% variation in thickness).
19. L. Armi and D. M. Farmer, *Prog. Oceanogr.* **21**, 1 (1988).
20. T. H. Ellison and J. S. Turner, *J. Fluid Mech.* **6**, 423 (1959).
21. J. E. Simpson, *Gravity Currents in the Environment and the Laboratory* (Halsted, New York, 1987).
22. W. Sturges, *J. Mar. Res.* **33**, 117 (1975).
23. L. J. Pratt, *J. Phys. Oceanogr.* **16**, 1970 (1986).
24. J. S. Turner, *Buoyancy Effects in Fluids* (Cambridge Univ. Press, Cambridge, 1973).
25. If the bottom boundary layer has a logarithmic profile, then the shear is $U_z = U_*/kz$, and the energy budget reduces to $\varepsilon = U_*^3/kz$, which can be evaluated from XDP data alone. The bottom stress estimated in this way is smaller by about a factor of 3 than the bottom stress estimated from the XCP data and profile method [similar to the result by R. K. Dewey and W. R. Crawford, *J. Phys. Oceanogr.* **18**, 1167 (1988)]. At this stage, we do not know whether this difference arises from physical, hydrodynamic effects that are not included in the simple model that is used in both techniques (thermal wind shear or form drag) or from the temporal intermittence of ε. Although the absolute values of bottom stress do not agree closely, there is nevertheless a high visual correlation between the two kinds of stress estimates.
26. There are ridges on the sea floor that run southeast to northwest and appear to steer the outflow through this turn. However, because the current turns at almost the rate of an inertial motion, we think that the ridges are depositional features that were produced by the current itself [also suggested in (*8*) for geological reasons], rather than geological features that just happen to lie beneath the outflow.
27. R. W. Griffiths, *Annu. Rev. Fluid Mech.* **18**, 59 (1986).
28. The bottom slope becomes much steeper in the western Gulf of Cadiz and may satisfy this inequality over some portions of the outflow that remain in contact with the bottom.
29. W. Zenk and L. Armi, *Deep-Sea Res.* **37**, 1805 (1990).
30. R. R. Dickson, E. M. Gmitrowicz, A. J. Watson, *Nature* **344**, 848 (1991).
31. P. M. Saunders, *J. Phys. Oceanogr.* **20**, 29 (1990).
32. J. H. Swift, *Deep-Sea Res.* **31**, 1339 (1984).
33. The entire subtropical North Atlantic has a significant net annual evaporation [R. W. Schmitt, P. S. Bogden, C. E. Dorman, *J. Phys. Oceanogr.* **19**, 1208 (1989)] that leads to a surface salinity maximum, although the salinities are considerably lower than those found in the Mediterranean Sea. This near-surface salinity maximum is advected toward the Norwegian Sea inflow by the North Atlantic current and appears to merge with the intermediate-level Mediterranean salinity maximum by a latitude of about 50°N. Thus, the high salinity of the inflow to the Norwegian Sea comes from both subtropical North Atlantic and Mediterranean sources, and, judging by the volume fluxes alone, the former may be the more important.
34. R. S. Pickart, *J. Phys. Oceanogr.* **22**, 1047 (1992).
35. P. C. Smith, *Deep-Sea Res.* **22**, 853 (1975).
36. C. Maillard, *Atlas Hydrologique de l'Atlantique Nord-est* [Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), Brest, France, 1986].
37. This research was supported by grants from the Office of Naval Research. We are grateful to J. H. Dunlap, T. W. Lehman, B. Brown, and J. Lynch for their help with the field experiment and the data analysis.

Simple Systems That Exhibit Self-Directed Replication

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

Biological experience and intuition suggest that self-replication is an inherently complex phenomenon, and early cellular automata models support that conception. More recently, 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 nontrivial self-replicating structures, some substantially smaller and simpler than those previously reported. The 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.

Mathematicians, computer scientists, and others have been studying artificial self-replicating structures or "machines" for over 30 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*, *5*). 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 (*6*). Additionally, it has been suggested that creating and using self-replicating devices will be important for atomic-scale manufacturing technology called nanotechnology (*7*). 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.

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

*To whom correspondence should be addressed.

Although the earliest work on artificial self-replicating structures or machines often used mechanical devices (*1*), subsequent work has been based largely on computational modeling, 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 herein by a period or blank space), whereas the other cells are said to be in an active state (*8*). 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 its immediate neighbor cells. Thus, any process of self-replication captured in a model like this must be an emergent behavior that arises 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 a redesign of specific components (9) or by an increase in cell state complexity (10), demonstrated that variations of the configuration could be used to simulate sexual reproduction (11), generalized von Neumann's basic result to other configurations and higher dimensional cellular spaces (12), established theoretical upper bounds on how rapid a population of self-replicating configurations could grow (13), and examined fundamental definitional issues (14) that continue to generate theoretical interest today (15). 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 can be done in a simpler fashion with cells that have only eight states (16). 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, the sheathed loop, was developed based on eight-state cells (Fig. 1B) (17). The term "sheathed" here indicates that this structure is surrounded by a covering or sheath (represented by X's in Fig. 1, A through C). For clarity, we have here designated names for self-replicating structures 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), the number of possible states in which a cell may be, and the type of neighborhood [V, von Neumann (18); M, Moore (19)]. For example, the sheathed loop in Fig. 1B is labeled SL86S8V because it spans 86 active cells, has strongly symmetric cell states with each cell assuming one of eight possible states, and its transition function is based on the von Neumann neighborhood (18).

In creating a sheathed loop, researchers abandoned the biologically implausible requirement of universal computability used in earlier models—in other words, they no longer required models to function as general purpose computers (20). For the purpose of avoiding 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 (17). Thus, sheathed loops are truely "information-replicating systems" in the sense of this term as used by organic chemists (21).

The original sheathed loop was a modified version of a periodic emitter, a storage element and timing device in Codd's model (16). 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 nonreplicating 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 if one stored in it a set of instructions that directed the replication process (17). 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 appears. 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 (18). Subsequently, two smaller self-replicating sheathed loops have been described, with one containing as few as 12 active cells (Fig. 1C) (22).

A
```
 XXXXXXXX
XOOOOOOOOX
XOXXXXXXOX
XOX    XOX
XOX    XOX
XOX    XOX
XOX    XOX
XOXXXXXXXOXXXXX
XOOOOOOOOOOOOOX
 XXXXXXXXXXXXX
```

B
```
 XXXXXXXX
XO+ OL OLX
X XXXXXX X
X+X    XOX
XOX    XOX
X X    XOX
X+X    XOX
XOXXXXXXXOXXXXX
X +O +O +OOOOOX
 XXXXXXXXXXXXX
```

C
```
 XX
XLOX
XL+X
 X*
```

D
```
-O+-O+-OL-OL
+          -
O          O
-          O
+          O
O          O
-          O
+          O
O          O
-          O
+          O
O-+O-+O-+O-+OOOO
```

E
```
O+-OL-OL
-      -
+      O
O      O
-      O
+      O
O      O
-+O-+O-+OOOO
```

F
```
OOO
O O
L++OO
```

G
```
OO
L+OO
```

H
```
OO<OO<LLOOOO
v          O
O          O
O          O
v          O
O          O
O          O
v          O
O          O
O          O
v          O
OO>OO>OO>OO^OOOO
```

I
```
OO<LLOOO
v      O
O      O
O      O
v      O
O      O
O      O
>OO>OO>OOOOO
```

J
```
OOO
O O
L>>OO
```

K
```
OO
L>OO
```

L
```
OOO
O OOO
L>^
```

Fig. 1. Self-replicating structures in two-dimensional cellular automata. Cells in the quiescent state are indicated by blank spaces. (**A**) Nonreplicating 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 (17); nonnumeric characters are used to represent cell states. The instruction sequence ++++++LL is embedded in the core of O's of the nonreplicating 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 (22). (**D** through **G**) Unsheathed self-replicating loops we have discovered that also use strongly rotation-symmetric cell states. Shown are UL48S8V, UL32S8V, UL10S8V, and UL06S8V, respectively. (**H** through **K**) Loops analogous to those in (D) through (G) except that some cell states are oriented (weak rotational symmetry). Shown are UL48W8V, UL32W8V, UL10W8V, and UL06W8V, respectively. (**L**) A small self-replicating loop UX10W8V where the construction arm extends from the "side" rather than from the "corner." Each loop shown here is a perfect square that is distorted by a character font that is taller than it is wide.

Unsheathed Loops

After studying the simplified cellular automata models of self-replicating structures developed in Codd's eight-state framework (16, 17, 22), 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 the 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 (*16*, p. 40; *22*). 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 (the surrounding covering represented herein by X's) can be discarded, consider the unsheathed version UL32S8V (Fig. 1E) of the original 86-component sheathed loop (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− that directs self-replication is similar (read the loop clockwise, starting at the lower right corner and omitting the "core" cells in state O). Roughly, each pair +− indicates that the construction arm should extend one cell, and the sequence L−L− indicates a left turn. 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 (Fig. 2, B through 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 (t) 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 also exhibits all of the other behaviors of the sheathed loop: it and its descendants 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 forms 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, and rotational symmetry requirements, among other characteristics, as the original sheathed loop and replicates in the same amount of time. However, it has only 177 rules compared to 207 for the sheathed loop (*23*) and is less than 40% of the size of the original sheathed loop (32 active cells versus 86 active cells, respectively); we have described the rules that form the transition function for UL32S8V (*24*).

Successful removal of the sheath makes possible the creation of a whole family of self-replicating unsheathed loops with the use of eight-state cells. Examples of these new self-replicating structures are shown ordered in terms of progressively decreasing size in Fig. 1, D through 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, these structures fall within Codd's framework (*16*). 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 with software we developed for this purpose (*24*). The smallest unsheathed loop in this specific group with eight-state cells, UL06S8V in Fig. 1G, is listed in Table 1; it is more than an order of magnitude smaller than the original sheathed loop SL86S8V (Table 1). Consisting of only six components and using the instruction sequence +L, UL06S8V replicates in 14 units of time (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 (Table 1) of which only

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 three 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. The successive arrival at the growth tip of +'s extends the emerging structure, and the arrival of L's cause left turns, which results in the 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.

A

```
0L-0L-00
-      0
+      0
0      0
-      +
+      -
0      0
-+0-+0-+0-+0
```

B

```
0L-00000
-      +
L      -
0      0
-      +
+      -
0      0    X
-+0-+0-+0-+0X
            X
```

C

```
               X
-0+-0+-0      X0+-
+      L       X 0
0      -         +
-      0         -
+      L         0
0      -         +
-      0         -
+0-+00000-L0-L0-+0
```

D

```
0L-0L-00  -0+-0+-0
-      0  +      +
+      0 X0X     -
0      0  X      0
-      +         L
+      -         -
0      0         0
-+0-+0-+0-+00000-L
```

E

```
       0
       +
       -
       0
0000+-0+  0+-0L-0L
0      -  -      -
-      0  +      0
L      +  0      0
0      -  -      0
-      0  +      0
L      +  0      0
0-+0-+0-  -+0-+0-+0000
```

Table 1. Replication time (iterations) and number of rules for sheathed and unsheathed loops. Dashes represent measurements that were not made.

Loop	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

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 that 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 (*22*), designated SL12S6V here, uses six-state cells, has 12 components (Fig. 1C), and, as indicated in Table 1, requires 60 state-change replication rules (*25*). We have created unsheathed loops, designated UL06S6V and UL05S6V, using six-state cells with half as many components and requiring only 46 or 35 state-change replication rules, respectively (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 as described (*24*). 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.

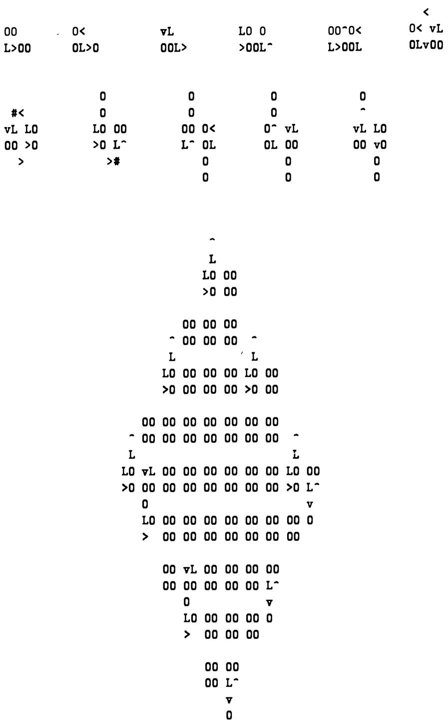

Fig. 3. Structure UL06W8V uses only five unique components, and its replication can be governed by either of the two relatively small sets of rules in Table 2. (**Top**) Starting at $t = 0$, the initial state shown at the upper left passes through a sequence of steps until at $t = 10$ an identical but 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. The growth of this colony continues indefinitely (this was verified by computer simulations out to 11 generations for all of the unsheathed loops described herein).

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 north, east, south, and west (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 that satisfied weak rotational symmetry: some cell states were directionally oriented (*2*, *9*). These oriented cell states were such that they permuted among one another consistently under successive 90° rotations of the underlying two-dimensional coordinate system (*26*). For example, the cell state designated ↑ in von Neumann's work is oriented and thus permutes to different cell states →, ↓, and ← 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 (*16*) and the simpler sheathed loops (*17*, *22*) are all based on more stringent criteria called strong rotational symmetry (*26*). With strong rotational symmetry, all cell states are viewed as being unoriented or rotationally symmetric. The transition functions for unsheathed loops (Fig. 1, D through G) also all use this strong rotational symmetry requirement (indicated by S). Their eight cell states are designated

$$.O\#L - *X+ \qquad (1)$$

where the period designates the quiescent state and # and * represent intermediate states of cells during the growth process (for example, indicating when two loops separate). All of these states are treated as being unoriented or rotationally symmetric by the transition function (*27*).

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. 1, H through K, whose initial state and instruction sequence are similar to those already described (Fig. 1, D through G). However, for the structures in Fig. 1, H

Table 2. Transition function for UL06W8V. Rule neighborhoods are as in expression 3 (no rotations listed).

CNESW → C′	CNESW → C′	CNESW → C′	CNESW → C′
	Replication rules		
. . . . . → .	> . O . L → L	O . . L . → O	. O > . . → .
. . . O . → .	O . . . > → >	. . O O O → ∧	. . > > O → .
. . L . . → .	. . . V . → .	O . ∧ > O → <	. < L . L → .
. > . . . → .	L . . O ∨ → O	O ∧ O . > → ∨	L # . O ∨ → O
. . . O O → .	O L L . O → O	O O L . O → O	. . . L < → .
O . O L . → O	. . . L L → .	O . < O ∧ → ∨	< . . . # → .
L O > . . → O	. . . . > → ∧	∧ . O O O → .	. . # ∨ . → .
> O O . L → L	. . . > . → O	. . . ∨ < → .	# . < L . → O
O . . > O → <	> . . . L → L	L < ∨ . O → O	. . . # . → O
O . O . > → >	L . > . O → O	∨ . O . L → .	. . L > O → .
O . . . O → O	O O O . > → >	O ∨ O . ∨ → >	. . O . O → .
O . < O . → ∨	O . L . O → O	. < ∨ ∨ < → .	O O . O L → O
< . . L O → L	L O ∧ . O → O	< . . . . → <	. . O L . → .
L < > . O → O	∧ . . . L → L	. . < < . → #	O O . O . → O
. . . > < → .	. . . ∧ O → <		
	Reduced set of replication rules		
∧ . O - - → .	O . - . > → >	O ∧ - - - → ∨	L - - - - → O
∨ . O - - → .	O ∨ O . - → >	. . . ∧ O → <	# - - - - → O
< . . . # → .	O O O . - → >	O . - > - → <	> - - . L → L
. . . . > → ∧	O . < O . → ∨	. . . > . → O	< . . L - → L
. . O O O → ∧	O . < O ∧ → ∨	. . . # . → O	. . < < - → #

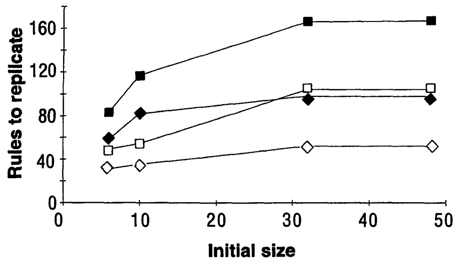

Fig. 4. Rules required for one replication versus initial size of the loop. Regardless of whether one considers all rules used to replicate (filled symbols) or only replication rules involving state changes (open symbols), comparable self-replicating loops using weak rotational symmetry (◆ and ◇) had substantially fewer rules in their transition functions than those using strong rotational symmetry (■ and □). This and Fig. 5 are based on the first eight rows of Table 1.

through K, weak symmetry is assumed, and the last four of the eight possible cell states

$$.O\#L\wedge > \vee < \qquad (2)$$

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, although 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. Although cells in such a model have eight possible states and are thus comparable in this sense with the work above on sheathed and unsheathed loops (Fig. 1, A through G), they also can be viewed as simpler in that they have only five distinct possible components. As can be seen in Table 1, relaxing the strong rotational symmetry requirement like this in the presence of oriented cell states or weak symmetry (as indicated by W in the structure labels) consistently led to transition functions that required fewer rules than the corresponding strong symmetry version; this is true by any of the measures in Table 1. 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), 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 five possible components (28). After several generations, the older, inactive structures are surrounded by persistently active, replicating progeny (Fig. 3, bottom), 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 part of Table 2. Each rule in Table 2 is of the form

$$\begin{array}{ccc} & N & \\ W & C & E \\ & S & \end{array} \longrightarrow C' \qquad (3)$$

(expressed as CNESW→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 those symmetry requirements that are in effect (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 (Fig. 5). This proportionality is effectively linear (slope is 5.27 and y intercept is −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. 1, E and J, using the Moore neighborhood (19). The resultant systems, designated UL32S6M and UL10W8M in Table 1, had the same replication time as identically structured UL32S8V and UL10W8V but required many 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 that 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 the complexity of the transition functions that support self-replication. As shown in Table 1 (state-change replication rules), 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.

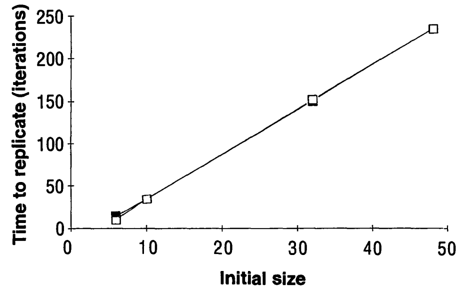

Fig. 5. Time to replicate versus initial size of the loop. Plots for strong (■) and weak (□) rotational symmetry are largely indistinguishable and effectively linear.

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 the transition function rules. We implemented a search program that takes as input a set of rules that represent a transition function, such as those in Table 2 (top), and produces as output a smaller set of reduced rules containing "don't care" or "wild card" positions (Table 2, bottom) (24). This program systematically combines the original rules, replacing multiple rules when possible with a single rule that contains positions where any cell state is permissible (designated by the underline). The introduction of such wild card 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 Table 1 (reduced total and reduced replication rules). For example, with UL06W8V the single new rule L____ → O that means "state L always changes to state O" replaces seven original replication rules, whereas the single rule >__L → L that indicates 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 wild card positions, it is possible to encode the replication process for unsheathed loop UL06W8V in only 20 rules (Table 2, bottom). 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 that underlie 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 (*21*), although recent results in experimental chemistry suggest this may someday be possible (*29*). However, we have shown here that there exist remarkably simple information-replicating systems, at least when they are viewed from an information-processing perspective (*30*). 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-Zhabotinsky autocatalytic reaction (*31*). 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 their 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 (*21, 29*) or the many innovative manufacturing techniques currently being developed in the field of nanotechnology (*32*) 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, *Am. Sci.* **46**, 255 (1958); H. Morowitz, *ibid.* **47**, 261 (1959).
2. J. von Neumann, *The Theory of Self-Reproducing Automata* (Univ. of Illinois Press, Urbana, 1966); A. Burks, in *Essays on Cellular Automata*, A. Burks, Ed. (Univ. 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).
5. C. Langton, C. Taylor, J. Farmer, S. Rasmussen, Eds., *Artificial Life II* (Addison-Wesley, New York, 1992).
6. J. Oró, S. Miller, A. Lazcano, *Annu. Rev. Earth Planet. Sci.* **18**, 317 (1990); C. Ponnamperuma, Y. Honda, R. Navarro-González, *J. Br. Interplanet. Soc.* **45**, 241 (1992).
7. K. Drexler, in (*4*), pp. 501–519.
8. 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") [E. Berlekamp, J. Conway, R. Guy, *Winning Ways for Your Mathematical Plays* (Academic Press, New York, 1982), vol. 2, chap. 25].
9. J. Thatcher, in *Essays on Cellular Automata*, A. Burks, Ed. (Univ. of Illinois Press, Urbana, 1970), chap. 5.
10. M. Arbib, *Inf. Control* **9**, 177 (1966).
11. P. Vitányi, *Math. Biosci.* **18**, 23 (1973).
12. J. Myhill, in *Views on General Systems Theory*, M. Mesarović, Ed. (Krieger, Huntington, NY, 1964), chap. 7; T. Ostrand, *J. Comput. Syst. Sci.* **5**, 623 (1971).
13. E. Moore, in *Proceedings of the Fourteenth Symposium on Applied Math*, R. Belman, Ed. (American Mathematical Society, Providence, RI, 1962), pp. 17–33.
14. R. Rosen, *Bull. Math. Biophys.* **21**, 387 (1959); E. Banks, in *Eleventh Annual Symposium on Switching and Automata Theory* (IEEE, New York, 1970), pp. 194–215; G. Herman, *Inf. Process. Lett.* **2**, 61 (1973); L. Priese, *J. Cybern.* **6**, 101 (1976).
15. A. Smith, in *Artificial Life II*, C. Langton, C. Taylor, J. Farmer, S. Rasmussen, Eds. (Addison-Wesley, New York, 1992), pp. 709–725.
16. E. Codd, *Cellular Automata* (Academic Press, New York, 1968).
17. C. Langton, *Physica D* **10**, 135 (1984); *ibid.* **22**, 120 (1986).
18. The von Neumann neighborhood of a cell consists of itself and the four contiguous vertical and horizontal cells (designated NESW for north, east, south, and west).
19. The Moore neighborhood of a cell consists of itself, its four contiguous vertical and horizontal cells (NESW), and its four adjacent diagonal cells (NE, NW, SE, and SW).
20. More formally, a Universal Turing Machine [J. Hopcroft and J. Ullman, *Introduction to Automata Theory, Languages, and Computation* (Addison-Wesley, Reading, 1979), chap. 7].
21. L. Orgel, *Nature* **358**, 203 (1992).
22. J. Byl, *Physica D* **34**, 295 (1989).
23. This is a slightly smaller number of rules for the sheathed loop SL86S8V than given in (*17*). Upon reimplementing the sheathed loop, we discovered that 12 of the original rules are never used.
24. J. Reggia, H. Chou, S. Armentrout, Y. Peng, technical report CS-TR-2965 (October 1992), available from Technical Report Librarian, Department of Computer Science, University of Maryland, College Park, MD 20742.
25. SL12S6V exactly recreates the sequence given in figure 3 of (*22*). To allow fair comparisons, we measured the number of transition function rules for SL12S6V in the same fashion as that for all of the other self-replicating structures described herein. Thus, the values listed for SL12S6V in Table 1 differ somewhat from those given in (*22*), 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.
26. We use the formal definition of rotational symmetry in cellular automata given in (*16*). Let $x = (i, j)$ be the integer coordinates of an arbitrary cell, $v^t(x)$ the state of cell x at time t, and f 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
$$p(v_o) = v_o$$
where v_o is the distinguished quiescent state; and if
$$v^{t+1}(x) = f\Big(v^t(x), v^t[x + (0, 1)], v^t[x + (1, 0)], v^t[x + (0, -1)], v^t[x + (-1, 0)]\Big)$$
then
$$p[v^{t+1}(x)] = f\Big(p[v^t(x)], p\{v^t[x + (-1, 0)]\}, p\{v^t[x + (0, 1)]\}, p\{v^t[x + (1, 0)]\}, p\{v^t[x + (0, -1)]\}\Big)$$
Function f has strong rotational symmetry if p is the identity function.
27. 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.
28. 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 herein.
29. J.-I. Hong, Q. Feng, V. Rotello, J. Rebek, Jr., *Science* **255**, 848 (1992); Q. Feng, T. K. Park, J. Rebek, Jr., *ibid.* **256**, 1179 (1992); G. von Kiedrowski, *Angew. Chem. Int. Ed. Engl.* **25**, 932 (1986); I. Amato, *Science* **255**, 800 (1992).
30. All of the new self-replicating structures described here are nontrivial, having a readily identifiable instruction sequence that both directs construction of a replica and is also copied onto the replica (*17, 24*). This can be seen, for example, for even the shortest instruction sequences such as < L in Fig. 3.
31. B. F. Madore and W. L. Freedman, *Science* **222**, 615 (1983); M. Gerhardt, H. Schuster, J. J. Tyson, *ibid.* **247**, 1563 (1990).
32. 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. M. Lieber, *ibid.* **257**, 375 (1992); C. Schneiker, in *Artificial Life*, C. Langton, Ed. (Addison-Wesley, New York, 1989), pp. 443–500; special issue on Engineering a Small World, *Science* **254**, 1300–1355 (29 November 1991).

33. Supported by NASA award NAGW-2805 and by a Faculty Research Fellowship Award from the University of Maryland (J.A.R.).