DETERMINISTIC NEURON - A MODEL FOR FASTER LEARNING

Farid Ahmed and Abdul Ahad S. Awwal
Wright State University
Computer Science & Engineering Department
Dayton, OH 45435.

Abstract

Training in most neural network architectures are currently being done by updating the weights of the network in a way to reduce some error measures. The well-known backpropagation algorithm and some other training algorithms use this approach. Obviously, this has been very successful in mimicking the way the biological neurons do their function. But the problem of slow learning and getting trapped in local minimas of error function domain deserve serious investigation. Various models are proposed with various levels of success to get rid of these two problems. In this work, we propose a deterministic model of the neuron, that guarantees faster learning by modifying the nonlinearity associated with each neuron. Only one such neuron is required to solve the generalized N-bit parity problem.

1 Introduction

The motivation to pursue the present work on modifying the nonlinearity of neuron is derived from our recent work [1, 2] on the character recognition using inner product associative memory. The interesting result obtained from that work was that a slight modification in the non-linearity of neuron in a neural network could have significant contribution to the performance of the net. Such inspiration led us to attack the linearly inseperable problems with properly chosen nonlinearity in the activation function. This type of problems, like the XOR problem cannot be solved by a two-layer (input and output layer) feedforward network. They need one or more hidden layers of neurons. The number of hidden layers and the number of neurons in each layer depend upon the particular problem at hand. But optimising this number has drawn much attention of research recently. Too few will not work and too many hidden units may tune the network to a particular application resulting in the poor generalization capability of the neural network. Kolmogorov [3] gave an upper bound of the number of hidden units. Kruglyak [4] found that the N-bit encoder problem can be solved with only two hidden units. Later on, Stork and Allen showed that only one hidden unit is needed for the N-bit encoder problem and two units are enough for the N-bit parity problem [5]. As an example of the applications of the proposed deterministic neuron, we here established that the N-bit parity problem can be solved with only one neuron in the hidden layer.

2 The Model

The feedforward network for the parity problem is shown in figure 1.

We follow the usual definition of the McCulloch-Pitts neuron model apart from the fact that we are assuming bipolar binary $(+1, -1)$ input instead of unipolar binary. The excitatory only synaptic connection vector to the hidden unit is set to 1's. The hidden unit acts only as a summing element. The sum, *net* is then processed by the novel nonlinearity of the activation function of the output unit.

The following table shows the relationship between the linear sum, *net* and the output y for the

 CH3306-8/93/0000-0845 $1.00

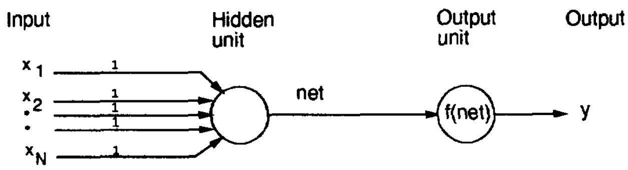

Figure 1: Feedforward Network for N-bit Parity Problem.

parity problem of $N = 2, 3, 4, 5$.

Table 1 : Linear sum vs Output

Linear sum, *net* for Input N=				Output
2	3	4	5	
−2	−3	−4	−5	−1
0	−1	−2	−3	+1
2	1	0	−1	−1
	3	2	1	+1
		4	3	−1
			5	+1

In the above table each column of the linear sum represents the different possible combinations of the sum of the bipolar binary inputs. For example, for $N = 3$, the 4 different possible values of the linear sums are −3, −1, 1 and 3. A very close look at the above table gave us the intuition of getting a periodic nature for the activation function of the output element. Following equation gives the novel activation function for our model.

$$f(net) = \sin(net + N - 1)\frac{\pi}{2} \quad (1)$$

where, N is the number of input variables for the parity problem.

Figure 2 shows the result for $N = 2$.

3 Discussion and Conclusion

We proposed a very novel nonlinearity for the activation function of the neuron that resulted in a deterministically derived network of *"smart"* neurons. The general N-bit parity problem was solved with only one such deterministic neuron. The result, we

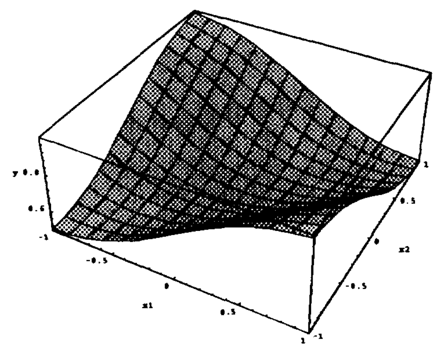

Figure 2: Output of the network for XOR problem

believe, will have far-reaching effect in the reduction of number of neurons and weight updates and thus ensuring faster learning in neural networks. Although such a model may not find its analogy in biological neuron, its departure from the conventional training via weight update can be justified, since it may be physically impossible to create immensly large number of neurons with dense interconnection as in biological system. Fewer *"smart"* neurons may outweigh conventional neurons in their capabilities.

4 References

1. F. Ahmed, A. A. S. Awwal, and P. Chen, "Character Recognition Using a Trinary Inner Product Associative Memory," IEEE 1992 NAECON Conference Proc., pp. 905-909, Dayton, Ohio, May 1992.

2. F. Ahmed and A. A. S. Awwal, "An adaptive opto-electronic neural network for associative pattern retrieval," to appear in Journal of parallel and distributed computing, 17, March, 1993.

3. Kolmogorov, A. N. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition, 1957, Doklady Akademi Nauk, USSR, 114, pp. 953-956.

4. Kruglyak, L. "How to solve the N-bit encoder problem with just two hidden units," Neural Computation, 2(4), 1990, pp. 399-401.

5. D. G. Stork and J. D. Allen, "How to solve the N-bit parity problem with Two hidden units," Neural Networks, vol 5, 1992, pp. 923-926.