Connectionist Network and its Application for Optimization

SUNITA CHAND

Department of Computer Application
Krishna Engineering College,Ghaziabad-201007,India

DR. MANU PRATAP SINGH

Department of Computer Science,
Dr. B. R. A. University, Khandari, Agra – 282002, India

Abstract

Connectionist network is the network of processing elements. These processing elements are connected with each other. It suggests that we may consider it as the fully connected network. The processing elements of this network are basically MP model neuron and normally, employee the bipolar non-linear output function. This network can be used as the associative memory if some constraints are imposed. The constraints of symmetric interconnections between the nodes and bipolar information processing are the normally used constraints. The associative memory feature of connectionist network has various applications in real world. The most widely used application is the optimization.

Keywords: Connectionist network ; MP model; The associative memory ;Optimization; feedback neural network; Hopfield Network

1. Introduction

The connectionist network can be implemented in the form of feedback neural network of Hopfield type, if the above said constraints are satisfied. On associating energy function with every output state of the network provides a powerful tool for accomplishing the task of optimization for the real world problem. There are many situations where a problem can be formulated as minimization or maximization of some cost function or objective function subject to certain constraints. It is possible to map such a problem onto a feedback network, where the units and connection strengths are identified by comparing the cost function of the problem with the energy function of the network expressed in terms of the state values of the units and the connections strengths. The solution to the problem lies in determining the state of the network at the global minimum of the energy function. In this process, it is necessary to overcome the local minima of the energy function. This is accomplished by adopting a simulated annealing schedule for implementing the search for global minimum.

2. Proposed Solution and Molding

The solution to an optimization problem by neural networks consists of the following steps:

(a) Express the objective function or cost function and the constraints of the given problem in terms of the variables of the problem:

Objective function (E) = cost + global constraints (1)

(b) Compare the objective function in equation 1 with the energy function as given in equation 2 of a feedback neural network of Hopfield type to identify the states and the weights of the network in terms of the variables and parameters appearing in the objective function.

Energy function: $$E = -\frac{1}{2}\sum_{i \neq j} w_{ij} s_i s_j \qquad (2)$$

(c) The solution to the optimization problem consists of determining the state corresponding to the global minimum of the energy function of the network. Assuming bipolar states for each unit, the dynamics of the network can be expressed as:

$$s_i(t+1) = \text{sgn}\left(\sum_{j \neq i} w_{ij} s_j(t)\right) \qquad (3)$$

(d) Direct application of the above dynamics in search of a stable state may lead to a state corresponding to a local minimum of the energy function. In order to reach the global minimum, bypassing the local minimum, the concept of stochastic unit is used in the activation dynamics of the network. In order to accomplish this for a stochastic unit the state of the unit is updated using a probability law, which is controlled by a temperature parameter (T). At low temperatures, the stochastic update approaches the deterministic update, which is dictated by the output function of the unit.

(e) The state of a neural network with stochastic units is described in terms of probability distribution. The probability distributions of the states at thermal equilibrium follow the Boltzmann-Gibb's law, namely

$$p(s_\alpha) = \frac{1}{Z} e^{-\frac{E\alpha}{T}} \qquad (4)$$

Where E_α is the energy of the network in the state S_α and Z is the partition function given by

$$Z = \sum_\alpha e^{-E\alpha/T} \qquad (5)$$

The network is allowed to relax to thermal equilibrium at a given temperature (T). Due to stochastic update the state of the network does not remain constant at thermal equilibrium. But the average value of the state of the network remains constant due to stationary of the probabilities $p(s_\alpha)$ of the states of the network at thermal equilibrium. The average value of the state vector is given by

$$\langle S \rangle = \sum_\alpha S_\alpha P(S_\alpha) \qquad (6)$$

(f) At higher temperatures many states are likely to be visited, irrespective of the energies of those states. Thus the local minima of the energy function can be escaped. As the temperature is gradually reduced, the states having lower energies will be visited more frequently. Finally, at T = 0, the state with the lowest energy will have the highest probability. Thus, the state corresponding to the global minimum of the energy function can be reached, escaping the local minima. This method of search for the global minimum of the energy function is called simulated annealing. Implementation of simulated annealing requires computation of stationary

probabilities at thermal equilibrium for each temperature in the annealing schedule. Moreover, the convergence to the global minimum is guaranteed only if the temperature parameter is reduced slowly starting from a high value initially. The state probabilities are computed by collecting the distribution of the states for a large number of cycles of updates of the states of the network at a given temperature. The cycles are repeated until the probabilities of states do not change substantially for different sets of cycles. Once the thermal equilibrium is reached, the temperature is changed to the next lower value. Thus the process of implementation of simulated annealing is very slow.

(g) In order to speed up the process of simulated annealing, the mean-field annealing approximation is used, in which the stochastic update of the binary/bipolar units is replaced by deterministic analog states. The basic idea of mean-field approximation is to replace the fluctuating activation value of each unit by its average value. That means x_i is replaced by $\langle x_i \rangle$.

where $\langle . \rangle$ represent the expectation or average of the random quantities. Likewise, in the average of the state of the *i*th unit given by

$$\langle s_i \rangle = \tanh\left(x_i / T \right) \qquad (7)$$

If x_i is replaced by $\langle x_i \rangle$, we get from Equation 6 and 7

$$\langle s_i \rangle = \tanh\left(\frac{1}{T} \sum_j w_{ij} \langle s_j \rangle \right) \qquad (8)$$

The mean-field approximation involves solving the following recursive equations involving the average values of the states of the units.

$$\langle s_i (t+1) \rangle = \tanh\left(\frac{1}{T} \sum_{j=1}^{N} w_{ij} \langle s_j (t) \rangle \right) \quad , \qquad \text{i = 1,2,….,N.} \qquad (9)$$

These are a set of coupled nonlinear deterministic equations. The equations are solved iteratively starting with some arbitrary values $\langle s_i (0) \rangle$ initially. Once the steady equilibrium values of $\langle s_i \rangle$ have been obtained, then the temperature is lowered. The next set of average states at thermal equilibrium are determined using the average state values at the previous thermal equilibrium condition as the initial values $\langle s_i (0) \rangle$ in the equations above for iterative solution. It must be note that due to deterministic set of equations involved in this computation, the computation will be much faster than in the case of simulated annealing. While the convergence to the global minima is not guaranteed in the mean-field approximation, it seems to yield good results. The set of equations of the mean field approximation is a result of minimization of an effective energy defined as a function of the temperature. This result in an alternative expression for equation 8 and is given by as:

$$\langle S_i \rangle = \tanh\left[-\frac{1}{T} \frac{\partial E(\langle s \rangle)}{\partial \langle S_i \rangle} \right] \quad (10)$$

Where the effective energy $E(\langle s \rangle)$ is the expression for energy of the Hopfield model using average for the state variables.

3. Conclusion and Applications:

Probably, the most studied problem in the context of optimization using the principles of neural networks in the shortest route finding for any network. This phenomenon can apply for determining the shortest routing path for ATM network. The optical fiber network over synchronous digital hierarchy operating on ATM mode appears as one of the best alternatives of the new high-speed telecommunications network. The network management requires planning studies so that the designed network allows efficient operation in real time. The quality of service (QoS) deteriorates drastically when the response timings are not met by the network. In realistic situations, a large number of audio, video and data calls with varying traffic characteristics being initiated by different sources simultaneously, can be varying for the available network bandwidth. The network needs to follow the shortest available path in the most efficient fashion and this process is called the shortest routing path. Neural-optimization techniques can provide the best solution in finding the shortest routing path for ATM networks. Here, the main emphasizes is to explore the different techniques of neural network for optimization and use this method for determining the shortest path for the data packets in ATM network.

REFERENCES

[1] Mangal M., Singh, M. P.(2007a), Analysis of Multidimensional XOR Classification Problem with Evolutionary Feed-forward Neural Networks, *International Journal on Artificial Intelligence Tools*, vol. 16, no.1, pp.111-120.

[2] Mangal M., Singh, M. P.(2007b), Analysis of Classification for the Multidimensional Parity-Bit-Checking Problem with Hybrid Evolutionary Feed-forward Neural Network., *Neurocomputing*, vol. 70, pp.1511-1524.

[3] Bishop,C.M.(1997), Neural Networks for Pattern Recognition. New York, Clarendon

[4] C.J.C. Burges(1995), "Handwritten digit string recognition ", in The Handbook of Brain Theory and Neural Networks (M.A. Arbid, ed.), Cambridge, MA: MIT Press, pp. 447-450.

[5] Yao, X., (1995) "Evolutionary artificial neural networks" in *Encyclopedia of Computer Science and Technology*, vol. 33, A. Kent and J. G. Williams, Eds. New York: Marcel Dekker, pp.137–170.

[6] Yao , X (1994)., "The evolution of connectionist networks", in *Artificial Intelligence and Creativity*, T. Dartnall, Ed. Dordrecht, The Netherlands: Kluwer, pp. 233–243.

[7] Hopfield, J. J., (1982), Neural networks and physical systems with emergent collective computational abilities". *Proceedings of the National Academy Sciences*, USA, Vol. 79, pp. 2554 - 2558.