Journal of Statistical Physics, Vol. 43, Nos. 1/2, 1986

Global Properties of Cellular Automata

Erica Jen[1,*]

Received August 7, 1985

Cellular automata are discrete mathematical systems that generate diverse, often complicated, behavior using simple deterministic rules. Analysis of the local structure of these rules makes possible a description of the global properties of the associated automata. A class of cellular automata that generate infinitely many aperiodic temporal sequences is defined, as is the set of rules for which inverses exist. Necessary and sufficient conditions are derived characterizing the classes of "nearest-neighbor" rules for which arbitrary finite initial conditions (i) evolve to a homogeneous state; (ii) generate at least one constant temporal sequence.

KEY WORDS: Cellular automata; discrete dynamical systems; local interactions; deterministic structures.

1. INTRODUCTION

Cellular automata are a class of simple mathematical systems that generate diverse, often complicated, behavior. First introduced by von Neumann and Ulam[13] as potential tools for studying biological self-reproduction, cellular automata have been reintroduced and used as mathematical models in a wide variety of contexts.[3] Typically, a cellular automaton consists of a lattice of sites whose values evolve deterministically according to local interaction rules. The site values are restricted to a finite set of integers, and the rules specify the value of a site as a function of the values of neighboring sites at the previous time step. Specifically, consider the class of automata defined on a one-dimensional (possibly infinite) set of

[1] Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545.

* Permanent address: Mathematics Department, University of Southern California, Los Angeles 90089.

0022-4715/86/0400-0219$05.00/0

sites x_i, each of which assumes any of the values $V = \{0,..., k-1\}$. Then the general form for a rule defining a particular automaton is given by

$$x_i^{t+1} = f(x_{i-r}^t,..., x_i^t,..., x_{i+r}^t) \qquad f\colon\ V^{2r+1} \to V \tag{1.1}$$

where $r \geqslant 0$ is a constant specifying the size of the neighborhood for each site, and each site x_i is assigned an initial value x_i^0.

Cellular automata thus represent a rather general class of discrete dynamical systems. Certain of these automata are clearly equivalent to other standard mathematical constructs, including shift-commuting maps and finite-difference schemes for solving partial differential equations. Other cellular automata, however, such as

$$x_i^{t+1} = x_{i-1}^t \text{ XOR } \max(x_i^t, x_{i+1}^t)$$

where XOR denotes addition modulo two, cannot easily be identified as discretizations of continuous systems, and may be regarded instead as systems that generate novel and potentially interesting mathematical behavior.

As a class of dynamical systems, cellular automata in fact exhibit a remarkable diversity of behavior. Even for fixed k (set of site values) and fixed r (size of neighborhood), the spatial and temporal sequences generated vary in nature from regular to seemingly random. In Refs. 15–19, Wolfram has provided an extensive catalog of the behavior associated with different choices of the rule combined with different initial conditions. In particular, on the basis of systematic computer simulation of a large number of automata, he conjectures that all automata belong to at least one of four classes, qualitatively characterized as follows:

"Class 1 evolution leads to a homogeneous state in which, for example, all sites have values 0;

Class 2 evolution leads to a set of stable or periodic structures that are separated and simple;

Class 3 evolution leads to a chaotic pattern;

Class 4 evolution leads to complex structures, often long-lived."

The above description of the classification scheme appears in Refs. 16 and 17. Figure 1 displays examples of the four classes. Although statistical quantities (such as dimension and entropy) have been used[11,15] to distinguish among the four classes, the classification scheme is essentially phenomenological in nature.

The diversity of behavior described above has motivated many attempts to model "complex" phenomena using cellular automata. In par-

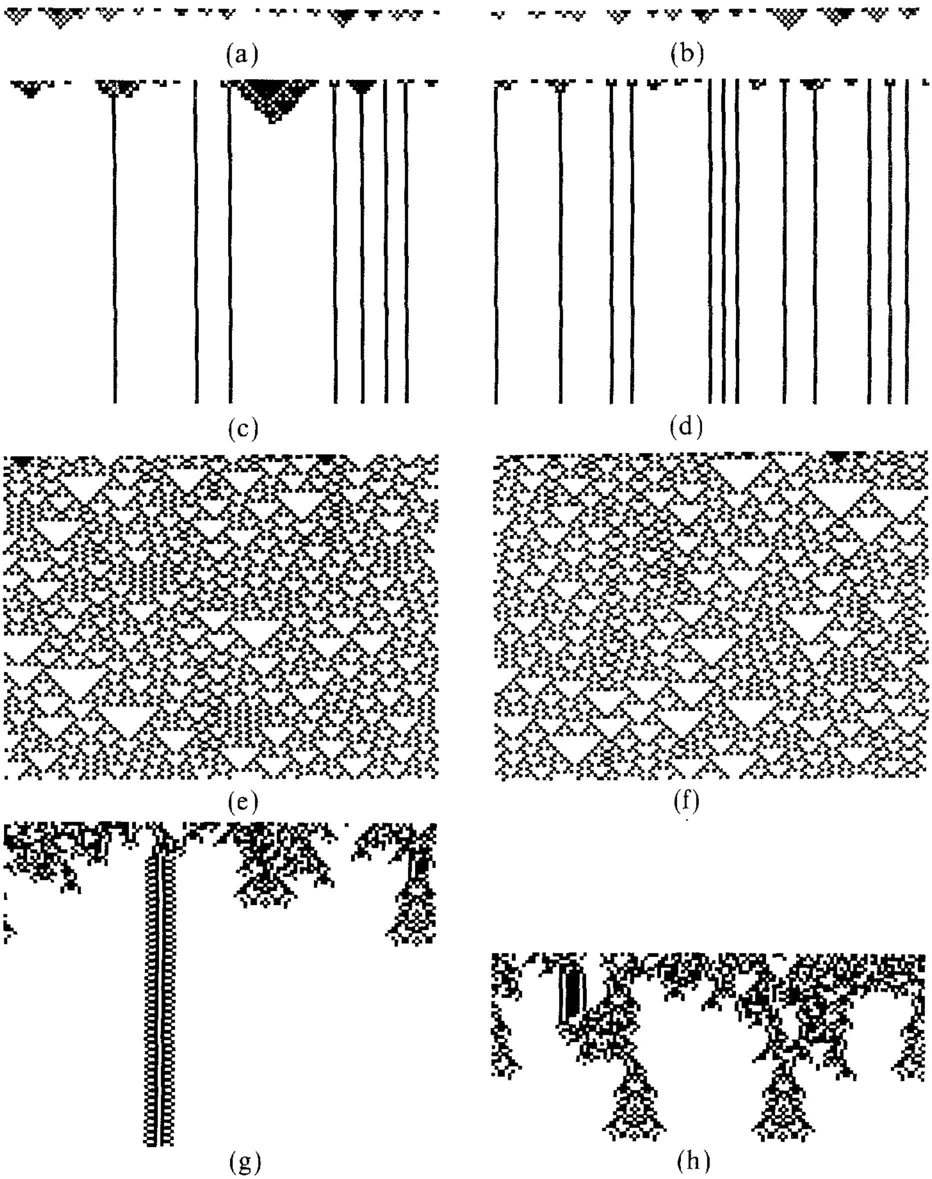


Fig. 1. Examples of the four types of automata behavior described by Wolfram's classification scheme.[15-19] Automata depicted are defined using rules with k (number of possible site values) = 3 and r (size of neighborhood) = 1. White squares represent value 0, grey squares value 1, and black squares value 2. The top row of each automaton provides the initial condition, and each subsequent row represents the state of the automaton at the next time step. Automata (a) and (b) are classified by Wolfram as Class 1; (c) and (d) as Class 2; (e) and (f) as Class 3; and (g) and (h) as Class 4.

ticular, cellular automata are viewed as prototypical mathematical models for systems consisting of a large number of simple, identical, and locally connected components. Examples of such systems include turbulent flow resulting from collisions of fluid molecules,[3] dendritic growth of crystals resulting from aggregation of atoms,[10] and patterns of electrical activity in simple neural networks resulting from neuronal stimulation.[1] Such problems are conventionally studied using continuous models based on partial differential equations. Cellular automata complement the continuous approach by providing alternative simulation tools characterized by discreteness, local interaction, simple construction, diverse range of behavior, and an inherently parallel form of evolution.

Despite the long-standing and brood-based interest in cellular automata, relatively few rigorous results describing automata behavior have been obtained. The properties of "additive" cellular automata, i.e., automata whose site values depend on the sum of neighboring site values at the previous time step, have been studied by Martin et al.[8] and Lind.[7] The equivalence of shift-commuting continuous maps to a subclass of cellular automata has been established by Hedlund.[4] Milnor[9,10] has considered the surjectivity of automata rules. Applications of concepts from information theory to cellular automata have been discussed by Waterman.[14] Other aspects of cellular automata are considered in Ref. 3. The general characterization of automata behavior remains for the most part, however, an open problem.

In this paper, an approach will be presented to the mathematical characterization of "elementary" cellular automata; i.e., automata whose sites assume either of the values $\{0, 1\}$, and whose rules depend only on nearest-neighbor interactions. In essence, the approach has two components, both of which are applicable to automata whose sites can assume more than two values, and whose rules involve more than nearest-neighbor interactions in one or more dimensions. The first focuses on the underlying structures of determinism implied, but often obscured, by the explicit formulation of the rule. Specifically, it is explicit in the form of a rule that the value x_i^{t+1} is defined as a function of the values x_{i-1}^t, x_i^t, x_{i+1}^t; the approach of this paper is to study the classification of rules for which the definition implies, in addition, that the values x_i^t and x_i^{t+1}, for example, determine x_{i+1}^t. In other words, the explicit determinism of cellular automata, which can be viewed as operating along a time-increasing path, forces in many cases an implicit determinism along some other path as well.

The second component of this paper's approach is the study of the tuples (w, x, y) and (x, y, z) into which a given pair $(x, y) \in \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ may "shift" by the appending of a new component

$w, z \in \{0, 1\}$ on the left or right. The set of possible 3-tuples is thereby classified into subsets, the elements of which share a particular (x, y). The deterministic structures discussed in the preceding paragraph will be shown to be direct consequences of the one-to-one, or many-to-one, nature of the rule restricted to these subsets. Furthermore, if (w, x, y) and (x, y, z) are "overlapping" tuples sharing a pair (x, y) and $\{wxyz\}$ appears as a spatial string of values at time t, then the values assigned by the rule to (w, x, y) and (x, y, z) determine part of the string at the next step. Shift transformations can thus be used to characterize the spatial and temporal sequences generated by automata rules.

The "deterministic structure" together with the "shift transformation" approach outlined above can be used to obtain rigorous results describing properties of cellular automata. In particular, the role of the one-to-one, versus many-to-one, nature of automata rules will be shown to be crucial in determining global behavior. Some of the results in this paper are motivated by Wolfram's phenomenological classification scheme, and serve both to support his conjectures and to clarify the dependence of automata behavior on the choice of initial conditions. For example, Section 5 presents necessary and sufficient conditions defining the class of elementary rules for which arbitrary finite initial conditions evolve to a homogeneous state. Section 5 also contains necessary and sufficient conditions characterizing the class of elementary rules for which constant temporal sequences are generated from arbitrary finite initial conditions. In addition to the results mentioned above, the analysis also makes possible the description of global properties not explicitly suggested by Wolfram's classification scheme. In Section 3, for instance, a class of rules is defined for which arbitrary finite initial conditions produce infinitely many temporal sequences that, although not necessarily "chaotic," are not periodic of any period. Section 4 provides conditions for the existence of "inverses" to elementary rules.

As is indicated by the above outline, the emphasis of this paper is on understanding the underlying structure of elementary cellular automata rules, and on defining the classes of rules possessing certain global properties. Taken as a whole, the results are suggestive both of the richness of cellular automata behavior, and of their intrinsic mathematical interest.

2. DETERMINISTIC STRUCTURES

"Elementary" cellular automata are defined[15,16] by rules of the form

$$x_i^{t+1} = f(x_{i-1}^t, x_i^t, x_{i+1}^t) \; f\colon \{0, 1\}^3 \to \{0, 1\} \qquad (2.1)$$

i.e., the sites can assume either of the values $\{0, 1\}$, and only nearest-neighbor interactions are considered. Hence a rule is equivalently defined by specifying the value assigned to each of the 2^3 possible 3-tuple configurations of site values; i.e., by specifying the a_j, $j = 0,..., 7$ such that

$$\begin{array}{cccccccc} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \end{array} \tag{2.2}$$

Since each $a_j \in \{0, 1\}$, there is a total of $2^{2^3} = 256$ possible rules.

Wolfram[15,16] has defined a labeling scheme according to which a rule is assigned a value

$$\text{rule number} = R = \sum_{j=0}^{7} a_j \cdot 2^j \tag{2.3}$$

where a_j is the value assigned to the 3-tuple corresponding to the number j in binary representation. For example, the rule defined by

$$x_i^{t+1} = (x_{i-1}^t + x_{i+1}^t) \bmod 2$$

can be rewritten in the form of (2.2) as

$$\begin{array}{cccccccc} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}$$

and then assigned a rule number

$$R = 0 \cdot 2^0 + 1 \cdot 2^1 + 1 \cdot 2^2 + 0 \cdot 2^3 + 1 \cdot 2^4 + 0 \cdot 2^5 + 0 \cdot 2^6 + 1 \cdot 2^7 = 150$$

Thus the rule numbers range from 0 to 255, with each rule uniquely labeled.

The explicit determinism of rules defined by (2.1) results from constraints on the value x_i^{t+1} assigned to the 3-tuple

$$\begin{array}{c} x_{i-1}^t x_i^t x_{i+1}^t \\ \downarrow \\ x_i^{t+1} \end{array}$$

(In the remainder of the paper, the above will be written as $x_{i-1}^t x_i^t x_{i+1}^t \to x_i^{t+1}$). It is possible, however, that a specific rule is defined in such a way as to imply an additional "deterministic structure." For example, it may be true that for all values of x_i^t, x_{i+1}^t

$$0 x_i^t x_{i+1}^t \to a \qquad \text{and} \qquad 1 x_i^t x_{i+1}^t \to \bar{a} \tag{2.4}$$

with $a \neq \bar{a}$. If relation (2.4) holds, then specifying the values of x_i^t, x_{i+1}^t, and x_i^{t+1} suffices to determine x_{i-1}^t. Rules for which (2.4) holds can be said to exhibit the deterministic structure

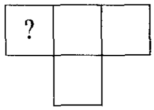
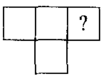
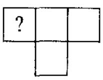
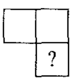
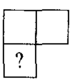
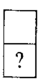
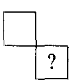
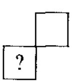
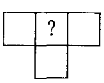
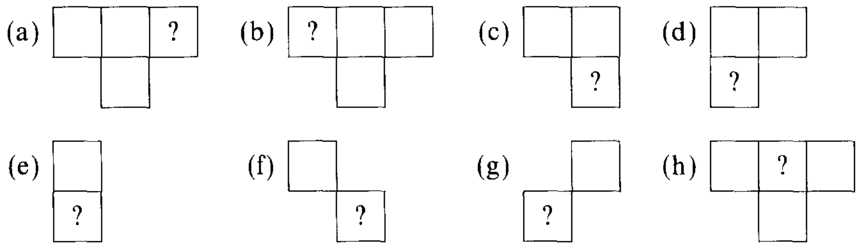


Table I

	Deterministic structure		Constraints on coefficients	Rule numbers
(a)		$x_{i-1}^t, x_i^t, x_i^{t+1} \Rightarrow x_{i+1}^t$	$a_j \neq a_{j+1}$ $j=0, 2, 4, 6$	85, 86, 89, 90, 101, 102, 105, 106, 149, 150, 153, 154, 165, 166, 169, 170
(b)		$x_i^t, x_{i+1}^t, x_i^{t+1} \Rightarrow x_{i-1}^t$	$a_j \neq a_{j+4}$ $j=0, 1, 2, 3$	15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, 180, 195, 210, 225, 240
(c)		$x_{i-1}^t, x_i^t \Rightarrow x_i^{t+1}$	$a_j = a_{j+1}$ $j=0, 2, 4, 6$	0, 3, 12, 15, 48, 51, 60, 63, 192, 195, 204, 207, 240, 243, 252, 255
(d)		$x_i^t, x_{i+1}^t \Rightarrow x_i^{t+1}$	$a_j = a_{j+4}$ $j=0, 1, 2, 3$	0, 17, 34, 51, 68, 85, 102, 119, 136, 153, 170, 187, 204, 221, 238, 255
(e)		$x_i^t \Rightarrow x_i^{t+1}$	$a_0 = a_1 = a_4 = a_5$ $a_2 = a_3 = a_6 = a_7$ $a_0 \neq a_2$	51, 204
(f)		$x_{i-1}^t \Rightarrow x_i^{t+1}$	$a_0 = a_1 = a_2 = a_3$ $a_4 = a_5 = a_6 = a_7$ $a_0 \neq a_4$	15, 240
(g)		$x_{i+1}^t \Rightarrow x_i^{t+1}$	$a_0 = a_2 = a_4 = a_6$ $a_1 = a_3 = a_5 = a_7$ $a_0 \neq a$	85, 170
(h)		$x_{i-1}^t, x_{i+1}^t, x_i^{t+1} \Rightarrow x_i^t$	$a_j \neq a_{j+2}$ $j=0, 1, 4, 5$	51, 54, 57, 60, 99,102, 105 108, 147, 150, 153, 156, 195, 198, 201, 204

Examples of other possible deterministic structures are represented below



(2.5)

The expressions of the above structures in terms of constraints on $\{x_i^t\}$ are given in Table I. In each case, the problem is to find the class of rules for which the particular deterministic structure holds. The analysis will be given in detail for two cases, and can be extended in a straightforward manner for other structures.

Notation. Let * be a "wild card" symbol denoting both of the values $\{0, 1\}$. The use of * in a relation signifies that the relation holds for both $* = 0$ or $* = 1$. The use of two *'s signifies that the relation holds for all 2^2 possible combinations of 0 and 1.

Consider case (a), which implies that x_{i+1}^t is determined by x_{i-1}^t, x_i^t, and x_i^{t+1}. Fix a choice of values for x_{i-1}^t and x_i^t, say, $x_{i-1}^t = x_i^t = 0$. Then the configurations

$$\begin{array}{ccc} 0\ 0\ 0 & \text{and} & 0\ 0\ 1 \\ \downarrow & & \downarrow \\ a_0 & & a_1 \end{array}$$

must be mapped to different values a_0 and a_1, respectively, if it is to be true that deterministic structure (a) holds. Similarly, for

$$\begin{array}{ccc} 0\ 1\ 0 & \text{and} & 0\ 1\ 1 \\ \downarrow & & \downarrow \\ a_2 & & a_3 \\ 1\ 0\ 0 & \text{and} & 1\ 0\ 1 \\ \downarrow & & \downarrow \\ a_4 & & a_5 \\ 1\ 1\ 0 & \text{and} & 1\ 1\ 1 \\ \downarrow & & \downarrow \\ a_6 & & a_7 \end{array}$$

$\{a_2, a_3\}$, $\{a_4, a_5\}$, and $\{a_6, a_7\}$ must assume pairwise differing values. Consequently, the deterministic structure will hold for any rule with a_j, $j = 0,..., 7$, such that

$$a_j \neq a_{j+1}, \qquad j = 0, 2, 4, 6 \tag{2.6}$$

Note that if relation (2.6) holds, the rule will be one-to-one in each of the subsets

$$\{000, 001\}, \{010, 011\}, \{100, 101\}, \{110, 111\} \tag{2.7}$$

Now consider case (d). Then, by the same reasoning as in the previous case, the configurations *00, *01, *10, and *11 must be mapped to pairwise equal values, where * represents the "wild card" symbol. Hence, it must be true that

$$a_j = a_{j+4}, \qquad j = 0, 1, 2, 3$$

In contrast to case (a), the rule will then be two-to-one in each of the subsets

$$\{000, 100\}, \{001, 101\}, \{010, 110\}, \{011, 111\} \tag{2.8}$$

The results of applying the above type of analysis to the remaining cases in (2.5) are summarized in Table I.

As is suggested by Table I, the lists of rule numbers corresponding to a particular deterministic structure obey periodic, or periodic-like, laws. The particular values of the periods are artifacts of the labeling system (2.3); the basic fact of periodicity is in itself, however, a reflection of the underlying structure. To understand the periodicity mechanism, consider, for example, class (h), defined by

$$a_j \neq a_{j+2}, \qquad j = 0, 1, 4, 5$$

Then any rule belonging to this class has a rule number computed from (2.3) as

$$R = a_0 \cdot 2^0 + a_1 \cdot 2^1 + \bar{a}_0 \cdot 2^2 + \bar{a}_1 \cdot 2^3 + a_4 \cdot 2^4 + a_5 \cdot 2^5 + \bar{a}_4 \cdot 2^6 + \bar{a}_5 \cdot 2^7$$

where $\bar{a}$ denotes "not a." There is a total of $2^4 = 16$ such rules. When a_0, a_1, a_4, a_5 assume all possible values, there will appear multiple "periodic" structures in the sequence of rule numbers R_i, $i = 1,..., 16$, sorted in increasing order. R_1 through R_4 will be the rules assigning the smallest possible value to $\sum_{j=4}^{7} a_j \cdot 2^j$. The differences $R_i - R_{i-1}$, $i = 2, 3, 4$, will be the differences in a sorted sequence of positive integers $\leqslant 12$ with two 1's in their binary representation; i.e., $R_i - R_{i-1} = 3$. R_5 through R_8 will be the rules assigning the next smallest value to $\sum_{j=4}^{7} a_j \cdot 2^j$. Their differences $R_i - R_{i-1}$, $i = 6, 7, 8$ will again be 3, and $R_5 - R_4$ will be equal to $(R_2 - R_1) \cdot 2^4 -$

(maximum possible $R_j - R_i$; $i, j < 4) = 3 \cdot 2^4 - (2^3 + 2^2 - (2^1 + 2^0)) = 39$. A similar analysis provides the sizes of the subsequent gaps. Clearly, the same mechanism produces the periodic, or periodic-like, structure of the rule numbers belonging to the other classes.

A useful tool for understanding the deterministic structures of this section and the results of later sections is the "directed shift" graph shown in Fig. 2. In the graph, an edge is drawn from the 3-tuple wxy to any tuple $xy*$ into which wxy can be shifted by the deletion of "w" and the appending of $* = \{0, 1\}$ on the right. Thus, for example,

$$001 \to 010, 011$$

$$110 \to 100, 101$$

In addition, for any particular rule, each node is assigned the appropriate a_j. Then the conditions for cases (a)–(h) can be re-expressed in terms of the graph. For example,

$$\text{case (c)} \qquad a_j = a_{j+1}, \qquad j = 0, 2, 4, 6$$

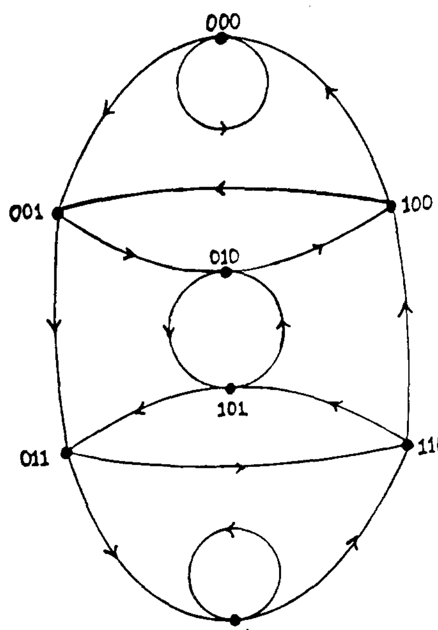


Fig. 2. Directed-shift graph. All possible 3-tuples wxy, with $w, x, y \in \{0, 1\}$, are represented as nodes in the graph. An edge is drawn from wxy to any tuple xyz into which wxy can be shifted by the deletion of w and the appending of $z \in \{0, 1\}$ on the right.

$\Leftrightarrow$ two tuples belonging to the same subset in (2.7) must be assigned equal values

$\Leftrightarrow$ two nodes in Fig. 2 with the same source must have equal values.

The directed shift graph will be useful in determining the sequences $\{x_i\}$, $\{y_i\}$, $i = 1,\ldots, n$ satisfying

$$\begin{array}{c} x_1 x_2 x_3 \cdots x_{n-2} x_{n-1} x_n \\ \downarrow \\ y_2 y_3 \cdots y_{n-2} y_{n-1} \end{array}$$

3. PERIODICITY OF SEQUENCES

Each constraint described in Section 2 induces, in a particular class of cellular automata, a deterministic structure in addition to that explicit in the formulation of the rules. The next sections will discuss some implications of those structures. It will be shown that the additional structure can be exploited in some cases to obtain results describing global properties of the associated class of automata. In this section, results of this type pertaining to the periodicity of temporal sequences will be derived.

In what follows, assume without loss of generality that $000 \to a_0 = 0$. Additional constraints, such as $001 \to a_1 = 1$ $(100 \to a_4 = 1)$, will sometimes be imposed to ensure left (right) propagation of nonzero values.

Definition. An initial condition $\{x_i^0,\ -\infty < i < \infty\}$ such that for some $-\infty < M \leqslant N < \infty$, $x_i^0 = 0$ for $i < M$, $i > N$, and $x_M^0 = x_N^0 = 1$, will be called an *arbitrary finite initial condition.*

Definition. The sequence $\{x_i^t,\ 0 \leqslant t < \infty\}$ is *periodic* if $\exists T_i, p_i$ such that $x_i^{t+p_i} = x_i^t$ for $t \geqslant T_i$.

Then it is easy to show

Theorem 1. Let $\{x_i^t\}$, $\{x_j^t\}$ be two periodic sequences with $i < j$. Then for $i < k < j$, the sequence $\{x_k^t\}$ must be periodic.

Proof. Consider the spatial string of values $\{x_k^t,\ i < k < j\}$ for any time t. The sites in the string can assume only the values $\{0, 1\}$, and therefore the string must repeat itself after at most 2^{j-i-1} time steps. Since the temporal sequences $\{x_i^t\}$ and $\{x_j^t\}$ are periodic with periods p_i and p_j, respectively, the entire "block" of sequences $\{x_k^t,\ i \leqslant k \leqslant j\}$ must be periodic with period $p \leqslant \mathrm{lcm}(p_i, p_j) \cdot 2^{j-i-1}$; i.e., there exists some T such that $x_k^t = x_k^{t+p}$ for $t \geqslant T$ and $i \leqslant k \leqslant j$.

Corollary 1. Let R be a rule that assigns $000 \to a_0 = 0$, $001 \to a_1 = 0$, and $100 \to a_4 = 0$. Then, with arbitrary finite initial conditions, the entire automaton is temporally periodic, i.e., there exist T and p such that $x_i^t = x_i^{t+p}$ for $t \geqslant T$ and all i.

Proof. The specified values of a_0 and a_1 ensure that the "left endpoint" value $x_M^0 = 1$ will not propagate to the left. Similarly, the "right endpoint" value $x_N^0 = 1$ is prevented from propagating to the right. Hence the sequences $\{x_i^t;\ t \geqslant 0\}$ are constant with all components equal to 0 for $i \leqslant M-1$ and $i \geqslant N+1$. Theorem 1 then implies the periodicity of the entire automaton.

The lemma and theorem that follow pertain to the temporal sequences generated by rules with a particular deterministic structure. Note that the deterministic structure considered implies that the rules are one-to-one in each of the subsets (2.7). The one-to-one nature of the rules will be used to establish that the temporal sequences are not periodic of any period.

Lemma 2. Let R be a rule belonging to class (a); i.e., satisfying $a_j \neq a_{j+1}$, $j = 0, 2, 4, 6$. Then given any two adjacent periodic sequences $\{x_i^t\}$, $\{x_{i+1}^t\}$, every sequence $\{x_{i+j}^t\}$ must be periodic with $T_{i+j} = \max(T_i, T_{i+1})$ and $P_{i+j} \leqslant \operatorname{lcm}(p_i, p_{i+1})$ for all $j \geqslant 2$.

Proof. If R belongs to class (a), then x_{i+2}^t is determined by x_i^t, x_{i+1}^t, and x_{i+1}^{t+1}. Hence, periodicity of $\{x_i^t\}$ and $\{x_{i+1}^t\}$ implies periodicity of $\{x_{i+2}^t\}$ with the stated properties. By induction, every sequence $\{x_{i+j}^t\}$ to the "right" must also be periodic for $j \geqslant 2$.

The theorem then follows:

Theorem 2a. Let R be a rule belonging to class (a) with $100 \to a_4 = 1$. Then, with arbitrary finite initial conditions, there can exist at most one periodic sequence $\{x_i^t\}$.

Proof. Suppose $\exists$ two periodic sequences $\{x_i^t\}$, $\{x_j^t\}$ with $i < j$. By Theorem 1, the sequence $\{x_{j-1}^t\}$ will also be periodic. Lemma 2 then implies that all sequences $\{x_{j+l}^t\}$, $l > 0$, to the "right" will also be periodic with the same $T = \max(T_{j-1}, T_j)$ and $p \leqslant \operatorname{lcm}(p_{j-1}, p_j)$. Finite initial conditions and $a_4 = 1$ imply, however, that there will be some J such that $x_{j+J}^t = 0$ for $T \leqslant t \leqslant T + p$, but $x_{j+J}^{T+q} = 1$ for some $q \geqslant p$. The contradiction implies that there cannot be two periodic sequences of any period.

The same result for rules of class (b) can be stated as:

Theorem 2b. Let R be a rule belonging to class (b) with $001 \to a_1 = 1$. Then, with arbitrary finite initial conditions, there can exist at most one periodic sequence.

The above two theorems support a conjecture on the part of Wolfram[19] that the "center" time sequence $\{x_0^t\}$ generated by Rule 30 is aperiodic. The rules for which the conditions of either Theorem 2a or b are satisfied are:

$$30, 86, 90, 150, 154, 210$$

Figure 3 depicts the evolution of a single nonzero site under Rule 30.

Theorems 2a and b can be generalized to the case of nonelementary cellular automata for which the rules depend on nearest-neighbor interactions, but the sites can assume values $V = \{0,..., k-1\}$ for arbitrary $k \geqslant 2$. Then the general form of a rule is given by

$$x_i^{t+1} = f(x_{i-1}^t, x_i^t, x_{i+1}^t); \qquad f\colon V^3 \to V \tag{3.1}$$

and there is a total of k^{k^3} rules. The conditions for a rule defined by (3.1) to possess deterministic structure (a) are given by

$$a_{l \cdot k^2 + m \cdot k^1 + n_i} \neq a_{l \cdot k^2 + m \cdot k^1 + n_j}; \qquad l, m, n_i, n_j \in V$$

for all l, m, and $n_i \neq n_j$. Similarly, the conditions for a rule to possess deterministic structure (b) are given by

$$a_{l_i \cdot k^2 + m \cdot k^1 + n} \neq a_{l_j \cdot k^2 + m \cdot k^1 + n}; \qquad l, m, n_i, n_j \in V$$

for all m, n, and $l_i \neq l_j$. Then the theorem analogous to Theorem 2 follows.

Theorem 3. Let R be a rule defined by (3.1) and either belonging to class (a) with $\{0,..., k-1\}00 \nrightarrow 0$, or belonging to class (b) with

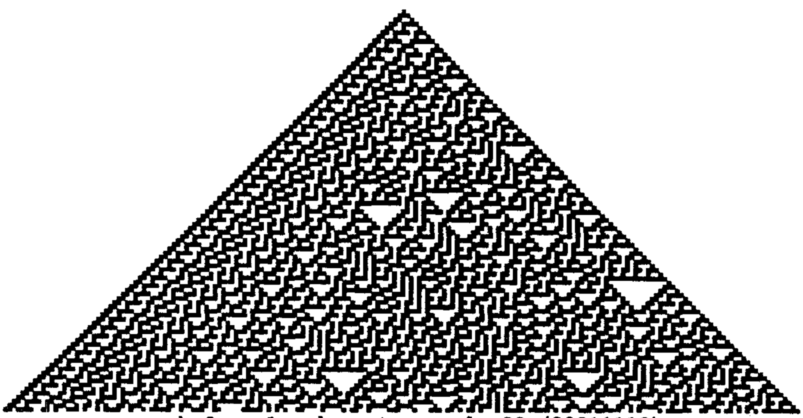


Fig. 3. Evolution of the cellular automaton defined by

$$x_i^{t+1} = x_{i-1}^t \text{ XOR } \max(x_i^t, x_{i+1}^t) \quad [\text{Rule } 30]$$

with an initial condition consisting of a single nonzero site. Site values 0 and 1 are represented by white and black, respectively.

$00\{0,..., k-1\} \nrightarrow 0$. Then with arbitrary finite initial conditions, there can be at most one periodic sequence.

Note that there will be $k!^{k^2}$ rules that exhibit deterministic structure (a), of which $(k!)^{k(k-1)}(k-1)!$ satisfy the requirements of right-propagating initial conditions and $000 \to a_0 = 0$. The same result holds for rules belonging to class (b).

Periodicity results for "diagonal" sequences may also be obtained. The next theorem establishes the periodicity of diagonal sequences for all elementary rules, and is accompanied by a corollary indicating the special periodicity properties for rules with a particular deterministic structure. The results are easily generalized to the nonelementary case.

Definition. A sequence $\{x^t_{i+t},\ t = 0, 1, 2,...\}$ will be called a *right diagonal* of x^0_i.

Definition. A right diagonal $\{x^t_{i+t}\}$ of x_i is *periodic* if $\exists J_i, p_i$ such that $x^t_{i+t} = x^{t+p_i}_{i+t+p_i}$ for $t \geqslant J_i$.

Analogous definitions can be made for left diagonal sequences.

Theorem 4. Let R be a rule defined by (2.1). Then the right and left diagonal sequences of the automaton generated by R are periodic.

Proof (by induction). Assume that rule R assigns $000 \to a_0 = 0$. Consider the right diagonal sequence $\{x^t_{i+t},\ t \geqslant 0\}$, to be denoted as $\{d^t_i\}$, and recall that $x^0_i = 0$ for $i > N$. The sequences $\{d^t_{N+1}\}$ and $\{d^t_{N+2}\}$ are then automatically constant with all elements equal to 0 and $J_{N+1} = J_{N+2} = 0$ since they lie "outside" the nonzero region of the automaton. By (2.1),

$$d^{t+1}_N = f(d^t_N, d^t_{N+1}, d^t_{N+2})$$

In particular

$$d^1_N = f(d^0_N, d^0_{N+1}, d^0_{N+2}) = f(d^0_N, 0, 0)$$

and

$$d^2_N = f(d^1_N, d^1_{N+1}, d^1_{N+2}) = f(d^1_N, 0, 0)$$

If $d^1_N = d^0_N$, then the sequence $\{d^t_N\}$ is periodic with period 1 and $J_N = 0$. If $d^1_N \neq d^0_N$, then either $d^2_N = d^1_N$ or $d^2_N = d^0_N$. In the former case, the sequence $\{d^t_N\}$ is periodic with period 1 and $J_N = 1$; in the latter, the sequence is periodic with period 2 and $J_N = 0$.

Assume now that for some $k \geqslant 1$, the diagonal sequences $\{d^t_{N-k+1}\}$ and $\{d^t_{N-k+2}\}$ are periodic with periods p_{N-k+1} and p_{N-k+2}, respectively, for $t \geqslant J_{N-k+1}$. Let $p = \operatorname{lcm}(p_{N-k+1}, p_{N-k+2})$. Then for $t \geqslant J_{N-k+1}$,

$$d^{t+1}_{N-k} = f(d^t_{N-k}, d^t_{N-k+1}, d^t_{N-k+2})$$

and

$$d_{N-k}^{t+p+1} = f(d_{N-k}^{t+p}, d_{N-k+1}^{t+p}, d_{N-k+2}^{t+p}) = f(d_{N-k}^{t+p}, d_{N-k+1}^{t}, d_{N-k+2}^{t})$$

If $d_{N-k}^{t} = d_{N-k}^{t+p}$ for $t = J_{N-k+1}$, then the sequence $\{d_{N-k}^{t}\}$ is periodic with period $p_{N-k} | p$ and $J_{N-k} = J_{N-k+1}$. If $d_{N-k}^{t} \neq d_{N-k}^{t+p}$ for $t = J_{N-k+1}$, but $f(1, d_{N-k+1}^{t+j}, d_{N-k+2}^{t+j}) = f(0, d_{N-k+1}^{t+j}, d_{N-k+2}^{t+j})$ for some $j < p$, then the sequence $\{d_{N-k}^{t}\}$ is periodic with period $p_{N-k} | p$ and $J_{N-k} = J_{N-k+1} + j$. If $f(1, d_{N-k+1}^{t+j}, d_{N-k+2}^{t+j}) \neq f(0, d_{n-k+1}^{t+j}, d_{N-k+2}^{t+j})$ for all $j < p$, then the sequence $\{d_{N-k}^{t}\}$ is periodic with period $p_{N-k} = 2p$ and J_{N-k+1}. Hence, all right diagonal sequences must be periodic. The proofs for left diagonal sequences, and for rules which assign $000 \to a_0 = 1$ are similar.

Corollary 4. Let R be a rule belonging to class (b) with $100 \to a_4 = 1$. Then with arbitrary finite initial conditions, the right diagonal sequences $\{x_{N-n+t}^{t}, t \geqslant 0\}$ will be periodic for all $n \geqslant 0$ with $J_{N-n} = n$ and periods which are powers of 2.

A symmetric corollary can be given regarding the periodicity of left diagonal sequences generated by rules belonging to class (a).

4. INVERSES FOR RULES

A second type of result that can be obtained from the analysis of deterministic structures together with the "shift transformations" described in Section 2 pertains to the existence of "inverses" for elementary cellular automata rules. Let $R\{x_i^t\}$ denote the sequence that results from applying rule R to $\{x_i^t, -\infty < i < \infty\}$. Then define the composition $R' \circ R$ of rules R and R' as $R'\{R\{x_i^t\}\}$.

Definition. Rule R^{-1} is the *inverse* of R iff $R^{-1}\{R\{x_i^t\}\} = \{x_i^t\}$.

Theorem 5a. Let R be a rule belonging to class (e), (f), or (g). Then R has an inverse R^{-1}, and $R\{000\} = R^{-1}\{000\}$.

Proof. Recall that "*" was defined in Section 2 to be a "wild card" symbol denoting both 0 and 1. For a rule belonging to class (e), $*0* \to a$ and $*1* \to \bar{a}$, with $a \neq \bar{a}$. Suppose $*0* = 0$. Then R' with $*1* \to 1$ and $*0* \to 0$ is the inverse of R. Hence the rule is its own inverse. For a rule belonging to class (f), $**0 \to a$ and $**1 \to \bar{a}$, with $a \neq \bar{a}$. Suppose $**0 = 0$. Then R' with $1** \to 1$ and $0** \to 0$ is the inverse of R. Rule $R = R^{-1}$ therefore belongs to class (g). The proof for the cases in which $*0* \to 1$, $**0 \to 1$, or $0** \to 1$ is similar.

The conditions in Theorem 5a can be shown to be necessary for a rule to possess an inverse. A lemma describing the constraints on rules with inverses follows.

Lemma 5. Define $\{b_i c_i d_i\colon R'\{b_i c_i d_i\}=0\}$. Suppose that for rule R, there exist x, y such that $x1y \to c_i$ for some i. Then the rule R has an inverse $R^{-1}=R' \Rightarrow$ either there exists no g such that $gx1 \to b_i$, or there exists no h such that $1yh \to d_i$.

Proof. Suppose $\exists g, h$ such that $gx1 \to b_i$ and $1yh \to d_i$. Then $R'\{R(gx1yh\}\}=R'\{b_i c_i d_i\}=0$, and therefore the rule is not invertible.

Theorem 5b. The only rules R with inverses R^{-1} are those satisfying the conditions of Theorem 5a.

Proof. Assume that rule R has an inverse R^{-1}, and $R\{000\}=0$. Then there are only two possibilities for R: (a) $*0* \to 0$, and (b) $*0* \to \{0, 1\}$. For (a), neither $**0$ nor $0**$ can be assigned the value 0. since either case implies $R^{-1}\{R\{*010*\}\}=R^{-1}\{000\}$, a violation of the lemma. Hence, $**0$ and $0**$ must be able to assume either value 0 or 1, and therefore the only rule with an inverse in this case is $*0* \to 0$, $*1* \to 1$. For case (b), $*0* \to \{0, 1\}$. Then if both $**0 \to \{0, 1\}$ and $**0 \to \{0, 1\}$, no inverse exists. Hence either $**0 \to 0$, $**1 \to 1$, or $0** \to 1$, $1** \to 1$.

5. HOMOGENEOUS STATES AND CONSTANT TEMPORAL SEQUENCES

The constraints discussed in Section 2 were global in the sense that a given relation, say, $0xy \to a$ and $1xy \to \bar{a}$ $(a \neq \bar{a})$ was to be satisfied by all values of x and y. In this section, the implications of less restrictive constraints will be discussed. In particular, the imposition of constraints satisfied for only specific choices of x and y, together with the analysis of tuples into which a given tuple may be "shifted," will be shown to define the classes of rules whose initial conditions either evolve to homogeneous state, or generate constant temporal sequences. As will become apparent, the constraints can be interpreted as requiring that the rules be two-to-one in at least some of the subsets (2.7) or (2.8).

Definition. An automaton is said to evolve to a *homogeneous state* if there exists $T < \infty$ such that $x_i^t = c$, c constant, for all $t \geqslant T$ and all i. (The constant c does not depend on either t or i.)

Theorem 6. A rule R evolves from all arbitrary finite initial conditions to a homogeneous state with $x_i^t = 0$ for all i and $t < \infty$ iff $a_0 = a_1 = a_2 = a_4 = 0$ and one of the following two conditions holds:

(i) $a_3 = 0$

(ii) $a_6 = 0$

Proof. First show that the conditions are necessary. As before, a_0 is assumed to be 0. (A symmetric theorem holds for the case $a_0 = 1$.) To prevent infinite propagation of the endpoint values $x_M^0 = x_N^0 = 1$, it is necessary that $001 \to a_1 = 0$ and $100 \to a_4 = 0$. If $010 \to a_2 = 1$, then a tuple 00100 appearing anywhere in the initial condition (for instance, at the "end" of the sequence) will be invariant under R. Hence $a_0 = a_1 = a_2 = a_4 = 0$. Next suppose $011 \to a_3 = 1$ and $110 \to a_6 = 1$. Then a tuple 001100 appearing anywhere in the initial sequence will be invariant. Therefore either $a_3 = 0$ or $a_6 = 0$ must be true to permit evolution to a constant state of 0.

Now show sufficiency. The proof proceeds by showing that at any time t, the last nonzero value x_i^t to the "right" or "left" must go to 0. First suppose $a_0 = a_1 = a_2 = a_3 = a_4 = 0$, and at time t, $x_i^t = 0$ for $i < M$, $x_M^t = 1$. Regardless of the value of x_{M+1}^t, it must be true that $x_M^{t+1} = 0$. Since $100 \to a_4 = 0$ prevents propagation on the right, the entire set of sites will evolve to 0 in a finite number of steps. The argument for the case $a_0 = a_1 = a_2 = a_4 = a_6 = 0$ is the same.

Remark. It has been assumed throughout that $000 \to a_0 = 0$. If this assumption is removed, then the conditions for evolution to a constant state with all site values equal to 1 are symmetric to those stated in Theorem 6.

The total number of rules for which arbitrary finite initial conditions evolve to a homogeneous zero state is therefore $2^3 + 2^3 - 2^2 = 12$. The rule numbers for which this behavior occurs are

$$0, 8, 32, 40, 64, 96, 128, 136, 160, 168, 192, 224$$

Note that the rules satisfy relaxed forms of the constraints for classes (c) and (d); namely, $\{00*, *00\} \to 0$ and either $01* \to 0$ or $*10 \to 0$, where "*" was defined in Section 2. Hence the rules are two-to-one in at least two of the subsets of (2.7) or (2.8).

Next consider the class of rules for which evolution from any finite initial condition generates at least one constant temporal sequence $\{x_i^t\}$. As in the case of evolution to a constant state, the rules exhibiting this behavior will again be two-to-one in some subsets of (2.7) or (2.8).

Definition. A sequence $\{x_i^t,\ t \geq 0\}$ is *constant* if $\exists T < \infty$ such that $x_i^t = C$, C constant for all $t \geq T$.

Definition. Denote the initial condition by $\{x_i^0,\ -\infty < i < \infty\}$ where for some M and N, $x_i^0 = 0$ for $i < M$, $i > N$, and $x_M^0 = x_N^0 = 1$. Then the *length* of the initial condition is defined to be $N - M + 1$.

Definition. A string $\{x_i^t,\ -\infty < i < \infty\}$ is *symmetric* at time t if $x_{i-j}^t = x_{i+j}^t$ for all $j \geqslant 0$ and some x_i^t.

It is easy to show

Lemma 7.1. Let $y_1,\ldots, y_n$ denote elements of $\{0, 1\}$. Then a rule R will generate constant temporal sequences if $*y_1 y_2 \to y_1,\ y_1 y_2 y_3 \to y_2,\ldots,\ y_{n-2} y_{n-1} y_n \to y_{n-1},\ y_{n-1} y_n{*} \to y_n$, and there exist t and i such that $x_{i+j}^t \times y_j$ for $j = 1,\ldots, n$.

Note a rule satisfying the condition of Lemma 7.1 will again satisfy relaxed forms of the constraints defining classes (c) and (d), and thus will be "two-to-one" in at least one subset of both (2.7) and (2.8).

The theorems that follow provide necessary and sufficient conditions for the generation of constant nonzero temporal sequences; i.e., sequences with $x_i^t = 1$ for $t \geqslant T$, T finite. Symmetric results hold for the case of constant zero temporal sequences. As a preliminary but fundamental result, the next lemma provides a condition for the generation of at least two adjacent constant temporal sequences by a rule exhibiting both left and right propagation.

Lemma 7.2. Suppose $a_1 = a_4 = 1$ and there exist sequences $\{x_i^t\}$, $\{x_{i+1}^t\}$ and some T such that for $t \geqslant T$, $x_i^t = x_i$ and $x_{i+1}^t = x_{i+1}$, where x_i, x_{i+1} are constants. Then there must exist $y_1, y_2, y_3, y_4 \in \{0, 1\}$ for which $*y_1 y_2 \to y_1$ and $y_3 y_4{*} \to y_4$.

Proof. Suppose there are no values y_1, y_2 for which $*y_1 y_2 \to y_1$. Then, in order for $\{x_i^t\}$ and $\{x_{i+1}^t\}$ to maintain constant values, the sequence $\{x_{i-1}^t\}$ must be constant for $t \geqslant T$. By induction, this establishes that all sequences $\{x_{i-j}^t\}$, $j \geqslant 0$, must also be constant for $t \geqslant T$. Finite initial conditions imply, however, that not all sequences to the "left" are constant, and hence there must exist y_1, y_2 such that $*y_1 y_2 \to y_1$. Similarly, there must exist y_3, y_4 such that $y_3 y_4{*} \to y_4$.

Theorem 7a. A rule R with $a_0 = 0$, $a_1 = a_4 = 1$ will generate from arbitrary finite initial conditions at least one constant nonzero temporal sequence iff one of the following conditions holds

(i) $\{*11, 11*\} \to 1$ $(a_3 = a_6 = a_7 = 1)$, and (11) either appears in the initial condition, or is generated under R;

(ii) $\{*10, 01*\} \to 1,\ 101 \to 0$ $(a_2 = a_3 = a_6 = 1,\ a_5 = 0)$, and (101) either appears in the initial condition, or is generated under R;

(iii) $\{010, 111\} \to 1,\ \{011, 110, 101\} \to 0$ $(a_2 = a_7 = 1,\ a_3 = a_5 = a_6 = 0)$, and the initial condition is symmetric with respect to some x_i^0;

(iv) $\{010, 101, 111\} \to 1,\ \{011, 110\} \to 0$ $(a_2 = a_5 = a_7 = 1,\ a_3 = a_6 = 0)$, and the initial condition is of odd length and generates, for some $t \geqslant 0$, a sequence symmetric with respect to some x_i^t.

Proof. Sufficiently of conditions (i) and (ii) follows from Lemma 7.1. Sufficiency of (iii) and (iv) is easy to show.

To show that the conditions are necessary, assume the existence of at least one constant nonzero sequence. Consider, for some time T, the "widest" blocks $\{x_1^t \cdots x_n^t\}$ such that $x_i^t = x_i$ is constant (possibly equal to 0) for $t \geqslant T$. For any T, multiple such blocks may exist, and each block must be finite in width since the initial conditions propagate in both directions.

Suppose there is a block of width $n \geqslant 2$. Then Lemma 7.2 implies that there exist y_1, y_2, y_3, y_4 satisfying $*y_1 y_2 \to y_1$, $y_3 y_4 * \to y_4$. Since $001 \to a_1 = 1$ and $100 \to a_4 = 1$, the only possible values are given by

(1) $y_1 y_2 = 10$, $y_3 y_4 = 01$

(2) $y_1 y_2 = 10$, $y_3 y_4 = 11$

(3) $y_1 y_2 = 11$, $y_3 y_4 = 11$

Subcase (1) implies $*10 \to 1$, $01* \to 1$, and $101 \to 1$. Subcase (2) implies $*10 \to 1$, $11* \to 1$, and $011 \to 1$, and hence $*11 \to 1$. Subcase (3) implies $\{*11, 11*\} \to 1$. Thus the conditions of the theorem have been shown to be necessary for blocks of width $n \geqslant 2$.

Now suppose there exists a nontrivial block of width $n = 1$; i.e., there exists a constant nonzero sequence $\{x_i^t\}$ neither of whose neighboring sequences is constant. Define the set

$$E = \{(xyz) \colon x = x_{i-1}^{t'},\ y = x_i^{t'},\ z = x_{i+1}^{t'} \text{ for some } t' \geqslant T\}$$

representing the set of "embeddings" of the sequence $\{x_i^t\}$ which can occur under the rule. Then the possibilities for the set E can be summarized as follows

(a) $E = \{(010), (111)\}$

(b) $E = \{(011), (110)\}$

(c) $E = \{(010), (011), (110)\}$ or $\{(011), (110), (111)\}$ or $\{(010), (110), (111)\}$

(d) $E = \{(*1*)\} = \{(010), (011), (110), (111)\}$

Note that any other set, e.g., $E = \{(011)\}$, would imply that one of the adjacent sequences is constant.

First consider case (a). It can be assumed that $x_{i-1}^t x_i^t x_{i+1}^t = 111$ and $x_{i-1}^{t+1} x_i^{t+1} x_{i+1}^{t+1} = 010$ for some $t \geqslant T$. Then $111 \to a_7 = 1$ and $010 \to a_2 = 1$. There must exist some $y \in \{0, 1\}$ for which $x_i^t x_{i+1}^t y = 11y \to 0$, and therefore $110 \to a=1$. Similarly, $011 \to a_3=0$. The only remaining undetermined a_j is a_5. Suppose $101 \to a_5=0$. Then the rule satisfies the conditions defining deterministic structures (a) and (b) from Section 2, and hence the symmetry of $\{x_{i-1}^t\}$ and $\{x_{i+1}^t\}$ implies the symmetry of the entire automaton with respect to $\{x_i^t\}$. It can be seen (by considering, for instance, the directed-shift graph of Fig. 2) that symmetry of the automaton generated with the specified values a_j requires symmetry of the initial condition, and thus the necessity of condition (iii) is established.

Now suppose for case (a) that $101 \to a_5=1$. Recall that the initial condition is denoted by $\{x_i^0\}$ where $x_i=0$ for $i>N$, $i<M$, and $x_N^0=x_M^0=1$. Assume, without loss of generality, that $x_{M+1}^0=1$. It can easily be verified that

$$x_{M-1-t}^t x_{M-t}^t x_{M+1-t}^t = \begin{cases} 010 & \text{for odd } t \geqslant 0 \\ 011 & \text{for even } t \geqslant 0 \end{cases}$$

Thus, it must be true that the diagonal elements $x_{M-1-t}^t=0$, $x_{M-t}^t=1$, and the diagonal sequence $\{x_{M+1-t}^t\}$ is periodic with period 2 for $t \geqslant 0$. Since both $100 \to a_4=1$ and $101 \to a_5=1$, it follows that the diagonal elements $x_{M+2-t}^t=1$ for $t \geqslant 2$. It is then straightforward to show (the proof proceeds along the lines of that for Theorem 4) that there exist values $T_{M+k}<\infty$ such that the diagonal sequences $\{x_{M+k-t}^t\}$ are periodic with period 2^P, $p \geqslant 0$, for odd $k \geqslant 3$, and constant with all elements equal to 1 for even $k \geqslant 4$ and $t \geqslant T_{M+k}$. The same result holds for the diagonal sequences $\{x_{N-k+t}^t\}$.

To show the necessity of condition (iv) for the above rule, assume the existence of a constant nonzero temporal sequence. It follows from the value specified for a_j that the rule violates the conditions defining deterministic structures (a) and (b) in that $\{*01, 10*\} \to 1$, but satisfies them otherwise. The proof proceeds by showing that the partial determinism suffices to induce symmetry in a region of the automaton "local" to the constant sequence. It can be assumed that $x_{i-1}^s x_i^s x_{i+1}^s = 111$ and $x_{i-1}^{s+1} x_i^{s+1} x_{i+1}^{s+1} = 010$ for some $s \geqslant 0$. Since $011 \to a_1=1$, $100 \to a_4=1$, and $101 \to a_5=1$, there is no $y \in \{0, 1\}$ for which either $10y \to 0$ or $y01 \to 0$, and therefore $x_{i-1}^{s+2} x_i^{s+2} x_{i+1}^{s+2} = 111$. It is easily verified that $x_{i-2}^t = x_{i+2}^t$ for $t \geqslant s$. Consider now the temporal "cone" of the automaton bordered by $\{x_{i+j}^s,\ -2 \leqslant j \leqslant 2\}$, and the diagonal sequences $\{x_{i-2-t}^{s+t},\ t \geqslant 0\}$ and $\{x_{i+2+t}^{s+t},\ t \geqslant 0\}$. It can be shown that no two adjacent 0's can occur in the temporal sequences within the defined region; i.e., $x_k^t=0 \Rightarrow x_k^{t-1}=x_k^{t+1}=1$, if x_k^{t-1}, x_k^t, and x_k^{t+1} are contained in the cone. It follows that the rule is deterministic in the cone, and moreover the cone must be symmetric with respect to $\{x_i^t\}$ for $t \geqslant s$. Periodicity of the diagonal sequences then implies that for some $S > s$, all diagonal sequences to the "left" and "right" of the cone must be alternatingly periodic of period 2^p, $p \geqslant 0$, and constant with all elements equal to 1 for $t \geqslant S$. As a result, the partial determinism of the rule implies symmetry of the entire automaton for $t \geqslant S$. Thus condition (iv) has been shown to be necessary.

Next consider case (b). As in case (a), it can be deduced that $011 \to a_3 = 0$, $110 \to a_6 = 0$, and $111 \to a_7 = 0$. If both $010 \to a_3 = 0$ and $101 \to a_5 = 0$, then the rule satisfies the conditions defining deterministic structures (a) and (b). It follows from the values specified for a_j, $j = 0, 1,..., 7$, that if $\{x_{i-1}^t\}$ and $\{x_{i+1}^t\}$ are "asymmetric" (i.e., $x_{i-1}^t \neq x_{i+1}^t$ for all $t \geqslant T$), then $\{x_{i-2}^t\}$ and $\{x_{i+2}^t\}$ are also asymmetric, but $\{x_{i-3}^t\}$ and $\{x_{i+3}^t\}$ must be symmetric with respect to $\{x_i^t\}$ for $t \geqslant T$. By induction, it is established that $\{x_{i-3j-1}^t\}$ and $\{x_{i+3j+1}^t\}$ must be asymmetric for all $j \geqslant 0$. The finiteness of initial conditions then produces a contradiction. It can be shown that a contradiction also results from choosing either $010 \to a_3 = 0$, $101 \to a_5 = 1$; or $010 \to a_3 = 1$, $101 \to a_5 = 1$. The case in which $010 \to a_3 = 1$, $101 \to a_5 = 0$ is covered by condition (ii).

Finally, it is straightforward to verify that choosing E to be a set included in case (c) or (d) leads either to a contradiction, or to one of the conditions shown to be necessary for the existence of a block of width $n \geqslant 2$.

Corollary 7.1. Suppose rule R satisfies condition (i) of Theorem 7a. Then the initial conditions that generate at least one constant nonzero temporal sequence include

(a) any string in which either (11) or $(1x_1 \cdots x_n 1)$ appears, where $x_1 = \cdots = x_n = 0$ and n is even;

(b) all possible strings, if $010 \to a_2 = 1$.

Corollary 7.2. Suppose rule R satisfies condition (ii) of Theorem 7a. Then the initial conditions that generate at least one constant nonzero temporal sequence include

(a) any string in which $(1x_1 \cdots x_n 1)$ appears, where $x_1 = \cdots = x_n = 0$ and n is odd;

(b) any string in which $(0x_1 \cdots x_n 0)$ appears, where $x_1 = \cdots = x_n = 1$ and $n \geqslant 3$ is odd, if $111 \to a_7 = 0$.

Remark. A test can be used to check whether a given initial condition will generate a symmetric string for some $t < \infty$, and thus satisfy condition (iv) of the theorem. Recall that the diagonal sequences $\{x_{M+k-t}^t, t \geqslant 0\}$ and $\{x_{N-k+t}^t, t \geqslant 0\}$ generated from rule R with $a_0 = a_3 = a_6 = 0$ and $a_1 = a_2 = a_4 = a_5 = a_7 = 1$ possess certain periodicity properties. Specifically, it can be shown that the diagonal sequence $\{x^t_{M+k-t}\}$ will be constant with all elements equal to 1 for even $k \geqslant 0$, and $\{x^t_{M+k+1-t}\}$ will be periodic of period $\leqslant 2^{L+1}$ for $t \geqslant \sum_{l=1}^{L} 2^l$, where $L = k/2$. The same results hold for the diagonal sequences $\{x^t_{N-k+t}\}$. For the automaton to be symmetric for some $t \geqslant S$, it must be true that the diagonal elements $x^t_{M+k-t} = x^t_{N-k+t}$ for $t \geqslant S$ and all $k \geqslant 0$. Hence, an initial condition will generate a symmetric string iff

$$x^t_{M+k-t} = x^t_{N-k+t} \qquad \text{for} \quad t \geqslant \sum_{l=1}^{L} 2^l \tag{5.1}$$

with $L = [k/2] =$ greatest integer $\leqslant k/2$, $k \geqslant 0$, and $x^0_{M+1} = x^0_{N-1}$. If condition (5.1) is violated for any $k \geqslant 0$, the automaton will not become symmetric for any t.

Theorems for the case where either $001 \to a_1 \neq 0$ or $100 \to a_4 \neq 0$ can be stated as follows. Corollaries analogous to those for Theorem 7a are easily obtained.

Theorem 7b. A rule R with $a_0 = a_1 = 0$, $a_4 = 1$ will generate from arbitrary finite initial conditions at least one constant nonzero temporal sequence iff one of the following conditions holds:

(i) $\{*11, 11*\} \to 1$ $(a_3 = a_6 = a_7 = 1)$, and (11) either appears in the initial condition, or is generated under R;

(ii) $01* \to 1$ $(a_2 = a_3 = 1)$

Proof. The proof proceeds along much the same lines as for Theorem 7a. The stable sequences for this case may be bounded on the left by an infinite number of 0 sequences, giving rise to condition (ii). The sufficiency of condition (ii) is proved by considering the "left-most" value $x_M = 1$ in the initial condition.

The next two theorems are stated without proof.

Theorem 7c. A rule R with $a_0 = 0$, $a_1 = 1$, $a_4 = 0$ will generate from arbitrary finite initial conditions at least one constant nonzero temporal sequence iff one of the following conditions holds

(i) $\{*11, 11*\} \to 1$ $(a_3 = a_6 = a_7 = 1)$, and (11) either appears in the initial condition, or is generated under R;

(ii) $*10 \to 1$ $(a_2 = a_6 = 1)$

Theorem 7d. A rule R with $a_0 = a_1 = a_4 = 0$ will generate from arbitrary finite initial conditions at least one constant nonzero temporal sequence iff one of the following conditions holds

(i) $\{*11, 11*\} \to 1$ $(a_3 = a_6 = a_7 = 1)$, and (11) either appears in the initial condition, or is generated under R;

(ii) $010 \to 1$ $(a_2 = 1)$, and (00100) either appears in the initial condition or is generated under R;

(iii) $\{011, 110\} \to 1$ $(a_3 = a_6 = 1)$, and (001100) either appears in the initial condition, or is generated under R.

6. SUMMARY

Elementary cellular automata are defined as automata whose sites can assume either of the values $\{0, 1\}$, and whose rules depend on nearest-neighbor interactions. Such automata are inherently deterministic in that the value of a site at any time is determined by the values of its neighboring sites at the previous time step; i.e., x_i^{t+1} is determined by x_{i-1}^t, x_i^t, and x_{i+1}^t. It has been shown that certain classes of elementary rules exhibit an additional determinism which results from the specific choice of values they assign to the set of possible 3-tuples (representing the set of possible tuples $(x_{i-1}^t, x_i^t, x_{i+1}^t)$.) These deterministic structures are directly tied to the one-to-one, versus many-to-one, nature of the rule restricted to the subsets $\{(x, y, 0), (x, y, 1)\}$ and $\{(0, x, y), (1, x, y)\}$ defined for all $x, y \in \{0, 1\}$. The subsets can be regarded as the subset of 3-tuples into which a given pair (x, y) can be shift transformed.

The analysis of deterministic structures and shift transformations makes possible the derivation of results describing certain global properties of cellular automata. These results are consistent with intuitive notions of differences between one-to-one and many-to-one mappings. In particular, it has been shown that the generations of either a homogeneous state or constant temporal sequences require that the underlying automaton rule be, at least in part, two-to-one. On the other hand, rules that are one-to-one in a well-defined sense generate aperiodic behavior.

Finally, in interpreting the results presented in this paper, it is useful to recall Wolfram's classification scheme described in the introduction. Elementary cellular automata are conjectured by Wolfram to fall into Class 1, 2, or 3; Class 4 automata have been observed only for k (number of possible values for each node) > 2, and r (size of neighborhood) > 1.[17] Theorem 6 defines the class of elementary rules for which arbitrary finite initial conditions evolve to a homogeneous state, and thus characterizes rules belonging to Class 1. Theorem 7 provides necessary and sufficient conditions for rules to generate constant temporal sequences, a property that usually, but not always, coincides with the qualitative one describing Class 2. (Recall that a rule belonging to Class 2 generates constant temporal sequences that are "separated" and "simple.") The rules satisfying the conditions of Theorem 2 generate infinitely many aperiodic sequences. Such rules clearly do not belong to either Class 1 or Class 2. Since the definition of aperiodicity overlaps, but does not coincide with, that of "chaotic" behavior, these rules constitute a class that presumably overlaps Class 3. Rule 150 is an example of a cellular automaton satisfying the conditions of both Theorems 2 and 6; given symmetric initial conditions, it generates one constant temporal sequence (in the "center"), and every other sequence is aperiodic. The implications of the results on the generation of constant temporal sequences for the use of cellular automata in pattern recognition are discussed in Ref. 5.

ACKNOWLEDGMENTS

I am grateful to Stephen Wolfram, whose insights have been essential in providing the motivation and the tools to study cellular automata, and whose suggestion that I investigate Rule 30 stimulated the development of the ideas in this paper. I wish also to express appreciation to George Zweig for numerous, extensive, and highly useful discussions.

REFERENCES

1. J. Anderson, *IEEE Trans. SMC* **13**:799 (1983).
2. D. Farmer, T. Toffoli, and S. Wolfram, eds., "Cellular Automata: Proceedings of an Interdisciplinary Workshop," *Physica* **10D** (1–2) (1984).
3. U. Frisch, B. Hasslacher, and Y. Pomeau, "A Lattice Gas Automaton for the Navier–Stokes Equation" (preprint).
4. G. Hedlund, *Math. Systems Theory* **3**:320 (1969).
5. E. Jen, *J. Stat. Phys.* **43**:243 (1986).
6. D. A. Lind, Reference (3).
7. O. Martin, A. Odlyzko, and S. Wolfram, *Comm. Math. Phys.* **93**:219 (1984).
8. J. Milnor, "Entropy of Cellular Automaton-Maps," Institute for Advanced Study preprint (May 1984).
9. J. Milnor, "Notes on Surjective Cellular Automaton-Maps," Institute for Advanced Study preprint (June 1984).
10. N. Packard, "Deterministic Lattice Models for Solidification and Aggregation" (preprint).
11. N. Packard, "Complexity of Growing Patterns in Cellular Automata" (preprint).
12. T. Toffoli, Reference (3).
13. J. von Neumann, *Theory of Self-Reproducing Automata*, A. W. Burks, ed. (Univ. Illinois Press, Urbana, 1966).
14. M. S. Waterman, Reference (3).
15. S. Wolfram, *Rev. Mod. Phys.* **55**:601 (1983).
16. S. Wolfram, *Physica* **10D**:1 (1984).
17. S. Wolfram, *Los Alamos Science* (Fall 1983).
18. S. Wolfram, *Comm. Math. Phys.* **96**:15 (1984).
19. S. Wolfram, *Phys. Rev. Lett.* **55**:449 (1985).