[29]

ATTRACTORS, BASIN STRUCTURES AND INFORMATION PROCESSING IN CELLULAR AUTOMATA

Kunihiko KANEKO
Institute of Physics, College of Arts and Sciences
University of Tokyo, Komaba, Meguro, Tokyo 153, JAPAN

One-dimensional cellular automata (CA) are investigated. An information theory for multi-attractor systems is constructed, and quantities are introduced to characterize the complexity of basin volumes, stability of attractors against noise, information storage in attractors, and connectivity among attractors by noise. These quantities and basin structures are calculated numerically for one-dimensional CA of various classes. The patterns of the main attractors are shown, focusing on the holographic memory of class-3 CA, successive changes of main attractors with size in class-4 CA, and soliton-like attractors in some CA.

1. Introduction

Spatially extended dynamical systems are important tools to understand the complex behavior in nature. The simplest model in such systems is a cellular automaton (CA), where the system is composed of a discrete time n and space i (lattice) and discrete variables $x_n(i)$'s on the lattice. It was introduced for the computer architecture and will also be useful to understand the qualitative nature in turbulence (in a wide sense) or neural networks or some other biological systems[1-3]. In this paper, the complexity in CA with a finite size is investigated.

Computer systems or artificial intelligence has recently been investigated from the viewpoint of a dynamical system theory. When we consider some kind of artificial intelligence on the basis of a dynamical system, informational aspects are important. As was pointed out by Rob Shaw[4], a dynamical system with chaos can be thought of as an information source.

Another aspect of a dynamical system is a storage of information. If a dynamical system has a large number of attractors as is commonly seen in the spatially extended systems, information can be stored in each attractor. Examples can be seen in the neural network models (see e.g., Hopfield[5]). In the following sections, an information theory for a multi-attractor system is constructed, mainly in connection with the CA with a finite lattice size. From the quantities introduced in the follow-

Stability of the storage is also important which is related to the problem of self-repair or retrieval[6]. We will consider the stability of each attractor against a noise and define the mutual information between attractors.

In this paper a one-dimensional cellular automaton with a lattice size N is investigated. If a state of each cell can take k values, the total number of states is k^N. Thus, the system finally settles down into a cycle.

From the viewpoint of the creation and storage of information, a cellular automaton can be classified into the following four types. The classification is essentially the same as the one by S. Wolfram[1,7], though the precise definition for the classification is not available at present. In the following, "a large number" means a quantity exponential to the system size, $(O(e^N))$, while "a small number" means a quantity less than some power of the system size $(o(N^a))$. The period of an attractor is said to be long if the period is $O(N)$, while it is short if it is bounded by $O(1)$.

(1) Small number of attractors with short periods: (No creation and small storage of information): class 1
(2) A large number of attractors with short periods: (No creation but large storage of information): class 2
(3) A small number of attractors with long periods: (Positive creation with small storage of information): class 3
(4) A large number of attractors with possible long periods: (Possible positive creation with large storage of information): class 4

In Sections 2 and 3, a framework to study the complexity of multi-attractor systems is introduced, where the complexity in the volume of basins and the dynamical aspects by the jumping process among attractors by a small noise (i.e., single site flip-flop) are investigated.

In Sections 4-7, the structure of attractors and basins are investigated for various classes of CA, mainly focusing on the number of attractors, the distribution of the volume of basins of attraction, the probability distribution at each attractor by the noise, and jumping process among attractors by the noise. The quantities introduced in Sections 2 and 3 are calculated.

2. Information Theory for Multi-attractor Systems

Here we study the number of attractors and the structure of the basin for each attractor in a CA with a finite size[9]. A quantity to characterize the complexity of the basin of attraction is introduced.

(a) Complexity of basins

Let us denote the number of attractors by M. Each attractor is denoted by $\{a_i\}$

($i = 1, 2, \ldots, M$). First, we examine how many initial configurations are attracted into each attractor a_i. The number of configurations divided by k^N gives the ratio of the basin volume for the attractor a_i, which is denoted by b_i ($\Sigma b_i = 1$). Let us define the complexity for basins by

$$C_B = -\Sigma b_i \ln b_i \quad ,$$

which characterizes the information for the initial state necessary to predict the final state. If each attractor has an equal volume of basins, $C_B = \ln M$ (maximum). In many cases studied here, the complexity is much less than $\ln M$, since the volume of basin of most attractors is very small. The distribution of the basin volume b_i itself is also important.

(b) Period of attractors and contraction ratio

Let us denote the period of each attractor by T_i. The summation $c = \Sigma T_i$ gives the volume of phase space utilized by CA after the transients have decayed out. If we start from all possible initial configurations, only a limited number of states remain after some iterations. The ratio of contraction for the process is given by $c/2^N$. For a discussion about the contraction of CA from a different point of view, see Hogg and Huberman[8].

3. Dynamical Process Among Attractors

(a) Jumping among attractors by noise[10]

Since most CA have more than one attractor, a unique invariant measure cannot be attained. In a real physical situation, existence of a small noise is expected. By the noise, a unique (or a small number of) measure is selected out. Here we consider the case with a very low noise. The noise takes only an integer value in CA, and the "low" noise here means that the rate of the application of noise is very low. In the low noise case, the dynamics of the system may be decomposed into the following two processes; (i) the state stays at the original attractors (for most of the time) and (ii) the state jumps out to some other attractors by the effect of a noise. The state of a CA stays at the original attractors $\{a_i$'s$\}$ most of the time and the transition among attractors by a noise occurs in a short time interval. Here, we neglect the time for the latter process as small compared with the time for the former. An example of a pattern for a stochastic CA is shown in Fig. 1, where the transitions among the attractors can be seen.

Let us define the transition matrix between attractors. If a low noise is applied on the attractor a_i, a jump from the attractor a_i to a_j occurs with some probability. Here the noise is a process $x \to x' = x + r \pmod{k}$ $(0 < r < k)$. The jumping

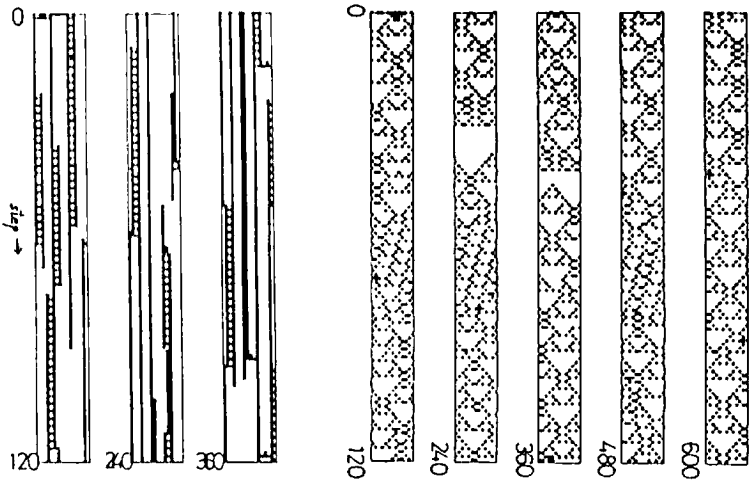


Fig. 1. Examples of the evolution of stochastic cellular automata. The flip-flop $0 \longleftrightarrow 1$ or $1 \longleftrightarrow 0$ by a noise occurs with the ratio p.
(a) Rule 108: noise $p = 0.05$: size $N = 16$
(b) Rule 146: noise $p = 0.05$: size $N = 13$

process depends on the state of the CA when the noise is applied (i.e., the phase of the oscillation with the period T_i; T_i possibilities) and on the lattice site at which the noise is applied (N (= system size) possibilities) and on the value of noise $r((k-1)$ possibilities). If the CA has only two states ($k = 2$), there is only one possibility for the noise (i.e., $1 \to 0$ or $0 \to 1$). In the following, we treat the case $k = 2$ (see Fig. 2 schematically).

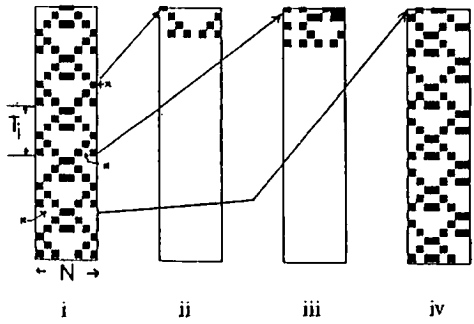


Fig. 2. Schematic representation of the transition by the noise on an attractor. Rule 146 and $N = 8$. The attractor (i) has a period 6 and there are 6×8 possibilities on the application of noise. In the figure three examples are shown, where the noise is applied on the position marked x. Two ((ii) and (iii)) go to the attractor all-0, while the other (iv) goes to the same attractor as (i) with 2 shifts.

P_{ij} is defined as the ratio of the transition from a_i to a_j. That is, the number of the events $a_i \to a_j$ for all the possible flip-flops by the noise, divided by the number of such possibilities $T_i N$.

The probability that a system is in the attractor a_i for a low noise case is given by

q_i = the i-th component of the eigenvector for the matrix P_{ji} corresponding to the eigenvalue 1.
(i.e., $\sum_i q_i P_{ij} = q_j$)

If the eigenvalue 1 is degenerate with the multiplicity m_p, the superposition of each eigenvector can be an invariant measure within the above low-noise limit approximation.

Some attractor a_i can be so weak that P_{ki} is zero for all k, (i.e., there is no flow from other attractors by a low noise) as can be seen in the next section. Such an attractor is repulsive against a noise and may be regarded as an extension of the notion of garden of Eden to attractors. If q_i is zero (which is a weaker condition than the above $P_{ki} = 0$-condition), the residence probability at the attractor a_i is zero, and such an attractor is unimportant in the low-noise limit.

The diagonal part P_{ii} is a measure for the strength of self-repair for the attractor a_i against a one-position flip-flop.

(b) Complexity in the jumping process among attractors by noise

Once the probability measure q_i is attained by a noise, we can define the complexity for the probability distribution for attractors by

$$C_M = -\Sigma q_i \ln q_i \quad ,$$

for a given invariant measure. The meaning of the quantity C_M is as follows: After the transients have decayed out for a low noise system, we make a measurement to determine at which attractor the system stays at that time. The information gain by the above measurement is C_M. The difference between C_B and C_M lies in that the former quantity is concerned with the knowledge about the initial state, while the latter is related to the observation for the aged system with a noise.

Another important quantity is a dynamical information gain by noise. Let us assume that we knew that a system had initially been at the attractor a_i and have observed that the system is now at the attractor a_j after a noise was applied. How much information has been obtained through this observation? We can get some information about the noise, i.e., the phase of the oscillation of CA when the noise is applied and the site where the noise is applied. The amount is given by $\ln (P_{ij}^{-1})$

Thus the dynamical information gain per noise is given by

$$C_D = -\sum_{ij} q_i P_{ij} \ln P_{ij} \quad ,$$

since the ratio for the event $a_i \longrightarrow a_j$ is $q_i P_{ij}$.

As is easily seen,

$$C_T = C_M - C_D$$

is non-negative. The quantity C_T corresponds to the mutual information[11] between attractors by noise.

If C_D is large, the information creation by noise is large. That is, the uncertainty about the attractor into which the system settles down after the addition of noise is large. It can also be stated that if the mutual information is large ($C_D \ll C_M$), the structure of the network of the transition among attractors is well organized, while the network of the transition is global and irregular if C_T is small.

(c) Method of the calculation in one-dimensional CA

As a simple example of the theory for multi-attractor systems in this section, one-dimensional cellular automata with two states (0 or 1) are investigated. The periodic boundary condition is used throughout this paper. The models are

(i) legal cellular automata with range 1[1]

(ii) totalistic cellular automata with range 2[7]

(iii) cellular automata with range 2 which have "soliton"-like excitations[12].

As a method for the coding for the rule, the rule number (for the model (i)) or the rule code (for the model (ii)) by Stephen Wolfram[1,7] is used, while the rule code by Aizawa, Nishikawa and the author[12] is used to characterize the rule for the model (iii) (see Appendix). In the following sections, we use the notation Rule *** for (i) and totalistic *** for (ii), where *** is a number which characterizes the rule or code. For (iii), we use $(l_1 l_2 l_3 l_4 l_5 l_6 l_7)(k_1 k_2 k_3 k_4 k_5)$, where k_i and l_i take 0 or 1.

The method of calculations is as follows: (i) Take a one-dimensional cellular automaton with a size N ($7 < N < 23$) and simulate it for all initial configurations (i.e., 2^N possibilities). (ii) Enumerate all possible attractors (find M, a_i ($i = 1, 2, \ldots, M$)), and their periods T_i's and list all the patterns. (iii) Calculate how many initial configurations are attracted into the attractor a_i. The number of such initial configurations divided by 2^N gives b_i, from which the basin complexity C_B is calculated. (iv) Take an attractor a_i and change a value of one lattice site for a_i ($0 \leftrightarrow 1$). There are $N \times T_i$ possibilities for this flip-flop. We simulate the CA starting from the configuration obtained by all these possible flip-flops and

check to which attractor a_j the state is attracted. The number of such configurations divided by $N \times T_i$ gives P_{ij}. The left eigenvector for P_{ij} corresponding to the eigenvalue 1 gives q_i. From P_{ij} and q_i, measure complexity C_M and dynamical complexity C_D are calculated.

Here, instead of obtaining all possible eigenvectors, we choose an initial vector $(b_1, b_2, b_3, \ldots, b_M)^T$ and multiply the matrix $\{P_{ij}\}$ many times till the set of vectors is settled down into the fixed point, from which we obtain $\{q_1, q_2, \ldots, q_M\}^T$. If the invariant measure is unique, this procedure gives the correct measure. If the measure is not unique (i.e., nonergodic), this procedure selects out one measure closest to the equipartition distribution for all the possible configurations. For a finite one-dimensional CA, such a nonergodic case seems to be rare except the following case; i.e., the all-0 attractor is sometimes disconnected from other attractors and there are two eigenvectors for the eigenvalue 1; one is $q_i^1 = 1$ for the all-0 attractor and $= 0$ otherwise, and the other is $q_i^2 = 0$ for all-0.

For the classification of attractors, the configurations which coincide by the spatial translation are regarded as the same attractor. For example, the patterns 11000001, 11100000 and 00111000 are regarded as the same.

The results for the various classes of CA are shown in the following three sections. See Ref. 13 for the preliminary report.

4. Class-1 and Class-2 CA

(a) Class-1 CA

As is expected, this case is trivial. The number of attractors M remains small (about $1 \sim 3$) even if N is increased. As N goes larger C_B, C_M, and C_D rapidly go to zero. For example, the possible attractor is all-0, and 0101010101 (which is possible only for N = even) for Rule 32.

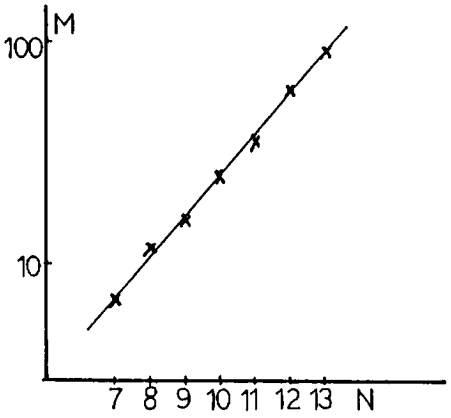
(b) Class-2 CA

The following rules were numerically investigated for $7 < N < 19$: elementary rules 132, 50, 108, and totalistic rules 24, and 104. Some examples for the number of attractors and three complexities are shown in Figs. 3 and 4. Let us discuss the common properties for these rules.

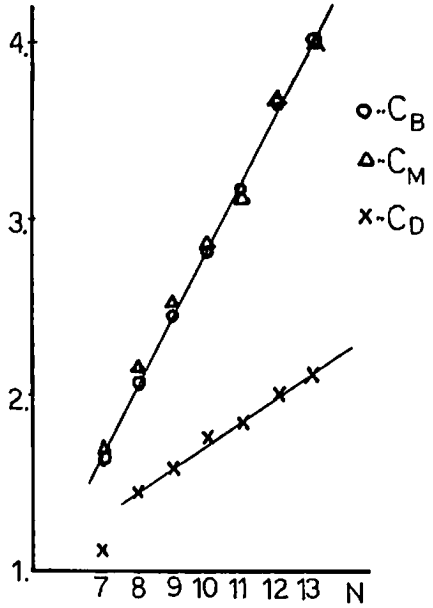
As is expected, the number of attractors increase as $\exp(\alpha \times N)$. The basin, measure and dynamical complexities increase as $a_B \times N$ + const., $a_M \times N$ + const., and $a_D \times N$ + const. where a's are some constants. See Table I for α, the values of the constants, and the types of basic oscillators.

In one type in class-2 CA such as the Rules 132, 108, and 50, a_B and a_M take almost the same values, while a_D takes a smaller value. In another type in CA, the "all-0" attractor is stable against a single flip-flop and there is a gate from other

(a)



(b)



(c)

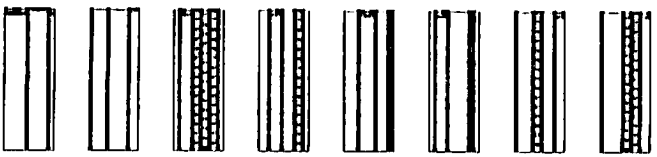


Fig. 3. Number of attractors M (a), C_B, C_M, and C_D (b) as a function of size N for the

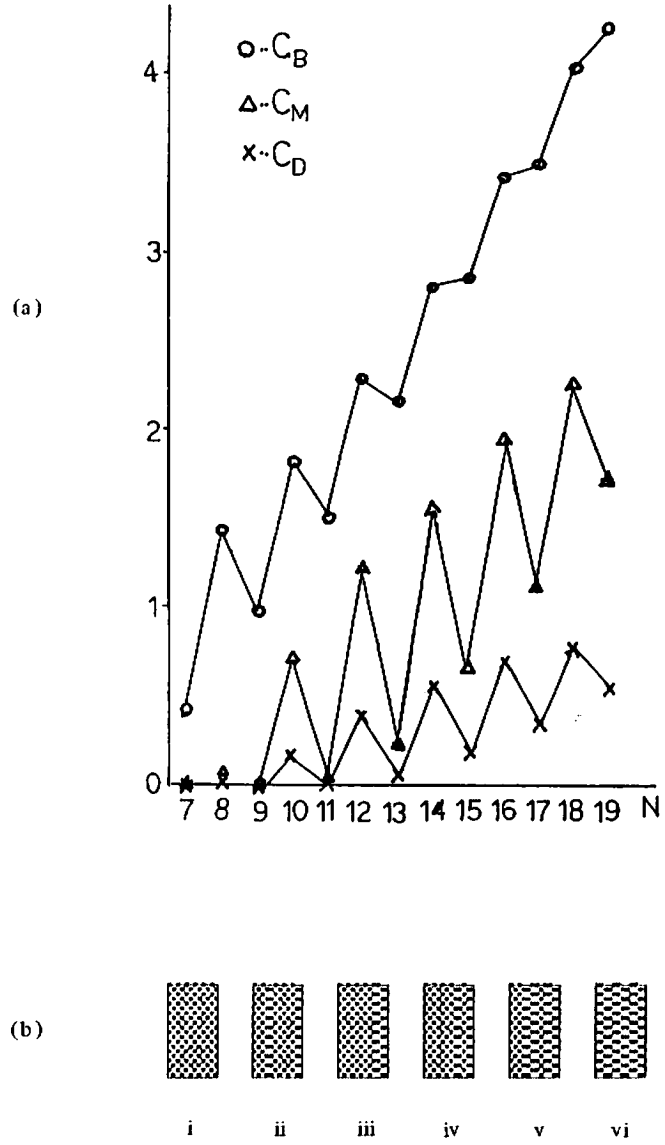


Fig. 4. C_B, C_M, and C_D (a) as a function of size N for Rule 50. Examples of attractors for $N = 12$ (b) (i-vi). In the paper, O is for C_B, Δ for C_M, and x for C_D.
The volume of the basin b_i is
(i) 15.6% (ii) 15.2% (iii) 12.3% (iv) 16.9% (max.) (v) 15.8%, and (vi) 0.1%.
The attractors with the basin volume larger than 3% are restricted to (i)–(v). The basin volume for all-0 is minimum (0.05%) and (vi) is the minimum except the all-0 attractor.

Table I. Properties of attractors for class-2 CA. $\beta = \alpha/\ln 2$, i.e., $M \sim 2^{\beta N}$.

Rule	β	a_B	a_M	a_D	Basic Oscillator and period
132	.5	1.3	1.3	0.8	010 ———1
108	.6	.4	.4	.13	010, 0110, ———1 101 ⟷ 111 ———2
104	.3	.05	0	0	0110 ———1 fall into all-0 by noise
tot24	.3	.08	0	0	01110 ———1 fall into all-0 by noise
50	.5	3.0	3.0	.6	0110 ⟷ 1001 in 010101-structure as kinks; period = 2

attractors to the all-0 state. Thus, the measure and dynamical complexities vanish, because all the other attractors fall into the attractor "all-0" by the effect of a noise. Thus $q_i = 1$ for the all-0 attractor and 0 for other attractors. The Rule 104 and totalistic 24 belong to this type.

The complexity of class-2 CA seems to be classified into the above two cases.

The class 2 behavior is understood by the superposition of local oscillators. If a local oscillator has a period t and spatial range r, the number of attractor is roughly given by $(t+1)^{N/r}$, since there are $(t+1)$ possibilities in each r region (put the oscillator or not and put it with which phase of oscillation). This argument is easily extended to the case where there are more than one type of local oscillators. The linear increase of complexity is explained in the same way.

The volume of each attractor changes as size in the following way. The ratio of the volume of the basin for the attractor "all-0" decreases as size and the basin volumes for the attractors with more oscillators increase successively as the size.

In class-2 CA, the increase of dynamical complexity is much smaller, which means that the transition among attractors is organized. An example of the transition matrix P_{ij} is shown in Table II. In the example, the possible change of the number of oscillators by a single flip is only ± 1. In the class-2 CA studied here, the transition is regular. Since an attractor in class-2 CA can be regarded as a superposition of local oscillators, single flip-flop cannot affect the global behavior, and the transition is rather limited.

Some attractors are "repulsive" by the noise, as can be seen in Table II (for such attractor a_i, P_{ki} is zero for all k).

If a local oscillator exists as a kink in a zigzag structure (see Fig. 4b), there appears a difference in the parity of the size N (see Fig. 4a). As the system size

Table II. Transition matrix P_{ij} for Rule 132 with $N = 9$ with all attractors $a_1 - a_{12}$. Here the number change by the transition is only ± 1.

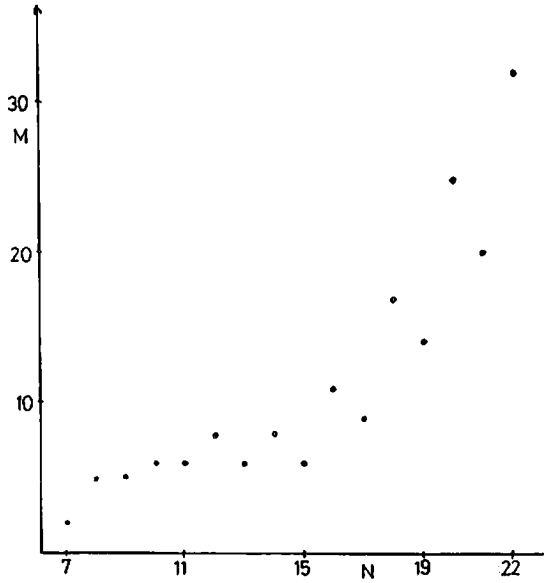
	1	2	3	4	5	6	7	8	9	10	11
1	0	1	0	0	0	0	0	0	0	0	0
2	$\frac{3}{9}$	0	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	0	0	0	0	0	0
3	0	$\frac{5}{9}$	0	0	0	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	0	0	0
4	0	$\frac{6}{9}$	0	0	0	0	$\frac{1}{9}$	$\frac{1}{9}$	0	0	0
5	0	$\frac{6}{9}$	0	0	0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	0	0	0
6	0	0	$\frac{4}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	0	0	0	0	$\frac{2}{9}$	0
7	0	0	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	0	0	0	0	$\frac{1}{9}$	0
8	0	0	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	0	0	0	0	$\frac{1}{9}$	0
9	0	0	0	1	0	0	0	0	0	0	0
10	0	0	0	0	0	$\frac{4}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	0	0
11	1	0	0	0	0	0	0	0	0	0	0

attractors: a_1 = 000000000 a_2 = 000000001 a_3 = 000000101
a_4 = 000000001 a_5 = 000010001 a_6 = 000010101
a_7 = 000100101 a_8 = 000101001 a_9 = 001001001
a_{10} = 001010101 a_{11} = 111111111
(all attractors are fixed points)

5. Class-3 CA

The class 3 behavior of CA is characterized by triangles with various sizes. Here we have investigated Rules 22, 18, 54, 146, and totalistic 12 and 22. The number of attractors and complexities as a function of system size are shown in Figs. 5-7, with some patterns of typical attractors (see also Fig. 8). See Ref. 13 for the Rule 54. Though the behavior is very complicated, the following points are common in class-3 CA.

(a)



(b)

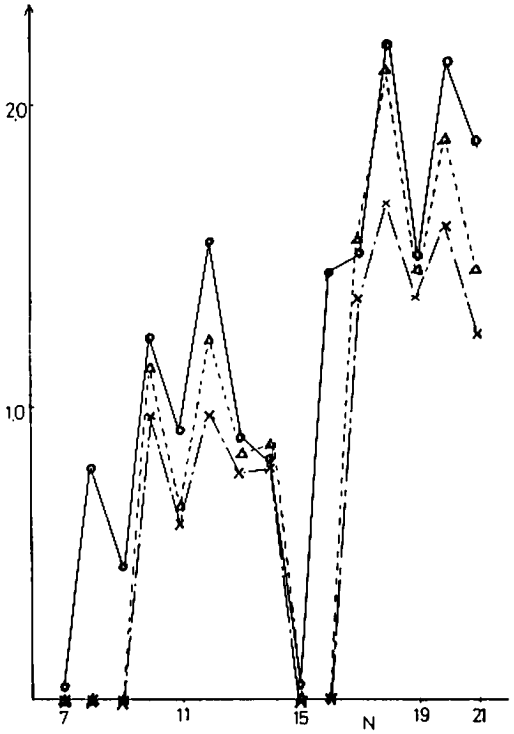


Fig. 5. Number of attractors M (a), C_B, C_M, and C_D (b) as a function of size N for the

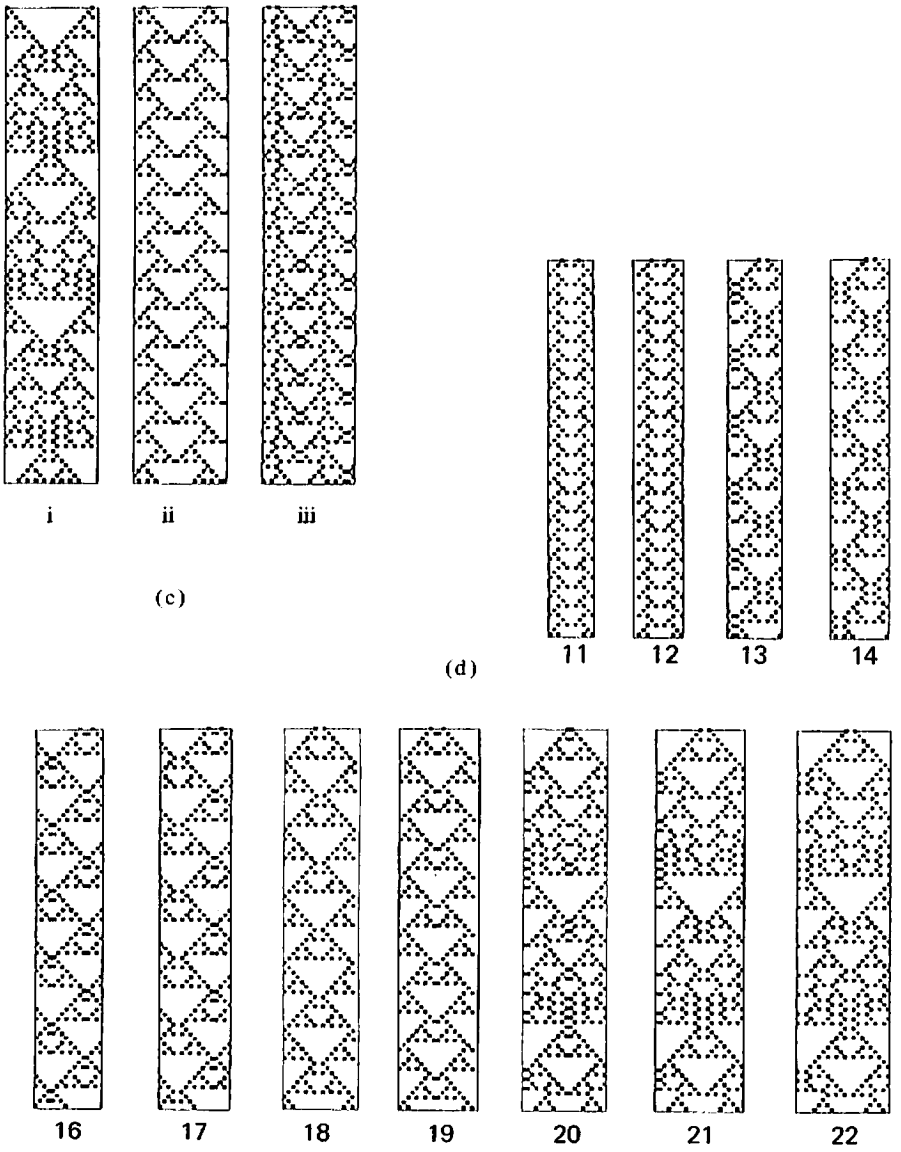
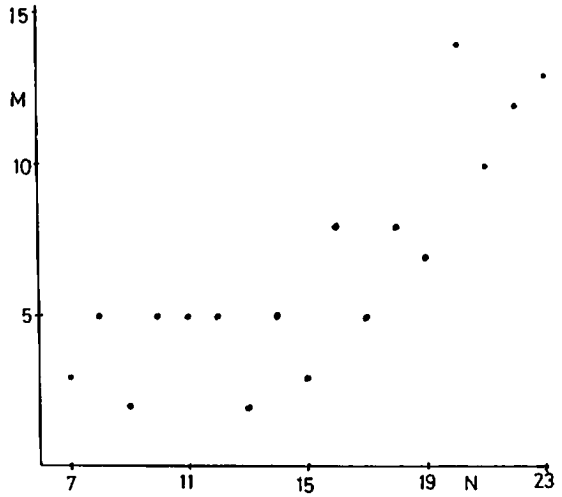


Fig. 5. (c) Examples of typical attractors for $N = 21$. The volume of the basin b_i is all-0:35% (i) 17.3% (ii) 11.0% (iii) 10.5%
The attractors with the basin volume larger than 10% are restricted to (i) – (iii).
In (d), the attractor with the largest basin volume (or second largest if the one with the largest volume is all-0) is shown for $10 < N < 23$.

The basin volume b_i for the depicted attractor is
70% ($N = 11$), 35% ($N = 12$), 70% ($N = 13$), 30% ($N = 14$),
5% ($N = 16$), 28% ($N = 17$), 12% ($N = 18$), 27% ($N = 19$),
18% ($N = 20$), 40% ($N = 21$), 17% ($N = 22$).

(a)



(b)

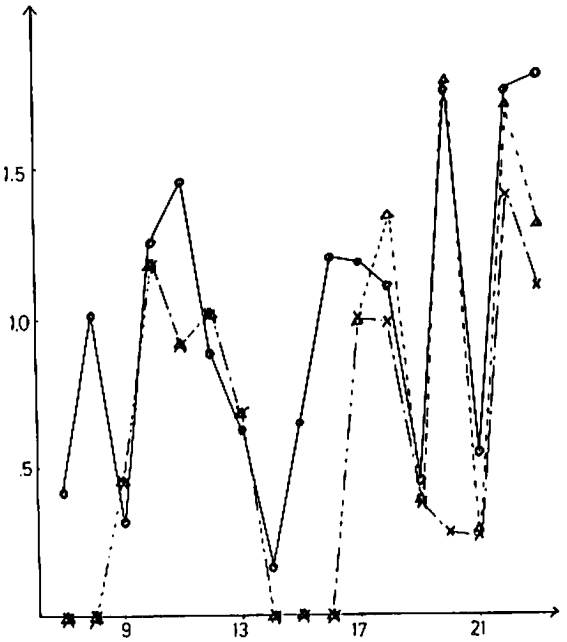


Fig. 6. Numbers of attractors M (a), C_B, C_M, and C_D (b) as a function of size N for the Rule 22.

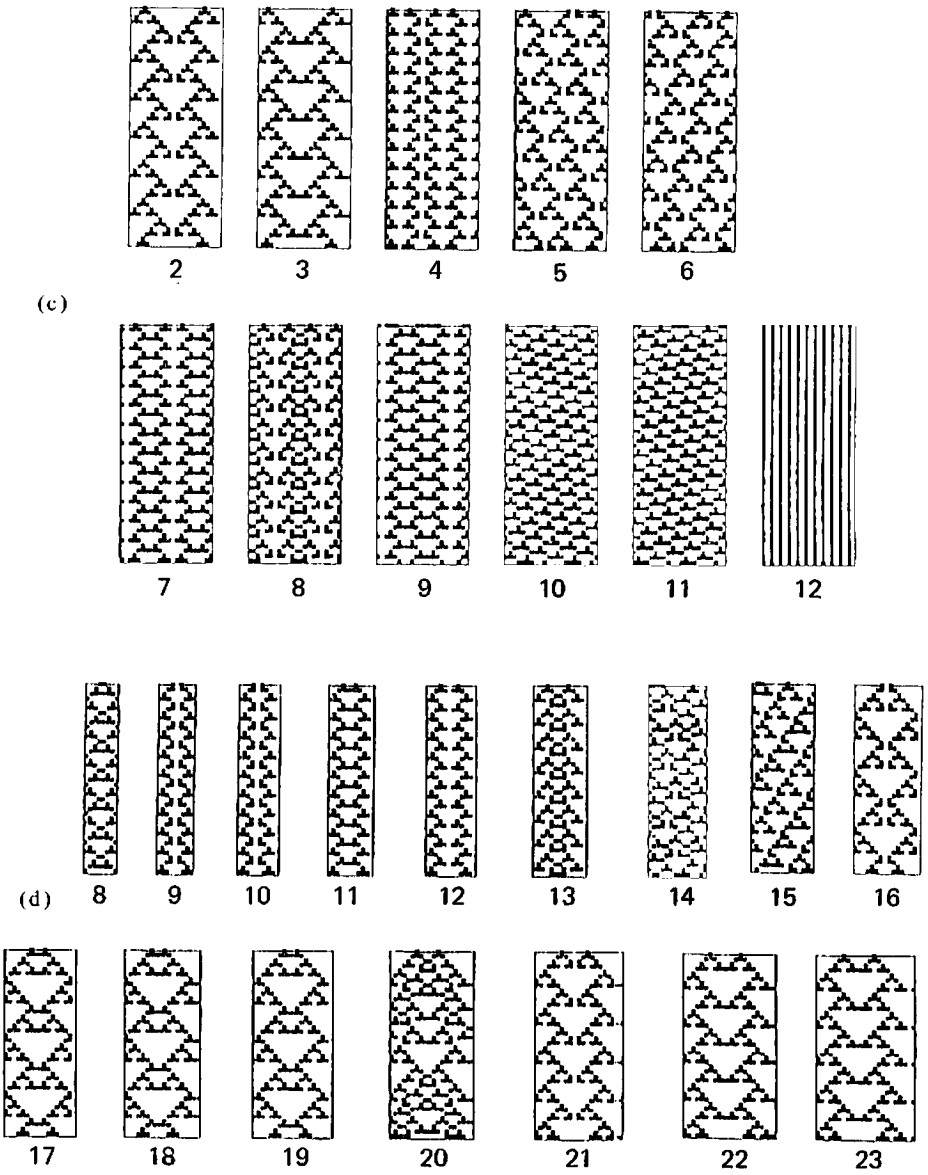


Fig. 6. (c) All attractors for $N = 22$. The volume of the basin b_i is (1) all-0: 0.4% (2) 30.3% (3) 39.4% (4) 6.5% (5) 9.6% (6) 9.6% (7) 0.9% (8) 0.8% (9) 2.3% (10) 0.01% (11) 0.01% (12) 5×10^{-5}%.
In (d), the attractor with the largest basin volume (or second largest if the one with the largest volume is all-0) is shown for $7 < N < 24$.

The basin volume b_i for the depicted attractor is
21.8% ($N = 8$), 89.6% ($N = 9$), 38.1% ($N = 10$), 35.9% ($N = 11$),
65.6% ($N = 12$), 32.4% ($N = 13$), 3.7% ($N = 14$), 10.4% ($N = 15$),
38.7% ($N = 16$), 40.0% ($N = 17$), 51.6% ($N = 18$), 91.3% ($N = 19$)

(e)

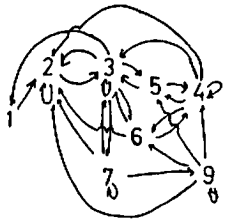


Fig. 6 (e) shows the transition loop by P_{ij} among attractors. Only the attractors with non-vanishing q_i are shown. The arrows indicate the possible transition between attractors by a single flip-flop noise. See Table III for P_{ij}.

(a)

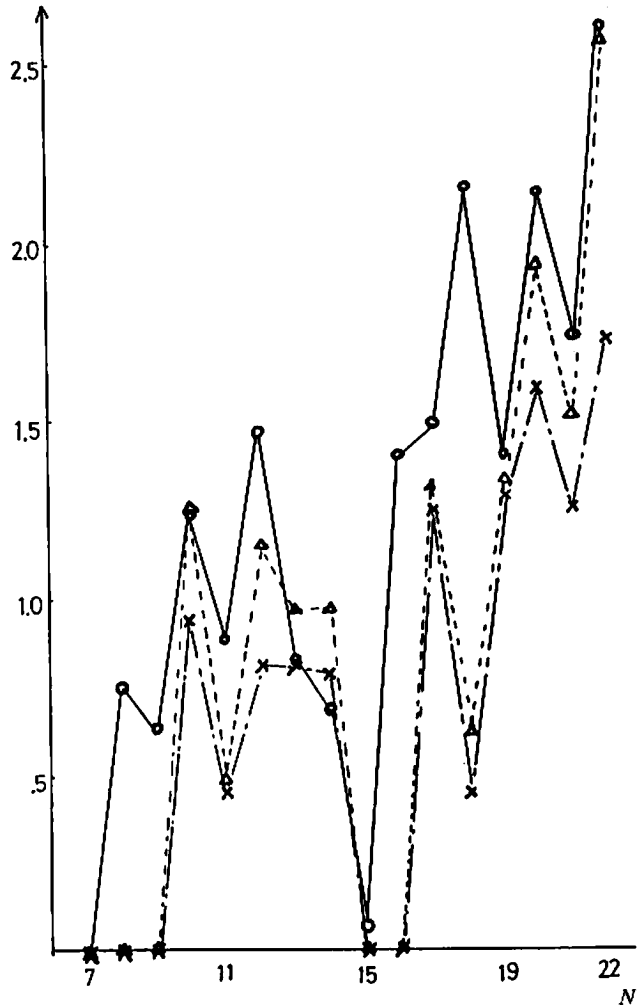


Fig. 7. (a) C_M, C_B, and C_D as a function of size N for the Rule 18. The number of attractors

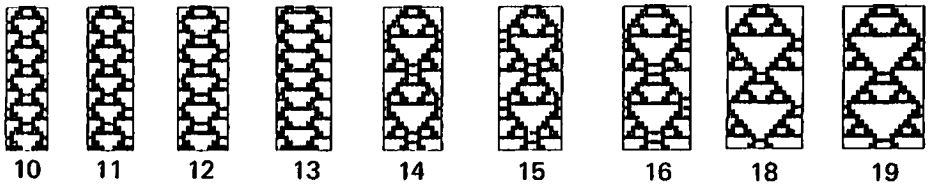


Fig. 8. The attractor with the largest basin volume (or second largest if one with the largest volume is all-0) is shown for $9 < N < 20$ for the totalistic rule 12.

The basin volume b_i for the depicted attractor is
19% ($N = 10$), 63% ($N = 11$), 39% ($N = 12$), 67% ($N = 13$),
38% ($N = 14$), 76% ($N = 15$), 36% ($N = 16$), 23% ($N = 18$),
47% ($N = 19$).

The attractor for $N = 17$ is omitted since it has only 1% volume.

(i) Number of attractors changes irregularly as the system size N. The increase in the size is at most bounded by some power of the system size N.

(ii) The attractors which have a large region of basins are the ones with triangle structures and the all-0 attractor. Among the attractors with triangle structures, the attractor with a larger size of triangles has a larger size of basin of attractions (see Figs. 5(c) and 6(c)). In Fig. 9, the basin volume is shown as a function of the size of the largest triangle in the attractor for the totalistic rule 22. In the example the basin volume is roughly proportional to the (size of the largest triangle

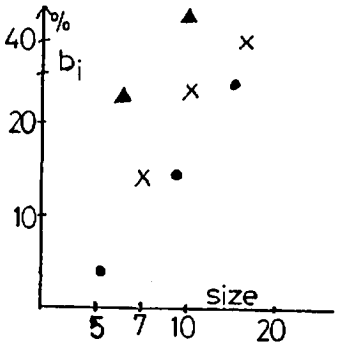


Fig. 9. Basin volume as a size of the largest triangle in the attractor for the totalistic rule 22. The size denotes the longest sequence of 0's in the attractor. Only the typical attractors with class-3-like triangle structures are chosen. • is for $N = 16$, ▲ for $N = 17$, and x for $N = 19$.

in the attractor).[1,5] For other rules, the similar behavior is observed, but the power fit is not so good (at least for small size N).

(iii) The attractors with the largest basin volume except the "all-0" are shown in Figs. 5(d), 6(d), and 8. The characteristic feature of such attractors is that they start from a simple seed. For example the attractors are formed from the seeds 101, 100001, or 11 for Rule 146, 11, 1111, or 11111 for totalistic rule 12, 1001, 10001, 100001, and 1000001 for Rule 22, and so on. The simplest seed which has a recurrence gives the attractor with the largest basin volume. This kind of choice is quite analogous to the choice of eigenfunction in the Schrodinger equation where the interference of phase is important. In our problem, the possible configuration of triangles is determined by a kind of interference effect.

The analogy with the wave mechanics also implies that the most probable attractor is related to the ground state of wavefunction in the sense that the number of nodes is smallest (in other words, it has only a small number of "seeds", i.e., the structure of 00000***0000). The attractor with smaller basin volume may be related to the excited state in the sense that it has more nodes.

(iv) The irregular behavior as a size change seems to depend on some number theoretic properties of the size and rules. For example, there occurs singular behavior around at $N = 2^k - i$ ($i = 0$, or 1, or -1 which depends on the rule). The number of attractors decreases at some values near $N = 2^k$ and all-0 attractor has a large measure. For Rule 146, the ratio for the attractor with all-0 has 99% measure at $N = 15$ ($= 2^4 - 1$). The complexities C_M and C_D vanish at $N = 15$ and 16 (see Fig. 5(b)). For Rule 22, C_M and C_D vanish at $N = 14$, 15, and 16 (see Fig. 6(b)), since all the attractors fall down into "all-0" attractor by the noise.

At these sizes, the measure for all-0 is close to 1, since the patterns from simple seeds fall into all-0 and the attractor with the large triangles cannot exist (see $N = 14$ and 15 of Fig. 6(d)).

(v) The complexities also change irregularly as the system size. They seem to increase slowly as the system size. Generally speaking, C_D is not so small compared with C_M. That is, the mutual information $C_M - C_D$ is small compared with the cases for the CA in other classes. Thus, the connectivity among attractors by a low noise is random.

An example of P_{ij} is shown in Table III, for Rule 22 with $N = 22$. After the transients, the measure for the attractor a_8, a_{10}, a_{11}, and a_{12} goes to zero and the transitions by small noise occur among attractors a_1, a_2, . . . , a_7 and a_9. The transition diagram is shown in Table III. Here the most probable loop of the transition is $a_2 \to a_3 \to a_2$, while the second most probable loop is $a_1 \to a_2 \to a_3 \to a_1$, and the next is $a_3 \to a_7 \to a_4 \to a_3$, and so on. We note that various transition loops are formed by the addition of a noise (Fig. 6(e)).

Table III. Transition matrix P_{ij} for Rule 22 with $N = 22$. The attractors are shown in Fig. 6.

	1	2	3	4	5	6	7	8	9	10	11	12
1	0	1	0	0	0	0	0	0	0	0	0	0
2	0	$\frac{5}{11}$	$\frac{6}{11}$	0	0	0	0	0	0	0	0	0
3	$\frac{1}{11}$	$\frac{2}{11}$	$\frac{6}{11}$	0	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{11}$	0	0	0	0	0
4	0	$\frac{1}{11}$	$\frac{4}{11}$	$\frac{3}{11}$	$\frac{1}{11}$	$\frac{1}{11}$	0	0	$\frac{1}{11}$	0	0	0
5	0	$\frac{6}{11}$	$\frac{4}{11}$	$\frac{1}{11}$	0	0	0	0	0	0	0	0
6	0	$\frac{6}{11}$	$\frac{4}{11}$	0	0	0	0	0	0	0	0	0
7	0	$\frac{1}{11}$	$\frac{2}{11}$	$\frac{4}{11}$	0	0	$\frac{2}{11}$	0	$\frac{2}{11}$	0	0	0
8	0	$\frac{4}{11}$	$\frac{4}{11}$	0	$\frac{1}{11}$	$\frac{1}{11}$	0	$\frac{1}{11}$	0	0	0	0
9	0	$\frac{3}{11}$	$\frac{9}{22}$	$\frac{1}{11}$	$\frac{3}{22}$	$\frac{1}{22}$	0	0	$\frac{1}{22}$	0	0	0
10	0	0	$\frac{7}{11}$	0	$\frac{3}{11}$	$\frac{1}{11}$	0	0	0	0	0	0
11	0	0	$\frac{7}{11}$	0	$\frac{1}{11}$	$\frac{3}{11}$	0	0	0	0	0	0
12	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	0	0	0	0	0	0	0

(vi) Another interesting quantity is the period of an attractor. The period of the main attractor increases slowly with an irregular change. The numerical observation shows: The main attractor starts from a small seed and grows with a constant speed till it comes back to the seed pattern after $O(N)$ steps (the position of the seed may be different from the original position). Thus, the increase of the period seems to be bounded at most by some power of the size, which is consistent with the numerical results. The longest period among all the attractors increases faster. It seems to be bounded by some power, though it shows a rapid jump at some values of N.

6. Class-4 CA

The class 4 behavior for CA characterized by Stephen Wolfram is long-time

transients and the existence of local oscillators and local propagating patterns and the sensitive dependence of patterns on the initial configurations. We have investigated here the totalistic-52, totalistic 20 and models S1-S2 (see Appendix). The results for other models with soliton-like excitations will be shown in the next section. The characteristic features for the basin structure of the class 4 systems may be summarized as follows:

(i) The number of attractors increase exponentially, though the increase is rather irregular. The pattern of attractor which has a large region of basins changes as size, though "all-0" or "all-1" has a large basin of attractions in many rules. The pattern of the attractor with the large basin of attraction except the all-0 or all-1 changes as the size (for some size, it is global and for other size it is local), which is a main difference from the other classes. As N is increased, attractors of essentially new type appear successively.

(ii) The basin complexity C_B takes a comparatively large value, which changes irregularly as size. The measure complexity C_M is much smaller than C_B, since the probability measure (by a noise) for "all-0" (or "all-1") is much larger than the ratio for the basin of attractions to such states. The dynamical complexity is much smaller, which means that the mutual information is rather large. In other words, the transition between attractors by noise is regularly structurized.

In the following we show three typical examples for class-4 CA.

(1) totalistic rule 52: (see Fig. 10).

The rule is symmetric about the transformation $0\langle-\rangle1$. The main attractors are all-0 and all-1 which have the same basin volume b_i and probability q_i by the symmetry. The ratio of the basin volume for the all-0 (or all-1) attractor is shown in Fig. 10(d). Both global and local attractors coexist, some of which are shown with the basin volume (Fig. 10(c)). The number of attractors increases exponentially with a large increasing rate.

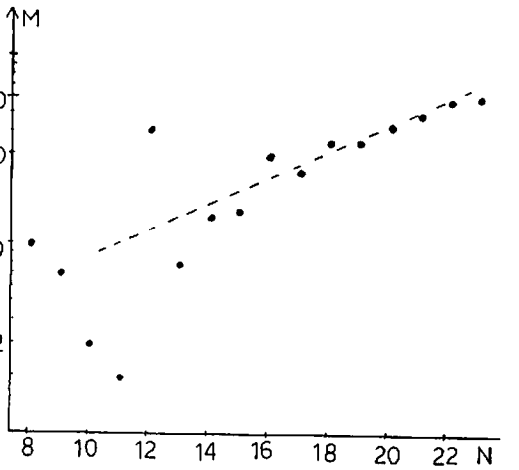
By the noise, the attractors fall into all-1 or all-0 attractor for almost all N. Since the probability for each attractor is 50%, the measure complexity takes $\ln 2$. The dynamical complexity vanishes for almost all N, since the all-1 and all-0 attractors are disconnected by a single flip-flop (see Fig. 10(e) for the transition loop by P_{ij}.)

(2) totalistic rule 20 (see Fig. 11):

The number of attractors increase rather irregularly around $10 < M < 20$, while it has a rapid jump at $N = 12$ and 16 ($M > 50$). The all-0 attractor has a large basin volume which is shown in Fig. 11(c). For $N > 14$, the state falls down into the all-0 attractor and the measure and dynamical complexities vanish.

In Fig. 11 (b), the main attractors for $N = 21$ are shown. A local irregular propagating pattern is remarkable. In Fig. 11 (c), the attractor pattern with the largest basin volume (or the second largest in the case that all-0 has the largest

(a)



(b)

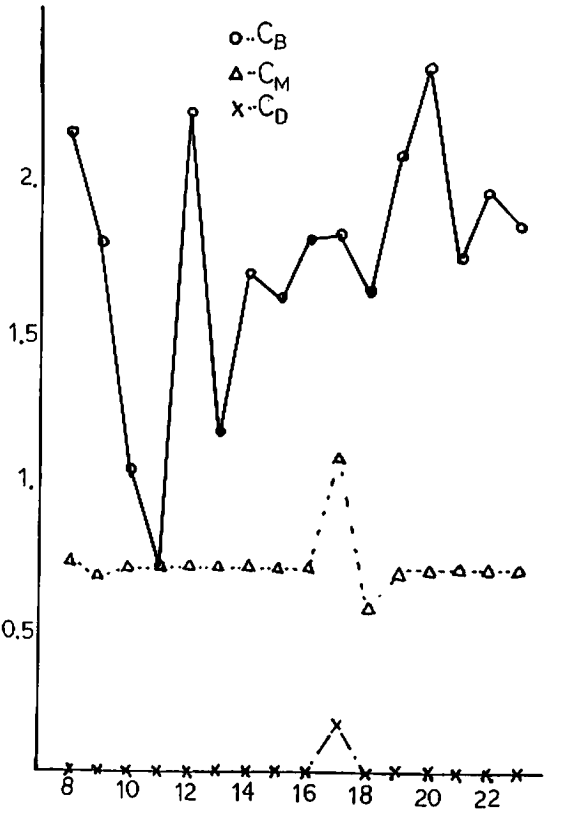
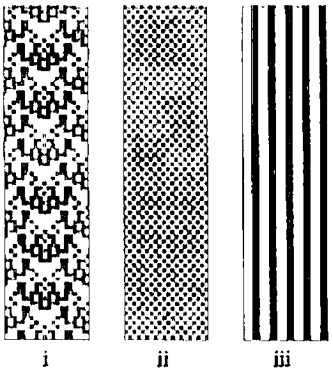
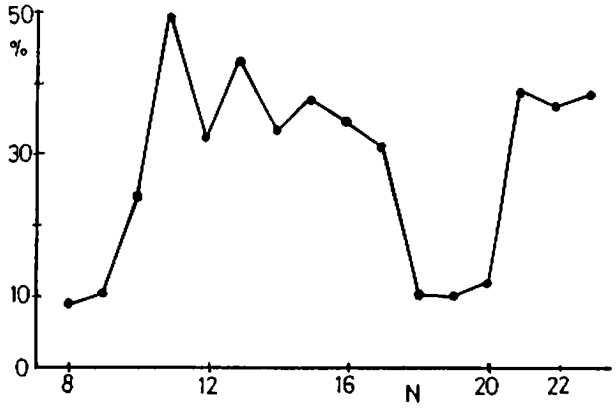


Fig. 10. Number of attractors M (a), C_B, C_M, and C_D (b) as a function of size N for the

(c)



(d)



(e)

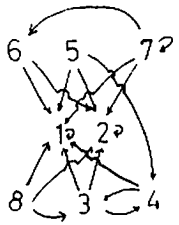


Fig. 10(c) Examples of typical attractors are shown for $N = 20$. The volume of the basin b_i is all-0: 11.9%, all-1 : 11.9%, (i) 23.3% (the attractor obtained by the $0(-)1$ transformation has the same basin volume), (ii) 4.8% (iii) 7.3%.
The attractors with the basin volume larger than 3% are restricted to (i)–(v).
(d) shows the basin volume all all-0 (or all-1) as a function of N.
(e) shows the transition loop by P_{ij} for totalistic rule 52 with $N = 13$. The arrows indicate the possible transition between attractors by a single flip-flop noise. Only the attractors

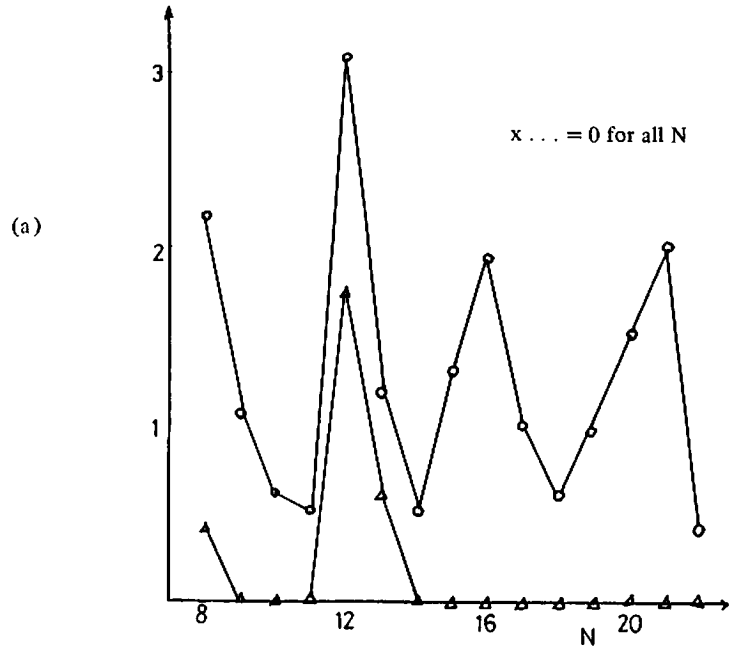


Fig. 11. C_B, C_M, and C_D (a) as a function of size N for the totalistic rule 20. Examples of typical attractors are shown for $N = 21$ (b).
The volume of the basin b_i is

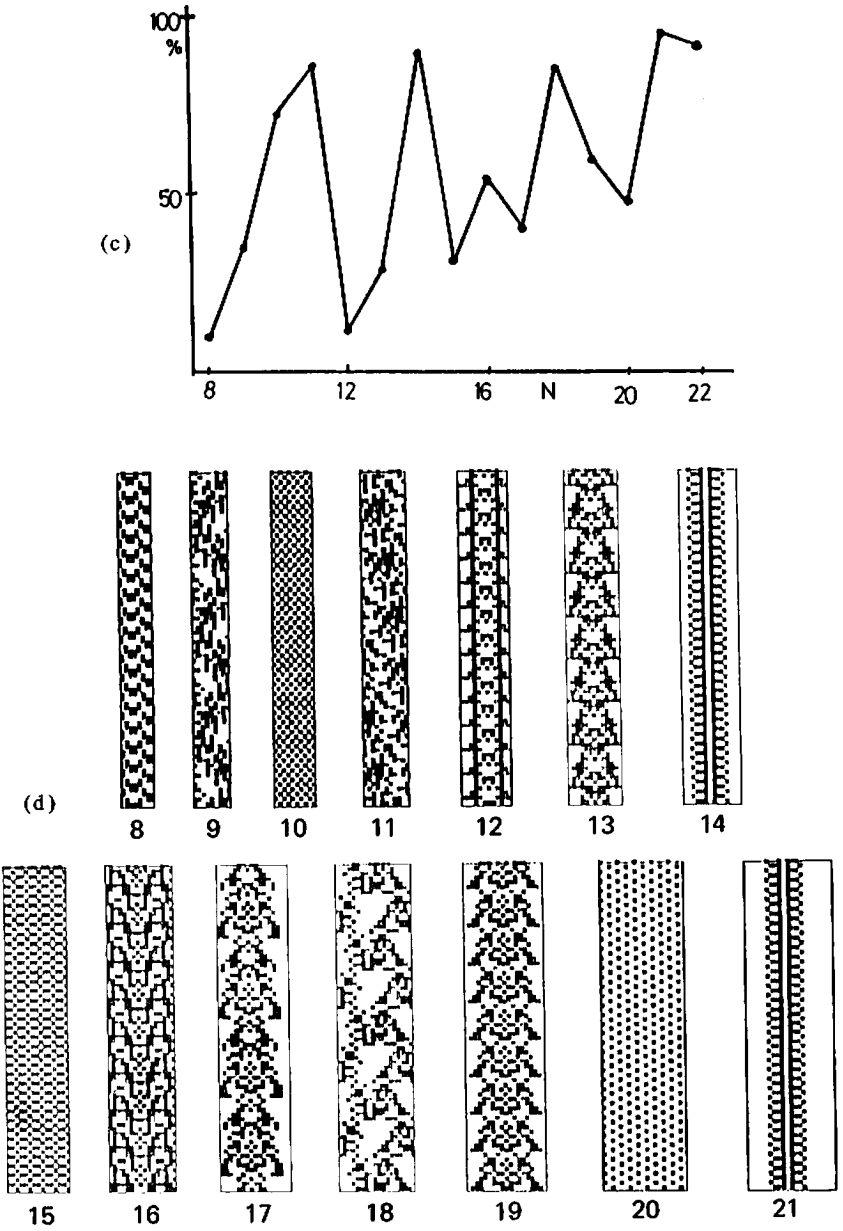


Fig. 11(c) shows the basin volume for all-0 as a function of N.
In (d), the attractor with the largest basin volume (or second largest if the one with the largest volume is all-0) is shown for $7 < N < 22$.
The basin volume b_i for the depicted attractor is

18.8%	$(N=8)$,	49.2%	$(N=9)$,	27.5%	$(N=10)$,	6.9%	$(N=11)$,
32.5%	$(N=12)$,	56.9%	$(N=13)$,	5.5%	$(N=14)$,	47.8%	$(N=15)$,
17.7%	$(N=16)$,	46.0%	$(N=17)$,	7.1%	$(N=18)$,	33.1%	$(N=19)$,

basin volume) is shown. We note that such attractors are global and irregular, and the patterns change their structures as the size in an irregular way, which is remarkably different from the class-3 case.

(3) model S1 $\{0001101\} - \{01011\}$ (see Fig. 12):

The number of attractors increases exponentially. The all-0 and all-1 attractors have large basin volumes. The measure entropy is much smaller than the basin volume, since the measure of all-0 (or all-1) attractor increases by the flip-flop. The dynamical entropy is much smaller and the transition by the flip-flop is regular.

(4) model S2 $\{0011101\} - \{01011\}$ (figures omitted) (for $N < 19$)

The following features are observed: (a) exponential increase of the number of attractors; (b) C_B takes about $1 \sim 1.5$ for $12 < N < 19$; (c) C_M is very small (for most N); (d) $C_D = 0$; (e) all-0 attractor has a large measure.

7. CA with Soliton-like Excitations

As is shown in the Appendix and Ref. 12, a class of CA with solitons shows an interesting behavior such as the integrable-like behavior or soliton turbulence. Here the following two typical examples will be investigated; one is for the integrable-like behavior, and the other for the turbulent-like behavior:

(1) Integrable-like behavior; model S3; $\{0000011\} - \{11000\}$

For this class, the basins for the state of superpositions of "solitons" go larger as the system size is increased. The important difference between this type of behavior and the usual integrable systems studied in the soliton theory is that our system is integrable only after the transients have decayed out. Thus, our system can be regarded as an "integrable system on an attractor".

For small N, however, the basin volume for the superposition of solitons is not necessarily large. Some bound states of solitons have large basin volumes as can be seen in Figs. 13, where the main attractors are shown for $N = 19$, 20, and 21. As the size is increased, however, the ratio for the superposition of soliton-states seems to increase. We have studied some attractors for $N = 30$ or 45 by choosing some samples of initial configurations. More than 80% of such initial configurations are attracted into the ensembles of solitons.

The number of attractors increase exponentially. The dynamical entropy is much smaller than the measure entropy, which shows that the transition among attractors is well-organized.

(2) Soliton turbulence; model S4; $\{0001011\} - \{10011\}$

Some CA show the soliton turbulence[12]. A typical pattern is shown in Fig. 15 (d), where the sensitivity on the phase of collisions of 1101-"solitons" make the turbulent-like phenomena. For the attractors for a CA with a small size, however, it is hard to find such patterns. The main attractors are (A) superposition of

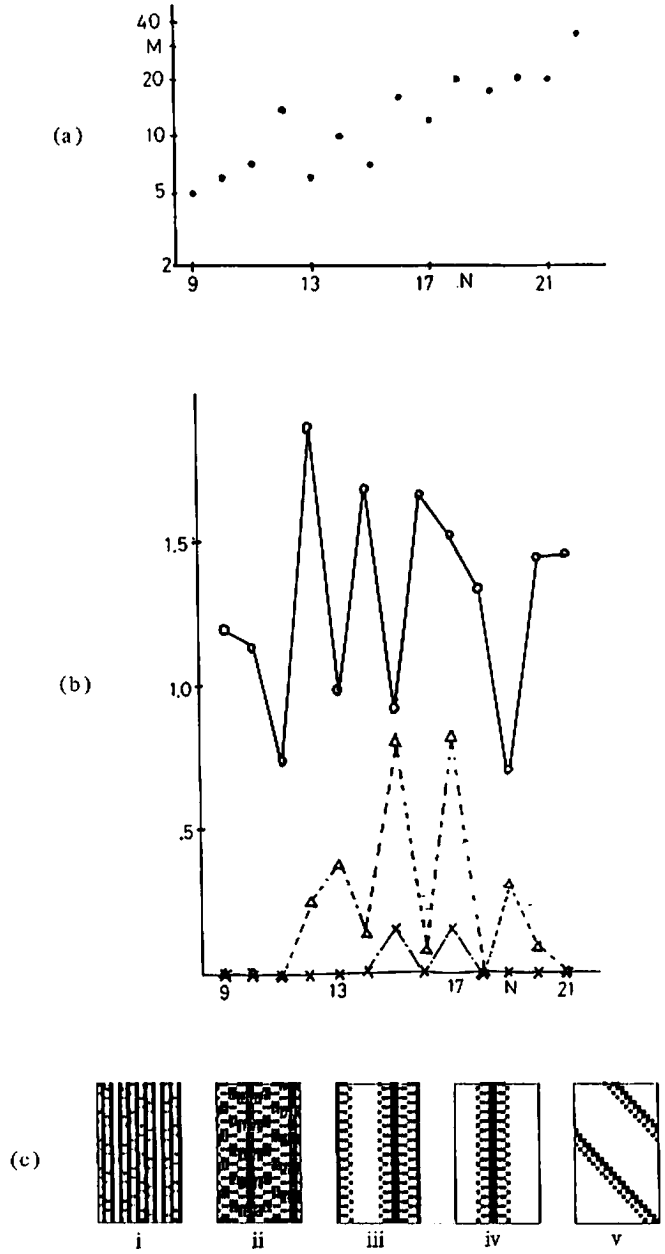
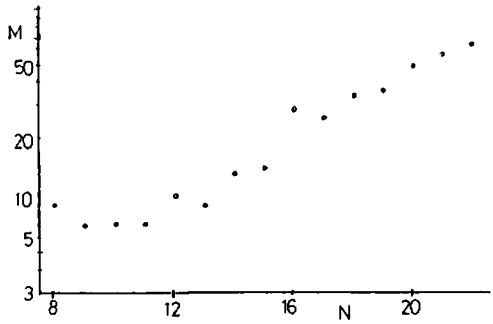
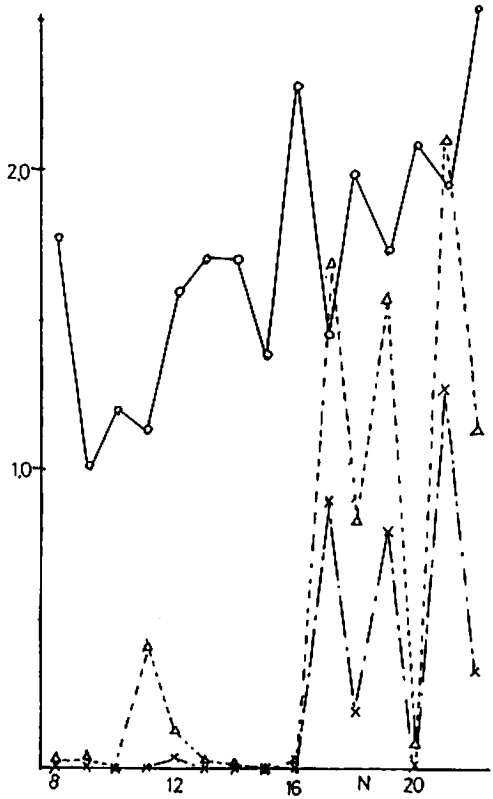


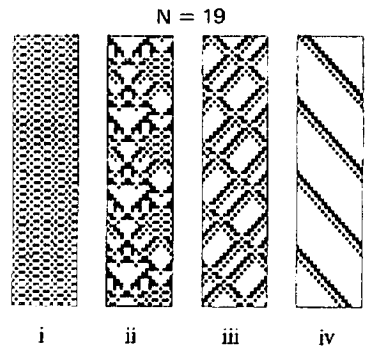
Fig. 12. Number of attractors M (a), C_B, C_M, and C_D (b) as a function of size N for the model S1. Examples of typical attractors for $N = 20$ (c).
The volume of the basin b_i is all-0: 39.5%; all-1: 30.3%; (i) 15.4% (ii) 10.4% (iii) 2.0% (iv) 1.1% (v) 0.3% (and the same volume for the attractor with the converse direction). The attractors with the basin volume larger than 0.3% are shown.

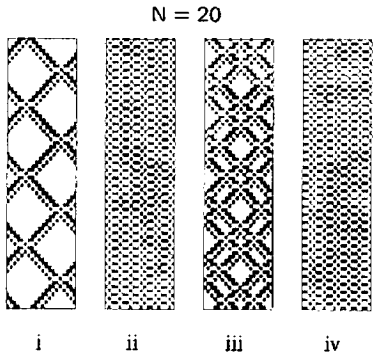


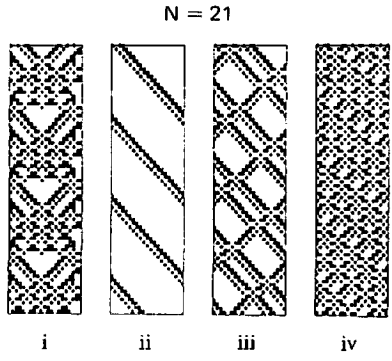
(a)



(b)







(c)

Fig. 13. Number of attractors M (a), C_B, C_M, and C_D (b) as a function of size N for the model S3.
In (c), the main attractors are shown
$N = 19$: (i) 52.2% (ii) 11.0% (iii) 5.3% (iv) 3.4%
$N = 20$: (i) 21.6% (ii) 19.4% (iii) 16.6% (iv) 3.5%

of some soliton-states (B) global periodic patterns (see Fig. 14). The latter patterns have the larger region of basins. Some collisions increase the number of solitons till they reach the global patterns (the latter attractors) with small periods and with more than 50% of 1's.

We have performed some simulations for $N = 30$ and 40 by choosing some initial configurations. Still, the attractor with the typical turbulent patterns have small measures (about 10%). The turbulent pattern is seen as the transients before the CA falls into the attractor in the type (B) above (see Fig. 14(d)). The transient time increases rapidly as the size. This observation may imply that the turbulent-like patterns in some CA may be characterized as the transients, the time for which diverges as the system size.

Here, (1) exponential growth of number of attractors (2) large C_B and very small C_D are observed again, which are typical in class-4 CA.

8. Discussion and Future Problems

We have investigated here the storage of information in the attractors of CA and the complexity of networks among attractors connected by a noise.

One important question is "what is the generic behavior for the CA with large N?" In class-1 CA, the attractor is trivial and shows no essential change as the size. In class-2 CA, the attractors are local. Thus the attractors at large N is essentially the superposition of the attractors of small size. In class-3 CA, the attractors are global and the period of the main attractors increases. The generic behavior at large N, however, is characterized by the attractors. The main attractors for a large size can be characterized by those for a small size, since there is self-similarity and the main attractor is generated by a simple seed.

In class-4 CA, however, it may be hard to predict the behavior of the CA with a large size from the result for the attractors, since the transients before the state falls into the attractors increase rapidly as size, as can be seen in the case for soliton-turbulence, where the turbulent behavior (for large N) may be attributed not to the attractors but to the transients.

The memory in the class-3 CA may be used as the holographic memory in the following three points. First, the attractor includes a self-similar triangle structure. Thus, we can construct a pattern similar to the original triangle structure even if the information of some parts is lost. Secondly, the interference effect in the triangle structures discussed in Section 5 is analogous with the interference in the wave mechanics. Thirdly, the attractor's network by the transition matrix P_{ij} is global.

It will be of interest to extend our approach to basin structures in other systems.

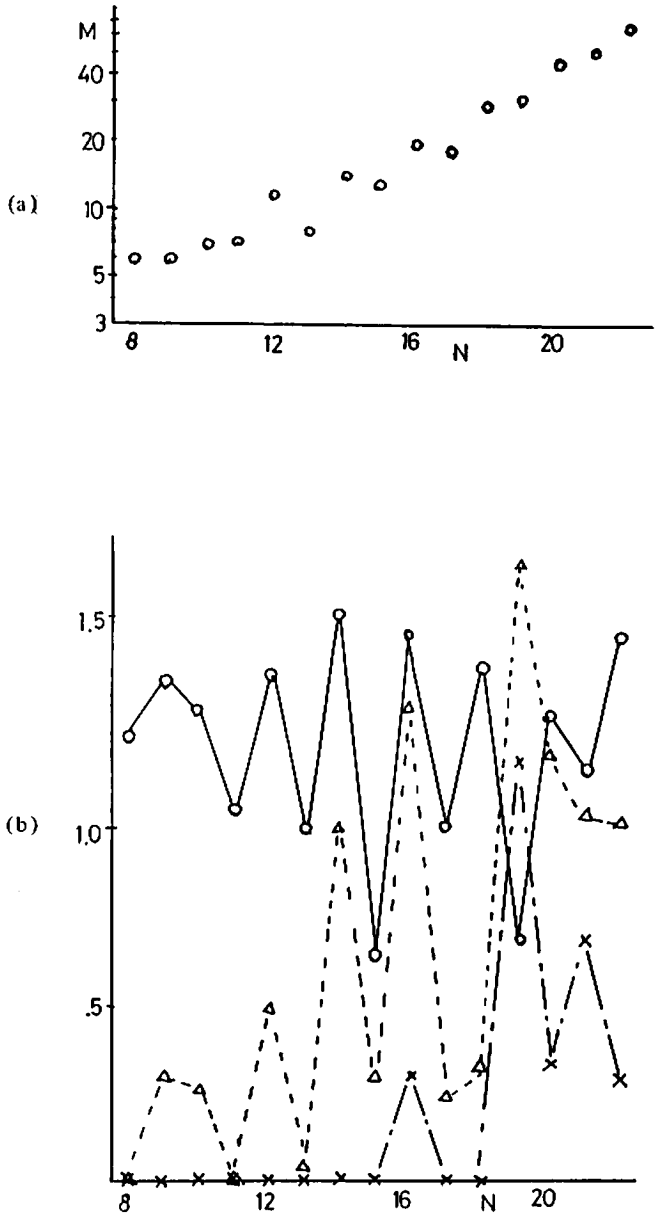


Fig. 14. Number of attractors M (a), C_B, C_M, and C_D (b) as a function of size N for the model S4.

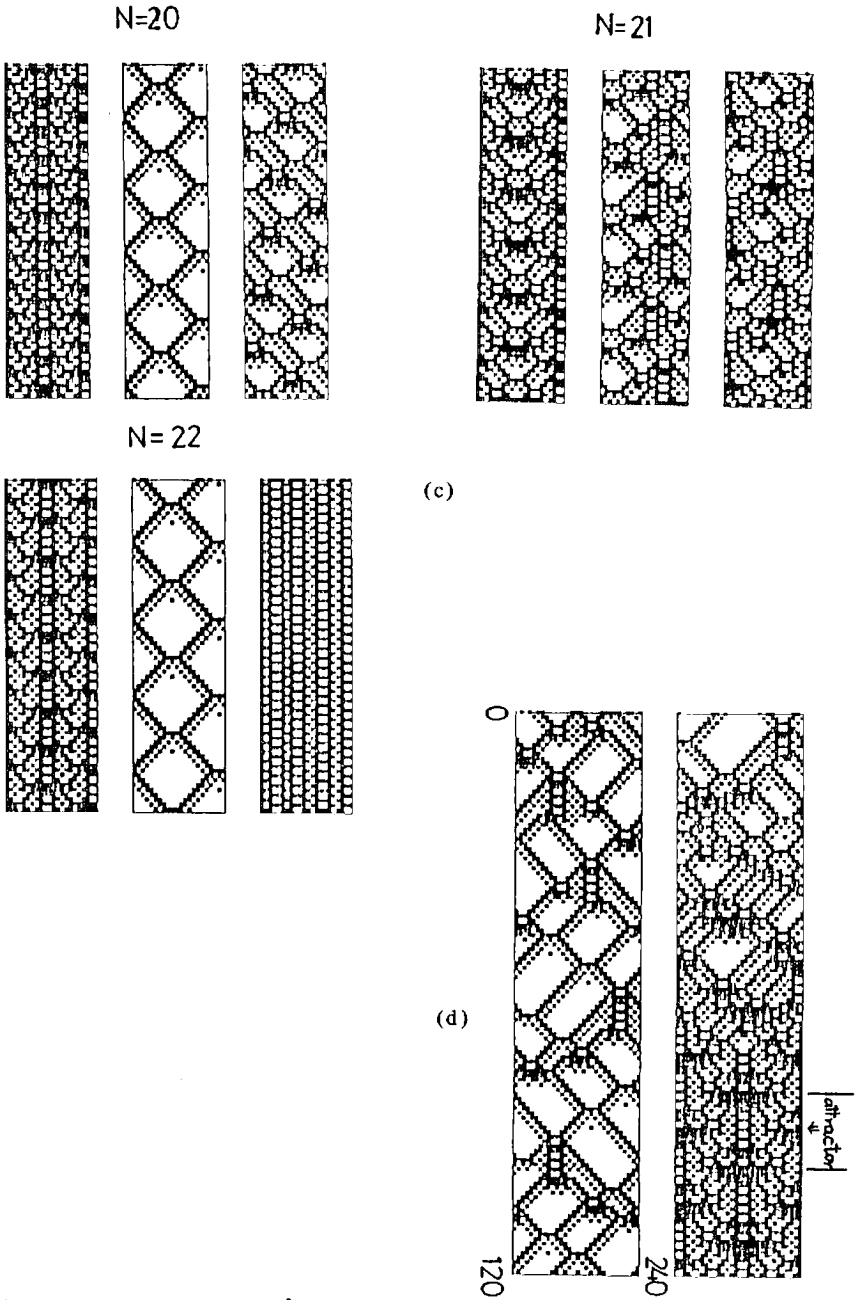


Fig. 14(c), the main attractors are shown
$N = 20$: (i) 52.6% (ii) 31.7% (iii) 3.6% (each)
$N = 21$: (i) 73.8% (ii) 6.2% (iii) 5.2%
$N = 22$: (i) 35% (ii) 13% (iii) 17.6%
In (d), an example of an evolution for $N = 30$ is shown. After the time step ~ 200, it falls into

transition by a noise can occur[14]. It is of interest to characterize the low-noise ordered states from the basin structure for the ordered states.

Another important example is a spin glass system, especially Sherington-Kirkpatrick model, where a lot of attractors (fixed points) are hierarchically organized[15].

Also, it is of importance to extend our approach to the dynamical system with continuous variables: Most dissipative systems such as low-dimensional maps[16] with small dissipation or high-dimensional maps such as coupled map lattices[17,18] can have a large number of attractors. The basin structure has recently been investigated intensively. The statistical properties such as the basin volume entropy and the jumping process among attractors by the noise can be investigated in the present paper's line.

Informational aspects in the dynamical system are important in the intelligent network system[19,20] such as the neural or immune network and some artificial intelligence systems.

In addition to the storage of information studied in this paper, the processing of information should be studied in the high-dimensional chaos or CA. In the CA with soliton-like excitations, the information can be transmitted by the "solitons", even if the state is turbulent. The information transmission is calculated by the mutual information flow, in a similar way as the coupled map lattice case[18].

Also, selective propagation of information will be of use in the intelligent system. In CA with soliton-like patterns, some specific patterns can easily propagate to other sites, while others decay out. The quantification of such selectivity will be of importance.

Acknowledgements

The author would like to thank Mr. Yukito Iba and Dr. Shinji Takesue for stimulating discussions and critical comments. He would also like to thank Dr. Yoji Aizawa, Dr. Norman Packard and Dr. Stephen Wolfram for useful discussions and critical comments. He is grateful to the Institute of Plasma Physics at Nagoya for the facility of FACOM M-200 and to the Research Institute of Fundamental Physics at Kyoto University for some financial support.

Appendix: CA with Soliton-like Excitations

In the class-4 cellular automata, the patterns which propagate with some speed is commonly seen. In Ref. 12, a class of CA with specific type of "soliton" (00101100) is investigated in detail. We impose the condition that the

00101100 should move right at the next step (i.e., 00010110). There are 2^{12} possible rules for the legal CA with 2-states and range = 2 which satisfies this type of 1011-soliton conditions.

Simulations for all these rules have been performed. The rule for the two-state CA with range 2 is coded by the 32 numbers $i_k = (0, 1)$; $(00000) \to i_0$, $(00001) \to i_1, \ldots, (11111) \to i_{31}$. The condition that the rule must be legal as stated by Wolfram and the above condition of the existence of 1011-soliton leaves only 12 parameters: $(l_1, l_2, l_3, l_4, l_5, l_6, l_7) - (k_1, k_2, k_3, k_4, k_5) \equiv (i_4, i_{10}, i_{14}, i_{17}, i_{21}, i_{27}, i_{31}) - (i_7, i_9, i_{15}, i_{19}, i_{23})$. Interesting behavior which does not belong to the usual class-4 type is soliton-like behavior. For some rules, the dynamics of system is governed only by the soliton-like excitations (1011) and their collisions. If the "solitons" pass through each other by collisions, the CA can be regarded as a kind of integrable system. In some other rules, the collision of solitons show the sensitive dependence on the phase of the collisions, which induces the turbulence as an ensemble of 1011-solitons. Some of course, show the usual class-4 behavior, while some show class-3 or class-2 behavior for most of the initial conditions. Here, the evolution of the following four rules are shown. See Ref. 12 for more details:

(1) S1: (0001101-01011): class-4 like (see Fig. 15.1)
(2) S2: (0011101-01011): class-4 like (see Fig. 15.2)
(3) S3: (0000011-11000): integrable-like (see Fig. 15.3)
(4) S4: (0001011-10011): soliton-turbulence (see Fig. 15.4)

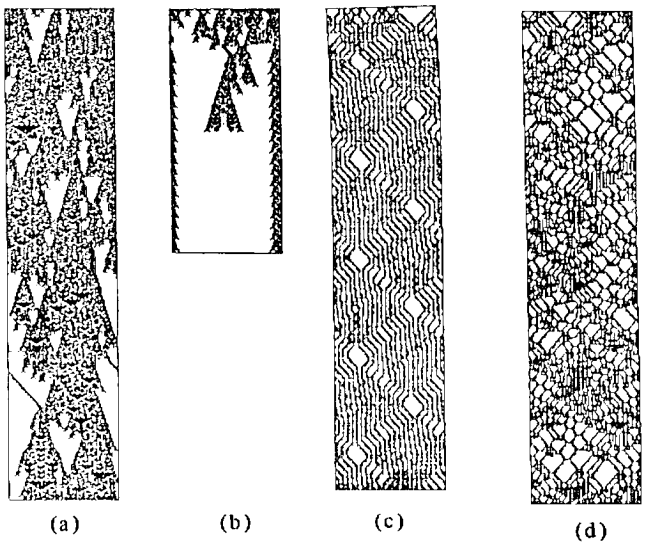


Fig. 15. Examples of the evolution of CA with "solitons". $N = 100$. Random initial configurations.

References

1. S. Wolfram, *Rev. Mod. Phys.* **55** (1983) 601; *Physica* **10D** (1984) 1.
2. *Physica* **10D** (1984), ed. D. Farmer, T. Toffoli, and S. Wolfram.
3. Dynamical Systems and Cellular Automata (1985), Academic Press, ed. J. Demongeot, E. Golès, and M. Tchuente.
4. R. Shaw, *Zeit. fur Naturforschung* **36a** (1981) 80.
5. See e.g., J. J. Hopfield, *Proc. Natl. Acad. Sci.* **79** (1982) 2554; **81** (1984) 3088.
6. T. Hogg and B. A. Huberman, *Proc. Natl. Acad. Sci.* **81** (1984) 6871.
7. S. Wolfram, in Ref. 2.
8. T. Hogg and B. A. Huberman, *Phys. Rev.* **A32** (1985) 2338.
9. See for the structures for Rule 90; O. Martin, A. Odlyzko, and S. Wolfram, *Comm. Math. Phys.* **93** (1984) 219.
10. The jumping among attractors was first investigated for the random Boolean networks, by S. Kauffman, *J. Theor. Biol.* 22 (1969) 437.
11. K. Matsumoto and I. Tsuda, *J. Phys.* **A18** (1985) 3561 and preprint.
12. Y. Aizawa, I. Nishikawa and K. Kaneko, in preparation.
13. K. Kaneko, in "Dynamical Systems and Nonlinear Oscillations" pp. 194-209 (ed. G. Ikegami, World Scientific, 1986).
14. K. Kaneko and Y. Akutsu, *J. Phys.* **A19** (1986) 69.
15. M. Mezard *et al.*, *Phys. Rev. Lett.* **52** (1984) 1156; D. Sherrington and S. Kirkpatrick, *Phys. Rev. Lett.* **35** (1975) 1792.
16. C. Grebogi, E. Ott, and J. Yorke, *Physica* **7D** (1983) 181; S. Takesue and K. Kaneko, *Prog. Theor. Phys.* 71 (1984) 35.
17. K. Kaneko, "Collapse of Tori and Genesis of Chaos in Dissipative Systems" (World Scientific, 1986) Chapter 7; J. P. Crutchfield and K. Kaneko, in preparation.
18. K. Kaneko, "Lyapunov Analysis and Information Processing in Coupled Map Lattices" to appear in *Physica* **D** and references cited therein.
19. J. D. Farmer, N. H. Packard, and A. S. Perelson, *Physica* **D**, to appear.
20. F. J. Varela, "Principles of Biological Autonomy", North-Holland, 1979.

Approaches to Complexity Engineering*

Stephen Wolfram
The Institute for Advanced Study, Princeton NJ 08540.

(December 1985; modified February 1986)

Principles for designing complex systems with specified forms of behaviour are discussed. Multiple scale cellular automata are suggested as dissipative dynamical systems suitable for tasks such as pattern recognition. Fundamental aspects of the engineering of such systems are characterized using computation theory, and some practical procedures are discussed.

The capabilities of the brain and many other biological systems go far beyond those of any artificial systems so far constructed by conventional engineering means. There is however extensive evidence that at a functional level, the basic components of such complex natural systems are quite simple, and could for example be emulated with a variety of technologies. But how a large number of these components can act together to perform complex tasks is not yet known. There are probably some rather general principles which govern such overall behaviour, and allow it to be moulded to achieve particular goals. If these principles could be found and applied, they would make new forms of engineering possible. This paper discusses some approaches to such forms of engineering with complex systems. The emphasis is on general concepts and analogies. But some of the specific systems discussed should nevertheless be amenable to implementation and detailed analysis.

In conventional engineering or computer programming, systems are built to achieve their goals by following strict plans, which specify the detailed behaviour of each of their component parts. Their overall behaviour must always be simple enough that complete prediction and often also analysis is possible. Thus for example motion in conventional mechanical engineering devices is usually constrained simply to be periodic. And in conventional computer programming, each step consists of a single operation on a small number of data elements. In both of these cases, much more complex behaviour could be obtained from the basic components, whether mechanical or logical, but the principles necessary to make use of such behaviour are not yet known.

Nature provides many examples of systems whose basic components are simple, but whose overall behaviour is extremely complex. Mathematical models such as cellular automata (e.g. [1]) seem

* Loosely based on an invited talk entitled "Cellular automaton engineering" given at the conference on "Evolution, Games and Learning" held at Los Alamos in May 1985. To be published in *Physica D*. More details will appear in due course.