Temporal Association in Asymmetric Neural Networks

H. Sompolinsky and I. Kanter

Department of Physics, Bar-Ilan University, 52100 Ramat-Gan, Israel

(Received 15 August 1986)

A neural network model which is capable of recalling time sequences and cycles of patterns is introduced. In this model, some of the synaptic connections, J_{ij}, between pairs of neurons are asymmetric ($J_{ij} \neq J_{ji}$) and have slow dynamic response. The effects of thermal noise on the generated sequences are discussed. Simulation results demonstrating the performance of the network are presented. The model may be also useful in understanding the generation of rhythmic patterns in biological *motor* systems.

Studies[1-3] of neural-network models of associative memory focussed mostly on systems with symmetric couplings, i.e., the connections between pairs of neurons satisfy $J_{ij} = J_{ji}$, $i \neq j$. In this case, the dynamics of the network is relatively simple: The system relaxes to states which are local minima of a global energy function, E. Once such a local minimum has been reached it remains absolutely stable (at zero temperature). To endow the network with properties of associative memory, the J_{ij} are designed so that a given set of states of the network (the embedded memories) are local minima of E. The synaptic connections in biological neural networks have a high degree of *asymmetry*, but the role of the asymmetry in the collective behavior of the networks has been unclear. Recent studies have indicated that adding a weak *random* asymmetry to a symmetric network increases the level of the internal noise but otherwise does not modify drastically the capability to store and retrieve information.[1,4,5]

Symmetric networks cannot provide *temporal association*. They lack the ability to retrieve a *sequence* of patterns using a single recalling input. In this paper, a model of a neural network with temporal association is proposed. The network has asymmetric bonds in which a sequence of patterns is embedded. In general, asymmetric bonds may give rise to cyclic or perhaps even chaotic flows in configuration space. However, there is no reason to expect that these flows enjoy the same kind of robustness (e.g., large basins of attraction, independence of the details of the dynamics, stability against noise) as that of stable states in symmetric networks. Evidently, a desirable motion would consist of a controlled set of transitions among "quasiequilibrium" states: The system stays in a state (or in its neighborhood) for a long but finite period of time after which a transition is made to the next quasiequilibrium state in the sequence. To accomplish this, we propose that the response of the neuron at one end of the asymmetric bond to a signal at the other end is not instantaneous but has a dynamic memory characterized by a time constant τ. This leads to the emergence of a sequence of patterns, each stable over a time period of the order of τ, as will

be shown below by analytical study and numerical simulations. The maximum length of the sequence that can be embedded in a network of size N is proportional to N. We argue that under certain circumstances the sequences are stable even in the presence of thermal fluctuations. This network provides perhaps the simplest mechanism by which sequences of transitions between quasiequilibrium states can be embedded in highly connected systems.

We consider a network consisting of N Ising spins $\{S_i\}_{i=1}^{N}$, where $S_i = +1$ (-1) represents the firing (nonfiring) state of the ith neuron. The embedded patterns are p states of the system $\{\xi_i^{\mu}\}_{i=1}^{N}$, $\mu = 1, 2, \ldots, p$. The patterns are random so that each ξ_i^{μ} takes the values ± 1 with equal probabilities. The couplings between the neurons are formed by two kinds of synapses. Synapses of the first kind are symmetric and are given by the "Hebb rule"

$$J_{ij}^{(1)} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^{\mu} \xi_j^{\mu}, \quad i \neq j, \tag{1}$$

as in the Hopfield model.[1] The second kind of synapses are

$$J_{ij}^{(2)} = \frac{\lambda}{N} \sum_{\mu=1}^{q} \xi_i^{\mu+1} \xi_j^{\mu}, \quad i \neq j, \tag{2}$$

where $q < p$. The asymmetric synapses $J_{ij}^{(2)}$ define an order among the q patterns. Cycles can be incorporated by definition of $\xi_i^{q+1} = \xi_i^1$, in (2). The model can be extended easily to contain several sequences and cycles, provided that they do not compete against each other, i.e., that a given pattern, μ, is connected to only one pattern, $\mu+1$. The addition of synapses of the form (2) was discussed a few years ago by Hopfield,[1] who noted that they are insufficient for the generation of stable sequences of patterns. If λ is small, transitions between the patterns do not occur at all and the patterns remain completely stable. If λ is large, the transitions occur too fast so that the patterns are not stable even for a limited amount of time, and the sequence is very quickly smeared. This problem is rectified in the present work by endowment of the synapses $J_{ij}^{(2)}$ with slow dynamic response.

The dynamic evolution of the system is defined by the alignment of spins with their local fields

$$S_i(t+1) = \text{sgn}[h_i(t)], \tag{3}$$

either sequentially or in parallel. The local fields consist of two contributions, $h_i(t) = h_i^{(1)}(t) + h_i^{(2)}(t)$. The first contribution, from $J_{ij}^{(1)}$, has the usual instantaneous time dependence,

$$h_i^{(1)}(t) = \sum_{j=1}^{N} J_{ij}^{(1)} S_j(t). \tag{4}$$

The contributions from $J_{ij}^{(2)}$ are

$$h_i^{(2)}(t) = \sum_{j=1}^{N} J_{ij}^{(2)} \bar{S}_j(t), \tag{5}$$

$$\bar{S}_j(t) = \int_{-\infty}^{t} dt' \, w(t-t') S_j(t'). \tag{6}$$

The function $w(t)$ represents a dynamic memory characterized by a time-decay constant, τ.[6] We assume that $w(t)$ is nonnegative and is normalized by $\int_0^\infty dt \, w(t) = 1$. Since $h_i^{(2)}(t)$ averages the spin state over time $\sim \tau$, its strength increases if the spin state is time persistent on the scale of τ. Thus, for appropriate values of λ, $h^{(2)}$ will induce transition to the pattern $\mu+1$ only after the system has stayed in the state μ for a time t_0 which is of order τ. In the following, we determine the range of λ for which stable sequences exist and the period t_0 between consecutive transitions. We will assume that $\tau \gg 1$, although in practice $\tau > 4$ is sufficient to generate sequences.[6]

The analysis of the emergent dynamic features of the network is simplified in the limit where p remains finite as $N \to \infty$. The local fields become in this case

$$h_i(t) = \sum_{v=1}^{p} \xi_i^v m^v(t) + \lambda \sum_{v=1}^{q} \xi_i^{v+1} \bar{m}^v(t), \tag{7}$$

where $m^v(t)$ are the overlap with the pattern v, $m^v(t) = N^{-1} \sum_{i=1}^{N} \xi_i^v S_i(t)$, and $\bar{m}^v(t)$ are time averages of m^v, as in Eq. (6). Only overlaps which are ~ 1 need to be considered.

Suppose that $\{S_i(t<0)\}$ are random and that $\{S_i(t=0)\} = \{\xi_i^1\}$, so that $m^v(t<0) = 0$ and $m^v(t=0) = \delta^{v1}$. After staying in this state for a time t, the local fields become $h_i^{(1)}(t) = \xi_i^1$, $h_i^{(2)}(t) = \lambda \xi_i^2 W(0,t)$, where we have defined

$$W(a,b) \equiv \int_a^b w(t) \, dt. \tag{8}$$

Note that $W(0,t)$ is a nondecreasing function of t and $W(0,\infty) = 1$. Therefore, if $\lambda < 1$ the system will stay in the initial pattern indefinetely. However, if $\lambda > 1$ a transition to the pattern ξ^2 will occur at time t_0 such that $W(0,t_0) = \lambda^{-1}$. This analysis is, however, valid only for the first state and depends on the specific initial conditions.

We want to study the persistence of a long sequence of transitions (assuming that the length q of the embedded

sequence is large). After a short transient the transitions will be *equally* spaced in time with "*steady-state*" period t_0. If we neglect the transients, the system will be in the pattern μ ($1 \ll \mu < q$) in the time interval $((\mu-1)t_0, \mu t_0)$. The transition of this state to the pattern $\mu+1$ will start when a finite fraction of spins flip to $\xi_i^{\mu+1}$. The first spins to flip are those with the strongest bias in favor of $\xi_i^{\mu+1}$. These are spins located at sites such that $\xi_i^{\mu+1} = -\xi_i^\mu$ and all $\xi_i^{\mu-k}$ happen to be equal to $\xi_i^{\mu+1}$. In other words, these spins have been parallel to $\xi_i^{\mu+1}$ at all times up to $(\mu-1)t_0$, and parallel to ξ_i^μ only at $t = (\mu-1)t_0 + \Delta t$, $0 < \Delta t < t_0$. Substituting these conditions into Eq. (7) one finds that, for these spins, the part of $h_i(t)$ parallel to ξ_i^μ is $\xi_i^\mu[1 + \lambda W(\Delta t, \Delta t + t_0)]$. The term $\lambda W(\Delta t, \Delta t + t_0)$ is the contribution of $h_i^{(2)}$ resulting from the state $\mu-1$, which appeared between $t - (\Delta t + t_0)$ and $t - \Delta t$. The part of h_i which is parallel to $\xi_i^{\mu+1} = -\xi_i^\mu$ is $-\lambda \xi_i^\mu[W(0,t) - W(\Delta t, \Delta t + t_0)]$. The transition will occur at the time for which these two contributions balance each other so that $\xi_i^\mu h_i(t)$ changes sign. Identifying this time as $t = \mu t_0$ (i.e., $\Delta t = t_0$), one finds $1 + \lambda W(t_0, 2t_0) = \lambda[W(0,t) - W(t_0, 2t_0)]$. If we assume $t \gg 1$ [i.e., $W(0,t) \simeq 1$] the following equation results:

$$W(t_0, 2t_0) \equiv \int_{t_0}^{2t_0} w(t) \, dt = \tfrac{1}{2}(1 - 1/\lambda). \tag{9}$$

This equation determines the duration t_0 of each state in the sequence, and its dependence on λ and on $w(t)$.

Note that in the derivation of Eq. (9) we have neglected the transition time t_{tr} between patterns. This is justified in the case of large τ since t_{tr} is short and does not increase with τ.

From Eq. (9) it is evident that the minimum value of λ for which stable sequences occur is $\lambda_{min} = 1$ as expected. When $\lambda \to \lambda_{min}$, $t_0 \to t_{max}$, t_{max} being the upper cutoff in w, i.e., $w(t > t_{max}) = 0$. The maximum value of λ for which long sequences are stable is determined by the maximum value of λ for which a solution to Eq. (9) exists.[7] This depends on the particular form of the memory function $w(t)$.

We now present three interesting examples: (1) *Step function*: $w(t) = \tau^{-1}$, $t < \tau$, and vanishes otherwise. In this case, Eq. (9) yields $t_0 = (\tau/2)(1 + 1/\tau)$, $\lambda > 1$, implying a short period $t_0 = \tau/2$ at large values of λ and a long period $t_0 = \tau$ at $\lambda = \lambda_{min} = 1$. (2) *Exponential decay*: $w(t) = \tau^{-1} e^{-t/\tau}$. Substituting in Eq. (9) yields

$$t_0 = \tau\{\ln 2 - \ln[1 - (2/\lambda - 1)^{1/2}]\}, \quad 1 < \lambda < 2.$$

Here $t_0 \to \infty$ as $\lambda \to 1$, whereas $t_0 \to \tau \ln 2$ as λ approaches its maximum value $\lambda_{max} = 2$. (3) *Time-delay*: $w(t) = \delta(t - \tau)$. Applying Eq. (9) to a smooth function $w(t)$ which approaches a delta function, one finds $t_0 = \tau$, $\lambda > 1$.

Note that Eq. (9) imposes a finite upper bound on λ in case (2), but not in cases (1) and (3). The reason is that the exponential decay lends a large weight to the im-

mediate times which leads to a rapid smearing of the transition if λ is big. This effect has been verified by numerical simulations. Simulations of the equations of motion (1)–(6) have been done with $N = 500$, $p = q + 1 = 10$, and $\tau = 8$. In the case of a step function, a regular stable sequence is obtained for $0.9 < \lambda < 5.5$[8] whereas in the case of an exponential decay, λ_{max} was found to be -1.8. Figure 1 presents the numerical results for both case (1) and case (2), with $\lambda = 2.5$.

The above analysis holds only in the limit of $N \to \infty$ and finite p. In particular, when $a \equiv p/N$ is finite as $N \to \infty$, the random overlaps among the patterns cannot be ignored. Simulations show that for small a, long sequences are still stable. However, keeping a finite modifies the bounds on the allowed range of λ. On one hand, $\lambda_{min}(a)$ is smaller than unity since fluctuations in the local fields are capable of inducing transitions even when $\lambda < 1$. On the other hand, increasing the value of λ enhances the relative level of the internal noise and therefore λ_{max} is expected to decrease as a function of a. As a increases, a critical value is reached above which sequences are not stable *for any value of* λ. These effects were verified by numerical simulations. Using a step function for $w(t)$ with $\tau = 8$, we find that, with $N = 500$, a sequence of 40 patterns can be embedded with $0.6 \lesssim \lambda \lesssim 1.4$, whereas for $p = q + 1 > 60$, sequences are not stable for any value of λ; see Fig. 2. This indicates that $a_c \sim 0.1$, in contrast to $a_c \sim 0.14$ in the symmetric ($\lambda = 0$) Hopfield model. The value of a_c depends on the form of $w(t)$ as well as on the lengths and number of the embedded sequences and cycles.

An important question is whether sequences exist also in the presence of fast stochastic noise. One can represent such noise by use of stochastic dynamics, in which if the spin S_i is aligned with its field $h_i(t)$ it may

still flip, with a probability $\exp(-2\beta | h_i |)$. Here $\beta^{-1} = T$ is the "temperature" of the network denoting the level of stochastic noise in the system. Sequences can be stable only if the time τ is much larger than the time, t_{eq}, for equilibration within the free-energy "valley" of the individual patterns but smaller than the "ergodic" time, t_{erg}, at which thermal hopping *between* the valleys occurs. These restrictions may not be severe since, in the symmetric case, when $T \lesssim \frac{1}{2}$, t_{erg} is extremely big (for large N) and t_{eq} is rather short. For $t_{eq} \ll \tau \ll t_{erg}$ a well-defined sequence of q states is expected to exist in our model, at low T. Each of these states is a quasiequilibrium state determined by an effective free energy which changes with time (on the scale of τ). (The μth state is characterized by a large value of m^μ.) At finite T, this state will also have a small overlap with $\{\xi^{\mu+1}\}$, and smaller overlaps with other patterns.

The present model has been discussed so far in the context of storage and associative recall of information. It might also be useful in understanding mechanisms for generating *rhythmic motoric* patterns (such as swimming and locomotion) by networks of neurons known as *central pattern generators* (CPG).[9] Understanding the principles underlying the function of CPG's in vertebrates as well as in invertebrates has been a longstanding challenge in neurobiology, in particular in cases where the networks contain mixtures of inhibitory and excitatory synapses.

A particular simple example is an oscillatory behavior in which the neurons alternate (coherently) between "bursts" of activity and inactivity. This means that the state of the network oscillates between a pattern $\{\xi_i\}$ and its antiphase $\{-\xi_i\}$. (Recall that the two states $\xi_i = \pm 1$ represent the firing and nonfiring states of the neuron i.) Such a behavior can be simply modeled by embedding the pattern $\{\xi_i\}$ in $J_{ij}^{(1)}$ and connecting it, in $J_{ij}^{(2)}$, with the pattern $\{-\xi_i\}$. Generalizing to several patterns
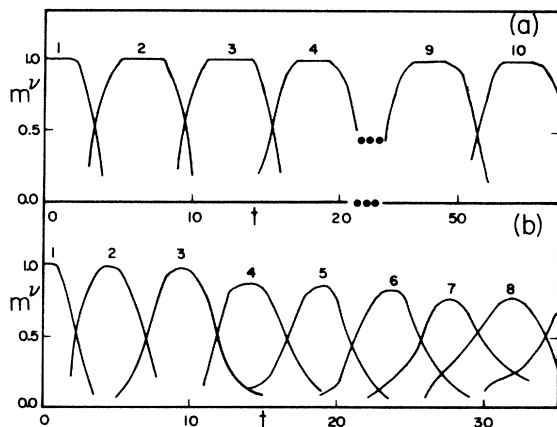
FIG. 1. Simulations of Eqs. (3)–(6) with parallel updating, and the parameters $N = 500$, $p = q + 1 = 10$, $\tau = 8$, and $\lambda = 2.5$. (a) $w(t)$ is a step function [case (1) in text]; (b) $w(t)$ decays exponentially [case (2) in text]. The curves represent the overlaps m^ν, at time t, of the state of the system with the patterns, ξ^ν. Numbers on top of the curves denote the index ν.
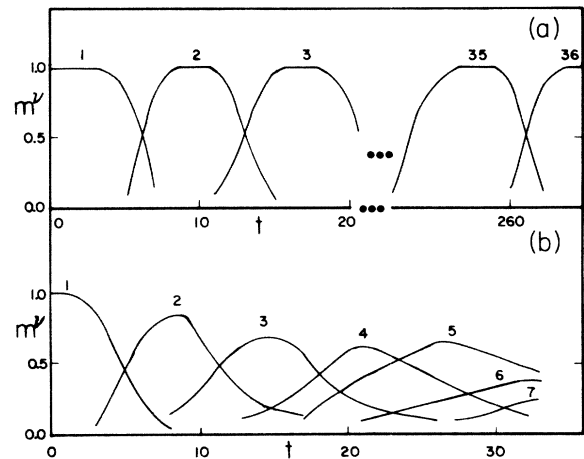
FIG. 2. Simulations of Eqs. (3)–(6) with $N = 500$, $\lambda = 1.0$, and $w(t)$ a step function with $\tau = 8$. (a) $p = q + 1 = 40$; (b) $p = q + 1 = 60$.

yields the *symmetric* network

$$J_{ij}^{(1)} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu, \quad J_{ij}^{(2)} = -\frac{\lambda}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu.$$

(10)

Thus, each pair of neurons is connected by a pair of excitatory and inhibitory synapses, one of which has a slow dynamic response $w(t)$, with characteristic time τ. Alternatively, this pair of synapses can be viewed as a single synapse which changes its sign as a function of time. This network is capable of generating p stable modes of oscillations. Each mode consists of oscillations between a pattern of activity, say $\{\xi_i^\mu\}$, and its antiphase $\{-\xi_i^\mu\}$. The period of the oscillations is of $O(\tau)$. The selection of the particular oscillation is done by, e.g., the intitial triggering input. It is interesting to note that synapses which differ in their response time by as much as a factor of 20 have been identified in certain CPG's.[10] Detailed comparison between the present model and several well-characterized pattern generators would be very interesting.

Note added— A very similar model for generating sequences has been proposed independently by Kleinfeld.[11] Other mechanisms have been recently proposed by Peretto and Niez and by Nebenzahl.[12]

One of us (H.S.) wishes to thank David Kleinfeld for drawing his attention to the problem of pattern generation in biological motor systems. Helpful comments on the manuscript by C. L. Henley, P. C. Hohenberg, J. J.

[1] J. J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982), and **81**, 3088 (1984).

[2] W. A. Little, Math. Biosci. **19**, 101 (1974); W. A. Little and G. L. Shaw, Math. Biosci. **39**, 281 (1978).

[3] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A **32**, 1007 (1985), and Phys. Rev. Lett. **55**, 1530 (1985), and Ann. Phys. (to be published).

[4] H. Gutfreund and Y. Stein (unpublished).

[5] J. A. Hetz, G. Grinstein, and S. A. Solla, to be published.

[6] If one uses a *sequential* updating of Eq. (3), the time scale τ has to be measured in number of updatings *per spin*.

[7] Equation (9) has to be supplemented by the requirement that

$$\lim_{\Delta t \to t_0} dW(\Delta t, \Delta t + t_0)/d(\Delta t) = w(2t_0) - w(t_0) < 0.$$

This condition guarantees that as $\Delta t \to t_0$, $\xi_i^\mu h_i$ changes sign from positive to negative and not vice versa.

[8] The finite upper bound on λ in this case is due to the fact that the simulations are done with finite N and τ.

[9] *Model Neural Networks and Behavior*, edited by A. I. Selverston (Plenum, New York, 1985).

[10] P. A. Getting and M. S. Dekin, in Ref. 9, pp. 3–20.

[11] D. Kleinfeld, to be published.

[12] P. Peretto and J. J. Niez, in *Disordered Systems and Biological Organization*, edited by E. Bienenstock (Springer-Verlag, Berlin, 1986); I. Nebenzahl, to be published.