THE NEURODYNAMICS OF CONTEXT REVERBERATION LEARNING

Kevin G. Kirby and Nancy Day

Department of Computer Science and Engineering
Wright State University, Dayton OH 45435

ABSTRACT

The associative learning of dynamical system trajectories requires construction of internal representations of input histories in a way that permits the set of such representations to be adaptively classified into states. In this paper we discuss two architectures for this type of learning; one connectionist, and one based on the continuous dynamics of intra-neuronal reaction and diffusion. Our simulations show that for learning a discrete-time sequence association, replacing the total connectivity in the connectionist network with nearest-neighbor connectivity increases efficiency by an order of magnitude. This suggests that the locality exhibited in molecular electronic architectures may yield a powerful learning device.

INTRODUCTION

Artificial neural network architectures have traditionally relied on massive connectivity to implement the learning of spatial and temporal associations. Some technologies, such as optical computing, show much promise for implementing the large number of connections required. On the other hand, if we are interested in harnessing molecular electronic systems for some kind of learning architecture, we may have elaborate dynamics at our disposal but are nevertheless forced to deal with mostly local interactions.

We will focus on a special case of spatio-temporal learning in which a system must learn associations between discrete-time signals. Unlike a static association, the output of the system at any given time not only depends on the current input, but on the history of inputs, One connectionist approach to this problem was attempted by Gallant [4]. Figure 1 shows the Gallant architecture for a simple association task in which a four-bit pattern is clamped onto the input units for four time steps, and must produce a prespecified binary sequence from the output unit during that interval. The network must adjust itself in accordance with an external error signal to learn an entire set of associations. This is done by applying a perceptron-like algorithm to the connection weights coming into the output unit. Performance of the learning system is measured in terms of the number of iterations of the output weight-updating algorithm required before all associations are learned perfectly. All other connection weights are fixed and random. The random connections are afferent to units in a large totally connected internal subnetwork. This means that inputs initiate a random reverberation of signals around the subnetwork. Randomness is indeed not a handicap, since (at least in the limit of large numbers of neurons in the subnetwork) the reverberation pattern will be different enough for each input history to allow the association algorithm at the output layer to adapt the weights based on correlations between signals coming in from the input lines, and signals coming in from the subnetwork. Thus we can say that the totally connected subnetwork is a repository of "context", which is accessed by a basically static learning algorithm to learn temporal associations. For this reason we label it a *context reverberation (CR) network.* For the 4-input, 4-step association, Gallant has needed on the order of 50 CR neurons, which means $40^2 = 1600$ internal connections. We believed that the random context reverberation function of the subnetwork might be served by non-connectionist molecular electronic systems, so we set out to find how essential the global connectivity really was.

LOCAL CONTEXT REVERBERATION DYNAMICS

A simulation system was set up in which we could control the size and connectivity level of the CR-subnetwork and monitor its learning times on a variety of sequence association sets. Figure 2 shows a comparison of three types of connectivity. Total connectivity yields the shortest learning times for a given number of CR neurons. Reducing connectivity to 7 nearest neighbors slows down learning somewhat, and reducing to 4 nearest neighbors (as if the CR-network was a rectangular grid) yields even slower learning. However, a tremendous number of connections have been eliminated. Thus, if we desire to have the associations learned in about 120 iterations, we can either use a 35-neuron totally connected net (with $35^2 = 1225$ connections), or a 55-neuron four-neighbor net (with only $4 \times 55 = 220$ connections).

In the above experiments, we averaged together the results of many different weight randomizations. Some were better than others, as the confidence levels of Figure 2 indicate. There was a correlation between the learning performance of a given set of CR-unit weights and the periodicity of the trajectory of the CR-units when a constant input was clamped on the unit (Figure 3). This suggests any externally forced aperiodic or long-period dynamical system would be a candidate for a context-reverberation system. There have been useful results established about periodicities of local random networks of boolean units [5] and threshold units [4], which would have direct relevance to the design of discrete CR-systems.

IMPLICATIONS AND CONCLUSIONS

We call the mechanism of this system *scrambled context*, because the dynamics in a sense "encrypts" the input histories while the learning algorithm "decrypts" them. (One finds networks using non-scrambled context in the connectionist robotics work of Jordan [3]). One continuous dynamical system with local interactions that might function as a CR subsystem would be the reaction-diffusion neuron [7,8]. Envisaged as an element of a selectional brain model [2], it can be seen as a model of intra-neuronal cyclic nucleotide dynamics [6], and as a model of continuous local computation based on the morphogenesis model of Turing [10]. Figure 4 sketches the dynamics of this system, in which a membrane can function as a repository of input history. A classical instantiation might be found in neural field dynamics [1], which if harnessed for a CR-system would employ *topographic* context. In this case the input history would set up topographic structures (such as stripes) to be used as inputs to association algorithms.

Scientifically, such systems motivate the re-introduction of dynamical systems theory into computability theory. Technologically, such systems motivate the study of molecular electronics for special-purpose components of artificial intelligence architectures.

CH2936-3/90/0000-1781 $01.00

REFERENCES

[1] S.-I. Amari, "Dynamics of Pattern Formation in Lateral-Inhibition Type Neural Fields", *Biol. Cybernetics*, Vol. 27, pp. 77-87: 1977.

[2] M. Conrad, "Evolutionary Learning Circuits", *J. Theor. Biol.* Vol. 46, pp. 167-188: 1974.

[3] M.I. Jordan, "Generic Constraints on Underspecified Target Trajectories", *Proc. IJCNN*, Vol I, pp. 217-225: 1989.

[4] S.I. Gallant and D.J. King, "Experiments with Sequential Associative Memories", *Cognitive Science Society Conference*, Montreal: 1988.

[5] Kauffman, S.A. "Principles of Adaptation in Complex Systems", In *Lectures in the Sciences of Complexity*, D. Stein, Ed., Addison-Wesley: 1989.

[6] K.G. Kirby and M. Conrad. "The Enzymatic Neuron as a Reaction-Diffusion Network of Cyclic Nucleotides", *Bull. Math. Biol.*, Vol. 46, pp. 765-783: 1984.

[7] K.G. Kirby and M. Conrad. "Intraneuronal Dynamics as a Substrate for Evolutionary Learning", *Physica D*, Vol. 22, pp. 205-215: 1986.

[8] K.G. Kirby, "Information Processing in the Lorenz-Turing Neuron", *Proc. IEEE EMBS*, Vol. 11, pp. 1358-1359: 1989.

[9] K.E. Kürten, "Dynamical Properties of Threshold Automata with Nearest-Neighbor Interactions on a Regular Lattice", *Proc IEEE ICNN*, Vol. I, pp. 37-43: 1988.

[10] A.M. Turing, "The Chemical Basis of Morphogenesis", *Phil. Trans. Royal Soc.*, Vol. B237, pp. 37-72: 1952.

This research was supported by a Research Incentive Award from the Wright State University Research Council. The authors may be contacted at the Department of Computer Science and Engineering, Wright State University, Dayton, OH 45435; (513) 259-1373.

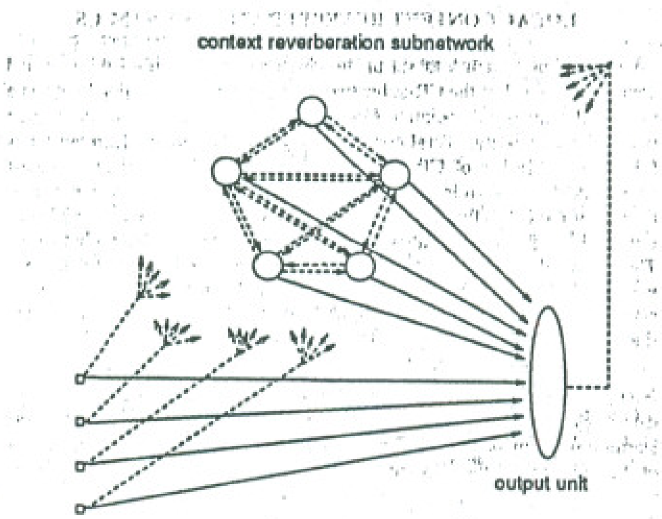


Figure 1. *A connectionist architecture for simple sequential learning. This one-output net was used by Gallant and King [4] for learning* spatial → temporal *associations. The input layer was clamped on a fixed pattern, and the output unit was to execute a pre-specified trajectory. Although only a five-unit "context-reverberation subnetwork" is depicted here, for four-bit input patterns 35 to 60 units were typically required. Dashed lines denote fixed random connections.*

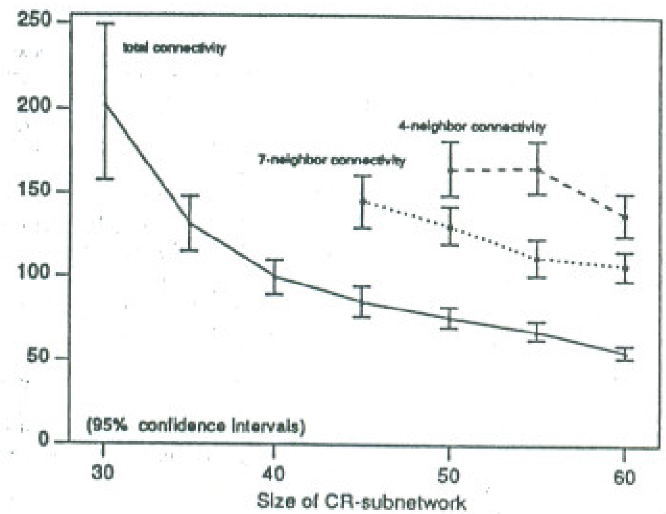


Figure 2. *The effect of reducing connectivity in the context-reverberation subnetwork. This data was obtained from 80 experiments per parameter choice, for four different training sets of eight 4-bit → 4-step associations. Although reducing connectivity increases the number of CR-units needed to learn the task in the same amount of time, the total number of connections required drops dramatically. This figure shows that it is nearly an order of magnitude more efficient to use local processing in the CR-subnetwork.*

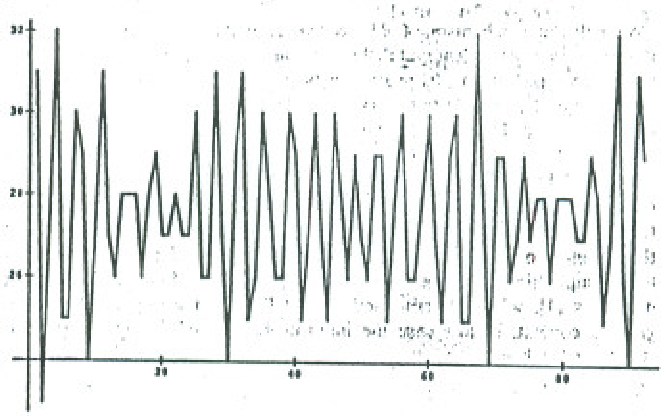
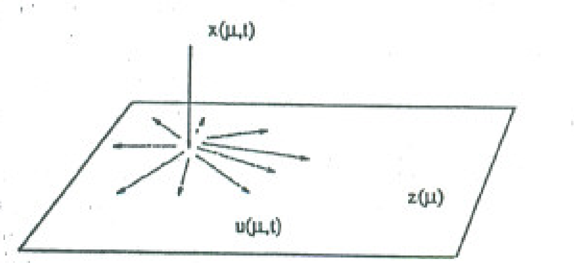


Figure 3. *Long periods are required for useful context-reverberation dynamics. Here is the firing intensity (number of active units) versus time for a 50-unit 4-connected subnetwork, when a 4-bit pattern is clamped on the inputs. This network was harvested from the best performing nets in the experiments of Figure 2. The trajectory begins with a transient of length 7, and then becomes oscillatory with period 84; the plot shows time steps 1-92. The dynamics is a superposition of many short-period ensembles.*



$$u = Ix - hy + \int \nabla \cdot D \nabla u + R(u)\, dt$$

$$y = \alpha \left[\int uz\, d^2\mu - \theta \right]$$

Figure 4. *The dynamics of the reaction-diffusion neuron. Input patterns arrive continuously onto the membrane, and stimulate an excitation signal* u, *which is governed by reaction-diffusion dynamics. A discontinuous process which converts patterns to discrete outputs is superimposed on the diffusion and mediated by the "excitase" distribution z. The results shown in the previous figures suggest that the local, continuous, nonlinear dynamics of such a neuron may enable it to function as an entire CR-subnetwork.*