Model of cortical organization embodying a basis for a theory of information processing and memory recall

(Hebb synapse/selective adaptive network/axial next-nearest-neighbor Ising model/fluctuations/synchronous time steps)

GORDON L. SHAW, DENNIS J. SILVERMAN, AND JOHN C. PEARSON

Physics Department, University of California, Irvine, CA 92717

Communicated by Frederick Reines, December 10, 1984

ABSTRACT Motivated by V. B. Mountcastle's organizational principle for neocortical function, and by M. E. Fisher's model of physical spin systems, we introduce a cooperative model of the cortical column incorporating an idealized substrate, the trion, which represents a localized group of neurons. Computer studies reveal that typical networks composed of cellular automata (8, 9) and an Ising model of cortical columns exhibit striking behavior, e.g., hundreds to thousands of quasi-stable, periodic firing patterns, any of which can be selected out and enhanced with only small changes in interaction strengths by using a Hebb-type algorithm.

In the spirit of Mountcastle's (1) organizational principle for neocortical function, and strongly motivated by Fisher's (2-5) model of physical spin systems, we have developed a cooperative mathematical model of the cortical column. Our model incorporates an idealized substrate, the trion, which represents a localized group of neurons. The trion model allows for a completely new framework for information processing and associative memory storage and recall: Small networks of trions with high symmetric interactions are found to yield hundreds to thousands of quasi-stable, periodic firing patterns (see Fig. 1). Experience or learning would then modify the interactions (away from the symmetric values) and select out the desired patterns [as in the selection principle of Edelman (6)]. Remarkably, we have found that relatively small modifications in trion interaction strengths (away from the symmetric values) via a Hebb-type algorithm (7) will enhance and select out any desired pattern. Conceptually, this suggests a radically different approach from those information processing models that start at the opposite extreme of a randomly connected neural network with no periodic firing patterns, and then [via Hebb-type modifications (7) in the synaptic interactions] reinforce specific firing patterns. Another exciting feature is that our model includes the known statistical fluctuations in the post-synaptic potentials. These fluctuations are essential for having the huge number of patterns. We believe that these phenomena are of interest to workers in the fields of neurophysiology and cellular automata (8, 9) and to molecular scale processors, as well as possible applications to a future generation of computers.

Despite the substantial theoretical efforts and results in modeling neural networks (see, e.g., references in refs. 10-12) the bases for the tremendous magnitudes of the processing capabilities and the memory storage capacities of mammals remain mysteries. We believe Mountcastle's (1) columnar organizing principle for the functioning of the neocortex will provide a basis for these phenomena, and we construct a mathematical model based on it. He proposed that the well-established cortical column (~500 μm in diameter) is the basic network in the cortex and comprises small irreducible processing units. The subunits are connected into columns or networks having the capability of complex spatial-temporal firing patterns. The creation and transformation of such patterns constitute the basic events of short-term memory and information processing. We strongly emphasize this assumption: that higher mammalian cortical processes involve complex spatial-temporal network neuronal firing patterns; this is in contrast to the assumption that the "coding" involves sets of neurons firing with high frequency.

A model presented by Little and developed by Little and Shaw (13-16) mapped the neural network problem onto a generalization of the (classical) Ising spin model of magnetism. Consider a network of interconnected neurons in which each neuron has two possible states, corresponding to whether or not it has recently fired. These states are updated synchronously in discrete time steps (τ), with the order of the refractory period (a few msec). The state of the system at time τ = 0 is probabilistically related to the state of the system at time τ = nτ, by the interactions between the neurons. The probabilistic feature simulates the known fluctuations in the post-synaptic potentials due to the statistical nature of the release of neurotransmitter. Examination of the solutions of the large fluctuation limit of this model revealed a subunit organization (17, 18) in which only a few levels (+, 0, −, with respect to background) of the output of each subunit of perhaps 30-100 neurons (19, 20) was important. (For example, in a group composed of 60 neurons, firing levels of +, 0, and − could correspond to 60-41, 40-21, and 20-0 neurons firing, respectively. More levels or nonequal and nonsymmetric spacing are easily accommodated.) The question we address in this paper is what qualitatively new phenomena can occur as a result of having such substructure.

The trion model is an attempt to abstract (in contrast to being derived) from the level of individual neurons to the next level or scale of phenomenological relevance. In making this change of scale, we have drawn from the previously discussed work of Mountcastle (1), from our theoretical studies (13-20), and from the ANNNI model work of Fisher (2-5). Fisher showed that a simple extension of the Ising model, the ANNNI (axial next-nearest-neighbor Ising) model, exhibits a large number of possible stable phases or ordered configurations of spins, whereas the nearest-neighbor Ising model has just two stable phases. The number of stable phases in these physical models is related to the memory capacity or processing capability in our neural model. Figure 1 in ref. 2 illustrates the infinite number of distinct spatially modulated phases in the ANNNI model. None of these phases is coexistent for a given set of parameters (except at a special multiphase point or line), whereas many of the analog quasi-stable periodic firing patterns can be excited for a giv-

Abbreviations: ANNNI model, axial next-nearest-neighbor Ising model; MP, magic pattern.

Fig. 1. An example of time evolution of firing activity in a network of six trions. Each square in a given row represents the firing level of a trion at a given time step.  ●, □, and ■ represent firing levels above background, at background, and below background, respectively. Time steps (rows) are consecutively ordered, beginning at the top of the column on the left and continuing from the bottom of that column to the top of the next column, etc. There are seven different periodic patterns present here that cycle two or more times. This example was derived from a Monte Carlo simulation based on Eq. 1, using the trion parameters of Table 1a, with $B = 10$.

There are three essential features (2–5) responsible for this additional complexity. (i) A strong positive coupling between nearest-neighbor spins within each layer, which causes them to align, thus providing essential stability against random thermal fluctuations. (ii) There are two levels of interaction between spins of different layers: a positive interaction (aligning) between nearest-neighbor layers and a negative interaction (reversing) between next-nearest-neighbor layers. (iii) Finite fluctuations are essential for the appearance of the large number of phases: in the zero fluctuation limit there are only two stable phases. We have incorporated into our trion model analogs of these three key features, and we have found that they are all essential for its rich behavior.

In our network composed of interconnected trions, each has three possible states, $S$, denoted by $+ (1)$, 0, $- (1)$, which represent a firing output above background, at background, and below background, respectively. Associated with each of the three trion states $S$ is a statistical weighting term $g(S)$, with $g(0) \gg g(+/–)$, which takes into account the number of equivalent firing configurations of the trion’s internal neuronal constituents (17–20). Thus, by effectively averaging over the states of its constituents and by using the approximation of only three distinct firing levels with $g(0) \gg g(+/–)$, the output of the trion gains crucial stability, which is analogous to feature $i$ described above in the ANNNI model. The trion states are updated synchronously in discrete time steps $\tau > \tau_0$. The state of the system at time $n\tau$ is probabilistically related to the states of the system at times $(n−1)\tau$ and $(n−2)\tau$ by the values of $g(S)$ and the temporal interactions between trions. The dependence on the two previous time steps is analogous to feature $ii$ described above in the ANNNI model, and the probabilistic factor is due to the random nature of synaptic transmission as well as other sources of noise. The probability $P_i(S)$ of the $i$th trion attaining state $S$ at time $n\tau$ is given by

$$P_i(S) = \frac{g_i(S)\exp[B\cdot M_i]}{\sum S_j \exp[B\cdot M_j]},$$

[1]

where $S_j$ and $S'_j$ are the states of the $j$th trion at times $(n−1)\tau$ and $(n−2)\tau$, respectively. $V_i$ and $W_i$ are the interactions between trions $i$ and $j$ between time $n\tau$ and times $(n−1)\tau$ and $(n−2)\tau$, respectively. $V_i^r$ is an effective firing threshold. $B$ is inversely proportional to the level of noise, “temperature,” or random fluctuations in the system. The deterministic limit is taken by letting $B$ approach infinity (analogous to the noise approaching zero), in which case the $S = 0$ states vanish.

Although our basic formalism (Eq. 1) is not derived from models of single neurons, the simplifying form readily follows from our previous work (17, 18), and the approximation that each trion corresponds to a group of 30–100 neurons (19, 20) for which $g(0)/g(+/–) \approx 500$. The key assumption concerning two discrete time steps $\tau$ clearly warrants further discussion. In analogy with the next-nearest-neighbor spatial interactions in Fisher’s spin model being a crucial feature in obtaining many stable phases, we find that interactions lasting two time steps led to a striking increase in the number and complexity of our quasi-stable periodic firing patterns. (Mathematically, the presence of interactions spanning two time steps leads to the description of the system by second-order difference equations, which are known to yield oscillatory solutions, while first-order equations only yield exponential solutions.) First, we note that the trion time step $\tau$ we have in mind is $\approx 50$ msec, a much larger time than the individual neuronal firing time of a few msec. To establish the plausibility of such a $\tau$ value, we note that the observation of periodic bursting in cortex has a long history (see, in particular, Morrell et al. (21, 22), who found multiphasic responses in cat visual cortex with peak separation of $\approx 50$ msec.) In addition to the simulation studies reported in this paper, we have conducted neurophysiological and experimental temporal and spatial studies (23) to test certain assumptions of the model. We present in Fig. 2 some of our data from cat primary visual cortex recordings from a group of 2–3 neurons, which show four equally spaced peaks in the post-stimulus histogram in response to a flashed bar (see also Figure 1 of ref. 23). These peaks are separated by $\approx 50$ msec, in close agreement with Morrell’s data (21, 22). Clearly, it would be very interesting to record simultaneously from two or more closely spaced microelectrodes to test our assumption of a discrete time step $\tau$. Although we have no evidence for interactions lasting two time steps, we are able to construct several different mechanisms for it. We believe these are both reasonable and testable. For example, we might imagine that the upper and lower cortical layers in a column separately are trion networks. If the firing of the upper layers is delayed from the bottom layers by $\tau/2$ (i.e., the peak burst firings were interleaved), then the interaction of the two networks with interactions lasting only one step is readily mapped onto our two time-step model. The experimental search for such an interleaving of bursts between upper and lower layers would be of great interest.

We have studied properties of the probabilistic time evolution of the states of the system as a function of $B$, $g(S)$, $V$, and $W$, as given by Eq. 1. Our calculations have considered...
Biophysics: Shaw et al.

terms $V_i$, $V_{i=1}$, $W_i$, $W_{i=2}$, $W_{i=2}$.

The principal finding has been that networks of a small number of trions with highly symmetric $V$s and $W$s can exhibit hundreds to thousands of quasi-stable periodic firing patterns. Since any of these patterns can be enhanced and selected out by small changes in the $V$s and $W$s (away from the symmetric values) by using a standard Hebb-type algorithm (7), we denote these patterns as MPs (magic patterns). Quasi-stable means that the probability of the MP repeating itself or cycling is relatively high. The MPs are found by computing the most probable temporal evolution of the trion states from each of the possible initial conditions and by determining whether that evolution leads back to the initial conditions with a high probability.

Networks of six trions with interactions spanning two time steps have $3^{6-6} = 531,441$ possible initial conditions (the number of configurations of the $S$s of the first two time steps.). We assume that the cycling of a MP for a few cycles is related to short-term memory or to some element of information processing (questions concerning these relationships, as suggested in the concluding paragraphs, will clearly involve much future work). Thus, we consider a large number of MPs to indicate a large storage capacity or processing capability. Not only do we have huge numbers of MPs, but they have the further very interesting properties that (i) in the absence of outside stimuli, a particular MP can persist for a few or many cycles depending on the noise parameter $B$ (see Table 1). (ii) The MPs can flip from one to another, even without outside stimuli (see Fig. 1). (iii) The results are not very sensitive to arbitrary substantial (10%-20%) changes in the interaction potential $V$ values (compare $a$ and $b$ in Table 1). (iv) Using a Hebb-type algorithm (7), relatively small (less than 5%) changes in the $V$ and $W$ values away from the symmetric values will enhance and select out any of the MPs. (Here we illustrate with a specific example how the Hebbian plasticity algorithm (see Eq. 2) enhances the MP’s probabilities of cycling even for the “complex” six cycles in a finite range of $B$.

The sample MP is that labeled $E_3$ in Table 2. For instance, the changes in the nearest-neighbor interactions for the third trion are $\Delta V_{34} = 0.1$ and $\Delta V_{32} = -0.02$, derived from Eq. 2 with $\varepsilon = 0.02$. The resulting new probabilities of cycling (using the complete set of changes in the $V$ and $W$ values including next nearest neighbors for all six trions) are 0%, 24%, 51%, 27%, and 0% for $B$ values of 20, 10, 8, 6, and 4, respectively, as compared with the corresponding probabilities of Table 1, which are 96%, 80%, 28%, 0%, and 0%, demonstrating the significant increase or enhancement of cycling probability due to the Hebbian change for $B$ values 8 and 6.)

![Fig. 2. Spike firing response of a cluster of three neurons in area 17 of a cat to a stationary flashed bar (2 1/2 by 4argarbar). Forty stimuli were presented with a 400-msec interstimulus interval and a 34-msec stimulus duration (denoted by horizontal bar), and data are displayed using a 2-msec bin. Note the presence of four peaks in the histogram. The interpeak interval is remarkably consistent at 50 ± 3 msec, which suggests a trion interburst interval of ~50 msec. These data are similar to those presented in ref. 23.](image)

### Table 1. Properties of two representative networks with 6 trions

<table>
<thead>
<tr>
<th>MP class</th>
<th>40</th>
<th>20</th>
<th>15</th>
<th>10</th>
<th>8</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (17)</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>94</td>
<td>89</td>
<td>89</td>
<td>56</td>
<td>2</td>
</tr>
<tr>
<td>B (72)</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>93</td>
<td>88</td>
<td>60</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>C (2)</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>90</td>
<td>88</td>
<td>69</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>D (2)</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>97</td>
<td>94</td>
<td>75</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>E (156)</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>80</td>
<td>80</td>
<td>28</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

The periodic firing patterns (MPs) are grouped into classes defined by the member MPs having the same probability of cycling for all values of $B$. Five of the 21 classes are shown in (A). The defining parameters are as follows: (a) $V_{i=2} = 1.0$, $V_{i=2} = 1.0$, $W_{i=2} = -1.0$, $W_{i=1} = -1.0$, $g(-1) = 1$, $g(0) = 500$, $g(+1) = 1$; (b) $V_{i=2} = 0.8$, $V_{i=2} = 1.0$, $W_{i=2} = -1.15$, $W_{i=1} = -1.1$, $g(-1) = 1$, $g(0) = 500$, $g(+1) = 1$. The number of MPs in each class is indicated in parentheses. Values represent the percent probability of each MP cycling or repeating for $B$ ranging from $B = 4$ (high noise level) to $B = 40$ (low noise level). By breaking the symmetry of the interactions some of the classes in (a) split into several classes in (b). This is denoted by subscripts the labels of the split classes in (b). The interactions between the trions in these networks have periodic boundary conditions, which is equivalent to a ring-like spatial connection of the trions. Only those MPs with a cycling probability $>10\%$ at $B = 10$ are included, and probabilities of cycling $<1\%$ are not listed. The total number of MPs in (a) is 1804 and in (b) is 883. One of the classes not shown is the background MP with all 0s, or a class by itself and having a probability of cycling of 98%.

In Table 1 we show the salient properties of two different networks of 6 trions, as defined by different $V$ and $W$ values. The particular $V$ and $W$ values of these two networks were chosen to illustrate typical model behavior (a more complete investigation of the parameter space will be presented elsewhere). Network $a$ has symmetrical interactions, and has the large total number of MPs of 1804. These 1804 MPs fall into 21 different classes (defined by the member MPs having the same probability of repeating for all values of $B$). Five of these classes are shown in Table 1. Network $b$ is the same as network $a$ but its interactions are asymmetrical, being within 20% of their corresponding values in network $a$. This substantial change in the interactions did not eliminate the capacity for large numbers of MPs nor did it introduce new MPs, thus illustrating the stability of the system. However, breaking the symmetry of the interactions did transform the structure and properties of the MP classes, creating greater diversity of response. Many of the classes in network $a$ split to form several classes in network $b$, as denoted by subscripts in the class labels in network $b$. The classes also show more interesting variability with the level of noise, especially the rapid decrease in the probability of repeating at low as well as at high levels of noise. The low noise extinction occurs because finite levels of noise are needed to sustain the trion $S = 0$ firing levels in asymmetrical networks (as can be seen from Eq. 1, $S = \pm 1$ will dominate as the noise goes to zero). This suggests that nonspecific inputs to the trion network might simulate noise, which could tune the network’s response properties—i.e., changes in $B$ will change which classes have the highest probability of repeating.  

Table 2. Representative MPs from classes in Table 1

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th></th>
<th>A2</th>
<th></th>
<th>B1</th>
<th></th>
<th>B2</th>
<th></th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Time steps are arranged vertically and trion numbers are arranged horizontally.

We have found, in general, as illustrated in the example in Table 1a, that highly symmetrical trion interactions lead to huge numbers of MPs. We speculate that these symmetrical interactions might be specified genetically, giving a "naive" network that could initially respond to many different input signals. Experience or learning (during a critical period [24]) could then modify the connections via a Hebb-type mechanism (7) to select out the appropriate responses or MPs. We assume that the cycling of the trion network through a firing pattern will produce small changes in the V and W values given by the Hebb-type algorithm (7, 17, 18).

\[
\Delta V_{ij} = \epsilon \sum_{\text{cycle}} S_i(T)S_j(T-1)
\]

\[
\Delta W_{ij} = \epsilon \sum_{\text{cycle}} S_i(T)S_j(T-2), \quad \epsilon > 0.
\]

By changing the V and W values away from their symmetric values by using Eq. 2, we find that any of the MPs can be enhanced and selected out (in the sense that the particular MP and some others in its class have their firing probabilities enhanced, whereas most others are suppressed). Although it is easy to show that Eq. 2 must enhance any one cycle, it seems remarkable to us that this works for every MP in Table 1 (and other examples of MPs we have investigated), and even for the very complex six cycles shown in Table 2 (for B values <10). We believe that our model might provide a specific framework for Edelman's group-selective theory of brain function (6).

We have carried out the same analysis on networks identical to those in Table 1a and b except that the statistical weight of the trion zero firing level, g(0), is set equal to zero. This is equivalent to a single neuron model (i.e., 1 neuron per trion), in which there are no S = 0 states. These networks are capable of supporting only 12 and 5 MPs, respectively. In general, we have found that only networks for which g(0)/g(-/+)) >> 1 have large numbers of MPs. In addition to this condition on the statistical weights and the necessity of finite fluctuations, large numbers of MPs require that the interactions span two time steps (i.e., some nonzero W) with both positive (excitatory) and negative (inhibitory) interactions.

We have investigated several other aspects of the model (we are exploiting the content-addressable nature of the MP network in our computer studies of these properties). Networks of seven and eight trions yield qualitatively similar numbers of MPs. The particular cyclic boundary conditions used do not appear to be a critical factor. (Recent models (25) of the spatial organization of orientation minicolumns in the primary visual cortex of cats and monkeys suggest that they are arranged in circular groupings, lending support to our periodic boundary conditions.) Networks with fixed boundary conditions also support many MPs, as do networks with g(+)/g(−) ratios not equal to one. Monte Carlo simulations of the temporal evolution of trion states, using Eq. 1 and a random number generator, are being used to study the dynamical properties of transitions between MPs (Fig. 1) and to study the effect of relaxing our assumption of synchronous firing. Introducing partial asynchrony (by updating at random one trion prior to the others) does not seriously degrade the quasi-stability of the MPs, though some reasonable semblance of a basic time step is necessary to fully exploit the potential of the model.

Major fundamental questions remain to be investigated: In working toward a theory of associative memory storage and recall, we need to know why reinforcing any particular MP via the Hebb algorithm (Eq. 2) will enhance it and select it out. This seems truly remarkable for some of the complex long-cycle MPs. Clearly, we need to know how outside stimuli excite the various MPs. As noted, certain MPs spontaneously flip to other MPs. This phenomenon should be relevant in obtaining and understanding hierarchies of associations, as well as the profound problem of obtaining long complex time sequences. Periodic driving inputs and the coupling of two or more trion networks should be considered. The nature of the internal processing within each trion must also be understood before a theory of information processing can be complete.

We conclude that the trion model allows for a completely new framework toward developing theories for information processing and for associative memory storage and recall. Even at these initial stages of the model, we believe that it will be very powerful in simulating multielectrode recordings in cortex, thus stimulating new ideas in designing and analyzing these important experiments. In particular, we believe that it will be crucial to both have several closely spaced (50-200 um separation) microelectrodes that can monitor neurons in several nearby trions and have stimuli of the sequential type used previously (23, 24) to excite the MPs. If, indeed, there is an approximately synchronous time step τ(30-100 msec) when groups of cortical neurons tend to burst, this should be exploited in the presentation of stimuli to the animal. Repetitive dynamic sequences spaced at time intervals τ might resonantly excite the spatial-temporal MPs. It also

appears potentially useful to do conditioning experiments of the type used by Morrell (21, 22) to excite these patterns.

We thank J. Patera and A. B. Scheibel for many helpful discussions, suggestions, and encouragement, and M. E. Fisher for helpful correspondence. We also thank J. Patera and R. Vasudevan for their help and participation in the early stages of the development of this work. We thank the Aspen Center for Physics, where part of this research was done. This work was supported in part by the University of California Irvine Focused Research Program in Brain Function.
