Hybrid Network Learning

Dr. Gnanambigai Dinadayalan
Departmentof Computer Science
Indira Gandhi Arts & Science College, Puducherry, India.

Dr. P. Dinadayalan
Department of Computer Science
K. M. centre for P.G. Studies, Puducherry, India.

Dr. K. Balamurugan
Department of Chemistry
Bharathidasan Govt. College for Women
Puducherry India.

ABSTRACT
This paper proposes Neural Network architecture for implementing associative memory. A new model has been developed that has good learning structure and high storage capacity. Hybrid Network Learning comprises interactive counter propagation network and associative memory. Interactive counter propagation network is used for pattern completion. The associative memory is applied for pattern association. Associative memory is content-addressable structure that maps a set of input patterns to a set of output patterns. Associative memory has been expressed in terms of Turing machine. Turing machine is a computing machine which is capable of finding the memory vector which most closely correlates to the input vector. It retrieves previously stored pattern that resembles the current pattern. The Turing machine structure is implemented using B-tree (Turing Tree). The experimental results show that the proposed approach has attained good performance in terms of speed and efficiency.

Keywords
Associative Memory, Learning, Training, Artificial Neuron, Patterns, B-tree, and Turing machine

1. INTRODUCTION
An artificial neural network is a mathematical model or computational model based on biological neural networks [4][5]. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. Content-addressed or associative memory [6] refers to a memory organization in which the memory is accessed by its content. Thus, reference clues are associated with actual memory contents until a desirable match or set of matches is found. Associative memory stands as the most likely model for cognitive memories, as well. Humans retrieve information best when it can be linked to other related information. Turing machines [3] are extremely basic abstract symbol-manipulating devices which, despite their simplicity, can be adapted to simulate the logic of any computer that could possibly be constructed. This paper is intended as a short introduction to neural network, associative memory and turing machine.

The remainder of the paper is organized as follows. In section 2, a preliminary survey of the previous work on neural networks, associative memory and Turing machine and are introduced. Section 3 introduces the concept of Hybrid Network Learning and possible approaches to implement the Hybrid Network Learning. In section 4 and section 5 describe implementation and performance evaluation of Hybrid Network Learning. Finally, we conclude the paper in section 6.

2. RELATED WORKS
We extend the approach to encompass a neural network approach coupling the power of the Turing machine [4] with the speed and storage efficiency of an associative memory. A brief description of neural networks, learning, associative memory and Turing machine are given below. Artificial neural networks [14] are massively parallel adaptive networks of simple nonlinear computing elements called neurons which are intended to abstract and model some of the functionality of the human nervous system in an attempt to partially capture some of its computational strengths. Learning and storage of memory (knowledge) are two main attributes of the neural networks. Learning is the acquisition of knowledge about the world. Most of the neural network structures undergo a learning procedure during which the synaptic weights are adopted. Algorithms for varying these connection strengths such that learning ensues are called learning rules. The objective of learning rules depends on the applications. Learning algorithms may be broadly categorized as supervised (error-based) or unsupervised (output based) [13]. Supervised learning algorithms employ an external reference signal (teacher) and generate an error signal by comparing the reference with the obtained response. Based on the error signal, a neural network modifies its synaptic connections to improve the system performance. In contrast, unsupervised learning algorithms do not incorporate a reference signal, and generally involve self-organization principles that rely only upon local information and internal control mechanisms in order to discover emergent collective properties. Memory is the retention or storage of that knowledge [13].

Neural networks have the potential of accurately describing the behaviour of extremely complex systems. The most commonly used network architectures for process modeling include the feed-forward network, the radial basis function network and the auto-associative network. Advanced network architectures are the dynamic, fuzzy, recurrent, and wavelet networks [1][4]. The feed-forward neural network was the first and arguably simplest type of artificial neural network devised. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network. Contrary to feed-forward networks, recurrent neural networks are models with bi-directional data flow. While a feed-forward network propagates data linearly from input to output, recurrent neural networks also propagate data from later processing stages to earlier stages. A recurrent neural network is a neural network where the connections between the units form a directed cycle. Recurrent neural networks must be approached differently from feedforward neural networks, both when analysing their behavior and training them. In a recurrent network[1][4][5], every neuron receives inputs from every other neuron in the network.

In its simplest form Associative Memory [7][9] is a memory system that allows one data item to be associated to another, so that access to one data item allows access, by association, to the other. In the neural network literature, Associative Memories are referred to as being auto-associative or hetero-associative. An auto-Associative Memory [14] allows recall of the same item that is put in. This type of network is some times called a clean-up network as it can be used to remove noise from a corrupted piece of data. A hetero-Associative Memory [14] allows recall of an associated item that is different from the input query. There are two major classes of neural network based Associative Memories. These are feed forward network based and recurrent network based. The feed forward Associative Memory networks operate by recalling data in one pass through a network where there are no recurrent connections. The recurrent Associative Memories operate by presenting a piece of data and iterating the network until the associated item is recalled.

A Turing machine [3] is a simple mathematical model of modern digital computers having strong computational capacity. The Turing machine is equivalent in computing power to the digital computer and also to all the most general mathematical notions computation. The Turing machine has the following features: it has an external memory which remembers arbitrarily long sequence of input, it has unlimited memory capacity and the model has a facility by which the input at left or right on the tape can be read easily. The Turing machine can produce certain output based on its input. The basic model of Turing machine has a finite control, and input tape that is divided into cells, and a tape head that scans one cell of the tape at a time. The tape has a leftmost cell but is infinite to the right. At any time, the Turing machine has a read/write head positioned at some cell on the tape.

3. HYBRID NETWORK LEARNING

In connection with the analysis we carried out towards the literature survey, we have started working on a new concept namely 'Hybrid Network Learning'. In the starting period, we studied and analyzed Turing machine approach with Associative memory. Fundamental concepts are presented and are studied with examples. Consequently, we aim towards a new technique, namely, 'Hybrid Network Learning'. This technique comprises of Counter-propagation network, Associative memory and Turing machine, applied for pattern classification, pattern recognition and pattern completion. Figure.1 illustrates the overall architecture of the Hybrid Network Learning. It consists of two parts: Interactive Counter propagation network and Associative Turing Machine.

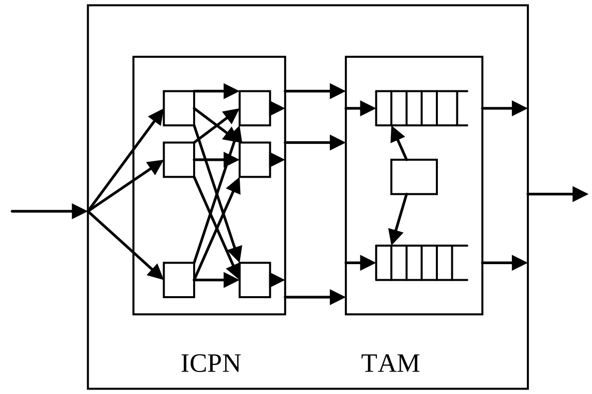

Figure 1. Structure of Hybrid Network Learning

Methods such as counter-propagation that combine network paradigms in building-block fashion may produce networks closer to the brain's architecture than any homogenous structure. The counter-propagation is used to train the given input pattern and produce a desired output. In counter-propagation if the input pattern given is correct or even if the input is partially incomplete or partially incorrect, the counter-propagation trains and produces the correct desired output. Then the desired output is compared with the target pattern. The target patterns are stored in the Turing Machine. The output of counter-propagation network is the input of Turing Machine. The following sections deal with counter propagation network and Associative Turing Machine.

3.1 Interactive Counter-Propagation Network (ICPN)

The Interactive counter-propagation network consists of three primary layers: input, layer, competition layer (hidden layer) and interpolation layer. The neurons in input layer serve only as fan-out points and no computation. The hidden layer is a Kohonen layer with competitive units that do unsupervised learning where as the output layer is a Grossberg layer, which fully connected with the hidden layer and does the supervised learning.

The Interactive counter-propagation is trained in two successive steps. The first step is performed between the input layer and the competition layer. This stage is called unsupervised competitive learning.

$NET_j = w_{1j}x_1 + w_{2j}x_2 + \ldots\ldots + w_{nj}x_n$ .... (1)

i.e., $NET_j = \sum_i w_{ij}x_i$

where NET_j is the NET output of Kohonen neuron j.

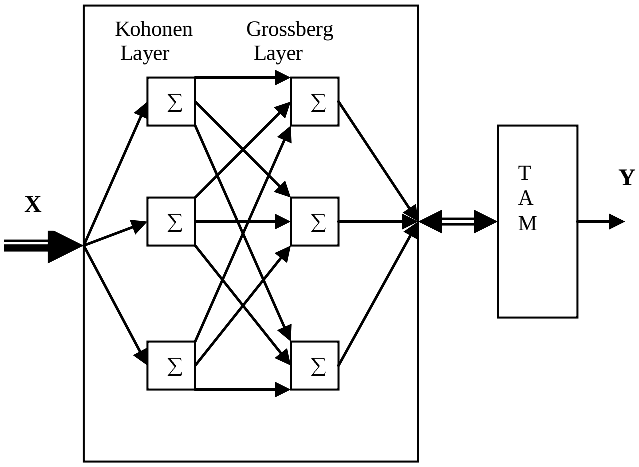

Figure 2. Structure of Interactive ICPN

The Kohonen neuron with the largest NET value is the winner. Its output is set to one; all others are set to zero. Here the network is allowed to develop its weights first, from input to the Kohonen layer nodes through unsupervised winner-take-all learning rule. In this rule for each iteration, the network is presented with all the input patterns of training data sets. The Kohonen layer classifies the input vectors into groups that are similar. This happens in the model in Kohonen's self-organization map. Self-organization means self-adaptation of a neural network. Without target outputs, the closest possible response to a given input signal is to be generated. Self-organization is also learning, but without supervision, it is a case of self-training.

The second step is performed between the competition layer (Kohonen Layer) and interpolation layer (Grossberg Layer). The input of the Grossberg layer is the output of the Kohonen Layer. The Grossberg layer functions in a familiar manner. The Grossberg Layer is relatively simple to train.

$$NET_j = v_{1jk1} + v_{2j}k_2 + \ldots\ldots + v_{nj}k_n = \sum v_{ij}k_i \quad \ldots\ldots (2)$$

where NET_j is the output of Grossberg neuron j, k_i the Kohonen-layer output vector and v_{nj} the Grossberg layer weight matrix.

An input vector is applied, the Kohonen outputs are established, and the Grossberg outputs are calculated as in normal operation. Each weight is adjusted only if it connects to a Kohonen neuron having a nonzero output. The Grossberg training is supervised. Supervised training requires input pattern with target pattern representing desired output. An input pattern is applied, the output of the network is calculated and compared to the corresponding target pattern (from Turing Associative Memory), and weights are changed according to an algorithm that tends to minimize the error. Grossberg layer produces the desired output which is the input of the Turing Associative Memory. Counter propagation network performs pattern completion or association tasks. That is, with missing information these memories are able to recall or complete an expected output.

3.2 Turing Associative Memory

This section considers how Turing machine can be used as associative memory device. Patterns stored in an associative memory are addressed by their contents. Associative memory is content-addressable memory as well as parallel search memory. The major advantage of Associative memory is its capability to performing parallel search and parallel comparison operations. Associative memory can be implemented by using Turing Machine.

We can modify the basic Turing machine by: increasing the number of read/write head, making the two dimensions and adding special purpose memory such as associative memory. All the above modifications in the Turing machine will speed up the operation of the machine. The structure of Associative Turing machine is modeled in figure 3. We construct Turing machine simulating associative memory having two tapes with two independent read/write heads. Each read/write head scans the input pattern simultaneously, thus increasing speed and efficiency of the machine. The header in the Turing machine is used for pattern match and read-write control. The headers allow parallel read or parallel write in the associative memory.

Each tape contains n number of tracks. The tape structure is arranged in a tabular form. It stores large amount of patterns. The associative memories are mainly used for fast search and ordered retrieval of large set of input patterns. The tabulation of patterns can be programmed into the cells of a Turing machine. The Turing machine is a very powerful machine that stores large amount of patterns. Such Turing Associative memory is used to associate one set of vectors with another set of vectors, say input and output patterns. The aim of an Associative Memory is, to produce the associated output pattern whenever one of the input patterns is applied to the Counter-propagation Network. Hybrid Network Learning replaces old stored patterns with new patterns.

The Turing Associative memory is a two tape with two headers and each head scans separately and functions simultaneously. The tapes in the Turing Associative memory are associative memories. The pattern scanning in Associative Memory is fast because it has two read/write headers to read and write patterns concurrently. Let us name tape1 and tape2 as M and M′ associative memories. The associative memory M consists of a related group of patterns whereas the associative memory M′ contains group of patterns which are complement of M. Here M and M′ are mutually exclusive, that is, there is no common pattern in M and M′. When an input pattern (from the output of CPN) is applied to the Turing Associative memory, searching begins in both memories M and M′ simultaneously. If M recognizes the input pattern, then Associative Memory accepts. If M rejects the input pattern, then the input pattern is definitely recognized by M′. As it is Universal Learning the input pattern is in either in M or M′. That is, the Associative Memory produces the output from M or M′. Clearly Associative Memory accepts M U M′. Thus, Hybrid Network Learning recognizes all types of input patterns.

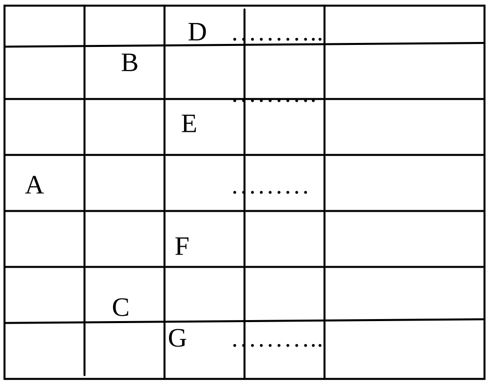

Figure 3. Structure of Turing Associative Memory

4. IMPLEMENTATION

The structure of Turing Machine is implemented using Turing Tree (T-tree). The T-tree structure is same as B-tree implementation. Since two-tape Turing machine is used, two equivalent T-trees are used. Each tape of Turing machine has n tracks. As the Turing machine has n number of tracks it consists of n number of braches in the T-tree. The root node of the T-tree has three components. The first two components have the range of values and third points to the next associated pattern. Each node of the T-tree contains a pattern.

A T-tree is a balanced tree data structure that behaves like a mapping but distributes patterns throughout a number of tree nodes. T-tree is a tree structure that keeps patterns sorted and allows pattern recognition. In T-trees, internal nodes can have a variable number of child nodes within some pre-defined range. When pattern is inserted or removed from a node, its number of child nodes changes. In order to maintain the pre-defined range, internal nodes may be joined or split.

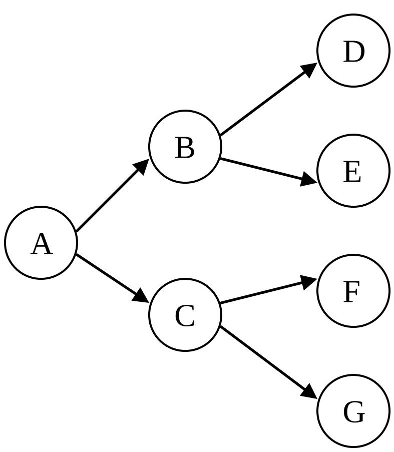

Figure 4. Implementation of Turing Associative Memory using T-tree

A T-tree is kept balanced by requiring that all leaf nodes are at the same depth. A leaf node is a node of a tree pattern structure that has zero child nodes. Often, leaf nodes are the nodes farthest from the root node. A non-leaf node is called an internal node. Some trees only store patterns in internal nodes, though this affects the dynamics of storing pattern in the tree. Pattern recognition process has similar accessing technique as in B-tree.

5. PERFORMANCE EVALUATION

To evaluate the performance of the Associative Learning Network, it has executed a large number of simulation experiments on the basis of the mathematical cost model. Initially through a series of preliminary experiments, it has determined the parameters of the model that have a considerable impact on the costs. The parameters that characterize the topology of the patterns are those that can affect in varying degrees the size and performance of the patterns. From the result of these experiments it was found that Hybrid Network Learning offers better performance than the traditional networks in nearly all cases. In this experiment 500 random patterns are used in building the Hybrid Network Learning.

6. CONCLUSION

We conclude that the Hybrid Network Learning is a good training network structure for learning. Learning required the existence of a systematic procedure for adjusting the network weights to produce the desired output. The Hybrid Network Learning is a combined works of counter-propagation network and Turing Machine. The Turing machine has been implemented using the concept of Associative memory. The Hybrid Network Learning is more powerful than other feed-forward network like perceptron and back-propagation network. The training session of the Hybrid Network Learning is better than traditional feed-forward networks. The memory size is very large and stores huge amount of patterns. The generalization capacity of the Hybrid Network Learning allows it to produce a correct output even when it is given an input vector that is partially incomplete or partially incorrect. Therefore, the Hybrid Network Learning is very efficient and effective for answering all kind of patterns.

7. REFERENCES

[1] D. Gnanambigai, P. Dinadayalan and R. Vasanthakumari, ‘‘Associative Learning Network with Turing Machine”, Proc. of the First International Conference on Data Engineering and Management, ICDEM, Tiruchirappalli, 2008, pp. 365-368.

[2] T. Kohonen, “Self-Organisation And Associative Memory”, Springer-Verlag, 1984.

[3] John E. Hopcroft, Jeffery D. Ullman, “Introduction to Automata Theory, Languages, and Computation”, NPH, 1994.

[4] B Kosko, ‘‘Adaptive Bidirectional Associative Memories,’’ Appl. Optics., vol. 26, no. 23, 1987, pp. 4947-4959.

[5] M. Anabda Rao, “Neural Networks – Algorithms and Applications”, NPH,2003.

[6] T. Kohonen, “Associative Memories: A System Theoretical Approach”, Springer-Verlag, Berlin, 1977.

[7] J Austin, ‘‘ A Review of RAM based Neural Networks”, Proc. of the Fourth International Conference on Microelectronics for Neural Networks and Fuzzy Systems., IEEE Computer Society Press, Turin, 1994, pp. 58-66.

[8] B Widrow, R. G. Winter and R. A Baxter, ‘‘Learning Phenomena in Layered Neural Networks,’’ IEEE 1st International Conference on Neural Networks, M Caudill and C Butler (Ed.), 1987, pp.II-411 - II-429.

[9] J Austin and T J Stonham, ‘‘An Associative Memory for use in Image Recognition and Occlusion Analysis,’’ Image and Vision Computing, vol. 5, no. 4, Nov. 1987, pp. 251-261.

[10] HINT81 Hinton, G. E., and Anderson, J. A., (Eds.) “Parallel Models of Associative Memory”, Lawrence Erlbaum Associates, 1981.

[11] HOPF82 Hopfield, J. J., "Neural networks and physical systems with emergent collective computational abilities", Proceedings of the National Academy of Sciences, Vol. 79, pp. 2,554-2,558, April, 1982.

[12] Bernard M. Moret, “The theory of Computation”, First edition, PEA, 1998.

[13] Dmitry O. Gorodnichy. "The Influence of Self-Connection On The Performance of Pseudo-Inverse Autoassociative Networks", "Radio Electronics. Computer Science. Control" journal,, Vol. 2, No. 2, pp. 49-58, 2001.

[14] Yingquan Wu and Dimitris A. Pados," A Feedforward Bidirectional Associative Memory", IEEE Transactions on Neural networks, vol. 11, No.4, July 2000.

[15] Tae-Dok Eom, Changkyu Choi, and Ju-Jang Lee, "Generalized Asymmetrical Bidirectional Memory", Machine Intelligence & Robotic Control, Vol. 1, No. 1, 43–45, 1999.

[16] H. Shi, Y. Zhao, and X. Zhuang, “A general model for bidirectional associative memories,” IEEE Trans on Systems, Man, and Cybernetics, vol. 28, no. 4, pp. 511–519, 1998.

[17] Z. Wang, ``A bidirectional associative memory based on optimal linear associative memory," IEEE Transactions on Computers, vol. 45, no. 10, pp. 1171--1179, 1996.