Journal of Statistical Physics, Vol. 45, Nos. 1/2, 1986

Long-Range Effects in an Elementary Cellular Automaton

Peter Grassberger[1]

Received March 20, 1986

We present evidence that one of the "elementary" one-dimensional cellular automata in the sense of Wolfram (rule 22 in Wolfram's notation) involves very complex long-range effects, similar to a critical phenomenon. This is in contrast to superficial evidence that would suggest that this rule leads to fairly simple behavior.

KEY WORDS: Cellular automata; correlations; entropies; long-range; complexity.

1. INTRODUCTION

One of the most interesting phenomena in the field of dynamical systems is the spontaneous generation of complex structure. This concerns first the time behavior of chaotic systems with few degrees of freedom. More intriguing yet are spatially extended systems that are able to generate spatial patterns. For recent attempts to define the concept of "complexity" in such cases mathematically, see Refs. 1 and 2.

In order to study the phenomenon of spatial pattern formation, it is crucial to devise simple models that can be simulated without too much numerical effort. A special role is played by cellular automata (CA), which are models with discrete-time, discrete-space lattices, and with discrete variables at each lattice site. They can be considered as kinetic spin systems, and we shall sometimes call "spins" the objects occupying lattice sites. In contrast to conventional spin system (e.g., the Glauber model), we do not require detailled balance for CAs, we consider deterministic evolution rules (although nondeterministic rules can also be studied[3-5]), and we have discrete time.

[1] Physics Department, University of Wuppertal, D-5600 Wuppertal 1, West Germany.

0022-4715/86/1000-0027$05.00/0

CAs were introduced quite early by von Neumann,[6] and have been studied ever since (for a good overview, see Ref. 7). But it was only recently that an effort was undertaken by Wolfram[8] to search systematically for interesting phenomena and to possible behavior in the simplest case of one-dimensional automata. He found essentially four types of behavior, ranging from very dull patterns to a behavior is conjectured to be able to simulate a universal Turing machine and thus to represent the highest possible (algorithmic) complexity.

Among these one-dimensional CAs, the simplest ones are those with two states per lattice site and with local evolution rules depending only on next neighbors. There are only 256 such rules, called "elementary" by Wolfram.[8] They do not seem to contain any case of type 4 behavior (universal computer). They are numbered in the following way: write the eight outcomes (in one time step) of local neighborhoods 111, 110,..., 000 in a row, and read them as the eight binary digits of the number of the rule. For instance, rule 22 is the rule defined by

$$s_{t+1}(i) = \begin{cases} 1 & \text{if} \quad s_t(i-1) + s_t(i) + s_t(i+1) = 1 \\ 0 & \text{else} \end{cases} \tag{1.1}$$

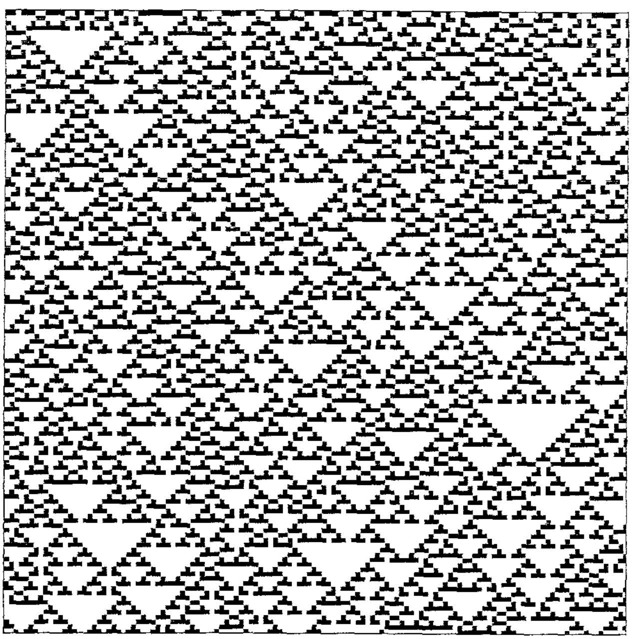

Fig. 1. Typical pattern generated by cellular automaton 22.

In the present paper, we shall study this rule 22 in detail. When starting from a random initial configuration, it seems first that the generated patterns in space-time are of limited complexity (see Fig. 1). We shall show in the following that this impression is misleading: there are very subtle correlations hidden in Fig. 1, associated with very long-range effects.

We should point out that there is no *a priori* reason why rule 22 should behave at all specially. Consequently, we must expect similar behavior in other deterministic spatiotemporal systems, such as reaction-diffusion systems, Benard cells, or chaotic semiconductor devices. In view of the subtlety of the effects discussed below, it might be very difficult to detect them there.

2. RULE 22

In this section, we shall study in detail the evolution under Eq. (1.1). In all cases, we shall start with random initial configurations.

(a) First, we shall discuss in what sense rule 22 shows a sensitive dependence on initial conditions. It is only due to this that we can expect ergodic behavior and that we are justified in discussing statistical aspects of a deterministic evolution.

Following Ref. 9, we consider two patterns $\{s_i(t)\}$ and $\{s_i'(t)\}$, which differ at time $t = 0$ only on one site. Sensitive dependence on initial conditions means that the region where the patterns are different widens with time. A typical run is shown in Fig. 2, where only the difference between the two patterns is plotted: a point is black if $s_i \neq s_i'$ there, and white otherwise. The Lyapunov exponent λ is defined[8] as the average speed by which the right and left fronts progress. The precise value of λ was estimated by letting the right front propagate as in Fig. 2, but forcing the left front to recede with the same speed. Then, we could use periodic boundary conditions (on lattices with up to 2400 sites) to perform very long runs with up to 10^7 time steps. The value thus obtained was $\lambda = 0.7660 \pm 0.0003$. More important than the precise number is that these calculations show beyond reasonable doubt that rule 22 does have a positive λ.

(b) Next, we computed the average density of "1" and the spatial correlation function

$$\langle s_i(t)\, s_{i+n}(t) \rangle - \langle s_i(t) \rangle^2 \tag{2.1}$$

by averaging over 120 lattices of size 11,000 with random start. The first 5000 iterations were discarded, and the next 9000 were used for the

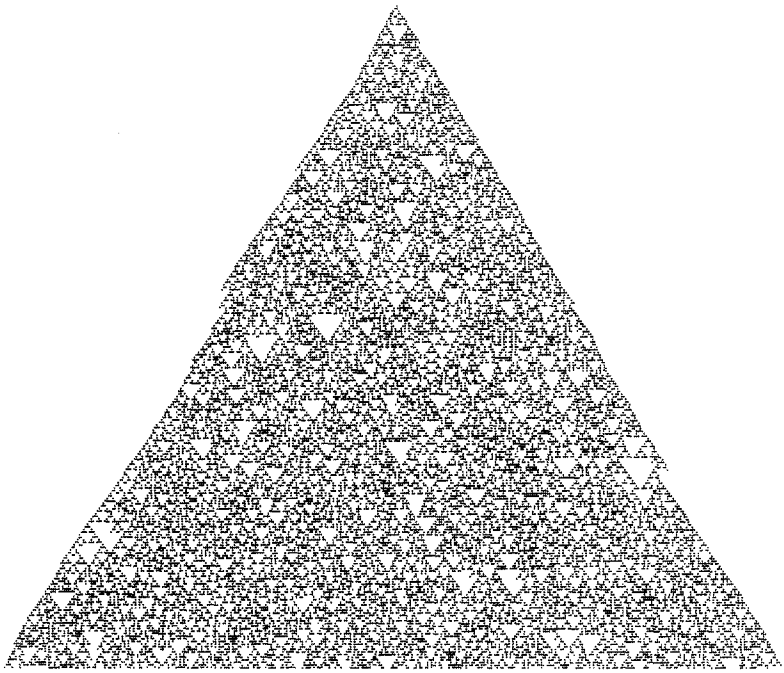

Fig. 2. Difference pattern between two patterns that at time $t=0$ agree everywhere except on a single point and are random otherwise.

averaging. The density was found to be 0.35096 ± 0.00001; correlations are shown in Fig. 3.

The fluctuations in Fig. 3 are not statistical (where not indicated, statistical errors are smaller than the dots). It is clear that, even apart from these fluctuations, the correlation does not decay exponentially as one might have guessed *a priori*. Plotting the data of Fig. 3 on a doubly logarithmic plot indicates that a power decay does not fit the trend of the data either: a better fit might be obtained, e.g., with an Ansatz involving an exponential of the square root of distance, but the fluctuations preclude a more definite statement. We might add that density and correlations seemed absolutely stationary after ~ 1000 iterations.

(c) The next calculations were done on lattices with finite width and with periodic boundary conditions. On such a lattice, any initial configuration must ultimately lead to an orbit periodic in time.

If the width is, e.g., $N = 60$, more than 80% of all configurations lead to the quiescent state ($s_i = 0$ for all i). Three typical runs are shown in Fig. 4. But the probability of ending in the quiescent state depends very

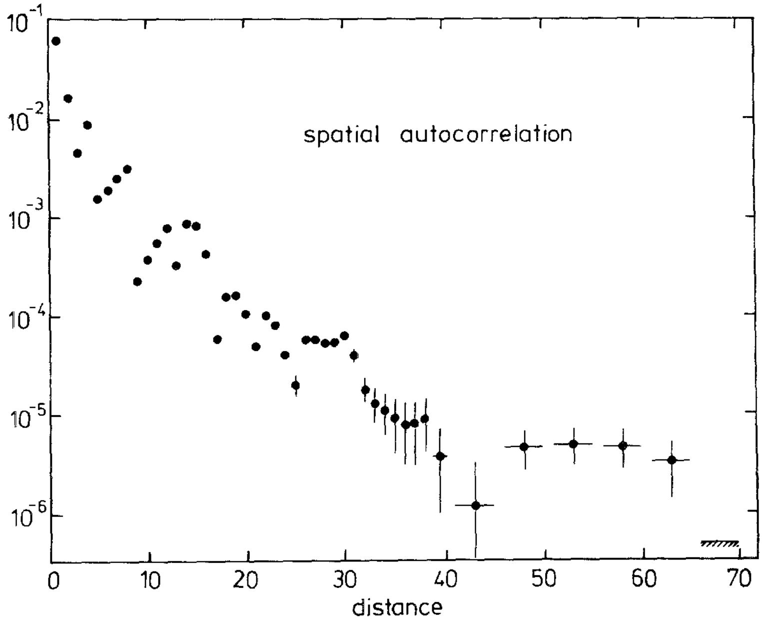

Fig. 3. Spatial correlation function $\langle s_i(t)\, s_{i+n}(t)\rangle - \langle s_i(t)\rangle^2$.

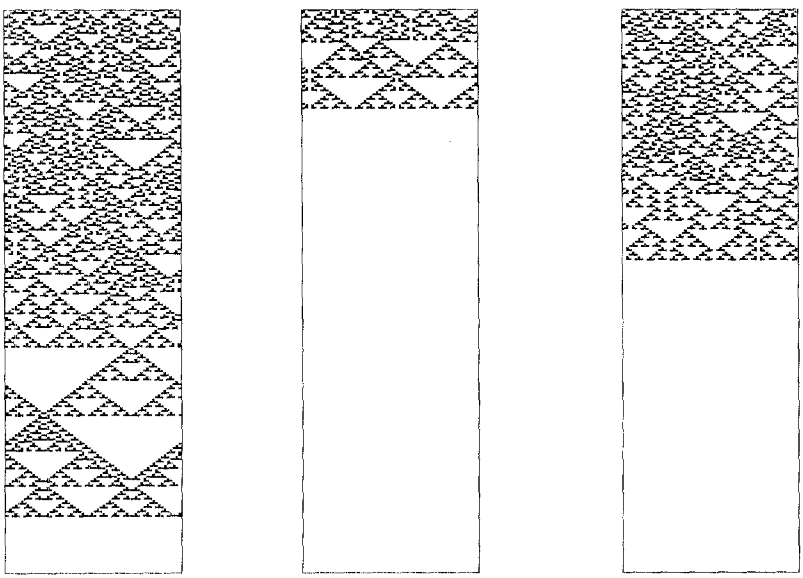

Fig. 4. Typical behavior of patterns on finite latices of width $N = 60$. On lattices of different widths, the behavior is in general completely different.

strongly and very irregularly on the width. Results, based on 2000 starting configurations for each value of the width N, are shown in Fig. 5. Similarly irregular is the average length of the cycle into which the orbits lead.

On the other hand, the rate with which randomly chosen starting configurations enter into cycles is surprisingly regular. For any fixed width N, the chance to be not yet on the final cycle decreases roughly exponentially,

$$P_N(t) = e^{-t/T(N)} \tag{2.2}$$

The time scales $T(N)$ entering here are plotted versus N in Fig. 6. More precisely, what is plotted in Fig. 6 is the difference $T_2 - T_1$, where T_1 is the time at which 50% of all configurations have entered a cycle and T_2 is the time at which 95% have entered it. We see from Fig. 6 that this time scale increase rather monotonically with N, roughly like

$$T(N) \propto e^{0.15N} \tag{2.3}$$

We consider the difference between Figs. 5 and 6 as very striking. Figure 5 shows that the precise value of N is very crucial in determining

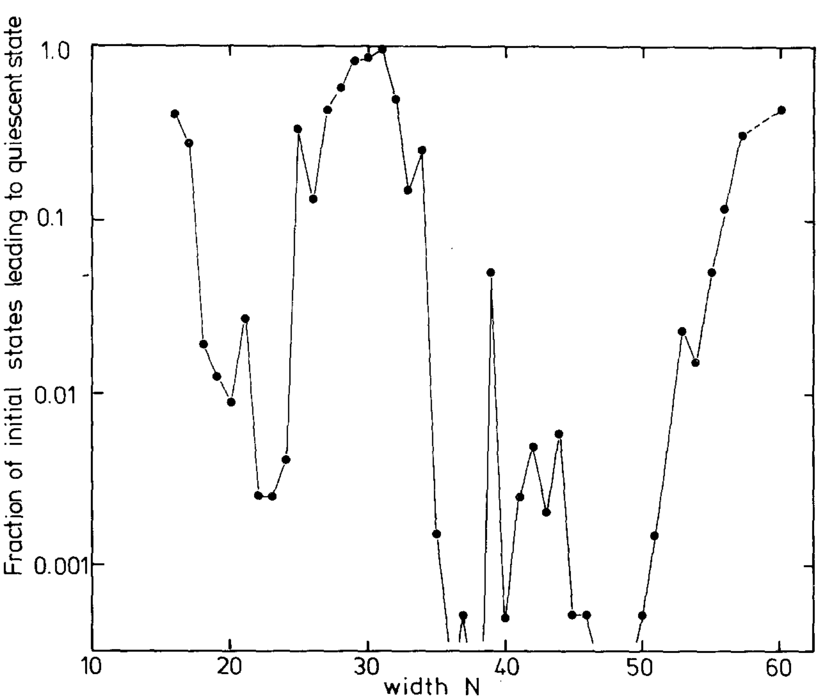

Fig. 5. Percentage of starting configurations leading ultimately into a quiescent state, on lattices of width N with periodic boundary conditions.

the final cycle into which typical orbits merge. Figure 6, on the other hand, suggests that during the transient time the exact value of N is not very important. Taken together, these results suggest that the precise value of N has a very subtle but still important influence, just as suggested by the correlations presented above.

(d) More sensitive to long-range effects than the correlation function (2.1) seem to be block entropies.

Let us consider a string $S = \{s_1, s_2, ..., s_N\}$ of N neighboring spins at fixed time t. The probability to find this string at a randomly chosen

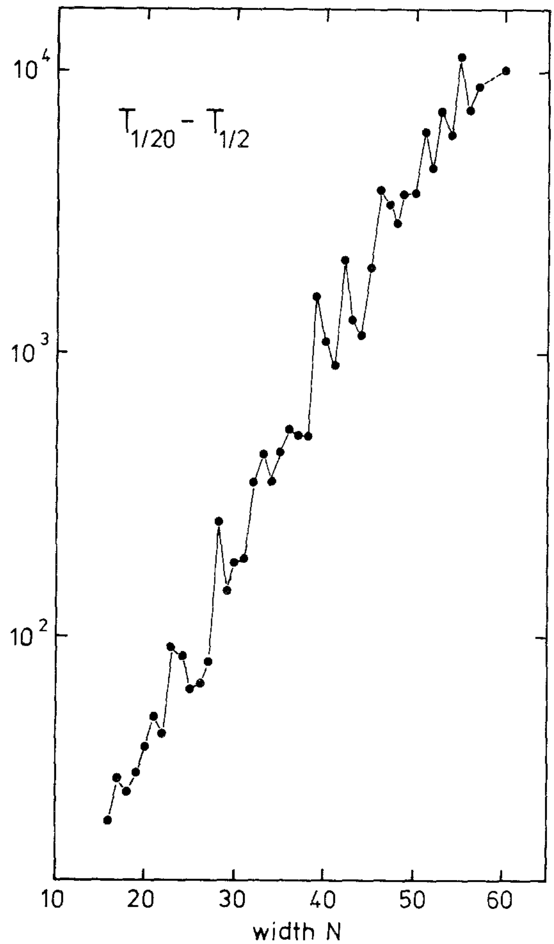

Fig. 6. Time scales for the average transient time during which trajectories have not yet entered the final cycle.

location is $p_N\{S\}$ for $t \to \infty$. The information stored in these N spins is then

$$H_N^{(x)} = -\sum_S p_N\{S\} \log p_N\{S\} \tag{2.4}$$

and spatial block entropies are defined as

$$h_N^{(x)} = H_{N+1}^{(x)} - H_N^{(x)} \tag{2.5}$$

The spatial entropy is then just

$$h^{(x)} = \lim_{N \to \infty} h_N^{(x)} \tag{2.6}$$

Temporal block entropies are defined in a similar way, but with an important difference.[9] Instead of a string, we consider a rectangular block of $N \times T$ spins, of spatial width N and of time length T. Informations $H_{N,T}^{(t)}$ are defined just as in Eq. (2.4), and block entropies are defined as

$$h_{N,T}^{(t)} = H_{N,T+1}^{(t)} - H_{N,T}^{(t)} \tag{2.7}$$

The temporal entropy is

$$h^{(t)} = \lim_{N,T \to \infty} h_{N,T}^{(t)} \tag{2.8}$$

It is well known[9] that $h_N^{(x)}$ decreases monotonically with N, while $h_{N,T}^{(t)}$ decreases with T for any N. Indeed, the difference

$$\delta h_N^{(x)} = h_N^{(x)} - h_{N+1}^{(x)} \tag{2.9}$$

is just the amount of information by which a spin s_{i+N} at site $i+N$ becomes less uncertain if the spin s_i gets known, in addition to all spins in between. We might call $\delta h_N^{(x)}$ the Nth-order mutual information in space.

Results for $h_N^{(x)}$ are shown in Fig. 7, while similar results for the temporal block entropies are shown in Fig. 8. Both are based on simulations of rather large lattices, with up to 36,000 lattice sites and up to 30,000 time steps. We see that both are very slowly but steadily decreasing (this decrease was overlooked in the much less extensive simulations of Ref. 10). On the other hand, the limit $N \to \infty$ seems to be reached for the temporal enropy already at $N=2$. This suggests that cylinder sets of width 2 provide generating partitions for the temporal entropy.

We have tried several parametrizations for the N (resp. T) dependence of the block entropies. By far the best fits were obtained with power laws

$$h_N^{(x)} = \text{const}/N^{0.06} \tag{2.10}$$

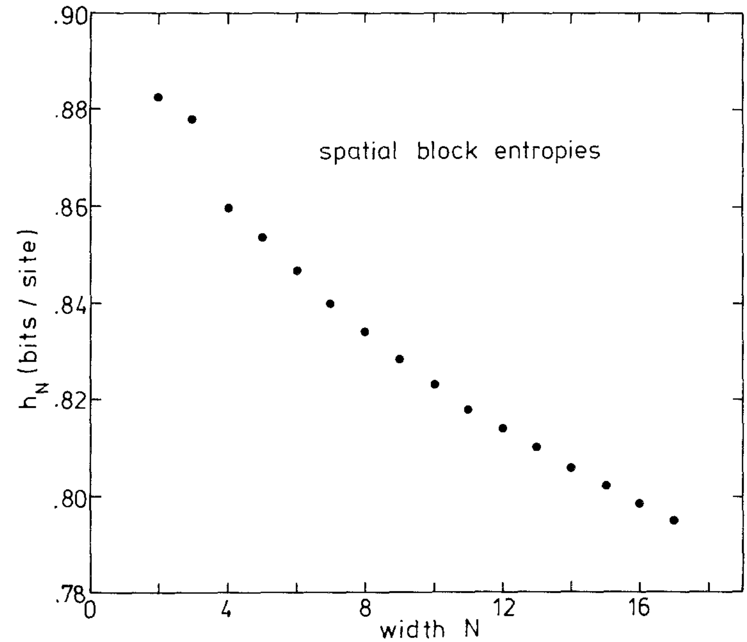

Fig. 7. Spatial block entropies $h_N^{(x)}$.

and

$$h_{N,T}^{(t)} = \text{const}/T^{0.18} \qquad \text{for} \quad N \geqslant 2 \tag{2.11}$$

Notice that these parametrizations imply that $h^{(x)} = h^{(t)} = 0$. This means that the pattern of Fig. 1 is not truly random, in contrast to its appearance, and in spite of the sensitive dependence on initial conditions indicated by the positive Lyapunov exponent. Instead of being random, it should be called "complex" by the criteria introduced in Ref. 2, since the block entropies decay slower than 1/(block length), implying very long and complex correlations.

(e) The above results show quite clearly that there are long-range effects, but they offer no explanation for them. We shall now present a last long-range effect which might also give some clue to the origin of these effects.

Cellular automata are special cases of dynamical systems, with some similarities to systems of few continuous degrees of freedom and with some striking dissimilarities. Indeed, every CA can be represented by a map that maps, e.g., the unit square onto itself.[10] The main common feature is that the unpredictability of chaotic evolution is due to a flow of information[11]

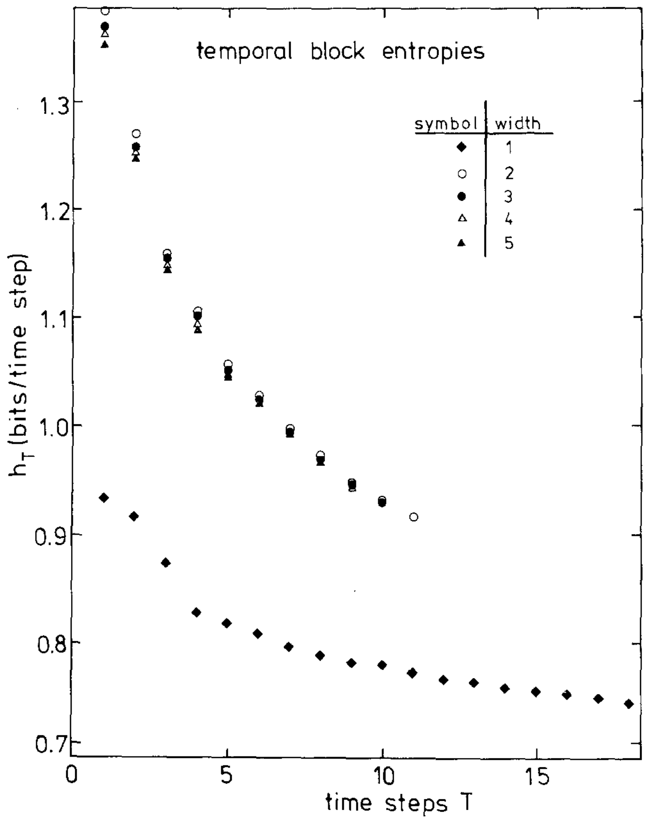

Fig. 8. Temporal block entropies $h^{(t)}_{N,T}$ for values of N from 1 to 5.

from "insignificant" digits of the continuous variables into the significant digits. The main difference between CAs and maps like the Henon map is that the latter is smooth, while the map corresponding, e.g., to rule 22 is highly fractal. As a consequence, nonlinear effects in smooth maps are only seen in the "significant" digits, while the information processing in the "insignificant" digits is a very simple flow only.[12] A consequence of this is a simple relation between temporal entropy, Lyapunov exponents, and dimension (which is essentially spatial entropy per digit).[12,13] Using the analogy between maps and CAs, this relation would lead in the present case to

$$H^{(t)}_{N,T} = H^{(x)}_{N+2\lambda T} \tag{2.12}$$

But this relation is in general wrong. It is violated numerically for rule 22. Instead, one has only an inequality[9]

$$H^{(t)}_{N,T} \leqslant H^{(x)}_{N+2\lambda T} \tag{2.13}$$

The reason for this seems to be that information is not just flowing through the sites of a CA, but is manipulated such that there is a loss of information at every site. On the one hand this loss implies that very long-range correlations can build up, since they are not disturbed by a permanent inflow of information, on the other hand it makes the block entropies converge to zero very slowly.

Let us be more specific after this rather vague introduction. Consider an infinite lattice in a stationary state characterized by spatial block probabilities $p_N\{S\}$. Consider now the case that two different configurations containing blocks $S(t)$ and $S'(t)$ lead to the same configuration at time $t+1$. At first sight, one might argue that this is impossible since then the number of different blocks would decrease and the state would not be stationary. Now that is not true, due to the infinity of the lattice, but we see that the situation is quite subtle.

It is best to formulate everything in terms of informations. If there were no correlations of range longer than N, and if two blocks S and S' of length N map onto the same block, then we would have an associated loss of information per time step and per lattice site of the order of $\delta I_N\{S, S'\}$, with

$$\delta I_N\{S, S'\} = p_N\{S\} \log \frac{p_N\{S\} + p_N\{S'\}}{p_N\{S\}} + p_N\{S'\} \log \frac{p_N\{S\} + p_N\{S'\}}{p_N\{S'\}} \tag{2.14}$$

(remember that $p_N\{S\}$ is the probability to find block S starting at a given lattice site). This leads for a string of length $M \gg N$ to an information loss $\propto M \cdot \delta I_N$ per iteration. This loss has to be counterbalanced by an inflow of information of at most 2 bits through the ends of the string. This is clearly impossible for sufficiently large M, showing that either our assumption of vanishing correlations beyond distance N is wrong, or δI_N must be zero.

For rule 22, we found that the shortest blocks leading to the same descendents are $S_1 = \{11011\}$ and $S_2 = \{11111\}$. Indeed, both $\cdots S_1 \cdots$ and $\cdots S_2 \cdots$ lead to the same string $\cdots 000 \cdots$. Both S_1 and S_2 seem to occur with nonvanishing probabilities in the stationary distribution. In order to apply Eq. (2.14), we consider all blocks of length $N = 2n + 5$ of the forms

$$S = \{LS_1R\}, \qquad S' = \{LS_2R\} \tag{2.15}$$

where L and R each represent blocks of length n. The total information losses

$$\Delta I_N = \sum_{L,R} \delta I_N\{S, S'\} \tag{2.16}$$

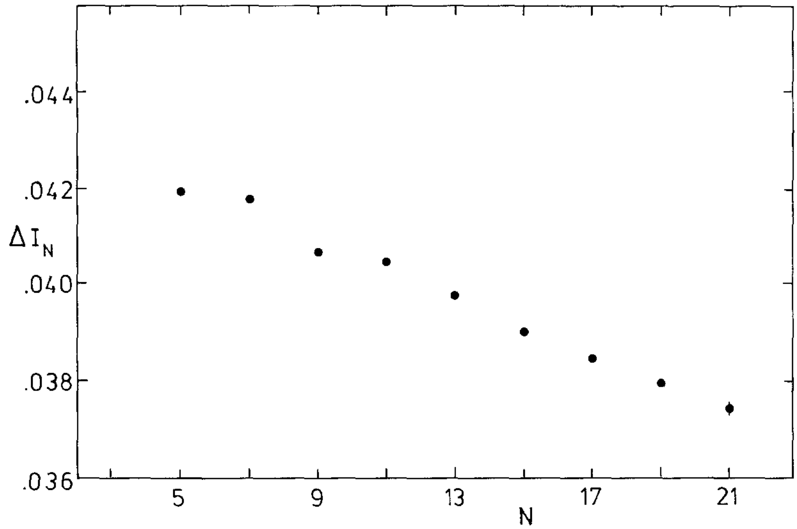

Fig. 9. Quantities ΔI_N [Eq. (2.16)] obtained by summing Eq. (2.14) over all sequences containing in their middle the strings 11011 (resp. 11111). These two sets of strings map onto the same strings in one time step. If there were no correlations of range $>N$, this would lead to an information loss of the order of ΔI_N per lattice site.

again obtained from simulations, are shown in Fig. 9. Statistical errors are smaller than the dots except for $N=21$.

According to our discussion, we would expect ΔI_N to approach zero as fast as the correlations do. The observed very slow decrease of ΔI_N thus indicates very long-range correlations. If long correlations were zero and ΔI_N different from zero, the distribution would contract to a small set of low entropy. We find thus that very long-range correlations are necessary to prevent the distribution from collapsing to something periodic. The result is neither periodic nor truly random.

3. DISCUSSION

We have presented numerical evidence that there are very important and nontrivial long-range effects in one seemingly very simple CA. This is so in spite of the very simple appearance of patterns created by this CA. There are other CAs that produce pictures looking much more complicated. We might thus conjecture that behavior similar to that found in the present paper should be quite ubiquitous in cellular automata and maybe also in other deterministic systems such as reaction-diffusion systems or hydrodynamic turbulence.

We have presented very heuristic arguments that this behavior results from a nearly contractive motion on the attractor of the system. The type of system we are considering is dissipative and irreversible on the entire phase space. But once an orbit is on the attractor, there can be no further contraction—otherwise the attractor would not yet be the attractor. In the studied case, this further contraction seems to be prevented only by the long-range correlations.

The resulting patterns are then neither periodic nor random, but somewhere in between. We propose to call them complex.[2] Note that the concept of "complexity" underlying this statement is not the usual one of algorithmic (or computational) complexity used, e.g., in Ref. 1. Instead, it is a probabilistic concept. According to algorithmic complexity, all systems able to simulate a universal Turing machine (such as Conway's game of life[14]) are in the same complexity class. Preliminary simulations[15] suggest that the game of life, started with random configurations, does not lead to similar long-range effects as those presented here and is thus in a lower probabilistic complexity class. On the other hand, there exist no indications that rule 22 can simulate a universal Turing machine. But more extensive simulations and better theoretical insight would be necessary to settle this problem.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge very interesting discussions with S. Wolfram. Also, I would like to thank J. Dias de Deus for inviting me to a very stimulating workshop on CAs in Lisbon (Dec. 1985). The present paper is partly based on a talk given there.

REFERENCES

1. S. Wolfram, *Commun. Math. Phys.* **96**:15 (1984).
2. P. Grassberger, Towards a quantitative theory of self-generated complexity, Wuppertal preprint (1986).
3. L. Schulmann and P. E. Seiden, *J. Stat. Phys.* **19**:293 (1978).
4. P. Grassberger, F. Krause, and T. von der Twer, *J. Phys. A* **17**:L105 (1984).
5. W. Kinzel, *Z. Phys. B* **58**:229 (1985).
6. J. von Neumann, *Theory of Self-Reproducing Automata* (A. W. Burks, ed.) (University of Illinois Press, 1966).
7. D. Farmer and S. Wolfram, eds., Proceedings of workshop on cellular automata, Los Alamos 1983 *Physica D* **10** (1984).
8. S. Wolfram, *Rev. Mod. Phys.* **55**:601 (1983).
9. S. Wolfram, *Physica D* **10**:1 (1984).
10. P. Grassberger, *Physica D* **10**:52 (1984).
11. R. Shaw, *Z. Naturforsch.* **36a**:80 (1981).
12. P. Grassberger, in *Chaos—An Introduction*, A. V. Holden, ed. (1986).
13. F. Ledrappier and L.-S. Young, Berkeley preprints (1984).
14. E. R. Berlekamp, J. H. Conway, and R. K. Guy, *Winning Ways, for Your Mathematical Plays* (Academic Press, 1982).
15. P. Grassberger, unpublished.