Journal of Statistical Physics, Vol. 51, Nos. 5/6, 1988

Dynamics of Simple Computer Networks

Michel Kerszberg[1] **and David Mukamel**[2]

Received February 18, 1988

We consider a simple two-dimensional layered automaton. Each processor in the automaton performs the same nonlinear, thresholdlike operation (so that the row-to-row evolution of the array can also be seen as the time development of a one-dimensional automaton). One row of the machine is reserved for input, another is singled out as output. We study the output space in detail, as restricted by the very wiring of the array, enumerating the output configurations, and characterizing them statistically. We demonstrate that input configurations flow to a set of zero measure in output space. The variations in output that are to be expected when input is subjected to perturbations are also examined.

KEY WORDS: Cellular Automata; non-linear dynamics; scaling; fixed points.

1. INTRODUCTION

In many natural and artificial systems, decision capacity based upon complex computations is distributed throughout a network of components. This is true, among others, of parallel data processing systems,$^{(1)}$ the central and peripheral nervous systems,$^{(2)}$ and biochemical as well as ontological development pathways.$^{(3)}$ The complication of these networks is such that, more often than not, a detailed description becomes both impractical and theoretically uninteresting. As against this, much can be learned from simplified models, and we propose in this paper to analyze in some detail the properties of such a model. Although the "majority model," as we call it, has been introduced in the context of developmental biology

[1] IFF der KFA Jülich, 5170 Jülich, West Germany, and Department of Nuclear Physics, Weizmann Institute of Science, 76100 Rehovot, Israel.

[2] Department of Nuclear Physics, Weizmann Institute of Science, 76100 Rehovot, Israel, and Laboratoire Louis Néel, CNRS, Grenoble 38041, France.

0022-4715/88/0600-0777$06.00/0

(a detailed account of its relevance there was given in ref. 4a), its significance will be evident also in the framework of the so-called neural networks[4b,5] if we simply mention that its elements are threshold automata[6] connected in multiple layers, in rough analogy with the cerebal cortex.[2] Two of the layers are privileged in that they are considered respectively as the "input" and the "output" layers.

Extensive work on automaton networks, classifying their dynamical properties, has been performed.[7] In spite of this, we believe that our model is the first simple, geometrically regular system whose input–output mapping can be characterized analytically in such detail as has been achieved here. Previous analytical results, closely connected to ours, have been obtained mainly in the context of random systems.[8] In Section 2 the model is defined in detail, and we comment briefly on the complex phenomena it exhibits. In Section 3 we discuss the properties of the input–output mapping. We show that the mapping is many-to-one, i.e., inputs can be grouped into basins of attraction characterized by a given output. In Section 3.1 we enumerate all attractors and show how their number scales with the size of the system. This property is shown to be relevant for many other quantities as well. Section 3.2 is devoted to the analysis of the statistical structure of attractors; in particular, we describe the typical attractors in this way, using the methods of statistical mechanics (by typical, we mean all attractors save for a set of measure zero). In Section 3.3 we show that most inputs map, in fact, on atypical attractors. The atypical attractors to which typical inputs flow we call the *accessible* attractors. In Section 3.4 we obtain bounds on the size of the accessible attractors and also calculate exactly the size of the largest attractors. Finally, in Section 4 we address briefly the question of the complicated interrelationships among some of the basins of attraction. We do this by analyzing the behavior of the output as the input is subjected to some diffusion process. Because of the complexity of the input–output mapping, even simple dynamics at the input level can lead to complicated behavior of the output.

2. THE MAJORITY MODEL

The computer array we are considering has a layered structure. Each layer consists of N processors, or automata, which receive their inputs from the layer immediately above. The internal state of the processors on the top layer is fixed: this set of states we call collectively the "input" of the array. In the simplest version (different ones will be alluded to in the following sections) each automaton processes three quantities, transferred respectively from the above left, above center, and above right automata (see

Fig. 1a). The inputs are bits (0 or 1); the output of the individual automaton is determined according to the "majority rule":

$$\text{if} \quad \sum \text{inputs} \leqslant 1, \qquad \text{then output} = 0$$
$$\text{else output} = 1$$

This rule is typical of *threshold* automata; here the threshold is 1. Analogous processing occurs in all of the M layers of which the machine consists; the bits required on the sides are supplied in one of two ways: the boundary conditions are either fixed (the bits are set to, say, 0) or periodic. The states of the Mth layer of processors we call the "output" of the array.

As simple as this model may seem, it exhibits a surprising richness of behavior in its input–output relationship.[4] As an illustration of this complexity, let us simply mention the fact (Fig. 1b) that with fixed 0 boundary conditions, say, the input string $I_1 = \{00101010100\}$ leads to output $O_1 = \{00000000000\}$ (provided $M \geqslant 5$), while $I_2 = \{00101110100\}$, differing from I_1 by only one bit, leads to $O_2 = \{00011111000\}$. There is a great variety of possible output states (their number decreases as M increases at fixed N); we term each O_i an "attractor" and the set of inputs $\{I_j^{(i)}\}$ leading to O_i its basin of attraction. Section 3 is devoted to counting, mainly in the $M \to \infty$ limit, all the possible attractors (whose number grows exponentially with N) and to determining the statistical structure and size of some basins of attraction.

```
0 0 1 0 1 0 1 0 1 0 0
0 0 0 1 0 1 0 1 0 0 0
0 0 0 0 1 0 1 0 0 0 0        a
0 0 0 0 0 1 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0

0 0 1 0 1 1 1 0 1 0 0
0 0 0 1 1 1 1 1 0 0 0
0 0 0 1 1 1 1 1 0 0 0        b
0 0 0 1 1 1 1 1 0 0 0
0 0 0 1 1 1 1 1 0 0 0
```

Fig. 1. The majority rule. (a) A finite array of majority rule processors. The state of the top processors (inputs) is forced to 0 or 1 as shown; each processor in the lower rows determines its own state as the one displayed by a majority of its three upper neighbors; the side processors (not depicted) are set to zero (b) Same array with input differing in one bit: five output bits are switched to 1.

This multiplicity of attractors and their complicated relationships are reminiscent of the situation in spin glasses and random automaton networks.[8,9] Additionally, as demonstrated above, the system displays in certain circumstances a strong dependence (in output) on its input conditions. If we think for a moment of the successive layers as the states at successive times of a *unique* layer, then "sensitive dependence on input" translates into a "strong dependence on initial conditions," as prevalent on the strange attractors of chaotic systems.[10]

3. STATIC INPUT–OUTPUT PROPERTIES

3.1. Number of Attractors and Scaling

When the depth M of the machine tends to infinity, it is fairly easy to compute exactly the number of possible output configurations, or attractors. Indeed, the result of successive applications of the majority rule can only be the disappearance of all *isolated* 0's or 1's. Thus, all we have to do is count the number of bit strings with no such isolated bits, and this will provides us with the number of possible output configurations. We illustrate the procedure for the case of periodic boundary conditions. This case has also been treated in ref. 11. The allowed attractors are then those strings that contain no isolated bits, save possibly for two identical isolated bits at each end. We begin by counting the number B^N of strings which *neither* start *nor* end with isolated bits. This we can obtain easily by first evaluating recursively the number of strings A^N that begin (at the left) with at least two identical bits, but may possibly end with one isolated (rightmost) bit. Assume indeed we have found the number of such strings with length $N-1$, and that a given such string ends with a bit x ($=0$ or 1). We now want to add a new bit to the right, say y. From the condition that there should be no *isolated* bit embedded in the new string, we find that if $x=y$, the addition of y is always permissible, while if $x\neq y$, the new N-string is allowed *provided* x was not isolated, i.e., provided the $(N-1)$-string ended with a pair. Thus, for instance,

$$A^N_{0,0}=A^{N-1}_{0,0}+A^{N-2}_{0,1} \tag{1a}$$

where $A^N_{a,b}$ denotes the number of N-strings beginning with at least two a's and ending with at least one b. Also,

$$A^N_{0,1}=A^{N-1}_{0,1}+A^{N-2}_{0,0} \tag{1b}$$

and $A^N_{1,1}=A^N_{0,0}$, $A^N_{1,0}=A^N_{0,1}$ by symmetry. We solve for the A^N by going to the variables $S^N = A^N_{0,0} + A^N_{0,1}$, $D^N = A^N_{0,0} - A^N_{0,1}$, in terms of which (1) becomes

$$S^N = S^{N-1} + S^{N-2} \tag{2a}$$

$$D^N = D^{N-1} - D^{N-2} \tag{2b}$$

Note that (2a) defines the Fibonacci series. We look for a solution $S^N = ax^N$, $D^N = by^N$, which yields

$$\begin{aligned} S^N &= a_1 \left(\frac{1+\sqrt{5}}{2}\right)^N + a_2 \left(\frac{1-\sqrt{5}}{2}\right)^N \\ D^N &= b_1 \left(\frac{1+i\sqrt{3}}{2}\right)^N + b_2 \left(\frac{1-i\sqrt{3}}{2}\right)^N \end{aligned} \tag{3}$$

Initial conditions must now be supplied. For $N=2$, strings beginning with 00 obviously also end with a 0; thus, $A^2_{0,0} = 1$, $A^2_{0,1} = 0$; on the other hand, when $N=3$, $A^3_{0,0} = 1$ (namely 000), while $A^3_{0,1} = 1$ (namely 001, which only *ends* with an isolated 1). Thus, $S^2 = 1$, $D^2 = 1$, and $S^3 = 2$, $D^3 = 0$. This allows us to compute

$$a_1 = \frac{\sqrt{5}}{5}, \qquad a_2 = -a_1, \qquad b_1 = \frac{i\sqrt{3}}{3}, \qquad b_2 = -b_1 \tag{4}$$

One then obtains

$$A^N_{0,0} = A^N_{1,1} = \frac{S^N + D^N}{2}, \qquad A^N_{0,1} = A^N_{1,0} = \frac{S^N - D^N}{2} \tag{5a}$$

Now the B^N are easily seen to be equal to the corresponding A^{N-1}: they result simply from repeating the rightmost bit, ensuring that the string begins *and* ends with a pair. Thus,

$$B^N_{a,a} = A^{N-1}_{a,a} \tag{5b}$$

It is now a simple matter to obtain the number C^N of atractors for a machine of width N. Certainly, all the $2(B^N_{0,0} + B^N_{0,1})$ N-strings are acceptable; however, we can take any of the $B^{N-2}_{0,0}$ (or $B^{N-2}_{1,1}$) strings of length $N-2$, supplement them on both sides with ones (or zeros), and obtain an acceptable attractor. There are $2B^{N-2}$ such attractors. Finally, from any $B^{N-1}_{0,1}$ string, one can form an attractor by adding either 1 at the right or 0 at the left (and similarly for the $B^{N-1}_{1,0}$ strings). This adds $4B^{N-1}_{0,1}$ strings to the count; thus

$$C^N = 2[B^N_{0,0} + B^N_{0,1} + B^{N-2}_{0,0} + 2B^{N-1}_{0,1} + \mathrm{mod}(N-1, 2)] \tag{6a}$$

The last term represents the attractors $\{0101...01\}$ and $\{1010...10\}$ present in even-N machines and not included previously. These attractors correspond to inputs consisting of repeated $\{01\}$ or $\{10\}$ substrings, which are reproduced after an even number of levels, while yielding repeated $\{10\}$'s or $\{01\}$'s, respectively, after an odd number of levels. Equation (6a) is correct, as can be checked by direct counting on large-M machines. Asymptotically, for large N, we have

$$C^N \sim \left(\frac{1+\sqrt{5}}{2}\right)^N \tag{6b}$$

It may be worth noting that, although boundary conditions can change the size of individual attractors drastically (see Section 3.4), they have little effect on the C^N: thus, for $N = 10$, (6a) gives 124 attractors, while in the case of fixed, all-0 boundary conditions there are 117, but with a—sometimes—very different appearance.

For large N, Eq. (6b) has an approximate extensive form. This is a rather pervasive feature in our model, as we now show. The phase space of the majority machine can be divided into sectors that become independent in the limit of infinite width and depth. This is because of the "canalization" property: two adjacent 0's or 1's at input are always propagated (Fig. 2a).

$$\begin{array}{cccccccc} ? & ? & ? & 0 & 0 & ? & ? & ? \\ ? & ? & ? & 0 & 0 & ? & ? & ? \\ ? & ? & ? & 0 & 0 & ? & ? & ? \end{array} \quad \text{a}$$

$$\begin{array}{cccccccc} ? & ? & 0 & 0 & 1 & 0 & ? & ? \\ ? & ? & 0 & 0 & 0 & ? & ? & ? \\ ? & ? & 0 & 0 & 0 & ? & ? & ? \end{array} \quad \text{b}$$

$$\begin{array}{cccccccc} ? & 1 & 1 & 0 & 1 & 0 & 1 & ? \\ ? & 1 & 1 & 1 & 0 & 1 & ? & ? \\ ? & 1 & 1 & 1 & 1 & ? & ? & ? \end{array} \quad \text{c}$$

Fig. 2. "Canalization" in the majority machine. (a) Two adjacent zeros or ones are transmitted all the way to the bottom, irrespective of the states of the other bits (?). We speak of "frozen" layers in this case. (b) If a 01 couple is itself flanked by two zeros (at right and left), the states of the second and further layers are frozen (with a triplet of zeros). (c) A 1010 group, flanked left and right by, say, two ones, leads to a frozen configuration starting from the third layer.

In order to understand the consequence of this for the structure of phase space, let us consider an infinitely deep and wide machine, and let us divide it in two by an imaginary vertical line. Two cases can exist: the twp borderline bits either are the same or are not. In the first case, they form a 00 or 11 pair, which, according to the majority rule, is transmitted from layer to layer without changes: thus, in this instance, there is no interaction between the two halves: the number of states in phase space is simply multiplicative, i.e., with self-explanatory notation, $W(0|0) \simeq W_L(0)\, W_R(0)$. Consider now the situation where we have, say, a 0 bit on the left with a 1 bit on the right. There are various possibilities. Suppose the configuration around the "joint" is now 0010; this generates a 000 triplet at the next level and we are back to the previous case (see Fig. 2b); similarly, *mutatis mutandis*, for 1011. Furthermore, 0011 generates two "propagating pairs," 00 and 11, and again the two halves will behave independently of each other. Thus, $W(00|10) \simeq W_L(00)\, W_R(10)$; $W(10|11) \simeq W_L(10)\, W_R(11)$, $W(00|11) = W_L(00)\, W_R(11)$. The only ambiguous situation is ...1010..., when the states of the left bits, say, can still influence the right-hand side. Here, too, however, partial decoupling occurs, since, e.g., after *two* layers ...110101... will yield ...11... again (Fig. 2c). We see that, the deeper the machine, the less the relative importance of the interaction between the two halves becomes, since a larger fraction of states are effectively frozen at the lower levels. For depth $M=1$, 1/2 of the total phase space of 2^N, $N \to \infty$ is "interaction-free." For $M=2$, the "free" fraction goes up to $1/2+6/16=7/8$; for $M=3$, the corresponding number is $1/2+6/16+6/64=31/32$, and so on. We conclude that, the deeper the machine, the better the approximation that considers its phase space to be made up of the direct product of two independent halves. More generally, for $M \to \infty$ we expect that, at the Mth layer, a property W such as, say, the number of available states, will have the form $W = W_L W_R$. Thus, one surmises

$$W \sim w^N \tag{7a}$$

where w is W per bit, e.g., phase space per bit.

In a different limit, that of $N \to \infty$, M finite, the only bits that can possibly affect a given output bit form its "reverse light cone" (Fig. 3); since this is *finite*, we expect again that two halves of a machine behave in an approximately independent way; and the extensivity property should thus be quite generally valid.

This conclusion can be extended rather simply into a scaling formula for certain quantities. Let $f(R, N)$ be the probability averaged over all input strings that if a certain number R of input bits are upset (inverted), the output remains unaltered. On grounds of the previous arguments, "half

$$\begin{array}{ccccccccc}
0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0
\end{array}$$

Fig. 3. For a machine of finite depth and infinite width, correlations extend only over a finite fraction of the machine, delimited by the "reverse light cone."

machines" can again be expected to behave more or less independently in this respect. We thus expect the function f to exhibit the following scaling behavior:

$$f(R, N) = [F(R/N)]^N \tag{7b}$$

We see in Fig. 4 that the function f is indeed of this form (7b), thus affording a sort of "law of corresponding states."

A variety of quantities can be easily evaluated using the extensivity and scaling properties as seen above; we discuss other examples as we proceed.

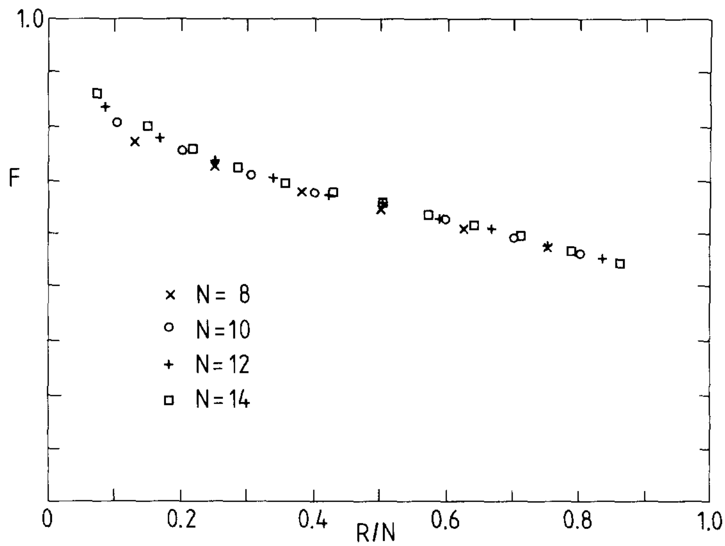

Fig. 4. Starting from any input configuration, inverting a fraction R/N of the input bits will conserve the output configuration with a probability $f \leqslant 1$. Shown is a plot of $F = f^{1/N}$ for various values of N: the points fall on the same curve, in accordance with the scaling law (7) (see text).

3.2. Statistical Structure of Attractors

In this section we demonstrate that for infinitely long chains all attractors (up to a set of zero measure) have a certain statistical structure. In order to carry out this analysis, we consider a chain of N bits. An attractor is characterized by L segments of bits where each segment is composed of either zeros or ones, but not both, and is of length $l \geqslant 2$. Let n_l be the number of times a segment of length l appears in this attractor. The set of integers $\{n_l\}$ satisfies the two constraints

$$\sum_{l=2}^{\infty} l n_l = N \tag{8a}$$

$$\sum_{l=2}^{\infty} n_l = L \tag{8b}$$

Let $\Omega\{n_l\}$ be the entropy associated with the set $\{n_l\}$. It is defined such that $I \equiv e^{\Omega}$ yields the number of attractors that are characterized by $\{n_l\}$. Simple combinatorial considerations give the following expression for I:

$$I \equiv e^{\Omega} = \frac{L!}{n_2! \cdot n_3! \cdots} \cdot 2 \tag{9}$$

In the thermodynamic limit ($N \to \infty$) one has

$$\Omega \simeq L \ln L - \sum_{l=2}^{\infty} n_l \ln n_l$$

In order to find the structure of a typical atractor in the thermodynamic limit ($N \to \infty$), one has to determine the set $\{n_l\}$ that maximizes Ω subject to the two constraints (8). To do that we consider

$$\tilde{\Omega} = L \ln L - \sum_{l=2}^{\infty} n_l \ln n_l - \alpha \sum_{l=2}^{\infty} n_l - \beta \sum_{l=2}^{\infty} l n_l$$

where α and β are Lagrange multipliers. Maximizing $\tilde{\Omega}$ with respect to n_l we find

$$n_l = A e^{-\beta l}, \qquad l = 2, 3, \ldots \tag{10}$$

with

$$A = e^{-(\alpha + 1)}$$

where α and β are determined (as functions of N and L) by the two constraints (8). Thus, we have shown that a typical attractor (i.e., all attractors up to a set of zero measure) are characterized by a particular distribution (10) of the number of segments n_l. To complete the analysis, we use the distribution function (10) to rewrite the entropy Ω in terms of N and L. We then determine $x \equiv L/N$ by maximizing Ω as a function of x. Using (10), we find that

$$\Omega = N[\beta + x \ln(L/A)]$$

With the use of the two constraints (8) one gets

$$\Omega = N\left(-\ln y + \frac{1-y}{2-y} \ln \frac{y^2}{1-y}\right)$$

where $y = e^{-\beta}$. The entropy Ω is maximized by

$$y \equiv \frac{\sqrt{5}-1}{2}$$

This yields

$$x \equiv \frac{L}{N} = \frac{\sqrt{5}-1}{2\sqrt{5}} \tag{11}$$

and

$$n_l = N\left(\frac{\sqrt{5}-1}{2\sqrt{5}}\right)\left(\frac{\sqrt{5}-1}{2}\right)^l \tag{12}$$

In summary, in the thermodynamic limit a typical attractor is characterized by segments of average length L/N given by (11). The number of segments of length l decreases exponentially with l according to the expression (12). One can easily check that the total number of attractors obtained in this way e^{Ω} agrees, in the large-N limit, with the exact expression obtained in the previous section for arbitrary N.

It is easy to see that in a typical attractor the segments are distributed in a statistically independendent way. In order to demonstrate this point, we consider a chain of L sites. On each site we define a variable s_i, which takes the values 2, 3,.... It represents the length of the ith segment in this chain. Consider, for example, the segment–segment correlation function. The number of configurations in which the i_1 and i_2 segments are of length l_1 and l_2, respectively, is given by

$$e^{\Omega_{12}} = \frac{(L-2)!}{n_2! \cdots (n_{l_1}-1)! \cdots (n_{l_2}-1)! \cdots} \cdot 2$$

Therefore, the joint probability of having segments of lengths l_1 and l_2 is given by

$$\frac{e^{\Omega_{12}}}{e^{\Omega}} = \frac{n_{l_1}}{L} \frac{n_{l_2}}{L-1} \simeq \frac{n_{l_1}}{L} \frac{n_{l_2}}{L}$$

which is the product of the probabilities for l_1 and l_2 segments.

3.3. Almost All Inputs Flow to Atypical Attractors

In this section we demonstrate that the entire input space (up to a set of zero measure) maps onto atypical attractors, that is, on a set of zero measure in atractor space. In order to carry out this analysis, we consider a typical input. A typical input is constructed by independently choosing bits 0 or 1 with probability 1/2. Now, the configurations in input space that yield segments of length two in output are necessarily of the form ...110011... or ...001100.... The probability for such input configurations is

$$2(1/2)^6 = 1/32$$

On the other hand, the probability to find a segment of length 2 in a typical output is

$$\frac{n_2}{L} = \left(\frac{\sqrt{5}-1}{2}\right)^2 \simeq 0.382$$

Since a typical input does *not*, in this particular case, yield the typical *output*, we conclude that most inputs flow to an atypical portion of output space, which we call its *accessible* portion.

3.4. Size of the Largest Attractor; Size of the Accessible Attractors

The largest attractors (those output strings corresponding to the largest number of possible inputs) are of great interest in applications, since they represent those outputs that are least sensitive to input distortion, and this is often a desirable characteristic.

For $M \to \infty$, these large attractors are those corresponding to output $\{000...000\}$ or $\{111...111\}$. It is easily seen that the basin of attraction of $\{000...000\}$, say, contains all input strings that have no adjacent 1's (these would automatically appear in output). Let us remark that this problem can be put into one-to-one correspondence with the problem of counting the total number of attractors, which has been solved above. If we take any allowed output configuration, say $\{...0001100111...\}$, and construct the "dual" string by applying the exclusve-OR(XOR) operation to each pair of its bits, we obtain $\{...001010100...\}$, a member of the $\{00...00\}$ attractor.

Let us, as an example, evaluate the size G^N of the $\{000...000\}$ attractor in the case of fixed, all-zero BC. The *only* condition on the relevant strings is that their 1's are strictly isolated: there are two strings for $N=1$ ($G^1=2$), namely $\{0\}$ and $\{1\}$, and three for $N=2$, i.e., $\{00\}$, $\{01\}$, and $\{10\}$. Furthermore, one N-string can always be built by adding a 0 to an $(N-1)$-string, while a 1 may be added only to strings not ending with 1 already. Thus, $G^N=G^{N-1}+G^{N-2}$, and the BC yield

$$G^N=\frac{3+\sqrt{5}}{2\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^N+\frac{-3+\sqrt{5}}{2\sqrt{5}}\left(\frac{\sqrt{5}-1}{2}\right)^N \tag{13}$$

The effect of the BC on the attractor can be drastic. Thus, for $N=10$, (13) gives $G^{10}=144$; on the other hand, one can check that the $\{111...111\}$ attractor, disfavored by the BC, has only 21 inputs to itself (note that $21=G^{10-4}$, and in general, the size of the all-1 attractors will G^{N-4}); while with periodic BC both attractors comprise 121 input strings each.

The size of the large attractors can in fact be calculated also for finite-depth machines. Thus, after one level of processing, the strings that flow to the $\{000...000\}$ attractor are those where 1's are framed by at last *two* 0's on each side: $\{...00100...\}$ is allowed, but clearly $\{...00101...\}$ is not. It is easy to count the input configurations obeying this condition; their number obeys the recursion relation for $G_1^N=G_1^{N-1}+G_1^{N-3}$. A solution of the form ax^N yields a cubic equation whose solution is $x=1.466...$. Thus, $G_1^N\sim(1.466...)^N$. At the second level, the input configurations that yield $\{000...000\}$ are all those already counted at level 1, plus those where one observes $\{...0010100...\}$. The recursion equation becomes

$$G_2^N=G_2^{N-1}+G_2^{N-3}+G_2^{N-5}$$

The resulting fifth-order algebraic equation gives $G_2^N\sim(1.570...)^N$. That the G_i ultimately converge to those given by the Fibonacci equation is clear, since the recursion relation for G_i, $i\geqslant 2$, is

$$G_i^N=G_i^{N-1}+G_i^{N-3}+\cdots+G_i^{N-2i-1} \tag{14}$$

and from this

$$G_i^{N-2}=G_i^{N-3}+\cdots+G_i^{N-2i-3}$$

Subtracting these last two equations yields

$$G_i^N = G_i^{N-1} + G_i^{N-2} - G_i^{N-2i-3}$$

and for $i \to \infty$, the last term becomes smaller and smaller—leaving the correct recursion relation for G_∞^N.

We now give a bound on the size of accessible attractors. We do this by showing that these attractors are almost all of the same size. Then the bound simply follows from dividing the total input space 2^N by the total number of attractors $[(1+\sqrt{5})/2]^N$. The argument is as follows. First, one can convince oneself that the size of the attractors depends on the correlation functions of the s_i only. If one can prove that the distribution of the s_i in the set of accessible attractors is self-averaging, then almost all members of this set will exhibit the correlations characteristic of the distribution. This would imply that almost all accessible atractors (up to a set of zero measure) indeed have the same size. To prove self-averaging, we show that the quantity

$$\frac{1}{N^2}\left\langle \sum_i s_i^2 - \left\langle \sum_i s_i \right\rangle^2 \right\rangle$$

decreases as $1/N$ for large N. This follows from the fact that the s_i have short-range correlations only. In particular, while s_i, s_{i+1}, and s_{i+2} are correlated, s_i and s_{i+j}, $\geqslant 3$, are uncorrelated. The point is that for given s_i and $s_{\pm 3}$, any local configuration in input that leads to a segment of length s_i will restrict $s_{i\pm 1}$ and $s_{i\pm 2}$, but not $s_{i\pm 3}$, and thus s_i and $s_{i\pm 3}$ are independent. Consider, e.g., the case $s_i = 3$. There are three local input configurations that lead with certainty to a substring i consisting of 3 ones, namely

$$\{0011100\}, \quad \{000101100\}, \quad \text{and} \quad \{00110100\}$$

These local configurations put some restrictions on s_{i+1} and s_{i-1}. For example, the second configuration restricts s_{i-1} to be at least 3 (so it cannot be 2) and the third one does the same thing for s_{i+1}. Both configurations prevent s_{i+1} and s_{i-1} from being equal to 2 simultaneously. Thus, correlations exist linking nearest and next nearest neighbors. On the other hand, there are no restrictions linking s_i and s_{i+j}, $|j| \geqslant 3$, that is,

$$\langle s_i s_{i+j} \rangle = \langle s_i \rangle \langle s_{i+j} \rangle, \qquad |j| \geqslant 3$$

Therefore, almost all accessible attractors have the same correlation functions, and thus the same size. It follows that the size of accessible attractors is $\geqslant [4/(1+\sqrt{5})]^N$.

4. ERROR-CORRECTING PROPERTIES; DYNAMICS

In applications concerned with artificial intelligence or evolutionary biology, the question of error sensitivity plays an essential role. For instance, if a machine is used in a pattern recognition context, sensitivity of the output string to input perturbations—e.g., bit "upsets" or inversions—is of primary importance in engineering considerations.

This sensitivity can be measured in the following convenient way. Consider a given machine, with a given initial input, to which a certain output corresponds. Assume now that the input is subjected to random bit upsets. The question we ask is: what is the probability that, if there are a bit inversions per unit time (the effect of which propagate instantaneously to the output), the output at time t is identical with output at time 0? This we call the autocorrelation function for this given output. It may seem that this is a complicated way to explore the phase space of the system, i.e., of mapping its attractors; however, a direct mapping is difficult in a 2^N-dimensional discrete space such as the one we are dealing with; furthermore, the interrelations between attractors are complex. Thus, if we look at an $N \times M = 12 \times 8$ machine with all-0 boundary conditions, we observe that minor changes in the input configurations may lead to a change of the output. For example, there are 492 points in input space—out of a total of 2^{12}—for which any change in a single bit will result in a change of the output. Thus, the picture in phase space is, it would seem, intractable unless we adopt some form of statistical search.

Simple cellular computers such as the majority machine can amplify the effect of correlations to an enormous extent. Unless these amplification effects remain isolated accidents, they do not bode well for the future of approximations that neglect either all correlations (see below) or even only some higher order correlations. Let us first evaluate some of the consequences of neglecting correlations altogether. Computing the 1-bit autocorrelation function is then easy indeed. Let us consider an input bit first; if the probability per unit time that this bit be upset is a, then the probability for the bit in question to have retained (or regained) its initial value at time t is the solution of

$$dP_0/dt = (1-a)P_0 + a(1-P_0) \tag{15}$$

with $P_0(0) = 1$, where the subscript 0 denotes layer 0 of the machine. Thus,

$$P_0 = \tfrac{1}{2}(1 + e^{-2at}) \tag{16}$$

We assume that all bits in input are upset with *independent* probabilities. Then we may evaluate simply P_1, the autocorrelation of a bit in layer 1.

For, if only one bit in a given input triplet is upset, then the chance is 1/2 that this upset has not changed the majority in the triplet; if two bits are upset, then again the probability is 1/2 that this has not led to a change in majority; while if all three bits were inverted, the majority has clearly changed. And so:

$$P_1 = P_0^3 + \tfrac{3}{2}P_0^2(1-P_0) + \tfrac{3}{2}P_0(1-P_0)^2 \tag{17}$$

This is an exact result. Now if all bit triplets at level 1 were changing independently of one another, all that one would have to do to get P_n would be to iterate (17). But this gives the nonsensical result that $P_i \to 1/2$ when $i \to \infty$ at any finite time, that is, each and every bit in output is randomized after even one inversion in input! Clearly, then, neglecting correlations is strictly meaningless. Let us now formulate the problem in terms general enough that useful aproximations will later be possible.

4.1. Formal Results

Let us consider an $N \times M$ machine. As a function of time, inversions are occurring in its input string at a rate a. At time t, the probability $P(R)$ that the number of inversions relative to the initial input is R will be given by

$$P[R(t)] = \binom{N}{R} (1-P_0)^R P_0^{N-R} \tag{18}$$

where P_0 is as defined in (16). If at $t=0$ input was I_0 and output O_0, at time t output will still be O_0 with a probability

$$P_M^N(I_0) = \sum_{R=0}^{N} P[R]\, f_{I_0}(R) \tag{19}$$

with $f_{I_0}(R)$ the "density" of the O_0 attractor around its point I_0, i.e., the number of input strings at Hamming distance R from I_0 that still flow to O_0, divided by the total number of strings at that distance. Let us rewrite (18) and (19), scaling variables so that $r = R/N$, taking N and R large; then

$$P_M^N(I_0) \simeq \int_0^1 \exp\{N[-r\ln r - (1-r)\ln(1-r) + r\ln(1-P_0) + (1-r)\ln P_0 + \ln F_{I_0}(r)]\}\, dr \tag{20}$$

where we have assumed that f is of the scaling form (7b). The integrand in (20) thus assumes a form suitable for a saddle-point approximation in the limit of large N. We now perform the calculation explicitly for particular cases.

4.2. Autocorrelation for the {000...000} Attractor

To use (20) concretely, one must be able to specify F_{I_0}. This will be done here for several cases. We first take $O_0 = \{000...000\}$, $M \to \infty$. Using the techniques illustrated in Section 3.2, one can easily determine the characteristics of a *typical* input I_0 leading to this attractor. The input will consist of isolated ones separated by segments comprising s_i zeros each. The number n_l of segments of length l can be seen to obey the distribution

$$n_l = \frac{3-\sqrt{5}}{2\sqrt{5}} N \left(\frac{1+\sqrt{5}}{2}\right)^{-l} \tag{21}$$

The total number of segments L is

$$L = N \frac{5-\sqrt{5}}{10} \tag{22}$$

Now, the density $f_{\{00...00\}}$ of the all-0 attractor around a typical input I_0 is given by the ratio of the number X of inversions in R bits that lead to other points inside the attractor, to the total number of possible ways of inverting R bits, i.e., $\binom{N}{R}$. Assuming a small overall density of inversions, so that no "interactions" between neighboring inversions must be considered, the places where inversions are allowed in I_0 are (i) the L locations where a 1 exists; and (ii) the $l-2$ locations inside a segment of l zeros.

Thus,

$$X = L + \sum_{l=3}^{\infty} n_l(l-2) \tag{23}$$

Using the distribution (21), X is easily evaluated:

$$X = N \frac{1+\sqrt{5}}{2\sqrt{5}} \equiv KN \tag{24}$$

It is then not difficult to see that for small R/N,

$$f_{I_0}\{00...00\} \simeq K^R = K^{rN} \tag{25}$$

with $r = R/N$. We now insert this expression f_{I_0} in (20) and perform a saddle point approximation. The saddle point equation is

$$\ln \frac{r}{1-r} - \ln \frac{1-P_0}{P_0} + \ln K = 0$$

and, for r small, $r \simeq (a/K)t$. Reinserting this in (20), we get

$$P_{\infty}^{\infty}(\{00...00\} \simeq \exp\left(-Na\frac{1-K}{K}t\right)$$

The exponent $Na(1-K)/K$ is less than Na (the corresponding exponent for the input correlation function), reflecting the error correction properties of the machine.

We now consider $f_{I_0}\{00...00\}$ for *atypical* inputs. For example, we examine the case $I_0 = \{00...00\}$. At a Hamming distance R from I_0, the representative input string has, at each site, a 1 with probability r; if we take $r \ll 1$, then the probability of finding a *pair* of 1's at a *given* location is r^2, and the occurrence of such pair is the most probable event that would destroy output 0_0. The probability that a given input bit pair is compatible with 0_0 is thus roughly $1-r^2$, and *all* pairs will be such with a probability

$$f_{\{000...000\}}(r) \simeq (1-r^2)^{N-1} \simeq (1-r^2)^N \tag{26}$$

(in the limit of N large). This is of the scaling form (7), and (20) then becomes

$$\begin{aligned} P_{\infty}^{N}(\{000...000\} &\approx \int_0^1 \exp\{N[-r\ln r-(1-r)\ln(1-r) \\ &\quad + r\ln(1-P_0)+(1-r)\ln P_0+\ln(1-r^2)]\}\,dr \end{aligned} \tag{27}$$

We now take $N \to \infty$ and evaluate (27) by the saddle-point method. The saddle-point equation is

$$-\ln\frac{r}{l-r}+\ln\frac{1-P_0}{P_0}-\frac{2r}{1-r^2}=0 \tag{28}$$

which becomes, with (16),

$$-\ln\frac{r}{l-r}+\ln\ \mathrm{tgh}\ at=\frac{2r}{1-r^2} \tag{29}$$

For short times and small r, this reduces to $r \simeq at$, which, when substituted in (27), yields

$$P_{\infty}^{\infty}(\{000...000\}, t) \simeq (1-a^2t^2)^N \simeq \exp(-Na^2t^2) \qquad \text{(short times)} \tag{30}$$

This Gaussian form is well confirmed by computer simulations (see Fig. 5). Note that, had we considered automata computing the majority of, say,

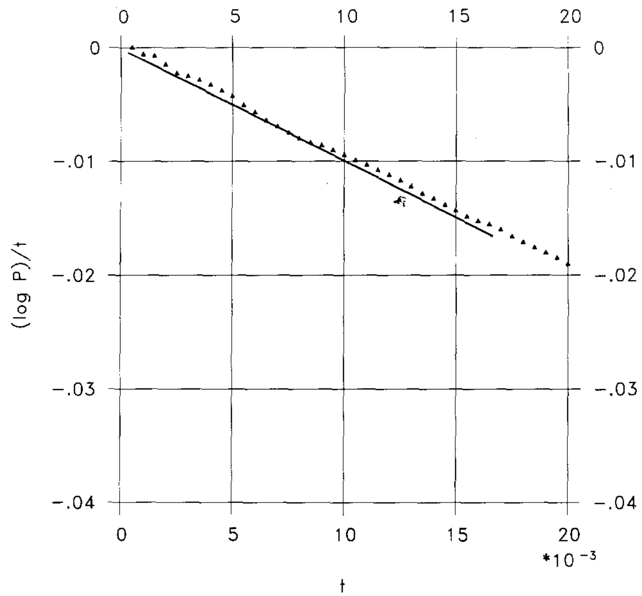

Fig. 5. Plot of $(\log P)/t$, where P is the probability that, starting from an all-0 input and proceeding in time with bit inversions, the output remains all-0 for an $N=100$, $M=\infty$ machine. The dependence expected is a Gaussian [Eq. (30)]. In the units used, we expect a straight line starting at (0, 0) and slope-1. This is well verified at short times.

five bits instead of three, the critical input fluctuation would then have consisted of *triplets* of 1's, and the autocorrelation decay would then have proceeded as $\exp(-t^3)$.

In the limit $t \to \infty$, (29) can again be solved and yields $r = 0.3236$; substituting in (20) results in $P_\infty^\infty(\{000...000\}, \infty) \simeq (0.8401)^N$. This is in reasonable agreement with the exact result $(0.8090)^N$ [derived from (13)], despite the crudeness of our approximations.

5. CONCLUSIONS

We have studied the behavior of extremely simple regular computing networks. We have calculated the statistical structure of both input and output spaces. The machine maps most of its input space into an atypical, measure-zero subset of its output space. Error correction capacity is observed, based on a number of rather complicated effects resulting from the surprising dependence of these networks on input correlations. It should prove very instructive to extend our work in a variety of directions; one can, for example, imagine networks where thresholds (here fixed) would vary (say randomly) from automaton to automaton, perhaps in some adaptive way. It would also be interesting to study longer range interactions. Finally, the networks considered to not involve feedback: clearly, the possibility of backpropagation effects could lead to dramatic consequences as far as network dynamics is concerned.

ACKNOWLEDGMENTS

We thank Paul Rujan for informative discussions. One of the authors (D. M.) thanks the members of Laboratoire Louis Néel, and especially Dr. B. Barbara, for their kind hospitality during a visit while this work was being concluded. We thank the Institute for Theoretical Physics, Santa Barbara, California for their support under NSF grant PHY82-17853. The work was also supported in part by the Israel Commission of Basic Research, and by the Einstein Center of Theoretical physics.

REFERENCES

1. R. Taylor, Artificial Intelligence and VLSI parallel computers, AI Conference, Munich, 28.11.86.
2. G. M. Edelman and V. B. Mountcastle, *The Mindful Brain* (MIT Press, 1978).
3. K. Sander, in *Evolution and Development* (Cambridge University Press, 1983), p. 137.
4. (a) Z. Agur and M. Kerszberg, *Am. Nat.* **129**:862 (1987); see also (b) B. A. Huberman and M. Kerszberg, *J. Phys. A* **18**:L331 (1985); M. Kerszberg, in *Disordered Systems and Biological Organization*, E. Bienenstock *et al.*, eds. (Springer-Verlag, 1986).
5. *Proceedings of the First IEEE International Neural Networks Conference* (San Diego, 14–21 June 1987).
6. E. Golès-Chacc, Thèse, Université de Grenoble (1985).
7. S. Wolfram, *Rev. Mod. Phys.* **55**:299 (1983).
8. B. Derrida, in *Chance and Matter* (Proceedings Les Houches Summer School, 1986); H. Flyvbjerg, in *Computer Simulation in Brain Science*, R. J. M. Cotterill, ed. (Cambridge University Press, 1988).
9. H. G. Schuster, *Deterministic Chaos* (Physik Verlag, 1984).
10. S. Redner, *J. Stat. Phys.* **25**:15 (1981); I. P. Goulden and D. M. Jackson, *Combinatorial Enumeration* (Wiley, 1983).
11. Z. Agur, A. S. Frenkel, and S. T. Klein, *Discr. Math.*, to appear.