

## A SIMPLIFIED HIPPOCAMPAL MODEL THAT LEARNS AND USES THREE KINDS OF CONTEXT

William B. Levy

Department of Neurological Surgery  
University of Virginia Health Sciences Center  
Charlottesville, Virginia 22908  
wbl@virginia.edu

### 1. ABSTRACT

Context plays a critical role in cognition. Previously Hirsh ('74), Kesner & Hardy (83), and Gray ('82) proposed that the hippocampus could learn context. Presented here is a simple hippocampal model that learns and uses three types of context: a context coming from the past, a context coming from the present, and a context concerned with the future. In all three of these situations, context is encoded by neurons that fire in a way analogous to hippocampal place cells. When these firing patterns do not appear, the network seems incapable of solving context dependent problems.

In psychology, the importance of context arises before the turn of the century (Boring, '50), in particular Titchner advocated a critical role for context in perception. Context is important to the networks that learn language in the schemes of Pollack ('90), Elman ('90), Jordan ('86), and Mozer ('92). Although not part of the usual terminology, context is at the heart of frames and schemas used by other cognitive psychologists. More to the point here, context learning seems to be part of hippocampal function (Hirsh, '74, Kesner & Hardy, '83, Gray, '82). Context learning is compatible with the Cohen and Eichenbaum theory of flexible memory (Cohen, '84; Eichenbaum et al., '92). Hirsh places the use of context at the center of proper encoding and recall of long-term memory. Context specifies the location of long-term memory storage. In this view, context is equivalent to episodic memory. Moreover, episodic memory associates disparate objects and events from single experiences; unfortunately, it is a lack of episodic memory that so hampers patients like H. M. and R. B. And, finally, even though not an explicit part of O'Keefe and Nadel's ('78) cognitive mapping theory (but see Nadel and Willner, '80), a coding that is analogous to hippocampal place cells (we call them context cells) — a coding that can be used to get from point A to point B — are context-based codes when viewed within the function of our model of the hippocampus.

To appreciate the role of context in memory, picture this one situation. You go to the hippocampal conference at Grand Cayman and for the first time you meet John Smith, a scientist from Seattle. A year or two later you visit the NIH and you see a vaguely familiar face; it's John Smith but you cannot remember his name. (As always striving for politeness as well as wishing to avoid embarrassment, you struggle to come up with a name to match the face.) If you can only remember the place, the circumstances, the episode where you met him, then you will have a chance of remembering the name. You well up a vague association of the conference room where you met and at the same time comes the hotel, the beach, and then ... "John, what a surprise seeing you here! How are you?"

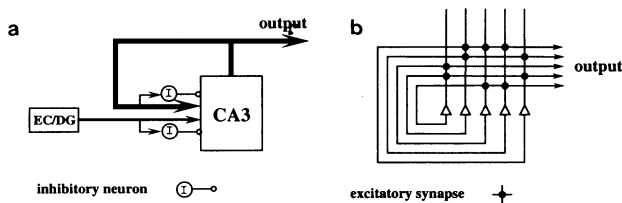
In other words, the storage of unique events is intimately associated with the surrounding circumstances (context). Of course, the idea of context-dependent memory is a couple of thousand years old as exemplified by the Roman's method of loci for memorizing long speeches. One sequence (the speech) is learned by associating it with another sequence of patterns (the sequence of statues you pass as you walk through a well-known museum) by using each successive statue and its locus as the context for successive words and phrases in the speech.

Less grandiose forms of context are useful in many other types of cognitive processes. Thus, many cortical regions would need to produce context codes. But, context-based codes do seem particularly important for hippocampal functions including setting up appropriately retrievable stores of memories.

## 2. THE SIMPLIFIED MODEL

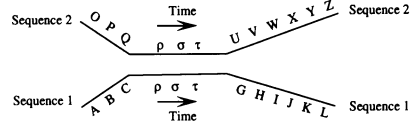
The problem of creating a code that appropriately reflects context can be seen as quite challenging when we consider the arbitrariness of patterns that might be associated - people with places or rhetoric on taxation to build some triremes with statues of Greek gods. Thus, it is pleasantly surprising that such a difficult problem can be solved (or at least begun to be solved) so simply. The extremely simplified, CA3-inspired network model which we have been studying (Levy, '89; Levy et al., '95; Levy & Wu, '96; Wu et al., '96) spontaneously and adaptively produces codes for context.

A generally accepted, gross hippocampal computational architecture (e.g. Levy, '89, Eichenbaum & Buckingham, '91; O'Reilly & McClelland, '94; Hasselmo & Schnell, '94) has three parts: 1) an input layer; 2) a recoder inspired by the hippocampal CA3 region, and 3) a decoder of the recoded signals inspired by the hippocampal CA1 region and its output targets. A CA3-like structure alone is all that is needed to create context codes, and to make this point, we do not include an explicit CA1 in our recent models (see Fig. 1a).



**Figure 1.** The Model. a. In the model the input layer is a combination of the entorhinal cortex and dentate gyrus. Accompanying this feedforward excitation is a proportional feedforward inhibition. The strong excitation of the network results from the recurrent connections, which is also accompanied by a feedback inhibition. The output of the network is the state of the excitatory CA3 cells themselves, and this is decoded by a simple cosine comparison. b. The recurrent excitatory synapses are sparse and randomly placed.

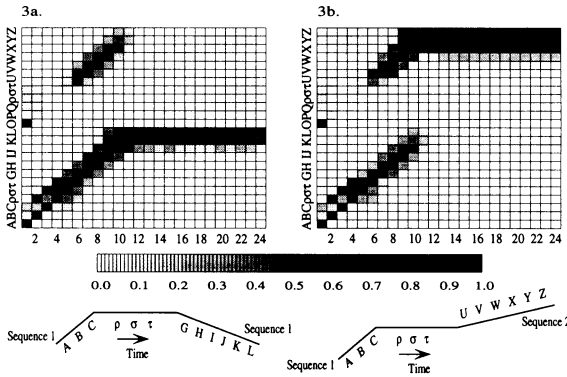
**Figure 2.** Pictorial representation of the environment. There are two sequences of 12 input patterns to learn. Each sequence contains three orthogonal segments. The two sequences share a subsequence of three patterns.



The specification of the CA3-like portion of the network has four essential aspects: (1) sparse recurrent excitatory connectivity (Fig. 1b) that produces more overall excitation than the external input; (2) a neuronal delay of at least one time step in converting an input to an output (i-o); (3) an associative modification rule that spans at least the i-o time step; and (4) some generic feedback inhibition that narrowly, but imperfectly compared to competitive networks, bounds total activity.

We have studied three abstract prediction problems to show that the network can learn and use three types of context: context coming from the past, context coming from the present, and context coming from the future. Context past is needed to solve sequence disambiguation (Minai et al., '94; Levy et al., '95; Wu et al., '96) (Fig. 2 and Fig. 3a). Context present is needed to solve the configural learning problem of transverse patterning (Levy et al., '96; Wu et al., '97). Context future is needed to solve goal finding problems (Levy et al., '95; Wu & Levy '96; Levy & Wu, '97) (Fig. 2 and Fig. 3b). In goal finding, a desired characteristic of the goal, but not its entire code, is part of the input. Note in Figure 3b how the network, when given a small fraction of the goal code, is able to overcome its natural tendencies (as shown by the sequence disambiguation result of Fig. 3a) and produce the appropriate path to the goal.

As opposed to simple sequence completion problems, these three context dependent problems require the network to construct cell firing patterns that we call local context firings. E.g., compare Fig. 1a and 1b of Wu et al. ('97) which shows two networks — the Fig.



**Figure 3.** Sequence prediction after learning the two partially overlapping sequences illustrated in Fig. 2. These similarity matrices use the cosine function. They compare the network states generated by each full pattern sequence after learning (ordinate) with the network states generated over time during testing (abscissa). Decode the output during testing by finding the darkest square in each column; the letter on the ordinate is the decoded answer. 3a. Sequence completion in the disambiguation problem is made difficult by the shared subsequence [ρστ]. Here we show the similarity values (used for decoding) when the network is transiently given pattern A. Appropriately enough for the learning and for the starting point, the states go to pattern L, a pattern essentially orthogonal to the representation of the other learned goal pattern, Pattern Z. 3b. When the same network is given the same transient input but two neurons of goal Z are also turned on, the network produces a sequence of representations leading to this partially specified goal. To create this path, the network must produce a novel sequence that appropriately combines its knowledge of the two separately learned sequences.

1b network learns the configural learning problem of transverse patterning, while the other does not. The repetitive cell firings of the recurrently driven neurons of Fig. 1b of Wu et al., ('97) are examples of local context neurons, the hypothesized analog of hippocampal place cells.

Finally, using as its inputs the compressed spontaneous replay of learned sequences coming out of the hippocampus (August & Levy, '96, '97), sequences of context codes could be learned by the cerebral cortex. These sequences — representing longer time spans — would be associated together by the cerebral cortex followed, recursively, by further compression in CA3 and association in cerebral cortex.

In summary, the CA3 computational model learns and uses three different kinds of context. The model reveals particular cell firing patterns that are, by hypothesis, the code for context. Further, we conjecture that this model is the appropriate neural basis of hippocampal theories that learn context and teach context to the cerebral cortex.

### 3. ACKNOWLEDGMENTS

This work was supported by NIH MH48161 and MH00622, by EPRI RP8030-08, and Pittsburgh Supercomputing Center Grant BNS950001 to WBL, and by the Department of Neurosurgery, Dr. John A. Jane, Chairman.

### 4. REFERENCES

- August, D. A. & Levy, W. B. (1996) Temporal sequence compression by a hippocampal network model. INNS World Congress on Neural Networks, 1299–1304.
- August, D. A. & Levy, W. B. (1997) Spontaneous replay of temporally compressed sequences by a hippocampal network model. *CNS'96*, this proceedings.
- Boring, E.G. (1950) A history of experimental psychology. 2nd edition. New York:Appleton Century Crofts.
- Cohen, N.J. (1984) Preserved learning capacity in amnesia: Evidence for multiple memory systems. In: *Neuropsychology of Memory* (L.R. Squire & N. Butters, Eds.), New York:Guilford Press, pp. 83–103.
- Eichenbaum, H. & Buckingham, J. (1991) Studies on hippocampal processing: Experiment, theory, and model. In: *Neurocomputation and Learning: Foundations of Adaptive Networks*. (M. Gabriel & J. Moore, Eds.). Cambridge: MIT Press, pp. 171–231.
- Eichenbaum, H., Otto, T. & Cohen, N.J. (1992) The hippocampus - what does it do? *Behav. Neural Biol.*, 57:2–36.
- Elman, J.L. (1990) Finding structure in time. *Cog. Sci.*, 14:179–211.
- Gray, J.A. (1982) *The neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system*. Oxford University Press:New York.
- Hasselmo, M.E. & Schnell, E. (1994) Laminar selectivity of the cholinergic suppression of synaptic transmission in rat hippocampal region CA1: computational modeling and brain slice physiology. *J. Neurosci.*, 14:3898–3914.
- Hirsh, R. (1974) The hippocampus and contextual retrieval of information from memory. *Behav. Biol.*, 12:421–444.
- Jordan, M.I. (1986) Attractor dynamics and parallelism in a connectionist sequential machine. In: *Proceedings of the Eighth Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Assoc. Inc., pp. 531–546.
- Kesner, R.P. & Hardy, J.D. (1983) Long-term memory for contextual attributes: Dissociation of amygdala and hippocampus. *Behav. Brain Res.* 8, 139–149.
- Levy, W.B. (1989) A computational approach to hippocampal function. In: *Computational Models of Learning in Simple Neural Systems*. (R.D. Hawkins & G.H. Bower, Eds.), New York: Academic Press, pp. 243–305.
- Levy, W. B & Wu, X. B. (1996) The relationship of local context codes to sequence length memory capacity. *Network* 7, 371–384.
- Levy, W. B & Wu, X. B. (1997) Predicting novel paths to goals by a simple, biologically inspired neural network. *CNS'96*, this proceedings.

- Levy, W. B., Wu, X. B. & Baxter R. A. (1995) Unification of hippocampal function via computational/encoding considerations. In: Proceedings of the Third Workshop: Neural Networks: from Biology to High Energy Physics. *International J. of Neural Sys.* 6 (Supp.), 71–80.
- Levy, W.B., Wu, X. B. & Tyrcha, J.M. (1996) Solving the transverse patterning problem by learning context present: A special role for input codes. INNS World Congress on Neural Networks, 1305–1309.
- Minai, A. A., Barrows, G. L., & Levy, W. B (1994) Disambiguation of pattern sequences with recurrent networks. INNS World Congress on Neural Networks, IV-176–181.
- Mozer, M.C. (1992) Induction of multiscale temporal structure. In: *Advances in Neural Information Processing Systems*, 4. (J.E. Moody, S.J. Hanson & R.P. Lippmann, Eds.), San Mateo, CA: Morgan Kaufman, pp. 275–282.
- Nadel, L. & Willner, J. (1980) Context and conditioning: A place for space. *Physiol. Psychol.* 8, 218–228
- O’Keefe, J & Nadel, L. (1978) *The Hippocampus as a Cognitive Map*. Oxford: Oxford Univ. Press.
- O’Reilly, R.C. & McClelland, J.L. (1994) Hippocampal conjunctive encoding, storage, and recall: avoiding a tradeoff. *Hippocampus* 4, 661–682.
- Pollack, J.B. (1990) Recursive distributed representations. *Artificial Intelligence* 46:77–105.
- Wu, X. B., Baxter, R. A. & Levy, W. B (1996) Context codes and the effect of noisy learning on a simplified hippocampal CA3 model. *Biol. Cybern.* 74, 159–165.
- Wu, X. B. & Levy, W. B (1996) Goal finding in a simple, biologically inspired neural network. INNS World Congress on Neural Networks, 1279–1282.
- Wu, X. B., Tyrcha, J. M. & Levy, W. B (1997) A special role for input codes in solving the transverse patterning problem. *CNS’96*, this proceedings.