

# Sequence Prediction and Cognitive Mapping by a Biologically Plausible Neural Network

Colin Prepscius and William B Levy  
Department of Neurological Surgery  
University of Virginia Health Sciences Center  
Charlottesville, VA 22908

*Cognitive mapping involves encoding a representation of an environment in memory. A cognitive map can be formed from collections of sequences of representations of environmental features that change in time. Sequence prediction is important to this theory because it enables efficient searching of the map. We propose a neural network inspired by the hippocampus that, by virtue of a recurrent hidden layer, predicts sequences at a faster rate than they are originally presented.*

## I. Introduction

### **Cognitive Mapping**

The role of the hippocampus in cognitive mapping has been studied extensively (see, for example, O'Keefe and Nadel, 1978). These studies propose that the hippocampus recodes and retains information about an organism's environment in such a way that this information can be used to navigate the real world (the rat in the maze being the classic example). According to Tolman (1932), a cognitive map enables an organism with three properties: 1) The ability to navigate between any two points, 2) The ability to find a detour when the most efficient path is blocked, and 3) The ability to find short cuts if better routes are suddenly opened.

Both Saypoff, Muller and Kubie (1992) and Schmajuk and Thieme (1992) treat cognitive mapping essentially in terms of graph theory; both use discrete (and orthogonal) representations of "places" and "views", in which the firing of each neuron represents a different single place or graph vertex. Adjacencies between places, or *edges* between graph *vertices* (in Schmajuk's case), or distances between them (in Saypoff's case) are represented by synaptic connections between neurons. During maze navigation with these models, all possible paths from the starting point to the goal point are mentally explored, until the shortest path is found. We propose that this is an essentially un-neural-like method of solving the problem, for the following reason: although this method may work well for small graphs (or small mazes), because the number of edges increases with the square of the number of vertices in a graph, in the real world exploring all possible paths would be difficult. Another limit on this method is that searching is carried out no faster than the rate of learning; because point B in the maze is encountered directly after point A during exploration, during mental navigation (or *prediction*), point B is not arrived at any faster from point A. Faster searches of the cognitive map are clearly desirable from the point of view of an animal's survival.

### **Sequence Prediction**

The role of sequence prediction in cognitive mapping is now clear. A succession of points or places an animal visits during exploration of its environment is a *sequence* which the animal needs to remember and recall later during mental navigation of the environment. Many sequences make up a cognitive map - this could be considered a generalization of Tolman's learning of expectancies (stimulus-response-stimulus sets).

The ability of neural networks to remember and recall sequences of varying degrees of difficulty has been studied extensively (see, for example, Mozer 1993). Biologically plausible networks of the type studied here (see below) also effectively recall a wide variety of sequences (Minai, Barrows and Levy, 1993, Minai and Levy 1993a). Several aspects of sequence learning have not been previously explored, to our knowledge: First, a truly fast-time prediction mechanism would not just recall a sequence in the order in which it is presented, but would predict the end of a sequence (or better yet, any arbitrary point in the sequence) by skipping over or compressing intervening points. Second, we propose that sequences in the real world change slowly in comparison to an animal's physical movement, or exploration of its environment. Thus, an animal would not encounter place A and then immediately place B, but would instead see a more or less continuous progression between the two. Third, an even more powerful ability for a neural network, based on these latter two abilities, is *temporal feature extraction*, which would enable an animal's cognitive map to mentally navigate between the most prominent patterns in a sequence, providing both a faster-than-real-time method of prediction, and a method of encoding these features into a cognitive map.

A neural network capable of fast-time prediction on a slowly-changing sequence is described below.

## II. Description of Neural Network

The network described here is illustrated in Figure 1, and is inspired by the architecture of the hippocampus (Levy 1989). The input layer, roughly corresponding to the entorhinal cortex, projects to two different layers, corresponding to the CA3 and CA1 layers of the hippocampus. The CA3 layer projects both back onto itself and onto the CA1 layer. The CA1 layer gives the final output.

The network studied is small; the CA3 region contains 512 neurons, and the CA1 region 100 (more CA1 neurons could be used but are not necessary for the inputs used here). By virtue of the recurrent connections between CA3 neurons, CA3 can recode each input pattern not only for each independent time-step but also in the context of its immediate history of neuron firings. This aspect of contextual recoding of an input signal has been studied by Minai, Levy, and Barrows, (1993) who found it to be a key factor in the remembrance of ambiguous sequences. This recoding property will also be important in the studies presented here.

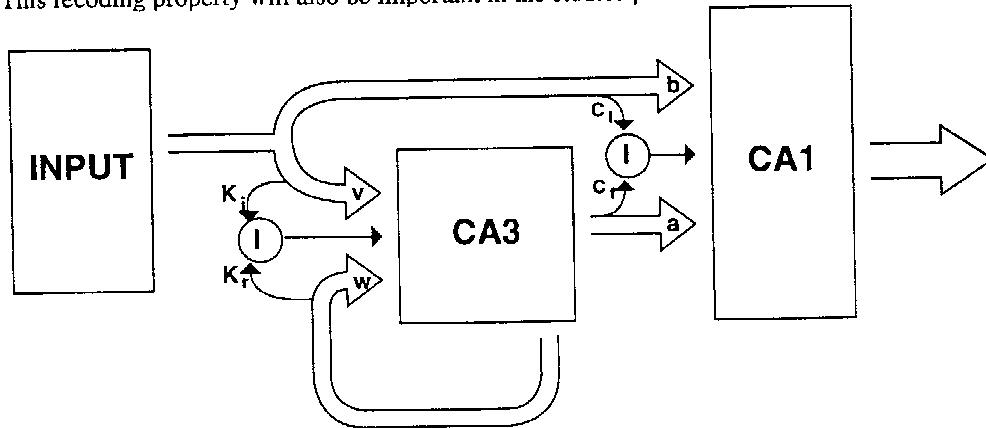


Figure 1 - Block diagram of the neural network inspired by the CA1 & CA3 regions of the hippocampus. Arrows represent projections between layers (connections). The small circles labelled I represent single inhibitory neurons.

Another important and biologically valid feature of the network is sparse, random connectivity. The recurrent CA3 projection (arrow w of Figure 1) is only 5% connected; neurons will therefore have connections to 25 other neurons in the CA3 layer. Connectivity from CA3 to CA1 (arrow a of Figure 1) is also 5%. Inputs project from the input layer to the CA3 (arrow v of fig.1) with a 1-2 connectivity, so if a total of 30 neurons are used in the input data, only 60 neurons in the CA3 are directly activated. Each input connection is strong enough to guarantee that its postsynaptic neurons will fire (arrow b of Figure 1). Each input connection is strong enough to guarantee that its postsynaptic neurons will fire (input designated by x in equations 1 and 2). The input is modulated via shunting inhibition through a single interneuron for each layer. The excitation,  $y_i$  of a CA3 neuron, and  $p_k$  from a CA1 neuron, is thresholded at a value  $\beta=0.9$ . Activity for the CA3 and CA1 neurons is described by (1, 2), where  $z_i(t)$  is the binary output of neuron i in the CA3 layer at time t, and  $q_k(t)$  is the output from a CA1 neuron, and  $h(x)$  denotes the Heaviside function.

$$\text{CA3 activation function : } y_i(t) = \frac{vx_{[i]}(t) + \sum_j w_{ij}z_j(t-1)}{\sum_j w_{ij}z_j(t-1) + K_r \sum_j z_j(t-1) + K_f \sum_j x_j(t)} \quad (1)$$

$$z_i(t) = h(y_i(t) - \beta) \quad 0 < \beta < 1$$

$$\text{CA1 activation function : } p_k(t) = \frac{bx_k(t) + \sum_j a_{kj}z_j(t)}{\sum_j a_{kj}z_j(t) + C_k \sum_j z_j(t) + C_f \sum_j x_j(t)} \quad (2)$$

$$q_k(t) = h(p_k(t) - \beta) \quad 0 < \beta < 1$$

### III. Sequence Prediction

#### How sequence prediction is accomplished

Sequence compression by the CA3 layer is illustrated below. A sequence consisting of three temporally overlapping patterns (A, B, C) is presented to the network repeatedly. After the network modifies its weights according to some associative scheme, the original sequence occurs during recall. That is, presentation of pattern A to the network would result in the recall of pattern B at a later time which would, in turn, recall pattern C at an even later time. However, the network does not necessarily end up conserving the original time course of the sequence. Any associative weight modification between A and B will be used during the next presentation of the sequence at the start of pattern A. Thus, pattern A will immediately recall a fraction of B. During training, this fraction grows until A effectively predicts the onset of B. Patterns B and C will exhibit the same behavior except that, eventually, pattern C will be predicted while pattern A is still occurring. That is, pattern A will immediately predict pattern C.

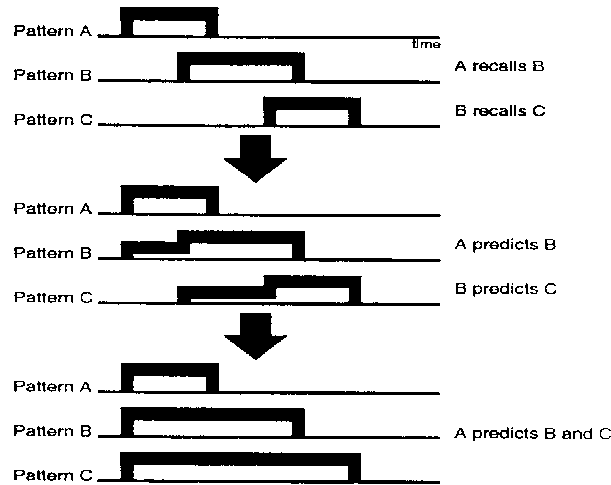


Figure 2 - Graphical representation of sequence compression.

Because we are running our networks discretely in time, a relevant question might be: how do these patterns persist for more than one time step? Because the CA3 layer is recurrently connected to itself, any pattern presented at time step 1 will persist with a decreasing representational strength and a different code at time step 2. This is, in fact, the contextual recoding explained above.

In these simulations, another factor assisting sequence compression is slowly-changing sequences - sequences of patterns that overlap spatially and thus give smooth transitions between patterns. Spatial overlap is, however, equivalent to temporal overlap. Other processing element characteristics that would support this overlapping effect include capacitance and delay lines, but they were not included in our model. Another relevant question might be: how can intervening patterns (such as pattern B above) be skipped entirely to avoid too much activity in the network? This is achieved through attractor properties of the network and through inhibition. Because pattern C ends up being presented to the network the most, it becomes a self-exciting "fixed point" - it is the end of the sequence - and consequently has the strength to inhibit the other patterns.

#### Results

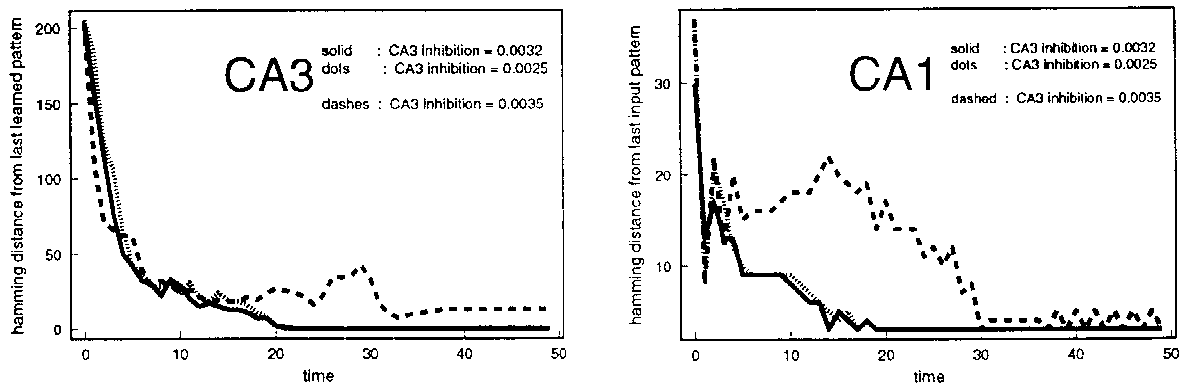
One illustrative sequence the network was trained on is 40 patterns in length. Each pattern contains eight on-bits, represented by the firing of 8 input neurons. The Hamming distances between successive patterns is 2 - any pattern will be missing 1 bit from the previous pattern and will also contain 1 additional on-bit. The network is presented with the entire sequence for 100 trials and allowed to modify its weights according to the postsynaptic learning rule (equ. 3) (Levy, Colbert and Desmond, 1990). The value  $\mu$ , the learning rate, is generally kept small in the model here, but Minai and Levy (1993a) have studied higher learning rates.

$$\text{Post-Synaptic Learning Rule : } w_{ij}(t) = w_{ij}(t-1) + \mu z_i(t)[z_j(t-1) - w_{ij}(t-1)] \quad (3)$$

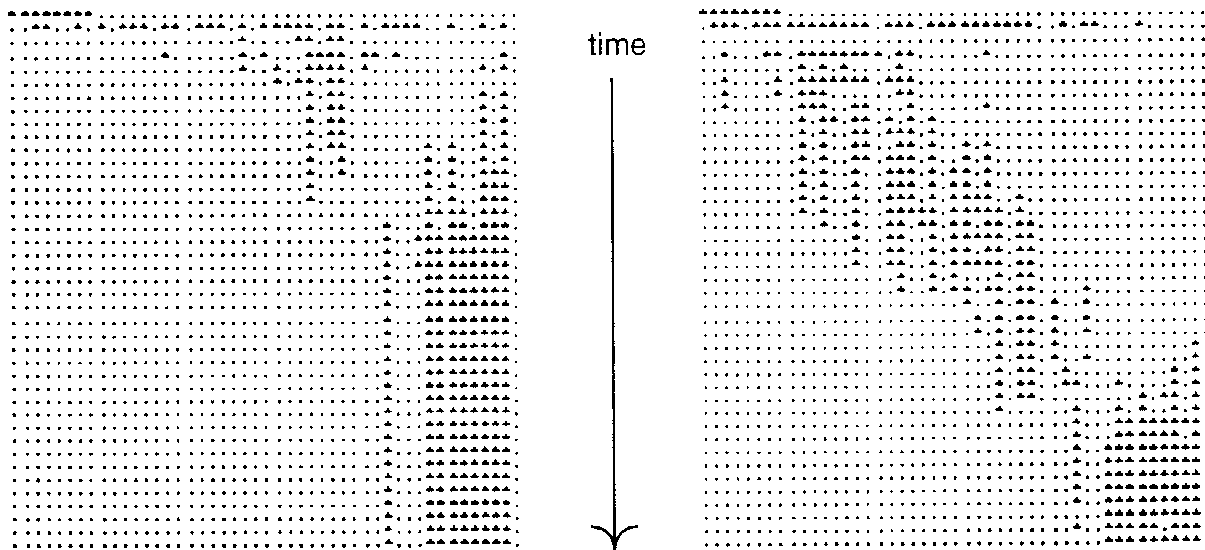
Before each trial, the network is reset randomly. During recall, the network is given the first pattern in the sequence as a prompt and allowed to relax.

When the network is given just the first pattern in the sequence at time step 1, the network exhibits one of two modes of behavior. In the first mode, the network exhibits *sequential recall*: it presents the patterns in the sequence in order at roughly the same rate as the patterns were presented (albeit with added noise), before settling at the last pattern (apparently a fixed point). The other mode is *predictive recall*, wherein the network converges to the last pattern at a significantly faster rate than when the sequence was presented. In all networks, predictive recall appears at values of inhibition (modulated through the interneurons) lower than the inhibition the network was trained at, and sequential recall appears at generally higher values of inhibition.

The speed of recall can be viewed graphically using the Hamming measure for distance between vectors. By plotting the distance from the current pattern to the last pattern as a function of time, the two different modes of recall can be distinctively pictured. Data for a typical network is given below (figure 3), for both the CA3 and CA1 layers, at different inhibition levels.



**Figure 3** - Recall of a length 40 sequence by a network at different CA3 inhibition levels in the CA3 (left) and CA1 (right) layers. The dashed lines illustrate sequential recall, the other two lines of each graph show predictive recall. Plotted in each graph is the Hamming distance of the representation at time  $t$  from the last pattern in the sequence.



**Figure 4** - Firing patterns of CA1 (presented as Hamming distances in Fig 3, right) illustrate predictive (left figure) and sequential (right figure) recall. Each \* represents the firing of a neuron. Each row represents the entire state of the CA1 at a particular point in time. Rows further down represent states at later times. Thus, the evolution of the CA1 layer through time can be traced directly.

Predictive and sequential recall can also be viewed directly by looking at the actual firings of the neurons over time (Figure 4). In fact, by doing so, several additional features of the representations pictured during recall become apparent: After the initial prompting input to the network, consisting of the first pattern in the sequence, at the next time step (the second timestep, that is) nearly all of the neurons involved in all subsequent patterns turn on, and then turn off at the third time step. This rapid change in activity levels is due to the inhibition not stabilizing the network quickly enough, and indicates, in one sense, that compression of the sequence has in fact progressed as far as possible- the first pattern does recall the last pattern (and all intervening patterns). The potential usefulness of this property is discussed in part IV. Following this initial burst, the network then recalls the sequence in predictive or sequential fashion as outlined above.

## **IV. Discussion**

### **Inhibitory Control**

Inhibition plays a vital role in controlling the speed of recall. Minai and Levy (1993b) found that for a similar network (sparsely connected, recurrent) which underwent no synaptic modification, inhibition was the controlling factor in setting the activity level of the network as it relaxed from a random initial starting point (see also Amari, 1972). In brief, low inhibition levels, when not allowing the network to completely turn on (a non-interesting fixed point), led to high-activity, high-frequency activity patterns. High levels of inhibition led to low-activity, low-frequency activity patterns. Although the authors of this paper pointed out that these analyses were not applicable generally to networks which undergo synaptic modification, and no similar study has been made of such networks, the same ideas seem applicable here.

On a macroscopic scale, inhibition in the CA3 controls the rate at which recurrent feedback decays. At low inhibition levels, at any given time-step, a larger number of context neurons is likely to turn on. This increases the chance that a portion of the last context pattern will be present. Since the last pattern is a fixed point of the system and reinforces itself through self-excitation, even a small portion of this pattern will quickly recall the rest of the pattern, inhibit the other previous intervening patterns, and lead to predictive recall. Higher inhibition levels reduce the chance of the last pattern appearing at any given time, and lead to sequential recall through learned patterns, rather than fast convergence.

### **Sequence Prediction and Cognitive Mapping**

Although these studies are preliminary, sequence prediction promises to be an important factor in cognitive mapping. An animal is hypothesized to explore its environment, coding afferent information in the context of its history. This information, in the form of temporal sequences, is later recalled during mental navigation of the assembled cognitive map. Based on various prompts, including the actual position of the animal and its desired position (driven perhaps by other hypothesized mechanisms, such as hunger and thirst), an appropriate sequence is recalled. The animal would need to vary the strength of inhibition (perhaps using one of the modulatory systems projecting to the hippocampus and perhaps reflected in theta rhythm) to affect the speed of recall. The networks studied here have the property that at the time step directly after the initial prompt, they recall a mixture of subsequent patterns in the sequence. This brief burst facilitates predictive recall due to the mechanism described above, but it is unclear whether there must be such a step. The usefulness of this single time-step can be readily imagined, however.

## **V. Conclusion**

The network studied here, while conserving biological plausibility, introduces an effective method for dealing with the memory and recall of time-varying sequences. Predictive recall is a property, first introduced by Levy (1989), that provides a useful mechanism for navigating cognitive maps. The complete cognitive mapping properties as defined by Tolman have not been realized so far, but will be explored in future work.

The theoretical framework for making additional steps in this direction, including learning of sequences that change at varying rates, learning multiple sequences, extraction and convergence to temporal features, and fast learning (Minai and Levy, 1993a), has been begun.

## Acknowledgements

This research was supported by the following grants: NIMH MH00622; NIMH MH48161; NSF MSS-9216372; and EPRI RP8030-08. It was also supported by the Department of Neurosurgery, University of Virginia, Dr. John A. Jane, Chairman.

## References

- Amari S. 1972. "Characteristics of Random Nets of Analog Neuron-Like Elements," *IEEE Transactions on Systems, Man, and Cybernetics*, VOL.SMC-2, No.5, November.
- Levy W. 1989. "A Computational Approach to Hippocampal Function," in Computational Models of Learning in Simple Neural Systems (R.D.Hawkins & G.H.Bower, eds.) [The Psychology of Learning and Motivation, vol. 23] San Diego: Academic Press, 243-305.
- Levy W, Colbert C, Desmond N. 1990. "Elemental Adaptive Processes of Neurons and Synapses: A Statistical/Computational Perspective," in: Neuroscience and Connectionist Theory, M. Gluck & D. Rumelhart (eds.), Hilldale, NJ, Lawrence Erlbaum Associates, 187-230.
- Minai A, Barrows G, Levy W. 1993. "Disambiguation of Pattern Sequences with Recurrent Networks," (in submission to *World Congress on Neural Networks*).
- Minai A, Levy W. 1993a. "Sequence Learning in a Single Trial," *Proceedings of the World Congress on Neural Networks*, Vol II.
- Minai A, Levy W. 1993b. "The Dynamics of Sparse Random Networks," *Biological Cybernetics*, in press.
- Mozer M. 1993. "Neural Net Architectures for Temporal Sequence Processing," (to appear in : A. Weigend & N. Gershenfeld (Eds.), Predicting the Future and Understanding the Past. Redwood City, CA: Addison Wesley Publishing).
- O'Keefe J, Nadel L. 1978. The Hippocampus as a Cognitive Map. Clarendon Press, Oxford.
- Saypoff R, Muller R.U, Kubie J.L. 1992. "How Place Cells Connected by Hebbian Synapses can Solve Spatial Problems," *Soc. Neurosci. Abstr.* 2. 1211.
- Schmajuk N, Thieme A. 1992. "Purposive Behavior and Cognitive Mapping: A Neural Network Model," *Biological Cybernetics*, 67, 165-174.
- Tolman EC. 1932. Purposive Behavior of Animals and Men. Irvington, New York.