Spatio-Temporal Sequence Processing with the Counterpropagation Neural Network

Lipo Wang

School of Computing and Mathematics
Deakin University
662 Blackburn Road
Clayton, Victoria 3168 Australia
Email: lwang@deakin.edu.au
URL: http://www.cm.deakin.edu.au/~lwang

ABSTRACT

We present a system that is capable of learning and retrieving spatio-temporal sequences using the forward-only counterpropagation neural network (FOCPNN). This artificial neural system has comparator units, a parallel array of FOCPNNs, and delayed feedback lines from the output of the system to the FOCPNN layer. The system has separate Conditioned Stimulus (CS) input channel and Unconditioned Stimulus (US) input channel, which is analogous to classical conditioning. During learning, pairs of sequences of spatial patterns are presented to the CS and the US input channels simultaneously and the system learns to associate patterns at successive times in sequence. During retrieving, an imperfect cue sequence, which may be obscured by spatial noise and temporal gaps, causes the system to output the stored spatio-temporal sequence. Compared with other existing temporal systems, this system shows computational advantages such as fast and accurate learning and retrieving, and ability to store a large number of complex sequences consisting of non-orthogonal spatial patterns.

1. INTRODUCTION

Recently there has been extensive research activity in temporal phenomena in neural networks (see, e.g., [1]-[17]). Grossberg's Avalanche model [2][3] has been mathematically proven to be capable of learning spatio-temporal sequences after an infinite number of presentations of the training sequences. The model uses delayed partial-differential-difference equations and is thus computationally expensive. Fukushima' spatio-temporal system [4] requires many iterations for sequence retrieval and non-orthogonal patterns retrieved by this system are often obscured by noise (spurious memories). Several authors have used time delays in Hopfield networks [5]-[7] to generate spatio-temporal sequences. These systems have rather low memory capacity. Guyon et al [8] proposed a temporal network that required *a priori* analytical expressions of all stored sequences. Time delays have also been incorporated into back-propagation networks to form recurrent networks [9], though back-propagation networks are known to have long learning times. Buhmann and Schulten [10] used noise to induce transitions between attractor patterns in Hopfield networks, thereby forming spatio-temporal sequences. There are other mechanisms for temporal processing, including time-dependent [11], asymmetric [12][13], and diluted higher order [14]-[16] weight matrices.

Despite the progress in spatio-temporal sequence processing with neural networks, difficulties such as slow and inaccurate learning and retrieving, strict orthogonality requirements, and limited memory capacity, need to be overcome. We present a system for processing spatio-temporal sequences using a dynamically-generated variant of the counterpropagation network [17]. We show that this approach

0-7803-3280-6/96/$5.00 ©1996 IEEE

eliminates the problems of existing temporal systems, thereby markedly improves the efficiency of spatio-temporal sequence processing [18].

2. SYSTEM SPECIFICATION AND PERFORMANCE

We first review the architecture and the learning algorithm of the forward-only counterpropagation neural network (FOCPNN) invented by Hecht-Nielsen [17]. The FOCPNN has three layers, i.e., an input layer, a competitive layer, and an output layer (Fig.1). When a CS input pattern $\vec{x}$ presented to the input layer during training, the competitive layer performs competitive learning and the neuron whose synaptic weights are most similar to the input pattern adjusts its weight vector according to

$$\vec{w}_i^{\text{new}} = (1-\alpha(t))\vec{w}_i^{\text{old}} + \alpha(t)\vec{x} \, . \tag{1}$$

The weights of the other neurons in the competitive layer remain unchanged. The learning rate α is a function of time [17]: it starts out with a high value and gradually decreases towards zero. All output weights connected to this winning neuron are modified towards the associated US training pattern (the "correct" output pattern) $\vec{y}$ in a similar fashion:

$$\vec{u}_i^{\text{new}} = (1-\alpha'(t))\vec{u}_i^{\text{old}} + \alpha'(t)\vec{y} \, . \tag{2}$$

All other weights in the output layer remain unchanged. Thus after training, the weights of each neuron in the competitive layer represent a cluster in the CS training patterns and the output weights connected to this neuron represent the associated US training pattern.

If a testing pattern $\vec{x}_k$ is presented to the FOCPNN after training, the neuron in the competitive layer with weights most similar to the testing pattern wins the competition and broadcasts the associated US pattern $\vec{y}_k$ to the output neurons.

We will make some small changes to the FOCPNN before using it in our system. In our variant of the FOCPNN, the competitive layer and all weights of the network are dynamically generated according to the competitive learning algorithm used in the ART network [19]. Compared to the original FOCPNN, this dynamically generated FOCPNN is more efficient to implement with software [20] and its storage capacity does not need to be pre-specified.

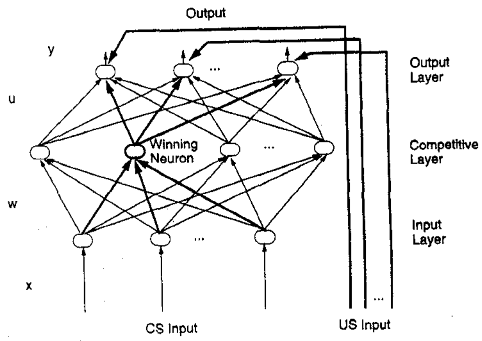

Fig.1 The forward-only counterpropagation neural network.

At the beginning of training, this FOCPNN variant has the desired numbers of input and output neurons corresponding to the dimensions of the CS and US patterns, but has no competitive neurons or weights. When the first pair of CS and US training patterns ($\vec{x}_1$ and $\vec{y}_1$) is presented to the network, the first competitive neuron is generated, with its weights being the CS training pattern ($\vec{w}_1 = \vec{x}_1$) and the weights connecting this neuron to all output neurons being the US training pattern ($\vec{u}_1 = \vec{y}_1$). For each subsequent pair of CS-US association, a competition is carried out in the competitive layer and the neuron whose weights are the most similar to the CS training pattern is found. If the similarity between the CS training pattern and the weight vector of this winning neuron is above a vigilance threshold, all incoming ($\vec{w}$) and outgoing ($\vec{u}$) weights connected to this neuron are modified according to eqs.(1) and (2). Otherwise, a new competitive neuron is generated in the same way in which the first competitive neuron is generated using the training vectors. Furthermore, we choose [21]

$$\alpha(t) = \alpha'(t) = 1/\tau_i \, , \tag{3}$$

where $(\tau_i - 1)$ is the number of times that the weights of neuron i has already won competition, so that the incoming and outgoing weights connected to neuron i are exactly the overall averages of the CS and US training patterns used to modify the weights of this neuron [21]:

$$\vec{w}_i(\tau_i) = \frac{1}{\tau_i}\sum_{\tau'=1}^{\tau_i}\vec{x}(\tau') , \qquad (4)$$

and

$$\vec{u}_i(\tau_i) = \frac{1}{\tau_i}\sum_{\tau'=1}^{\tau_i}\vec{y}(\tau') . \qquad (5)$$

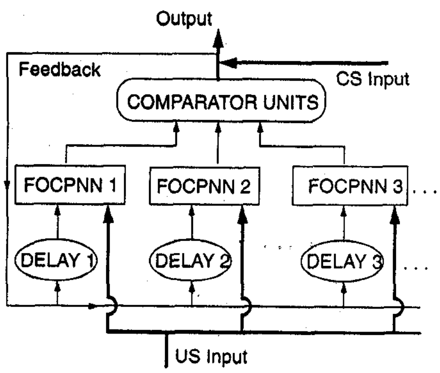

Fig.2 The present general system for spatio-temporal processing.

The present system for spatio-temporal processing is shown in Fig.2. The system consists of N comparator units, a parallel array of N_L FOCPNNs, and time delays that feed the overall output of the system back to the neural network layer. There are N output neurons and N input neurons in each FOCPNN. The time delay associated with the l-th FOCPNN delays the signal by l time steps with respect to the current time, where $l = 1, 2, ..., N_L$. There are two separate input channels: the Conditioned Stimulus (CS) and the Unconditioned Stimulus (US) channels, in analogy with classical conditioning: after repeated presentations of a US together with a CS, the CS alone can generate the response caused by the US.

During learning, pairs of sequences of spatial patterns are presented to the CS and the US input channels of the system simultaneously, one spatial pattern at each time step. The system outputs external signals during learning. Memory is achieved through the hetero-associations of the FOCPNNs.

During retrieving after learning, a small piece of a stored sequence, which may or may not be obscured by spatial noise and/or temporal gaps, is presented to the system through the CS input channel only. The US input channel of the system is not used during retrieving. A conflict threshold θ_c is assigned to the comparator units to allow for an overall "don't know" answer: if a fraction larger than θ_c of the comparator units outputs conflicting answers, the system outputs an overall "don't know" answer. Otherwise each comparator unit carries out a weighted average over the outputs of the corresponding neurons in all the FOCPNNs that carry signals, i.e.,

$$S_i(t) = \phi\left[\frac{\sum' a_l O_{il}(t)}{\sum' a_l}\right] , \; i = 1, 2, ..., N, \qquad (6)$$

where $S_i(t)$ is the output of the i-th comparator unit of the system, $O_{il}(t)$ is the state of the i-th output neuron in the l-th FOCPNN, $\sum'$ means a sum over only signal-carrying FOCPNNs, and the function $\phi(x)$ rounds up x to the nearest gray shade value.

In the present implementation, we choose both dimensions of the CS and US spatial patterns, as well as the numbers of neurons in the input and output layers in the FOCPNNs, to be 11×11. Three such FOCPNNs are used ($N_L = 3$). We choose the conflict threshold for the comparator units to be 1/3 and $a_1 = a_2 = a_3$ in eq.(1). The system is trained to store twenty sequences, three of which are shown in Fig.3, and is then tested in situations shown in Fig.4. Each training sequence is presented only once to the system. Note that in Fig.3 pattern I, which is the same as pattern 1, appears in both sequences (a) and (b), and pattern J appears more than once within sequence (c), which is sometimes called a higher-order or complex sequence. Fukushima [4] used sequences (a) and (b) in training and testing his spatio-temporal processing system.

In Fig.4(a), all retrieved images in sequence (a) are noise-free, whereas some retrieved images, i.e., E and F, are obscured by noise for Fukushima's system. The ability of outputting a "don't know" answer often can significantly reduce error rate in practical applications (see Fig.4(b)), whereas Fukushima's system outputs meaningless sequences in this kind of situation. When pattern D is presented to Fukushima's system, the retrieval of sequence (a) is very difficult: it takes many iterations and many retrieved images are imperfect. In comparison, our system retrieves sequence (a) accurately and quickly (Fig.4(d)). Fukushima's system has not been tested with an unknown sequence (Fig.4(e)), but can be expected to yield meaningless output since it is unable to give a "don't know" answer. The cases with temporal gaps (Fig.4(f)) and complex sequences (Fig.4(g)) have not been tested in Fukushima's system.

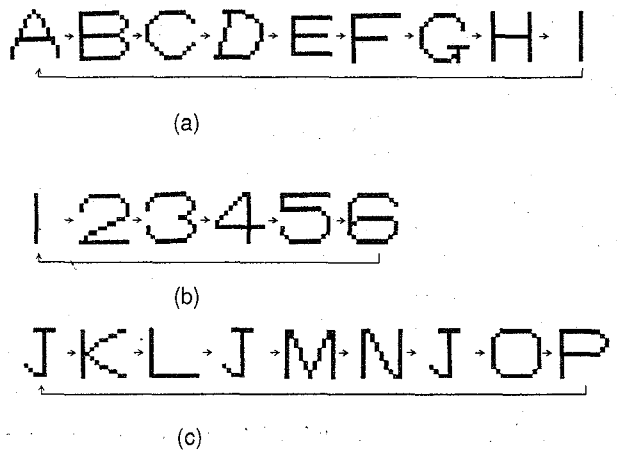

Fig.3 Three examples of spatio-temporal sequences used in learning.

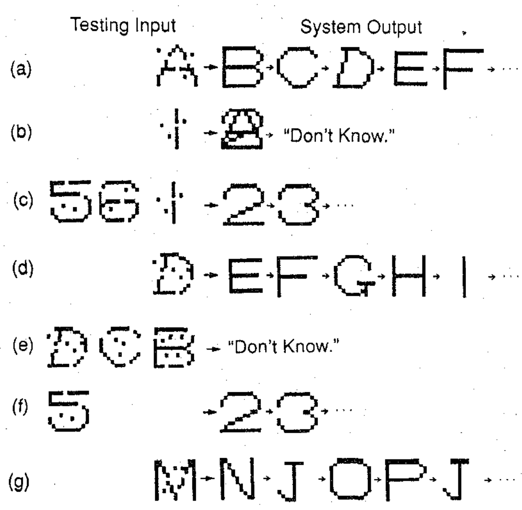

Fig.4 Testing results. (a) Retrieval of sequence (a). (b) Response to an ambiguous testing sequence. (c) Unambiguous retrieval of sequence (b). (d) Response to a non-orthogonal pattern. (e) Response to an unknown sequence. (f) Response to a sequence with gaps. (g) Retrieval of a higher-order or complex spatio-temporal sequence.

3. SUMMARY

We have presented a system that can learn, recognize, and generate spatio-temporal sequences, using a variant of the counterpropagation network. After learning, the system is able to recognize and generate the whole sequence after being presented with a small piece, which may or may not be obscured by spatial noise and may or may not contain temporal gaps, of a stored sequence. Compared to other existing temporal systems, advantages of the present system include short learning time, fast and accurate retrievals, and ability to store a large number of complex sequences consisting of non-orthogonal spatial patterns. These computational properties are desirable in practical applications such as real time speech processing and robotic control.

4. REFERENCES

[1] L. Wang and D.L. Alkon, Ed., *Artificial Neural Networks: Oscillations, Chaos, and Sequence Processing.* IEEE Computer Society Press, Los Alamitos, California, 1993.

[2] S. Grossberg, "On learning of spatiotemporal patterns by netwroks with ordered sensory and motor components, I: Excitatory components of the cerebellum." *Studies Applied Math.*, vol. 48, 1969, pp.105-132.

[3] S. Grossberg, "Some networks that can learn, remember, and reproduce any number of complicated space-time patterns." *Studies Applied Math.*, vol.49, 1970, pp.135-166.

[4] K. Fukushima, "A model of associative memory in the brain." *Kybernetik*, vol. 12, 1973, pp.58-63.

[5] H. Sompolinsky and I. Kanter, "Temporal association in asymmetric neural networks." *Phys. Rev. Lett.*, vol. 57, 1986, pp.2861-2864.

[6] D. Kleinfeld, "Sequential state generation by model neural networks." *Proc. Natl. Acad. Sci. (USA)*, vol. 83, 1986, pp.9469-9473.

[7] D.W. Tank and J.J. Hopfield, "Neural computatation by concentrating information in time." *Proc. Natl. Acad. Sci. (USA)*, vol. 84, 1987, pp.1896-1900.

[8] I. Guyon, L. Personnaz, J.P. Nadal, G. Dreyfus, "Storage and retrieval of complex sequences in neural networks." *Phys. Rev. A*, vol. 38, 1988, pp.6365-6372.

[9] A.C. Tsoi and A.D. Back, "Locally recurrent globally feedforward networks: a critical review of architectures." *IEEE Transactions on Neural Networks*, vol. 5, 1994, pp. 229-239.

[10] J. Buhmann and K. Schulten, "Noise-driven temporal association in nueral networks." *Europhys. Lett.*, vol. 4, 1987, pp.1205-1209.

[11] S. Dehaene, J.P. Changeux, J.P. Nadal, "Neural networks that learn temporal sequences by selection." *Proc. Natl. Acad. Sci. (USA)* vol. 84, 1987, pp.2727-2731.

[12] A.C.C. Coolen and Th.W. Ruijgrok, "Image evolution in Hopfield networks." *Phys. Rev. A*, vol. 38, 1988, pp.4253-4255.

[13] H. Nishimori, T. Nakamura, M. Shiino, "Retrieval of spatio-temporal sequence in asynchronous neural network." *Phys. Rev. A*, vol. 41, 1990, pp.3346-3354.

[14] Lipo Wang, "Oscillatory and chaotic dynamics in neural networks under varying operating conditions," *IEEE Trans. Neural Networks* (in press).

[15] Lipo Wang, E.E. Pichler, and J. Ross, "Oscillations and chaos in neural networks: an exactly solvable model." *Proc. Natl. Acad. Sci. (USA)*, vol. 87, 1990, pp.9467-9471.

[16] Lipo Wang and J. Ross, "Chaos, multiplicity, crisis and synchronicity in higher order neural networks." *Phys. Rev. A*, vol. 44, 1991, pp.R2259 - 2262.

[17] R. Hecht-Nielsen, *Neurocomputing*, Adison-Wesley, New York, pp.147-152, 1990.

[18] A related system using another neural network has been presented by L. Wang and D.L Alkon, "An artificial neural network system for temporal-spatial sequence processing", *Pattern Recognition*, vol. 28, no. 8, pp. 1267-1276, 1995.

[19] G. A. Carpenter and S. Grossberg, "Neural dynamics of category learning and recognition: Attention, memory consolidation, and amnesia." in *Brain Structure, Learning, and Memory*, J. Davis, R. Newburgh, and E. Wegman (eds.), AAAS Symposium Series, pp.233-290, 1986.

[20] G. L. Tarr and S. K. Rogers, M. Kabrisky, M. E. Oxley, and K. L. Priddy, "Effective neural network modelling in C," *Arificial Neural Networks*, T. Kohonen, K. Mäkisara, O. Simula, and J. Kangas (Eds.), Elsevier Science, North-Holland, pp. 1497-1500, 1991.

[21] L. Wu and F. Fallside, "On the design of connectionist vector quantizers," *Computer Speech and Language*, vol. 5, pp. 207-229, 1991.