Synaptic Learning Models of Map Separation in the Hippocampus

Mark C. Fuhs, David S. Touretzky

Computer Science Department and Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

When rats trained to forage in one environment are exposed to a second, highly similar environment, their hippocampal place code exhibits a partial remapping in the new environment that becomes more complete with repeated exposures (Tanila, Shapiro, and Eichenbaum, 1997; Bostock, Muller, and Kubie, 1991). If the perforant path projection to CA3 functions as a pattern completion mechanism, and the DG projection via the mossy fibers performs pattern separation (O'Reilly and McClelland, 1994), then partial remapping can be understood as the combined effect of these two projections. We investigated learning rules that could be responsible for the gradual separation of two maps, and found that, while simple Hebbian learning and Hebbian covariance learning would not produce the separation effect, BCM learning was one rule that would.

Key words: Hippocampus, LTP, Cognitive maps, Plasticity

1 Introduction

Physiological evidence suggests that place cells in the dentate gyrus, CA3, and CA1 of the rat hippocampus represent different environments with different spatial activity patterns. These activity patterns, or place fields, can be clearly observed within the first ten minutes of exposure to a novel environment, and in general do not change with repeated exposure to the environment. When presented with two similar environments, cells initially have similar fields in both, but with repeated exposure to both environments, the fields become more distinct [2,4]. Because no changes were made to the environments, the

changes in the responses of place cells most likely reflect plasticity in the neural circuitry of the rat. We therefore investigated what properties of synaptic plasticity would be necessary to engender these experience-dependent changes.

2 Map Separation in the Hippocampus

In the Bostock, Muller and Kubie [2] experiment, rats foraged randomly for food in a cylindrical arena with either a white or black cue card. In the beginning, rats were repeatedly exposed to the cylinder with a white cue card. Then, place cells in the hippocampus were recorded while the rats were exposed first to the white card, followed by the black card. Rats were exposed to the two conditions each day for several days. Initially, place cell firing fields in the white card and black card environments were very similar, suggesting that the hippocampus used a single representation for both environments. However, upon repeated exposure to the two environments, place fields were found to diverge. Some cells were active in only one of the two environments; others were active in both environments, but with place fields that were topologically unrelated to each other.

This plasticity is consistent with two hypotheses about partial remapping. One is that individual place cells independently differentiated themselves between the two environments, without regard for the behavior of other place cells. The other, consistent with multiple map models of the hippocampus, is that all of the place cells changed their representations simultaneously, reflecting a switch to a new map for the second environment. Unfortunately, at present there is insufficient data available to distinguish between these two hypotheses.

In the Shapiro, Tanila and Eichenbaum [4] experiment, rats explored a four-arm radial maze for a brain stimulation reward. Each arm of the maze was marked with a distinct texture and odor, which served as "local" cues for the rat's location. The curtains surrounding the maze bore four distinctive objects, which served as "distal" cues. After repeatedly exploring in the first configuration of cues, rats were repeatedly exposed to two environments successively: the Standard environment, and a second environment, known as the Double Rotation environment, in which the maze arms were rotated 90 degrees counterclockwise and the distal cues were rotated 90 degrees clockwise. The place fields of cells recorded while the rat was in each of the environments were categorized into two types: those that rotated with one of the cue constellations, and those that either disappeared completely or appeared in a position inconsistent with the rotation of any of the cues. Though the remapping between the two environments remained partial throughout the experiment, the percentage of fields rotating with either distal or local cues decreased as the rat received more exposure to the two environments. Correspondingly, the per-

centage of cells with unrelated fields in the two environments increased with repeated exposure to both environments.

3 Modeling Map Separation

To simplify the computational requirements for simulation, we modeled the simpler of the two tasks, the Bostock *et al.* task. To do this, we created a virtual representation of the environments. Each environment contained either a white or a black cue card and the maze wall, represented by a single point. We simulated the entorhinal cortex with 1000 units, each tuned to a bearing and distance of one of the cues with a gaussian response function.

Our model of information processing was the dual pathway architecture. Specifically, information from the entorhinal cortex is propagated to CA3 both directly via the perforant path and indirectly, through the dentate gyrus, whose mossy fiber projections provide input to CA3. We modeled DG using 20,000 sigmoidal units which received input from EC; strong global inhibition kept all but the most active cells from firing, resulting in activity patterns that changed significantly with small changes in EC activity. To match physiological constraints, our model's mossy fiber projection from DG to CA3 was significantly more sparse than the perforant path projection; correspondingly, individual mossy fiber synapses were ten times as strong as perforant path synapses. For the 1,000 CA3 units, we used a linear firing rate model, with a non-zero threshold for firing. Additionally, each layer also incorporated a global inhibitory unit to provide competitive dynamics.

Physiological evidence indicates that LTP occurs in both the mossy fiber and perforant path synapses in CA3. However, the plasticity associated with the mossy fiber synapses appears to be nonassociative and is therefore unlikely to be useful in directly storing spatial information. We therefore explored what various forms of Hebbian plasticity in the perforant path inputs to CA3.

4 Learning Rules

The first learning rule we considered was simple Hebbian learning, $\Delta w_{ij} = x_i \cdot y_j$, where x_i is the activity of presynaptic unit i, y_j is the activity of postsynaptic unit j, and w_{ij} is the efficacy of the synapse between them. To bound the magnitude of the weights, a per-neuron weight vector normalization was performed after each epoch of learning.

The effect of this learning rule was the opposite of what we desired. Neurons

with fields of different strengths in the two environments developed stronger fields in both environments. This behavior follows from the fact that this learning rule maximizes the variance of the activity of the neuron; spatially selective activity in two environments yields greater variance in activity than activity in only one environment.

The second learning rule we considered was Hebbian covariance learning, $\Delta w_{ij} = (x_i - \overline{x}_i) \cdot \left(y_j - \overline{y}_j\right)$, where $\overline{x}_i$ is the (time averaged) mean activity of the presynaptic neuron and $\overline{y}_j$ is the mean activity of the postsynaptic neuron. This rule would occasionally cause a cell to differentiate its place fields in the two environments; however, this behavior was relatively rare. Furthermore, this learning rule causes weights to adapt when either the presynaptic or postsynaptic neuron is active. Thus, all of a neuron's weights will change with exposure to any of the environments in which it is active. This results in previous environments being forgotten at the same rate at which new environments are acquired. In our simulations, exposure to a second environment almost completely obliterated the memory of the first. This observation is incompatible with the existence of multiple maps in hippocampus.

The third learning rule we considered was the BCM [1] learning rule:

$$\Delta w_{ij} = x_i \cdot \Phi\left(y_j, \theta\right)$$
$$\theta = \overline{y}_j^2$$
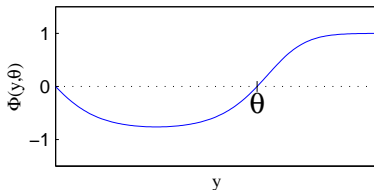
This learning rule has two important properties. First, learning only occurs when both the presynaptic and postsynaptic cells are active. This avoids interference problems between environments because only weights for features present in an environment will be modified. Additionally, the BCM learning rule defines a sliding threshold θ which determines the boundary between *sufficiently* active and *insufficiently* active. Strong postsynaptic activity will strengthen active synapses; weak postsynaptic activity will weaken active synapses. Thus, place fields which are strongly active in one environment but more weakly active in the second environment typically remain strong in the first environment but are further weakened in the second. This results in cells becoming more discriminating about the environment in which they are active.

In simulations, we classified the fields of cells into one of four categories (see Figure 1). The first category contained cells that were active in neither of the two environments. This was a result of the global inhibition suppressing their activity. The second and third categories contained cells that were active in either the first but not the second environment or the second but not the first. The fourth category contained cells that were active in both environments.
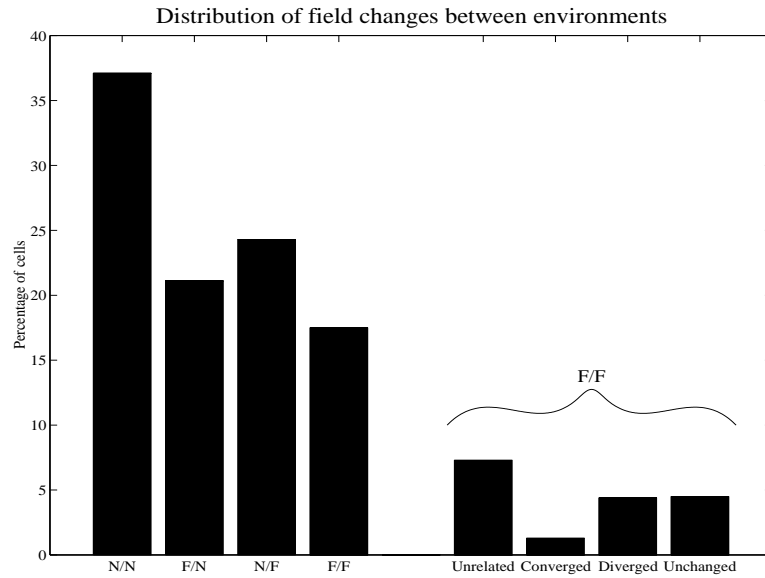
4

Fig. 1. The distribution of place cell behaviors in our model. Env1/Env2: F = Field present, N = No field present.

The fields of cells in the fourth category were analyzed for learning effects. Some fields were present in both environments but were either unrelated in the two environments, or were unchanged due to learning. Of the cells whose fields changed due to learning, most fields diverged in the two environments with repeated exposure to both. Specifically, fields which were initially correlated in the two environments were typically weakened in the second environment, resulting in a strong place field in the first environment, and little or no field in the second environment. A minority of cells that exhibited very strong fields in both environments were strengthened by the learning process, converging to the same field in both environments.

5   Discussion

The map separation seen in our model is most specifically attributed to the orthogonalization properties of DG and the BCM learning rule. The DG serves to orthogonalize input patterns, resulting in far more significant changes to the mossy fiber input to CA3 between environments than to the perforant path input. This has both direct and indirect consequences. The direct consequence is that, since CA3 cell activity is based on a contribution from the dentate, activity patterns in CA3 will naturally be less overlapping than if CA3 activity were solely a function of perforant path input. The cause of many of the CA3 cells being active in only one of the two environments was the contribution of activity from DG. The indirect consequence is that, even if a CA3 cell receives enough input from the perforant path to be active in both environments,

the DG contribution often biases the strength of the activity enough to affect whether the activity of the field falls mostly on the positive or negative portions of the BCM curve. Thus, the small differences in activity due to the dentate are magnified by the BCM learning rule to produce a map separation effect.

There are two interesting predictions made by the model. First, while maps in both environments will undergo some change as a result of learning, the map of the newer environment will change more significantly. This asymmetry is due to the prolonged initial exposure of the rat to the first environment. This exposure strengthens weights, reinforcing the map for the first environment. These stronger weights make weakening fields in the first environment more difficult than in the second environment. Thus, a cell with a field in each environment is more likely to lose this field in the second environment. A second prediction made by the model is that DG fields should show remapping prior to CA3 or CA1 fields, since it is DG's novel representation of the second environment that in part engenders differentiation between the two environments.

References

[1] E L Bienenstock, L N Cooper, P W Munro, Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex, J. Neurosci 2 (1982) 32-48.

[2] E Bostock, R U Muller, J L Kubie, Experience-dependent modifications of hippocampal place cell firing, Hippocampus 1 (2) 193-206.

[3] R C O'Reilly, J L McClelland, Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off, Hippocampus 4 (6) (1994) 661-682.

[4] M L Shapiro, H Tanila, H Eichenbaum, Cues the hippocampal place cells encode: Dynamic and hierarchical representation, Hippocampus 7 (6) 624-642.

Mark C. Fuhs is third-year doctoral candidate in the Computer Science Department and the Center for the Neural Basis of Cognition at Carnegie Mellon University. His current research involves understanding the mathematical and computational underpinnings of spatial and contextual information processing.

David S. Touretzky is a Senior Research Scientist in the Computer Science Department and the Center for the Neural Basis of Cognition at Carnegie Mellon University. He received his PhD in Computer Science from Carnegie Mellon in 1984. Dr. Touretzky's research interests include representations of space in the rodent brain and computational models of animal learning.