Parallel computation in coupled chemical kinetic systems

A. Hjelmfelt[†], F. W. Schneider[†], and J. Ross[*‡]
[*]Corresponding author
[†]Institute of Physical Chemistry, University of Würzburg, Marcusstrasse 9/11,
D-8700 Würzburg, Germany
[‡]Dept. of Chemistry, Stanford University, Stanford, CA 94305

Abstract

We construct a network of open, bistable reaction systems coupled by mass transfer. The transfer rates are determined by a Hebb type learning rule. The programmable network can store patterns consisting of high and low concentrations in each bistable system. The network is a parallel computer; it recognises patterns similar, but not necessarily identical, to stored patterns. Bistability and mass transfer are present in biological systems, and as shown here, provide a possible basis for the processing of information.

All living beings must process information to one extent or another. This processing must be primarily of chemical origin; be it by macroscopic dissipative processes such as biochemical switches or transport of ions across a membrane, or by molecular properties, such as different conformations of a protein. The maintenance of stationary states far from equilibrium requires dissipation (entropy production). The choice of macroscopic dissipative processes for the construction of computational devices of possible interest in biological studies needs substantiation. One argument in its favor is the consumption of ATP used to drive kinetic processes which in the brain is about 10kg per day.[1] In three prior articles[2-4] suggestions were made for the chemical implementation of computing machines: the chemical coupling of reaction mechanisms[5] far from equilibrium whose stationary states have properties of a McCullouch-Pitts neuron,[6] led to the construction of logic gates, a finite state machine which generalizes to a universal Turing machine,[7] and a parallel neural network computer. In this communication we show a new class of parallel chemical computers based on chemical kinetics systems with multiple stable stationary states coupled by mass transfer. We demonstrate the computational capacity of such networks by storing patterns in it and solving pattern recognition problems.

Consider a set of chemical reactions occurring homogeneously in an enclosure, which may be a cell compartment, a neuron, or a stirred reaction vessel. Each enclosure is an open system: reactants enter and products, intermediates and unreacted material exit. We use a set of chemical reactions

which is bistable. There are many reaction mechanisms and physical processes in which bistability is known including enzymatic reactions[8] and neuronal response.[9,10] For simplicity of illustration we use an inorganic chemical reaction, the iodate-arsenous acid reaction,[11,12]

$$IO_3^- + 3H_3AsO_3 \rightleftharpoons I^- + 3H_3AsO_4, \tag{1}$$

for which the kinetics are adequately described by the temporal variation of a single variable, the concentration of I^-. Bistability occurs in this isothermal reaction: for given range of input flows of reactants (given constraints) the homogeneous reaction system at steady state may have either a high or low $[I^-]$. The two stable stationary states are not equally probable except at one set of values of the constraints.

The systems communicate with each other by mass transfer; either diffusional, in which case it is reciprocal, or active transport, in which case we require it to be reciprocal. The temporal evolution of the ith system, $i = 1, \ldots, N$ in the network is[12-14]

$$\begin{aligned}\frac{d[I^-]_i}{dt} &= -k_B[I^-]_i^3 + \{k_B([I^-]_o + [IO_3^-]_o) - k_A\}[I^-]_i^2 + \{k_A([I^-]_o + [IO_3^-]_o) - k\}[I^-]_i \\ &\quad + k[I^-]_o + \lambda \sum_{j \neq i} k_{ij}([I^-]_j - [I^-]_i),\end{aligned} \tag{2}$$

in which $[I^-]_o$ and $[IO_3^-]_o$ are the concentrations of iodide and iodate in the reactant flows that maintains each system away from equilibrium, k_A and k_B are effective rate coefficients, and k is the reactant flow rate. The mass transfer coefficients, $k_{ji} = k_{ij} \geq 0$, are reciprocal and chosen by a learning rule, and $\lambda \geq 0$ is a scaling parameter which may be adjusted freely to optimize the performance of the network. There exists a Liapunov function for chemical systems described by eqn. 2[14] which is an evolution criterion towards stable stationary states (analogous to the Gibbs free energy of a system approaching equilibrium). If each system has only one stable stationary state then the only stable stationary state of N coupled systems is spatially homogeneous. However, since the systems are bistable, there are 2^N possible stable stationary states of the network; the bistability of a single system is essential to the network operation. Each system represents a pixel,

and thus the stable stationary states of the network are patterns of low and high $[I^-]$, two of which are homogeneous.

Patterns are stored in this chemical network analogously to such storage in neural networks. Similarities and differences between the present work and neural networks of the Little[15] or Hopfield[16-18] type are discussed later. Let R_i^p be the activity of the ith system in pattern p; $R_i = 1$ for the high $[I^-]$ state and 0 for the low $[I^-]$ state. We use a Hebbian learning rule[19] to determine the mass transfer coefficients. If two systems i and j are in the same state in a stored pattern, either both have high $[I^-]$ or both have low $[I^-]$, then the connection weight is increased; otherwise the connection weight is decreased. The contributions to the connection strength between systems i and j is summed over each stored pattern. The learning rule is

$$k_{ij} = \frac{1}{N}\theta[\sum_p (2R_i^p - 1)(2R_j^p - 1)], \tag{3}$$

in which $\theta(x) = x$ if $x \geq 0$ and $\theta(x) = 0$ if $x < 0$. Thus, only if two systems are in the same state in the majority of patterns will they be connected. This learning rule stores both the pattern and its negative image.[16]

All simulations are performed with eqn 2 but to gain insight into these coupled equations an approximation can be performed. Suppose a number of random patterns are stored in the network by the learning rule, and the network is presented a pattern with a few pixels in error relative to the stored pattern p', then the time evolution of the ith system in the network is approximately given by[20]

$$\frac{d[I^-]_i}{dt} \approx f([I^-]_i) + \frac{\lambda}{2N}[([I^-]_i^{p'} - [I^-]_i) + 2\alpha(P)(\frac{[I^-]^{high} + [I^-]^{low}}{2} - [I^-]_i)], \tag{4}$$

in which $[I^-]^{high}$ and $[I^-]^{low}$ are the iodide concentrations of the two bistable states. The first term, $f([I^-]_i)$, is the right hand side of eqn 2 without the summation and drives each system to one of the two stable stationary states of the uncoupled system. The second term, $([I^-]_i^{p'} - [I^-]_i)$ drives system i to its state in pattern p'. The last term drives the system to homogeneity, and its multiplier[20] $\alpha(P)$ grows monotonically with the number of stored patterns, P, from $\alpha(2) = 0$ and is

about 0.5 when $P = 5$. Thus, when few patterns are stored in the network, they are expected to be successfully recalled.

The computational process carried out with eqn. 2 consists first of the storage of a number of patterns in the network by the learning rule, eqn. 3; thus the network is programmable. Second, the entire network is given initial conditions (i.e. each system is given an initial I^- concentration of one of the two stable stationary states of an uncoupled system). These initial conditions constitute the presented pattern.[21] If the presented pattern is similar to a stored pattern, then the pattern recognition process consists of the temporal evolution of the network from the presented pattern to the stable stationary state corresponding to the similar stored pattern; the network corrects errors in a recognized pattern so that it more closely resembles the most similar stored pattern. If the presented pattern is not recognized, the network evolves to a homogeneous state. The number of errors in recognized patterns are not always reduced to zero; extraneous steady states develop due to the "mixing" of stored patterns.[16] The network may correct many errors but may recall an extraneous pattern instead of a stored pattern. The number of such extraneous states increases as the number of stored patterns increases.

We characterize the chemical computer by examining the percentage of presented patterns, with a known number of errors relative to a stored pattern p',[21] which result in recall of patterns in the following classes: 1) p' or its associated extraneous patterns, 2) other stored patterns or their associated extraneous patterns, or 3) a homogeneous final state in the case of nonrecognition. The minimum number of pixels which must be reversed to convert an extraneous pattern to a stored pattern (i.e. the minimum Hamming distance) determines with which stored pattern an extraneous pattern is associated. To measure the extent of the error correction in recognised patterns, we calculate the average number of errors, relative to p', from the recalled patterns in class one.

In Fig. 1 we show a typical temporal evolution of the numerical solution of eqns 2.[13] The concentration of I^- in each system as a function of time is indicated by the shading. Three patterns

were stored in a network of 36 coupled systems, and a stored pattern is recalled perfectly at $t = 10$ from a presented pattern with 10 pixels in error, relative to the recalled pattern. The total experimental time corresponding to the computer simulation is about 5 hours (the numerical solution of eqn. 2 requires only a few seconds).[22] If we look at the time series for the 13th system, we see that it is initially in the darker state but by the third time step it has been corrected to the lighter state. On the other hand, system 12 is initially in the lighter state but is corrected at the seventh time step to be in the darker state. Systems initially in the wrong state exert an influence on those in the correct state. System 14 is initially in the light state, and in the recalled pattern it is also in the light state, however between t=1 and 3 it begins to move toward the dark state, but then at t=4 it returns to the light state.

What is the probability of a stored pattern being recalled? To study this we store three patterns in a network of 72 systems, and the network is presented with a pattern containing a certain number of errors relative to the pattern p'.[21] After we calculate the steady state of the network, we place it in one of the three classes. Plotted in fig. 2 are the percentages of the final states in each class as a function of the number of initial errors (relative to p'). Presented patterns with ten or fewer errors almost always result in the recall of patterns in class 1. The homogeneous state (class 3) is never found (in 100 trials) and patterns in class 2 are rarely recalled. The recall of patterns in class 1 is still the most likely result for patterns with 14 to 22 initial errors, but patterns in class 2 are occasionally recalled, and homogeneity is the second most common final state. For presented patterns with large numbers of initial errors homogeneity is the most likely final state. Presented patterns with more than 36 initial errors are attracted to the negative image of p'. If presented patterns are chosen randomly less than 5% will have 25 or fewer errors relative to one of the stored patterns or their negatives. From fig 2 most of these 95% with 25 or more errors will result in homogeneous final states. The basins of attraction of the stored patterns, although substantial, are small compared to the total number of possible presented patterns. The stored patterns are well

separated in concentration space and surrounded by large areas of nonrecognizable patterns.

Not all recalled patterns are error free, but the average number of errors is substantially reduced by the operation of the network. In fig. 3 is we plot the average number of errors in the final pattern against the number of errors in the presented pattern calculated from those final states in class 1. The pattern p' is almost always recalled perfectly if the presented pattern has less than 14 errors. For example, 99% of the presented patterns with 10 initial errors result in the recall of patterns in class 1 (fig 2), and of these recalled patterns the average number of errors relative to p' is 0.17, or an improvement of a factor of about 50. Only part of the presented patterns with large numbers of errors result in the recall of patterns in class 1, but for those in class 1 there is a substantial reduction in the number of errors. The network typically corrects most of the errors or rejects a pattern as nonrecognizable.

Smaller networks also possess pattern recognition abilities, albeit to a lesser extent than large networks. A network of 72 systems would be difficult to implement experimentally, but a network with 13 systems and 32 connections is feasible.[23] From eqn 3 we see that a network with two stored patterns has on the average only 1/4 of the possible connections as nonzero. Thus, a network of 13 systems with two stored patterns has an average of $13 \times 12/4 = 39$ connections. If we restrict the stored patterns to having no more than 7 systems in one state then the network has on the average about 33 connections and from these we use only sets of stored patterns that require no more than 32 connections. The results are summarized in table 1. The pattern p' is recalled almost perfectly if there is only one error in the presented pattern. For two initial errors, 94% of the presented patterns result in a recall of patterns in class 1 with an average of 0.37 errors, an improvement of a factor of about 5.

We show that simple nonlinear chemical systems coupled by linear mass transport can perform programmable information processing such as pattern recognition. The chemical network has many similarities and some differences with a neural network of the Little[15] or Hopfield[16–18] type: patterns

are stored in both the chemical and the Hopfield network by a Hebbian learning rule, but the connection weights (the k_{ij}) may have either sign in a Hopfield network; the chemical systems must be bistable, but the neurons in a Hopfield network are typically monostable; in both, stored patterns are stable steady states and are recalled when the network is initialized in their basins of attraction. In an electrical realization of a neural network,[17] the neurons are amplifiers, the connections are wires, the connections weights are resistors; their analogs in the chemical computer are the bistable reaction mechanisms, mass transfer, and the mass transfer rates. The chemical network shares many of the desirable features of neural network models: both are robust in the presence of noise, both retain some computational power when damaged, and in both the computational abilities are not strongly dependent on model parameters. Because the connection weights can be either positive or negative in the Hopfield network as well as in the chemical network described in refs. 2-4, they perform better than the network presented here. The chemical implementation of parallel computers given here and in refs. 2-4 provides a chemical basis of neural networks.

There are many biological reaction mechanisms and biological systems with multiple stationary states; mass transfer among compartments in biological systems is ubiquitous. These are the necessary components of the parallel computer presented here, and the components are at least available in living systems. The predictions for the simple chemical case are verifiable.

References and notes

[1]E. A. Nessholme and A. R. Leech, *Biochemistry for Medical Science* (Wiley, N.Y., 1983), p. 146.

[2]A. Hjelmfelt, E.D. Weinberger, and J. Ross, PNAS-USA **88**, 10983 (1991).

[3]A. Hjelmfelt, E.D. Weinberger, and J. Ross, PNAS-USA **89**, 383 (1992).

[4]A. Hjelmfelt and J. Ross, PNAS-USA **89**, 388 (1992).

[5]M. Okamoto, T. Sakai, and K. Hayashi, Biosystems **21**, 1 (1987).

[6]W. McCulloch and W. Pitts, Bull. Math. Biophys. **5**, 115 (1943).

[7]M. Minsky, *Computation: Finite and Infinite Machines* (Prentice-Hall, Englewood Cliffs, NJ, 1967)

[8]P. E. Rapp, J. Exp. Biol. **81**, 281 (1979).

[9]L. Wang and J. Ross, PNAS-USA **87**, 988 (1989), and references therein.

[10]R. Guttman, S. Lewis, and J. Rinzel, J. Physiol. **305** 377 (1980).

[11]P. De Kepper, I.R. Epstein, and K. Kustin, J. Amer. Chem. Soc. **103**,6121 (1981).

[12]T. Pifer, N. Ganapathisubramanian, and K. Showalter, J. Chem. Phys. **83**, 1101 (1985).

[13]The necessary constants, take from ref. 12, are: $k_A = 0.21 M^{-1} s^{-1}$, $k_B = 21000 M^{-2} s^{-1}$. We choose $X_0 = 2.0 \times 10^{-5} M$, $Y_0 = 7.1 \times 10^{-4} M$, $k = 7.18 \times 10^{-3} s^{-1}$. The multiplier λ is 0.00085 for fig. 1, 0.0013 for figs. 2 and 3, and 0.0023 for fig. 4.

[14]K.L.C. Hunt, J. Kottalam, M.D. Hatlee, and J. Ross, J. Chem. Phys. **96**, 7019 (1992).

[15]W. Little (1974), Math. Biosci. **19** 101 (1974).

[16]J.J. Hopfield, PNAS-USA **79**, 2554 (1982).

[17]J.J. Hopfield, PNAS-USA **81**, 3089 (1984).

[18]J.J. Hopfield and D. Tank, Sci. **233**, 625 (1986).

[19]D. O. Hebb, *The Organization of Behavior* (Wiley, NY, 1949).

[20]Since the stored patterns are random, roughly half the systems in pattern p' are in the same state

as system i in pattern p'. The expected values of k_{ij} in each case are

$$< k_{ij} > \;=\; \sum_{n=P/2-1}^{P-1} \binom{P-1}{n} \frac{(2(n+1)-P)}{2^{P-1}N} \qquad \text{if } [I^-]_i^{p'} = [I^-]_j^{p'} \tag{5}$$

$$< k_{ij} > \;=\; \sum_{n=P/2}^{P-1} \binom{P-1}{n} \frac{(2n-P)}{2^{P-1}N} \qquad \text{if } [I^-]_i^{p'} \neq [I^-]_j^{p'}. \tag{6}$$

Eqn. 5 is greater than eqn. 6 by a factor of $1/N$, and $\alpha(P)$ in eqn 4 is the value of the eqn. 6 multiplied by N.

[21]Stored patterns are generated randomly, and one of them is designated p'. A certain number of randomly chosen pixels are reversed (i.e. a system which should be in the high $[I^-]$ state is set to the low $[I^-]$ state or vice versa) in the pattern p' to generate the presented pattern.

[22]The average time for a stored pattern to emerge depends on the reaction kinetics, and the iodate-arsenous acid reaction is notoriously slow.

[23]As an example of what is experimentally feasible we cite the work of Laplante and Erneux (J.-P. Laplante and T. Erneux, J. Phys. Chem. **96**, 4931 (1992)) where they performed experiments with a system of 16 coupled bistable systems (the chlorite-iodide reaction in a continuous flow stirred tank reactor) with the possibility of 32 connections. We examine a network which may be implemented with this hardware.

[24]A. H. gratefully acknowledges financial support from the Alexander von Humboldt-Stiftung. J. R. gratefully acknowledges support from the National Science Foundation.

Figure Captions

Fig 1: Sample time series obtained from a chemical network with 36 coupled systems and three stored patterns. The x-axis gives the system number and the y-axis gives the time. The shading gives the I^- concentration with pure black and white representing the concentration of I^- in the high and low iodide states of an uncoupled system. Because systems in the high iodide state and the low iodide state are coupled in the network, the concentrations of iodide in the recalled pattern (at $t = 10$) do not match exactly those of a collection of isolated systems, and the white and black in the recalled pattern are not as intense as in the presented pattern (at $t = 0$) in which the systems have iodide concentrations set to those of uncoupled systems. The presented pattern has 10 errors relative to the stored pattern, and at $t = 10$ a stored pattern is recalled perfectly.

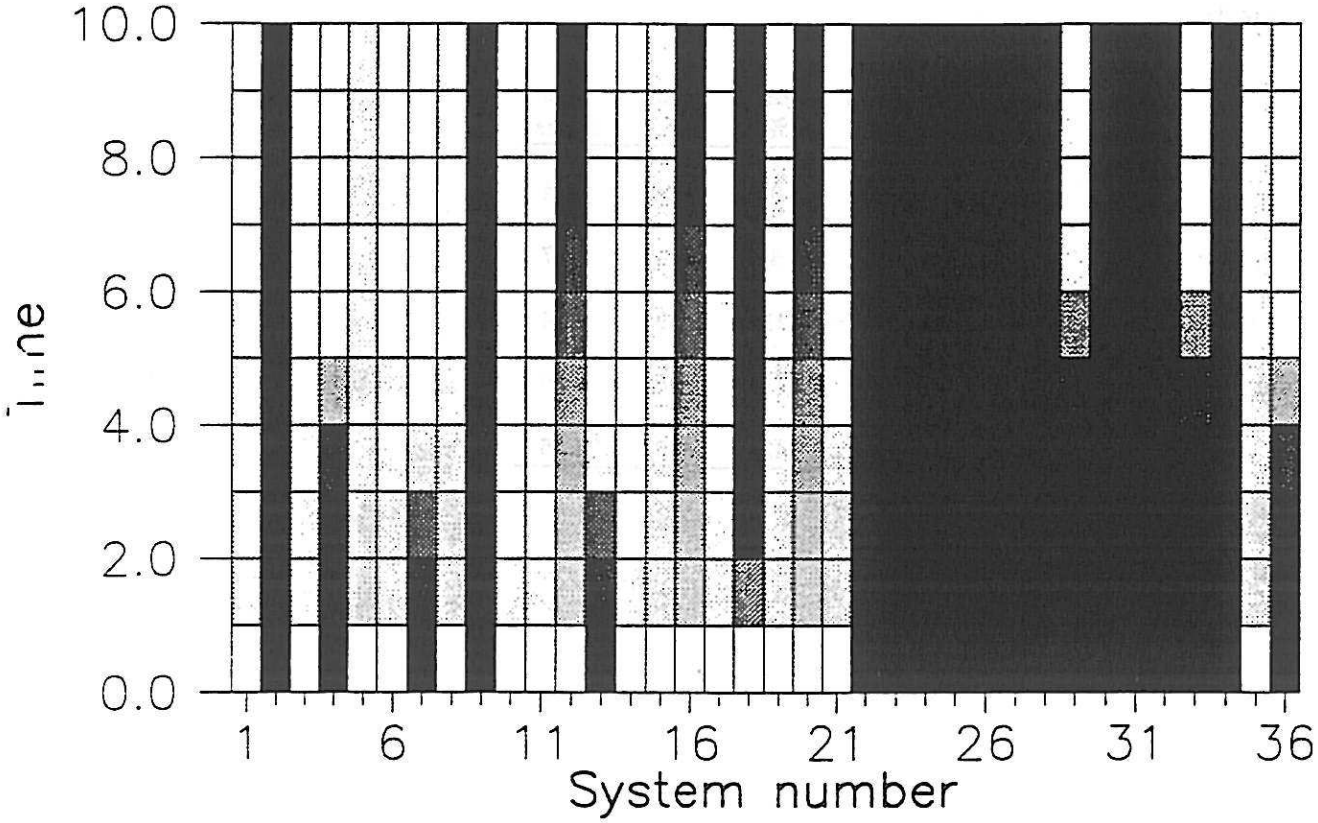
Fig 2: Percentage of final patterns which are homogeneous (class 3), most closely related to p' (class 1), or most closely related to the other stored patterns (class 2) against the number of errors, relative to p', in the presented pattern. The network has 72-systems and 3 stored patterns. The stored patterns are generated randomly, and 100 different sets of three stored patterns and a presented pattern were simulated for each number of initial errors.

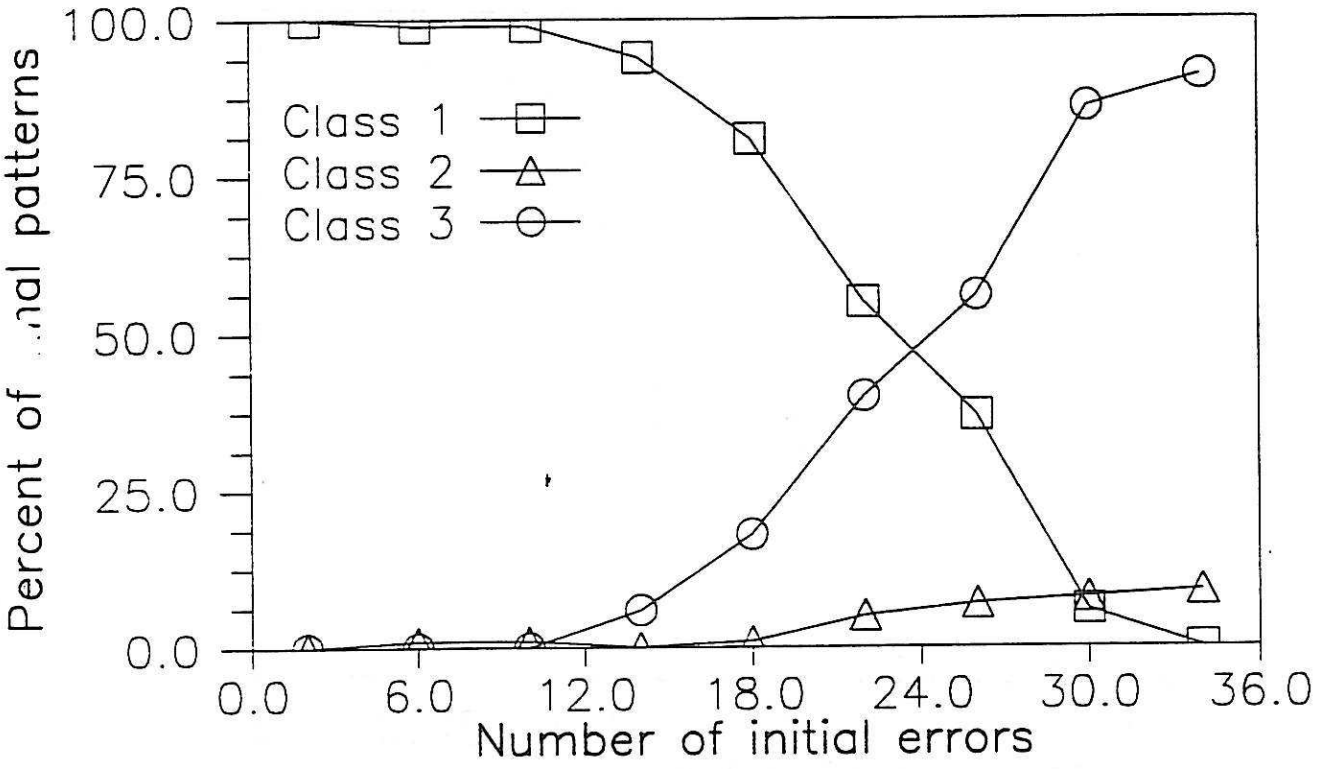
Fig 3: Average number of errors, relative to p', in the recalled patterns calculated from the final patterns in class 1 in fig. 2 against the number of errors in the presented patterns.

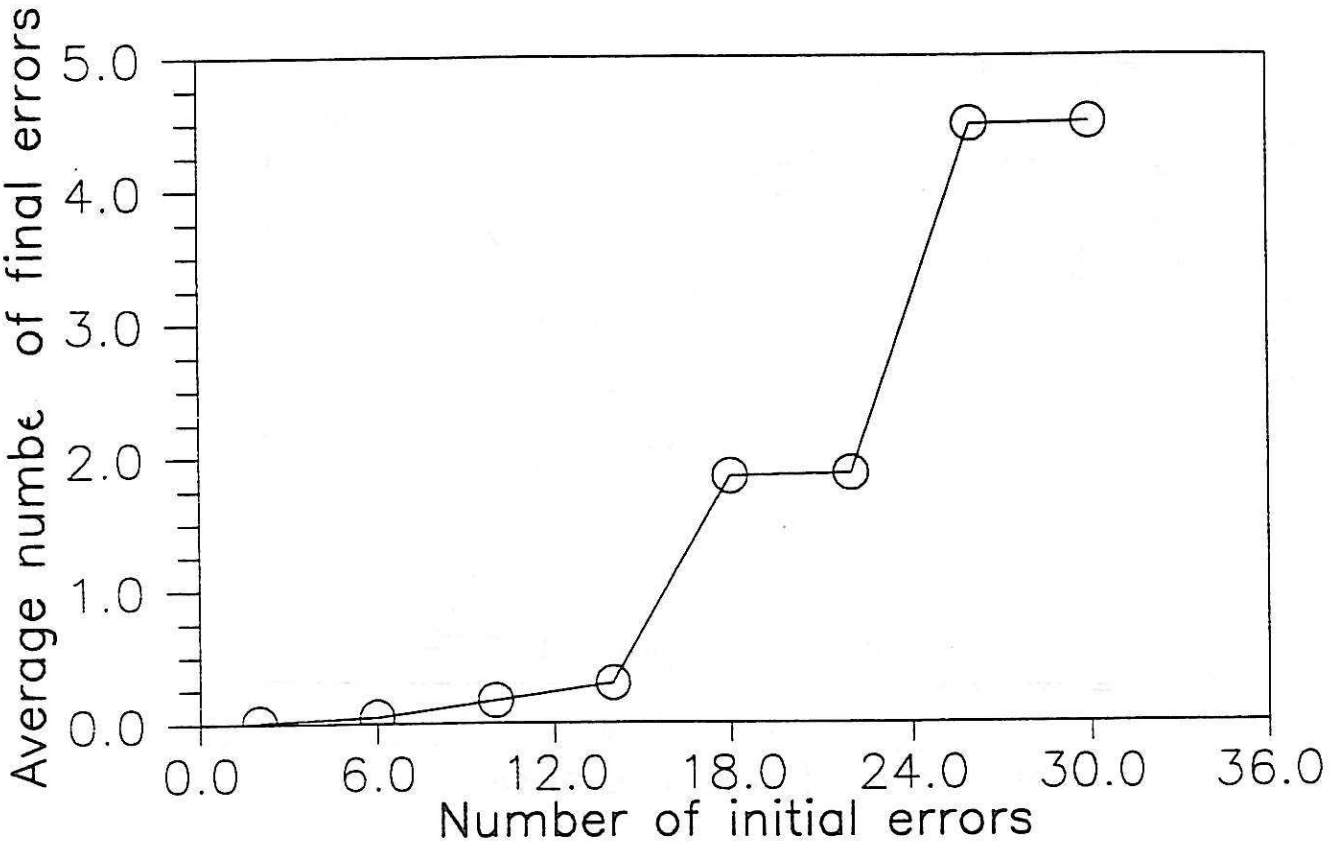
Table 1: Results for a network with 13 systems, a maximum of 32 connections, and two stored patterns with 7 systems in one state and 6 systems in the other state. The classes are as described in Fig. 2.

Table 1:

initial errors	class 3 (%)	class 2 (%)	class 1 (%)	final errors
1	0	0	100	0.04
2	0	6	94	0.37
3	0	21	79	0.84
4	3	41	56	1.79
5	8	42	50	3.36


10.0
8.0
6.0
4.0
2.0
0.0
1
6
11
16
21
26
31
36
System number


100.0
75.0
50.0
25.0
0.0
Percent of ..nal patterns
Class 1
Class 2
Class 3
0.0
6.0
12.0
18.0
24.0
30.0
36.0
Number of initial errors


Average numbe of final errors
5.0
4.0
3.0
2.0
1.0
0.0
0.0
6.0
12.0
18.0
24.0
30.0
36.0
Number of initial errors