Appendix IX

Artificial Neural Network

One efficient way of solving complex problems is following the lemma "divide and conquer". A complex system may be decomposed into simpler elements, in order to be able to understand it. Also simple elements may be gathered to produce a complex system (Bar Yam, 1997). Networks are one approach for achieving this. There are a large number of different types of networks, but they all are characterized by the following components: a set of nodes, and connections between nodes.

The nodes can be seen as computational units. They receive inputs, and process them to obtain an output. This processing might be very simple (such as summing the inputs), or quite complex (a node might contain another network...). The connections determine the information flow between nodes. They can be unidirectional, when the information flows only in one sense, and bidirectional, when the information flows in either sense. The interactions of nodes though the connections lead to a global behaviour of the network, which cannot be observed in the elements of the network. This global behaviour is said to be emergent. This means that the abilities of the network supercede the ones of its elements, making networks a very powerful tool.

One type of network sees the nodes as 'artificial neurons'. These are called artificial neural networks (ANNs). An artificial neuron is a computational model inspired

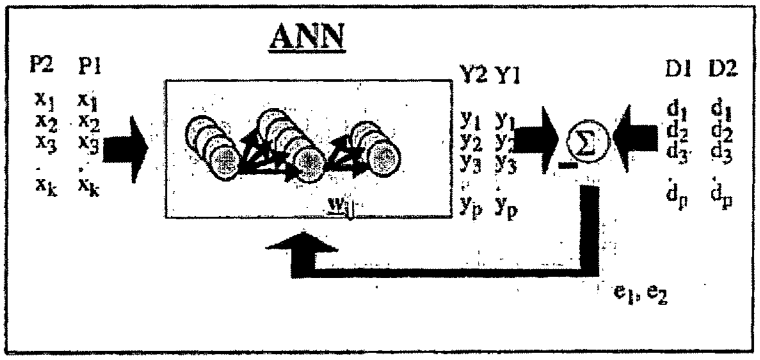

Figure IX.1 The Style of Neural Computation

in the natural neurons. Natural neuronsreceive signals through synapses located on the dendrites or membrane of the neuron. When the signals received are strong enough (surpass a certain threshold), the neuron is activated and emits a signal though the axon. This signal might be sent to another synapse, and might activate other neurons.

An ANN is developed with a systematic step-by-step procedure which optimizes a criterion commonly known as the learning rule. The input/output training data is fundamental for these networks as it conveys the information which is necessary to discover the optimal operating point. In addition, its non linear nature makes neural network processing elements a very flexible system.

Basically, an ANN is a system that receives an input, process the data, and provides an output (Refer Figure IX.1). Once an input is presented to the neural network, and a corresponding desired or target response is set at the output, an error is composed from the difference of the desired response and the real system output.

In neural network design, the engineer chooses the network topology, the trigger function or performance function, learning rule and the criteria for stopping the training phase. So, it is difficult to determine the size and parameters of the network as there is no rule or formula to do the same. The best that can be done for success with the network is to play with it. The problem with this method is when the system does not work properly it is hard to refine the solution. Despite this issue, neural networks based solution is very efficient in terms of development, time and resources.

There are many ways to implement ANNs. It is difficult to find optimal network architecture, considering the uniqueness of each system or problem. There are numerous neural network simulation softwares which allow fast development of neural networks. These softwares provide menus and graphics to define the network in terms of layers and cells in each layers, the propagation rule, activation rule, output function and learning algorithm. They allow feeding of input/output matched pairs termed as patterns for learning and validation.

The permissible error for the validation set can also be specified. The weights and bias are updated and the network is tuned. The learning terminates either on the basis of number of cycles permitted for learning and validation or on achievement of error values less than the target values specified.

Such simulation softwares can be either executable or open source softwares. The executables like Easy NN, Neural Works etc. come with binary code and provide predefined functionality. On the other hand the open source softwares, in addition to providing some predefined functionality, come with the source code and permit the modification and extension of the software definition. SANN, Genesis etc. are examples of such neural network simulation softwares with open source.

The Biological model

ANNs born after McCulloc and Pitts introduced a set of simplified neurons in 1943. These neurons were represented as models of biological networks into conceptual components for circuits that could perform computational tasks. The basic model of the artificial neuron is founded upon the functionality of the biological neuron. By definition, "Neurons are basic signaling units of the nervous system of a living being in which each neuron is a discrete cell whose several processes are from its cell body"

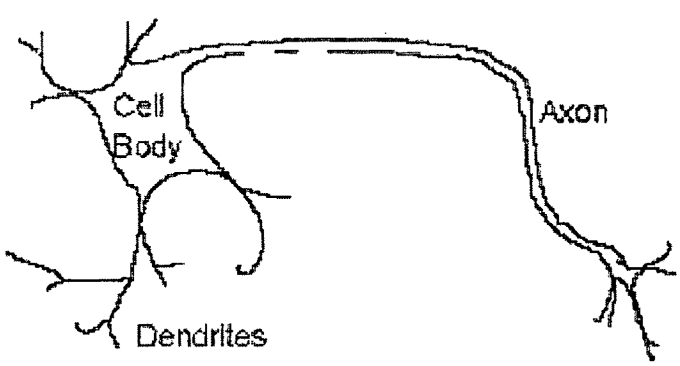

Figure IX.2 Natural Neuron

The biological neuron (Figure IX.2) has four main regions to its structure. The cell body, or soma, has two offshoots from it. The dendrites and the axon end in pre-synaptic terminals. The cell body is the heart of the cell. It contains the nucleolus and maintains protein synthesis. A neuron has many dendrites, which look like a tree structure, receives signals from other neurons.

A single neuron usually has one axon, which expands off from a part of the cell body. This is called the axon hillock. The axon's main purpose is to conduct electrical signals generated at the axon hillock down its length. These signals are called action potentials.

The other end of the axon may split into several branches, which end in a pre-synaptic terminal. The electrical signals (action potential) that the neurons use to convey the information of the brain are all identical. The brain can determine which type of information is being received based on the path of the signal.

The brain analyzes all patterns of signals sent, and from that information it interprets the type of information received. The myelin is a fatty issue that insulates the axon. The non-insulated parts of the axon area are called Nodes of Ranvier. At these nodes, the signal traveling down the axon is regenerated. This ensures that the signal travel down the axon to be fast and constant.

The synapse is the area of contact between two neurons. They do not physically touch because they are separated by a cleft. The electric signals are sent through chemical interaction. The neuron sending the signal is called pre-synaptic cell and the neuron receiving the electrical signal is called postsynaptic cell.

The electrical signals are generated by the membrane potential which is based on differences in concentration of sodium and potassium ions and outside the cell membrane.

Biological neurons can be classified by their function or by the quantity of processes they carry out. When they are classified by processes, they fall into three categories: Unipolar neurons, bipolar neurons and multipolar neurons.

Unipolar neurons have a single process. Their dendrites and axon are located on the same stem. These neurons are found in invertebrates.

Bipolar neurons have two processes. Their dendrites and axon have two separated processes too.

Multipolar neurons: These are commonly found in mammals. Some examples of these neurons are spinal motor neurons, pyramidal cells and purkinje cells.

When biological neurons are classified by function they fall into three categories. The first group is sensory neurons. These neurons provide all information for perception and motor coordination. The second group provides information to muscles, and glands. There are called motor neurons. The last group, the interneuronal, contains all other neurons and has two subclasses. One group called relay or protection interneurons. They are usually found in the brain and connect different parts of it. The other group called local interneurons are only used in local circuits

The Mathematical Model

When modeling an artificial functional model from the biological neuron, three basic components must be taken into consideration. First, the synapses of the biological neuron are modeled as weights. The synapse of the biological neuron is the one which interconnects the neural network and gives the strength of the connection. For an artificial neuron, the weight is a number, and represents the synapse. A negative weight reflects an inhibitory connection, while positive values designate excitatory connections. The following components of the model represent the actual activity of the neuron cell. All inputs are summed altogether and modified by the weights. This activity is referred as a linear combination. Finally, an activation function controls the amplitude of the output. For example, an acceptable range of output is usually between 0 and 1, or it could be -1 and 1. Mathematically, this process is described in Figure IX.3

From this model the interval activity of the neuron can be shown to be:

$$v_k = \sum_{j=1}^{p} w_{kj} x_j$$

The output of the neuron, y_k, would therefore be the outcome of some activation function on the value of v_k.

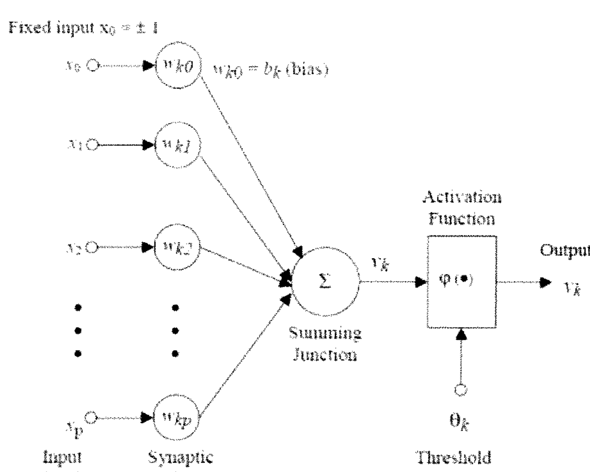

Figure IX.3 Mathematical Model of ANN

Neural Network Topologies

The pattern of connections between the units and the propagation of data can be categorized into two categories.

- **Feed-forward neural networks** where the data ow from input to output units is strictly feed-forward. The data processing can extend over multiple (layers of) units, but no feedback connections are present, that is, connections extending from outputs of units to inputs of units in the same layer or previous layers.
- **Recurrent neural networks** that do contain feedback connections. Contrary to feed-forward networks, the dynamical properties of the network are important. In some cases, the activation values of the units undergo a relaxation process such that the neural network will evolve to a stable state in which these activations do not change anymore. In other applications, the change of the activation values of the output neurons is significant, such that the dynamical behaviour constitutes the output of the neural network.

Training of ANNs

A neural network has to be configured such that the application of a set of inputs produces (either 'direct' or via a relaxation process) the desired set of outputs. Various methods to set the strengths of the connections exist. One way is to set the weights explicitly, using a priori knowledge. Another way is to 'train' the neural network by feeding it teaching patterns and letting it change its weights according to some learning rule. (Refer Figure IX.4)

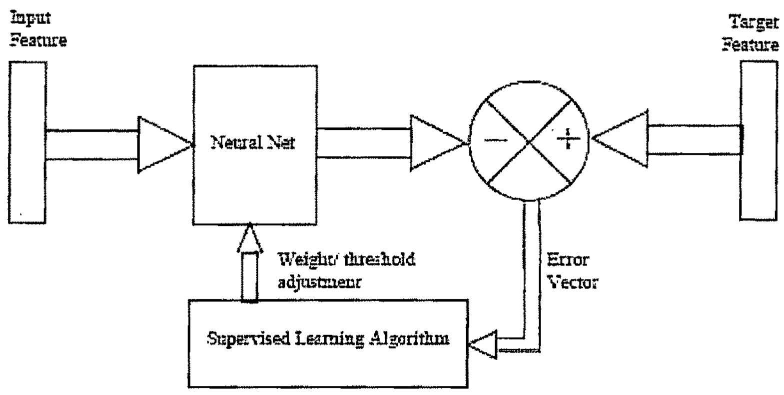

Figure IX.4 Training of ANN

We can categorise the learning situations in two distinct sorts. These are:

- **Supervised learning or Associative learning** in which the network is trained by providing it with input and matching output patterns. These input-output pairs can be provided by an external teacher, or by the system which contains the neural network (self-supervised).
- **Unsupervised learning or Self-organisation** in which an (output) unit is trained to respond to clusters of pattern within the input. In this paradigm the system is supposed to discover statistically salient features of the input population. Unlike the supervised learning paradigm, there is no a priori set of categories into which the patterns are to be classified; rather the system must develop its own representation of the input stimuli.
- **Reinforcement Learning** This type of learning may be considered as an intermediate form of the above two types of learning. Here the learning machine does some action on the environment and gets a feedback response from the environment. The learning system grades its action good (rewarding) or bad (punishable) based on the environmental response and accordingly adjusts its parameters. Generally, parameter adjustment is continued until an equilibrium state occurs, following which there will be no more changes in its parameters. The self organizing neural learning may be categorized under this type of learning.