TR－H－106

Eliminating spurious memories using a network of chaotic elements

Shin Ishii

1994. 10. 27

ATR人間情報通信研究所

〒619-02 京都府相楽郡精華町光台2-2　☎07749-5-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1011
Facsimile: +81-7749-5-1008

Eliminating spurious memories using a network of chaotic elements

Shin Ishii

ATR Human Information Processing Research Laboratories

2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

(TEL) +81 7749 5 1069 (FAX) +81 7749 5 1008

(E-mail) ishii@hip.atr.co.jp

Abstract

A Globally Coupled Map (GCM) model is a network of chaotic elements that are globally coupled with each other. We have already proposed an associative memory system based on the GCM, which has a better ability than the Hopfield network. This success is obtained through the mechanism that a network state can escape from spurious memories with its chaotic dynamics. Therefore, our approach is not to reduce spurious memories, rather, it is to escape from them.

In this paper, we propose a modified associative memory system, in which spurious memories are noticeably reduced. This is achieved by modifying the chaotic dynamics of the system, and not by modifying its learning rule. With this improvement, our system's memory capacity and the basin volume are expanded in a great deal. Some experimental results in comparison with those of a neural network employing a nonmonotonic output function are also shown.

Keywords: *choas, globally coupled map, associative memory, nonmonotonic dynamics, spurious memory*

1 Introduction

Recently, there have been many studies on artificial neural network models with nonequilibrium dynamics. They have been encouraged by recent biological experimental results of mammalian brains. In the cat visual cortex, for example, stimulus-specific synchronized oscillations have been reported by Eckhorn et al. (1988) and Gray and Singer (1989). Sakarda and Freeman (1987) reported that, in the rabbit olfactory bulb, limit cycle activities occur for perceptible specific odors but chaotic activities occur for novel odors.

Based on the above-mentioned biological results, the spatiotemporal complexity in recent nonequilibrium neural network models has mainly been attributed to the network's asymmetric connections, i.e., excitatory and inhibitory connections (Lie and Hopfield 1989; Yao and Freeman 1990; Grossberg and Somers 1991). A nonequilibrium neural network model was also proposed by Aihara et al. (1990); it was deduced from experiments with squid giant axons. This model's spatiotemporal complexity is not generated by the network structure, rather by the dynamics of each single neuron. Neural network models with asymmetric connections have also been investigated theoretically (Amari 1972a, b; Sompolinsky and Kanter 1986; Sompolinsky et al. 1988).

From a physical viewpoint, studies have been made on nonlinear coupled oscillators (Kuramoto 1991). The main interest of these studies has been the spatiotemporal complexity of certain physical systems such as spin-glasses. Kaneko proposed several models, which were originally "fuzzificated" versions of the Cellular Automaton model investigated by Wolfram (1984). Kaneko's models are based on coupled chaotic elements. Each element evolves in time according to the logistic map, and the couplings are of the nearest neighbor type (Kaneko 1984), or of the global coupling type (Kaneko 1990). The latter model is called "Globally Coupled Map (GCM) model", and many of its interesting characteristics such as glassy attractors have been reported.

All of the above-mentioned models intend to "mimic" biological neural network models or physical systems; further breakthroughs are needed to properly achieve this.

On the other hand, from a technical viewpoint, it is important to implement nonequilibrium information processing like in the human brains. For example, let's think about the associative memory (Kohonen 1977), which is a key technology utilized in many technical fields such as pattern recognition

and database retrieval. Hopfield (1984) proposed a neural network approach to the associative memory, in which, an association process corresponds to minimization of the network's Lyapunov function. In this sense, his network employs equilibrium dynamics. His model also employs the auto-correlation learning rule (Hebb 1949). Since his work, many associative memory systems based on nonequilibrium neural network models have been proposed, many of which employ auto-correlation or mutual correlation learning rules. For example, Nara et al. (1993) proposed an associative memory system based on an asymmetric neural network with a mutual correlation learning rule. Hayashi (1994) proposed a system based on excitatory and inhibitory connections with a recurrent neural network learning method. Adachi et al. (1993) proposed a system based on the model proposed by Aihara et al. (1990) with an auto-correlation learning rule.

Among these, we have already proposed an associative memory system (Ishii et al. 1994) based on Kaneko's GCM (Kaneko 1990). In our system, each element evolves in time according to a cubic map $f(x) = \alpha x^3 - \alpha x + x$ instead of the logistic map $f(x) = 1 - \alpha x^2$ employed in the original GCM. The learning rule employed in our system is of the auto-correlation type. However, in our system, both the memory capacity and the basin volume for each memory are larger than in the Hopfield network (Hopfield 1984) employing the same learning rule. This result indicates that the chaotic dynamics employed in our system is more efficient than the steepest descent like dynamics employed in the Hopfield network. Nevertheless, even in our system, spurious memories, i.e., the network's equilibrium points that do not correspond to any of the proper memories, do exist.

In this paper, we intend to eliminate spurious memories in our associative memory system only by modifying its dynamics. If we employ a strong learning method such as diagonal learning (Kohonen 1977), the system might be improved. However, our major interest lies in the dynamics. As a result, in our modified system, spurious memories are noticeably reduced, giving an even larger memory capacity and an even larger basin volume for each memory than the original system. Our research is inspired by Morita's work (Morita 1993), in which the Hopfield network was extremely improved by employing a nonmonotonic output function. Therefore, we compare our modified system with a variation of Morita's system experimentally. We give a somewhat intuitive interpretation as to why such improvement can be achieved in our new system.

2 Associative memory based on GCM

We have already proposed a modified GCM model called "S-GCM" (Ishii et al. 1994), which is designed for information processing applications. Our S-GCM employs the cubic map S-MAP instead of the logistic map employed in the original GCM. This modification makes it easy for each unit to represent one bit, i.e., -1 or 1. Our S-GCM has attractors called "cluster frozen attractors". Therefore, our S-GCM can easily represent binary spatial patterns as its attractors. We have also described several of S-GCM's characteristics (Ishii et al. 1994). They are as follows.

- Representation stability. The S-GCM falls into a cluster frozen attractor over a wide range of parameters, i.e., a representation of information is stable over parameters.
- Retrieval ability. In the S-GCM, represented information can be preserved or broken by controlling the system's parameter values.

These characteristics have shown that our S-GCM model can be applied to associative memory systems. In this section, we introduce our basic associative memory system based on the S-GCM (Ishii et al. 1994).

[System $\mathcal{CF}$]

$$x_i(t+1) = (1-\epsilon) f_i(x_i(t)) + \frac{\epsilon}{N} \sum_{j=1}^{N} f_j(x_j(t)) \tag{1}$$

$$f_i(x) = f(x;i) = \alpha_i x^3 - \alpha_i x + x \qquad x \in [-1,1] \tag{2}$$

In Eq. (1), $x_i(t)$ denotes the i-th unit's value at time t, and N is the number of units. Each unit's dynamics is almost entirely given by the cubic function S-MAP (2); the portion described as a summation in Eq. (1) is defined as feedback from the "mean-field". In this system, each unit's strength of chaos α_i is diverse, and evolves in time, like:

$$\alpha_i' = \alpha_i + (\alpha_i - \alpha_{min}) \tanh(\beta E_i') \tag{3}$$

$$E_i' = -x_i \sum_{j=1}^{N} \sigma_{ij} x_j \tag{4}$$

$$\sigma_{ij} = \frac{1}{N}\sum_{k=1}^{M} \xi_i^k \xi_j^k \tag{5}$$

where $\xi_i^k \in \{1, -1\}$ denotes the i-th bit of the k-th memorized pattern and M is the number of memorized patterns. That is, $\{\xi^1, ..., \xi^M | \xi^k \in \{1, -1\}^N\}$ is a set of memorized binary patterns. β is a constant value. The learning method defined by equation (5) is of the auto-correlation type.

Let us show an association process. When $\mathcal{CF}$ with 100 units memorizes five 100-bit binary patterns "A","J","P","T", and "S", and the input binary pattern is a 35% reversed pattern of "A", it associates "A" after scores of transitions. Figure 1 shows this association process. In this figure, highly chaotic motions are observed at the early association stage. As time elapses, these motions become quiet, and the association is completed successfully when the system falls into a 4-cluster frozen attractor.

It is well known that the memory capacity of the simplest Hopfield network can theoretically be estimated at $0.138N$ (Amit et al. 1987). Experimental results have shown that our system has a larger memory capacity, and it is estimated at $0.186N$ (Ishii et al. 1994). In the following, we compare our system with the Hopfield network in terms of "success rate", which indicates how successfully the network can associate a target pattern, when the distance of initial states is known. Each network incorporates 100 units and memorizes five alphabet patterns. The distance of an initial state $\mathbf{x}(0)$ from a target pattern ξ is determined using initial overlap $ol(\mathbf{y}(0), \xi)$. Here, the overlap of a state $\mathbf{x}$ from a target pattern ξ is given by

$$ol(\mathbf{y}, \xi) = \frac{1}{N}\sum_{i=1}^{N} y_i \xi_i \tag{6}$$

where $\mathbf{y}$ is a binary vector derived from the state vector $\mathbf{x}$. Figure 2(a) and 2(b) show the results obtained for the target pattern "A" and "T", respectively. These figures indicate that our system has a higher success rate, i.e., a larger basin volume for each memorized pattern than the Hopfield network.

3 System to eliminate spurious memories

3.1 Modification description

Our system has a better ability as an associative memory system than the Hopfield network, although the two models employ the same auto-correlation learning rule. The reason is considered as follows; in our system, with its chaotic dynamics in an early association stage, a network state can escape from spurious memories. On the other hand, in the Hopfield network, with its steepest descent like dynamics, a network state can not escape from a spurious memory, i.e., a local minimum of the Lyapunov function. Nevertheless, even in our system, spurious memories exist, which inhibit the system from having an even larger memory capacity or an even larger basin volume.

Now we propose a new system $\mathcal{CF}^+$, which is a dynamical system defined by Eqs. (1),(2),(5), and the following equations.

$$\alpha_i' = \alpha_{mid} + (\alpha_{mid} - \alpha_{min}) \tanh(\beta E_i') \tag{7}$$

$$E_i' = -x_i U_i' \tag{8}$$

$$U_i' = U_i + \sum_{j=1}^{N} \sigma_{ij} O(x_j; \alpha) - \frac{U_i}{\tau'} \tag{9}$$

$$\tau' = \tau^* \tau \tag{10}$$

$$O(x;\alpha) = \begin{cases} x & \text{if } \alpha > \alpha_u \\ 0 & \text{if } \alpha < \alpha_l \\ \frac{\alpha - \alpha_l}{\alpha_u - \alpha_l} x & \text{otherwise} \end{cases} \tag{11}$$

where α_i' denotes a new value of the i-th unit's strength of chaos. The domain of α is set to be between two parameters $\alpha_{min} \leq \alpha \leq \alpha_{max}$, and the other two parameter values are determined to be $\alpha_{min} < \alpha_l < \alpha_u < \alpha_{max}$. τ is a parameter increasing gradually from 1 in time, i.e., its initial value is 1 and $\tau^* \geq 1$. Here, the evolution described in Eqs. (7),(8),(9),(10), and (11) is done once in every 32 time steps, i.e., at $t = 32, 64, 96...$ In this sense, we describe this evolution without using t.

Let us show the meaning of this modification, in short. Eqs. (3) and (4) have been changed into Eqs. (7),(8),(9),(10), and (11). However, Eq. (7) has almost the same meaning as Eq. (3). This alteration is only for an improvement of the system's ability, in fact. Additionally, if the output

function $O(x;\alpha)$ is an identical function, Eq. (9) is almost equivalent to a Euler difference equation of the Hopfield network's internal potential

$$\frac{dU}{dt} = -\frac{U}{\tau} + \sum_{j=1}^{N} \sigma_{ij} x_j. \tag{12}$$

If τ is constant 1, Eq. (9) resembles the definition of an internal potential (see Eq. (13)) in the mean field theory (Peterson and Anderson 1987), in which the network converges to a minimum of the network's Lyapunov function more rapidly than the Hopfield network (Sato 1994). On the other hand, in the Hopfield network employing Eq. (12), the network's equilibrium points become close to a corner of the domain hypercube as τ becomes large. From this observation in resembling models, we consider that in our system, the system's mode becomes stable as τ increases, which contributes to suppressing chaotic motions in a later association stage. Further discussion will be done in the next subsection.

Therefore, the most important modification is the existence of the output function O. Figure 3 shows $O(x)/x$ against α, schematically. As this figure shows, O is an identical function, when α is relatively large, i.e., the unit's state is very chaotic. On the other hand, O always returns 0, when α is relatively small, i.e., the unit's state is stable. Namely, O suppresses the output of a unit having a low α value. In section 5, we will discuss the meaning of this output function.

3.2 Movement

Let us show some examples of association processes.

Figure 4(a) and (b) show time series of overlap in the modified system, when the initial overlap is set to be various values. When the overlap value is equal to 1, the system state is equivalent to the target pattern. On the other hand, when the overlap value is around 0, the state has no correlation with the target. In Figure 4(a) and 4(b), the $\mathcal{CF}^+$ with $N = 256$ units memorizes $M = 32$ random binary patterns ($r = M/N = 0.125$), and memorizes $M = 64$ patterns ($r = 0.250$), respectively. The other parameters are set as $\alpha_{max} = 4.0, \alpha_{mid} = 3.85, \alpha_{min} = 3.40, \alpha_u = 3.70, \alpha_l = 3.45, \epsilon = 0.10, \tau^* = 1.02$, and $\beta = 2.0$, and initial values are set as $\alpha_i = 3.50, U_i = 0.0$, and $\tau = 1.0$. Initial values for x are set as the initial overlap with the target being the specific value. The detailed manner for this coding is described in the previous paper

(Ishii et al. 1994). As Figure 4(a) shows, when the number of memorized patterns is relatively small, the system can easily associate a target even from initial states that are far from the target, and the time interval needed for each association is very short. On the other hand, as Figure 4(b) shows, when the number of memorized patterns is relatively large, the basin of attraction becomes small. In this case, if its initial state is relatively close to the target, the system can associate the target; if its initial state is far from the target, the system fails to associate it. The critical distance is about 0.56 (72 in Hamming distance) in this figure's case, and it turns out to be large enough.

For comparison, Figure 5(a) and (b) show time series of overlap in the original system $\mathcal{CF}$. Parameters are set as $\alpha_{max} = 4.0, \alpha_{min} = 3.40, \epsilon = 0.10$, and $\beta = 0.10$, and initial values are set as $\alpha_i = 3.50$. Figure 5(a) is for a relatively small number of memories ($r = 0.125$), and Figure 5(b) is for a relatively large number of memories ($r = 0.250$). As mentioned in section 2, we estimate the memory capacity of the $\mathcal{CF}$ to be about $0.186N$, which is smaller than the number of memories in the case Figure 5(b) shows. However, actually, with such a large number of memories, basins of attraction can exist, although they are small enough. Nevertheless, in our old system the memory capacity and the basin volume are much smaller than in the modified system. Furthermore, the time interval needed for each association process is much longer than in the modified system. The association time depends on the parameter β, in fact. If β is large, the association time becomes small. However, such a system almost always fails to make a proper association.

Figure 6 shows a time series of the values of all units in $\mathcal{CF}^+$. Let's compare it with an association process of the $\mathcal{CF}$ shown in Figure 1. In our old system, when the system is able to make a proper association, the state becomes stable in a 4-cluster frozen attractor. On the other hand, in our new system, even when the system is able to successfully associate a memorized pattern, some chaotic movements remain. Since such chaotic movements are "localized" in two clusters and separated from each other as Figure 6 shows, we can extract a binary representation from the state by a coarse-viewing method.

If τ is constant 1, the system $\mathcal{CF}^+$ becomes more unstable. Figure 7(a) and (b) show time series of overlap, when τ^* is set to be 1. As these figures show, the system's "global" chaotic motions remain, and the state often becomes far from the target because of its chaotic motions. If we utilize this phenomenon, it might be possible to make some chaotic associative memory

systems. Accordingly, the gradual enlarging of τ described by Eq. (10) contributes to stabilizing an association process of our new system, as mentioned in the previous subsection.

4 Evaluation

In this section, we evaluate several systems including our new system $\mathcal{CF}^+$. They are:

- Hopfield network in the mean field theory equation (MFT)
- MFT employing a nonmonotonic output function (NM)
- Our old system ($\mathcal{CF}$)
- Our new system ($\mathcal{CF}^+$)

A rough description of the MFT is as follows.

$$x_i(t+1) = \tanh(\beta u_i(t+1)) = \tanh\left(\beta \sum_j \sigma_{ij} x_j(t)\right) \tag{13}$$

where parameter β is set to be 2.0 in the following experiments. Here, we choose the MFT (Peterson and Anderson 1987) instead of the continuous-time version, which is Hopfield's original model (Hopfield 1984). The reason is that the MFT is easier to treat and more efficient than the continuous-time model (Sato 1994). Moreover, we choose the MFT equation with a nonmonotonic output function instead of Morita's original continuous-time model or his partial reverse method (Morita 1993) for the same reason. A rough description of the NM is as follows.

$$x_i(t+1) = g(\sum_j \sigma_{ij} x_j(t)) \tag{14}$$

$$g(u) = \left(\frac{1-\exp(-cu)}{1+\exp(-cu)}\right)\left(\frac{1+\kappa\exp(c'(|u|-h))}{1+\exp(c'(|u|-h))}\right) \tag{15}$$

Since the Hopfield network employing a nonmonotonic output function g does not have a Lyapunov function in general, it is not proper to regard this model as a mean field theory equation. However, we consider that this model

works as well as Morita's original model. Parameters are set as $c = 50, c' = 15, h = 0.5$, and $\kappa = 0$. All of the parameter values except κ are the same as Morita's (Morita 1993). In Morita's original paper, $\kappa = -1$ was used. In Figure 8, the association success rate of the NM is shown with various κ values, when each initial state is a 1-bit reversed pattern of a target, with the number of units $N = 128$ and the number of memorized patterns $M = 32$. From observing this figure and for analysis simplicity, we determine $\kappa = 0$. Figure 9 shows the function shape of g with these parameter values.

Parameter values of the $\mathcal{CF}$ and the $\mathcal{CF}^+$ used in the following experiments are the same as those shown above.

4.1 Memory capacity

In this subsection, we show an experimental result for the memory capacity of the four systems. Here, "memory capacity" means how many random patterns the network of N elements can store so that every memory has its own basin of attraction.

Figure 10 shows the simulation results. Each network succeeds in memorizing M random patterns, if the probability of bit-wise flips after some transient period is less than 1.5%, when the initial state is set to be one of the memorized patterns. According to a theoretical analysis (Amit et al. 1987) of the simplest Hopfield network (Hopfield 1982), if the probability of bit-wise flips exceeds 1.5% before reaching an attractor, all memories become useless, i.e., a phase transition occurs. In our old system, new system, and the NM, we can not observe such a distinguished phase transition. In these systems, memories become useless much more slowly as the number of memories increases. Figure 11 shows the rate of bit-wise flips in $\mathcal{CF}^+$ after some transient period, with $N = 512$ and various numbers of memories. Moreover, in these three systems, there sometimes remain some chaotic movements. Therefore, this probability value is not so meaningful to determine the memory capacity; it is just a criterion for comparison.

As Figure 10 shows, in each of the four systems, the memory capacity (rate) does not depend on the number of units N. In the MFT, it is estimated at about $0.125N$, which agrees well with the theoretical result. The $\mathcal{CF}$ has a much larger memory capacity of $0.186N$, which is about 50% larger than that of the MFT. The NM has an even larger one of $0.216N$, which is 73% larger than that of the MFT. Moreover, in our modified system, it is about $0.273N$, which is more than twice that of the MFT, and about 30%

larger than that of the old system. According to Morita's estimation (Morita 1993), the Hopfield network employing a nonmonotonic output function has about a $0.32N$ memory capacity, which is much larger than our experimental result. This difference is maybe due to the difference of the memory capacity criterion. Actually, as mentioned in the previous section, in our old system, even when the number of memorized patterns exceeds the capacity, basins of attraction can exist, although they are narrow enough. This is also the case in the NM and the new system.

4.2 Basin volume

In this subsection, we compare the four systems in terms of the basin volume.

Each of Figures 12(a),(b),(c), and (d) show the success rate of the four systems. "Success rate" means how successfully the network can associate a target pattern, when the distance of initial states is known as an overlap value. Each initial state is taken to have a smaller distance from the target than from any of the other memories. Figure 12(a) is for $N = 512, M = 64$ $(r = 0.125)$. In Figure 12(b), $N = 512, r = 0.250$; in Figure 12(c), $N = 256, r = 0.125$; and in Figure 12(d), $N = 256, r = 0.250$. According to our estimation in the previous subsection, $r = 0.250$ exceeds the memory capacity of the $\mathcal{CF}$ and the NM. However, actually, some small basin of attraction exists near a target even in those systems, as Figure 12(b) and (d) show.

As these figures show, the success rate of our new system is larger than that of the MFT, $\mathcal{CF}$, and the NM in every case. With a relatively small number of memories, our new system can associate a target pattern from very far initial states with a high success rate. Therefore, we conjecture that the basin volume in our new system is much larger than in any of the three other systems.

4.3 Total basin volume

In this subsection, we investigate the total basin volume of the four systems. Here, "total basin volume" means how large in total are the basins of attraction for all proper memories in the whole domain area. Or, it means the probability that a binary vector chosen at random belongs to a basin of attraction of any proper memory. Table 1 shows the results in percent with several conditions.

Table 1

	MFT	$\mathcal{CF}$	NM	$\mathcal{CF}^+$
$N = 64, r = 0.125$	14.3	3.5	99.8	96.5
$N = 256, r = 0.125$	0.0	0.0	81.4	23.4
$N = 256, r = 0.250$	0.0	0.0	0.4	0.3

Table 1 shows that the $\mathcal{CF}$ has a smaller total basin volume than the MFT, and that the $\mathcal{CF}^+$ has a smaller one than the NM. Additionally, as the number of units increases, the total basin volume becomes small in the four systems.

As the number of units increases, the domain area becomes huge, in an exponential order. On the other hand, the number of available memories increases much more slowly, in a linear order of N. Therefore, as N increases, the density of available memories becomes small. We consider this to be the reason why the total basin volume rate is small in the four systems with large N value. In the three systems except the MFT, some chaotic motions may remain, when the system fails to make a proper association. However, in most failure cases in Table 1, the system becomes stable at an illegal state, i.e., a spurious memory. Namely, Table 1 shows that in the $\mathcal{CF}^+$, the basin volume for spurious memories is larger than in the NM. This result indicates that there are more spurious memories in the $\mathcal{CF}^+$ than in the NM. It seems that it contradicts the result shown in the previous subsection. Now, let us briefly discuss this phenomenon.

According to a neuro-dynamics analysis by Amari and Maginu (1988), in the simplest Hopfield network, basins of attraction have a strange shape, i.e., like a starfish. Therefore, even if the initial state is very close to a target, the network may fail to associate it. The above-mentioned contradictory result suggests that this is also the case in the NM. In the NM, the probability that a random initial state belongs to a basin of a proper memory is larger than in the $\mathcal{CF}^+$. However, because of the starfish like shape of basins, it often fails to make a proper association even when it starts at a close initial state from a target, which lowers its success rate. Actually, the NM often associates an unexpected memory that is not the closest memory to its initial state[1]. On the other hand, we consider that in the $\mathcal{CF}^+$, basins of attraction have a

[1]Notice that in the experiments in section 4.2, all initial states are taken to be nearest from a target memory. Therefore, even though the NM can associate another memory besides the target, the result is an unexpected one.

sphere like shape. Therefore, it can associate a proper memory with a high success rate from an initial state with a relatively small distance from the target, although there are many spurious memories in its domain area. In this sense, we can say that in our $\mathcal{CF}^+$, the "effective" basin volume is larger than in the NM.

Accordingly, we can conclude as follows. In our new system, the memory capacity and the effective basin volume are much larger than in the old system. The success rate is also much higher. This is because in our new system, the dynamics is so improved that spurious memories located close to proper memories are almost entirely wiped out.

5 Discussion

In this section, we discuss the reason why such an improvement is achieved in our modified system. For this discussion, first we show an intuitive interpretation of why spurious memories occur in the MFT.

In the MFT, each unit evolves in time according to Eq. (13). When an internal potential u_i is large, we say the i-th unit is "obstinate". An obstinate unit is not easily influenced in its output by fluctuations in the values of other units. Because of the output saturation of the sigmoidal output function (see Figure 13) of the MFT, even if there is some small change in the summation $\sum_j \sigma_{ij} x_j$, the output is not altered. This "obstinacy" tends to be generated when some part of a state matches one of the memorized patterns. On the other hand, with the sigmoidal output function, the output value of such an obstinate unit is large, i.e., the unit's influence on the other units is large. As for a state at an early association stage, when one part becomes obstinate by matching a memorized pattern, and another part also becomes obstinate by matching another memorized pattern, their value does not change each other, and they attract the remaining part with their large output. In such a case, the state becomes stable at a mixed pattern of two or more memorized patterns. This is a so-called spurious memory in the MFT. In the case of the Hopfield network, this situation can be more serious, since an internal potential might take a large value, and at that time, the unit would not be able to change in its value any more.

Accordingly, spurious memories occur because of the strong influence of the units that are obstinate , i.e., too stubborn to be influenced by the other units. If we want to reduce spurious memories, one way is to lower the in-

fluence of obstinate units on the other units. In the NM, we employ the output function g shown in Figure 9, which does this control. With this control, the NM can successfully eliminate spurious memories as Table 1 shows. However, an association result may differ from the expected pattern, since using the nonmonotonic output function g is nothing but ignoring some units' output. This interpretation suggests the reason why basins of attraction take a strange shape in the NM, as mentioned in the previous section. We consider that this description also holds in the Hopfield network employing a nonmonotonic output function, although Morita (1993) showed another intuitive interpretation for his system's success.

In the $\mathcal{CF}$, parameter alpha evolves in time according to Eqs. (3) and (4). When E_i is small, α_i becomes small, and that means the unit becomes stable in its value. In such a case, we say the unit is obstinate. When a unit becomes obstinate, the unit is not easily influenced in its value by other units, since the unit's chaos is weak. On the other hand, such an obstinate unit strongly influences the other units through α alteration defined by Eqs. (3) and (4). This is almost the same situation with the MFT. An obstinate unit would not easily be influenced in its value, although it strongly influences the other units' value. This is the reason why spurious memories occur in our old system $\mathcal{CF}$.

Therefore, if we lower the influence of obstinate units on the other units, spurious memories can be eliminated. Hence, we introduce the function O to suppress the influence of a unit being obstinate, i.e., having a low α value. With this control, our new system $\mathcal{CF}^+$ can noticeably reduce spurious memories.

We do not know the reason why the basin shape in our $\mathcal{CF}^+$ is better than in the NM. However, our conjecture is as follows. In the $\mathcal{CF}^+$, the control done by the function O affects a unit's value x in an indirect manner. O controls the parameter α through Eqs. (7),(8), and (9), and α alters x through Eqs. (1) and (2). This indirect control achieves a partial "annealing" like mechanism, which inhibits the basins from becoming a starfish like shape. This is not the case in the NM with its rather direct control by the nonmonotonic output function.

6 Conclusion

In our old associative memory system $\mathcal{CF}$, both the memory capacity and the basin volume are larger than in the Hopfield network employing the same learning rule. This success is obtained through the mechanism that a network state can often escape from spurious memories with its chaotic dynamics. However, even in our old system, spurious memories do exist.

In this paper, we give an intuitive interpretation as to why spurious memories occur both in the Hopfield network and in our old system, in terms of the "obstinacy" of a unit. With this observation, we consider that if we lower the influence of an obstinate unit on the other units, we can reduce spurious memories. Hence, we propose a new modified associative memory system, which does this control through controlling the strength of chaos. With this control, our new system can almost entirely wipe out spurious memories located close to proper memories. As a result, our new system has an even larger memory capacity and an even larger "effective" basin volume than not only the old system but also the MFT employing a nonmonotonic output function.

Experimental results indicate that our new system's success rate is very high, even when the number of memorized patterns is relatively large. This success is more noticeable than in the MFT with a nonmonotonic output function. Accordingly, the system can generally associate a proper memory, which is actually the most important ability in associative memory systems from a technical viewpoint.

References

[1] Adachi, M., Aihara, K., & Kotani, M. (1993). "An analysis of associative memory dynamics with a chaotic neural network." *Proceedings of the International Symposium on Nonlinear Theory and its Applications* (Hawaii, Dec., 1993), 1169-1172.

[2] Amari, S. (1972a). "Characteristics of random nets of analog neuron-like elements." *IEEE Transactions on System, Man and Cybernetics*, **SMC-2**, 643-657.

[3] Amari, S. (1972b). "Learning and patterns and pattern sequences by self-organizing nets of threshold elements." *IEEE Transactions on Computers*, **C-12**, 1197-1206.

[4] Amari, S. & Maginu, K. (1988). "Statistical neurodynamics of associative memory." *Neural Networks*, **1**, 63-73.

[5] Amit, D.J., Gutfreund, H., & Sompolinsky, H. (1987). "Statistical mechanics of neural networks near saturation." *Annals of Physics*, **173**, 30-67.

[6] Aihara, K., Takabe, T., & Toyoda, M. (1990). "Chaotic neural networks." *Physics Letters A*, **144**, 333-340.

[7] Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitboeck, H.J. (1988). "Coherent oscillations: A mechanism of feature linking in the visual cortex? Multiple electrode and correlation analysis in the cat." *Biological Cybernetics*, **60**, 121-130.

[8] Gray, C.M., & Singer, W. (1989). "Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex." *Proceedings of the National Academy of Science, USA*, **86**, 1689-1702.

[9] Grossberg, S., & Somers, D. (1991). "Synchronized oscillations during cooperative feature linking in a cortical model of visual perception." *Neural Networks*, **4**, 453-466.

[10] Hayashi, Y. (1994). "Oscillatory neural network and learning of continuous transformed patterns." *Neural Networks*, **7**, 219-231.

[11] Hebb, D.O. (1949). *The Organization of Behavior*. Wiley, New York.

[12] Hopfield, J.J. (1982). "Neural networks and physical systems with emergent collective computational abilities." *Proceedings of the National Academy of Sciences, USA*, **79**, 2554-2558.

[13] Hopfield, J.J. (1984). "Neurons with graded response have collective computational properties like those of two-state neurons." *Proceedings of the National Academy of Sciences, USA*, **81**, 3088-3092.

[14] Ishii, S., Fukumizu, K., & Watanabe, S. (1994). "A network of chaotic elements for information processing." *submitted to Neural Networks.*

[15] Kaneko, K. (1984). "Period-doubling of kink-antikink patterns, quasiperiodicity in antiferro-like structures and spatial intermittency in coupled logistic lattice." *Progress of Theoretical Physics*, **72**, 480-486.

[16] Kaneko, K. (1990). "Clustering, coding, switching, hierarchical ordering, and control in a network of chaotic elements." *Physica D*, **41**, 137-172.

[17] Kohonen, T. (1977). *Associative Memory - A System-Theoretical Approach*, Springer-Verlag, Berlin Heidelberg.

[18] Kuramoto, Y. (1991). "Collective synchronization of pulse-coupled oscillators and excitable units." *Physica D*, **50**, 15-30.

[19] Lie, Z., & Hopfield, J.J. (1989). "Modeling the olfactory bulb and its neural oscillatory processing." *Biological Cybernetics*, **61**, 379-392.

[20] Morita, M. (1993). "Associative memory with nonmonotone dynamics." *Neural Networks*, **6**, 115-126.

[21] Nara, S., Davis, P., & Totsuji, H. (1993). "Memory search using complex dynamics in a recurrent neural network model." *Neural Networks*, **6**, 963-973.

[22] Peterson, C., & Anderson, J.R. (1987). "A mean field theory learning algorithm for neural networks." *Complex Systems*, **1**, 995-1019.

[23] Sakarda, C.A., & Freeman, W.J. (1987). "How brains make chaos in order to make sense of the world." *Behavioral and Brain Sciences*, **10**, 161-195.

[24] Sato, M. (1994). "The asynchronous MFT equation converges faster than the Hopfield network." *submitted to Neural Computation.*

[25] Sompolinsky, H., & Kanter, I. (1986). "Temporal association in asymmetric neural networks." *Physical Review Letters*, **57**, 2861-2864.

[26] Sompolinsky, H., Crisanti, A., & Sommers, H.J. (1988). "Chaos in random neural networks." *Physical Review Letters*, **61**, 259-262.

[27] Yao, Y., & Freeman, W.J. (1990). "Model of biological pattern recognition with spatially chaotic dynamics." *Neural Networks*, **3**, 153-170.

[28] Wolfram, S. (1984). "Cellular automata as models of complexity." *Nature*, **311**, 419-424.

Figure 1
An association process of the $\mathcal{CF}$, with 100 units and 5 alphabet memories "A","J","P","T", and "S".
The upper figure : The initial pattern is a 35% reversed pattern of "A". As time elapses, the system state becomes close to the target pattern "A", and finally the system is able to successfully associate the target.
The lower figure : A time series of all the units plotted every 32 time steps. The abscissa denotes the transition time ($t = 32, 64, 96, ..., 32 \times 50$) and the ordinate denotes the units' value. This graph corresponds to the binary spatial patterns shown in the upper figure.

Figure 2
Association success rate of the $\mathcal{CF}$ (solid line) and the Hopfield network (dotted line), when the initial overlap is set to be various values. Each network incorporates 100 units and memorizes five alphabet patterns.
(a) The target is the pattern "A". (b) The target is the pattern "T".

Figure 3
The function shape of the function O. O is an identical function, when α is relatively large. O always returns 0 , when α is relatively small, Namely, O suppresses the output of a unit having a low α value.

Figure 4
Time series of overlap in the $\mathcal{CF}^+$. $\alpha_{max} = 4.0, \alpha_{mid} = 3.85, \alpha_{min} = 3.40, \alpha_u = 3.70, \alpha_l = 3.45, \epsilon = 0.10, \tau^* = 1.02$, and $\beta = 2.0$. In addition, initial values $\alpha_i = 3.50, U_i = 0.0$, and $\tau = 1.0$. The abscissa denotes the time ($t = 32, 64, 96, ..., 32 \times 100$).
(a) The number of memories is relatively small, $N = 256, r = 0.125$. The system can easily associate a target even from a far initial state in a short time period.
(b) The number of memories is relatively large, $N = 256, r = 0.250$. When the initial state is relatively close to a target, the system can associate it. When the initial state is far from a target, the system fails to associate it.

Figure 5
Time series of overlap in the $\mathcal{CF}$. $\alpha_{max} = 4.0, \alpha_{min} = 3.40, \epsilon = 0.10$, and $\beta = 0.10$. In addition, initial values $\alpha_i = 3.50$. The abscissa denotes the time ($t = 32, 64, 96, ..., 32 \times 500$).
(a) The number of memories is relatively small, $N = 256, r = 0.125$. When the initial state is very close to a target, the system can associate it, although it takes a long time period to complete the association.
(b) The number of memories is relatively large, $N = 256, r = 0.250$. Although this memory number exceeds the memory capacity of $\mathcal{CF}$, actually, the system can associate a target by starting from some initial states.

Figure 6
A time series of the values of all units in the $\mathcal{CF}^+$ plotted every 32 time steps. The abscissa denotes the time ($t = 32, 64, 96, ..., 32 \times 100$). Even when the system can successfully associate a memorized pattern, some chaotic movements remain. However, the chaotic motions tend to disappear, helped by the τ control defined by Eq. (10).

Figure 7
Time series of overlap in the $\mathcal{CF}^+$, with τ set to be constant 1. All parameter values are the same as in Figure 4(a) and (b) except τ^*. The abscissa denotes the time ($t = 32, 64, 96, ..., 32 \times 100$). Comparing with Figure 4(a) and (b), in this system, "global" chaotic motions remain, which do harm to the making of a proper association.
(a) $N = 256, r = 0.125$. (b) $N = 256, r = 0.250$.

Figure 8
The association success rate of the NM with various κ values. Each initial state is a 1-bit reversed pattern of a target, with $N = 128$ and $M = 32$. Notice that if $\kappa = 1$, the NM becomes exactly equivalent to the MFT. In a mean field theory approach, $\kappa = -1$, which is equal to the value used by Morita, is not a good parameter value.

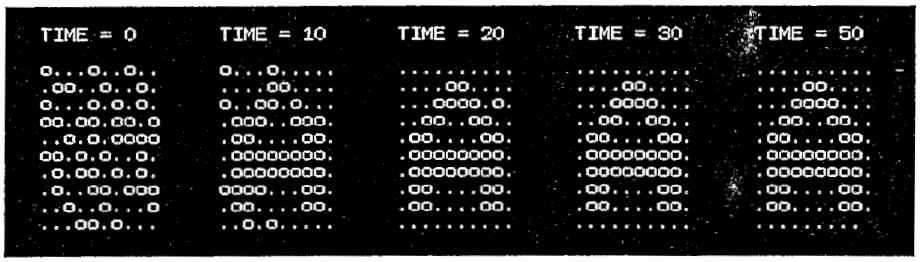
Figure 9
The function shape of the nonmonotonic output function used in the NM system. Parameter values are $c = 50, c' = 15, h = 0.5$, and $\kappa = 0$. Using this output function is nothing but ignoring the output of units having a large internal potential.

Figure 10
The memory capacity of the MFT, the $\mathcal{CF}$, the NM, and the $\mathcal{CF}^+$, with various number of units. It seems that the capacity does not depend on the number of units. The estimated capacity values are, $0.125N$ for the MFT, $0.186N$ for the $\mathcal{CF}$, $0.216N$ for the NM, and $0.273N$ for the $\mathcal{CF}^+$.

Figure 11
The rate of bit-wise flips in $\mathcal{CF}^+$ after some transient period, with $N = 512$ and various numbers of memories. This figure does not show a distinguished phase transition, which is observed in the Hopfield network.

Figure 12
The success rate of the MFT, the $\mathcal{CF}$, the NM, and the $\mathcal{CF}^+$, when the distance of initial states is known as an overlap value.
(a) $N = 512, r = 0.125$. Our $\mathcal{CF}^+$ can associate a target in 100% even from far initial states. (b) $N = 512, r = 0.250$. Our $\mathcal{CF}^+$ has a large effective basin volume, even with a relatively large number of memories. It seems that the $\mathcal{CF}$ and the NM also have a small basin volume, although the number of memories in this figure case exceeds their capacity. (c) $N = 256, r = 0.125$. (d) $N = 256, r = 0.250$.

Figure 13
Schematic figure of the sigmoidal output function employed in the MFT and the Hopfield network. When a unit's internal potential is large, the unit is regarded as being "obstinate" in the following sense: its output value is large, although it is not easily influenced in its value by the other units.



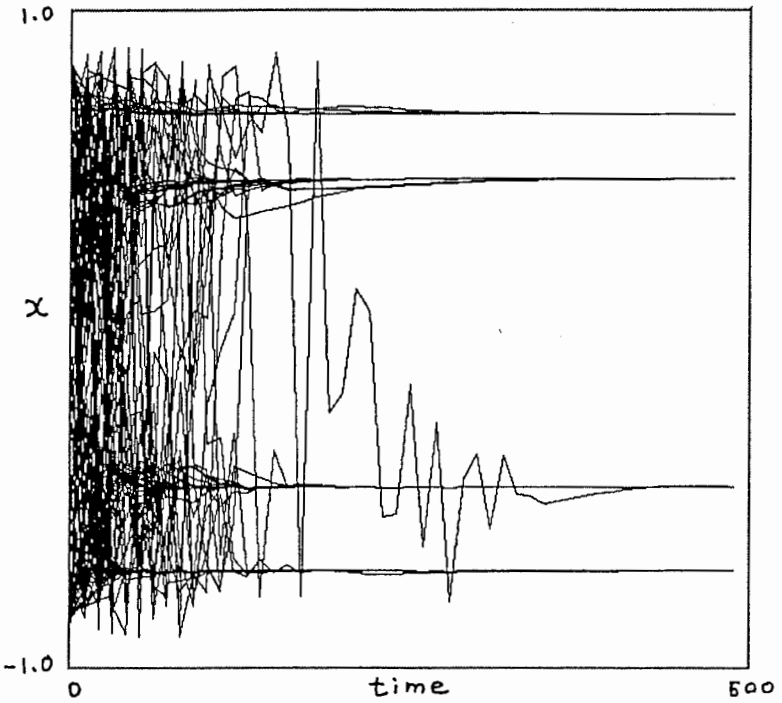


Figure 1

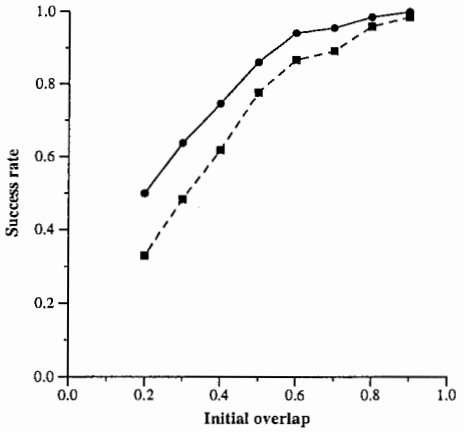


Figure 2(a)

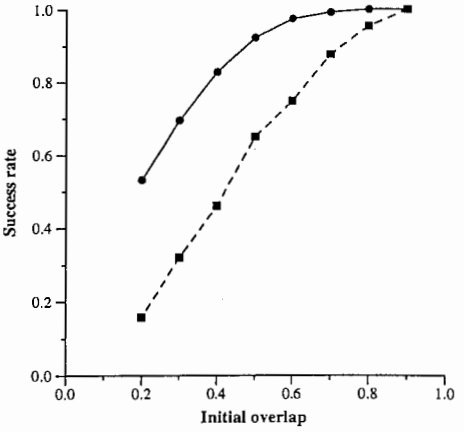


Figure 2(b)

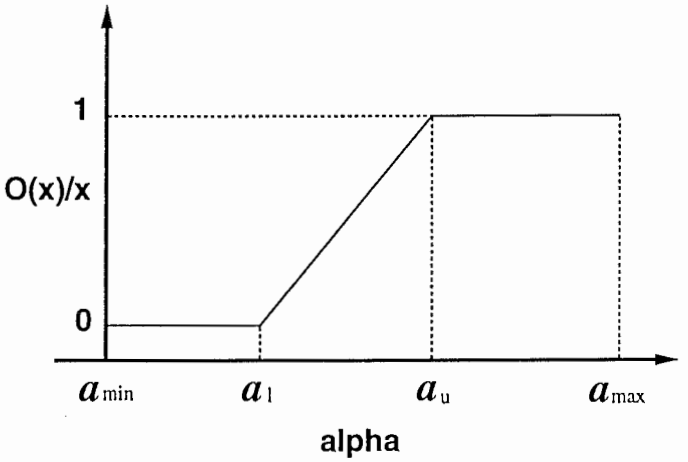


Figure 3

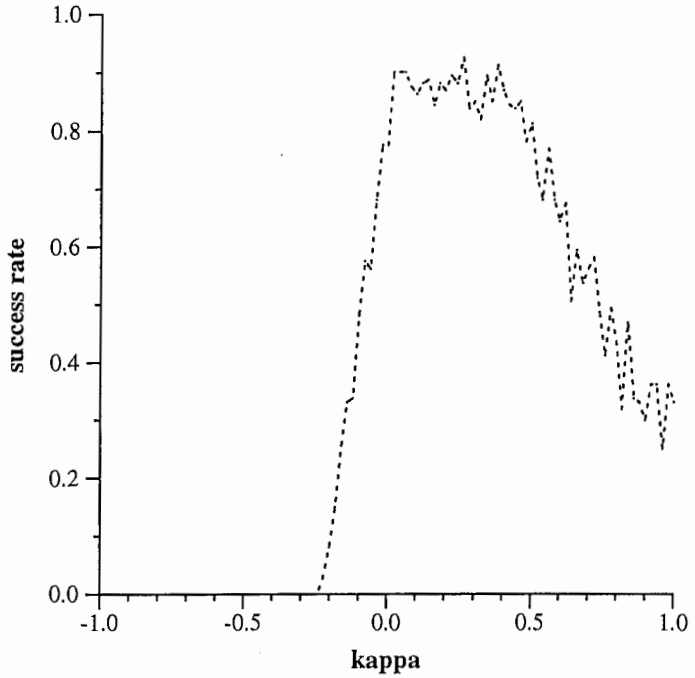


Figure 8

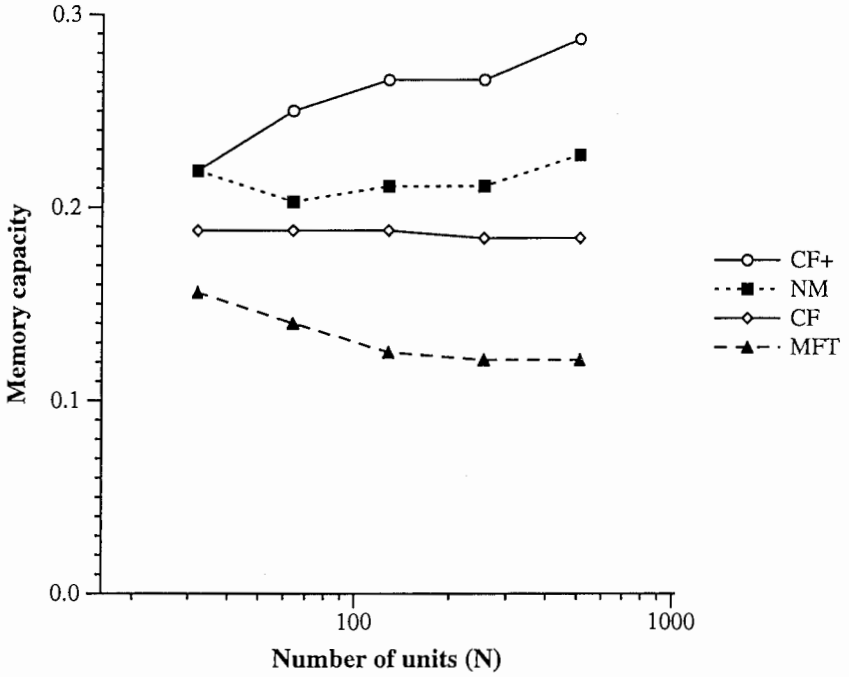

0.3
0.2
0.1
0.0
Memory capacity
100
1000
Number of units (N)
CF+
NM
CF
MFT

Figure 10

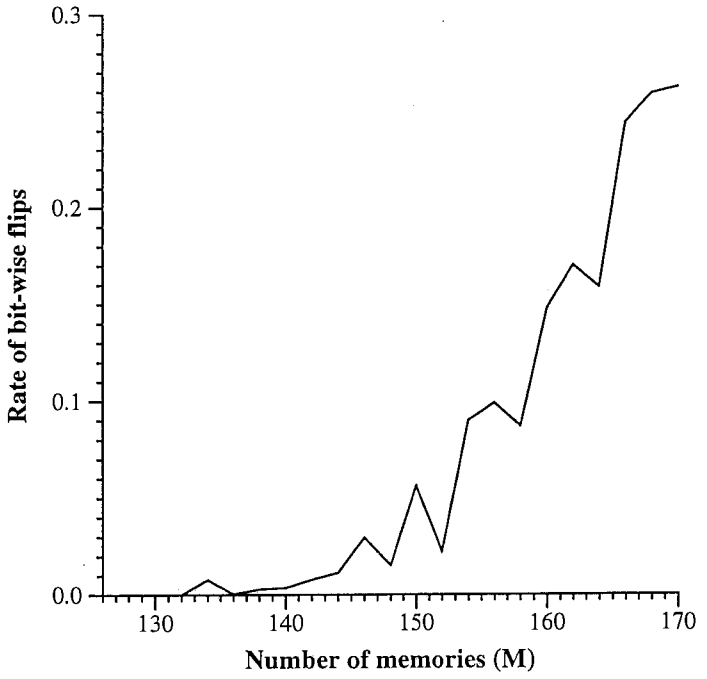

0.3
0.2
0.1
0.0
Rate of bit-wise flips
130
140
150
160
170
Number of memories (M)

Figure 11

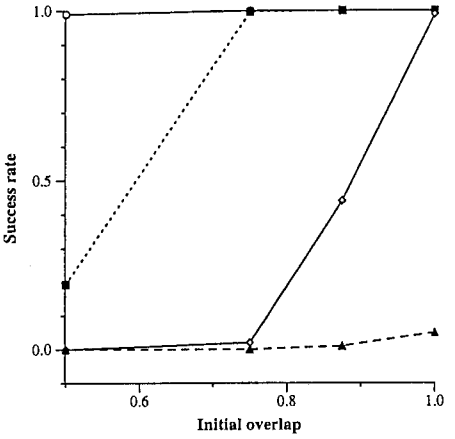


Figure 12(a)

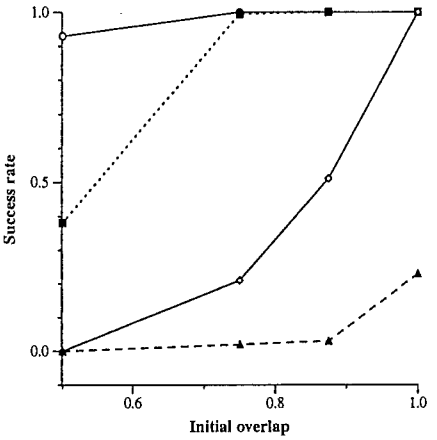


Figure 12(c)

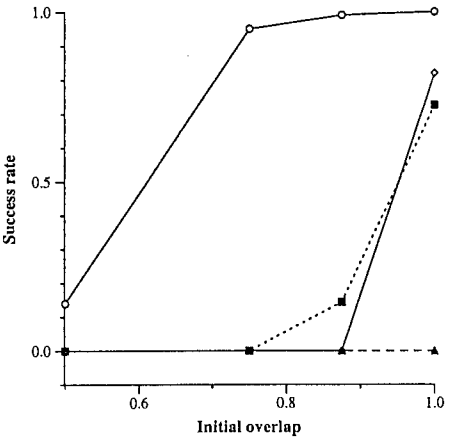


Figure 12(b)

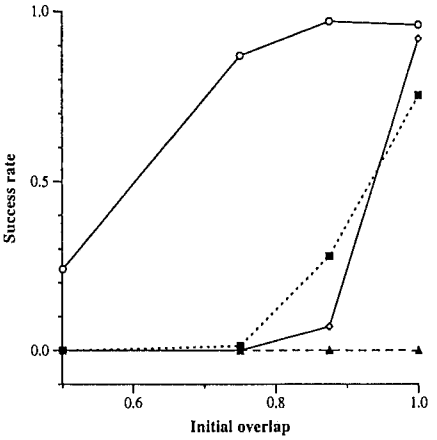


Figure 12(d)

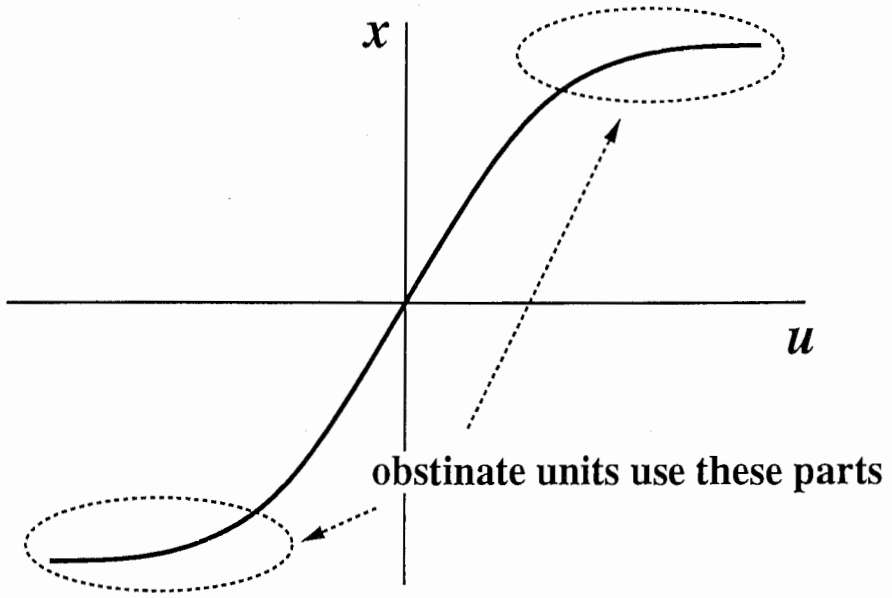


Figure 13

Figure 4(a)

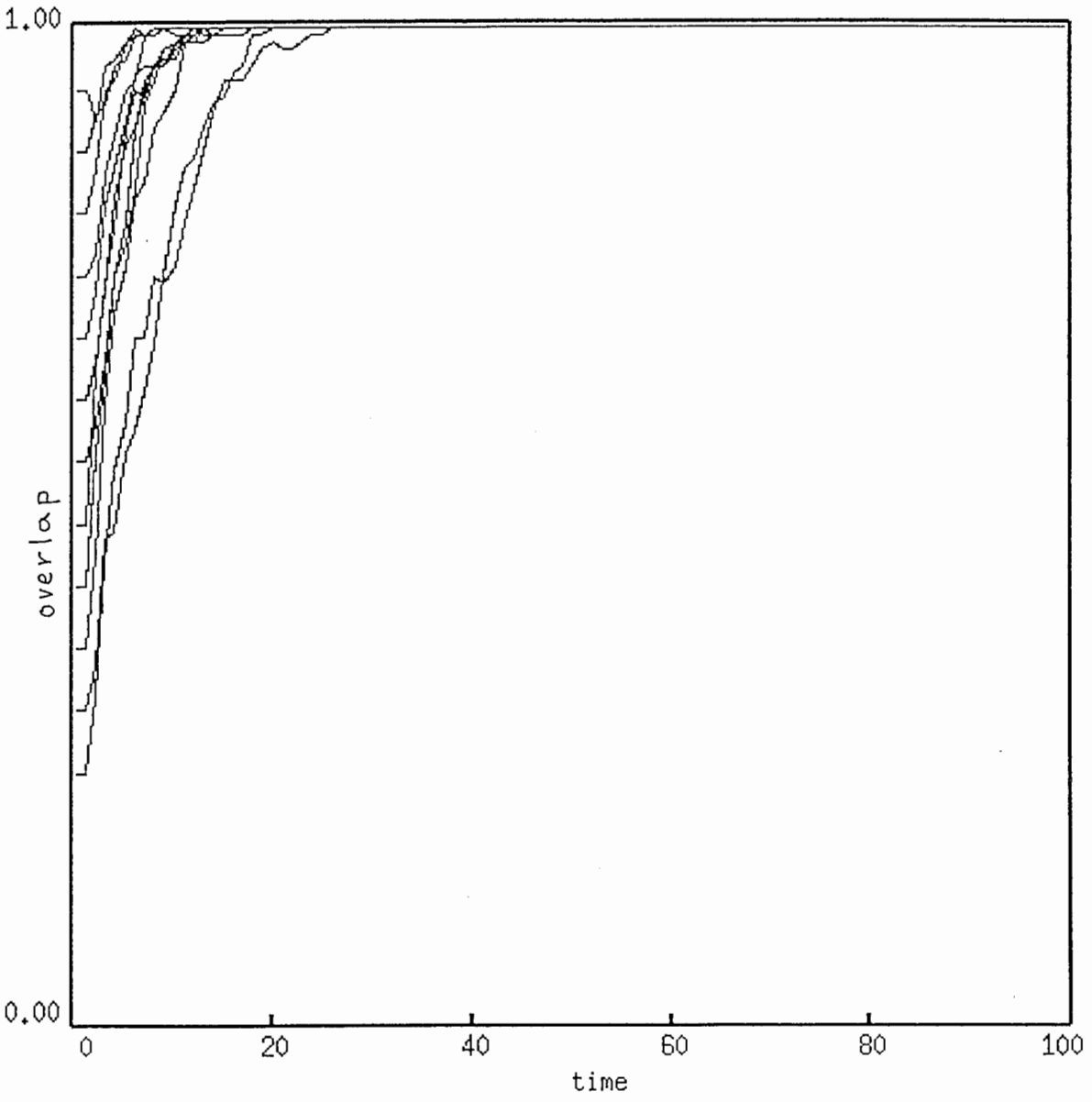


Figure 4(b)

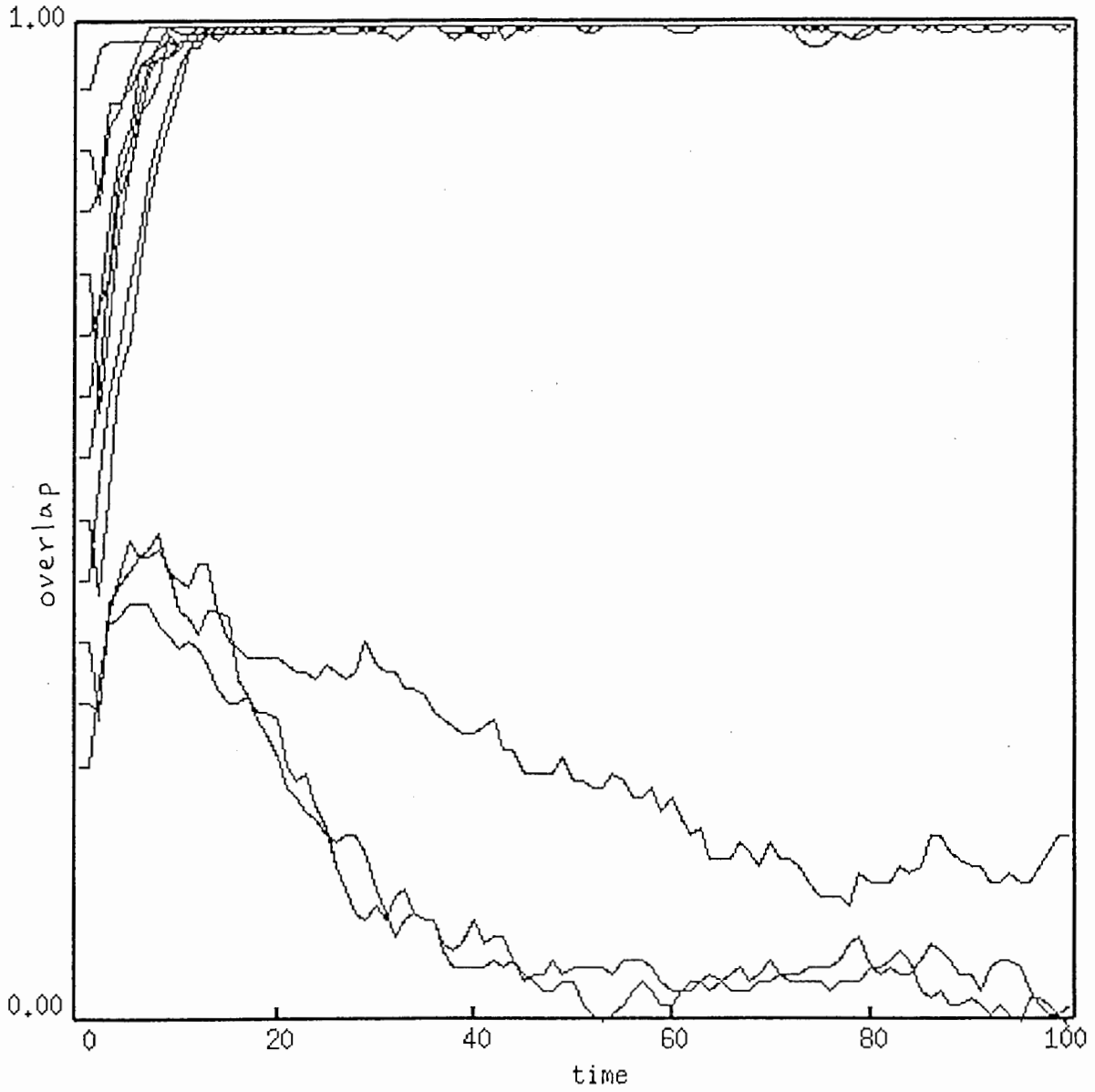


Figure 5(a)

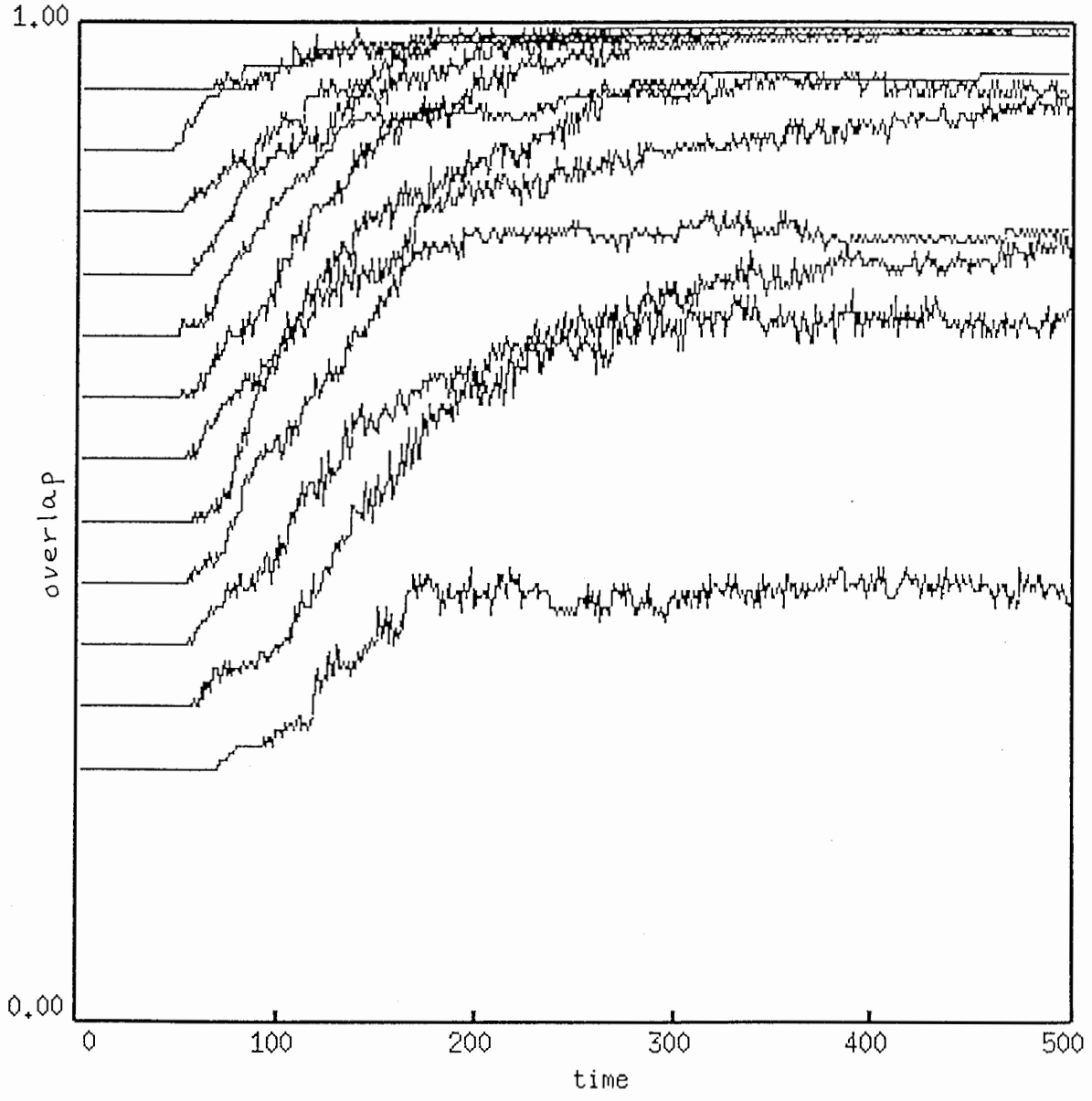


Figure 5(b)

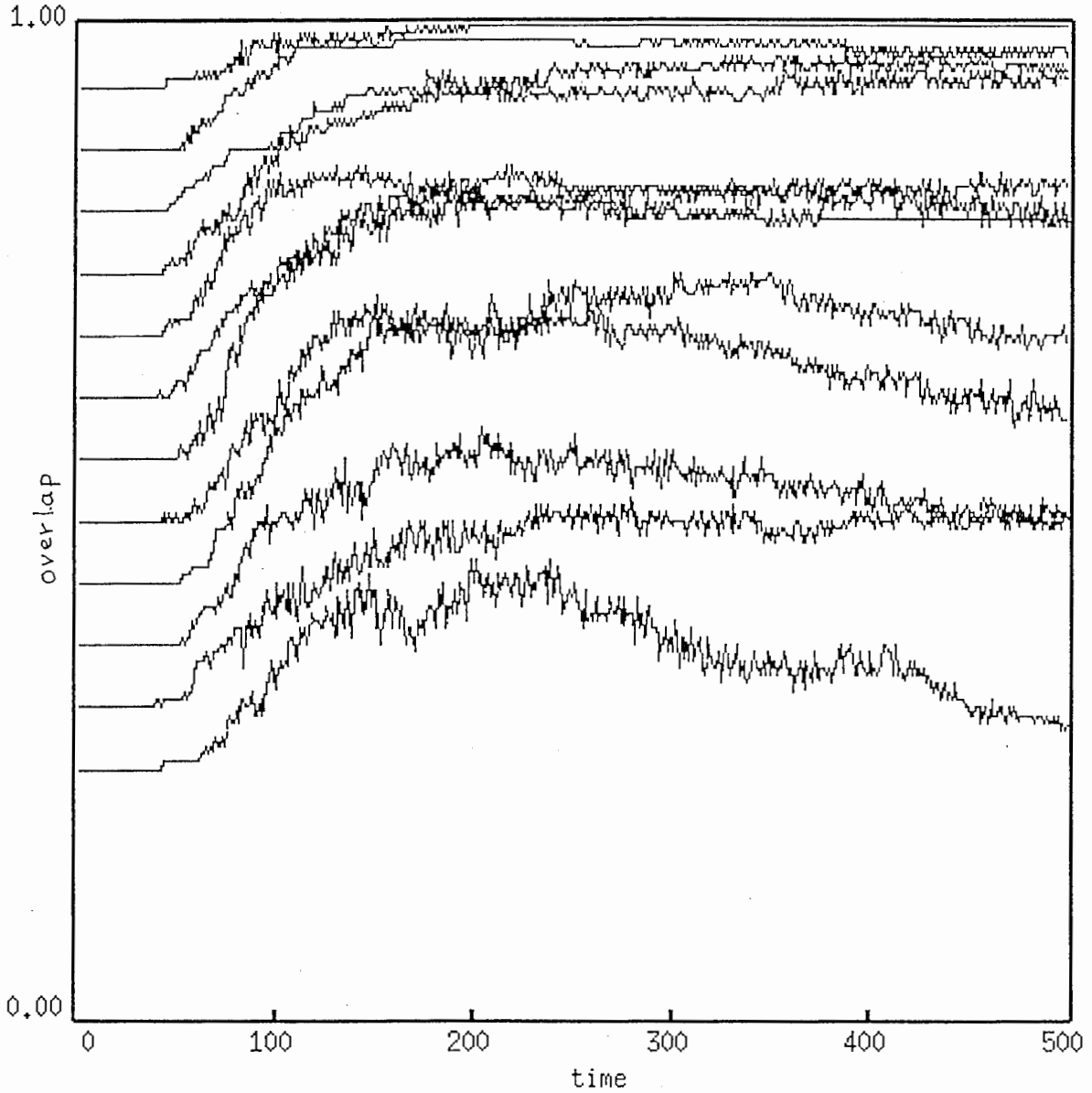


Figure 6

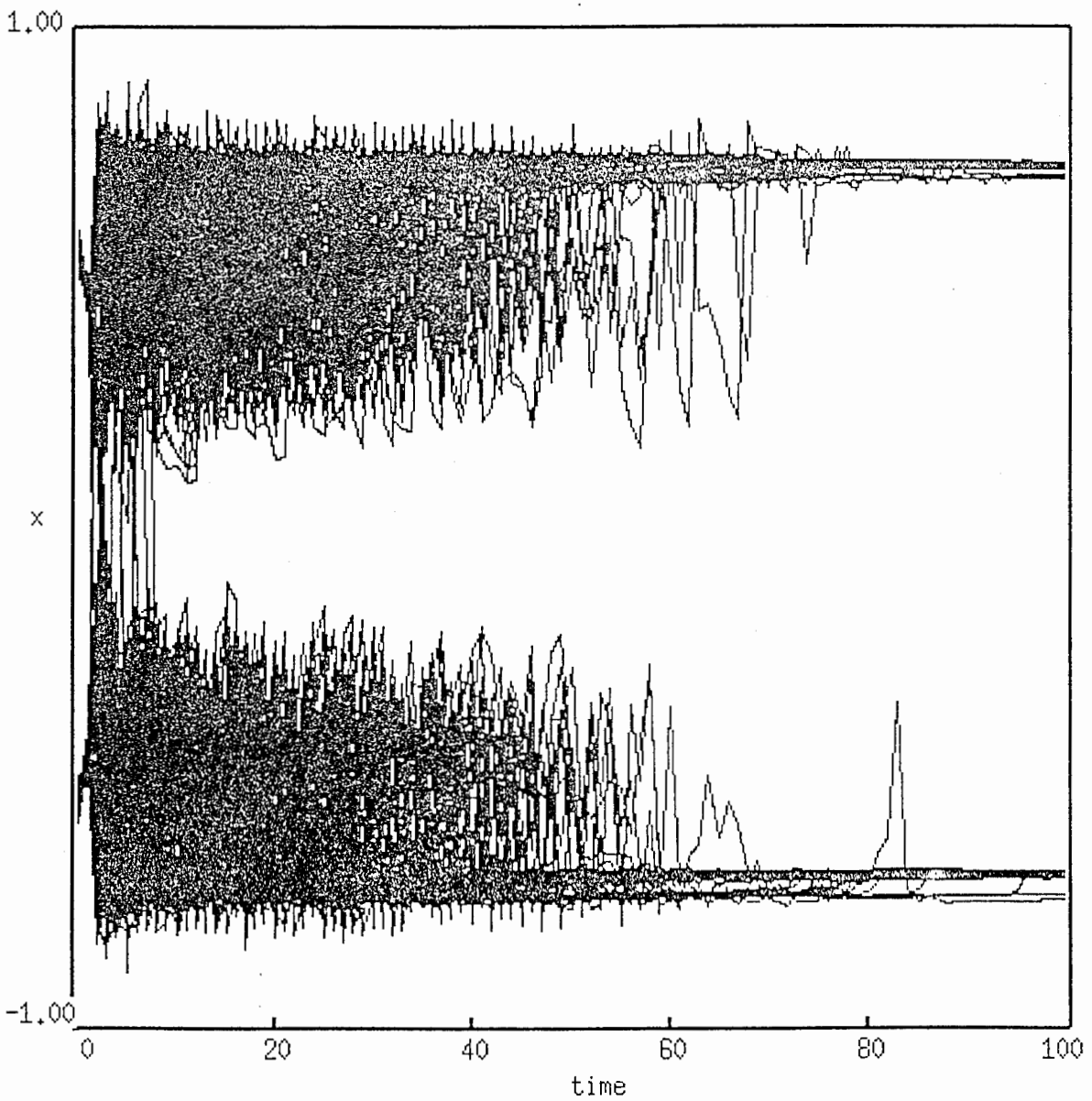


Figure 7(a)

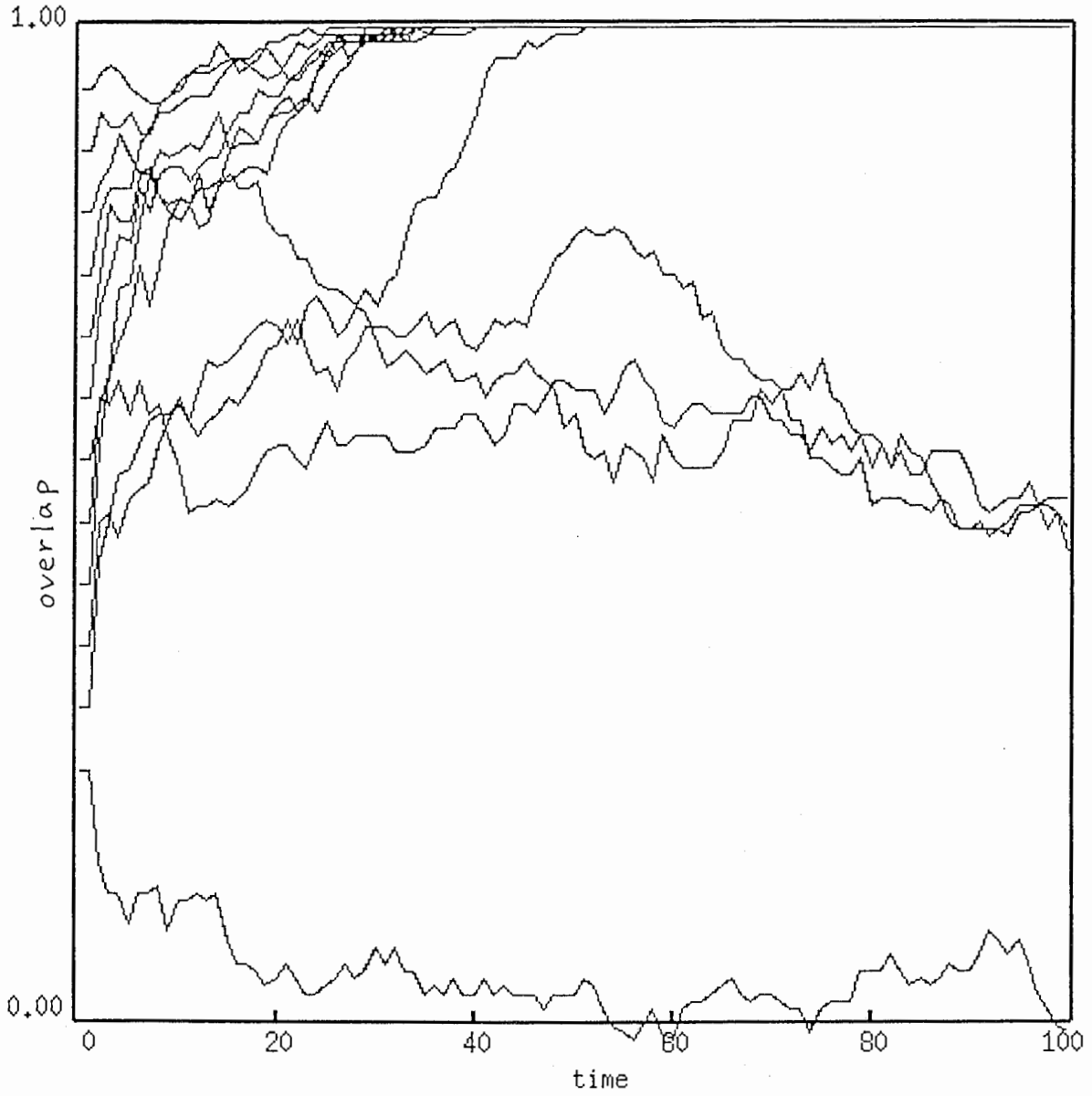


Figure 7(b)

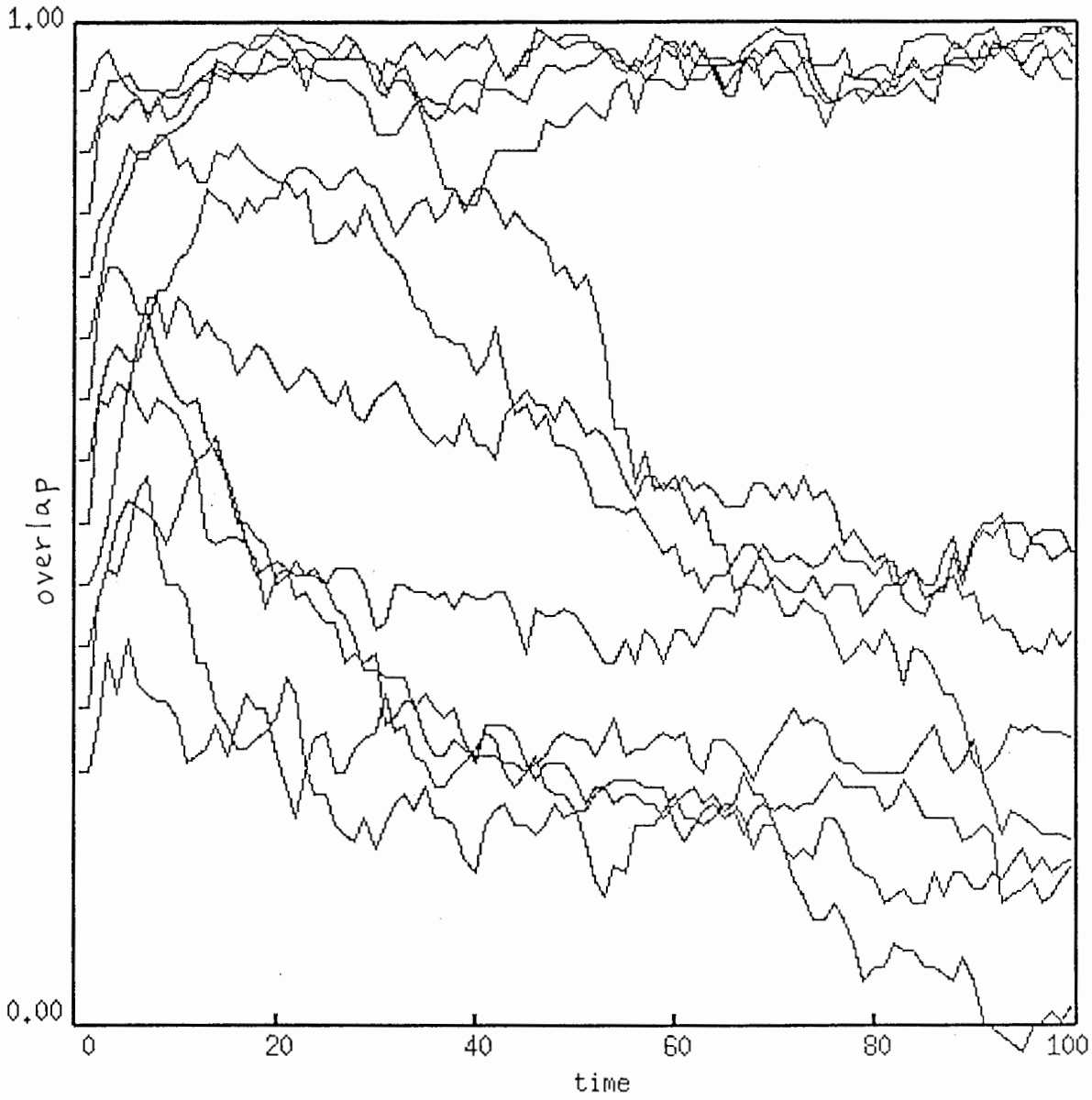


Figure 9