

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/101015>

Please be advised that this information was generated on 2020-10-25 and may be subject to change.

ORIGINAL CONTRIBUTION

Retrieval of Pattern Sequences at Variable Speeds in a Neural Network With Delays

TOM M. HESKES AND STAN GIELEN

University of Nijmegen

(Received 23 April 1990; revised and accepted 20 May 1991)

Abstract—We discuss a neural network for the storage and retrieval of sequences of patterns. Assuming a distribution of delays, short relative to the duration of each of the single patterns, and assuming adaptive thresholds, we show that a sequence of patterns, learned at a particular speed, can be replayed at variable speeds. The retrieval speed can be adjusted by a threshold adaptation constant. The effect of noise on the transitions between successive pattern states is studied.

Keywords—Neural networks, Delays, Temporal sequences, Variable retrieval speed, Dynamical threshold, Hebbian learning, Pattern overlap, Motor control.

1. INTRODUCTION

The ability of Hopfield-type networks (Hopfield, 1982) to store and retrieve sequences of patterns has been discussed by several authors. These networks are characterized by a competition between stability of a pattern and transition towards the next pattern. On a microscopic level, this controversy is often accomplished and controlled by couplings between neurons, made up of two different contributions: a stabilizing, usually symmetric part, and a destabilizing, asymmetric part. The various models differ in the way the couplings are built. One approach for storage and retrieval of sequences of patterns was introduced by Peretto and Niez (1986), who assumed fast synaptic plasticity. The pattern generator of Sompolinsky and Kanter (1986) and Kleinfeld (1986), assuming couplings formed by two kinds of synapses, is an example of another type of approach. We will follow a slightly different strategy as proposed by Coolen and Gielen (1988) and Herz, Kühn, and Van Hemmen (1989). These authors postulated a broad distribution of delays in the connections between neurons, present during both retrieval and learning

phase, and show that their network is capable of regenerating a stored sequence of patterns.

Yet the theories using the latter approach have some properties that are not very plausible from a biological point of view. First, these models need a distribution of delays exceeding the duration of an individual pattern within a sequence to “measure” the lifetime of a pattern during the learning session. We will assume that the delays are much smaller than the duration of a pattern. Especially in biological systems, this is a more realistic assumption, since transmission delays are in the order of milliseconds, which is much larger than the width of the autocorrelation function of sensory stimuli and motorprograms. Second, biological nervous systems are able to scale the patterns in time so as to generate patterns at various speeds, different from the “learning” speed. The known models, using a broad distribution of delays, can only change the lifetime of a pattern during retrieval by scaling all delays in the network. This would correspond to variable transmission delays. In our model we will make the biologically more plausible assumption that the delays are fixed.

We assume a fully connected one-layered network with a distribution of transmission delays, all of which are shorter than the duration of an individual pattern within a sequence. With such relatively small delays, there is no reason why the network, once in a particular attractor state, should leave this state. Therefore, we use dynamical thresholds, which change as a function of the input to the neuron in the past, similar to those introduced by Horn and Usher (1989), to drive the network into another state. We show

Acknowledgements: The authors gratefully acknowledge the support of the Dutch Foundation for Neural Networks. Moreover, we thank Ton Coolen for critically reading a previous version of this manuscript.

Requests for reprints should be sent to Tom M. Heskés, Department of Medical Physics and Biophysics, University of Nijmegen, Geert Groote Plein Noord 21, 6525 EZ Nijmegen, The Netherlands.

that this new state has to be the next pattern in the stored sequence. The lifetime of a pattern during retrieval depends on the adaptation rate of the thresholds, which is variable.

2. STORAGE OF PATTERN SEQUENCE

We will store a sequence of patterns in a fully connected one-layered neural network consisting of N neurons. The neurons are represented as Ising spins $s_i(t)$, with $s_i(t) = +1$ if neuron i fires at time t and $s_i(t) = -1$ if neuron i is at rest. We consider a cycle of p patterns, described by the corresponding spin states $\{\xi_i^\mu | 1 \leq i \leq N; 1 \leq \mu \leq p\}$, each of duration Δ . For the time being, the patterns are pseudo orthogonal, which means that the ξ_i^μ are randomly chosen from $\{-1, +1\}$ with equal probability. Since it takes a time τ_{ij} for a state change of neuron j to affect the state of neuron i , the connectivity of neuron j to the neuron i according to a Hebbian type learning rule (Hebb, 1949), is defined by:

$$J_{ij}(\tau_{ij}) \propto \int_0^{\rho\Delta} dt s_i(t)s_j(t - \tau_{ij}). \quad (1)$$

We assume a distribution of short delays: $\tau_{ij} < \Delta \forall i, j$. With a proper normalisation the connectivities in eqn (1) can be rewritten as:

$$J_{ij}(\tau_{ij}) = \frac{1}{N-p} \times \left\{ \left(1 - \frac{\tau_{ij}}{\Delta}\right) \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu + \frac{\tau_{ij}}{\Delta} \sum_{\mu=1}^p \xi_i^\mu \xi_j^{\mu-1} \right\}. \quad (2)$$

It is easily verified that for $\tau_{ij} = 0 \forall i, j$ the connectivities reduce to the standard Hopfield connectivities to store a set of p discrete patterns as attractors in the network.

We will assume a continuous distribution of delays $\rho(\tau)$, satisfying $\rho(\tau) = 0$ for $\tau < 0$ and $\tau > \tau_{\max}$. For simplicity, we will also assume that $\rho(\tau_{\max} - \tau) = \rho(\tau)$, such that:

$$\int_0^{\tau_{\max}} d\tau \rho(\tau)\tau = \frac{\tau_{\max}}{2}. \quad (3)$$

In fact we could make many other (asymmetric) choices and get even better performance of our model.

3. RETRIEVAL OF PATTERN SEQUENCE

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

$$s_i(t+1) = \text{sgn}[h_i(t)] = \text{sgn}[I_i(t) - \Theta_i(t)], \quad (4)$$

with $\Theta_i(t)$ the time dependent neuronal threshold and $I_i(t)$ the local input containing the contributions

of all other spins:

$$I_i(t) = \sum_{j \neq i} J_{ij}(\tau_{ij})s_j(t - \tau_{ij}). \quad (5)$$

The neuronal ‘‘decision time’’ in eqn (4) is our unit of time 1. The choice $\Delta > \tau_{\max} > 1$ is most plausible. We will work in the thermodynamic limit $N \rightarrow \infty$.

To explain the dynamics of the network we start by imposing a part of the learned sequence for at least a duration τ_{\max} before the system is set free to evolve, starting at a time t after the last transition to the state corresponding to pattern v . Then the local input at time t reads:

$$I_i(t) = \frac{1}{N-p} \sum_{\mu=1}^p \xi_i^\mu \sum_{j \neq i} \times \left[\left(1 - \frac{\tau_{ij}}{\Delta}\right) \xi_j^\mu + \frac{\tau_{ij}}{\Delta} \xi_j^{\mu-1} \right] s_j(t - \tau_{ij}). \quad (6)$$

The expectation value of the local input at this time t is computed by taking the ensemble average over the distribution of the delays, thereby assuming that the delays and the pattern states are independent. The sum over the neurons j can be split into two parts: one for all j satisfying $t \geq \tau_{ij}$ and one for all j satisfying $t < \tau_{ij}$, such that we may substitute ξ_j^v or ξ_j^{v-1} for $s_j(t - \tau_{ij})$. The full derivation is given in the Appendix. For $t \geq \tau_{\max}$, the expectation value of the local input is given by:

$$\begin{aligned} \langle I_i(t) \rangle &= \int_0^{\tau_{\max}} d\tau \rho(\tau) \left(1 - \frac{\tau}{\Delta}\right) \xi_i^v \\ &\quad + \int_0^{\tau_{\max}} d\tau \rho(\tau) \frac{\tau}{\Delta} \xi_i^{v-1} \\ &= \left(1 - \frac{\tau_{\max}}{2\Delta}\right) \xi_i^v + \frac{\tau_{\max}}{2\Delta} \xi_i^{v-1}. \end{aligned} \quad (7)$$

The expression for the local input becomes more complicated for $t < \tau_{\max}$:

$$\begin{aligned} \langle I_i(t) \rangle &= \int_t^{\tau_{\max}} d\tau \rho(\tau) \left(1 - \frac{\tau}{\Delta}\right) \xi_i^{v-1} \\ &\quad + \int_t^{\tau_{\max}} d\tau \rho(\tau) \frac{\tau}{\Delta} \xi_i^v + \int_0^t d\tau \rho(\tau) \\ &\quad \times \left(1 - \frac{\tau}{\Delta}\right) \xi_i^v + \int_0^t d\tau \rho(\tau) \frac{\tau}{\Delta} \xi_i^{v-1}. \end{aligned} \quad (8)$$

Eqn (8) can be rewritten in a more general way by:

$$\langle I_i(t) \rangle = g_{-1}(t)\xi_i^{v-1} + g_0(t)\xi_i^v + g_1(t)\xi_i^{v+1}. \quad (9)$$

If the distribution of the delays is known, the functions $g_k(t)$ can be calculated. As an example we take the homogeneous distribution $\rho(\tau) = 1/\tau_{\max}$. Figure 1 shows the functions $g_k(t)$ for $\tau_{\max}/\Delta = 1/4$. Initially $g_{-1}(t)$ is relatively large: the system tends to return to the previous state. When time t increases, the fraction of delays τ_{ij} with $\tau_{ij} > t$ decreases, resulting in a decrease of $g_{-1}(t)$ and an increase of $g_0(t)$ and $g_1(t)$. At $t \approx \tau_{\max}/2$, $g_0(t)$ equals $g_{-1}(t)$. Since all de-

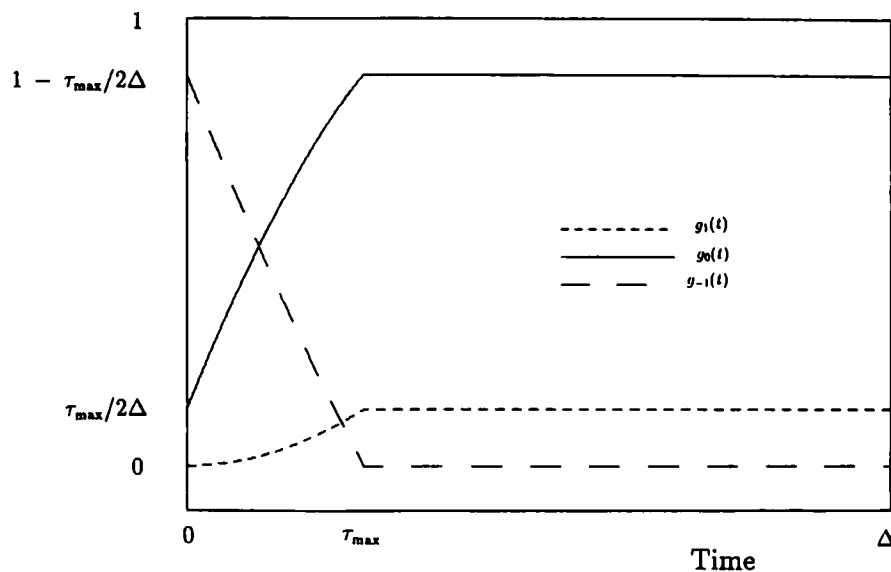


FIGURE 1. Three different contributions to the expectation value of the local input for a homogeneous distribution of delays. It is assumed that the network changed from state $\nu - 1$ to ν at time $t = 0$. The function $g_k(t)$ reflects the tendency of the network to stay in pattern state $\nu + k$.

lays are shorter than τ_{\max} , $g_{-1}(t) = 0$ for $t > \tau_{\max}$. The fact that $g_0(t) > g_1(t) \forall t$, means that the tendency to maintain the system in the new state ν is always larger than the tendency to go to the next state $\nu + 1$.

Up to now, there is no mechanism to drive the system toward the next pattern in the sequence. Once in a particular attractor state, the system will stay there forever. It is even worse: we have to impose the next pattern for at least a duration $\tau_{\max}/2$ to prevent the spins from flipping back to their previous state. Therefore, we will introduce dynamical thresholds, as proposed by Horn and Usher, to solve this problem.

4. DYNAMICAL THRESHOLDS

The neuronal threshold proposed by Horn and Usher (1989) was variable depending on the previous states of the neuron. With a variable threshold, an attractor state becomes unstable, forcing a transition of the network to another state. To reproduce a particular sequence of patterns, "pointers" similar to those introduced by Sompolinsky and Kanter (1986) were used. We will show that pointers are not necessary and that the couplings in our network provide enough information to regenerate the stored cycle at a variable speed by using time dependent thresholds slightly different from those proposed by Horn and Usher.

The duration of an individual pattern during retrieval will be distinguished by two distinct periods: the reset period and the adaptation period. The reset period, which will be discussed later, is needed to force the system to stay in the new state after the

transition and to reset the thresholds. Therefore, at the beginning of the adaptation period the thresholds are zero. During the adaptation period they are modified according to:

$$\frac{d}{dt} \Theta_i(t) = c\Theta_i(t) + \alpha(s_i(t) - I_i(t)). \quad (10)$$

Here α is the threshold adaptation constant and c the threshold decay constant, for convenience set equal to 0 in the rest of our paper. Equation (10), which describes the dynamics of the thresholds, is different from that proposed by Horn and Usher. It assumes that the threshold only changes if there is a difference between the state of the neuron and its local input. Assuming that the reset time is at least equal to τ_{\max} and defining $t = 0$ at the beginning of the adaptation period of pattern ν , we see:

$$\langle \Theta_i(t) \rangle = \alpha \frac{\tau_{\max}}{2\Delta} (\xi_i^\nu - \xi_i^{\nu+1})t. \quad (11)$$

For $\xi_i^{\nu+1} = \xi_i^\nu$, there is, on the average, no threshold adaptation, whereas for $\xi_i^{\nu+1} = -\xi_i^\nu$ the thresholds are adapted in the direction favouring a transition to the new spin state $\xi_i^{\nu+1}$. The expectation value of the time t^* to reach the critical threshold value for transition reads:

$$\langle t^* \rangle = \frac{1}{\alpha} \left(\frac{\Delta}{\tau_{\max}} - 1 \right). \quad (12)$$

So the duration of the adaptation period depends on the threshold adaptation value α , which is globally variable. This last equation only holds in the absence of noise. The effect of noise will be discussed later.

During the reset period the dynamics of the thresholds is described by:

$$\frac{d}{dt} \Theta_i(t) = -F(t - t_i) \Theta_i(t), \quad (13)$$

with t_i defined as the time at which neuron i changed his state. The exact shape of the function $F(t - t_i)$ is not important. In our simulations we assume that the dynamics governed by eqn (13) forces neuron i to retain its new state for at least a time $\tau_{\max}/2$ and resets the threshold to zero within a reset time equal to τ_{\max} .

5. THE EFFECT OF NOISE

So far, we have discussed the performance of our model in the absence of noise. In this section we will investigate the effect of noise due to the pseudo orthogonality of the stored patterns of the transition between two successive patterns. With noise, all neurons will not change their state corresponding to the next pattern simultaneously. As a result, the transition is not infinitely fast. We will assume that the transition is still fast enough, such that the transition time is negligible in comparison with t^* , the duration of the adaptation period.

The deviations of the local inputs from their expectation values as computed in the previous section are caused by the pseudo orthogonality of the patterns. During the adaptation period we may replace eqn (6) by:

$$I_i = \langle I_i \rangle + R_i, \quad (14)$$

with R_i a rest term, equal to:

$$R_i = \frac{1}{N-p} \sum_{\mu \neq \nu} \xi_i^\mu \sum_{j \neq i} \xi_j^\mu \times \left[\left(1 - \frac{\tau_{ij}}{\Delta} \right) \xi_j^\mu + \frac{\tau_{ij}}{\Delta} \xi_j^{\mu-1} \right] \xi_j^\nu. \quad (15)$$

For large N and p , we may replace R_i by a Gaussian noise term χ_0 with standard deviation σ_0 , obeying (see Appendix):

$$\sigma_0^2 = \langle R_i^2 \rangle \approx \frac{p}{N} \int_0^{\tau_{\max}} d\tau \rho(\tau) \times \left[\left(1 - \frac{\tau}{\Delta} \right)^2 + \left(\frac{\tau}{\Delta} \right)^2 \right]. \quad (16)$$

Through the local input, the noise terms affect the thresholds, as described by eqn (10). Equation (11) can now be rewritten:

$$\Theta_i(t) = \begin{cases} \alpha(\tau_{\max}/\Delta + \chi) \xi_i^\nu t & \text{for } \xi_i^{\nu+1} \neq \xi_i^\nu \\ \alpha \chi t & \text{for } \xi_i^{\nu+1} = \xi_i^\nu \end{cases} \quad (17)$$

with χ a time-independent noise term. The standard deviation of χ is not equal to σ_0 . In fact, the noise on the thresholds increases with each transition. There are two reasons for this: first the thresholds of spins

that did not flip still suffer by the noise which was present in the previous pattern state; and second, the adaptation periods of spins that did flip are reset at different times. If the transition is fast, as it should be, the second noise factor can be neglected, relative to the first one. Given the standard deviation σ_ν during the pattern state ν , we can calculate the standard deviation $\sigma_{\nu+1}$:

$$\sigma_{\nu+1}^2 \approx \frac{1}{2} \sigma_\nu^2 + \sigma_0^2. \quad (18)$$

We see that the standard deviation is bounded by:

$$\sigma \approx \sqrt{2} \sigma_0. \quad (19)$$

To obtain proper transitions, this standard deviation σ has to be small, to prevent erroneous spin flips induced by the noise. For the rest of the paper we will therefore assume:

$$\sigma \ll \frac{\tau_{\max}}{\Delta}. \quad (20)$$

To get an idea of the shape and the duration of the transition, we will derive an approximate equation for the dynamics of the system in terms of pattern overlaps. Again, we start with the system in state ν , at the beginning of the adaptation period. The pattern overlaps are given by the equation:

$$q^\mu(t) = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu s_i(t). \quad (21)$$

We write the expectation value of the local input $I_i(t)$ in terms of the pattern overlaps with the patterns:

$$\begin{aligned} \langle I_i(t) \rangle &= \left\langle \sum_j J_{ij}(\tau_{ij}) s_j(t - \tau_{ij}) \right\rangle \\ &= \left\langle \int d\tau \rho(\tau) \sum_j J_{ij}(\tau) s_j(t - \tau) \right\rangle \\ &= \left\langle \int d\tau \rho(\tau) \sum_\mu q^\mu(t - \tau) \right. \\ &\quad \left. \times \left[\left(1 - \frac{\tau}{\Delta} \right) \xi_i^\mu + \frac{\tau}{\Delta} \xi_i^{\mu+1} \right] \right\rangle. \end{aligned} \quad (22)$$

In the summation over μ only the terms ν and $\nu + 1$ are nonzero. We neglect the term containing $\xi_i^{\nu+2}$, which is small compared to the basic noise χ_0 . Since in the time domain of the transition from state ν to $\nu + 1$, neuron i hardly ever has a spin opposite to both ξ_i^ν and $\xi_i^{\nu+1}$, we may write:

$$q(t) \equiv q^{\nu+1}(t) = 1 - q^\nu(t), \quad (23)$$

with $q(t)$ just the fraction of already flipped spins at time t . Focussing on the case $\xi_i^{\nu+1} \neq \xi_i^\nu$, we find (see Appendix):

$$\langle I_i(t) \rangle = \xi_i^\nu \times \left\{ 1 - \frac{\tau_{\max}}{\Delta} - \int d\tau \rho(\tau) q(t - \tau) \left[2 - 3 \frac{\tau}{\Delta} \right] \right\}. \quad (24)$$

The real local field is just this expectation value plus a noise term:

$$I_i(t) = \langle I_i(t) \rangle + \chi_0 \\ \equiv \xi_i^y \left[1 - \frac{\tau_{\max}}{\Delta} - Q(t) - \chi_0 \right], \quad (25)$$

with $Q(t)$ equal to the integral in eqn (24). The noise on the thresholds depends on the noise on the local fields and on the history of the neurons. For simplicity, we take the history dependent noise into account by assuming that the noise on both the local fields and the thresholds is Gaussian with a standard deviation σ . With these assumptions, the thresholds for spins that should flip, read:

$$\Theta_i(t) = \alpha \xi_i^y \left[\left(\frac{\tau_{\max}}{\Delta} + \chi \right) t + \int_0^t ds Q(s) \right]. \quad (26)$$

If we define the critical noise $\chi^*(t)$ as the amount of noise where the threshold equals the local input ($I_i(\chi^*, t) = \Theta_i(\chi^*, t)$), it reads:

$$\chi^*(t) = -\frac{\tau_{\max}}{\Delta} + \frac{1 - Q(t) - \alpha \int_0^t ds Q(s)}{1 + \alpha t}. \quad (27)$$

If the noise term χ for a particular neuron is greater than the critical noise term $\chi^*(t)$, either this neuron has already flipped its spin or it will flip its spin in the next step. With $\Phi(\chi)$ the Gaussian noise distribution function, the fraction of flipped spins obeys:

$$q(t+1) = \int_{\chi^*(t)}^{\infty} d\chi \Phi(\chi) = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{\chi^*(t)}{2\sigma_0} \right) \right]. \quad (28)$$

This shows that the fraction of flipped spins $q(t)$ depends, through $\chi^*(t)$, on its own history (see eqns (24)–(28)): the already flipped spins trigger the other spins.

Equation (28) is difficult to analyze mathematically, but easy to simulate on a computer. Assuming a homogeneous distribution of delays, we find transitions like those displayed in Figure 2. In Figure 2a the standard deviation σ_0 is small, whereas in Figure 2b it is large ($\sigma_0 = .25 \times \tau_{\max}/\Delta$). A larger standard deviation has two consequences: a shorter adaptation period and a less sharp transition. In our calculations we assumed that the duration of the transition is relatively short, at least shorter than the reset period. Simulations show that if the condition expressed in eqn (20) is fulfilled, our assumption is correct, except in extreme cases when the threshold adaptation constant is very small.

6. SIMULATION RESULTS

To test our theories, we have simulated our model on a computer. A sequence of three pseudo orthogonal patterns, each with duration $\Delta = 25$, is stored in a fully connected one-layered network consisting of 400 neurons, using eqn (1). The transmission de-

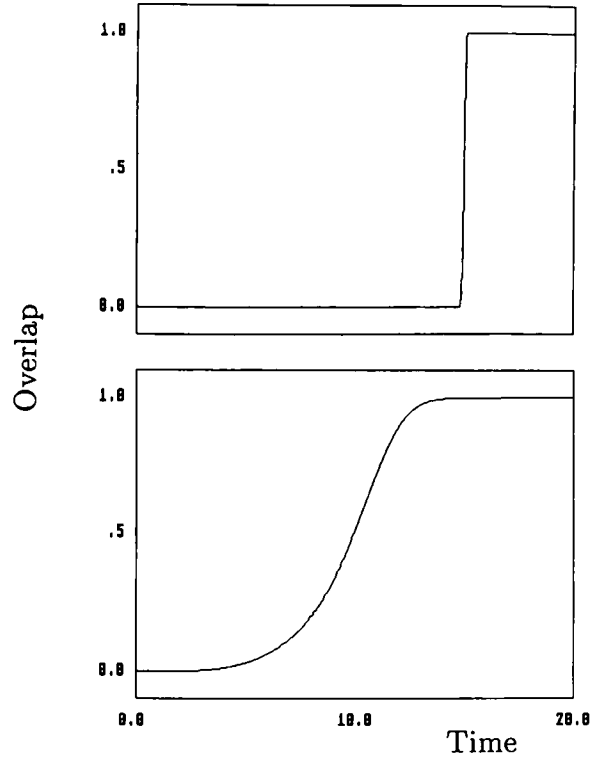


FIGURE 2. Transition between two successive patterns. The origin of time is the beginning of the adaptation period preceding the transition of pattern state ν to $\nu + 1$. The overlap of the network with pattern state $\nu + 1$ is shown as a function of time for two different values of the standard deviation σ_0 . (a) Small standard deviation: $\sigma_0 = .0025 \times \tau_{\max}/\Delta$; (b) Large standard deviation: $\sigma_0 = .25 \times \tau_{\max}/\Delta$. Lifetime of a pattern during the learning stage $\Delta = 25$. Maximum delay $\tau_{\max} = 10$.

lays in the network are homogeneously distributed between zero and $\tau_{\max} = 10$. At the beginning of the retrieval phase we impose the first pattern, slightly distorted with noise, on the network. The network reconstructs this pattern and replays the stored cycle. In Figure 3 the overlap q^2 with the second pattern is shown as a function of time for three different values of the threshold adaptation constant α . The duration of the cycle decreases with increasing α , in agreement with eqn (12).

To store and replay sequences of biased patterns with bias $a = \langle \xi_i^\mu \rangle \forall \mu$, the learning rule is generalized to:

$$\delta J_{ij}(\tau_{ij}) \propto s_i(t)[s_j(t - \tau_{ij}) - a]. \quad (29)$$

It is quite easy to prove that with a proper normalisation of the couplings, the dynamics with biased patterns is equivalent to the dynamics with unbiased patterns, except for an increase in the standard deviation of the noise. This increase is due to the fact that a smaller fraction of the thresholds is reset at every transition. Equation (18) has to be rewritten:

$$\sigma_{i+1}^2 \approx \frac{1 + a^2}{2} \sigma_i^2 + \sigma_0^2, \quad (30)$$

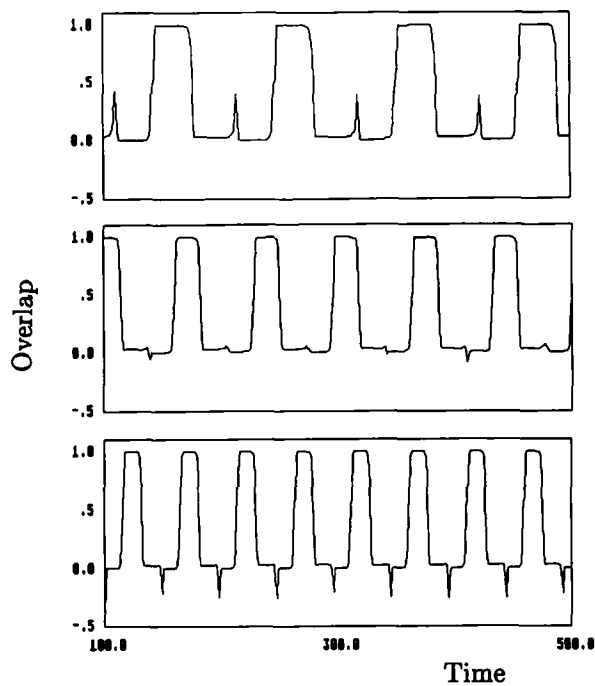


FIGURE 3. Overlap with the second pattern of a cycle consisting of three patterns as a function of time for different threshold adaptation constants α . (a) Slow replay: $\alpha = 1/2\tau_{\max}$; (b) Normal replay: $\alpha = 1/\tau_{\max}$; (c) Fast replay: $\alpha = 2/\tau_{\max}$; Number of neurons $N = 400$. Lifetime of a pattern during the learning stage $\Delta = 25$. Maximum delay $\tau_{\max} = 10$.

leading to:

$$\sigma \approx \sqrt{\frac{2}{1-a^2}} \sigma_0. \quad (31)$$

The equations describing the dynamics of the transition need some modifications, but qualitatively there is no difference. Simulation results with cycles of three biased patterns ($a = -0.5$), stored in a neural network consisting of 625 neurons and replayed at three different speeds, are shown in Figure 4.

7. CONCLUSIONS AND DISCUSSION

The aim of this study was to investigate a neural network model that was able to regenerate sequences of previously stored patterns at a variable speed. Examples of such behaviour have been found in many biological systems, in particular in those related to motor control. For example, when a complex movement has been learned, it can be generated easily at any desired speed by scaling all phases of the movement equally (Carter & Shapiro, 1984).

The nature of biological systems imposed some constraints on our model, such as the fact that transmission delays between neurons may be different for various neurons, but fixed in time. The results of our study demonstrate that adaptation of the threshold

of neurons is sufficient to allow regeneration of a stored sequence at a variable speed. The fact that the threshold of neurons can be modulated by neurotransmitters, which affect the intrinsic membrane properties of the neuron, has been demonstrated in neurophysiology (Hounsgaard, Hultborn, & Kiehn, 1986). The effect is that the temporal properties of the neuron can be adjusted, giving it transient or more sustained firing properties. This is just what is required in our model. In addition, the recruitment threshold of motoneurons in human arm muscles is modulated separately for each muscle when going from isometric contractions (zero velocity) to slow movements (Tax, Denier van der Gon, Gielen, & Van den Tempel, 1989).

In our theoretical analysis, we assumed *synchronous* processing of the spins without noise, that is, at *zero temperature*. Asynchronism can be modelled by choosing a unit i at random at each time step $1/N$, computing the local field h_i , and aligning the spin with this local field. If the maximum transmission delay τ_{\max} is much larger than the unit of time 1, asynchronous processing leads to the same results as synchronous processing. A small nonzero temperature leads to noise that can be incorporated in the analysis of Section 5.

Another aspect of our model which deserves some discussion, concerns the behaviour of the neuronal

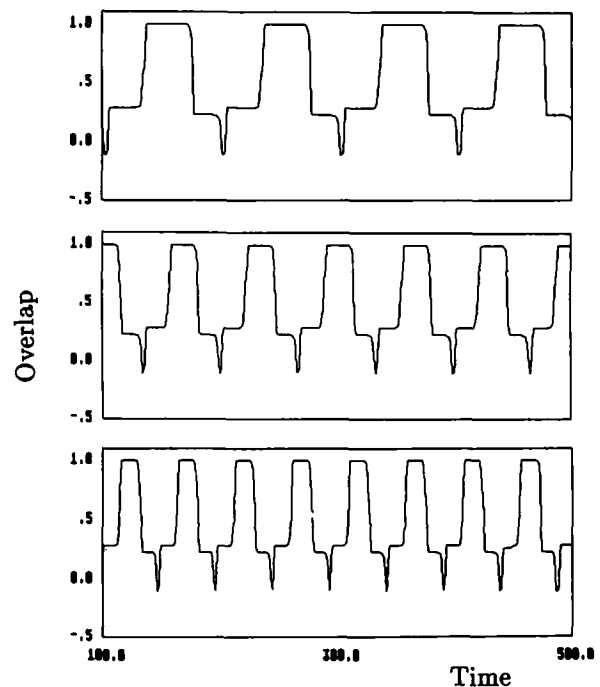


FIGURE 4. Overlap with the second pattern of a cycle consisting of three patterns with a bias $a = -0.5$ as a function of time. (a) Slow replay: $\alpha = 1/2\tau_{\max}$; (b) Normal replay: $\alpha = 1/\tau_{\max}$; (c) Fast replay: $\alpha = 2/\tau_{\max}$. Number of neurons $N = 625$. Lifetime of a pattern during the learning stage $\Delta = 25$. Maximum delay $\tau_{\max} = 10$.

threshold when the neuron changes from one state to another. As explained in the text, the threshold should change after this transition, but not too fast, lest the system fall back to the previous state. This does not seem to be a severe restriction since a sudden reset of the threshold is not physiologically plausible. The time needed for a reset of the threshold is on the order of $\tau_{\max}/2$. Considering that time delays in a biological neural network are on the order of milliseconds, this seems a reasonable assumption.

The Hebbian-type learning rule defined in eqn (29) allows the correct storage of patterns with the same mean activity but otherwise uncorrelated. This restriction on the class of patterns that can be stored might be too severe from a biological point of view. Guyon, Personnaz, Nadal, and Dreyfus (1988) showed that with a local iterative Widrow–Hoff-type learning rule (see Diederich & Oppel, 1987) the storage and retrieval of time sequences of correlated patterns can be realized. It might be interesting to study whether this learning rule, combined with a broad distribution of delays and dynamical thresholds, facilitates the regeneration of sequences of correlated patterns at different speeds.

REFERENCES

- Carter, M. C., & Shapiro, D. C. (1984). Control of sequential movements: Evidence for generalized motor programs. *Journal of Neurophysiology*, **52**, 787–796.
- Coolen, A. C. C., & Gielen, C. C. A. M. (1988). Delays in neural networks. *Europhysics Letters*, **7**, 281–285.
- Diederich, S., & Oppel, M. (1987). Learning of correlated patterns in spin-glass networks by local learning rules. *Europhysics Letters*, **58**, 949–952.
- Guyon, I., Personnaz, L., Nadal, J. P., & Dreyfus, G. (1988). Storage and retrieval of complex sequences in neural networks. *Physical Review A*, **38**, 6365–6372.
- Hebb, D. O. (1949). *The organization of behaviour*. New York: Wiley.
- Herz, A., Kühn, R., & van Hemmen, J. L. (1989). Hebbian learning reconsidered: Representation of static and dynamic objects in associative neural nets. *Biological Cybernetics*, **60**, 457–467.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, **79**, 2554–2558.
- Horn, D., & Usher, M. (1989). Neural networks with dynamical thresholds. *Physical Review A*, **40**, 1036–1044.
- Hounsgaard, J., Hultborn, H., & Kiehn, O. (1986). Transmitter-controlled properties of α -motoneurons causing long-lasting motor discharge to brief excitatory inputs. In H. Freund, U. Büttner, B. Cohen, & J. Noth (Eds.), *Progress in brain research*, **64**, (pp. 39–49). Amsterdam: Elsevier.
- Kleinfeld, D. (1986). Sequential state generator by model neural networks. *Proceedings of the National Academy of Sciences of the United States of America*, **81**, 9496–9473.
- Peretto, P., & Niez, J. J. (1986). Collective properties of neural networks. In E. Bienenstock, F. Fogelman-Soulié, & G. Weisbuch (Eds.), *Disordered systems and biological organization* (pp. 171–185). Berlin: Springer-Verlag.
- Sompolinsky, H., & Kanter, I. (1986). Temporal association in asymmetric neural networks. *Physical Review Letters*, **57**, 2861–2864.
- Tax, A. A. M., Denier van der Gon, J. J., Gielen, C. C. A. M., & Van den Tempel, C. M. M. (1989). Differences in the activation of m. biceps brachii in the control of slow isotonic movements and isometric contractions. *Experimental Brain Research*, **76**, 55–63.

APPENDIX

For the model without noise, the expectation value for the local input to neuron i can be evaluated in the following way:

$$\begin{aligned}
 \langle I_i(t) \rangle &= \left\langle \sum_{j \neq i} J_{ij}(\tau_{ij}) s_j(t - \tau_{ij}) \right\rangle \\
 &= \left\langle \int_0^{\tau_{\max}} d\tau \rho(\tau) \sum_{j \neq i} J_{ij}(\tau) s_j(t - \tau) \right\rangle \\
 &= \left\langle \int_0^{\tau_{\max}} d\tau \rho(\tau) \sum_{j \neq i} J_{ij}(\tau) \xi_j^{\tau-1} + \int_0^{\tau} d\tau \rho(\tau) \sum_{j \neq i} J_{ij}(\tau) \xi_j^{\tau} \right\rangle \\
 &= \left\langle \int_0^{\tau_{\max}} d\tau \rho(\tau) \frac{1}{N-p} \sum_{\mu=1}^p \xi_i^{\mu} \sum_{j \neq i} \right. \\
 &\quad \times \left[\left(1 - \frac{\tau}{\Delta}\right) \xi_j^{\mu} + \frac{\tau}{\Delta} \xi_j^{\mu-1} \right] \xi_j^{\tau-1} \right\rangle \\
 &\quad + \left\langle \int_0^{\tau} d\tau \rho(\tau) \frac{1}{N-p} \sum_{\mu=1}^p \xi_i^{\mu} \sum_{j \neq i} \right. \\
 &\quad \times \left[\left(1 - \frac{\tau}{\Delta}\right) \xi_j^{\mu} + \frac{\tau}{\Delta} \xi_j^{\mu-1} \right] \xi_j^{\tau} \right\rangle \\
 &= \int_0^{\tau_{\max}} d\tau \rho(\tau) \left(1 - \frac{\tau}{\Delta}\right) \xi_i^{\tau-1} + \int_0^{\tau_{\max}} d\tau \rho(\tau) \frac{\tau}{\Delta} \xi_i^{\tau} \\
 &\quad + \int_0^{\tau} d\tau \rho(\tau) \left(1 - \frac{\tau}{\Delta}\right) \xi_i^{\tau} + \int_0^{\tau} d\tau \rho(\tau) \frac{\tau}{\Delta} \xi_i^{\tau+1}. \quad (A1)
 \end{aligned}$$

For $t \geq \tau_{\max}$, the integration for $t > \tau > \tau_{\max}$ can be left out:

$$\begin{aligned}
 \langle I_i(t) \rangle &= \int_0^{\tau_{\max}} d\tau \rho(\tau) \left(1 - \frac{\tau}{\Delta}\right) \xi_i^{\tau} + \int_0^{\tau_{\max}} d\tau \rho(\tau) \frac{\tau}{\Delta} \xi_i^{\tau+1} \\
 &= \left(1 - \frac{\tau_{\max}}{2\Delta}\right) \xi_i^{\tau} + \frac{\tau_{\max}}{2\Delta} \xi_i^{\tau+1}. \quad (A2)
 \end{aligned}$$

The standard deviation of the Gaussian noise term follows from (we assume $N \rightarrow \infty$ and $p \ll N$):

$$\begin{aligned}
 \langle R_i^2 \rangle &= \left\langle \left\{ \frac{1}{N-p} \sum_{\mu \neq i} \xi_i^{\mu} \sum_{j \neq i} \left[\left(1 - \frac{\tau_{ij}}{\Delta}\right) \xi_j^{\mu} + \frac{\tau_{ij}}{\Delta} \xi_j^{\mu-1} \right] \xi_j^{\tau} \right\}^2 \right\rangle \\
 &= \left\langle \frac{1}{N^2} \int d\tau \rho(\tau) \sum_{\mu} \sum_j \right. \\
 &\quad \times \left[\left(1 - \frac{\tau}{\Delta}\right)^2 + \left(\frac{\tau}{\Delta}\right)^2 + \left(1 - \frac{\tau}{\Delta}\right) \frac{\tau}{\Delta} \xi_j^{\mu} \xi_j^{\mu+1} \right] \right\rangle \\
 &= \frac{p}{N} \int d\tau \rho(\tau) \left[1 - 2 \frac{\tau}{\Delta} + 2 \left(\frac{\tau}{\Delta}\right)^2 \right]. \quad (A3)
 \end{aligned}$$

For the homogeneous distribution $\rho(\tau) = 1/\tau_{\max}$, this equation leads to:

$$\langle R_i^2 \rangle = \frac{p}{N} \left[1 - \frac{\tau_{\max}}{\Delta} + \frac{2}{3} \left(\frac{\tau_{\max}}{\Delta}\right)^2 \right]. \quad (A4)$$

If the effect of noise is incorporated, the expectation value of the local input in terms of pattern overlaps reads:

$$\begin{aligned}
 \langle I_i(t) \rangle &= \left\langle \int d\tau \rho(\tau) \sum_{j \neq i} J_{ij}(\tau) s_j(t - \tau) \right\rangle \\
 &= \left\langle \int d\tau \rho(\tau) \frac{1}{N} \sum_{\mu} \xi_i^{\mu} \sum_j \right. \\
 &\quad \times \left[\left(1 - \frac{\tau}{\Delta}\right) \xi_j^{\mu} + \frac{\tau}{\Delta} \xi_j^{\mu-1} \right] s_j(t - \tau) \right\rangle
 \end{aligned}$$

$$\begin{aligned}
&= \left\langle \int d\tau \rho(\tau) \sum_{\mu} q^{\mu}(t-\tau) \left[\left(1 - \frac{\tau}{\Delta}\right) \xi_i^{\mu} + \frac{\tau}{\Delta} \xi_i^{\mu+1} \right] \right\rangle \\
&= \left\langle \int d\tau \rho(\tau) q^{\nu}(t-\tau) \left[\left(1 - \frac{\tau}{\Delta}\right) \xi_i^{\nu} + \frac{\tau}{\Delta} \xi_i^{\nu+1} \right] \right\rangle \\
&\quad + \left\langle \int d\tau \rho(\tau) q^{\nu+1}(t-\tau) \left[\left(1 - \frac{\tau}{\Delta}\right) \xi_i^{\nu+1} + \frac{\tau}{\Delta} \xi_i^{\nu+2} \right] \right\rangle \\
&= \int d\tau \rho(\tau) \left\{ [1 - q(t-\tau)] \left[\left(1 - \frac{\tau}{\Delta}\right) \xi_i^{\nu} + \frac{\tau}{\Delta} \xi_i^{\nu+1} \right] \right. \\
&\quad \left. + q(t-\tau) \left(1 - \frac{\tau}{\Delta}\right) \xi_i^{\nu+1} \right\}, \quad (\text{A5})
\end{aligned}$$

with $q(t) \equiv q^{\nu+1}(t) = 1 - q^{\nu}(t)$. For $\xi_i^{\nu} = \xi_i^{\nu+1}$, this leads to:

$$\begin{aligned}
\langle I_i(t) \rangle &= \xi_i^{\nu} \int d\tau \rho(\tau) \left[1 - q(t-\tau) \frac{\tau}{\Delta} \right] \\
&= \xi_i^{\nu} \left[1 - \int d\tau \rho(\tau) q(t-\tau) \frac{\tau}{\Delta} \right]. \quad (\text{A6})
\end{aligned}$$

Whereas, for $\xi_i^{\nu} \neq \xi_i^{\nu+1}$, we find:

$$\begin{aligned}
\langle I_i(t) \rangle &= \xi_i^{\nu} \int d\tau \rho(\tau) \\
&\quad \times \left\{ 1 - 2 \frac{\tau}{\Delta} - q(t-\tau) \left[2 - 3 \frac{\tau}{\Delta} \right] \right\} \\
&= \xi_i^{\nu} \left\{ 1 - \frac{\tau_{\max}}{\Delta} - \int d\tau \rho(\tau) q(t-\tau) \left[2 - 3 \frac{\tau}{\Delta} \right] \right\}. \quad (\text{A7})
\end{aligned}$$

NOMENCLATURE

t : time.

N : number of neurons in the network.

$s_i(t)$: spin of neuron i .

τ : transmission delay.

τ_{ij} : transmission delay for signals from neuron j to i .

J_{ij} : connectivity strength of neuron j to neuron i .

p : number of patterns.

ξ_i^{μ} : spin of neuron i in pattern state μ .

τ_{\max} : maximum delay in the network.

Δ : lifetime of a single pattern in the sequence during the learning stage.

$\rho(\tau)$: distribution function of delays.

$h_i(t)$: local field of neuron i .

$\Theta_i(t)$: threshold of neuron i .

$I_i(t)$: local input of neuron i .

$g_k(t)$: the tendency of the system to proceed to the k -th pattern following on the present pattern.

c : threshold leakage constant.

α : threshold adaptation constant.

t^* : duration of the adaptation period.

R_i : difference between the local input of neuron i and its expectation value.

χ_0, χ : noise term.

σ_0, σ : standard deviation of χ_0 and χ , respectively.

σ_{ν} : standard deviation of noise with system in pattern state ν .

$q^{\mu}(t)$: pattern overlap with pattern μ .

$q(t)$: pattern overlap with new pattern state after the transition.

$Q(t)$: weighted integral of $q(t)$.

$\chi^*(t)$: critical noise.

$\Phi(\chi)$: distribution function of noise χ .

a : mean activity in the network.