ELSEVIER

Neurocomputing 32–33 (2000) 493–499

NEUROCOMPUTING

www.elsevier.com/locate/neucom

Entorhinal/dentate excitation of CA3: A critical variable in hippocampal models

S. Polyn, X.B. Wu, W.B. Levy*

Health Science Center, Department of Neurological Surgery, University of Virginia, Box 420, Charlottesville, VA 22908, USA

Accepted 13 January 2000

Abstract

We investigate how the strength of entorhinal cortical inputs during training affects learned performance using computer simulations of a minimal computational model of hippocampal region CA3. After the model learns two partially overlapping sequences, it is tested on two contradictory prediction problems — disambiguation and goal-finding. Relative to total activity, the activity level of entorhinal inputs during learning profoundly affects performance on each task. The optimal input levels differ for the two sequence prediction problems, but a small region of overlap exists where both tasks can usually be performed successfully. This sensitivity to relative input activity suggests critical tests of the model.

Keywords: Hippocampus; CA3; Recurrent network; Sequence learning; Disambiguation

1. Introduction

A biologically plausible, neural-like network model of the hippocampus solves a range of problems, including sequence completion and prediction [3,4]. A fundamental aspect of hippocampal models is the relationship between the feedback excitation of CA3 recurrent neurons versus the feedforward excitation provided by the entorhinal cortex both directly and indirectly through the dentate gyrus [1,3,7]. Although we have argued that the feedback influence for firing CA3 neurons must be greater than the feedforward influence [3,4], we have not characterized the sensitivity of our hippocampal model to the relationship between these two excitatory influences. Here the relevant comparisons are made on two contradictory sequence prediction

* Corresponding author.

0925-2312/00/$ - see front matter
PII: S0925-2312(00)00204-6

problems; disambiguation, where the model must be able to distinguish between two overlapping sequences, and goal-finding, where the model must transition from one sequence to the other (Fig. 2). The performance of the model is dependent upon the relative proportion of feedforward and recurrent excitation. Because the model shows a well-defined sensitivity to this relative excitation, explicit parametric descriptions of this sensitivity can make strong predictions for testing by experimental neuroscientists.

2. Methods

The network, a model of region CA3 of the hippocampus, is a sparsely connected (10%), recurrent excitatory system (for more details see [4]). In addition to the recurrent connections, each of the 4096 McCulloch-Pitts neurons receives an external input that represents the entorhinal cortex and dentate gyrus (EC/DG) inputs. This single input is hypothesized to be very powerful if activated (Fig. 1).

The model uses a post-synaptic associative modification rule [2] that is disabled during testing. For details of the computations performed by the elements of the network, see [8] (Shon et al., this volume). The activity of the network (the number of neurons active on a given time-step) is analyzed in terms of the fraction of external inputs $I(t)$ and the fraction of neurons fired by recurrent activity $R(t)$ on that time-step. Without the time variable, the term I refers to the average $I(t)$ across all time-steps for that trial. Activity fluctuations are controlled by two inhibitory interneurons and a resting conductance (K_R, K_I and K_0, respectively). Across training the total activity fluctuates due to imprecise activity control. To make the simulations more reliable, the value of feedback inhibition (K_R) is varied after each trial to maintain more nearly constant activity levels.

Each training trial consists of one of the two sequences, each 12 patterns long. Each individual input pattern of the sequences corresponds to a set of neurons in the network that are activated for a single time-step by an external connection. As Fig. 2 shows, each input sequence contains three orthogonal subsequences. The patterns

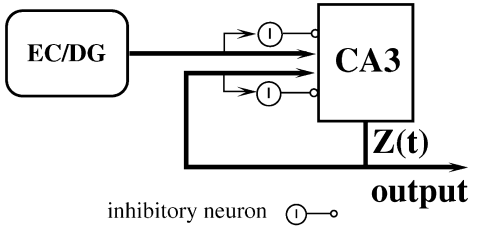

Fig. 1. The hippocampal model. Area CA3 of the hippocampus is modeled by a set of 4096 neurons. A different EC/DG input projects to each neuron of the CA3 system, forcing that CA3 neuron to fire upon activation of its one and only EC/DG input.

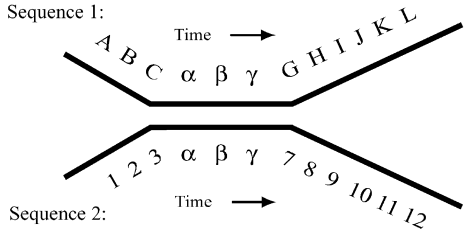

Fig. 2. A schematic of the input sequences. Two input sequences are presented to the network. Each sequence is composed of 12 patterns. The EC/DG input patterns are broken up into three orthogonal subsequences — the initial subsequence, the shared subsequence and the tail subsequence. A pattern consists of a set of neurons being turned on for one time-step. The patterns within a subsequence have a slow shift rate, that is, each pattern has 85% of its neurons in common with the patterns preceding and following it.

within each subsequence have a slow shift rate; that is, each pattern has several neurons in common with the patterns immediately preceding and immediately following it. Notably, the two sequences share a common, central subsequence of three patterns (α, β, and γ of Fig. 1). There are 350 presentations of each sequence, with the presentation of the two sequences interspersed.

Testing consists of sequence completion in response to a partial input. A cosine comparison (the normalized dot product of two vectors) is used to compare the network states that evolve in response to the test input with those produced by a complete sequence of input patterns. By definition, the largest value at each time-step is the decoded network state at that moment. The networks are first tested on the disambiguation task. From each starting point (1 or *A* of Fig. 2), the network must predict the appropriate path to the end of the learned sequence (12 or *L*, respectively). In this problem, the test input is only activated for a single time-step. In the goal-finding task, the network must predict the path to an alternate goal (e.g. from 1 to *L*) using a partial goal description. The partial goal description consists of a subset (8–26%) of the neurons from the goal pattern (e.g. *L* when starting at pattern 1). These neurons are active throughout testing and produce an *induced attractor* [5]. There are at least two ways a simulation can solve the goal-finding problem. It can reiterate a learned sequence and transition smoothly from sequence 1 to sequence 2, or it can jump to its goal without passing through the choice point or its neighbors. When the simulation jumps to the goal state it does not reproduce the patterns around the choice point (e.g. γ of Fig. 1).

There are two sources of randomness in the model, the initial connectivity of the neurons, and the initial randomization of activity before each trial of training. Because simulations have different initial connectivities, there is a range of performance across simulations. For each combination of activity level and input size, a set of 10 simulations (i.e. 10 initial connectivities) are run and tested. Each of these 10 trained networks is tested with 10 different initial activities on both problems.

3. Results and discussion

The simulations suggest several general hypotheses about information processing in the hippocampus. First, for any one task, there exists an optimal level of input activity. Second, at this optimum, the recurrent activation is stronger (i.e. fires more neurons) than the input activity. Third, although the disambiguation and goal-finding tasks have different optimal levels of input activity, there is an in-between level of activity that is not far from the optima of either task.

Fig. 3a shows that, as a function of total activity, there is an optimal level of input activation, and this level yields 100% performance on the disambiguation task. As total activity increases from 3% to 6%, the simulations become more robust, in the sense that a large percentage of simulations can perform the disambiguation problem across a larger range of inputs. For each total activity level the simulations begin to fail at a certain level of input activity. The input size relative to the total activity becomes important.

In the region of best performance on the disambiguation task, the recurrent connections account for a larger percentage of the total activity than the input connections. The disambiguation performance curves as a function of total activity become quite similar when graphed against I (Fig. 3b). Apparently, the ratio of external input to total activity (I) is important and determines whether a network can encode a sequence usefully [4]. The ideal relative level of input activity for the disambiguation task is between 10% and 20% of the total network activity (i.e., $0.1 < I < 0.2$).

In contrast to the disambiguation task, goal-finding has a slightly different optimal level of input, as seen in Fig. 4. The performance levels have been optimized for path

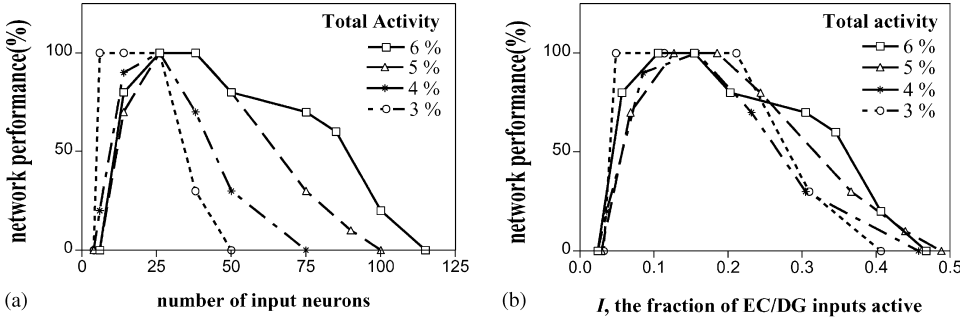

Fig. 3. (a) Input strength interacts with total activity to affect disambiguation performance. The strength of the EC/DG input affects the probability a simulation will solve the disambiguation problem. Network performance measures the percentage of networks, out of a set of ten, that perform disambiguation successfully. The curves extend further to the right when activity is raised. (b) Normalization of input activity aligns disambiguation performance curves. If the abscissa from Fig. 3a is changed to I, the fraction of total network activity accounted for by the input neurons, the performance curves for each activity level line up more closely. Apparently, the optimal level for I is between 0.1 and 0.2 for the disambiguation problem.

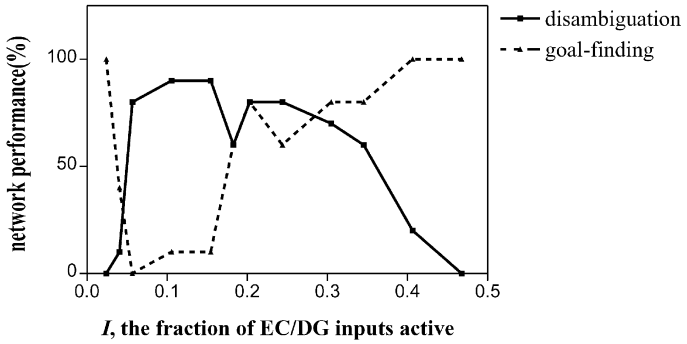

Fig. 4. Disambiguation and goal-finding are differentially affected by level of I. There are different optimal levels of I for different tasks. As I increases, disambiguation performance reaches a peak before goal-finding performance. While goal-finding continues to climb as I increases beyond 0.3, at that value of I the networks are jumping to the goal state rather than following a path. Simulations were run at $\sim 6\%$ activity for both tasks.

following goal-finding. Here the ideal I for goal-finding is between 0.2 and 0.3. Although not perfectly overlapping with disambiguation, a compromise can be reached at $I = 0.2$, where both disambiguation and goal-finding are near optimal.

4. How the network solves the problems

There are two kinds of information that are dynamically represented by the pattern of cell firing at any given time. The external activity $I(t)$ represents the immediate present while the recurrent activity $R(t)$ carries information forward from the past (also called context past). The $R(t)$ activity at each time t is excited by the previous activities, $I(t-1)$ and $R(t-1)$. If an $I(t-1)$ is large relative to its companion $R(t-1)$, then less information from the more distant past, i.e. before $t-1$, will be coded in the present $R(t)$.

Fig. 4 shows that disambiguation is stronger than goal-finding for $0.1 < I < 0.2$. A simulation is able to find a code that will take it through each sequence reliably. In these cases, the R activity is able to carry information about the initial unshared subsequence across the shared subsequence (α, β, and γ of Fig. 2).

To quantify this carrying of information across the shared subsequence, we examined the coding distinction between a simulation's learned cell firing patterns for γ when they appeared in one sequence versus the other (γ_1 vs. γ_2, Fig. 5) for different levels of external activity. A simulation that disambiguates the two sequences must code γ differently for each sequence. For $0.1 < I < 0.2$, the cosine value of γ_1 vs. γ_2 is less than 0.4. If I is larger, ($0.2 < I < 0.3$) the γ's become more similar, and disambiguation performance begins to decline. Apparently, as the cosine value increases beyond 0.4, $I(t-1)$ comes to dominate over $R(t-1)$ in determining $R(t)$ activity.

On the far left of Fig. 4, where I is very low (< 0.05), the model fails on the disambiguation task. Here, Fig. 5 shows that the cosine similarity between γ_1 and

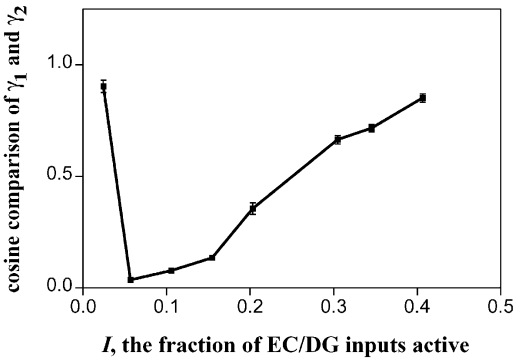

Fig. 5. γ_1 and γ_2 show increasing similarity as I increases. For disambiguation to be successful, the end pattern of the shared subsequence (γ of Fig. 2), must be coded differently when it appears in one sequence versus the other (see text). The similarity of cell firing to γ in the two situations increases as I increases, except at the first point of the graph which is explained in the text. In order for the two γ's to be coded differently, the network must retain some information about the initial subsequence across the first two patterns, α and β, of the shared subsequence. A cosine comparison is used to compare the codes used by the network for γ. Simulations were run at $\sim 6\%$ activity.

γ_2 for this I is about 0.9. The high cosine value of this anomalous point is most likely due to the $R(t)$ activity of the simulation falling to near-zero values during training on the initial subsequence, and rising for presentation of the shared subsequence (which is presented twice as often during training). The $R(t)$ activity generated by the shared subsequence is apparently unaffected by the small $I(t-1)$ influence of the tail subsequence, because the cosine similarity of the two sequences extends all the way to the end of the sequences.

For a slightly higher value of I (0.056), the simulations fail at disambiguation two out of ten times despite γ_1 and γ_2 being quite different (cosine ~ 0.1). In these cases, the tail ends of the two sequences are coded similarly (cosine values greater than 0.8 were seen when comparing end of sequence activation patterns between the two sequences) by the end of training. Here, the external activity $I(t-1)$ is strong enough to overcome the network's propensity [6] for following random sequences but not strong enough to overcome the similarity of the overlapping subsequence. $R(t)$ activities for the time-step following γ are heavily influenced by the $I(t-1)$ activity generated at γ. The $R(t)$ activity of G and 7 are thus driven to be similar to one another (Fig. 5). With repeated training, this tendency towards similarity cascades forward. After many training cycles, both the shared and tail subsequences of the two sequences are represented by a highly similar set of active neurons.

For values of I greater than 0.3, $I(t-1)$ activity overwhelms $R(t-1)$ activity in determining $R(t)$ and the subsequences become coded too similarly for disambiguation to be successful at all.

In order to solve the goal-finding problem to our satisfaction, the simulation must be able to both follow a path and reach the goal. For the extreme values of I ($I < 0.05$; $I > 0.3$) the simulation is able to do goal-finding by jumping to the final goal pattern,

but it is unable to follow a path, so this performance is unacceptable. In the range of optimal disambiguation performance ($0.1 < I < 0.2$), goal-finding is not performed well by the simulations. As Fig. 5 shows, γ_1 and γ_2 are coded differently for these values of I (less than 0.2 similarity). Goal-finding performance improves as the cosine comparison between γ_1 and γ_2 gets larger ($0.2 < I < 0.3$). Here, the sub-sequences are coded similarly enough that, with the prompt from the partial goal description, a network can switch paths. However without that prompt, disambiguation would be performed. In this region of I, where disambiguation and goal-finding can be performed simultaneously by a single network, the networks are using a truly flexible code to represent the two sequences. Thus, only by achieving a balance of input activity and recurrent activity can the network be trained to solve both tasks.

Acknowledgements

This work was supported by NIH MH48161 to WBL and MH57358 to N. Goddard, and by the Department of Neurosurgery.

References

[1] M.E. Hasselmo, B.P. Wyble, G.V. Wallenstein, Encoding and retrieval of episodic memories: role of cholinergic and GABAergic modulation in the hippocampus, Hippocampus 6 (6) (1996) 693–708.
[2] W.B. Levy, Associative encoding at synapses, Proceedings of the Fourth Annual Conference of the Cognitive Science Society, 1982, pp. 135–136.
[3] W.B. Levy, Acomputational approach to hippocampal function, in: The Psychology of Learning and Motivation, Vol. 23, Academic Press, San Diego, 1989, pp. 243–305.
[4] W.B. Levy, A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks, Hippocampus 6 (1996) 579–590.
[5] W.B. Levy, X.B. Wu, R. Baxter, Unification of hippocampal function via computational/encoding considerations, Int. J. Neural Systems 6 (Supp. 1995) 71–80.
[6] A.A. Minai, W.B. Levy, Predicting complex behavior in sparse asymmetric networks, Neural Inform. Process. Systems 5 (1993) 556–563.
[7] E.T. Rolls, Functions of the primate hippocampus in spatial processing and memory, in: Neurobiology of Comparative Cognition, Erlbaum, New Jersey, 1990, pp. 339–362.
[8] A.P. Shon, X.B. Wu, W.B. Levy, Using computational simulations to discover optimal training paradigms, Neurocomputing (1999).