The Journal of Neuroscience, November 1994, 14(11): 6435–6445

Correlations of Cortical Hebbian Reverberations: Theory versus Experiment

Daniel J. Amit,[1,2] Nicolas Brunel,[1] and M. V. Tsodyks[2,3]

[1]INFN, Sezione di Roma, Istituto di Fisica, Universita di Roma, La Sapienza, Roma, Italy, [2]Racah Institute of Physics, Hebrew University, Jerusalem, and [3]Center for Neural Computation, Hebrew University, Jerusalem, Israel

Interpreting recent single-unit recordings of delay activities in delayed match-to-sample experiments in anterior ventral temporal (AVT) cortex of monkeys in terms of reverberation dynamics, we present a model neural network of quasi-realistic elements that reproduces the empirical results in great detail. Information about the contiguity of successive stimuli in the training sequence, representing the fact that training is done on a set of uncorrelated stimuli presented in a fixed temporal sequence, is embedded in the synaptic structure. The model reproduces quite accurately the correlations between delay activity distributions corresponding to stimulation with the uncorrelated stimuli used for training. It reproduces also the activity distributions of spike rates on sample cells as a function of the stimulating pattern. It is, in our view, the first time that a computational phenomenon, represented on the neurophysiological level, is reproduced in all its quantitative aspects.

The model is then used to make predictions about further features of the physiology of such experiments. Those include further properties of the correlations, features of selective cells as discriminators of stimuli provoking different delay activity distributions, and activity distributions among the neurons in a delay activity produced by a given pattern. The model has predictive implications also for the dependence of the delay activities on different training protocols.

Finally, we discuss the perspectives of the interplay between such models and neurophysiology as well as its limitations and possible extensions.

[Key words: reverberations, delay memory tasks, temporal correlations, inferotemporal cortex, neural network model, realistic neurons, cognitive neurophysiology, learning]

The experiments of the Miyashita group (Miyashita, 1988; Miyashita and Chang, 1988; Sakai and Miyashita, 1991) on the delay activity in anterior ventral temporal (AVT) cortex of monkeys trained to perform delayed matching-to-sample tasks have disclosed significant correlations in the internal representations of stimuli chosen to be uncorrelated, when those are presented during training in a fixed sequence. The experiments themselves have very significant implications both for the underestimated domain of neurophysiology—single-unit recording—as well as for cognitive psychology (for a detailed discussion, see, e.g., Amit, 1992, 1993a,b). A detailed theoretical model for such experiments can much enhance their implications, as well as provide precise clues for tests and extensions.

In the first experiment (Miyashita and Chang, 1988) the monkey is trained to recognize and match pairs out of a set of visual patterns, with 16 sec separating sample from match. One observes selective enhancements of neural spike rates that persist for 16 sec after the removal of the stimulus. This fact constitutes strong evidence for (nonergodic) attractor dynamics (defined below) in the corresponding neural module of some 1 mm^2 of anterior ventral temporal cortex.[1] Persistent spike rate distributions in the absence of a driving stimulus must be maintained by local recurrent synaptic feedback. Indeed, the anatomy of cortex indicates (Braitenberg and Schuz, 1991) that the level of local feedback in a column of 1 mm^2 is sufficiently high to be able to maintain stimulus-specific delay activities (for more detailed discussion of this point, see, e.g., Amit, 1992). In a neural network, a Hebbian assembly, attractor dynamics implies that when the network is stimulated by an afferent stimulus its dynamics drive it, following the removal of the stimulus, to one of a restricted set of stable spike rate distributions. Each of these attractors has a basin of attraction that is the totality of stimuli leading the network to the same rate distribution. The set of attractors, Hebbian reverberations, is a property of the neural module. It is determined by the synaptic structure, formed in the learning process, which maintains the particular reverberation by means of the structured feedback. In this sense, each of the stable rate distributions is an internal representation of the set of stimuli leading to it. Similar concepts have also been recently introduced to account for activity distributions of neurons in the hippocampus of performing rats (McNaughton and Nadel, 1990).

A second experiment (Miyashita, 1988) provides information about the structure of these internal representations, that is, about the different spike rate distributions during the delay. The visual stimuli were presented, during training, in a fixed sequential order. It was then discovered that despite the facts that care was taken to ensure the absence of spatial correlations between the learned visual stimuli, and during testing stimuli were presented in a random order, the delay spike rate distributions displayed correlations. The reverberations corresponding to stimuli that had been nearer to each other in the training

Received July 19, 1993; revised Mar. 21, 1994; accepted Apr. 21, 1994.

The work of M.V.T. was supported in part by grants from the Ministry of Science and Technology of Israel and from Intel Corp.

Correspondence should be addressed to Daniel J. Amit at the Racah Institute of Physics.

[1] Some of the concepts used in this presentation go by different names in different communities. The main sets of equivalences are as follows: delay activity ≡ reverberation ≡ attractor; stimulus ≡ pattern; cell assembly ≡ neural module ≡ neural network. We will sometimes insert equivalent terminology in parentheses.

sequence had more highly correlated rate distributions. What the monkey's brain seems to be doing during learning is converting the temporal correlations, represented by the fixed sequential order, into spatial correlations between the corresponding internal representations (see also, e.g., Griniasty et al., 1992).

The description of the dynamics of an attractor neural network (ANN) as a set of stable attractors maintained by a structured (learned) feedback in the synaptic structure has been the central theme of the Hopfield program (Amit, 1989). While providing a very robust detailed description of associative memory in neural networks, it suffers from two serious shortcomings when confronted quantitatively with the above-mentioned experiments. The first is that experiment exhibits wide rate distributions in the delay activities, while the models have attractors with very sharp rate distributions—neurons are either quiescent or close to saturation. The second is the fact that in the Hopfield model, when uncorrelated stimuli are embedded into the synaptic matrix, the resulting attractors are also uncorrelated.

Clearly, neural networks with discrete elements cannot reproduce wide rate distributions in attractors; but the simple extension of the neural elements to integrate-and-fire spiking units is sufficient to produce analog distributions of spike rates in delay activities as well as attractors with realistically low spike rates (Amit and Tsodyks, 1991a,b, 1992). This modification of the models had been carried out in the context of autoassociative networks, that is, networks whose reverberations are very close in structure to the memorized stimuli. As such, this modification does not deal with the second issue, that of the appearance of correlations in the attractors stimulated by uncorrelated stimuli. The reason for this automatic auto-associativity is the fact that synaptic modifications, due to a stimulus presented for learning, depend exclusively on the neural activity correlations in that stimulus.

This problem was confronted in a recent study (Griniasty et al., 1992). It was found that attractor neural networks that also connect, in their synaptic structure, information about contiguous stimuli (patterns) learned in a sequence have correlated delay activities (attractors, reverberations) even though the learned stimuli are uncorrelated. For simplicity, the context was narrowed to networks of discrete elements; hence, rate distributions could not be directly confronted with experiment. Yet the model presents several attractive features.

(1) The delay activity distribution (an attractor) corresponding to a given learned stimulus (i.e., the delay activity provoked in the neural assembly by the presentation of that stimulus) is correlated with the delay activity corresponding to other stimuli until there is a separation of several patterns in the sequence of the learned patterns, despite the fact that the synaptic matrix connects only consecutive patterns (nearest neighbors) in the sequence.

(2) The correlation distance of the attractors, and the amplitudes of the correlations are robust to the parameters of the model.

(3) The fact that the empirical correlations of delay activities (attractors) can be reproduced theoretically by coupling only contiguous patterns in the learned sequence of stimuli concords nicely with naive scenarios of learning in the presence of attractors (Griniasty et al., 1992; Amit, 1993a).

(4) The appearance of such correlations between the different delay activities is a transcription, during the learning process, of temporal correlations in the training information into spatial (activity distribution) correlations of the internal representations of the different stimuli. In other words, this is an embryo of context sensitivity (see, e.g., Amit, 1993a).

Given the potential of these experiments in establishing a direct bridge between cognitive neurophysiology and modeling, we have undertaken to bring the two modifications together: to combine a network of quasi-realistic neural elements and a synaptic matrix that allows information about contiguous stimuli, in a training sequence of fixed order, to be imprinted. In some sense it is an inquiry into the domain of validity and robustness of the surprising result found by Griniasty et al. (1992). On the other hand, it is the challenge of finding the characteristics that would bring the model to a level of quantitative detailed agreement with as much information as is given by the short accounts of the experiments. The mere interpretation of delay activities as dynamic attractors and of their correlations as attractor correlations produced by learning in the presence of attractors leads directly to cognitive and neurophysiological predictions (Griniasty et al., 1992; Amit, 1993a). The construction of the detailed model leads, as will be shown below, to new experimental predictions.

Materials and Methods

Overview

We have considered a network of integrate-and-fire neurons operating in the presence of high levels of nonselective uncorrelated noise originating in the global spontaneous activity. The neuron is represented by its current to spike rate transduction function, which includes the effect of noise due to spontaneous activity (see, e.g., Amit and Tsodyks, 1991a,b). Such neurons are taken to represent the excitatory neurons of the network, the pyramidal cells. It is in the synaptic matrix connecting these neurons that learning is manifested. The synaptic matrix, representing the training process, was constructed to represent the inclusion of the information about the contiguity of patterns in the training sequence (Griniasti et al., 1992). Inhibition is taken to have fixed synapses and its role is to react in an inhibitory way, proportional to the mean level of activity in the excitatory network, so as to control the overall activity in the network. For comparison we have also studied a more traditional network that describes learning among all synapses, excitatory as well as inhibitory. Quite surprisingly, the results are essentially the same.

The delay activities are investigated by presenting to the neural module (cell assembly) one of the uncorrelated stimuli as a set of afferent currents into a subset of the neurons. These currents are removed after a short time and the network is allowed to follow the dynamics as governed by the feedback represented in the synaptic matrix. Eventually, the network arrives at a stationary distribution of spike rates. This is the delay activity distribution corresponding to the stimulus which has excited the network.

Below we describe the details of the model. It's performance was simulated on a Sparc 10 SUN station. The results of simulations were compared with the results of experiments of the Miyashita group (Miyashita, 1988; Miyashita and Chang, 1988; Sakai and Miyashita, 1991).

The excitatory network

The network is composed of N excitatory neurons and an associated inhibitory network. The excitatory neuron i ($i = 1,\ldots,N$) is characterized at time t by its incoming current $I_i(t)$ and its firing rate $V_i(t)$. Its afferent current is composed of feedback from the other excitatory neurons, hyperpolarizing current from the inhibitory network $-T_i(t)$, and an occasional external current $H_i^{\text{ext}}(t)$ representing the stimulus. The dynamics of the excitatory neurons are

$$\tau_{\text{exc}}\dot{I}_i = -I_i + \frac{1}{fN}\sum_{j\neq i} J_{ij}V_j - T_i + H_i^{\text{ext}}, \qquad (1)$$

where J_{ij} is the efficacy of the synapse connecting the presynaptic neuron j to the postsynaptic neuron i. f is the mean coding level of the stimuli (the mean fraction of active neurons per pattern). The incoming current

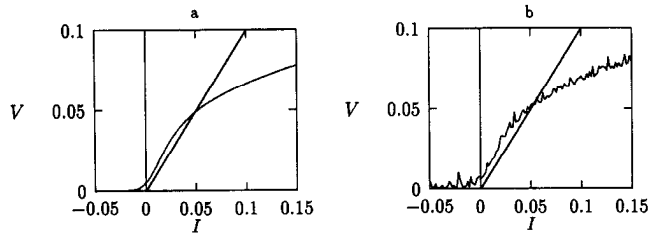

Figure 1. *a*, Spike frequency versus mean afferent current of an integrate-and-fire neuron. *b*, Same as *a* with superimposed spontaneous random activity of width 0.3% of saturation, used to model an excitatory element of the network. Zero current is spontaneous input only. Frequency axis in fraction of saturation frequency of the neuron. The parameters in *a* are given in the text.

into neuron i is converted into a spike rate via

$$V_i = \phi(I_i), \tag{2}$$

where ϕ is the current-to-rate transduction function for an excitatory neuron. The implemented network has $N = 4000$ excitatory neurons, a coding rate $f = 0.01$, and a current decay time constant $\tau_{exc} = 10$ msec. The transduction function of the neurons is shown in Figure 1. The figure represents the neuron's spike frequency versus the incoming mean synaptic current over and above the mean contributed by the global afferent spontaneous activity. Thus, the zero point on the current axis is the point at which the neuron senses only the effect of spontaneous activity. The fact that the frequency is not zero at this point is due to the fluctuating nature of the afferent spontaneous current.

Moreover, to the transduction function of the excitatory neurons we have added, by hand, a noisy component. In other words, for a given current input the frequency was calculated by adding to the frequency given by the smooth curve in Figure 1*a* a random term leading to Figure 1*b*. This extra component represents the fact that our model network has not been tuned enough to maintain autonomously a stable spontaneous activity. The noise added to the transduction function maintains the spontaneous activity artificially. Without it, neurons whose activity is not enhanced by the stimulus tend to have zero spike rates. (We return to this question in the concluding section). Note that in the transduction function used here the rate goes to 1 for very large values of the afferent current.

For this transduction function the neuronal response at $I = 0$ is $V \sim 0.005S$, where S is the saturation frequency of a pyramidal cell. The stable rates, as indicated by the upper intersection of ϕ with the straight line in Figure 1, are $V \sim 0.06S$. It would be $V \sim 30$ sec^{-1} if $S = 500$ sec^{-1} (see, e.g., the Discussion). Underlying the derivation of ϕ is an integrate-and-fire neuron with an absolute refractory period $\tau_{ARP} = 2$ msec, an RC time constant for the depolarization $\tau = 4\tau_{ARP}$, the depolarization threshold $\theta = 2.04$, the stationary depolarization due to spontaneous activity $\mu = 2$ (2% below threshold), and the remaining noise due to fluctuations in the spontaneous afferent $\sigma = 0.02$ (half of the distance between the threshold and the depolarization due to the spontaneous activity) (Amit and Tsodyks, 1991a,b). We performed simulations with different values of the parameters of the transduction function. In all cases the behavior of the network is similar. The only difference is in the absolute spike rates at which the selective neurons are active.

The inhibitory response

A separate inhibitory network is composed of N_{inh} inhibitory neurons. The inhibitory reaction is unstructured: every inhibitory neuron receives the same current from the excitatory neurons, and the entire inhibitory network becomes equivalent to a single inhibitory neuron characterized by its afferent current $I^{inh}(t)$ and its spike rate V^{inh}. The dynamics of the synaptic current into every inhibitory neuron are given by

$$\tau_{inh}\dot{I}^{inh} = -I^{inh} + \frac{1}{fN}\sum_i V_i, \tag{3}$$

with V_i the activity of neuron i in the excitatory network, and

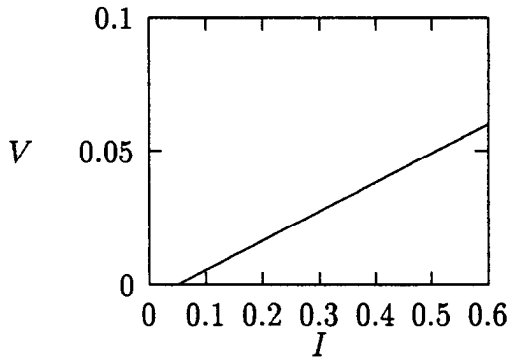

Figure 2. Spike rate versus synaptic current for the inhibitory neuron. It has a threshold $\theta^{inh} = 0.05$ in units of current and a slope $A = 0.1$. When the entire excitatory network operates at the imposed spontaneous rate level, the inhibitory neuron is 0.25 above threshold and has a frequency of 0.025 of saturation.

$$V^{inh} = \psi(I^{inh}), \tag{4}$$

where ψ is the inhibitory current-to-rate transduction function.

In other words, the inhibitory neuron is driven by a current proportional to the mean activity in the excitatory network. The inhibitory response is the same to every excitatory neuron: for all i, T_i in Equation 1 is given by

$$T = \frac{N_{inh}}{fN} V^{inh}. \tag{5}$$

The numerator in the coefficient is due to the replacement of N_{inh} inhibitory neurons by a single one. The denominator is part of the normalization of the inhibitory–excitatory synaptic connection.

Note that we are restricting the model to hyperpolarizing inhibition and no shunting. Equations 1, 3, and 5 together with the transduction functions ϕ and ψ describe the dynamics in full, once the synaptic matrix connecting the excitatory neurons, J_{ij}, is given.

The parameters we use are $\tau_{inh} = 2$ msec, $N_{inh} = fN = 40$. The inhibitory transfer function was chosen to be threshold linear (Treves, 1990). We checked that taking different inhibitory response functions, such as an "integrate-and-fire" one, yields the same results. It appears quite essential that the integration time of the inhibitory neuron be shorter than the integration time of the excitatory neuron (see also, e.g., Amit et al., 1991). The number of inhibitory neurons may seem low, given that typically there are 20% as many inhibitory neurons in the cortex as excitatory ones. But our inhibitory neuron is schematized to be connected to all excitatory neurons, while it should connect to 2–3% of them. When the connectivity is taken into account the number is actually slightly too high. This difference can be considered as a small rescaling of the inhibitory–excitatory synaptic strengths.

The learned synaptic matrix

The synaptic matrix reflects the learning of a sequence of p binary patterns (stimuli) ($\{\eta_i^\mu = 0, 1\}$; $i = 1,\ldots,n$; $\mu = 1,\ldots,p$) with coding level f, presented in a fixed order during the training session (Miyashita, 1988; Griniasti et al., 1992). The sequence is considered periodic; that is, pattern $p + 1$ is identical with pattern 1. The binary form of the stored memories is a symbolic indication of the neurons whose rates are (or are not) elevated by the stimulus. The actual rates actually maintained by the neurons in the network are analog (see, e.g., Amit and Fusi, 1993, for a discussion of this point of view).

Learning in the network is envisaged to have occurred only between the excitatory units. The effect of inhibition is only to control general activity levels in the network. We model the synaptic matrix in a Willshaw-like fashion (Willshaw et al., 1969); that is, the synaptic efficacies take only the values 0, 1, or a. The elements of the synaptic matrix J_{ij} are taken to be $J_{ij} = 1$ if there exists at least one pattern μ for which $\eta_i^\mu = \eta_j^\mu = 1$ (this is the original Willshaw prescription); $J_{ij} = a$ if $J_{ij} = 0$, by way of the Willshaw prescription, and there exists at least one pair of consecutive patterns μ, $\mu + 1$ for which $\eta_i^\mu = \eta_j^{\mu+1} = 1$ or $\eta_j^\mu = \eta_i^{\mu+1} = 1$; $J_{ij} = 0$ otherwise.

The parameter a represents the strength with which the contiguity of two patterns in the learned sequence is imprinted during learning (see, e.g., Griniasti et al., 1992). In the simulation we used $a = 0.5$. We checked that their behavior was robust in an interval around this value.

Here we have restricted ourselves to a synaptic matrix that is prescribed a priori. A matrix that performs in a similar way can be generated also by a quasi-realistic learning process. This process will be discussed in a forthcoming report (Brunel and Amit, 1994). A learning process that involves contiguous stimuli in a sequence was also considered by Foldiák (1991), but there the context does not include attractor dynamics, and consequently it would be hard to envisage how it could mix stimuli that are presented many seconds apart, as is the case in the experiments discussed here.

Sources of noise in the network

Fast noise on the excitatory gain function. As was explained earlier, this was added to correct for the fact that the spontaneous activity is not a stable mode of the system. Thus, at every integration step after the total synaptic current is computed, we generate a random number of Gaussian distribution with mean zero and width 0.3% of saturation spike rate, and add its absolute value to the frequency resulting from the gain function of Figure 1*a*, leading to Figure 1*b*. It is the noisy spike rates that then reenter the dynamics.

Inhomogeneity in inhibitory synapses. To mimic potential inhomogeneities in the inhibitory response, without paying the price of slowing down very significantly the simulation, we have generated a fixed random number, W_i, for each excitatory neuron from a Gaussian law of mean 1 and width 0.2, and substituted the inhibitory response T_i by TW_i, where T is given by Equation 5.

Procedure

The dynamics of the simulated network are fully described by Equations 1, 2, 3, and 5. During the simulations, (1) the stimuli were presented to the network by injecting an external current of strength $H^{ext} = H = 0.2$, during a short period t_p, into the set of neurons corresponding to one of the stimuli. (2) Then the external currents were stopped and the network flowed to the attractor. (3) Due to the presence of the dynamic noise mentioned above, the network does not reach a fixed point. The attractor is typified by the fact that the network dynamics make small fluctuations about a state in which a selective set of neurons have elevated frequencies. The arrival at such a state was detected by measuring the time-averaged mean activity in some specific populations of neurons. These populations are the neurons which are activated in any given stimulus μ, and their mean activities are

$$m_\mu(t) = \frac{1}{fN} \sum_{i=1}^{N} \eta_i^\mu V_i(t). \tag{6}$$

When the time averages of *all* the m^μs, over a moving window of 20 msec, did no longer vary significantly, the dynamics were stopped. (4) Then we recorded the time-averaged activities of all the neurons in the excitatory network during the last time window. The statistical properties of the delay activities were analyzed off line.

Correlations between attractors—definitions

Standard correlations. The delay activity of neuron i in the attractor provoked by stimulus μ is denoted by V_i^μ. We consider the distribution of mean rates V_i^μ in a sample S of N_s neurons as a random variable. Its mean is

$$\overline{V_\mu} = \frac{1}{N_s} \sum_{i \in S} V_i^\mu, \tag{7}$$

its variance is

$$\overline{\Delta V_\mu^2} = \frac{1}{N_s} \sum_{i \in S} (V_i^\mu)^2 - \overline{V_\mu}^2, \tag{8}$$

and the covariance of a pair of such random variables, corresponding to a given pair of attractors μ and ν, is

$$\mathrm{Cov}_{\mu\nu} = \frac{1}{N_s} \sum_{i \in S} V_i^\mu V_i^\nu - \overline{V_\mu}\,\overline{V_\nu}.$$

The correlation between the activity distributions in the two attractors μ and ν is

$$C_{\mu\nu} = \frac{\mathrm{Cov}_{\mu\nu}}{\sqrt{\overline{\Delta V_\mu^2}\,\overline{\Delta V_\nu^2}}}. \tag{9}$$

The mean correlation between two attractors at distance k is defined as

$$C_k = \frac{1}{p} \sum_{\mu} C_{\mu,\mu+k}, \tag{10}$$

where p is the total number of memorized attractors.

Kendal rank coefficients (KRC). The KRCs are calculated independently for each recorded neuron. For a given neuron one computes

$$U_{\mu\nu}^k = \mathrm{sign}[(V_i^\mu - V_i^\nu)(V_i^{\mu+k} - V_i^{\nu+k})], \tag{11}$$

with k ($1 \le k \le p/2$) fixed, where V_i^μ is the mean activity of neuron i in attractor μ. The KRC of neuron i at distance k is the mean value of the elements of the matrix U^k; that is,

$$R_k = \frac{2}{p(p-1)} \sum_{\mu<\nu} U_{\mu\nu}^k \tag{12}$$

(see, e.g., Snedecor and Cochran, 1969). The KRCs are then averaged over any given sample of recorded neurons.

Results

Performance of quasi-realistic models

Here we report that the phenomenon of the conversion of temporal correlations (contiguity of stimuli in the training sequence) into spatial correlations of neural delay activity distributions in a cell assembly persists when the model network is composed of quasi-realistic neural elements. We have tested the structure of the delay activity distributions (attractors, reverberations) of a local neural module composed of integrate-and-fire neurons. We find that the reverberations are correlated over a finite range of neighbors in the training sequence, much as in the experiment of Miyashita (1988), and well beyond the nearest neighbors information embedded in the synaptic matrix. The network has been investigated both by simulation as well as by approximate analytical means. The results are quite consistent. The theoretical considerations are relegated to a forthcoming report (Brunel 1994) so as not to burden unduly the present account.

To obtain an intuition into the nature of the results, we start with a qualitative description of the functioning of the network. When the network is subject to a given external stimulus, a subset of neurons is activated. We refer to this subset as the pattern. The actual level of activation of these neurons is determined and maintained by the strength of the input. In an auto-associative network, this activation pattern is maintained by the feedback in the synaptic interaction, after the removal of the stimulus. It is the distribution of neurons driven by the stimulus which is maintained by the feedback. The rates are determined by the network's dynamics (Amit and Tsodyks, 1991a,b). As was mentioned above, this property of the autoassociative model is due to the fact that the synaptic efficacies are enhanced only by the activity correlations of the corresponding pair of neurons in each learned pattern separately.

In the present model the synaptic efficacies are enhanced also by the activity correlations of pairs of neurons activated in two consecutive stimuli during training. The relative strength of the contributions to the synapses from the same pattern and from neighboring ones is a parameter a of the model (see Materials and Methods and Griniasty et al., 1992). Consequently, when a particular stimulus is presented for retrieval, following the

removal of the stimulus, neurons belonging to the successive and preceding patterns (i.e., those that would have been driven by the presentation of either of these two stimuli, and not by the present one) tend to get activated. It is the activity of the neurons directly stimulated, mediated by the cross-stimulus terms in the synaptic efficacies, which is exciting the neurons of the neighboring patterns. For some range of the parameter a, the indirectly activated neurons have a lower level of activity than the primary ones. This activation spreads from first neighbors in the training sequence to second neighbors, and so on, until the network reaches a stable attractor. The role of unstructured inhibition is to preclude the activation, at elevated frequencies, of neurons belonging to too many patterns. The qualitative form of the attractor would be like Figure 3. The spread of activation drops to spontaneous activity levels after a few neighbors in the sequence.

In Figure 3 we plot the spike rate of neurons versus the serial position number (SPN) of the corresponding pattern in the training sequence. If, for clarity of the argument, we ignore the possibility that neurons may belong to two different patterns, then the SPN classifies a set of neurons in the network, and all those would have the indicated spike rate. Under this approximation, Figure 3 is a full description of the reverberation. The central peak is the activity of the neurons belonging to the stimulated pattern, the next lower activities are those of all neurons belonging to the two nearest neighbor patterns, and so on. Different reverberations would be described by a similar figure, shifted to be centered around the stimulated pattern. A second attractor is indicated in Figure 3 as the dot-shaded cluster under the light line. It is chosen to be centered around a pattern not much removed in the sequence from the first, to emphasize the origin of attractor correlations in this type of attractors.

In fact, given that the attractors are of this form, one can directly conclude that they will be correlated if the corresponding stimuli are not much removed from each other. Attractors whose rate distributions overlap have groups of active neurons in common and hence are correlated. The correlations, which are directly related to the area of the overlap, will decrease monotonically with the separation of the two stimuli, much like the experimental data (crosses) shown in Figure 5. Moreover, the experimental Figure 6*b* has a form similar to that of Figure 3. The correspondence between the two figures is the following: the experimental figure corresponds to the rate distribution on a single *selective* cell, but every selective cell belongs to at least one of the patterns. Hence, it will have the highest rate if it belongs to the stimulating pattern and lower rates, which can be read from Figure 3, as one moves away from the stimulating pattern in the sequence. These simple considerations capture quite well the results of the detailed simulations and of the analysis.

The situation is somewhat more complicated when neurons belong to more than one of the uncorrelated patterns, as is observed in Figure 6*a*. Such occurrences are not too frequent when the fraction of neurons activated by each stimulus is low. We discuss rate distributions on such neurons below.

Numerical experiments and results

The uncorrelated patterns composing the sequence used for "training" were presented, one by one, as stimuli to the network. Each stimulus was presented as a set of external afferent currents that persist briefly (80 msec) and then are removed, allowing the neuronal module to find its natural pattern of delay activities. Time in the simulation is physical neural time, introduced via the integration time constants (see, e.g., Materials and Methods). The activities of all the neurons in the network in the delay period (the reverberating state, the attractor) were recorded and analyzed off line. We have dealt only with spike frequencies, both because the network dynamics had been set in these terms and because the attention of the experiment was on rates in single-unit recordings.

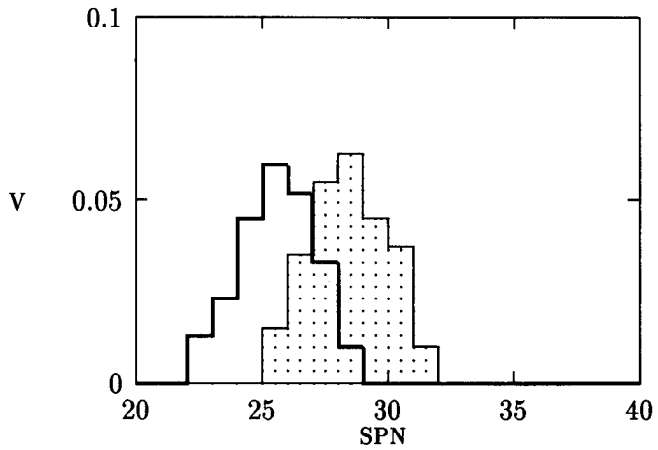

Figure 3. Schematic structure of attractors (see text): spiking rate of neuron classes versus the serial position number (SPN) of pattern in the training sequence. Two reverberations are sketched, those evoked by stimuli 25 (*heavy line*) and 28 (*light line, dot shaded*).

The end of a run was called when the temporal average, over a window of 20 msec, of the average frequencies of neurons in the foreground of all the memorized patterns stopped varying. This would typically take place after 100–200 msec, as can be seen in Figure 4. We plot the rates of five neurons as a function of time, during and following the presentation of the stimulus. One can observe neurons that are driven by the stimulus and remain active in the delay period with a lower frequency. Others are driven and are inactive following the removal of the stimulus, and some neurons are inactive during stimulus presentation and become active during the delay period. This type of behavior, corresponding to "error correction," should be compared with Figure 3 of Sakai and Miyashita (1991) (see also Discussion in Amit, 1992). Note that the term "error correction" is used here to indicate that the stimulus presented, at this phase in the task, is not identical to the one that was presented during learning, but is still in the basin of attraction of the attractor that was created. In such cases a neuron may behave differently during stimulation and during the delay. Since we do not attribute any specific meaning to the detailed structure of the representations, we cannot properly speak about errors and error correction.

At the end of each run we have single-unit recordings from all neurons in the network (rates in the delay period) corresponding to each of the attractors provoked by the corresponding stimulus. We then identify the "selective" units, that is, those units that have a rate significantly above the spontaneous rate in response to at least one stimulus. Among those, we select at random a sample of 50 neurons, similar in size to that selected in the experiment. These neurons are analyzed to produce the type of end products reported in the experiment of Miyashita (1988). The KRCs for this sample in the simulation (diamonds) and, for comparison, the corresponding experimental data

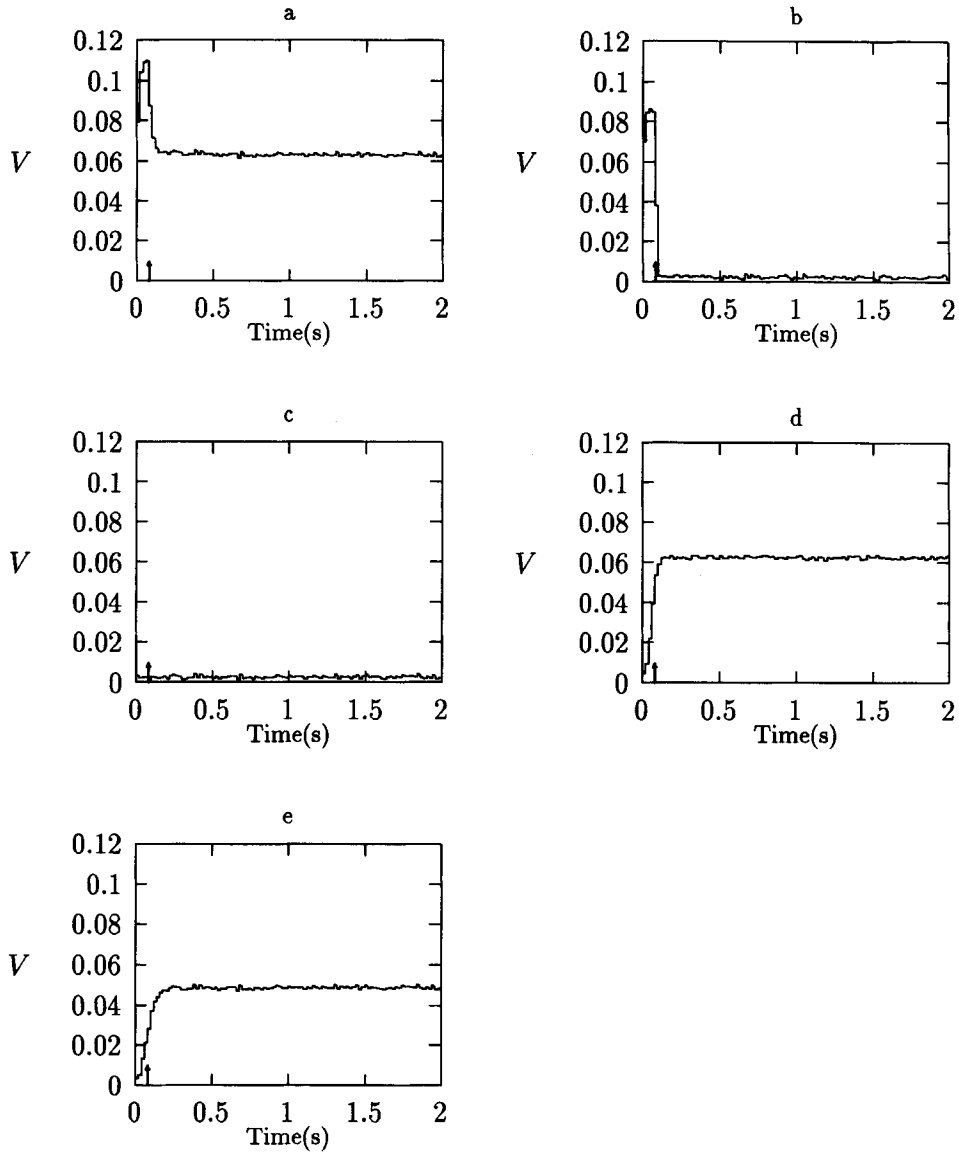

Figure 4. Spike frequency versus time of five neurons. Time indicated by *arrow* is when stimulus is removed: *a*, driven by stimulus active in delay (foreground neuron of this stimulus); *b*, driven by stimulus and inactive in reverberation (error corrected by network dynamics); *c*, not driven by stimulus and inactive in delay (background neuron for this pattern); *d*, not driven by stimulus and active in delay with rate as in *a* (error correction, foreground of stimulus); *e*, undriven by stimulus and active in delay (background of presented stimulus but foreground of stimulus nearby in the sequence).

(crosses) (Fig. 3*c* of Miyashita, 1988) are shown in Figure 5*a*. Next, a subsample is selected of those selective neurons for which the first KRC is significant ($P < 0.05$) (see, e.g., Snedecor and Cochran, 1969; Miyashita, 1988). Those correspond to neurons with first neighbor KRCs greater than 0.2 and they comprise about one-half of the neurons of the sample. Their KRCs are compared with the experimental results in Figure 5*b*. Note that the apparent discrepancy between the theoretical and experimental results is not significant. It is due to the representation of errors as standard errors. The two sets of data are consistent; that is, they are within the variance of each other.

We do not know if the crossings of the experimental and theoretical curves in Figure 5 are significant. Our guess is that it is due to fluctuations and may appear also in plotting, on the same graph, the correlations for two different monkeys.

We have also analyzed the spike rate distributions in the delay period in sample individual cells, corresponding to Figure 3, *a* and *b*, and Figure 1, *c* and *e* of Miyashita (1988). This was done in two ways. First, the rate distribution on individual neurons as a function of the serial position of the stimulus in the training sequence (SPN) was produced for two cells (see, e.g., Fig. 7). The two cells were chosen to give distributions similar to the two cells chosen in the experiment (Fig. 6). We see in each that the cell represented on the left is one that participates in two stimuli and hence has two peaks. Participating in two stimuli implies, due to the correlations generated in the learning process, participation in the delay activities of neighboring stimuli in the sequence. This is expressed by the fact that the peaks have a finite width.

Next we constructed the delay frequency histogram for a given cell when the network is stimulated by both learned patterns and unlearned ones. For the latter we have used stimuli uncorrelated with any of the patterns coded into the synaptic matrix. These results are presented in Figure 9 for comparison with the experimental results reproduced in Figure 8. Both in the experiment and in the model all "new" stimuli leave the network with all neurons at spontaneous activity levels during the delay period. Most learned stimuli leave a given neuron indifferent (spontaneous activity). Only a few stimuli will produce elevated rates. The comparison is quite satisfactory.

Additional information in model network

KRC sensitivity to number of memories

Certain features of the KRCs are peculiar to the special way in

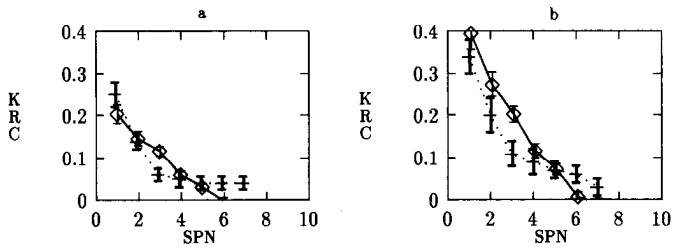

Figure 5. KRCs versus separation in the training sequence. Symbols are average KRCs over cell sample. Error bars are SEs (rms/$\sqrt{n}$). *n*, number of neurons in sample. +, experiment; ◇, simulation. *a*, Samples of selective neurons (experiment, 57 cells; simulation, 72 cells). *b*, Subsample of neurons with enhanced first neighbor KRCs (see text) (experiment, 28 cells; simulation, 25 cells). Experimental data are from Figure 3*c* of Miyashita (1988).

which they are defined, and are not descriptive of the dynamics. In particular, the KRCs are very sensitive to the number of memorized stimuli, and the canonical correlations are not. In other words, the same neural assembly storing a high number of learned stimuli will have very similar structure of the correlations between the corresponding delay activities to those of an assembly storing a low number of memories, but very different KRCs. Moreover, for the same network with the same synaptic structure, choosing samples of delay activity distributions of a different size leads to very different KRCs.

This simple fact is predictive. In other words, our model indicates that the reverberations showing the degree of KRCs reported in the experiments should have a very high level of standard correlation coefficients.

In Figure 10 we present an example of the dependence of the KRCs on the number of attractors used in the statistical analysis. To do this we have taken 20 of the delay activity distributions obtained in the simulation that produced Figure 5*b*. We use the same number of sample neurons. The behavior of the standard correlations versus SPN remains, of course, the same, since the attractors are unchanged. A detailed analysis of this difference will be presented in a forthcoming report (Brunel 1994).

Standard correlations between reverberations

It is quite significant that in order to reproduce the KRCs of Miyashita (1988) the reverberations have much higher correlation coefficients. In fact, the correlations of delay activities corresponding to successive stimuli in the training sequence can become as high as 0.89 (see, e.g., Fig. 11), compared with 0.2–0.3 for the corresponding KRCs. We have therefore checked

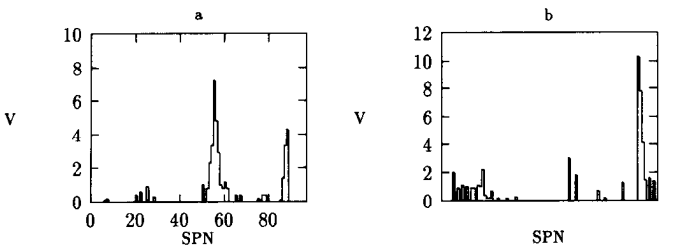

Figure 6. Average delay discharge rate versus serial position of the stimulus in two selective cells (experiment). Activity for learned images is reproduced from Figure 3, *a* and *b*, of Miyashita (1988). Note that the first cell participates in the clusters representing two different learned stimuli.

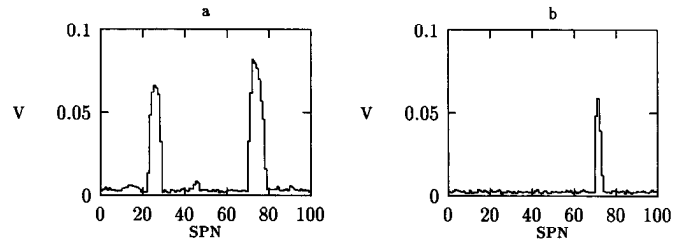

Figure 7. Average delay discharge rate versus serial position separation in two selective cells (simulation).

that even at this high level of correlation, single-unit recordings of several randomly chosen, selective neurons can distinguish between the different attractors. In our view it is the conventional correlations between the different delay activities that determine the ability of the local module to communicate the results of its computation down the line. We therefore present these correlations as an item for further experimental confrontation.

For comparison with the KRCs discussed above, we present in Figure 11 the averaged standard correlations between attractors, defined in Materials and Methods, versus serial position number for the same dynamical model and an identical set of memorized patterns. We find that the correlation coefficient of attractors corresponding to stimuli at separation 1 in the sequence is as high as 0.89. This raises the issue of the distinguishability of the two attractors based on single-unit recordings of a relatively small number of neurons.

The answer can be read from Figure 7. Considering the single narrow peak in Figure 7*b*, one observes that the rates in the side columns of the peak are significantly lower than in the central column. Hence, in this case this neuron distinguishes between the different stimuli. This neuron belongs to the foreground group of only one of the pure, uncorrelated stimuli that provoke the activities represented by the high columns. It belongs to the foreground of the stimulus for which it gives the highest rate. It is also recruited into the attractors that are provoked by the neighboring stimuli. Neurons are sampled if they are selective, that is, if they manifest a high rate in response to at least one stimulus. Thus, typically, every selective neuron will distinguish between the different reverberations.

There are two provisos, first, that the frequencies have to be averaged on a time scale long compared to the time scale of fluctuations of the spontaneous activity. But more important, one should try to keep away from situations such as the wide peak in Figure 7*a*. Note that the neuron represented in this figure cannot distinguish a few attractors near the center of the peak.

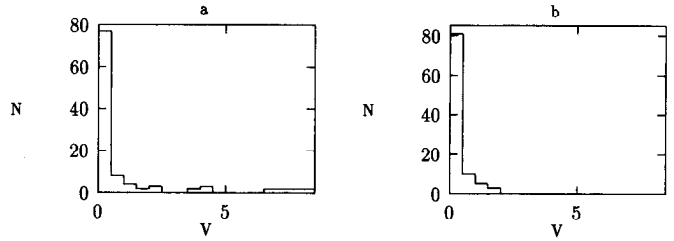

Figure 8. Spike rate distribution of average firing rate in delay period for 97 learned (*a*) and new (*b*) pictures in a given cell (experiment). Reproduced from Figure 1, *c* and *e*, of Miyashita (1988).

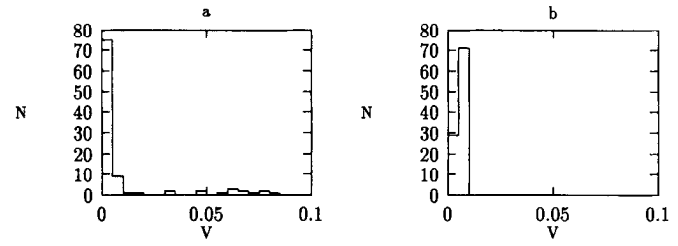

Figure 9. Spike rate distribution of average firing rate in delay period for 100 learned (*a*) and new (*b*) stimuli in a given cell (simulation): number of stimuli N versus rate V, in units of saturation spike rate. Bin width is 0.005.

The frequencies are too close. A closer look at this situation discloses that this peak is unusually wide, and the frequencies at its center are atypically high. Such neurons should be avoided as classifiers of reverberations!

This may sound somewhat counterintuitive, since neurons responding strongly to stimuli are usually strong favorites; but Figure 7*a* makes the point quite clearly. What the wide peak represents is the merging of two or more narrow peaks. This can happen if a neuron happens to belong to the foreground of two of the uncorrelated stimuli that are neighbors in the training sequence, in contrast to the situation represented by the second, narrow, peak in this same figure. This peak implies that the same neuron belongs to the foreground of yet another pattern that is *not* close to the others in the training sequence, and hence causes no classification problems.

Coming back to the wide peak, one observes that a neuron of this type will have especially high frequency and this frequency will be essentially equal for the two attractors, or the two driving stimuli. The second conclusion is rather obvious and is due to the symmetric life this neuron lives in both stimuli, during learning as well as during activation. It is the first part that needs some explaining. If a neuron belongs to the foreground of two stimuli, then it will tend to form excitatory synapses, during learning, with the neurons that belong to both foregrounds, once when the first stimulus is presented and once when the second is presented. If, moreover, those two stimuli are neighbors in the training sequence, the attractor provoked by any one of the two stimuli excites the neurons of both foregrounds. Neurons that belong to both foregrounds receive excitatory input from the union of the two foregrounds. Consequently, they have high and equal frequencies. These conclusions are very weakly dependent on the particular model. Note also

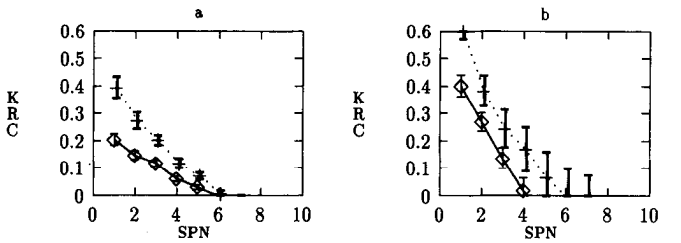

Figure 10. Dependence of KRCs on number of reverberations used in the statistical analysis. *a*, High and low KRCs for 100 patterns. *b*, The same for 20 delay activity distributions. Note that both types of KRCs start higher and fall off faster for smaller numbers of analyzed delay activity distributions. The complete set of reverberations is identical in *a* and *b*.

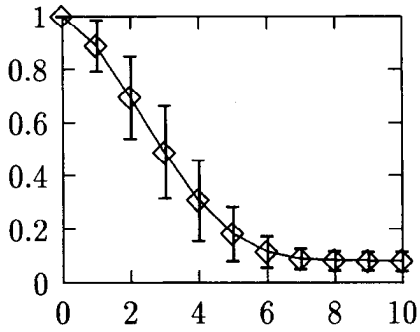

Figure 11. Standard correlations versus SPN in the simulations. The delay activity distributions leading to these correlations are identical to those that led to the KRCs of Figure 5. Error bars are SDs, not SEs as in Figure 5.

that this is yet another conclusion for confrontation between theory and experiment.

Fortunately, these neurons are not very frequent for uncorrelated stimuli of low coding rate, as are the ones both in the experiment and in the model. The probability that a neuron from the foreground of one stimulus participates in the foreground of another of p uncorrelated patterns of coding rate f may be as high as $[1 - (1 - f)^p]$, but that a neuron in the foreground of a pattern participates in another one out of a small group of patterns, neighbors of the given pattern, has a probability of qf, where q is the number of relevant neighbors (4–5). Hence, a relatively small sample of selective neurons will produce neurons that do not have frequencies too high and that will distinguish among the neighboring reverberations.

Spike rate distributions in delay activities

In addition to the attractor correlations, we have also investigated the activity distribution among the neurons of the assembly in a given attractor, Figure 12. Those are not available in the published material and can serve as predictions about the biological reality in AVT. The spike rate distributions in delay activities, *in a given attractor* (Fig. 12), have a large fraction of the neurons at spontaneous activity levels. The rest have a rather wide distribution about the mean activity in the attractor. One finds a high peak about the spontaneous rate and a wide peak, including about 5–6% of the neurons, around the central frequency of neurons participating in the active group recruited from the foreground of stimuli mixed into the attractor. This would be yet another quantity worth measuring and comparing.

Given the nature of the network, operating at the level of currents and spike rates, we could not, of course, reproduce the time structures of the spike trains, though there should be no special difficulty reducing the modeling level one storey down to have an actual representation of spike events (see, e.g., Amit et al., 1991).

Finally, we have tested the network with "new" stimuli, that is, stimuli uncorrelated with any of the memories imprinted into the synaptic structure. For those stimuli the network goes into a state in which all neurons have activities below 0.01, that is, three times the width of the imposed spontaneous activity. We have then tested the same network with "new" stimuli in the absence of the imposed spontaneous activity. In this case all neurons relax to zero frequency when the stimulus is re-

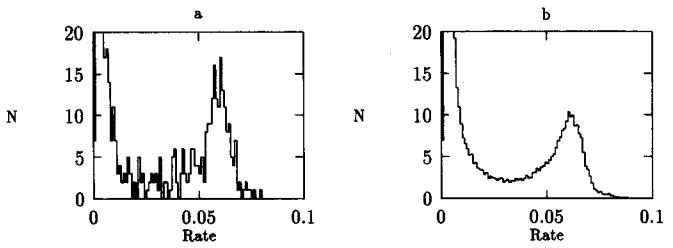

Figure 12. Spike rate distribution (number of neurons, N) in a reverberation. *a,* In a sample attractor. *b,* Averaged over all 100 attractors. Rates are as fraction of the saturation spike rate.

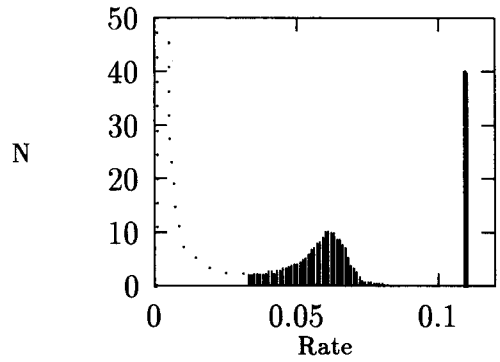

Figure 13. Comparison of sizes of active neuron populations in stimulus and in reverberation: histograms of frequency populations. *Full line* is for stimulus (frequency not to scale). *Dotted line* is delay activity reverberation. *Shaded area* is population of neurons with spike rates greater than 0.5 maximum rate. Includes about 5% of neurons.

moved. This is a clear indication that our analysis of the instability of spontaneous activity and the need for its imposition is fundamentally correct.

Coding rates in stimuli and in attractors

The salient fact that training is done with a fixed sequence of stimuli presented in a fixed order (Miyashita, 1988) is expressed by allowing information about contiguous patterns in the sequence to be coded into the synaptic structure. In the network we have implemented, the synapses connecting excitatory neurons can take one of three values: 0, 1, and a (see Materials and Methods).

This type of model is suited for effectively storing memories of low coding level, that is, those in which the fraction of network neurons activated by a given stimulus is low. We take it to be about 1%. The fraction of neurons active in the delay period is higher. It reaches several times the level of coding in a single stimulus. For example, Figure 13 superposes a histogram of neural activity in a state driven by a stimulus and in a delay activity attractor. The sharp peak on the right represents the population of neurons driven by a stimulus. The shaded area under the curve is the population of neurons in a given reverberation with a spike rate (chosen arbitrarily) higher than one-half the maximum rate in the distribution. It includes about 5% of the neurons.

Discussion

Outlook

The results of the Miyashita experiments as well as the availability of a theoretical tool for their reproduction carry much promise. The close approach, on a very detailed level, of the behavior of the model to the available empirical observations implies very strongly that the attractor picture is adequate for describing at least part of higher cortical functions (see also Zipser et al., 1993). This holds for the computational component as well as for the learning one. In fact, the type of cortical phenomena described here make the boundary between the two much less well defined than in current paradigms.

This is not the place to expand on the potential psychological implications of the attractor picture of cortical function. Much of it has been foreshadowed by Hebb, and some recent considerations, in a more specific context, have been touched upon by Griniasti et al. (1992) and Amit (1993a,b). What concerns us here is the dialectic between modeling and neurophysiology. The existence of the model creates the ground for formulating hypotheses for much more detailed experimentation and experimental data analysis.

All along we have tried to suggest direct predictions of the model. We have mentioned the level of actual correlations versus SPN of the rate distributions in the delay periods, the rate distribution among neurons in a given reverberation and its relation to the attractor structure as a well defined mixture of pure uncorrelated patterns, the relation of neuron spike rate to its function as reverberation discriminator and the corresponding probabilities, and the dependence of the KRCs either on the selection of subsets of delay activity distributions from the total for data processing, or in training with less than 100 patterns, and so on.

Those are at the simplest level. At the next level one can use the attractors and their correlations as hypotheses for the connection between psychophysical phenomena and neurophysiological phenomena. Those can be tested on the model and then looked for in experiment. Some are discussed in Amit (1993a).

Next, moving one level away one can test the learning hypothesis underlying the model. It consists of assuming that uncorrelated stimuli that arrive at the module during training first create uncorrelated attractors. Those attractors carry structural information from one stimulus to the consecutive one, to allow for the imprinting of the connecting terms in the synaptic matrix. Such a picture implies many consequences for possible effects related to the delay activities generated by different learning scenarios: at some stage uncorrelated delay activities should show up, and they should be reinforced by extending the training time. The structure of the correlation coefficients should strongly depend on the mixture of subsequences of ordered and disordered presentations. Some steps in this direction have already been undertaken in studying the effect of sequences of ordered pairs of stimuli unordered among themselves (Sakai and Miyashita, 1991). Since in the proposed model the prescribed matrix can be replaced by a learning dynamic, the hypotheses can first be developed and tested on the model.

Open questions

The extensions of the application of the model depend on the solidity of its foundations and those involve both its experimental and theoretical dimensions. We therefore turn to a discussion of the exposed flanks.

Experimental front

Undoubtedly the reverberations observed in the Miyashita experiments are correlated. What has not been demonstrated is either that the stimuli as such are uncorrelated or that they

remain uncorrelated when they finally arrive at AVT. The first part of the question relates to the relation between the distributions of activities produced by the visual stimuli at the retina. The second concerns the possibility that even activities uncorrelated at the retina arrive correlated at AVT, following the elaborate preprocessing.

If the stimuli are correlated at the entrance, much of the cognitive import of the phenomenon is lost. If they are uncorrelated on the retina but arrive correlated at AVT, only the learning part of the interpretation is damaged. In other words, it is not the generation of correlated attractors from uncorrelated input that is taking place in AVT. We would argue that in that case some other, earlier, module will have to do what was ascribed to AVT in terms of learning.

A careful study of the correlations during stimulation is of central importance.

The model

As complex as the model has become it is still simple enough to be studied analytically and may be useful in clarifying the potential influence of the multiplicity of factors involved. It is therefore important to emphasize its potential weaknesses.

Synaptic matrix

At this stage, it seems to us that the most unrealistic feature of the model is the detailed synaptic matrix, both the learned excitatory–excitatory part and the coupling to the inhibition. Concerning the first, we are encouraged by two facts: The dynamics change little when the structured matrix connecting the excitatory neurons changes between rather remote cases, the one described and the one of Tsodyks and Feigel'man (1988). Moreover, our preliminary results indicate that a rather naive dynamic learning process leads to a synaptic matrix rather similar to the one used in this study. One is left with the impression that the detailed structure of this synaptic matrix is not very crucial, provided it has good attractors for the uncorrelated stimuli, in the absence of sequential effects; that sequential effects couple nearest neighbors in the synaptic matrix; and that inhibition is in tune with the excitatory system.

We report results only for this network, though we have experimented with a very different network as well: the synaptic matrix reported in Griniasti et al. (1992) for 0–1 neurons. This is a matrix with synapses of large analog depth, based on the model of Tsodyks and Feigel'man (1988) and Buhmann et al. (1989). The results are not very different, which indicates that the performance is rather robust to details of the synaptic matrix, provided the synaptic matrix embeds a reasonable expression of the learning hypothesis.

After structuring the synaptic values on the basis of the learned set of *uncorrelated* stimuli, the network has a naturally sparse synaptic connectivity. For the parameters used in the simulation (100 learned patterns and coding level 1%; see Materials and Methods), every neuron is connected to some 3% of the other neurons in the module, though one should keep in mind that in the model this connectivity is correlated with the learned patterns. We have not simulated the network with unstructured low connectivity, but experience with discrete models indicates that attractor dynamics are robust to very high levels of random dilution (see, e.g., Derrida et al., 1987; Tsodyks, 1988). The network comprised 4000 excitatory neurons to allow for the storage of a large number of memories with low coding levels. This has turned out to be essential since the special correlation coefficients (KRC) are very sensitive to the number of random variables from which they are calculated (see Additional information in model network, above, and Brunel 1994).

The model presented has a synaptic structure totally determined by the learning process. This implies that synapses are not only strengthened and weakened by learning, but are also formed in the process. It is more likely though that the connectivity in the relevant neural module is predetermined in development. If it is so, learning has to take place on a reduced population of synapses. This would imply additional noise in the network's dynamics and a consequent reduction of memory capacity. Certain levels of synaptic dilution can be tolerated by our model, but others may be too high for storing the connected 100 memories. However, things become easier as the number of neurons increases and we are, in our simulations, at least a factor of 25 below the size of the biological cortical column.

As far as the coupling to the inhibition is concerned, the main defects we perceive are (1) that the localized structure of the inhibition is sacrificed and (2) that there are no internal dynamics to the inhibition. The first has the effect that we cannot describe potential effects of the stimulus on inhibitory neurons. Such effects seem to appear in the recordings presented in Figure 3*d* of Sakai and Miyashita (1991). They will probably reappear when the scope of the simulation is enlarged. We expect, though, that localizing the inhibition will not have significant effects on the global behavior of the network. The second issue seems rather innocuous. It is just the local nature and the dilute presence of inhibitory neurons that would justify the lack of inhibitory–inhibitory coupling.

Neurons

Our neurons are oversimplified. None of the colorful phenomena of cable theory are included (Segev, 1992). Some arguments to support the possibility that this may be a justified approximation in cortex have been advanced in Amit and Tsodyks (1992), based on the smoothing effects of the immense afferent flux due to spontaneous activity in the entire cortex. Even if these arguments are not exact, they do leave room for hope that the complex effects of nonlinearities on the dendrites are small perturbations.

This leads to the question of the spontaneous activity. The present model does not deal with this phenomenon satisfactorily, as we have clearly emphasized in the text. It is our view that spontaneous activity in the cortex is self-maintaining—it generates itself and stabilizes itself. It is a global unstructured attractor at low spike rates. Our neural elements do generate the spontaneous activity, but do not maintain it. That is why we had to introduce it by hand. This is one direction in which the model must and will be improved.

This may have to do with a better treatment of the inhibition about which we have been particularly cavalier. We did test that more realistic inhibition functions do not disturb the results on the attractors, but to fix up the problem of the spontaneous activity we may have to enter into more detail.

There is a question about the absolute values of the spike rates at which the reverberations stabilize. They are all related to the saturation rate, which in turn is determined in this type of model by the inverse of the absolute refractory period and by the membrane's integration time constant. If in fact the saturation rates are 500 spike/sec, our rates are too high, perhaps by a factor of 2. This is not a large gap, considering that the values of the two time constants are not known very precisely

in vivo. On the other hand, it is possible that with a larger number of neurons (100,000 compared to our 4000) one may be able to stabilize reverberations with lower rates, even if saturation rates are at 500. And then there is always the escape route of adaptation, which may reduce the effective saturation frequency.

References

Amit DJ (1989) Modeling brain function. New York: Cambridge UP.

Amit DJ (1992) In defense of single electrode recording. Network 3:385.

Amit DJ (1993a) The Hebbian paradigm reintegrated: local reverberations as internal representations. Brain Behav Sci, in press.

Amit DJ (1993b) Cognitive neuro-physiology: an empirical basis for neural modelling and cognitive psychology. In Neural computing research and applications (Orchard G, ed). Bristol: IOP Publishing.

Amit DJ, Fusi S (1993) Dynamic learning in neural networks with material sysnapses. Neural Comput, in press.

Amit DJ, Tsodyks MV (1991a) Quantitative study of attractor neural network retrieving at low spike rates. I: Substrate—spikes, rates and neuronal gain. Network 2:259.

Amit DJ, Tsodyks MV (1991b) Quantitative study of attractor neural network retrieving at low spike rates. II: Low-rate retrieval in symmetric networks. Network 2:275.

Amit DJ, Tsodyks MV (1992) Effective neurons and attractor neural network in cortical environment. Network 3:121.

Amit DJ, Evans MR, Abeles M (1991) Attractor neural networks with biological probe neurons. Network 1:381.

Braitenberg V, Schuz A (1991) Anatomy of the cortex. Berlin: Springer.

Brunel N (1994) Analysis of dynamics in networks converting temporal into spatial correlations. Network, submitted.

Buhmann J, Divko R, Schulten K (1989) Associative memory with high information content. Physiol Rev A39:2689.

Derrida B, Gardner E, Zippelius A (1987) An exactly soluble asymmetric neural network model. Europhys Lett 4:167.

Foldiák P (1991) Learning invariance from transformation sequences. Neural Comput 3:194.

Griniasty M, Tsodyks MV, Amit DJ (1992) Conversion of temporal correlations between stimuli to spatial correlations between attractors. Neural Comput 5:1.

McNaughton BL, Barnes CA (1990) From cooperative synaptic enhancement to associative memory: bridging the abyss. Neurosciences 2:403.

McNaughton BL, Nadel L (1990) Hebb–Marr networks and the neurobiological representation of action in space. In Neuroscience and connectionist theory (Gluck MA, Rumelhart DE, eds). Hillsdale, NJ: Erlbaum.

Miyashita Y (1988) Neuronal correlate of visual associative long-term memory in the primate temporal cortex. Nature 335:817.

Miyashita Y, Chang HS (1988) Neuronal correlate of pictorial short-term memory in the primate temporal cortex. Nature 331:68.

Sakai K, Miyashita Y (1991) Neural organization for the long-term memory of paired associates. Nature 354:152.

Segev I (1992) Single neurone models: oversimple, complex and reduced. Trends Neurosci 15:414.

Snedecor GW, Cochran WG (1969) Statistical methods. Ames, IA: Iowa State University.

Treves A (1990) Graded-response neurons and information encodings in auto-associative memories. Physiol Rev A42:2418.

Tsodyks MV (1988) Associative memory in asymmetric diluted neural network with low level of activity. Europhys Lett 3:203.

Tsodyks MV, Feigel'man MV (1988) The enhanced storage capacity in neural networks with low activity level. Europhys Lett 46:101.

Tweney RD, Heiman GH, Hoemann HW (1977) Effects of visual disruption on sign intelligibility. J Exp Psychol [Gen] 106:255.

Willshaw D, Buneman OP, Longuet-Higgins H (1969) Non-holographic associative memory. Nature 222:960.

Zipser D, Kehoe B, Littlewort G, Fuster J (1993) A spiking network model of short-term active memory. J Neurosci 13:3406.