A Model of Clipped Hebbian Learning in a Neocortical Pyramidal Cell

Bruce Graham and David Willshaw

Centre for Cognitive Science, University of Edinburgh
2 Buccleuch Place, Edinburgh EH8 9LW, Scotland, UK
E-mail: *B.Graham@ed.ac.uk & D.Willshaw@ed.ac.uk*
Phone: +44 131 650 4408; Fax: +44 131 650 6626

Abstract. A detailed compartmental model of a cortical pyramidal cell is used to determine the effect of the spatial distribution of synapses across a dendritic tree on the pattern recognition capability of the neuron. By setting synaptic strengths according to the *clipped Hebbian* learning rule used in the associative net neural network model, the cell is able to recognise input patterns, but with a one to two order of magnitude decrease in performance compared to the computing units in the network model. Performance of the cell is optimised by particular forms of input signal, but is not altered by different pattern recognition criteria.

1 Introduction

The experimental phenomena of long-term potentiation (LTP) and long-term depression (LTD) suggest certain synaptic pathways within the mammalian central nervous system are modified in a *Hebbian* fashion (for a review see [4]). Neural network models of associative memory use Hebbian learning for pattern storage. Such memories are often proposed as models of cortical networks and their capacities are used to indicate the potential storage capacity of cortical regions, such as the hippocampus [1, 8, 11, 13]. However, the computing units in such models are idealised versions of biological neurons and do not include many potential sources of noise inherent in real neurons.

Here we compare the memory performance of an output unit in an associative net model [14, 15] with that of a realistic compartmental model of a cortical pyramidal cell. The same set of input patterns is associated with the output unit and the pyramidal cell by setting synaptic strengths according to a *clipped Hebbian* learning rule. During pattern recall input patterns are presented to the output unit, or the cell, as cues, and the unit, or cell, should respond preferentially to the patterns with which it was coactive during pattern storage. The decision to respond is based on the weighted sum of the inputs, or *dendritic sum*. The output unit receives a dendritic sum that is simply the number of active inputs to which the unit is connected. On the other hand, the neuron receives input signals distributed across its dendritic tree. These signals vary over time and must travel to the cell body (soma) to be integrated into the final dendritic sum. The temporal nature of the signals and their spatial distribution introduces noise into the dendritic sums that is not present in the idealised associative net model. Measuring performance by the signal-to-noise ratio (S/N) between the dendritic sums due to the patterns that should be recognised and the sums due to those that should not be recognised reveals a one to two order of magnitude reduction in the S/N in the biological case. Thus current model estimates of cortical capacity are likely to be inflated.

2 The Models

In the associative net [14, 15], input and output units with binary activity are connected by synapses that are also binary. Pairs of patterns are stored in the net using a *clipped Hebbian* learning rule that changes the weight of a synaptic connection from 0 to 1 if both input and output units are active for the same pattern pair. Though extremely simple, this learning rule provides both high memory capacity and information efficiency if the patterns are sparse (few active inputs per pattern) [8, 15]. We use a net with 4096 input units, which is of the order of the number of excitatory inputs to a cortical pyramidal cell. Input patterns contain 400 randomly chosen active units. Output activity is set so that an output unit is associated with about 10% of the input patterns. Following pattern storage the weights from a representative output unit are used as the weights of the synapses of the pyramidal neuron. Thus this output unit and the neuron both should recognise the same input patterns during recall.

The neuron model is based on the morphology of a layer 5 pyramidal cell from the visual cortex of a cat [7], as shown in Figure 1. Model parameters are derived from those used by Bernander [2, 3] for the same cell (see figure legend). The neuron is divided into 441 isopotential compartments with the lengths and diameters of each cylindrical compartment being determined by the cell morphology (using the computer programme *ntscable*, written by J.C. Wathey at the Salk Institute). The length of each compartment is less than one-tenth of the space constant. Simulation of the cell is performed using the simulator NEURON [9], which uses backward Euler integration to solve equations for the membrane potential in each compartment at each time step. In most experiments the entire cell is passive, but when action potentials are required as output signals, fast sodium channels and delayed-rectifier potassium channels are added to the soma, with the Hodgkin-Huxley-style kinetics and densities of Bernander et al. [3].

The 4096 inputs to the cell are in the form of excitatory synapses distributed uniformly across the apical dendritic tree in layers 1 to 3 (Fig. 1). During pattern recall the binary activity of an input pattern is represented by trains of action potentials at those synapses corresponding to active inputs. All the spike trains are independent, Poisson-distributed and of the same mean frequency. Only those synapses given a weight of 1 by the *clipped Hebbian* learning during pattern storage will produce a postsynaptic response. The postsynaptic conductance due to a presynaptic action potential is modelled by an alpha function with a maximum conductance of 0.5nS and a time to peak of 1.5m$secs$ [2]. This conductance gives rise to an excitatory postsynaptic potential (EPSP) which

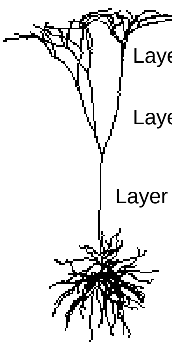

Fig. 1. Morphology of the layer 5 pyramidal cell. The basic passive membrane properties are: capacitance $C_m = 1\mu F/cm^2$; axial resistance $R_a = 200\Omega cm$; membrane resistance $R_m = 20,000\Omega cm^2$ in the soma and $R_m = 50,000\Omega cm^2$ in the dendrites (to approximate the effect of low frequency background excitatory and inhibitory activity on the cell); reversal potential is $-66\mathrm{m}V$.

travels passively along the dendrite until it reaches the soma. The depolarization of the membrane at the soma due to all the EPSPs is the weighted sum of the inputs, or dendritic sum.

The ability of either the output unit or neuron to correctly discriminate the input patterns during pattern retrieval is determined by how much overlap there is between the dendritic sums due to the patterns that should be recognised (*high* patterns) and the sums due to the patterns that should not be recognised (*low* patterns). This is conveniently measured by the signal-to-noise (S/N) ratio between these two classes of dendritic sums [6]:

$$\rho = \frac{(\mu_h - \mu_l)^2}{1/2(\sigma_l^2 + \sigma_h^2)} \tag{1}$$

where μ_h (μ_l) is the mean and σ_h^2 (σ_l^2) is the variance of the dendritic sums from the *high* (*low*) input patterns. If the mean values of the two classes of dendritic sum are well separated and their variances are small, the S/N will be high and the output unit or neuron is likely to discriminate the *high* and *low* patterns correctly.

3 Results

For a given number of stored patterns, the S/N of the output unit, or the cell, was obtained by presenting each stored input pattern as a cue and recording the resultant dendritic sum. The S/N was then calculated from the sample mean and variance of the dendritic sums due to the *high* and *low* input patterns. The S/N of the output unit as a function of the number of stored patterns is shown by the solid line in Figure 2 (plot "net"). It decreases exponentially as the number of stored patterns increases.

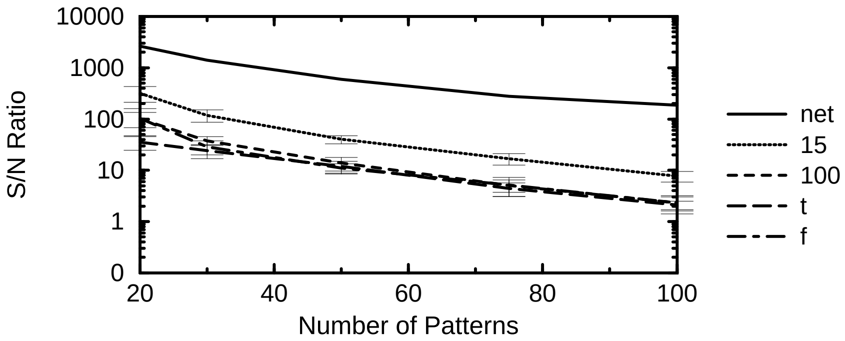

Fig. 2. S/N of the associative net and the pyramidal cell. Legend: net - associative net; 15 - random spikes at 15Hz; 100 - random spikes at 100Hz; t - time of first emitted spike; f - number of emitted spikes. Each pyramidal cell plot is the mean data from 10 runs at each number of stored patterns, with a different spatial distribution of synapses for each run. Error bars show 95% confidence limits.

Figure 2 also shows the S/N of the pyramidal cell when the dendritic sum is the soma depolarization recorded 100m*secs* after the initiation of input activity for input spike trains at 15H*z* (plot "15") or 100H*z* (plot "100"). The S/N of the output unit varies between 8 times and 90 times the pyramidal cell S/N for a given number of stored patterns. Conversely, a S/N of 320, say, is produced by at most 20 patterns in the pyramidal cell, but by 72 patterns in the associative net.

The random inputs at 15H*z* provide approximately 3 times the S/N of inputs at 100H*z*. The higher S/N of lower input frequencies is further demonstrated in Figure 3, where the S/N of 30 stored patterns is plotted against input frequency. This is due to saturation of the pyramidal cell synapses at high frequencies. At 100H*z*, 400 active synapses is enough to depolarize the entire apical dendritic tree to near the synaptic reversal potential.

For actual pattern recognition an output signal must be generated that indicates whether or not the input pattern has been recognised. When sodium and potassium channels are added to the soma, random 100H*z* input triggers action potentials. Figure 2 shows the S/N when the dendritic sum is the first emitted spike time (plot "t") or the number of spikes emitted in the first 200m*secs* of a simulation (plot "f"). This generation of action potentials converts the soma depolarization into an output signal without affecting the S/N. The S/N obtained from just a single spike is as good as that due to the average frequency of firing

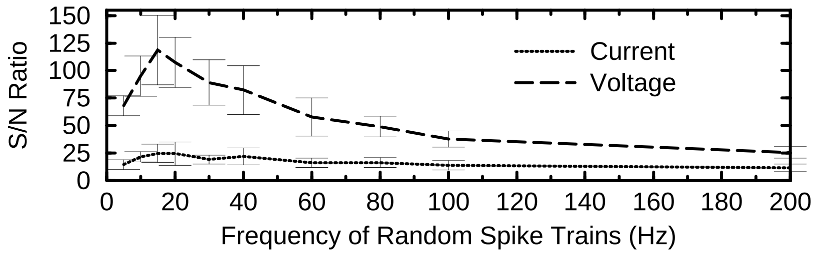

Fig. 3. S/N of soma depolarization for 30 stored patterns with different frequencies of random input. Each plot is the mean data from 10 runs at each number of stored patterns, with a different spatial distribution of synapses for each run. Error bars show 95% confidence limits.

over time. The first spike due to a *high* input pattern is generated about 40m*secs* after the onset of input activity at 100H*z*. This is the minimum time for the cell to recognise an input pattern.

4 Discussion

We have used computer simulations to estimate the signal-to-noise (S/N) ratio of a cortical pyramidal neuron when excitatory synaptic strengths are set according to a *clipped Hebbian* learning rule. For a variety of measurements of the weighted sum of the inputs, the S/N of the neuron is one to two orders of magnitude less than that of output units in the associative net model.

The performance of the neuron is influenced by the form of input signal used to represent an input pattern. For the continuous spike trains used here, lower firing frequencies are better than high frequencies due to synaptic saturation. This is a possible reason for the low average firing rates observed in the mammalian cortex. A single synchronous spike on each input line provides the same performance as the optimum continuous firing frequency (results not shown).

A number of different output signals provide the same recognition capability. Somatic depolarization is converted into action potentials without loss of S/N, and the time of the first emitted spike is as good a recognition indicator as the average firing frequency over a period of time. Thus recognition can be fast.

The use of S/N as a performance measure allows us to ignore issues of threshold setting across multiple neurons and the effect of partial connectivity. Variability in connectivity and in dendritic morphology across cells will introduce further noise into the dendritic sums and make threshold setting difficult.

The synapses in our model are distributed randomly and uniformly across a passive dendritic tree. Pyramidal dendrites are not passive [10] and the effect of active processes on the S/N needs to be explored. Some work has shown how nonlinear processes such as NMDA receptors and calcium channels in the dendrites can influence Hebbian learning of synaptic strengths [5] and amplify the signal from localised groups of synapses [12].

In conclusion, the spatial distribution of synapses across a dendritic tree introduces considerable noise that reduces a neuron's pattern recognition capability compared to a point unit in an artificial neural network. This form of noise is not included in network models that have been used to provide estimates of the storage capacity of the mammalian hippocampus [1, 8, 11, 13]. Thus these estimates are likely to be inflated.

Acknowledgement: to the MRC for financial support under PG 9119632.

References

1. M. Bennett, W. Gibson, and J. Robinson. Dynamics of the CA3 pyramidal neuron autoassociative memory network in the hippocampus. *Phil. Trans. Roy. Soc. Lond. B*, 343:167–187, 1994.
2. O. Bernander. *Synaptic Integration and Its Control in Neocortical Pyramidal Cells.* PhD thesis, California Institute of Technology, 1993.
3. O. Bernander, C. Koch, and R. Douglas. Amplification and linearization of distal synaptic input to cortical pyramidal cells. *J. Neurophys.*, 72:2743–2753, 1994.
4. T. Brown, E. Kairiss, and C. Keenan. Hebbian synapses: biophysical mechanisms and algorithms. *Ann. Rev. Neurosci.*, 13:475–511, 1990.
5. T. Brown, Z. Mainen, A. Zador, and B. Claiborne. Self-organization of Hebbian synapses in hippocampal neurons. In R. Lippmann, J. Moody, and D. Touretzky, editors, *Neural Information Processing Systems 3*, pages 39–45, San Mateo, California, 1991. Morgan Kaufmann.
6. P. Dayan and D. Willshaw. Optimising synaptic learning rules in linear associative memories. *Biol. Cybern.*, 65:253–265, 1991.
7. R. Douglas, K. Martin, and D. Whitteridge. An intracellular analysis of the visual responses of neurones in cat visual cortex. *J. Physiol.*, 440:659–696, 1991.
8. B. Graham and D. Willshaw. Capacity and information efficiency of the associative net. *Network*, 8:35–54, 1997.
9. M. Hines. A program for simulation of nerve equations with branching geometries. *Int. J. Biomed. Comput.*, 24:55–68, 1989.
10. D. Johnston, J. Magee, C. Colbert, and B. Christie. Active properties of neuronal dendrites. *Ann. Rev. Neurosci.*, 19:165–186, 1996.
11. D. Marr. Simple memory: a theory for archicortex. *Phil. Trans. Roy. Soc. Lond. B*, 262:23–81, 1971.
12. B. Mel. Synaptic integration in an excitable dendritic tree. *J. Neurophys.*, 70:1086–1101, 1993.
13. A. Treves and E. Rolls. Computational analysis of the role of the hippocampus in memory. *Hippocampus*, 4:374–391, 1994.
14. D. Willshaw. *Models of distributed associative memory.* PhD thesis, University of Edinburgh, 1971.
15. D. Willshaw, O. Buneman, and H. Longuet-Higgins. Non-holographic associative memory. *Nature*, 222:960–962, 1969.

This article was processed using the LaTeX macro package with LLNCS style