IEEE International Conference on Neural Networks, 1997, I-368-371.

A Simple, Biologically Motivated Neural Network Solves the Transitive Inference Problem

William B Levy and Xiangbao Wu
Department of Neurological Surgery
University of Virginia Health Sciences Center
Charlottesville, Virginia 22908, USA
E-mail: wbl@virginia.edu, xw3f@virginia.edu

Abstract

Configural learning problems can be resolved by both rats and humans if they are not too difficult (e.g. Alvarado and Rudy 1992, 1995). The configural learning problem which we explore here is transitive inference. Transitive inference (learn the four pairs A>B, B>C, C>D, D>E, then test with the novel pair B?D) was once viewed as a logical problem. However, it is now acknowledged that when the stimuli are appropriate even three year old humans can solve this problem and, as well, so can pigeons and rats. Thus, even though the problem is a simple exercise in logic, there is reason to suspect that mammals, or for that matter, neural networks will solve such a problem without recourse to any explicit syllogistic reasoning. In fact, by casting the input stimuli in a form appropriate for a sequence learning neural network, a hippocampal-like network can solve the transitive inference problem. Furthermore, performance is appropriately disrupted by turning the linear sequence of relationships into a nonlinear (circular) relationship.

1. Introduction

Configural learning problems are characterized by the need to separate and possibly interact substimuli (atomic stimuli) of a set of stimuli complexes (e.g., A is an atomic stimulus of the complex AB). Because such tasks require subjects to acquire nonspatial relationships between stimuli and because an intact hippocampus is required for such problems, configural learning is regarded as a hippocampally dependent function (Dusek and Eichenbaum 1996) that reflects something apart from the cognitive mapping theory of O'Keefe and Nadel's (1978) of hippocampal function. Because we are studying a hippocampal-like neural network model, it is natural to challenge the network with such problems.

2. The Problem of Transitive Inference

Transitive inference (Fig. 1a) is a classical problem in cognitive psychology and was once considered a suitable problem for judging the development of logical abilities in humans (see references in Siemann and Delius 1994; Wynne 1996). Transitive inference is based on a relationship such as "weighs more than;" the problem is to infer from A weighing more than B and B weighing more than C, that A weighs more than C. Such problems can be solved by logical inference. However, one should not assume this is how the problem is solved because rats, pigeons, and three year old humans can solve such problems if the stimuli are in an appropriate form (Wynne 1996). Thus, we consider the minimal neural-like network that might solve such problems. Psychologists now use at least 4 pairs of 5 atomic items (i.e. distinct patterns) as shown in Fig. 1a. After learning which atomic stimulus (A, B, C, D) is rewarded in each of these four pairs, the subject gets the novel BD combination and must choose one of the these two atomic stimuli. Of course, B is the right answer. (One also tests other pairs, but combinations involving an end element (A or E) are quite easy because A has always been correct and E has always been the incorrect choice.)

There is no single accepted control experiment for transitive inference. However, we would advocate the example shown in Fig. 1b. Note that we have discarded the atomic stimulus E, and in the fourth stimulus pair replaced it by the atomic stimulus, A. To understand the difference between the two learning sets, think of the original problem (i.e. transitive inference, Fig. 1a) as a row of five individuals such that A is to the left of B, B is to the left of C, and so on, while in the control example (Fig. 1b) think of individuals A, B, C, and D as sitting at a round table where we are informed who is to the immediate left of whom; then, for each problem, we are asked whether B is to the left or the right of D. Of course for our control, Fig. 1b, there is no right answer.

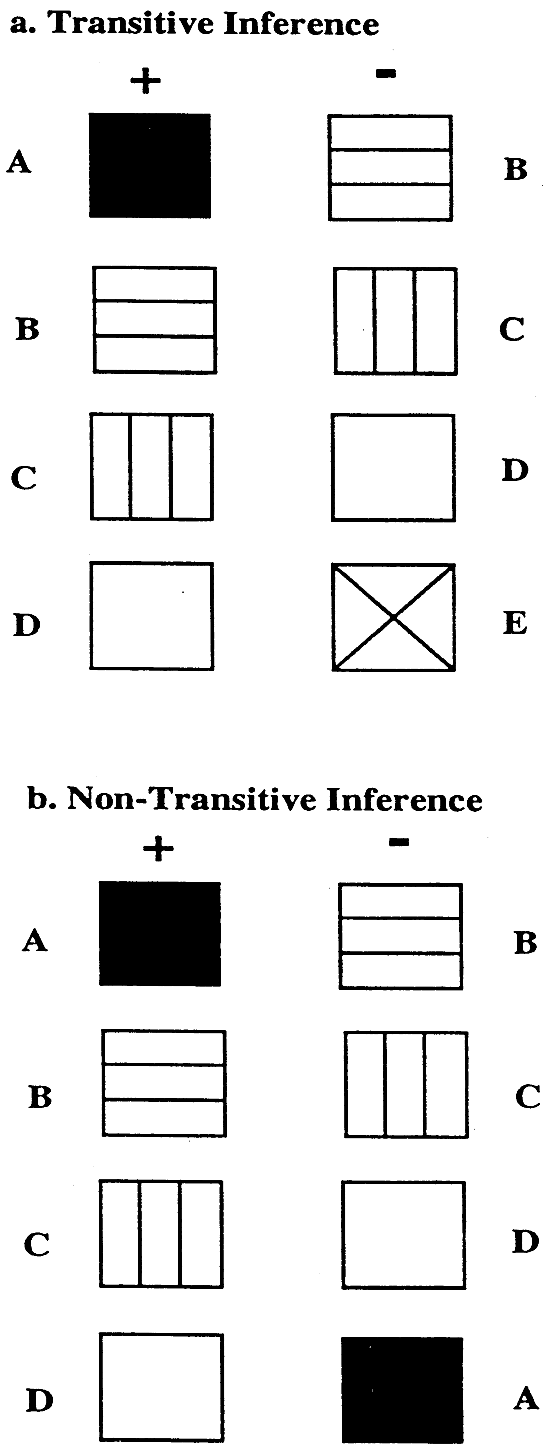

Figure 1. The two problem sets showing the input pairs (e.g., A+B-). Correct and incorrect stimuli are represented as + and - respectively. In both transitive inference (a.) and non-transitive inference (b.), the test input is the novel pair BD. For transitive inference B is the right answer, but there is no logically correct answer for the BD pair in the non-transitive inference situation.

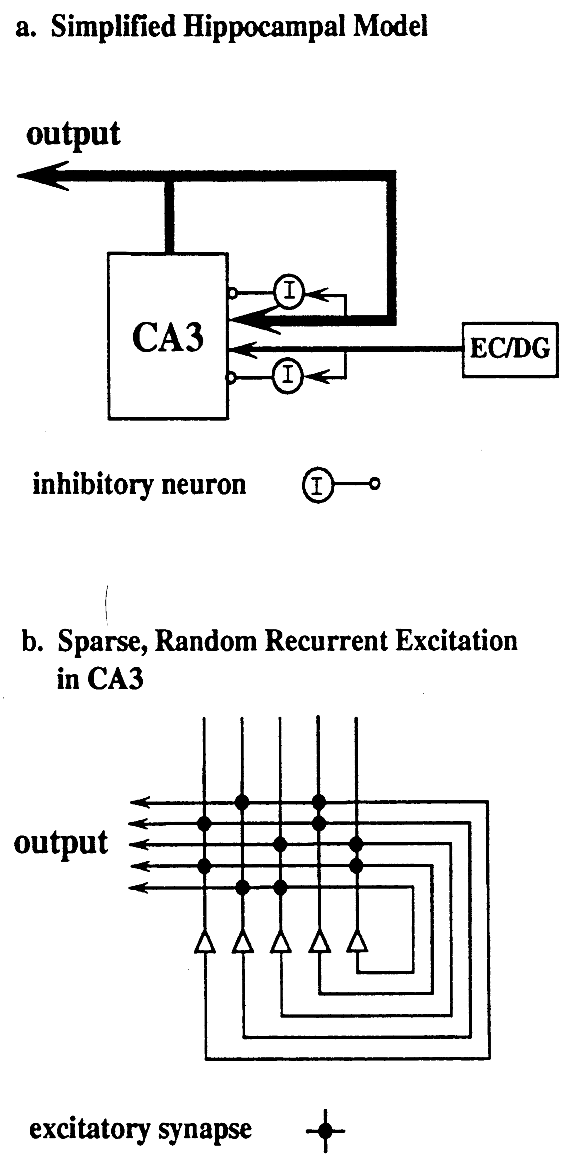

Figure 2. The Model. a. In the model the input layer is a combination of the entorhinal cortex and dentate gyrus. Accompanying this feedforward excitation is a proportional feedforward inhibition. The strong excitation of the network results from the recurrent connections, which is also accompanied by 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.

3. The Network

The hippocampal model is essentially a model of region CA3 (Fig. 2). The input layer corresponds to a combination of the entorhinal cortex and dentate gyrus. To make the system's operation as transparent as possible, decoding is performed by similarity comparisons rather than a CA1-subiculum-entorhinal decoding system. The CA3 model is a sparsely (10%) interconnected feedback network of 512 neurons where all direct connections are excitatory and the network elements are McCulloch-Pitts neurons. There is an interneuron mediating feedforward inhibition, and one mediating feedback inhibition. Inhibition is of the divisive form, but the system is not purely competitive because of a slight delay. Synaptic modification develops over training. The process controlling synaptic modification is a local, self-adaptive postsynaptic rule that includes both potentiation and depression aspects (Levy and Steward 1979; Levy 1982). The network computations are all local and are contained in three equations: spatial summation adjusted by inhibition; threshold to fire or not; and local Hebbian synaptic modification (see Levy and Wu 1996, Wu et al., 1996 for details).

4. Results

In studying transitive inference we use the same network that learned transverse patterning (Levy et al., 1996), and we construct a similar set of input codings. Thus, the inputs are sequences of stimulus, response, and reinforcement. In this case, however, a staged learning procedure is employed, as in the experiments of Dusek and Eichenbaum (1996). That is, there are five phases of training. Phase 1 consisted of the presentation of 10 trials of each pair in serial (i.e., ten trials of A+B- followed by ten trials of B+C-, then ten trials of C+D- and ten trials of D+E-). While in Phases 2, 3 and 4, trials were presented in blocks of 5, 3 and 1 trial respectively. The last phase of training is just random pair presentation of the four premise pairs.

To test the reconfiguration capabilities of an animal or a network model, novel pairings are presented. In particular, BD and AE are represented (Dusek and Eichenbaum 1996). Fig. 3 shows a comparison between experimental results of Dusek and Eichenbaum (1996) and our network simulation results. The BD comparison rather than the AE comparison is critical here because the AE comparison is not affected by the hippocampal inactivating lesions (Dusek and Eichenbaum 1996). From Fig. 3, one can see that B is the typically chosen answer (percent correct about 80%) for the BD trial (which is not presented during learning). Likewise, A is the typical answer for the AE trial.

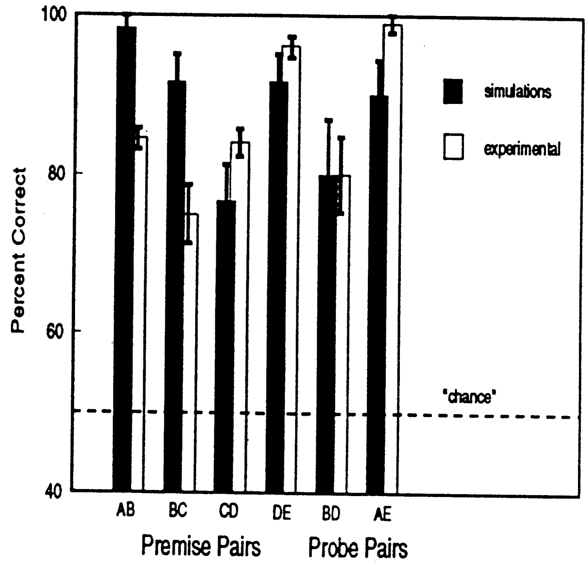

Figure 3. A comparison between experimental results and simulations of the transitive inference problem. The open rectangles are the experimental results of Dusek and Eichenbaum (1996). The filled rectangles are the simulation results. Note that, as in the rat experiment, the model performs significantly above chance on the BD comparison. Error bars represent the standard error of the mean over 30 simulations. Dashed lines represent chance performance level.

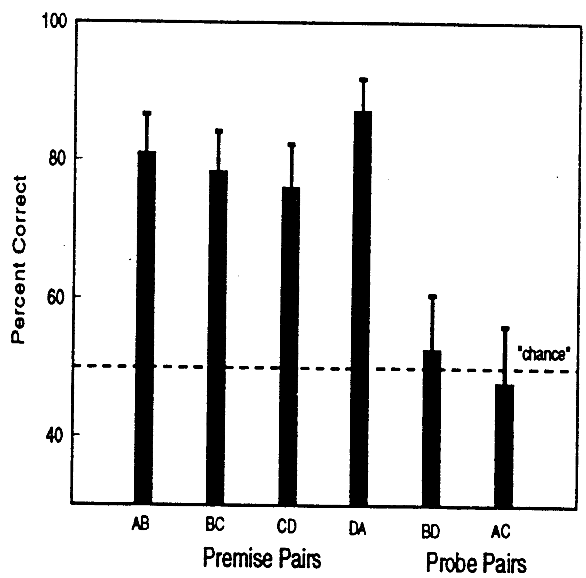

Figure 4. Simulations of the non-transitive inference problem. Note that, the network

randomly chooses B or D without noticeable preference on the BD comparison. Error bars represent the standard error of the mean over 30 simulations. Dashed lines represent chance performance level.

Finally, turning the transitive inference problem into a nonlinear problem (Fig. 1b), which is essentially unsolvable, leads to the network failing to make consistent responses (as it should fail to do) on the BD comparison. That is, the learning set now consists of AB, BC, CD, and DA, where in the (DA) pairing, D is the correct answer (+) and A is the incorrect answer (-). In this case the network's responses seem logical to the BD test pairing. Specifically, the network randomly chooses B or D without any noticeable preferences (percent correct about 50%, see Fig. 4). Likewise, there is no preference over A and C (Fig. 4). (Randomization derives from the random initialization activity, an initialized procedure used throughout all simulations.) As yet there are no experimental results for comparison.

5. Discussion

By turning configural problems into sequence prediction/goal finding problems (Levy et al. 1995, Wu and Levy 1996), we have shown that a sequence learning neural network model of hippocampal region CA3 has configural learning abilities similar to humans and rats. Such results strengthen the validity of this model as the essence of hippocampal computations. Additional results indicate the critical nature of the codes the network constructs. Indeed, just by looking at the learned firing patterns, we can predict if the network will succeed or fail when tested. Specifically, when BD evokes enough ($\sim$ 12%) C neuronal firings, the B answer is selected. Finally, the fact that such a minimal network model appears to have logical abilities might encourage psychologists to consider neural network models of human cognition in place of higher level models.

6. 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.

References

[1] M. C. Alvarado and J. W. Rudy. Some properties of configural learning: An investigation of the transverse-patterning problem, *J. Exp. Psychol.:Animal Behav. Processes* 18:145-153, 1992.

[2] M. C. Alvarado and J. W. Rudy. Rats with damage to the hippocampal-formation are impaired on the transverse-patterning problem but not on elemental discriminations, *Behav. Neurosci.* 109:204-211, 1995.

[3] J. A. Dusek and H. Eichenbaum. The hippocampus and memory for orderly stimulus relations, submitted.

[4] W. B Levy. Associative encoding at synapses, *Proceedings of the Fourth Annual Conference of Cognitive Science Society*, pp.135-136, 1982.

[5] W. B Levy and O. Steward. Synapses as associative memory elements in the hippocampal formation, *Brain Res.* 175:233-245, 1979.

[6] W. B Levy and X. B. Wu. The relationship of local context codes to sequence length memory capacity, *Network* 7:371-384, 1996.

[7] W. B Levy, X. B. Wu, and R. A. Baxter. Unification of hippocampal function via computational/encoding considerations. In: *Proceedings of the Third Workshop: Neural Networks: from Biology to High Energy Physics. Int. J. Neural Sys.* 6(Supp.):71-80, 1995.

[8] W. B Levy, X. B. Wu, and J. M. Tyrcha. Solving the transverse patterning problem by learning context present: A special role for input codes, *INNS World Congress on Neural Networks*, pp.1305-1309, 1996.

[9] J. O'Keefe and L. Nadel. The hippocampus as a cognitive map, *Oxford University Press, Oxford U.K.*, 1978.

[10] M. Siemann and J. D. Delius. Processing of hierarchic stimulus structures has advantages in humans and animals, *Biol. Cybern.* 71:531-536, 1994.

[11] X. B. Wu and W. B Levy. Goal finding in a simple, biologically inspired neural network, *INNS World Congress on Neural Networks*, pp.1279-1282, 1996.

[12] X. B. Wu, R. A. Baxter, and W. B Levy. Context codes and the effect of noisy learning on a simplified hippocampal CA3 model, *Biol. Cybern.* 74:159-165, 1996.

[13] C. D. L. Wynne. A minimal model of transitive inference, In: *Models for Action* (C. D. L Wynne and J. E. R. Staddon, eds.) Lawrence Erlbaum Assoc., in press.

The 1997 IEEE International Conference on Neural Networks

June 9 - 12, 1997
Westin Galleria Hotel,
Houston, Texas, USA

The 1997 IEEE International Conference on Neural Networks Proceedings

IEEE Catalog Number: 97CH36109
ISBN Softbound: 0-7803-4122-8
ISBN Casebound: 0-7803-4123-6
ISBN Microfiche: 0-7803-4124-4
Library of Congress: 97-72856

Additional Proceedings may be ordered from:

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
(800) 678-IEEE