Quantum Series®

ENGINEERING

QUANTUM *Series*

Sem - 5 & 7 CS & IT

Application of Soft Computing

- Topic-wise coverage of entire syllabus in Question-Answer form.
- Short Questions (2 Marks)

Includes solution of following AKTU Question Papers

2015-16 • 2016-17 • 2017-18 • 2019-20

CONTENTS

RCS 071 : Application of Soft Computing

1 UNIT

Neural Networks-I (Introduction and Architecture)

CONTENTS

1

UNIT

Neural Networks-I (Introduction and Architecture)

CONTENTS

PART-1

Neuron, Nerve Structure and Synapse.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 1.1. **Describe nerve structure and synapse in brief.**

AKTU 2013-14(Sem-III), Marks 05

OR

Explain neuron with its structure.

OR

Explain the structure of a biological neuron with the help of diagram.

AKTU 2017-18(Sem-IV), Marks 07

Answer

A neuron is a small cell that receives electro-chemical signal from its various sources and transmits electrical impulses to other neurons.

Nerve structure :

1. A neuron is composed of nucleus (a cell body) known as soma. Dendrites are long irregularly shaped filaments attached to the soma.
2. The dendrites behave as input channels, all input from other neurons arrive through the dendrites.

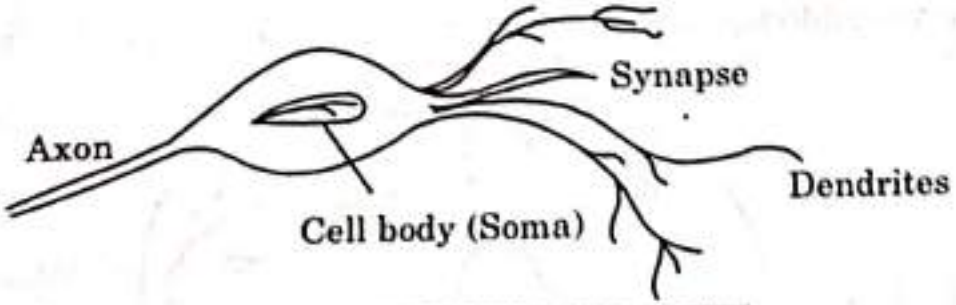


Fig. 1.1.1. Structure of neuron.

3. The axon is electrically active and serves as an output channel.
4. Axons are non-linear threshold device which produce a voltage pulse called action potential or spike that lasts for about a millisecond.
5. If the cumulative inputs received by the soma raise the internal electric potential of the cell known as membrane potential, then the neurons fires by propagating the action potential down the axon to excite other neurons.

Synapse :

1. Synapses (nerve endings) are elementary structural and functional units that mediate the interactions between neurons.
2. The axon terminates in a specialised contact called synapse or synaptic junction that connects the axon with the dendritic links of another neuron.
3. The synaptic junction which is a very minute gap at the end of the dendritic link contains a neuro-transmitter fluid.
4. Each dendritic link can have many synapses acting on it thus bringing about massive interconnectivity.
5. The size of the synapse is believed to be related to learning. Thus, synapses with larger area are thought to be excitatory while those with small area are believed to be inhibitory. Again it is the increased neuronal activity which is thought to be responsible for learning and memory.

Que 1.2. Define structure of human brain with the help of neuron structure. **AKTU 2014-15(Sem-IV), Marks 05**

Answer

1. The concept of neurons as the fundamental constituent of the brain has made the study of its functioning comparatively easier.
2. A neuron is a small cell that receives electro-chemical signals from its various sources and in turn responds by transmitting electrical impulses to the neurons.
3. Brain contains about 10^{10} basic units called neurons. Each neuron in turn, is connected to about 10^4 other neurons.
4. While some of the neurons perform input and output operations, the remaining from a part of an interconnected network of neurons which are responsible for signal transformation and storage of information.

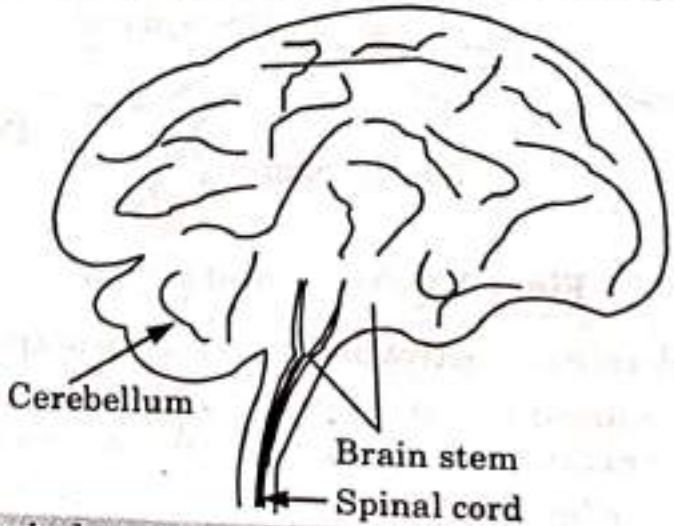


Fig. 1.2.1. Physical structure of the human brain cross-sectional view.

Structure of neuron : Refer Q. 1.1, Page 1–2F, Unit-1.

PART-2

Artificial Neuron and its Model.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 1.3. **Define an artificial neural network. State the properties of the processing element of an artificial neural network ?**

AKTU 2013-14(Sem-IV), Marks 05

Answer

1. Artificial Neural Networks (ANNs) is a simplified model of the biological nervous system and performs the computing task as the human brain.
2. It is a massively parallel distributed processing system made up of highly interconnected neural computing elements that have the ability to learn and thereby acquire knowledge and make it available to use.

Properties of processing element :

1. An artificial neural network is composed of many simple processing elements (called artificial neurons) whose functionality is loosely based on the neurons in human.
2. Processing element learns via a process of adjustments to the connections between the processing elements and element parameters.
3. The processing elements are connected into a particular network pattern, with different patterns serving different functional purposes.
4. The output of processing element in artificial system is electrical only, and may be either analog or digital, or a hybrid.

Que 1.4. **What is artificial intelligence ? What are three technologies in artificial intelligence ?**

AKTU 2014-15(Sem-III), Marks 05

OR

What are the three technologies involved in artificial intelligence ?

AKTU 2014-15(Sem-IV), Marks 05

Answer

Artificial Intelligence :

1. Artificial Intelligence (AI) is the branch of computer science concerned with the study and creation of computer systems that behave in some form of intelligence.
2. Artificial intelligence is the branch of computer science that deals with ways of representing knowledge using symbols rather than numbers, or heuristic method for processing information.

Technologies in artificial intelligence :

1. **Robotics :** This involves developing the mechanical or computer devices that perform tasks requiring a high degree of precision.
2. **Natural language processing :** It helps computers to understand and reacts to statements and commands made in natural language such as English.
3. **Neural networks :** It is a branch of artificial intelligence where the creation of neural network model leads to a computer program which is able to think and learn.

Que 1.5. Define an artificial neural network. State the characteristics of an artificial neural network.

AKTU 2016-17(Sem-IV), Marks 10

OR

What are the characteristics of neural network ?

AKTU 2014-15(Sem-IV), Marks 05

Answer

Artificial neural network : Refer Q. 1.3, Page 1–4F, Unit-1.

Characteristics of ANNs :

1. The NNs can map input patterns to their associated output patterns.
2. The NNs learn by examples. Thus, NN architectures can be trained with known examples of a problem before they are tested for their inference capability on unknown instances of the problem. They can identify new objects that are previously untrained.
3. The NNs possess the capability to generalize. Thus, they can predict new outcomes from past trends.
4. The NNs are robust systems and are fault tolerant. They can therefore, recall full patterns from incomplete, partial or noisy systems.
5. The NNs can process information in parallel, at high speed, and in a distributed manner.

Que 1.6. **Compare and contrast biological neuron and artificial neuron.**

AKTU 2013-14(Sem-IV), Marks 05

Answer

S.No.	Criteria	Biological neuron	Artificial neuron
1.	Processing	Highly parallel, slow but more efficient than artificial neurons.	Highly parallel, fast but lower quality than biological neuron.
2.	Size	10^{11} neurons and 10^{15} interconnections.	10^2 to 10^4 nodes.
3.	Learning	It learns by recalling information.	It cannot learn by recalling information.
4.	Fault tolerance	Performance can go lower with even a minute damage.	It is designed for best performance, therefore can tolerate any kind of fault.
5.	Storage capacity	It saves the data in the form of synapse.	It saves the data in continuous memory locations.

Que 1.7. **What are the fundamental building blocks of the biological neural network ? Discuss.**

AKTU 2014-15(Sem-IV), Marks 05

Answer

In human brain, the elementary nerve cell called a neuron is the fundamental building block of biological neural network. A biological neural network has three major regions :

1. **Soma or cell body :** It contains the cell's nucleus and other vital components called organelles which perform specialized tasks.
2. **A set of dendrites :** It forms a tree like structure that spread out from the cell. The neuron receives its input electrical signal along these set of dendrites.
3. **Axon :** It is tabular extension from the cell (Soma) that carries an electrical signal away from Soma to another neuron for processing.

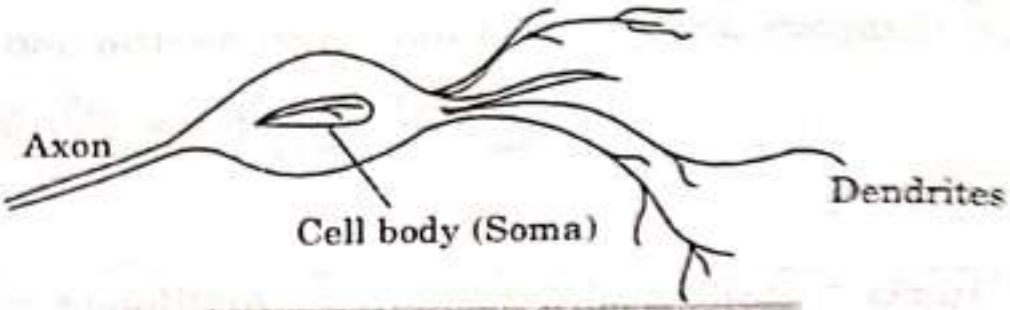


Fig. 1.7.1. Structure of neuron.

PART-3

Activation Functions.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 1.8. What are the activation function ? Explain its use in neuron model.

AKTU 2013-14(Sem-III), Marks 05

Answer

1. An activation function is the basic element in neural model. It is used for limiting the amplitude of the output of a neuron.
2. It is also called squashing function, because it squashes (limits) the permissible amplitude range of the output signal to some finite value.
3. Typically, the normalized amplitude range of the output of a neuron is written as the closed unit interval [0, 1] or [– 1, 1].

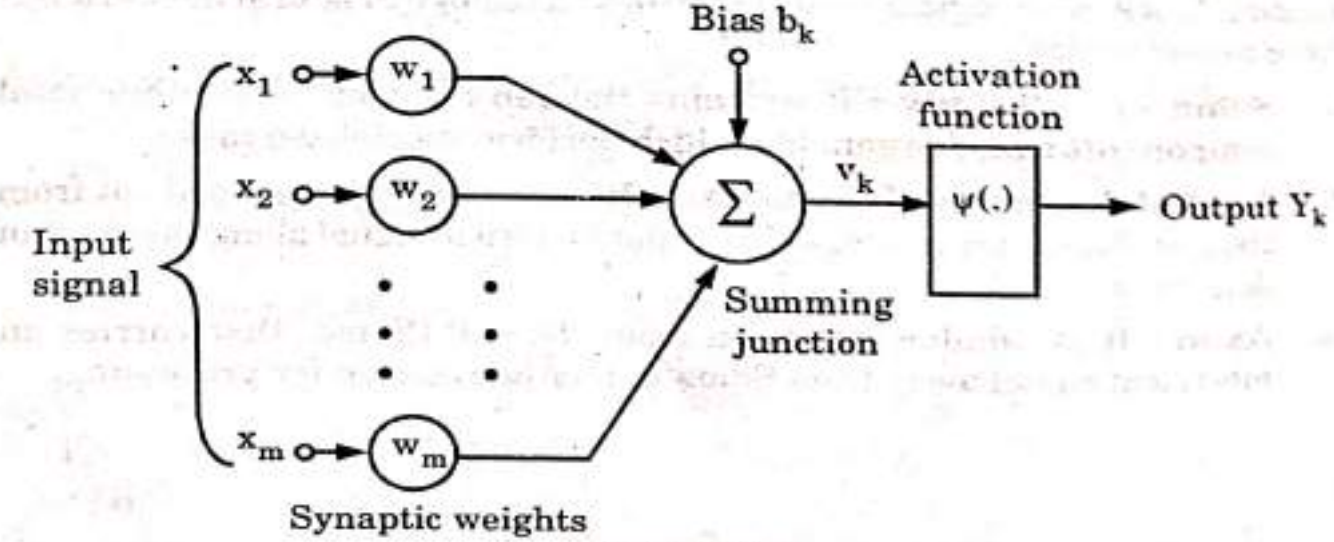


Fig. 1.8.1.

4. In mathematical term the output y_k

$$y_k = y(u_k + b_k) \qquad ...(1.8.1)$$

where $$u_k = \sum_{j=1}^{m} w_j x_j$$

called linear combiner output where y is called activation function, b_k is bias used for the effect of applying an affine transformation to the output u_k of the linear combiner in the model.

5. $$V_k = u_k + b_k \qquad ...(1.8.2)$$

V_k is called induced local field or activation potential of neuron k.

6. Hence, from eq. (1.8.2)

$$Y_k = y(V_k)$$

Use of activation function is neuron model :

1. It helps neuron model in learning.
2. It helps in functional mappings between the inputs and response variable.
3. Its main purpose is to convert an input signal of a node in a neuron model to an output signal.

Que 1.9. **What is the necessity of activation function ? List the commonly used activation functions.**

AKTU 2013-14(Sem-IV), Marks 05

OR

What is activation function ? What are different types of activation function ?

AKTU 2014-15(Sem-III), Marks 05

Answer

Activation function : Refer Q. 1.8, Page 1–7F, Unit-1.

Necessity of activation function :

1. It helps to on calculate exact output of Artificial Neural Network (ANN).
2. It introduces non-linear properties to our network.

Different types of activation function :

1. **Signum function :**
   i. The signum function is also called quantizer function in which the function ϕ is defined as

$$\phi(I) = +1, I > \theta$$
$$= -1, I \leq \theta$$

   where θ is threshold value.
   ii. In signum function, if the value of V is greater than θ, then output is +1 else it is –1.

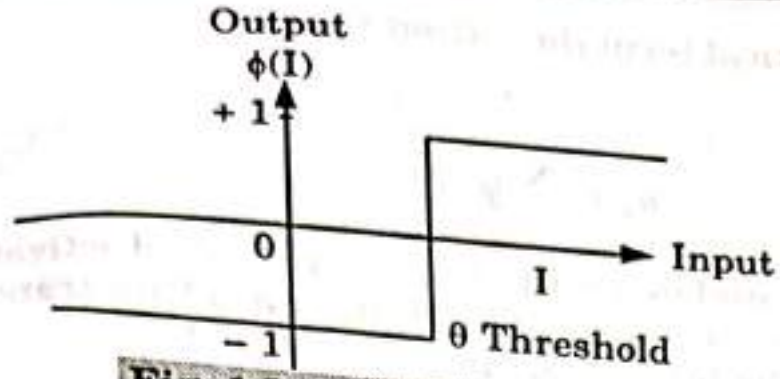


Fig. 1.9.1. Signum function.

2. Sigmoidal function :

i. This function is a continuous function that varies gradually between the asymptotic values 0 and 1 or –1 and +1 and is given by

$$\phi(I) = \frac{1}{1+e^{(-\alpha I)}}$$

where 'α' is slope parameter of sigmoid function.

ii. By varying the parameter 'α', we obtain sigmoid function of different slope. The slope at the origin equals $\alpha/4$.

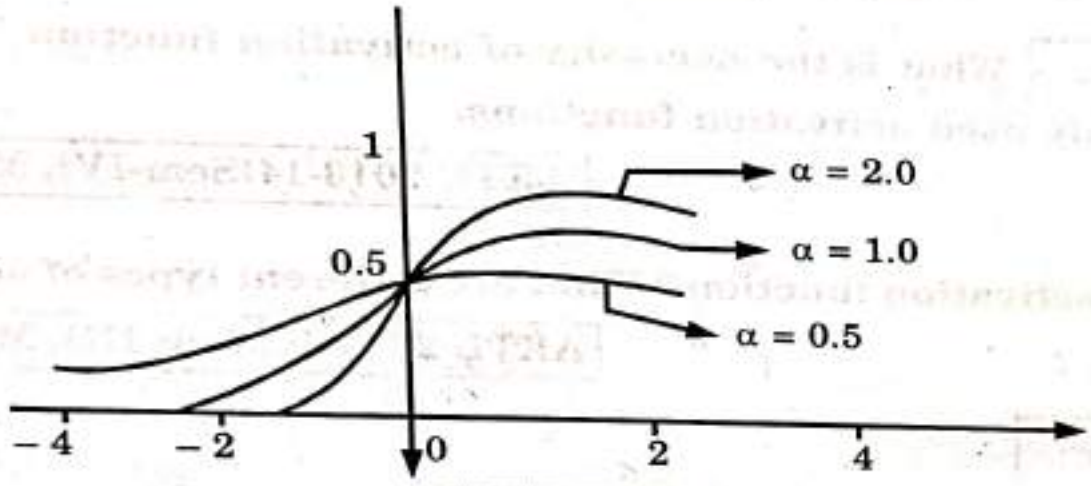


Fig. 1.9.2.

3. Identity function : It is a linear function and can be defined as

$$\phi(I) = I \text{ for all } I$$

The output remains the same as input. The input layer uses the identity activation function.

4. Binary step function : This function can be defined as

$$\phi(I) = \begin{cases} 1 & \text{if} \quad I \geq \theta \\ 0 & \text{if} \quad I < \theta \end{cases}$$

where θ represents the threshold value. This function is most widely used in single layer network to convert the net input to an output that is a binary (1 or 0).

5. Bipolar step function : This function can be defined as

$$\phi(I) = \begin{cases} 1 & \text{if } I \geq \theta \\ -1 & \text{if } I < \theta \end{cases}$$

where θ represent the threshold value. This function is also used in single layer network to convert the net input to an output that is bipolar (+1 or – 1).

Que 1.10. Derive an activation function for thresholding function.

AKTU 2014-15(Sem-IV), Marks 10

Answer

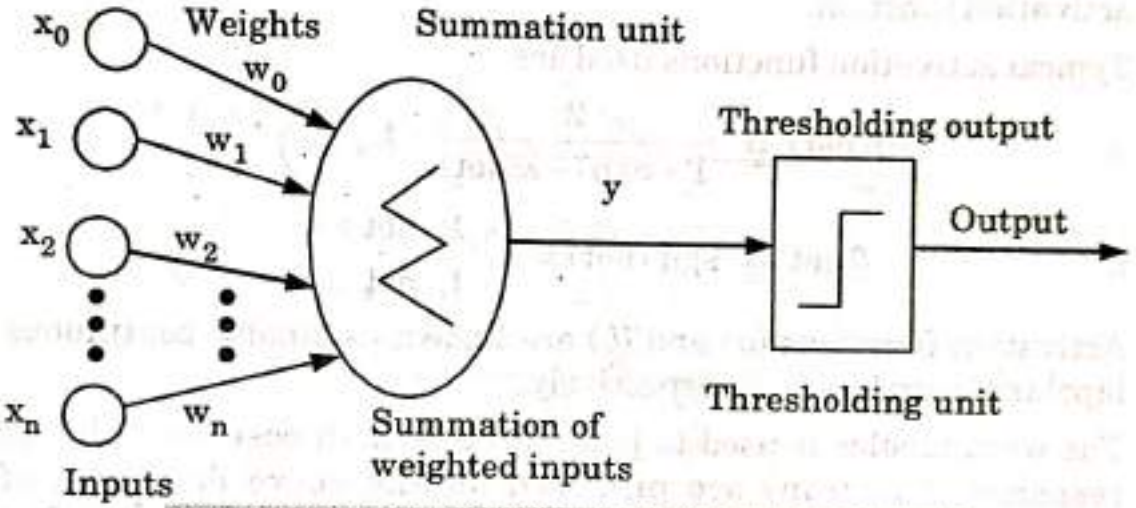


Fig. 1.10.1. Simple model of an artifical neuron.

1. Let, $x_1, x_2, x_3, ..., x_n$ are the n inputs to the artificial neuron. $w_1, w_2, ..., w_n$ are the weights attached to the input links.
2. It is this acceleration of retardation of the input signals that is modeled by the weights.
3. Thus, weights here are multiplicative factors of the inputs to account for the strength of the synapse. Hence, the total input I received by the Soma of the artificial neurons is

$$I = w_1x_1 + w_2x_2 + ... + w_nx_n$$

$$= \sum_{i=1}^{n} w_1x_i \qquad ...(1.10.1)$$

4. To generate the final output y, the sum is passed on to a non-linear filter f called activation function, or transfer function, or squash function which releases the output

$$y = \phi(I) \qquad ...(1.10.2)$$

5. A commonly used activation function is thresholding function. In this, the sum is compared with a threshold value θ. If the value of I is greater than θ, then the output is 1 else it is 0.

$$y = \phi\left(\sum_{i=1}^{n} w_i x_i - \theta\right)$$

Que 1.11. **Write the expression for bipolar continuous and bipolar binary activation function.** **AKTU 2015-16(Sem-IV), Marks 05**

Answer

1. The neuron as a processing node performs the operation of summation of its weighted inputs, or the scalar product computation to obtain net.
2. Subsequently, it performs the non-linear operation $f(\text{net})$ through its activation function.
3. Typical activation functions used are :

a. $$f(\text{net}) \triangleq \frac{2}{1 + \exp(-\lambda\, \text{net})} - 1$$

b. $$f(\text{net}) \triangleq \text{sgn}(\text{net}) = \begin{cases} +1, & \text{net} > 0 \\ -1, & \text{net} < 0 \end{cases}$$

3. Activation functions (a) and (b) are known as bipolar continuous and bipolar binary function respectively.
4. The word bipolar is used to point out that both positive and negative response of neurons are produced for the above definition of the activation function by shifting and scaling the bipolar activation function defined in a and b.

PART-4

Neural Network Architecture : Single Layer and Multilayer Feed Forward Networks, Recurrent Network.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 1.12. **Draw a single layer feed forward network and explain its working functions.** **AKTU 2013-14(Sem-III). Marks 05**

Answer

Working function :

1. Single layer feed forward network comprises of two layers, namely the input layer and the output layer.

2. The input layer neurons receive the input signals and the output layer neurons receive the output signals.
3. The synaptic links carrying the weights connect every input neuron to the output neuron but not vice-versa.
4. Such a network is said to be feed forward in type or acyclic in nature.
5. Output layer performs computation.
6. The sum of the product of weights and the inputs is calculated in each node.
7. The input layer transmits the signals to the output layer.

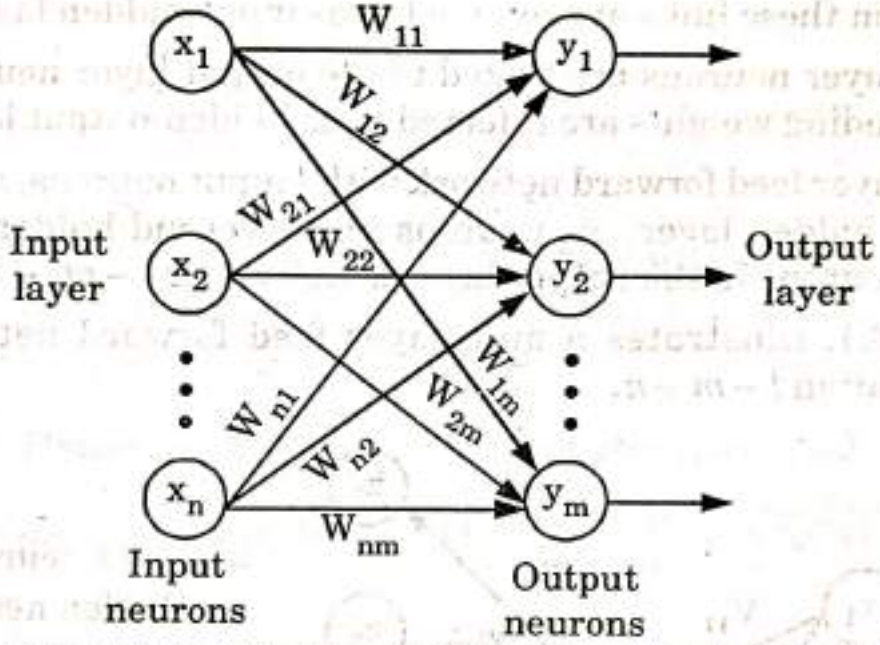


Fig. 1.12.1. Single layer feed forward network.

Que 1.13. Write short notes on :

i. Single layer feed forward network

ii. Multilayer feed forward network

iii. Recurrent network **AKTU 2014-15(Sem-III), Marks 05**

OR

What is neural network architecture ? Explain the different type of neural network architectures. **AKTU 2017-18(Sem-IV), Marks 07**

Answer

Neural network architecture : Neural network architecture refers to the arrangement of neurons into layers and the connection patterns between layers, activation functions and learning methods.

Different types of neural network architecture :

i. **Single layer feed forward network :** Refer Q. 1.12, Page 1–11F, Unit-1.

ii. Multilayer feed forward network :

1. Multilayer feed forward network is made up of multiple layers.
2. Architectures of this network consist of an input, an output layer and have one or more intermediary layers called hidden layers.
3. The computational units of the hidden layer are known as the hidden neurons or hidden units.
4. The hidden layer helps in performing useful intermediary computations before directing the input to the output layer.
5. The input layer neurons are linked to the hidden layer neurons and the weights on these links are referred to as input-hidden layer weights.
6. Hidden layer neurons are linked to the output layer neurons and the corresponding weights are referred to as hidden-output layer weights.
7. A multilayer feed forward network with l input neurons, m_1 neurons in the first hidden layer, m_2 neurons in the second hidden layer and n output neurons in the output layer is written as $l - m_1 - m_2 - n$.
8. Fig. 1.13.1, illustrates a multilayer feed forward network with a configuration $l - m - n$.

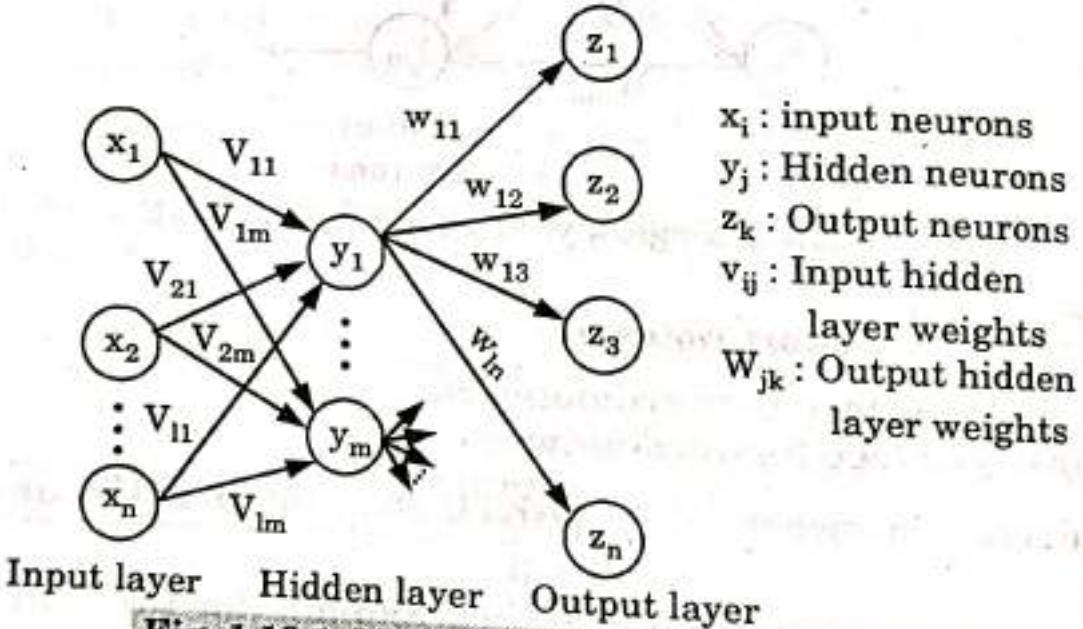


Fig. 1.13.1. Multilayer feed forward network.

iii. Recurrent network :

1. A Recurrent Neural Networks (RNNs) is a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence. This allows recurrent network to exhibit temporal dynamic behaviour.
2. RNNs can use their internal state (memory) to process sequences of inputs.
3. RNNs are designed to take a series of input with no predetermined limit on size.

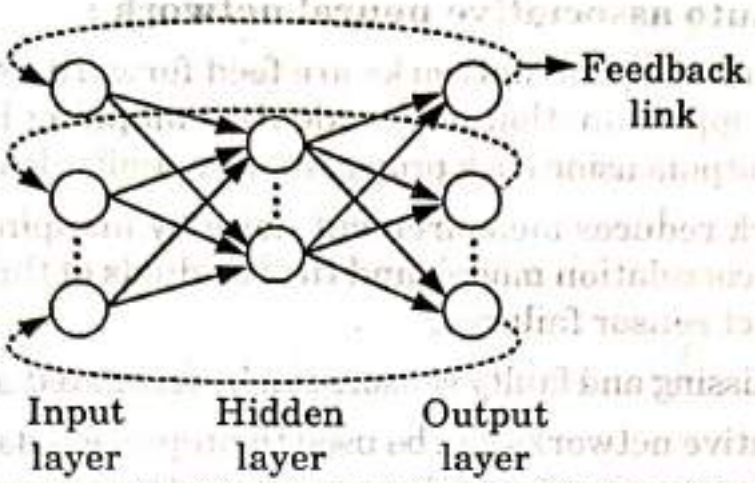


Fig. 1.13.2. A recurrent neural network.

4. RNNs can take one or more input vectors and produce one or more output vectors and the output (s) are influenced not just by weights applied on inputs like a regular NN, but also by a hidden state vector representing the context based on prior input(s)/output(s). So, the same input could produce a different output depending on previous inputs in the series.

Que 1.14. Draw neural network architecture and explain auto associative properties in it. **AKTU 2013-14(Sem-III), Marks 05**

Answer

Neural network architecture :

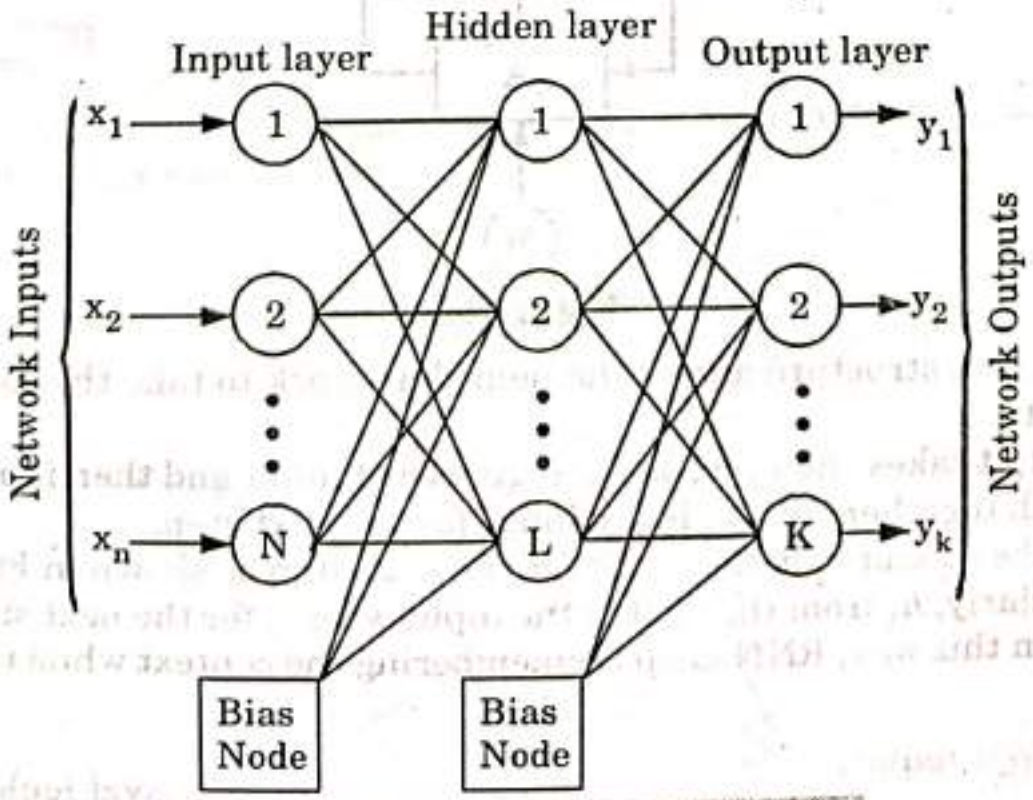


Fig. 1.14.1. Artificial neural network.

Properties of auto associative neural network :

1. Auto associative neural networks are feed forward network trained to produce an approximation of the identity mapping between network inputs and outputs using back propagation or similar learning procedures.
2. This network reduces measurement noise by mapping inputs into the space of the correlation model, and the residuals of this mapping can be used to detect sensor failures.
3. Values for missing and faulty sensors can be estimated using the network.
4. Auto associative networks can be used to preprocess data so that sensor-based calculations can be performed correctly even in the presence of large sensor biases and failures.

Que 1.15. How recurrent network work ? Compare with multilayer neural network. **AKTU 2013-14(Sem-III), Marks 05**

Answer

1. In Recurrent Neural Network (RNN), all the inputs are related to each other.
2. RNN remembers all these relations while training itself.
3. In order to achieve it, the RNN creates the networks with loops in them, which allows it to persist the information as shown in Fig. 1.15.1.

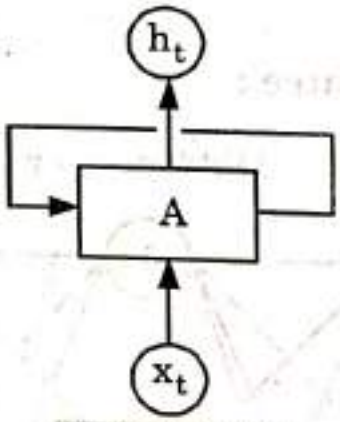


Fig. 1.15.1.

4. This loop structure allows the neural network to take the sequence of input.
5. First, it takes the x_0 from the sequence of input and then it outputs h_0 which together with x_1 is the input for the next step.
6. So, the h_0 and x_1 is the input for the next step as shown in Fig. 1.15.2. Similarly, h_1 from the next is the input with x_2 for the next step and so on. In this way, RNN keeps remembering the context while training.

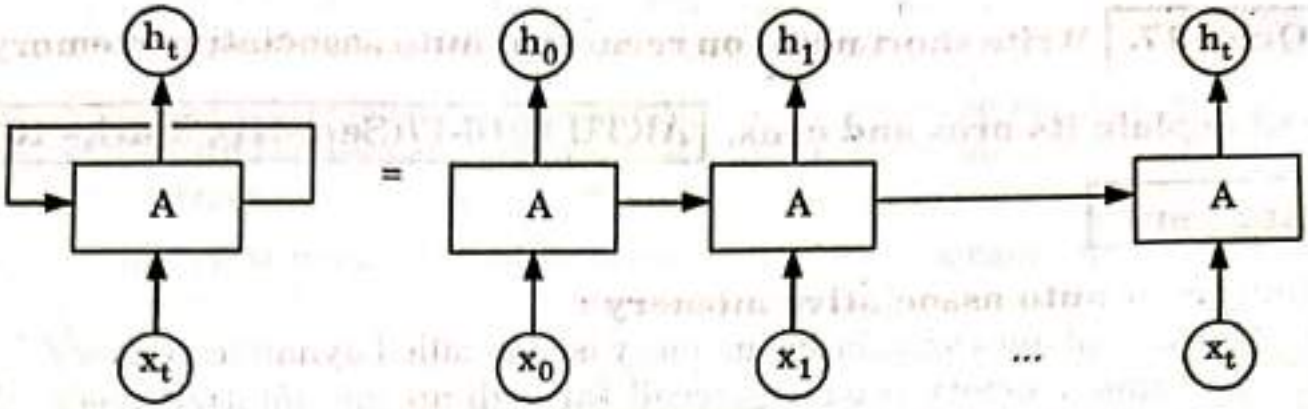


Fig. 1.15.2.

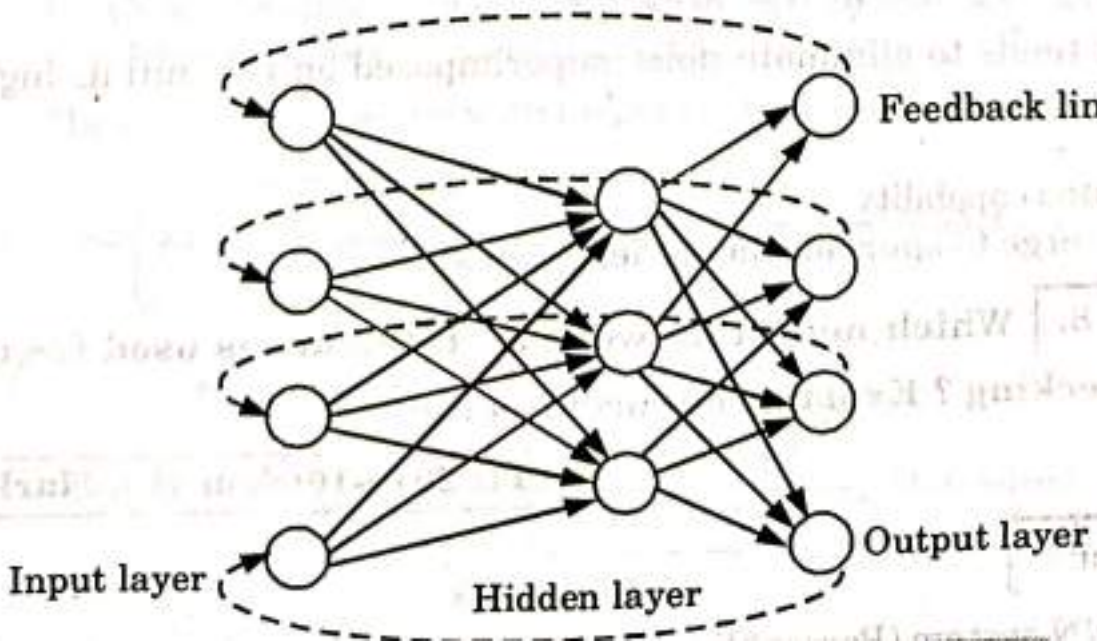
Comparison :

S. No.	Recurrent neural network	Multilayer neural network
1.	Data and calculations flow in a backward direction, from the output data to the input.	Data and calculations flow in a single direction, from the input data to the outputs.
2.	It contains feedback links.	It does not contain feedback links.
3.	It is used for text data, speech data.	It is used for image data, time series of data.

Que 1.16. Construct a recurrent network with four input nodes, three hidden nodes and four output nodes that has lateral inhibition structure in the output layer. **AKTU 2013-14(Sem-IV), Marks 05**

Answer

A recurrent network with four input nodes, three hidden nodes and four output nodes are constructed as follows :



Que 1.17. **Write short notes on recurrent auto associative memory and explain its pros and cons.** **AKTU 2016-17(Sem-III), Marks 10**

Answer

Recurrent auto associative memory :

1. Recurrent auto associative memory is also called dynamic memory.
2. Dynamic memory produces recall through an input/output feedback mechanism, which takes time. Fig. 1.17.1, illustrates the dynamic associative memory.

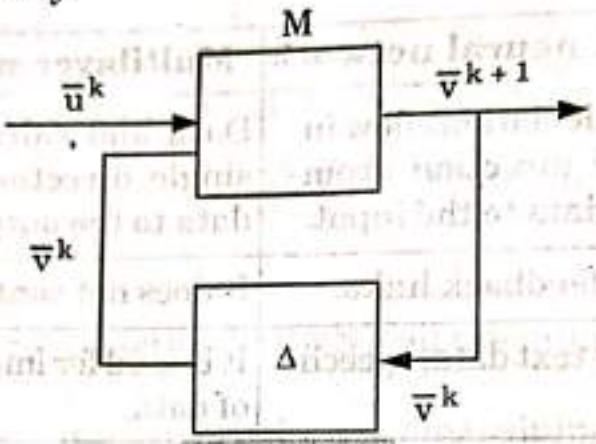


Fig. 1.17.1.

3. The operator M operates at present instant k on the present input $\overline{u}^k$ and output $\overline{v}^k$ to produce the output in the next instant $k + 1$.
4. Δ is a unit delay needed for cyclic operation.
5. Pattern is associated to itself (*i.e.*, auto).
6. It dynamically evolved and finally converges to an equilibrium state according to the recursive formulae,

$$\overline{v}^{(k)} = M(\overline{u}^{(k)})$$

and

$$\overline{v}^{(k+1)} = M'(\overline{u}^{(k)}, \overline{v}^{(k)})$$

Pros : It tends to eliminate noise superimposed on the initializing input vector.

Cons :

i. Limited capability

ii. Converge to spurious memories (states).

Que 1.18. **Which neural network architecture is used for online spell checking ? Explain that architecture.**

AKTU 2015-16(Sem-IV), Marks 10

Answer

1. PENN system (Personalized Error correction using Neural Networks) is used for detecting misspelled words.

2. In particular, we use a neural network that is trained using observations of the specific corrections that a typist makes.
3. Corrections that are made enough times are characterized as possible errors, and the word used to replace the errant word is then defined as the corrected-word associated with the specific error.
4. If the typist makes the same error again, the program subtly suggests the corrected-word that is associated with the error.
5. Since the actual behaviour of the typist is used to determine potential errors, the problems associated with traditional spell checking methods do not apply.

PART-5

Various Learning Techniques, Perception and Convergence Rule.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 1.19. Briefly discuss classification of learning algorithms.

AKTU 2014-15(Sem-III), Marks 05

Answer

Classification of learning algorithm :

1. **Supervised learning :**
   i. Learning is performed with the help of a trainer.
   ii. In ANN, each input vector requires a corresponding target vector, which represents the desired output.
   iii. The input vector along with target vector is called training pair.

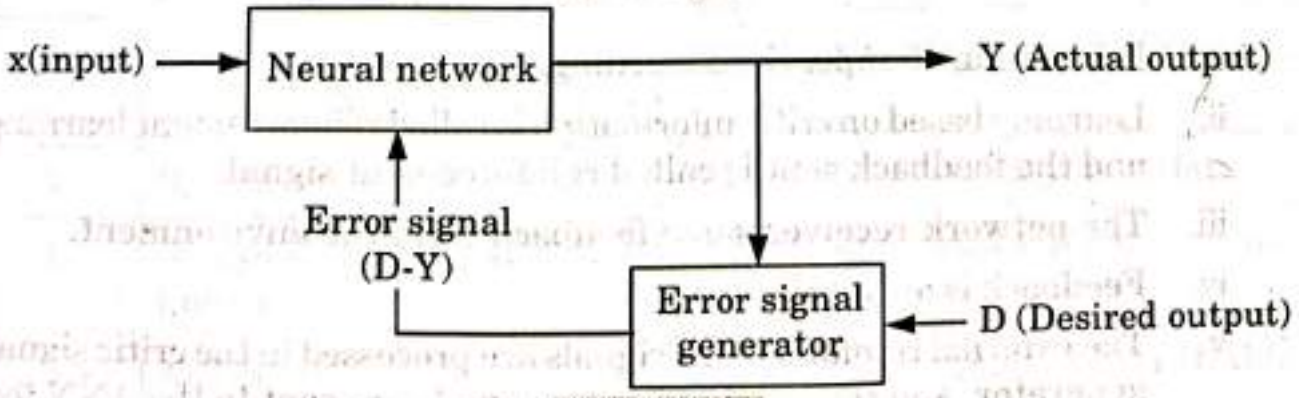


Fig. 1.19.1.

   iv. The input vector results in output vector.

v. The actual output vector is compared with desired output vector.

vi. If there is a difference means an error signal is generated by the network. It is used for adjustment of weights until actual output matches desired output.

2. Unsupervised learning :

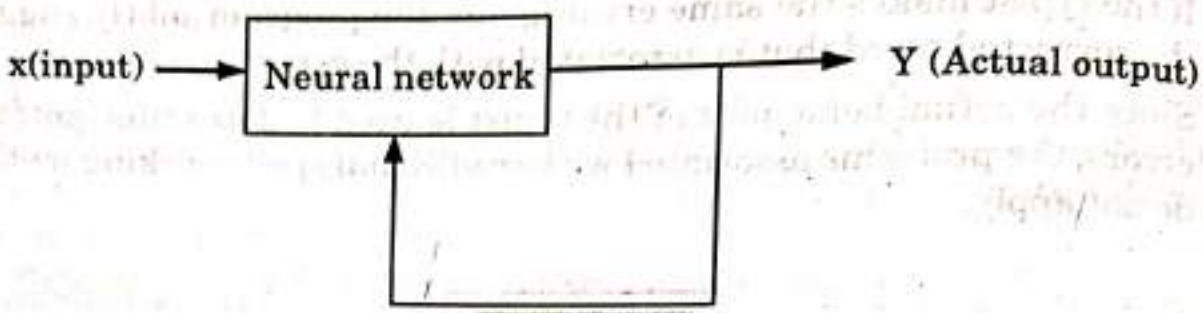


Fig. 1.19.2.

i. Learning is performed without the help of a trainer.

ii. In ANN, during training process, network receives input patterns and organizes it to form clusters. It is observed that no feedback is applied from environment to inform what output should be or whether they are correct.

iii. The network itself discovers patterns, regularities, features/ categories from the input data and relations for the input data over the output.

iv. Exact clusters are formed by discovering similarities and dissimilarities so called as self-organizing.

3. Reinforcement learning :

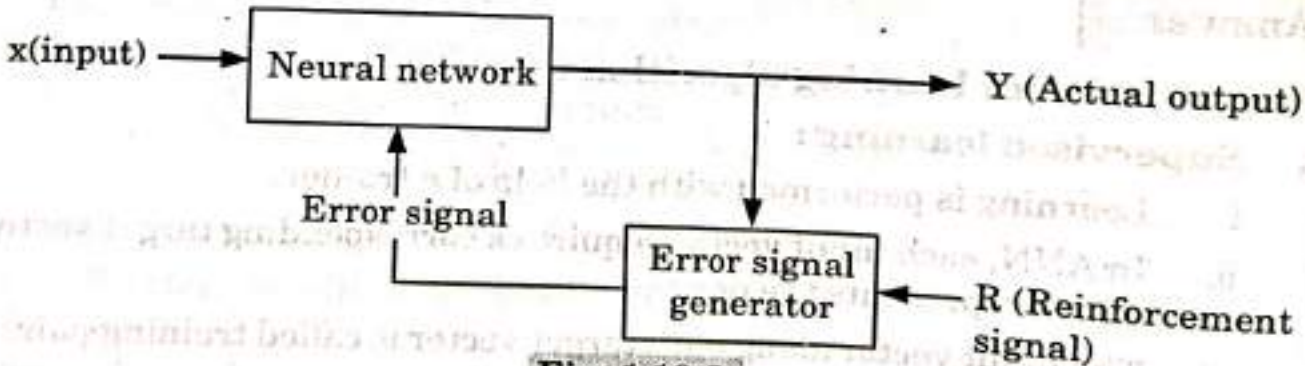


Fig. 1.19.3.

i. It is similar to supervised learning.

ii. Learning based on critic information is called reinforcement learning and the feedback sent is called reinforcement signal.

iii. The network receives some feedback from the environment.

iv. Feedback is only evaluative.

v. The external reinforcement signals are processed in the critic signal generator, and the obtained critic signals are sent to the ANN for adjustment of weights properly to get critic feedback in future.

Que 1.20. **What is meant by learning ? How is supervised learning different from unsupervised learning ?**

AKTU 2013-14(Sem-IV), Marks 05

OR

What is reinforcement learning ? Discuss basic difference between learning factors or classifications.

AKTU 2014-15(Sem-IV), Marks 05

OR

What is called supervised and unsupervised training ? Discuss with suitable example.

AKTU 2017-18(Sem-III), Marks 10

AKTU 2015-16(Sem-III), Marks 05

OR

Draw an artificial neural network. Explain supervised and unsupervised learning in artificial neural network.

AKTU 2016-17(Sem-III), Marks 10

Answer

Artificial neural network : Refer Q. 1.14, Page 1–14F, Unit-1.

Learning :

1. Learning is a process by which the free parameters of neural network are adapted through a process of stimulation by the environment in which the network is embedded.
2. The type of learning is determined by the manner in which the parameter changes take place.

Difference between supervised and unsupervised learning :

S. No.	Supervised learning	Unsupervised learning
1.	It uses known and labeled data as input.	It uses unknown data as input.
2.	Computational complexity is very complex.	Computational complexity is less.
3.	It uses offline analysis.	It uses real time analysis of data.
4.	Number of classes are known.	Number of classes are not known.
5.	Accurate and reliable results.	Moderate accurate and reliable results.

Supervised, unsupervised and reinforcement learning : Refer Q. 1.19, Page 1–18F, Unit-1.

Example of supervised learning :

1. Construction of a predictive model for acute myocardial infarction by using proteomic measurements and clinical variables.
2. Prediction of cardiovascular event risk.
3. Prediction of in-stent rest enosis from plasma metabolites.

Example of unsupervised learning :

1. Construction of predictive representations of patients in an unsupervised fashion from electronic health records.
2. Subtyping of congestive heart failure with preserved ejection fraction.
3. Subtyping of type 2 diabetes mellitus from electronic medical records.

Que 1.21. Explain perceptron rule. State perceptron convergence theorem.

AKTU 2014-15(Sem-III), Marks 05

Answer

Perceptron rule :

1. Perceptrons are trained on desired behaviour. The desired behaviour can be summarized by a set of input, output pairs

$$p_1t_1, p_2t_2, ..., p_Qt_Q$$

where p is an input to the network and t is the corresponding correct (target) output.

2. The objective of perceptron rule is to reduce the error e, which is the difference between the neuron response a, and the target vector t.
3. The perceptron learning rule (learnp) calculates desired changes to the perceptrons weights and biases given an input vector p, and the associated error e.
4. The target vector t must contain values of either 0 or 1.
5. Each time learnp is executed, the perceptron has a better chance of producing the correct outputs.
6. If a bias is not used, learnp works to find a solution by altering only the weight vector w to point toward input vectors to be classified as 1, and away from vectors to be classified as 0. This results in a decision boundary that is perpendicular to w, and which properly classifies the input vectors.
7. There are three conditions that can occur for a single neuron once an input vector p is presented and the network's response a is calculated :

Case 1 : If an input vector is presented and the output of the neuron is correct ($a = t$, and $e = t - a = 0$), then the weight vector w is not altered.

Case 2 : If the neuron output is 0 and should have been 1 ($a = 0$ and $t = 1$, and $e = t - a = 1$), the input vector p is added to the weight vector w. This makes

the weight vector point closer to the input vector, increasing the chance that the input vector will be classified as a 1 in the future.

Case 3 : If the neuron output is 1 and should have been 0 ($a = 1$ and $t = 0$, and $e = t - a = -1$), the input vector p is subtracted from the weight vector w. This makes the weight vector point farther away from the input vector, increasing the chance that the input vector is classified as a 0 in the future.

Perceptron convergence theorem : It state that if there is a weight vector w^* such that $f(w^*p(q)) = t(q)$ for all q, then for any starting vector w, the perceptron learning rule will converge to a weight vector (not necessarily unique and not necessarily w^*) that gives the correct response for all training patterns, and it will do so in a finite number of steps.

Que 1.22. Explain Hebbian learning.

AKTU 2015-16(Sem-III), Marks 10

AKTU 2015-16(Sem-IV), Marks 05

Answer

Hebbian learning :

1. The Hebbian learning rule is a learning rule that specifies how much the weight of the connection between two units should be increased or decreased in proportion to the product of their activation.
2. This rule states that the connections between two neurons might be strengthened if the neurons fire simultaneously.
3. Three major points were stated as a part of Hebbian learning mechanism :
   i. Information is stored in the connections between neurons in neural networks, in the form of weights. In this the input-output pattern pairs (x_i, y_i) are associated by the weight matrix W

$$W = \sum_{i=1}^{n} X_i Y_i^T$$

   ii. Weight change between neurons is proportional to the product of activation values for neurons.

$$\Delta w \propto x.y \Rightarrow \Delta w = \beta.x.y$$

   iii. As learning takes place, simultaneous or repeated activation of weakly connected neurons incrementally changes the strength and pattern of weights, leading to stronger connections.
4. The Hebbian rule works well as long as all the input patterns are orthogonal or uncorrelated.

Que 1.23. Explain delta rule for pattern association.

AKTU 2015-16(Sem-IV), Marks 05

Answer

1. The delta rule is an iterative learning process used for input patterns that are linearly independent but not orthogonal.
2. The delta learning rule is valid for continuous activation functions in the supervised training mode.
3. The learning signal for this rule is called delta is defined as follows :

$$r = \triangleq [di - f(/Wi^{t}x)]\, f'(W_i^{t}x)$$

4. The term $f'(W_i^{t}x)$ is the derivative of the activation function $f(\text{net})$ computed for net = $W_i^{t}x$.
5. This learning rule can be readily derived from the condition of least squared error between 0_i and d_i. Calculating the gradient vector with respect to W_i of the squared error defined as

$$E \triangleq \frac{1}{2}(d_i - 0_i)^2$$

which is equivalent to

$$E = \frac{1}{2}[d_i - f(W_i^{t}x)]^2$$

6. The delta rule is used for neural network training.
7. This rule parallels the discrete perception training rule. It also can be called the continuous perception training rule. The delta learning rule can be generalised for multilayer networks.

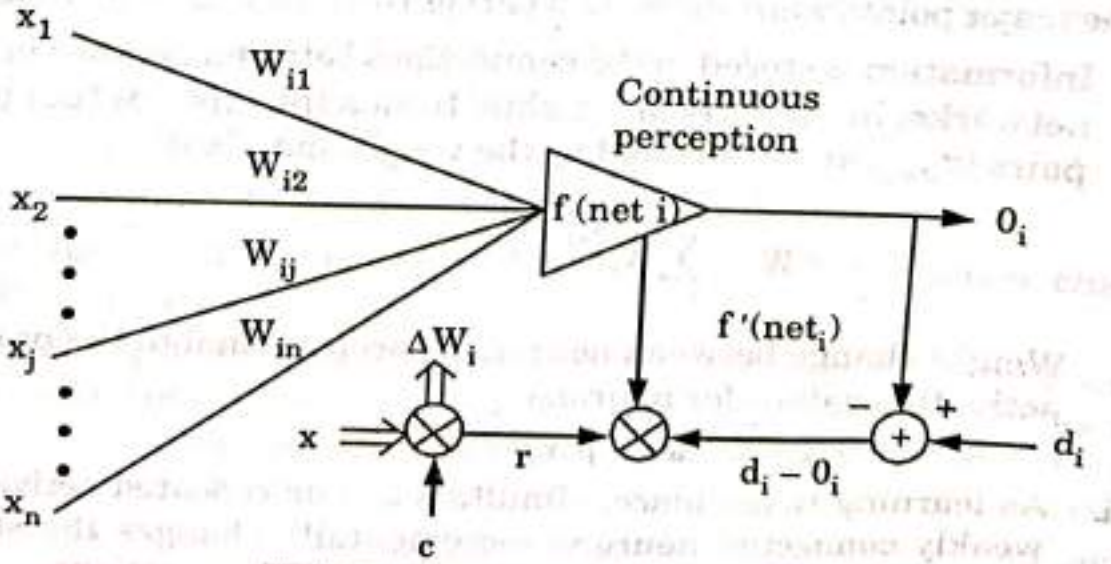


Fig. 1.23.1. Delt learning rule.

Que 1.24. Explain different learning rule in neural network.

Answer

Different learning rule in neural network are :

1. **Hebbian learning rule :** Refer Q.1.22, Page 1–22F, Unit-1.
2. **Perceptron rule :** Refer Q.1.21, Page 1–21F, Unit-1.
3. **Delta learning rule :** Refer Q.1.23, Page 1–22F, Unit-1.
4. **Correlation learning rule :**
   1. The correlation learning rule based on a similar principle as the Hebbian learning rule.
   2. It assumes that weights between responding neurons should be more positive, and weights between neurons with opposite reaction should be more negative.
   3. The correlation rule is the supervised learning.
   4. In mathematical form the correlation learning rule is as follows:

$$\Delta W_{ij} = \eta x_i d_j$$

Where d_j is the desired value of output signal, x_i is the input signal, W_{ij} is the change in weight.

PART-6

Auto-associative and Hetero-associative Memory.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 1.25. **What is hetero-associative memory network ? Explain.**

AKTU 2013-14(Sem-IV), Marks 05

OR

What is Hetero and Auto-associative memory ?

AKTU 2014-15(Sem-III), Marks 05

OR

What is hetero-associative memory ? Describe in context of neural network.

AKTU 2013-14(Sem-III), Marks 05

Answer

Hetero-associative memory :

1. Hetero-associative memory is the associative memory where the retrieved pattern is different from the input pattern not only in content but also different in type and format.
2. In hetero-associative memory, the training input and target output vectors are different.

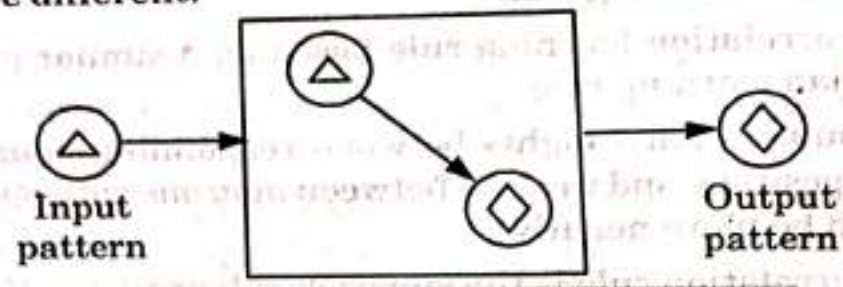


Fig. 1.25.1. Hetero-associative memory.

Auto-associative memory :

1. Auto associative memory is a single layer neural network in which the input training vector and the output target vectors are the same.
2. The weights are determined so that the network stores a set of patterns.
3. It retrieves a previously stored pattern that most closely resembles the current pattern.

Hetero-associative neural network :

1. Hetero-associative network consists of one layer of weighted interconnections.
2. Input has 'n' units and output has 'm' units and there is a weighted interconnection between input and output.
3. Weight is determined by Hebb rule or Delta rule.
4. It is more useful for the association of patterns.

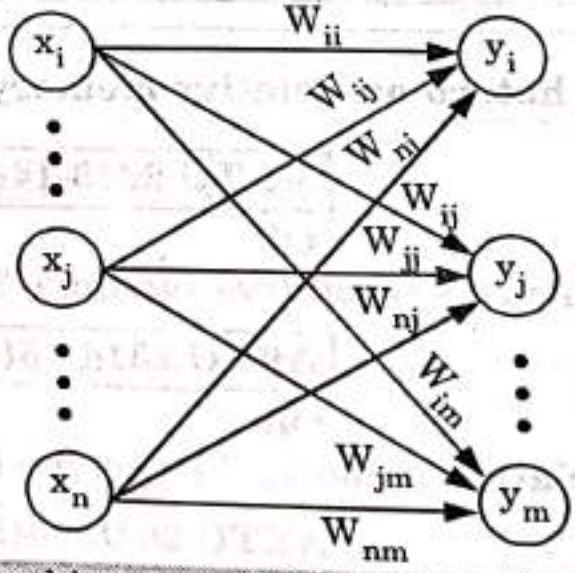
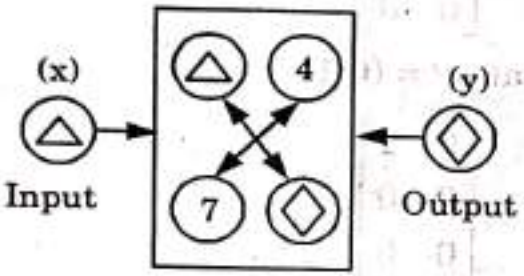
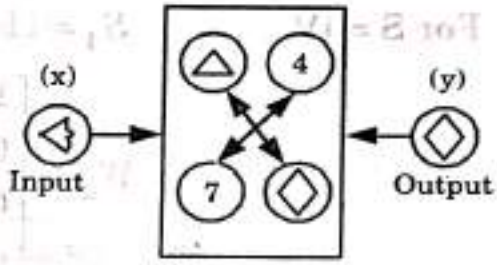


Fig. 1.25.2. Architecture of a hetero-associative neural net.

Que 1.26. **Differentiate between hetero-associative and auto-associative memory.**

AKTU 2017-18(Sem-IV), Marks 10

Answer

S. No.	Hetero-associative memory	Auto-associative memory
1.	In hetero associative memory, the associated pattern pair (x, y) is different.	In auto associative memory, the associated pattern pair (x, y) is same patterned.
2.	This model recalls a output 'y' on a given input x or vice-versa. 	In this model, given a distorted or a partial input pattern x as input, the whole pattern with perfect form y is recalled as output. 
3.	Useful for the association of patterns.	Useful for image refinement.
4.	Hetero associative correlation memory is known as hetero correlators.	Auto associative correlation memories are known as auto correlators.

Que 1.27. **For the given input vectors $S = (S_1, S_2, S_3, S_4)$ and output vectors $t = (t_1, t_2)$, find the weight matrix using hetero- associative training algorithm.**

$S = (S_1, S_2, S_2, S_4)$	$t = (t_1, t_2)$,
I = (1, 0, 1, 0)	(1, 0)
II = (1, 1, 0, 0)	(1, 0)
III = (1, 1, 1, 0)	(0, 0)
IV = (1, 0, 0, 0)	(0, 1)

AKTU 2015-16(Sem-III), Marks 10

AKTU 2015-16(Sem-IV), Marks 10

Answer

The weight matrix for $S = I$ is calculated as :

$$S_1 = (1\ 0\ 1\ 0) \text{ and } t = (1\ 0)$$

$$W_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} [1\ 0] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Similarly, for S = II,

$$S_2 = (1\ 1\ 0\ 0) \text{ and } t = (1\ 0)$$

$$W_2 = \begin{bmatrix}1\\1\\0\\0\end{bmatrix}[1\ 0] = \begin{bmatrix}1 & 0\\1 & 0\\0 & 0\\0 & 0\end{bmatrix}$$

For S = III, $S_3 = (1\ 1\ 1\ 0)$ and $t = (0\ 0)$

$$W_3 = \begin{bmatrix}1\\1\\1\\0\end{bmatrix}[0\ 0] = \begin{bmatrix}0 & 0\\0 & 0\\0 & 0\\0 & 0\end{bmatrix}$$

For S = IV, $S_4 = (1\ 0\ 0\ 0)$ and $t = (0\ 1)$

$$W_4 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}[0\ 1] = \begin{bmatrix}0 & 1\\0 & 0\\0 & 0\\0 & 0\end{bmatrix}$$

The weight matrix to store all four pattern pairs is sum of the weight matrices to store each pattern pairs separately, *i.e.*,

$$= W_1 + W_2 + W_3 + W_4$$

$$= \begin{bmatrix}1 & 0\\0 & 0\\1 & 0\\0 & 0\end{bmatrix} + \begin{bmatrix}1 & 0\\1 & 0\\0 & 0\\0 & 0\end{bmatrix} + \begin{bmatrix}0 & 0\\0 & 0\\0 & 0\\0 & 0\end{bmatrix} + \begin{bmatrix}0 & 1\\0 & 0\\0 & 0\\0 & 0\end{bmatrix} = \begin{bmatrix}2 & 1\\1 & 0\\1 & 0\\0 & 0\end{bmatrix}$$

Que 1.28. Explain the topology and learning in Bidirectional Associative Memory.

AKTU 2015-16(Sem-III), Marks 10

Answer

Topology in BAM :

1. The Bidirectional Associative Memory (BAM) is a simplified version of the backpropagation neural network.
2. A BAM contains two layer of neurons namely input layer (X) and an output layer (Y).
3. Layer are fully connected to each other once the weight have been established, input in layer X presents the patterns in layer Y and vice-versa.
4. A formalization of the backpropagation neural network's topology is shown as :

Topology : $T = (F, L)$

Framework : $F = \{c_0, c_1\}$

where c_0 and c_1 represent the input and output layers respectively.

Cluster layer : $c_1 = \{n_{i,j}\}$

Input layers : $c_0 = \{n_{0,0}, n_{0,1}, n_{0,2}, n_{0,3}\}$

Output layer : $c_2 = \{n_{1,0}, n_{1,1}, n_{1,2}\}$

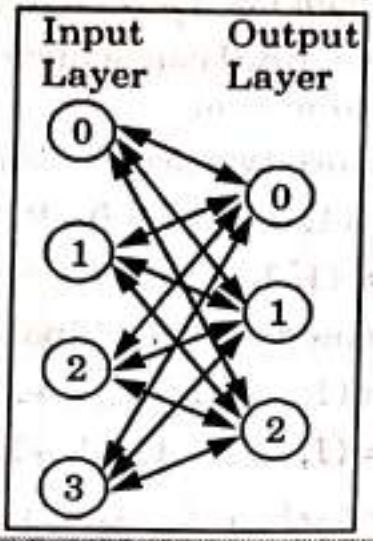


Fig. 1.28.1. BAM network.

Interconnecting linkage :

$$L = \{w_{0,i \leftrightarrow 1,j}\}$$

$$L = \begin{Bmatrix} w_{0,0 \leftrightarrow 1,0}, w_{0,1 \leftrightarrow 1,0}, w_{0,2 \leftrightarrow 1,0}, w_{0,3 \leftrightarrow 1,0}, \\ w_{0,0 \leftrightarrow 1,1}, w_{0,1 \leftrightarrow 1,1}, w_{0,2 \leftrightarrow 1,1}, w_{0,3 \leftrightarrow 1,1}, \\ w_{0,0 \leftrightarrow 1,2}, w_{0,1 \leftrightarrow 1,2}, w_{0,2 \leftrightarrow 1,2}, w_{0,3 \leftrightarrow 1,2}, \end{Bmatrix}$$

Learning in BAM :

1. BAM is created by adapting the nonlinear feedback of the Hopfield model to a hetero associative memory.
2. It is used to store N bipolar pairs (x_p, y_p) where $p = 1, 2, \ldots\ldots N$, where $x_p = (x_{1p} \ldots\ldots x_{j1p})^T$, $y_p = (y_{1p} \ldots\ldots y_{j2p})^T$ and $x_{ip}, y_{jp} \in [-1, 1]$.
3. The first layer has J_1 neurons and the second has J_2 neurons.
4. BAM learning is accomplished with a simple Hebbian rule.
5. The weight matrix is

$$W = YX^T = \sum_{p=1}^{N} y_p x_p^T$$

where X and Y are matrices that represent the sets of bipolar vector pairs to be associated.

6. This learning rule leads to poor memory storage capacity, is sensitive to noise, and is subject to spurious steady states during recall.
7. The retrieval process is an iterative feedback process that starts with $X^{(0)}$ in layer 1 :

$$Y^{(v+1)} = sgn\,(WX^{(v)}),\ X^{(v+1)} = sgn\,(WY^{(v+1)})$$

8. For any real connection matrix, one of fixed points (X_f, Y_f) can be obtained from this iterative process.
9. A fixed point has the properties :

$$X_f = sgn\,(W^T y_f),\ Y_f = sgn\,(W^T x_f)$$

10. With the iterative process, BAM can achieve both hetero associative and auto associative data recollection.

For example : Let us store two associations, A1 : B1 and A2 : B2

$$A1 = (1, 0, 1, 0, 1, 0),\ B1 = (1, 1, 0, 0)$$
$$A2 = (1, 1, 1, 0, 0, 0),\ B2 = (1, 0, 1, 0)$$

These are then transformed into the bipolar forms as :

$$X1 = (1, -1, 1, -1, 1, -1),\ Y1 = (1, 1, -1, -1)$$
$$X2 = (1, 1, 1, -1, -1, -1),\ Y2 = (1, -1, 1, -1)$$

We calculate $M = \sum X_i^T Y_i$ where X_i^T denotes the transpose. So,

$$M = \begin{bmatrix} 2 & 0 & 0 & -2 \\ 0 & -2 & 2 & 0 \\ 2 & 0 & 0 & -2 \\ -2 & 0 & 0 & 2 \\ 0 & 2 & -2 & 0 \\ -2 & 0 & 0 & 2 \end{bmatrix}$$

Que 1.29. **Consider three orthogonal vectors [1 –1 1– 1] [– 1 1 1 –1] [1 1 –1 – 1]. Find the weight matrix to store all the three orthogonal vectors and test the response of the net for each of the input vectors given.**

AKTU 2015-16(Sem-III), Marks 10

Answer

The weight matrix can be calculated as the sum of three weight matrix of the given vectors

$$W = W_1 + W_2 + W_3$$

Weights are calculated using Hebb's outer product rule

$$W_1 = S_1^T(p)\, S_1(p)$$
$$S_1(p) = [1 \quad -1 \quad 1 \quad -1]$$
$$W_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} [1 \quad -1 \quad 1 \quad -1]$$

$$= \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$

$$W_2 = S_2^T(p)\, S_2(p)$$

$$S_2(p) = [-1 \quad 1 \quad 1 \quad -1]$$

$$W_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix} [-1 \quad 1 \quad 1 \quad -1]$$

$$= \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$W_3 = S_3^T(p)\, S_3(p)$$

$$S_3(p) = [1 \quad 1 \quad -1 \quad -1]$$

$$W_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} [1 \; 1 \; -1 \; -1] = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$$

Hence, total weight of three vectors to be stored is

$$W = W_1 + W_2 + W_3$$

$$= \begin{bmatrix} 0 & -1 & 1 & -1 \\ -1 & 0 & -1 & 1 \\ 1 & -1 & 0 & -1 \\ -1 & 1 & -1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 & -1 & 1 \\ -1 & 0 & 1 & -1 \\ -1 & 1 & 0 & -1 \\ 1 & -1 & -1 & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & 1 & -1 & -1 \\ 1 & 0 & -1 & -1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & 1 & 0 \end{bmatrix}$$

$$W = \begin{bmatrix} 0 & -1 & -1 & -1 \\ -1 & 0 & -1 & -1 \\ -1 & -1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

First vector :

Step 1 : To initialize weight matrix

$$W = \begin{bmatrix} 0 & -1 & -1 & -1 \\ -1 & 0 & -1 & -1 \\ -1 & -1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

Step 2 : For testing the first stored input vector follow steps 3 to 5.

Step 3 : $X = [1 \quad -1 \quad 1 \quad -1]$

Step 4 : $y_{-in} = XW$

$$= [1 \quad -1 \quad 1 \quad -1] \begin{bmatrix} 0 & -1 & -1 & -1 \\ -1 & 0 & -1 & -1 \\ -1 & -1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

$$= [1-1 \quad 1-1]$$

Step 5 : $y = f[1-1 \quad 1-1] = (1-1 \quad 1-1)$

The response vector y is the same as stored vector. Hence the input vector is recognized as "known vector".

Second vector :

Step 1 : For testing the second stored input vector do step 2 to 4

Step 2 : $X = [-1 \quad 1 \quad 1 \quad -1]$

Step 3 : $y_{-in} = XW$

$$= [1 \quad -1 \quad 1 \quad -1] \begin{bmatrix} 0 & -1 & -1 & -1 \\ -1 & 0 & -1 & -1 \\ -1 & -1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

$$= [-1 \quad 1 \quad 1-1]$$

Step 4 : $y = f[-1 \quad 1 \quad 1-1] = [-1 \quad 1 \quad 1-1]$

The response vector y is the same as the stored vector. Hence, the input vector is recognized as "known vector".

Third vector :

Step 1 : For testing the third vector follow steps 2 to 4.

Step 2 : $X = [1 \quad 1-1-1]$

Step 3 : $y_{-in} = XW$

$$= [1 \quad 1 \quad -1 \quad -1] \begin{bmatrix} 0 & -1 & -1 & -1 \\ -1 & 0 & -1 & -1 \\ -1 & -1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

$$= [1 \quad 1 - 1 - 1]$$

Step 4 : $y = f[1 \quad 1 \quad -1 - 1] = [1 \quad 1 - 1 - 1]$

The response vector y is the same as the stored vector. Hence the input vector is recognized as "known vector".

Que 1.30.

i. Use the Hebb rule of discrete BAM, find the weight matrix to store the following (binary) input-output pattern pairs.

$S(1) = (1, 1, 0)$ $t(1) = (1, 0)$

$S(2) = (0, 1, 0)$ $t(2) = (0, 1)$

ii. Using binary and bipolar step functions as the activation functions test the response of the network on each of the input pattern ?

AKTU 2015-16(Sem-III), Marks 05

AKTU 2016-17(Sem-III), Marks 10

Answer

i. To find weight matrix :

$S(1) = (1, 1, 0)$ $t(1) = (1, 0)$

$$w_1 = \begin{bmatrix} 1 \\ +1 \\ -1 \end{bmatrix} [1 - 1] = \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$S(2) = (0, 1, 0)$ $t(2) = (0, 1)$

$$w_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} [-1 \quad 1] = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}$$

To store weight matrix, $w = w_1 + w_2 = \begin{bmatrix} 2 & -2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

ii. For binary activation function :

For first input pattern $x = [1 \quad 1 \quad -1]$

From S to t, $t = S(1)$

$$w = [1 \quad 1 \quad -1] \begin{bmatrix} 2 & -2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$= [2 + 0 + 0 - 2 + 0 - 0]$$

$$f[2 - 2] = [1 \ 0]$$

The output [1 0] resembles pattern 1. Thus this is able to recognize the stored pattern.

From t to S,

$$t(1)w^T = [1 \ -1] \begin{bmatrix} 2 & 0 & 0 \\ -2 & 0 & 0 \end{bmatrix}$$

$$= [4 \ 0 \ 0]$$

$$= [1 \ 0 \ 0]$$

which is the pattern 1. Hence this is able to recognize the stored pattern.

For second input pattern = [– 1 1 – 1]

From S to t,

$$S(2) = (-1 \ 1 \ -1) \quad t(2) = (-1 \ 1)$$

$$t = S(2) \quad w = [0 \ 1 \ 0]$$

$$t_{-inj} = [0 \ 1 \ 0] \begin{bmatrix} 2 & -2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$= [0 \ 0]$$

which is the correct response for pattern 2.

From t to S,

$$S = t(2)w^T$$

$$S_{in-i} = [0 \ 1] \begin{bmatrix} 2 & 0 & 0 \\ -2 & 0 & 0 \end{bmatrix}$$

$$S_i = f[-2 \ 0 \ 0] = [-1 \ 0 \ 0]$$

which is also the correct response.

For bipolar activation function :

From S to t,

$$S = [1 \ 1 \ -1]$$

$$t_{in-j} = [1 \ 1 \ -1] \begin{bmatrix} 2 & -2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$t_j = f[2 \ -2] = [1 \ -1]$$

which is the correct response of first input patter $S(1)$.

Now presenting t input pattern

$$S_{in-j} = tw^T$$

$$S_i = f(S_{in-i}) = [1 \ -1] \begin{bmatrix} 2 & 0 & 0 \\ -2 & 0 & 0 \end{bmatrix}$$

$$= f[4 \ 0 \ 0]$$

$$S_i = [1 \ 0 \ 0]$$

which is the correct response. Hence, it is a known vector.

Presenting $S(2)$ input pattern,

$$S = [-1 \quad 1 \quad -1]$$

$$t_{in-j} = [-1 \quad 1 \quad -1] \begin{bmatrix} 2 & -2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = [-2 \quad 2]$$

$$t_j = f[-2 \quad 2] = [-1 \quad 1]$$

which is the correct response.

Presenting t input pattern,

$$t = [-1 \quad 1]$$

$$S_{i-in-i} = [-1 \quad 1] \begin{bmatrix} 2 & 0 & 0 \\ -2 & 0 & 0 \end{bmatrix}$$

$$= [-4 \quad 0 \quad 0]$$

$$S_j = [-1 \quad 0 \quad 0]$$

which is the correct response.

Que 1.31.

a. Use the Hebb rule to find the weight matrix to store the following (binary) input output pattern.

$$X(1) = (1\,0\,1) \quad r(1) = (1\,0)$$
$$X(2) = (0\,1\,0) \quad r(2) = (0\,1)$$

b. Using the binary step function (with threshold 0) as the activation function for both layers, test the response of your network in both directions of each binary training pattern. Initial activation of the other layers is set to zero.

AKTU 2017-18(Sem-III), Marks 10

Answer

i. To find weight matrix :

$X(1) = (1, 0, 1) \quad r(1) = (1, 0)$

$$w_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} [1 \; -1] = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$w_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} [-1 \quad 1] = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}$$

To store weight matrix

$$w = w_1 + w_2 = \begin{bmatrix} 2 & -2 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$$

b. For first input pattern $X = [1 \quad -1 \quad 1]$
From X to r,

$$r = X(1)$$

$$w = [1 \quad -1 \quad 1] \begin{bmatrix} 2 & -2 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$$

$$= [2 + 2 + 2 - 2 - 2 - 2]$$

$$f[6 - 6] = [1 \quad 0]$$

The output [1 0] resembles pattern 1. Thus this is able to recognize the stored pattern.
From r to X,

$$r(1)w^T = [1 \quad -1] \begin{bmatrix} 2 & -2 & 2 \\ -2 & 2 & -2 \end{bmatrix}$$

$$= [4 \quad -4 \quad 4]$$

$$= [1 \quad -1 \quad 1]$$

which is the pattern 1. Hence this is able to recognize the stored pattern.
For second input pattern = $[-1 \quad 1 \quad -1]$
From X to r,

$$X(2) = (-1 \quad 1 \quad -1) \quad r(2) = (-1 \quad 1)$$

$$r = X(2) \quad w = [0 \quad 1 \quad 0]$$

$$= [-6 \quad . \quad 6]$$

$$r_j = f[-6 \; 6] = [-1 \; 1]$$

which is the correct response.
From r to X,

$$r = [-1 \; 1]$$

$$r(2)w^T = [-1 \quad 1] \begin{bmatrix} 2 & -2 & 2 \\ -2 & 2 & -2 \end{bmatrix}$$

$$= [-4 \quad 4 \quad -4]$$

$$= [-1 \quad 1 \quad -1]$$

which is the correct response.

Que 1.32. **Describe briefly the architecture of Hopfield Network.**

AKTU 2016-17(Sem-III), Marks 10

Answer

1. A Hopfield network is a recurrent artificial neural network based on associative neural network models.
2. Hopfield networks are associated with the concept of simulating human memory through pattern recognition and storage.
3. Hopfield model used to solve optimization problems such as travelling salesman, task scheduling.

4. A Hopfield network operates in a discrete line fashion or in other words, it can be said that the input and output patterns are discrete vector, which can be either binary (0, 1) or bipolar (+ 1, – 1) in nature.
5. The network has symmetrical weights with no self-connections *i.e.*, $w_{ij} = w_{ji}$ and $w_{ij} = 0$.

Hopfield network architecture :

1. Hopfield network model consists of neurons with one inverting and one non-inverting output.
2. The output of each neuron should be the input of other neurons but not the input of self.
3. Weight/connection strength is represented by w_{ij}.
4. Connections can be excitatory as well as inhibitory. It would be excitatory, if the output of the neuron is same as the input, otherwise inhibitory.
5. Weights should be symmetrical, *i.e.* $w_{ij} = w_{ji}$.
6. The output from Y_1 going to Y_2, Y_i and Y_n has the weights w_{12}, w_{1i} and w_{1n} respectively. Similarly, other arcs have the weights on them.

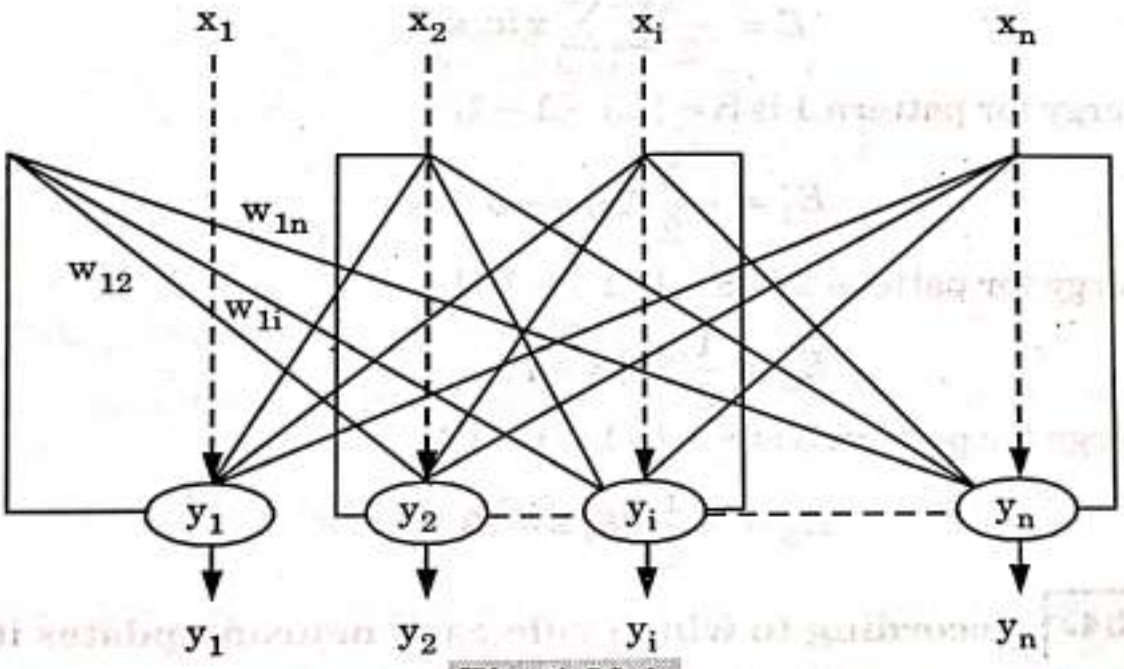


Fig. 1.32.1.

Que 1.33. Draw the architecture of Hopfield net. Design Hopfield net for 4 bit bipolar pattern The training pattern are sample S1[1, 1, – 1, – 1] II Sample S2[– 1, 1, – 1, 1] III sample S3[– 1, –1, – 1, 1]. Find the weight matrix and energy for three input samples.

AKTU 2017-18(Sem-III), Marks 10

Answer

Architecture of Hopfield network : Refer Q. 1.32., Page 1–35F, Unit-1.

Numerical :

$$x = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$$

$$w = \frac{1}{4}\sum_{i=1}^{i} x_i x_i^t = x^t x$$

$$= \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & -1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 1 & 1 & -3 \\ 1 & 3 & -1 & -1 \\ 1 & -1 & 3 & -1 \\ -3 & -1 & -1 & 3 \end{bmatrix}$$

Calculating energy of the network for each pattern,

$$E = \frac{-1}{2}\sum_{i=1}^{4}\sum_{j=1}^{4} x_i w_{ij} x_i^t$$

Energy for pattern 1 is S = [1 1 – 1 – 1]

$$E_1 = -\frac{1}{2}[18] = -9$$

Energy for pattern 2 is S = [– 1 1 – 1 1]

$$E_2 = \frac{1}{2}[22] = -11$$

Energy for pattern 3 is S = [– 1 – 1 – 1 1]

$$E_3 = -\frac{1}{2}[26] = -13$$

Que 1.34. According to which rule each neuron updates its state in Hopfield network ? What is the dynamic behaviour of Hopfield network ?

AKTU 2015-16(Sem-III), Marks 05

Answer

Activity rule :

1. A Hopfield network's activity rule is used by each neuron to update its state, as it is a single neuron with the threshold activation function :

$$x(a) = \Theta(a) \equiv \begin{cases} 1, & a \geq 0 \\ -1, & a < 0 \end{cases}$$

2. Since there is feedback in a Hopfield network (every neuron's output is an input to all the other neurons) we will have to specify an order for the updates to occur. The updates may be :

a. Synchronous updates : All neurons compute their activations.

$$a_i = \sum_j w_{ij} x_j$$

then update states simultaneously to

$$x_i = \Theta(a_i)$$

b. Asynchronous updates :

i. One neuron at a time computes its activation and updates its state.

ii. The sequence of selected neurons may be fixed sequence or a random sequence.

Dynamic behaviour of Hopfield network :

1. We start the development of the theory by assuming a recurrent or feedback network of the form shown in Fig. 1.34.1.

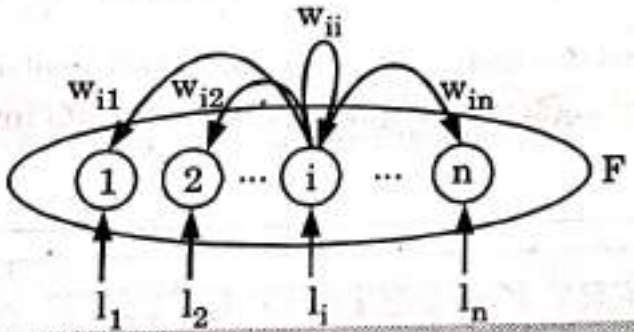


Fig. 1.34.1. Hopfield autoassociative memory architecture.

2. Layer F comprises n neurons which receive external input I_i where $i = 1, \ldots, n$.

3. Each neuron feeds back signals $\delta_i(x_i)$, through weights w_{ij}, $i = 1, \ldots, n$; $j = 1, \ldots, n$.

4. The simplest form of dynamics is the additive activation dynamics model.

$$\dot{x}_i = \underbrace{-A_i x_i}_{\text{Passive decay}} + \underbrace{\sum_{j=1}^{n} w_{ji} \delta_j(x_j)}_{\text{Signal feedback}} + \underbrace{I_i}_{\text{External input}} \qquad i = 1, \ldots, n \qquad \ldots(1.34.1)$$

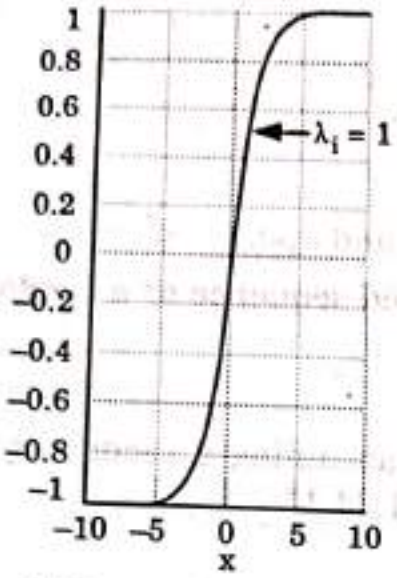
where we assume that individual neurons have distinct sigmoidal characteristics

$$\delta_i(x) = \frac{1 - e^{-\lambda_i x}}{1 + e^{-\lambda_i x}} \qquad \ldots(1.34.2)$$

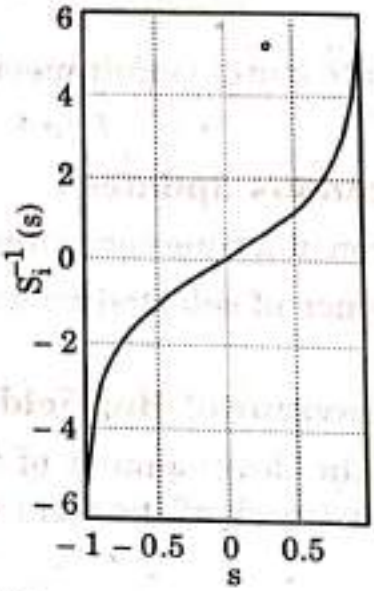
with an inverse function

$$\delta_i^{-1}(s) = \frac{1}{\lambda_i} \ln\left(\frac{1+s}{1-s}\right) = \frac{1}{\lambda_i} \delta^{-1}(s) \qquad \ldots(1.34.3)$$

where $s = \delta(x) = \frac{1 - e^{-x}}{1 + e^{-x}}$. The signal function and its inverse have been plotted in Fig. 1.34.2. Usually the gain scale factor λ_i is common for all neurons.



(a) Bipolar sigmoidal function



(a) Inverse of bipolar sigmoidal function

Fig. 1.34.2. The sigmoidal signal function and its inverse for $\lambda = 1$.

VERY IMPORTANT QUESTIONS

Following questions are very important. These questions may be asked in your SESSIONALS as well as UNIVERSITY EXAMINATION.

Q. 1. What is the necessity of activation function ? List the commonly used activation functions.

Ans: Refer Q. 1.9.

Q. 2. What is artificial intelligence ? What are three technologies in artificial intelligence ?

Ans: Refer Q. 1.4.

Q. 3. Define structure of human brain with the help of neuron structure.

Ans: Refer Q. 1.2.

Q. 4. Define an artificial neural network. State the characteristics of an artificial neural network.

Ans: Refer Q. 1.5.

Q. 5. Compare and contrast biological neuron and artificial neuron.

Ans: Refer Q. 1.6.

Q. 6. What are the fundamental building blocks of the biological neural network ? Discuss.

Ans: Refer Q. 1.7.

Q. 7. Draw neural network architecture and explain auto associative properties in it.

Ans: Refer Q. 1.14.

Q. 8. Construct a recurrent network with four input nodes, three hidden nodes and four output nodes that has lateral inhibition structure in the output layer.

Ans: Refer Q. 1.16.

Q. 9. Briefly discuss classification of learning algorithms.

Ans: Refer Q. 1.19.

Q. 10. What is meant by learning ? How is supervised learning different from unsupervised learning ?

Ans: Refer Q. 1.20.

Q. 11. What is hetero-associative memory ? Describe in context of neural network.

Ans: Refer Q. 1.25.

Q. 12. Differentiate between hetero-associative and auto-associative memory.

Ans: Refer Q. 1.26.

2
UNIT

Neural Network-II (Back Propagation Network)

CONTENTS

PART-1

Architecture : Perceptron Model, Solution, Single Layer Artificial Neural Network, Multilayer Perception Model.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 2.1. What is the multilayer perceptron model ? Explain it.

AKTU 2013-14(Sem-III), Marks 05

Answer

1. Multilayer perceptron is a class of feed forward artificial neural network.
2. Multilayer perceptron model has three layers; an input layer, and output layer, and a layer in between not connected directly to the input or the output and hence, called the hidden layer.
3. For the perceptrons in the input layer, we use linear transfer function, and for the perceptrons in the hidden layer and the output layer, we use sigmoidal or squashed-*S* function.
4. The input layer serves to distribute the values they receive to the next layer and so, does not perform a weighted sum or threshold.
5. The input-output mapping of multilayer perceptron is shown in Fig. 2.1.1 and is represented by

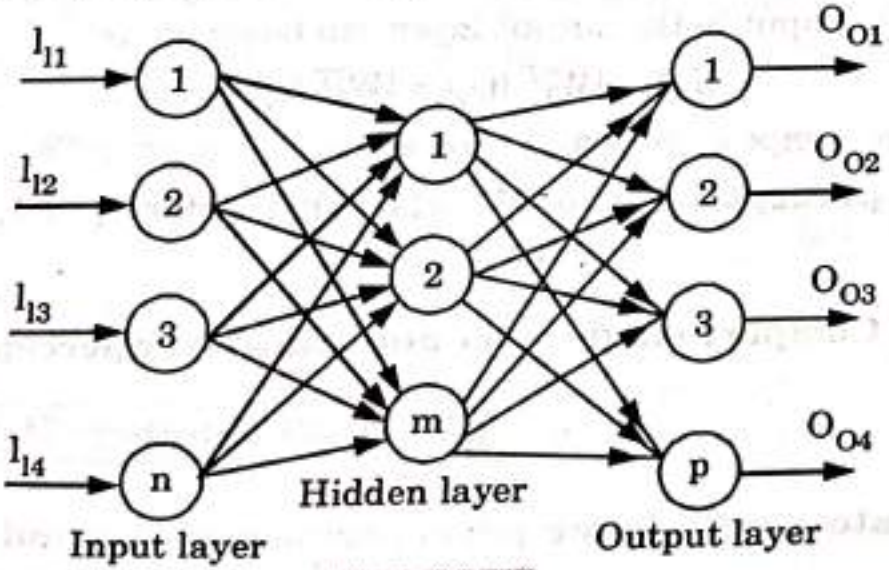


Fig. 2.1.1.

6. Multilayer perceptron does not increase computational power over a single layer neural network unless there is a non-linear activation function between layers.

Que 2.2. Explain single layer and multilayer perceptron neural network.

AKTU 2017-18(Sem-IV), Marks 10

Answer

Single layer perceptron neural network :

1. In single layer perceptron neural network, a single neuron can perform certain simple pattern detection problems and we need larger networks to offer greater computational capabilities.
2. The input layer consists of 'n' neurons and the output layer consists of 'm' neurons. W_{ij} indicate the weight of the synapse connecting ith input neuron to the jth output neuron.
3. The inputs of the input layer and the corresponding outputs of the output layer are given as :

$$I_I = \begin{Bmatrix} I_{I1} \\ I_{I2} \\ \vdots \\ I_{In} \end{Bmatrix}_{n \times 1}; O_O = \begin{Bmatrix} O_{O1} \\ O_{O2} \\ \vdots \\ O_{Om} \end{Bmatrix}_{m \times 1} \quad ...(2.2.1)$$

4. Assume, we use linear transfer function for the neurons in the input layer and the unipolar sigmoidal function for the neurons in the output layer.

$$\underset{n \times 1}{\{O_1\}} = \underset{m \times 1}{\{I_1\}} \text{ (linear transfer function)} \quad ...(2.2.2)$$

$$I_{oj} = W_{1j} I_{I1} + W_{2j} I_{I2} + ... + W_{nj} I_{IN} \quad ...(2.2.3)$$

5. Hence, the input to the output layer can be given as

$$\underset{m \times 1}{\{I_O\}} = \underset{m \times n}{[W]^T} \{O_I\} = [W]^T \underset{n \times 1}{\{I_I\}}$$

Multilayer perceptron neural network : Refer Q. 2.1, Page 2–2F, Unit-2.

Que 2.3. Compare single layer and multilayer perceptron model.

AKTU 2013-14(Sem-III), Marks 05

OR

Differentiate single layer perceptron method and multilayer perceptron method.

AKTU 2016-17(Sem-III), Marks 10

Answer

S. No.	Single layer perceptron model	Multilayer perceptron model
1.	It does not contain any hidden layer.	It contains one or more hidden layer.
2.	It can learn only linear function.	It can learn linear and non-linear function.
3.	It requires less training input.	It requires more number of training inputs.
4.	Learning is faster.	Learning is slower.
5.	Faster execution of the final network.	Slower execution of the final network.

Que 2.4. Discuss the different features of single layer perception.

AKTU 2013-14(Sem-IV), Marks 05

Answer

Features of single layer perception :

1. Input is multi-dimensional (*i.e.*, input can be a vector).
2. Input nodes are connected to a node or multiple nodes in the next layer.
3. The output values of a perceptron can take on only one of two values (0 or 1) due to the hard-limit transfer function.
4. Perceptrons can only classify linearly separable sets of vectors.
5. A signal layer perceptron can only learn linear functions.

Que 2.5. What is the significance of error signal in perceptron network ? Explain.

AKTU 2013-14(Sem-IV), Marks 05

Answer

1. In perceptron model, neuron k is driven by a signal vector $x(n)$ produced by one or more layers of hidden neurons, which are themselves driven by an input vector applied to the source nodes of the neural network.
2. The argument denotes discrete time involved in adjusting the synaptic weights of neuron k.
3. The output signal of neuron k is denoted by $y_k(n)$. This output signal, representing the only output of the neural network, is compared to a desired response or target output, denoted by $d_k(n)$.

4. Consequently, an error signal, denoted by $e_k(n)$ is produced. By definition,

$$e_k(n) = d_k(n) - y_k(n)$$

5. The error signal $e_k(n)$ actuates a control mechanism, the purpose of which is to apply a sequence of corrective adjustments to the synaptic weights of neuron k.

6. The corrective adjustments are designed to make the output signal $y_k(n)$ come closer to the desired response $d_k(n)$ in a step by step manner.

7. This objective is achieved by minimizing the cost function of index of performance, $e(n)$ defined in terms of the error signal $e_k(n)$ as :

$$e(n) = \frac{1}{2} e^2{}_k(n)$$

where, $e(n)$ is the instantaneous value of error energy.

8. The step-by-step adjustments to the synaptic weights of neuron k are continued until the system reaches steady state. At that point, the learning process is terminated. The learning process described here is referred to as error-correction learning.

Que 2.6. **How linear separable task is defined for two dimensional spaces ? Discuss XOR problem.** **AKTU 2014-15(Sem-IV), Marks 05**

Answer

Linear separability and the XOR problem :

1. Consider two-input patterns (X_1, X_2) being classified into two classes as shown in Fig. 2.6.1.

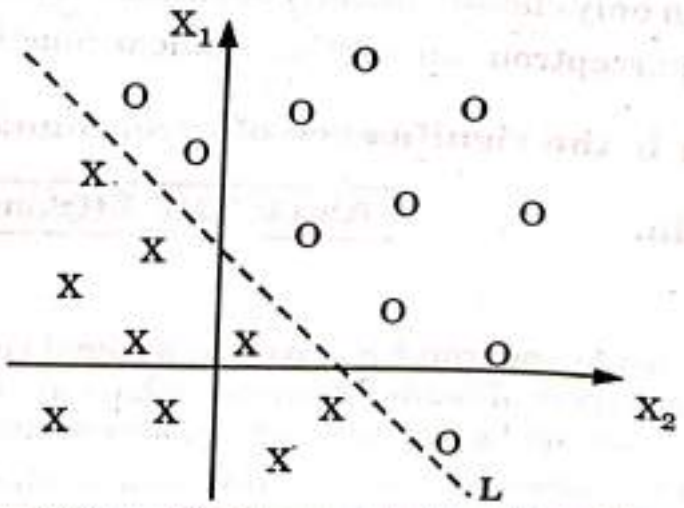


Fig. 2.6.1. Linearly separable pattern.

2. Each point with either symbol of X or O represents a pattern with a set of values (X_1, X_2). Each pattern is classified into one of two classes.

3. These classes can be separated with a single line L. They are known as linearly separable patterns.
4. Linear separability refers to the fact that classes of patterns with n-dimensional vector $X = (X_1, X_2, X_3, \ldots\ldots, X_n)$ can be separated with a single decision surface. Here, the line L represents the decision surface.
5. The processing unit of a single-layer perceptron network is able to categorize a set of patterns into two classes.
6. Conversely, the two classes must be linearly separable for the perceptron network to function correctly.
7. An example of linearly inseparable pattern is a logical exclusive-*OR* (*XOR*) function. As shown in Fig. 2.6.2 is the illustration of *XOR* function that two classes, 0 for black dot and 1 for white dot, cannot be separated with a single line. The solution seems that patterns of (X_1, X_2) can be logically classified with two lines L_1 and L_2.

X_1	X_2	X
0	0	0
0	1	1
1	0	1
1	1	0

$Y = X_1 \oplus X_2$

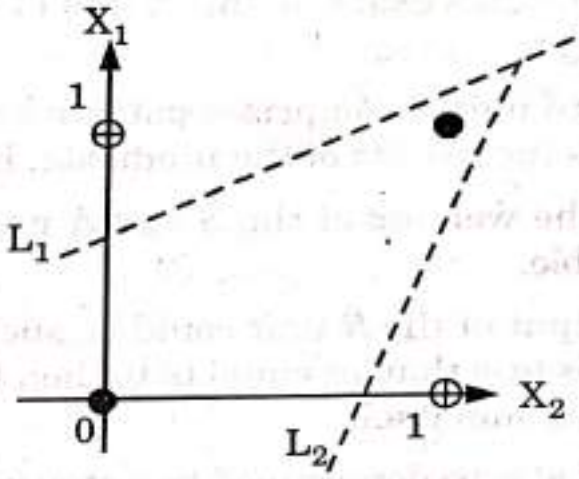


Fig. 2.6.2. Exclusive-OR Function.

Que 2.7. **Give artificial neural network architecture. What is Rosenblatt's perceptron model ?**

AKTU 2014-15(Sem-IV), Marks 05

Answer

Artificial neural network architecture : Refer Q. 1.14, Page 1–14F, Unit-1.

Rosenblatt's perceptron model :

1. The perceptron is a computational model of the retina of the eye and hence, is named as perceptron. The network comprises three units, the sensory unit S, association unit A, and response unit R.
2. The S unit comprising 400 photodetectors receives input images and provides a 0/1 electric signal as output. If the input signals exceed a threshold, then the photodetector outputs 1 else 0.

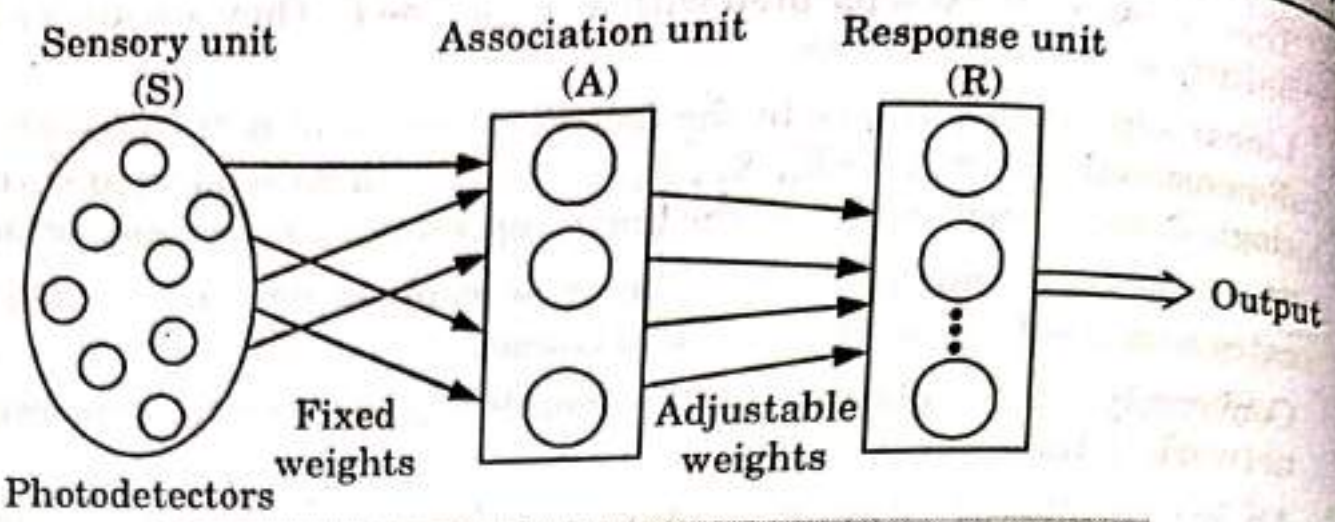


Fig. 2.7.1. Rosenblatt's original perceptron model.

3. The photodetectors are randomly connected to the association unit A.
4. The *A* unit comprises feature demons or predicates.
5. The predicates examine the output of the *S* unit for specific features of the image.
6. The third unit *R* comprises pattern recognizers or perceptrons, which receives the results of the predicate, in binary form.
7. While the weights of the *S* and *A* units are fixed, and those of *R* are adjustable.
8. The output of the *R* unit could be such that if the weighted sum of its inputs is less than or equal to 0, then the output is 0 otherwise it is the weighted sum itself.
9. It could also be determined by a step function with binary values (0/1) or bipolar values (–1/1).
10. Thus, in the case of a step function yielding 0/1 output values, it is defined as

$$y_j = f(\text{net}_j) = 1, \text{ if net}_j > 0$$
$$= 0, \text{ otherwise}$$

where $$\text{net}_j = \sum_{i=1}^{n} x_i w_{ij} \quad \ldots(2.7.1)$$

Here, x_i is the input, w_{ij} is the weight on the connection leading to the output units (*R* unit), and y_j is the output.

Que 2.8. **Explain the following Neural Network Architecture in detail :**

i. Rosenblatt's Perceptron Model

ii. McCulloch-Pitts Model **AKTU 2016-17(Sem-IV), Marks 15**

OR

Explain McCulloch-Pitts neuron model and write disadvantage of it.

AKTU 2015-16(Sem-III), Marks 10

AKTU 2015-16(Sem-IV), Marks 10

Answer

i. Rosenblatt's Perceptron Model : Refer Q. 2.7, Page 2–6F, Unit-2.

ii. McCulloch-Pitts Model :

1. McCulloch-Pitts neuron allows binary 0 or 1 states only, *i.e.*, it is binary activated.
2. These neurons are connected by direct weighted path. The connected path can be excitatory or inhibitory.
3. Excitatory connections have positive weights and inhibitory connections have negative weights.
4. There will be same weights for the excitatory connection entering into particular neuron. The neuron is associated with the threshold value.

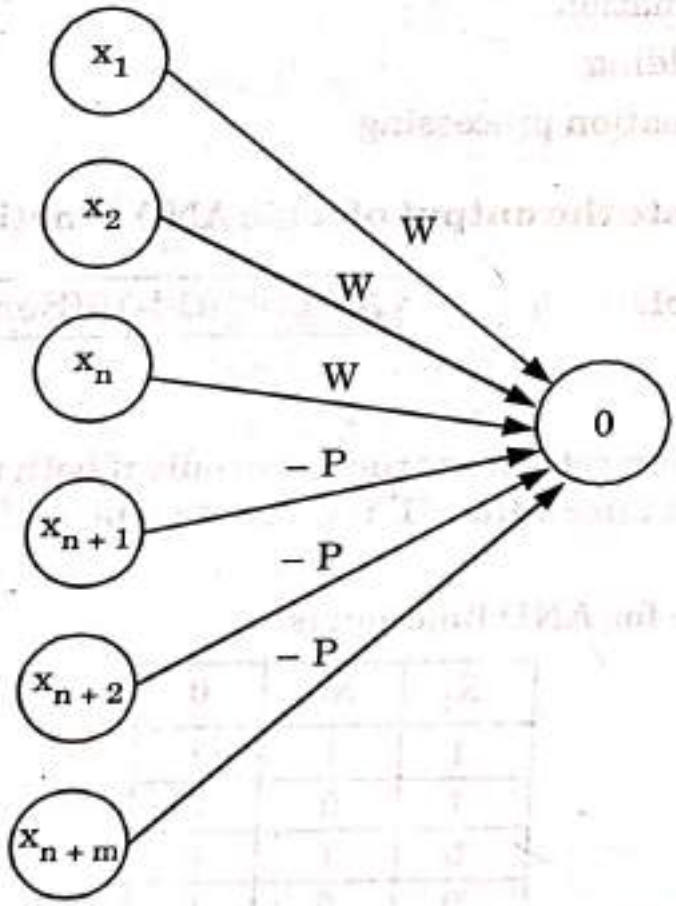


Fig. 2.8.1. Architecture of a McCulloch-Pitts neuron.

5. The neuron fires if the net input to the neuron is greater than the threshold.
6. The threshold is set so that the inhibition is absolute, because, non-zero inhibitory input will prevent the neuron from firing. It takes only one time step for a signal to pass over one connection link.

7. '0' is the McCulloch-Pitts neuron, it can receive signal from any number of other neurons.
8. The connection weights from $X_1, ..., X_n$ are excitatory, denoted by 'W' and the connection weights from $X_{n+1} .... X_{n+m}$ are inhibitory denoted by '$-P$'
9. The McCulloch-Pitts neurons Y has the activation function,

$$f(\text{net}) = 0^{k+1} = \begin{cases} 1 \text{ if } \sum_{i=1}^{n} W_i X_i k \geq T \\ 1 \text{ if } \sum_{i=1}^{n} W_i X_i k < T \end{cases}$$

where T is the threshold and net 0 is the total input signal received by the neuron 0.

Disadvantages of McCulloch-Pitts model : There are some features which are missing in McCulloch-Pitts neuron model which are :

1. Non-binary input and output
2. Non-linear summation
3. Smooth thresholding
4. Temporal information processing

Que 2.9. Generate the output of logic AND function by McCulloch Pitts neuron model.

AKTU 2014-15(Sem-III), Marks 05

Answer

1. The AND function returns a true value only if both the inputs are true, else it returns a false value. '1' represents true and '0' represents false value.
2. The truth table for AND function is

X_1	X_2	0
1	1	1
1	0	0
0	1	0
0	0	0

3. A McCulloch-Pitts neuron to implement AND function is shown in Fig. 2.9.1. The threshold on unit 0 is 2.

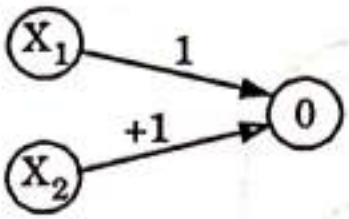


Fig. 2.9.1. McCulloch-Pitts neuron to perform AND function.

4. The output 0 is

$$0^{k+1} = f(\text{net})$$

The net input is given by

$$\text{net} = \sum_i \text{weight} \times \text{input}$$

$$\text{net} = 1 \times X_1 + 1 \times X_2$$

$$= X_1 + X_2$$

5. From this the activation of output neuron can be formed.

$$0^{k+1} = f(\text{net}) = \begin{cases} 1 \text{ if } \sum_{i=1}^{n} W_i X_i \; k \geq 2 \\ 1 \text{ if } \sum_{i=1}^{n} W_i X_i \; k < 2 \end{cases}$$

Now present the input

i. $X_1 = X_2 = 1, Y_{in} = X_1 + X_2 = 1 + 1 = 2$

$0^{k+1} = f(\text{net}) = 1$. Since net = 2

ii. $X_1 = 1, X_2 = 0, Y_{in} = X_1 + X_2 = 0 + 1$

$0^{k+1} = f(\text{net}) = 0$. Since net = $1 < 2$

This is same when $X_1 = 0$ and $X_2 = 1$

iii. $X_1 = 0, X_2 = 0, Y_{in} = X_1 + X_2 = 0 + 0 = 0$

Hence $0^{k+1} = f(\text{net}) = 0$. Since $Y_{in} = 0 < 2$.

Que 2.10. What is ADALINE ?

AKTU 2014-15(Sem-IV), Marks 05

Answer

1. ADALINE (Adaptive Linear Neural Element Network) is a neural network architecture that makes use of supervised learning.
2. In ADALINE, there is only one output neuron and the output values are bipolar (–1 or + 1). The inputs x_i could be binary, bipolar or real valued.

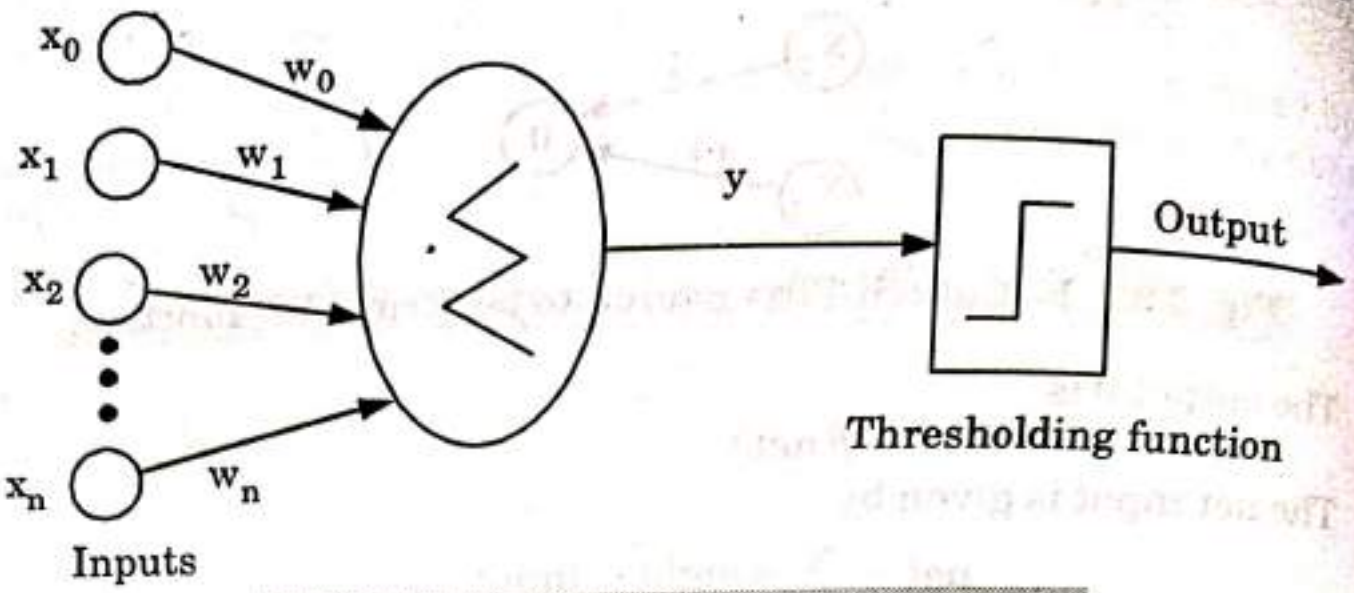


Fig. 2.10.1. Simple ADALINE network.

3. The bias weight is w_0 with an input link of $x_0 = +1$. If the weighted sum of inputs is greater than or equal to 0 then the output is 1 otherwise it is -1.

$$w_i^{new} = w_i^{old} + \alpha(t - y)\, x_i$$

α = Learning coefficient
t = Target output
y = Computed output
x_i = Input

4. The supervised learning algorithm adopted by the network is similar to the perception learning algorithm.
5. ADALINE network is successful and useful applications because it is used virtually in all high speed modems and telephone switching systems to cancel echo in long distance communications.

Que 2.11. **Write short note on Adaline and Madaline networks.**

AKTU 2015-16(Sem-III), Marks 05

Answer

Adaline network : Refer Q. 2.10, Page 2–10F, Unit-2.

MADALINE network :

1. A MADALINE (Many ADALINE) network is created by combining a number of ADALINES. The network of ADALINES can span many layers.
2. Fig. 2.11.1 illustrates a simple MADALINE network.
3. The learning rule adopted by MADALINE network is termed as 'MADALINE Adaptation Rule' (MR) and is a form of supervised learning.
4. In this method, the objective is to adjust the weights such that the error is minimum for the current training pattern, but with as little damage to the learning acquired through previous training patterns.
5. It solves the problem of non-linear separatibility.

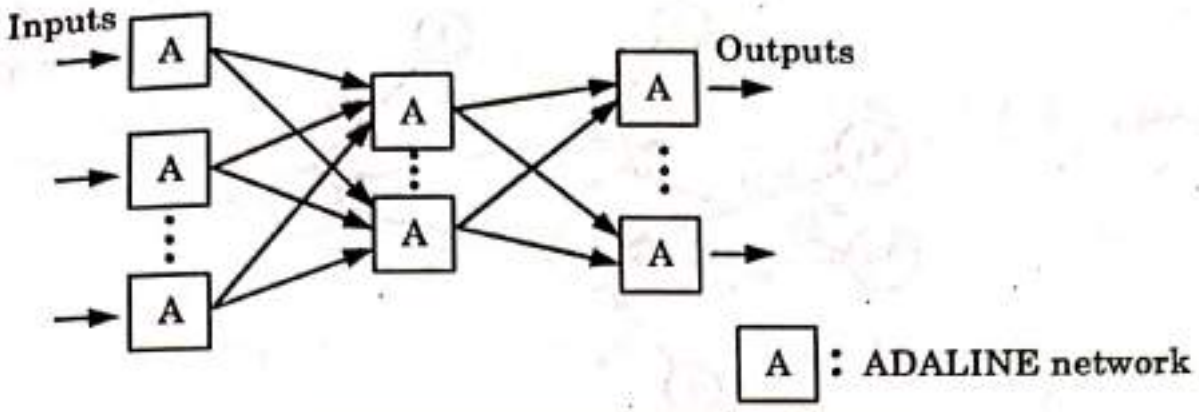


Fig. 2.11.1. MADALINE network.

Que 2.12. **Form a Madaline network for XOR function with bipolar input and targets using MR1 algorithm.**

AKTU 2017-18(Sem-III), Marks 10

OR

Can a two input Adaline compute the XOR function ? How will you solve the same by using Madaline ?

AKTU 2016-17(Sem-III), Marks 10

OR

Draw a network for solving Exclusive OR problem.

AKTU 2015-16(Sem-IV), Marks 10

Answer

The Boolean function XOR is not linearly separable as its positive and negative instances cannot be separated by a line or hyperplane. Hence a single layer perceptron can never compute the XOR function. Therefore a two input Adeline cannot compute the XOR function.

XOR function using Madaline :

The truth table for XOR function is

x_1	x_2	x_3
1	1	–1
1	–	1
–1	–1	–1

The architecture of the Madaline network can be given as :

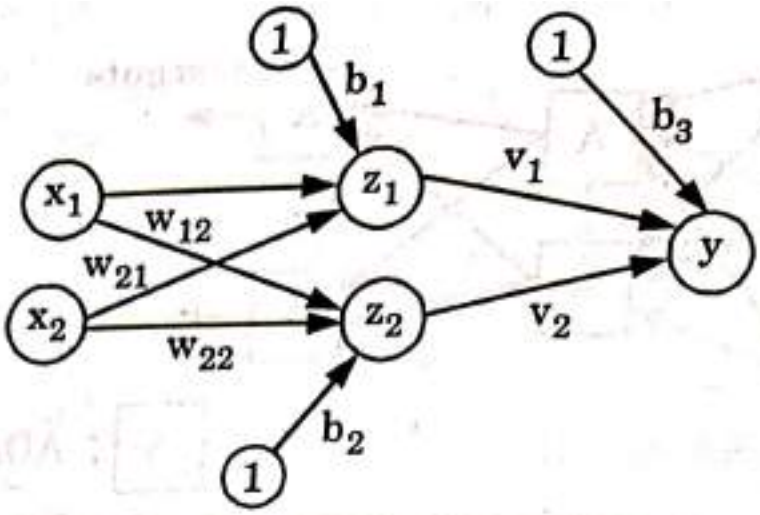


Fig. 2.12.1.

The algorithm dealt is followed and XOR function is generated. It is found that within 2 epochs, the weight converges.

Initial values :

Learning rate $\alpha = 0.5$

$$w_{11} = 0.05,\ w_{21} = 0.2,\ b_1 = 0.3$$
$$w_{12} = 0.1,\ w_{22} = 0.2,\ b_2 = 0.5$$
$$v_1 = v_2 = b_3 = 0.5$$

The training algorithm is followed.

For first epoch, first input pair (1, 1) : 1,

Step 1 : Weights and bias are initialized.

$$v_1 = v_2 = b_3 = 0.5$$
$$w_{11} = 0.05,\ w_{21} = 0.2,\ b_1 = 0.3$$
$$w_{12} = 0.1,\ w_{22} = 0.2,\ b_2 = 0.15$$
$$\alpha = 0.5$$

Step 2 : Begin training

Step 3 : For input and training pair (1, 1) : 1, perform steps 4 to step 8.

Step 4 : $x_i = S_i$

$$x = (1, 1)$$

Step 5 : Calculate the net input

$$z_{-in1} = b_1 + x_1 w_{11} + x_2 w_{21}$$
$$= 0.3 + 1 \times 0.05 + 1 \times 0.2$$
$$= 0.55$$
$$z_{-in2} = b_2 + x_1 w_{12} + x_2 w_{22}$$
$$= 0.15 + 1 \times 0.1 + 1 \times 0.2$$
$$= 0.45$$

Step 6 : Find output z_1 and z_2 apply activations

$$z_1 = f(z_{in1}) = 1$$
$$z_2 = f(z_{in2}) = 1$$

Step 7: Calculate net input of output

$$y_{-in} = b_3 + v_1 z_1 + v_2 z_2$$
$$= 0.5 + 1 \times 0.5 + 1 \times 0.5$$
$$= 1.5$$

Find the output, and apply activations.

$$y = f(y_{in}) = 1$$

Step 8 : Here $t = -1$ and $y = 1$

Hence $t \neq y$, so weights are updated on both z_1 and z_2 since both have positive net input

$$\begin{aligned} w_{11}\text{ (new)} &= w_{11}\text{ (old)} + \alpha\,(-1 - z_{in1}) \times 1 \\ &= 0.05 + 0.5\,(-1 - 0.55) \times 1 \\ &= -0.725 \\ w_{12}\text{ (new)} &= w_{12}\text{ (old)} + \alpha(-1 - z_{in2}) \times 2 \\ &= 0.1 + 0.5\,(-1 - 0.45)\,1 \\ &= -0.625 \\ b_1\text{ (new)} &= b_1\text{(old)} + \alpha(-1 - z_{in1}) \\ &= 0.3 + 0.5\,(-1 - 0.55) \\ &= -0.475 \\ w_{21}\text{(new)} &= w_{21}\text{(old)} + \alpha(-1 - z_{in1}) \times 2 \\ &= 0.2 + 0.5\,(-1 - 0.55)1 \\ &= -0.575 \\ w_{22}\text{ (new)} &= w_{22}\text{ (old)} + \alpha(-1 - z_{in2}) \times 2 \\ &= 0.2 + 0.5(-1 - 0.45)\,1 \\ &= -0.525 \\ b_2\text{ (new)} &= b_2\text{ (old)} + \alpha\,(-1 - z_{in2}) \\ &= 0.15 + 0.5\,(-1 - 0.45) = -0.575 \end{aligned}$$

After four epochs of training, the final weights are found to be :

$$w_{11} = -0.73 \qquad w_{12} = 1.27$$
$$w_{21} = 1.53 \qquad w_{22} = -1.33$$
$$b_1 = -0.99 \qquad b_2 = -1.09$$

Que 2.13. Draw and explain the multiple perceptron with its learning algorithm.

AKTU 2015-16(Sem-III), Marks 10

AKTU 2015-16(Sem-IV), Marks 05

Answer

1. The perceptrons which are arranged in layers are called multilayer (multiple) perceptron.
2. This model has three layers : an input layer, output layer and one or more hidden layer.
3. For the perceptrons in the input layer, the linear transfer function used and for the perceptron in the hidden layer and output layer, the sigmoidal or squashed-S function is used. The input signal propagates through the network in a forward direction.
4. In the multilayer perceptron bias $b(n)$ is treated as a synaptic weight driven by fixed input equal to + 1.

$$x(n) = [+1, x_1(n), x_2(n), \ldots\ldots\ldots x_m(n)]^T$$

where n denotes the interation step in applying the algorithm.

5. Correspondingly we define the weight vector as :

$$w(n) = [b(n), w_1(n), w_2(n)\ldots\ldots\ldots, w_m(n)]^T$$

6. Accordingly the linear combiner output is written in the compact form

$$V(n) = \sum_{i=0}^{m} w_i(n)x_i(n) = w^T(n)\, x(n)$$

Architecture of multilayer perceptron :

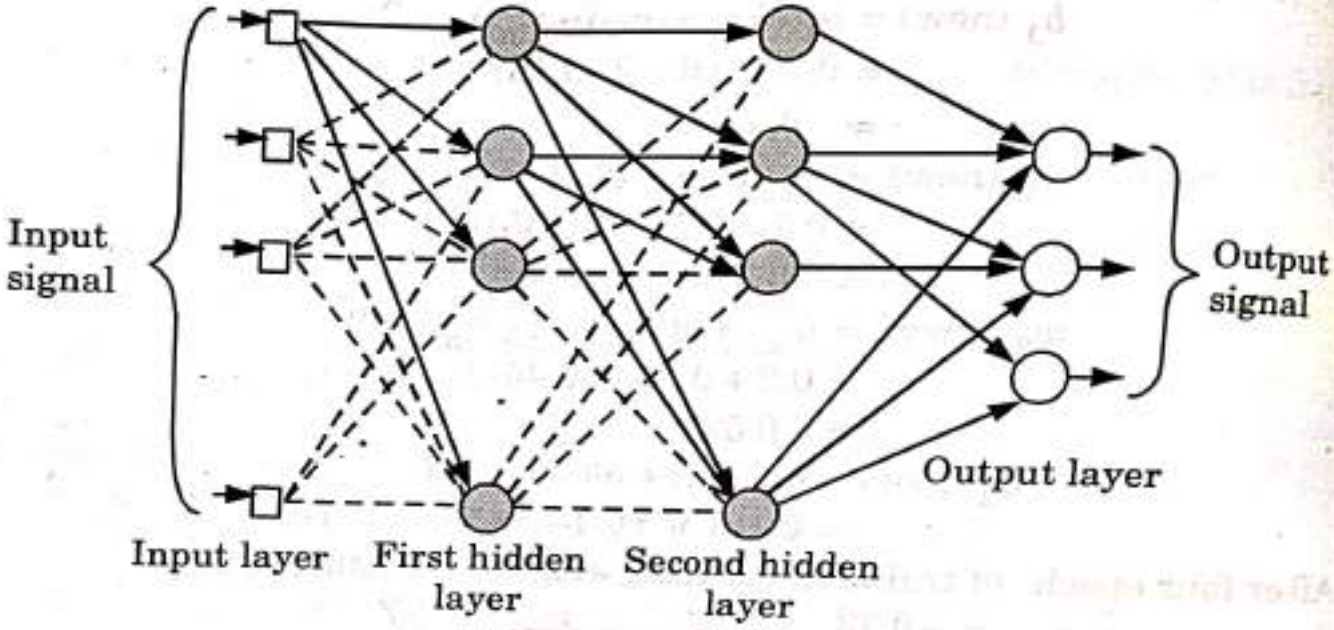


Fig. 2.13.1.

7. Fig. 2.13.1 shows the architectural model of multilayer perceptron with two hidden layer and an output layer.
8. Signal flow through the network progresses in a forward direction, from the left to right and on a layer-by-layer basis.

Learning algorithm :

1. If the nth number of input set $x(n)$, is correctly classified into linearly separable classes, by the weight vector $w(n)$ then no adjustment of weights are done.

$$w(n + 1) = w(n)$$

If $w^T x(n) > 0$ and $x(n)$ belongs to class G_1.

$$w(n + 1) = w(n)$$

If $w^T x(n) \leq 0$ and $x(n)$ belongs to class G_2.

2. Otherwise, the weight vector of the perceptron is updated in accordance with the rule.

Que 2.14. Which is the most common radial basis function ? Explain that function.

AKTU 2015-16(Sem-IV), Marks 10

AKTU 2017-18(Sem-III), Marks 10

Answer

1. A Radial Basis Function Network (RBFN) is a particular type of neural network :

   a. Gaussian function :

   $$\phi(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right) \text{ width parameter } \sigma > 0$$

   b. Multi-quadric function :

   $$\phi(r) = (r^2 + \sigma^2)^{1/2} \text{ parameter } \sigma > 0$$

   c. Generalized multi-quadric functions :

   $$\phi(r) = (r^2 + \sigma^2)^{\beta} \text{ parameters } \sigma > 0, 1 > \beta > 0$$

   d. Inverse multi-quadric function :

   $$\phi(r) = (r^2 + \sigma^2)^{-1/2} \text{ parameter } \sigma > 0$$

   e. Generalized inverse multi-quadric functions :

   $$\phi(r) = (r^2 + \sigma^2)^{-\alpha} \text{ parameters } \sigma > 0, \beta > 0$$

2. There are a number of choices for the radial basis function R_i, but Gaussian function is the most commonly used function.
3. Radial functions are a special class of function. Their characteristic feature is that their response decreases (or increases) monotonically with distance from a central point.
4. The centre, the distance scale, and the precise shape of the radial function are parameters of the model, all fixed if it is linear.
5. A typical radial function is the Gaussian which is

   $$h(x) = \exp\left(-\frac{(x-c)^2}{r^2}\right)$$

6. Its parameters are its centre c and its radius r.
7. A Gaussian RBF monotonically decreases with distance from the centre.

PART-2

Back Propagation Learning Methods, Effects of Learning Rule, Coefficient, Back Propagation Algorithm, Factors Affecting Back Propagation Training, Applications.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 2.15. **What are the back propagation learning methods ?**

AKTU 2013-14(Sem-III), Marks 05

Answer

Learning methods of back propagation :

1. **Static back propagation :**
   a. It is kind of back propagation network which produces a mapping of a static input for static output.
   b. It is useful to solve static classification issues like optical character recognition.
   c. The mapping is rapid in static back propagation.
2. **Recurrent back propagation :**
   a. Recurrent back propagation is feed forward until a fixed value is achieved.
   b. After that, the error is computed and propagated backward.
   c. It is non-static in recurrent back propagation.

Que 2.16. **Discuss the effect of learning rule co-efficient.**

AKTU 2013-14(Sem-III), Marks 05

Answer

1. The range of learning coefficient depends on the number and types of input patterns.
2. The learning coefficient cannot be negative because this would cause the change of weight vector to move away from ideal weight vector position.
3. If the learning coefficient is zero, no learning takes place and hence, the learning coefficient must be positive.
4. If the learning coefficient is greater than 1, the weight vector will overshoot from its ideal position and oscillate.
5. This coefficient must be smaller when there are many input patterns because the step length is controlled by learning coefficient.
6. A better selection of learning rate is possible if more information is available about the input pattern.

Que 2.17. **What are the drawbacks of back propagation algorithms ?**

AKTU 2013-14(Sem-III), Marks 05

OR

Write down the advantages and disadvantages of back propagation networks.

AKTU 2013-14(Sem-IV), Marks 05

Answer

Advantage of back propagation networks/algorithm :

1. It is fast, simple and easy to program.
2. It has no parameters to tune (except for the number of input).
3. The batch update of weight exists, which provides a smoothing effect on the weight of correction terms.
4. The computing time is reduced if weight chosen is small at the beginning.

Disadvantage/Drawbacks of back propagation network/algorithm :

1. The actual performance of back propagation on a particular problem is clearly dependent on the input data.
2. Back propagation can be sensitive to noisy data and outliers.
3. Fully matrix-based approach is used for back propagation instead of a mini-batch.
4. Once a network learns one set of weights, any new learning causes catastrophic forgetting.

Que 2.18. How hidden layer computation is done in back propagation learning ? Explain.

AKTU 2013-14(Sem-IV), Marks 05

Answer

1. Consider a sigmoidal function or squashed-S function, the output of the pth hidden neuron is given by

$$O_{Hp} = \frac{1}{(1+e^{-\lambda(I_{Hp}-\theta_{Hp})})} \quad \ldots(2.18.1)$$

where O_{Hp} is the output of the pth hidden neuron, I_{Hp} is the input of the pth hidden neuron, and θ_{Hp} is the threshold of the pth neuron.

2. A non-zero threshold neuron is computationally equivalent to an input that is always held at – 1 and the non-zero threshold becomes the connecting weight values as shown in Fig. 2.18.1.

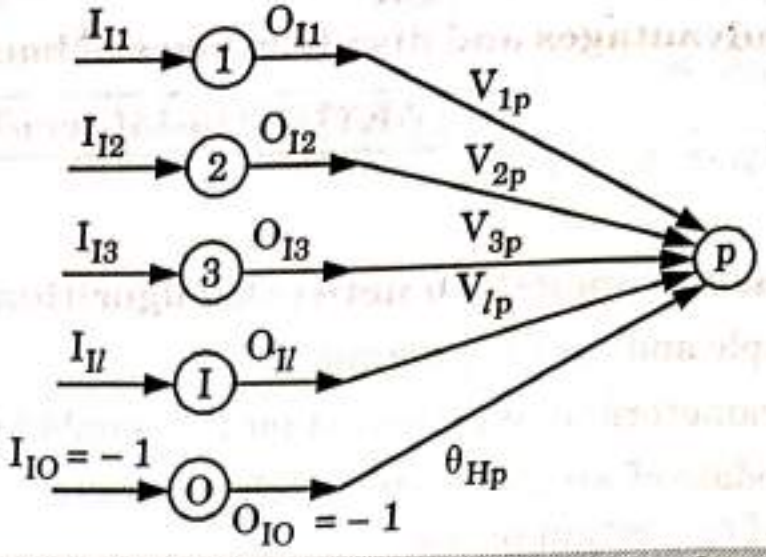


Fig. 2.18.1. Treating threshold in hidden layer.

But in our derivations we will not treat threshold as shown in Fig. 2.18.1. Now, output to the hidden neuron is given by

$$\{O\}_H = \begin{Bmatrix} - \\ - \\ \dfrac{1}{(1+e^{-\lambda(I_{Hp}-\theta_{Hp})})} \\ - \\ - \end{Bmatrix} \quad \text{...(2.18.2)}$$

3. Treating each component of the input of the hidden neuron separately, we get the output of the hidden neuron as given by eq. (2.18.1).

The input to the output neurons is the weighted sum of the outputs of the hidden neurons to get I_{Oq} (*i.e.*, the input to the qth output neuron)

$$I_{Oq} = W_{1q}O_{H1} + W_{2q}O_{H2} + \ldots + W_{mq}O_{Hm}$$
$$(q = 1, 2, 3, \ldots, n)$$

4. Denoting weight matrix or connectivity matrix between hidden neurons and output neurons as $[W]$, we can get input to the output neuron as

$$\{I\}_O = [W]^T \{O\}_H$$
$$n \times 1 \quad n \times m \quad m \times 1$$

Que 2.19. **State the importance of back propagation algorithm.**

AKTU 2013-14(Sem-IV), Marks 05

Answer

Importance of back propagation algorithm :

1. Backward propagation algorithm has high computational efficiency.
2. Backward propagation algorithm is an important mathematical tool for improving the accuracy of predictions in data mining and machine learning.

3. Artificial neural networks use back propagation as a learning algorithm to compute a gradient descent with respect to weights.
4. Along with classifiers such as Naive Bayesian filters and decision tress, the back propagation algorithm has emerged as an important part of machine learning applications that involve predictive analytics.

Que 2.20. Explain Back propagation learning algorithm in detail.

AKTU 2014-15(Sem-IV), Marks 05

OR

Write the algorithm for back propagation for back propagation training and explain about the updation of weight.

AKTU 2016-17(Sem-III), Marks 10

OR

Write the algorithm for back propagation training method.

AKTU 2017-18(Sem-IV), Marks 07

OR

Write various steps of the back propagation algorithm.

AKTU 2015-16(Sem-III), Marks 05

AKTU 2015-16(Sem-IV), Marks 05

Answer

Back propagation learning algorithm :

Step 1 : Normalize the inputs and outputs with respect to their maximum values. For each training pair, assume there are 'l' input given by $\underset{l \times 1}{\{I\}_I}$ and 'n' outputs $\underset{n \times 1}{\{O\}_O}$ in a normalized form.

Step 2 : Assume the number of neurons in the hidden layer lies between $l < m < 2l$.

Step 3 : $[V]$ Represents the weight of synapses connecting input neurons and hidden neurons and $[W]$ represents weights of synapses connecting hidden neurons and output neurons. Initialize the weights to small random values from – 1 to 1. For general problems, λ can be assumed as 1 and the threshold values can be taken as zero.

$$[V]^o = [\text{Random weights}]$$

$$[W]^o = [\text{Random weights}]$$

$$[\Delta V]^o = [DW]^o = [0] \quad \ldots(2.20.1)$$

Step 4 : For the training data, present one set of inputs and outputs by using linear activation function as :

$$\underset{l \times 1}{\{O\}_I} = \underset{l \times 1}{\{I\}_I} \qquad \ldots(2.20.2)$$

Step 5 : Compute the inputs to the hidden layer by multiplying corresponding weights of synapses as :

$$\underset{m \times 1}{\{I\}_H} = \underset{m \times l}{[V]^T} \cdot \underset{l \times 1}{\{O\}_I} \qquad \ldots(2.20.3)$$

Step 6 : Let the hidden layer units evaluate the output using the sigmoidal function as :

$$\{Q\}_H = \underset{m \times 1}{\begin{bmatrix} \cdot \\ \cdot \\ \dfrac{1}{(1+e^{-I_{Hi}})} \\ \cdot \\ \cdot \end{bmatrix}} \qquad \ldots(2.20.4)$$

Step 7 : Compute the inputs to the output layer by multiplying corresponding weights of synapses as

$$\underset{n \times 1}{\{I\}_O} = \underset{n \times m}{[W]^T} \quad \underset{m \times 1}{\{O\}^H} \qquad \ldots(2.20.5)$$

Step 8 : Let the output layer units evaluate the output using sigmoidal function as

$$\{O\}_O = \left\{ \begin{matrix} \cdot \\ \cdot \\ \dfrac{1}{(1+e^{-I_{oj}})} \end{matrix} \right\} \qquad \ldots(2.20.6)$$

Step 9 : Calculate the error and the difference between the network output and the desired output as for the ith training set as

$$E^p = \frac{\sqrt{\Sigma(T_j - O_{Oj})^2}}{n} \qquad \ldots(2.20.7)$$

Step 10 : Find $\{d\}$ as

$$\{d\} = \underset{n \times 1}{\begin{bmatrix} \cdot \\ \cdot \\ (T_k - O_{Ok})O_{Ok}(1 - O_{Ok}) \\ \cdot \\ \cdot \end{bmatrix}} \qquad \ldots(2.20.8)$$

Step 11 : Find [Y] matrix as

$$[Y] = \{O\}_H \quad \{d\} \qquad \text{...(2.20.9)}$$

$$m \times n \quad m \times 1 \; 1 \times n$$

Step 12 : Find $[\Delta W]^{t+1} = \alpha[\Delta W]^t + \eta[Y]$

$$m \times n \quad m \times n \quad m \times n$$

Step 13 : Find $\{e\} = [W] \quad \{d\}$

$$m \times 1 \quad m \times n \; n \times 1$$

$$\{d^*\} = \begin{Bmatrix} \cdot \\ \cdot \\ e_i(O_{Hi})(1 - O_{Hi}) \\ \cdot \\ \cdot \end{Bmatrix}_{m \times 1 \, m \times 1} \qquad \text{...(2.20.10)}$$

Find [X] matrix as :

$$[X] = \{O\}_I \quad \{d^*\} = \quad \{I\}_I \quad \{d^*\} \qquad \text{...(2.20.11)}$$

$$1 \times m \quad 1 \times 1 \quad 1 \times m \; 1 \times 1 \quad 1 \times m$$

Step 14 : Find $[\Delta V]^{t+1} = \alpha[\Delta V]^t + \eta[X]$

$$1 \times m \quad 1 \times m \quad 1 \times m$$

Step 15 : Weight update,

$$[V]^{t+1} = [V]^t + [\Delta V]^{t+1}$$

$$[W]^{t+1} = [W]^t + [\Delta W]^{t+1}$$

Step 16 : Find error rate as

$$\text{error rate} = \frac{\Sigma E_p}{nset}$$

Step 17 : Repeat steps 4-16 until the convergence in the error rate is less than the tolerance value.

Que 2.21. **Find new weights through backpropagation from the network shown in Fig. 2.21.1. Input pattern is (0.4, – 0.7) and target output is 0.1. Use the learning rate $\alpha = 0.25$ and Bipolar sigmoidal activation function.**

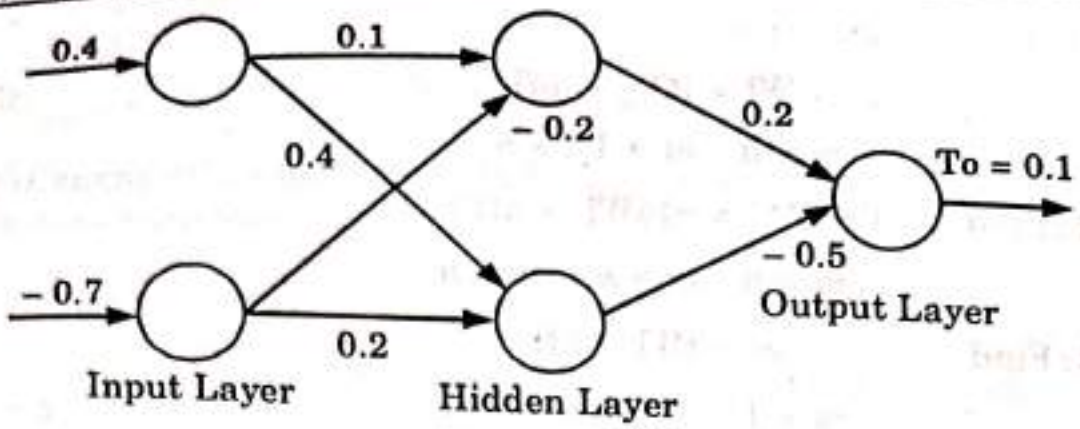


Fig. 2.21.1.

AKTU 2014-15(Sem-III), Marks 10

Answer

Step 1 : $[O]_1 = [I]_I = \begin{bmatrix} 0.4 \\ -0.7 \end{bmatrix}$

Step 2 : Initialize the weights as

$$[V]^0 = \begin{bmatrix} 0.1 & 0.4 \\ -0.2 & 0.2 \end{bmatrix}_{2\times 2} ; [W]^0 = \begin{bmatrix} 0.2 \\ -0.5 \end{bmatrix}_{2\times 1}$$

Step 3 : Find $[I]_H = [V]^T [O]_I$ as $= \begin{bmatrix} 0.1 & -0.2 \\ 0.4 & 0.2 \end{bmatrix} \begin{bmatrix} 0.4 \\ -0.7 \end{bmatrix} = \begin{bmatrix} 0.18 \\ 0.02 \end{bmatrix}$

Step 4 : $$[O]_H = \begin{bmatrix} \frac{1}{1+e^{-0.18\times 0.25}} \\ \frac{1}{1+e^{-0.02\times 0.25}} \end{bmatrix} = \begin{bmatrix} \frac{1}{1+e^{-0.0450}} \\ \frac{1}{1+e^{-0.050}} \end{bmatrix} = \begin{bmatrix} 0.5112 \\ 0.5125 \end{bmatrix}$$

Step 5 : $[I]_O = [W]^T [O]_H = (0.2 - 0.5) \begin{bmatrix} 0.5112 \\ 0.5125 \end{bmatrix} = -0.15401$

Step 6 : $$[O]_O = \left[\frac{1}{1+e^{0.15401}} \right] = 0.5384$$

Step 7 : Error $= (T_O - O_O)^2 = (0.1 - 0.5384)^2 = 0.19219$

Step 8 : Let us adjust the weights

First find $d = (T_O - O_{O1})(O_{O1})(1 - O_{O1})$

$= (0.1 - 0.5384)(0.5384)(0.4616) = -0.10895$

$[Y] = [O]_H (d)$

$$= \begin{bmatrix} 0.5112 \\ 0.5125 \end{bmatrix} (-0.10895) = \begin{bmatrix} -0.0556 \\ -0.0558 \end{bmatrix}$$

Step 9 : $[\Delta W]^1 = \alpha[W]^0 + \eta[Y]$ (Assume $\eta = 0.6$)

$$= 0.25 \begin{bmatrix} 0.2 \\ -0.5 \end{bmatrix} + 0.6 \begin{bmatrix} -0.0556 \\ -0.0558 \end{bmatrix} = \begin{bmatrix} -0.01664 \\ -0.09152 \end{bmatrix}$$

Step 10 : $[e] = [W]\,[d] = \begin{bmatrix} 0.2 \\ 0.5 \end{bmatrix}(-0.10895) = \begin{bmatrix} -0.02179 \\ -0.054475 \end{bmatrix}$

Step 11 : $[d^*] = \begin{bmatrix} (-0.02179) & (0.5112) & (1-0.5112) \\ (0.054475) & (0.5125) & (1-0.5125) \end{bmatrix}$

$$= \begin{bmatrix} -0.005444 \\ -0.01361 \end{bmatrix}$$

Step 12 : $[X] = [O]_I \langle d^* \rangle = \begin{bmatrix} 0.4 \\ -0.7 \end{bmatrix}[-0.005444 \quad -0.01361]$

$$= \begin{bmatrix} -0.002177 & 0.005444 \\ 0.0038108 & -0.009527 \end{bmatrix}$$

Step 13 : $[\Delta V]^1 = \alpha[\Delta V]^0 + \eta[X] = \begin{bmatrix} 0.023694 & 0.103266 \\ -0.04772 & -0.053266 \end{bmatrix}$

Step 14 : Updated weights

$$[V]_1 = \begin{bmatrix} 0.1 & 0.4 \\ -0.2 & 0.2 \end{bmatrix} + \begin{bmatrix} 0.023694 & 0.103266 \\ -0.04772 & -0.053266 \end{bmatrix}$$

$$= \begin{bmatrix} 0.12369 & 0.50326 \\ -0.24772 & 0.146734 \end{bmatrix}$$

$$[W]_1 = \begin{bmatrix} 0.2 \\ -0.5 \end{bmatrix} + \begin{bmatrix} 0.01664 \\ -0.09152 \end{bmatrix} = \begin{bmatrix} 0.18336 \\ -0.59152 \end{bmatrix}$$

Que 2.22. **If the net input to an output neuron is 0.64, calculate its output when the activation function is :**

i. **Binary sigmoidal, assume $\lambda = 1$**

ii. **Bipolar sigmoidal, assume $\lambda = 1$**

AKTU 2014-15(Sem-III), Marks 05

Answer

i. For binary activation function,

$$f(\text{net input}) = \frac{1}{1+e^{-\lambda \text{net}}} = \frac{1}{1+e^{-0.64}} = 0.6548$$

ii. For bipolar activation function,

$$f(\text{net input}) = \frac{2}{1+e^{-\lambda \text{net}}} - 1 = \frac{2}{1+e^{-0.64}} - 1 = 0.3095$$

Que 2.23. **Discuss the factors affecting the training of back propagation neural network.** **AKTU 2016-17(Sem-IV), Marks 10**

Answer

The training of a back propagation network is based on the choice of the various parameters. Also the convergence of the back propagation network is based on learning factors :

1. **Initial weights :** The ultimate solution may be affected by the initial weights of a multilayer feed-forward network. They are initialized at small random values.
2. **Learning rate :** The learning rate affects the convergence of the back propagation network. A large value of the learning rate may speed up the convergence but might result in overshooting, while a smaller value has vice-versa effect.
3. **Momentum factor :** The gradient descent is very slow if the learning rate is small, and oscillates widely if the value of learning rate is large. One very efficient and commonly used method that allows a large learning rate without oscillations is by adding a momentum factor to a normal gradient descent method.

Que 2.24. Discuss the selection of various parameters in BPN.

AKTU 2016-17(Sem-IV), Marks 10

Answer

Selection of various parameters in BPN (Back Propagation Network) :

1. **Number of hidden nodes :**
   i. The guiding criterion is to select the minimum nodes which would not impair the network performance so that the memory demand for storing the weights can be kept minimum.
   ii. When the number of hidden nodes is equal to the number of training patterns, the learning could be fastest.
   iii. In such cases, Back Propagation Network (BPN) remembers training patterns losing all generalization capabilities.
   iv. Hence, as far as generalization is concerned, the number of hidden nodes should be small compared to the number of training patterns (say 10:1).
2. **Momentum coefficient (α) :**
   i. The another method of reducing the training time is the use of momentum factor because it enhances the training process.

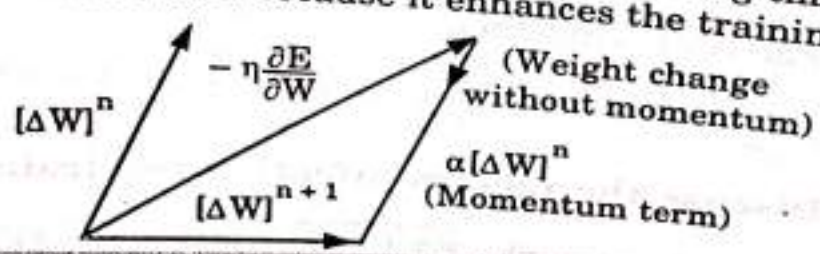


Fig. 2.24.1. Influence of momentum term on weight change.

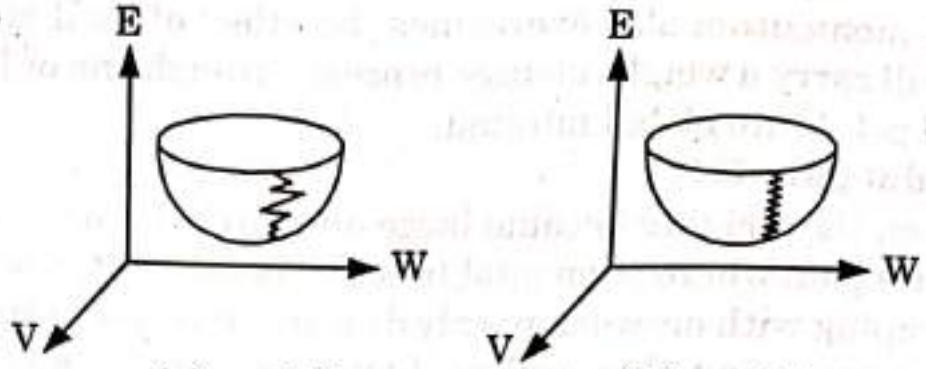
ii. The momentum also overcomes the effect of local minima.
iii. It will carry a weight change process through one or local minima and get it into global minima.

3. **Sigmoidal gain (λ) :**
   i. When the weights become large and force the neuron to operate in a region where sigmoidal function is very flat, a better method of coping with network paralysis is to adjust the sigmoidal gain.
   ii. By decreasing this scaling factor, we effectively spread out sigmoidal function on wide range so that training proceeds faster.
4. **Local minima :**
   i. One of the most practical solutions involves the introduction of a shock which changes all weights by specific or random amounts.
   ii. If this fails, then the solution is to re-randomize the weights and start the training all over.
   iii. Simulated annealing used to continue training until local minima is reached.
   iv. After this, simulated annealing is stopped and BPN continues until global minimum is reached.
   v. In most of the cases, only a few simulated annealing cycles of this two-stage process are needed.
5. **Learning coefficient (η) :**
   i. The learning coefficient cannot be negative because this would cause the change of weight vector to move away from ideal weight vector position.
   ii. If the learning coefficient is zero, no learning takes place and hence, the learning coefficient must be positive.
   iii. If the learning coefficient is greater than 1, the weight vector will overshoot from its ideal position and oscillate.
   iv. Hence, the learning coefficient must be between zero and one.

Que 2.25. Discuss how learning rate coefficient affects the back propagation training. **AKTU 2017-18(Sem-IV), Marks 10**

Answer

1. Learning rate coefficient determines the size of the weight adjustments made at each iteration and hence influences the rate of convergence.
2. Poor choice of the coefficient can result in a failure in convergence.
3. We should keep the coefficient constant through all the iterations for best results.
4. If the learning rate coefficient is too large, the search path will oscillate and converges more slowly than a direct descent as shown in Fig. 2.25.1(a).
5. If the coefficient is too small, the descent will progress in small steps significantly increasing the time to converge as shown in Fig. 2.25.1(b).



(a) $\eta = 1.5$ (b) $\eta = 0.2$

Fig. 2.25.1. Convergence paths for different learning coefficients.

Que 2.26. **What is learning rate ? What is its function ?**

AKTU 2015-16(Sem-IV), Marks 05

Answer

1. Learning rate is a constant used in learning algorithm that define the speed and extend in weight matrix corrections.
2. Setting a high learning rate tends to bring instability and the system is difficult to converge even to a near optimum solution.
3. A low value will improve stability, but will slow down convergence.

Learning function :

1. In most applications the learning rate is a simple function of time for example L.R. = $1/(1 + t)$.
2. These functions have the advantage of having high values during the first epochs, making large corrections to the weight matrix and smaller values later, when the corrections need to be more precise.
3. Using a fuzzy controller to adaptively tune the learning rate has the added advantage of bringing all expert knowledge in use.
4. If it was possible to manually adapt the learning rate in every epoch, we would surely follow rules of the kind listed below :
   a. If the change in error is small, then increase the learning rate.
   b. If there are a lot of sign changes in error, then largely decrease the learning rate.
   c. If the change in error is small and the speed of error change is small, then make a large increase in the learning rate.

Que 2.27. **Discuss some application areas of back propagation networks.**

AKTU 2013-14(Sem-IV), Marks 05

Answer

Application areas of back-propagation network :

1. Load forecasting problems in power systems
2. Image processing
3. Fault diagnosis and fault detection
4. Gesture recognition, speech recognition

5. Signature verification
6. Bioinformatics
7. Structural engineering design

Que 2.28. Find the weights required to perform the following classification using perceptron network. The vectors (1, 1, 1, 1) and (– 1, 1, – 1, – 1) are belonging to the class (so have target 1), vectors (1, 1, 1, – 1) and (1, – 1, – 1, 1) are not belonging to the class (so have target value – 1), Assume learning rate is 1 and weights is 0.

AKTU 2016-17(Sem-IV), Marks 10

Answer

1. The initial weights are assumed to be zero and the learning rate at 1. The updation is done according to perception learning rule.
2. If $y \neq t$, weight change, $\Delta w = \propto t x_i$ and $\Delta b = \propto t$
   New weights are,
$$w_{(new)} = w_{(old)} + \Delta w$$
$$b_{(new)} = b_{(old)} + \Delta b$$
3. If $t = y$, no weight change.
   By using formula we form the table.
   where, $y_{-in} = b + \sum_i x_i w_i$ and $y = f(y_{in})$ is the activation applied.

Input					Net	Output	Target	Weight Changes				Weights					
x_1	x_2	x_3	x_4	1	y_{in}	y	t	Δw_1	Δw_2	Δw_3	Δw_4	Δb	w_1	w_2	w_3	w_4	b
													(0	0	0	0	0)
Epoch 1 :																	
1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
1	1	1	–1	1	3	1	–1	–1	–1	–1	1	–1	0	0	0	2	0
–1	1	–1	–1	1	0	0	1	–1	1	–1	–1	1	–1	1	–1	–1	1
1	–1	–1	1	1	–1	–1	–1	0		0	0	0	–1	1	–1	–1	1
Epoch 2 :											initial →		0	0	0	2	0
1	1	1	1	1	2	1	1	0	0	0	0	0	0	0	0	2	0
1	1	1	–1	1	–2	–1	–1	0	0	0	0	0	0	0	0	2	0
Initial →													–1	1	–1	–1	1
–1	1	–1	–1	1	5	1	1	0	0	0	0	0	–1	1	–1	–1	1
1	–1	–1	1	1	–1	–1	–1	0	0	0	0	0	–1	1	–1	–1	1

5. The final weights from Epoch 1 are used as the initial weights for Epoch 2. Thus the output is equal to target by training for suitable weights.

VERY IMPORTANT QUESTIONS

Following questions are very important. These questions may be asked in your SESSIONALS as well as UNIVERSITY EXAMINATION.

Q. 1. Compare single layer and multilayer perceptron model.

Ans. Refer Q. 2.3.

Q. 2. Give artificial neural network architecture. What is Rosenblatt's perceptron model ?

Ans. Refer Q. 2.7.

Q. 3. Explain McCulloch-Pitts neuron model and write disadvantage of it.

Ans. Refer Q. 2.8.

Q. 4. What are the back propagation learning methods ?

Ans. Refer Q. 2.15.

Q. 5. How hidden layer computation is done in back propagation learning ? Explain.

Ans. Refer Q. 2.18.

Q. 6. What are the drawbacks of back propagation algorithms ?

Ans. Refer Q. 2.17.

Q. 7. Explain Back propagation learning algorithm in detail.

Ans. Refer Q. 2.20.

Q. 8. Discuss the selection of various parameters in BPN.

Ans. Refer Q. 2.24.

Q. 9. Discuss how learning rate coefficient affects the back propagation training.

Ans. Refer Q. 2.25.

Q. 10. Discuss some application areas of back propagation networks.

Ans. Refer Q. 2.27.

3
UNIT

Fuzzy Logic-I (Introduction)

CONTENTS

PART-1

Basic Concepts of Fuzzy Logic, Fuzzy Sets and Crisp Sets.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 3.1. **Define fuzzy logic and its importance in our daily life. What is role of crisp sets in fuzzy logic ?**

AKTU 2013-14(Sem-III), Marks 10

Answer

1. Fuzzy logic is an approach to computing based on "degrees of truth" rather than "true or false" (1 or 0).
2. Fuzzy logic includes 0 and 1 as extreme cases of truth but also includes the various states of truth in between.
3. Fuzzy logic allows inclusion of human assessments in computing problems.
4. It provides an effective means for conflict resolution of multiple criteria and better assessment of options.

Importance of fuzzy logic in daily life :

1. Fuzzy logic is essential for the development of human-like capabilities for AI.
2. It is used in the development of intelligent systems for decision making, identification, optimization, and control.
3. Fuzzy logic is extremely useful for many people involved in research and development including engineers, mathematicians, computer software developers and researchers.
4. Fuzzy logic has been used in numerous applications such as facial pattern recognition, air conditioners, vacuum cleaners, weather forecasting systems, medical diagnosis and stock trading.

Role of crisp sets in fuzzy logic :

1. It contains the precise location of the set boundaries.
2. It provides the membership value of the set.

Que 3.2. **Define classical set and fuzzy sets. State the importance of fuzzy sets.**

AKTU 2014-15(Sem-III), Marks 05

Answer

Classical set :

1. Classical set is a collection of distinct objects.
2. Each individual entity in a set is called a member or an element of the set.
3. The classical set is defined in such a way that the universe of discourse is splitted into two groups as members and non-members.

Fuzzy set :

1. Fuzzy set is a set having degree of membership between 1 and 0.
2. Fuzzy sets $\tilde{A}$ in the universe of discourse U can be defined as set of ordered pair and it is given by

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x) \mid x \in U)\}$$

Where $\mu_{\tilde{A}}$ is the degree of membership of x in $\tilde{A}$.

Importance of fuzzy set :

1. It is used for the modeling and inclusion of contradiction in a knowledge base.
2. It also increases the system autonomy.
3. It act as an important part of microchip processor-based appliances.

Que 3.3. Compare and contrast classical logic and fuzzy logic.

Answer

S. No.	Crisp (classical) logic	Fuzzy logic
1.	In classical logic an element either belongs to or does not belong to a set.	Fuzzy logic supports a flexible sense of membership of elements to a set.
2.	Crisp logic is built on a 2-state truth values (True/False).	Fuzzy logic is built on a multistate truth values.
3.	The statement which is either 'True' or 'False' but not both is called a proposition in crisp logic.	A fuzzy proposition is a statement which acquires a fuzzy truth value.
4.	Law of excluded middle and law of non-contradiction holds good in crisp logic.	Law of excluded middle and law of contradiction is violated.

Que 3.4. **What is the difference between crisp set and fuzzy set?**

AKTU 2015-16(Sem-IV), Marks 05

OR

Define a fuzzy set. How it is different from a crisp set ?

AKTU 2017-18(Sem-IV), Marks 07

Answer

Fuzzy set : Refer Q. 3.2, Page 3–2F, Unit-3.

S. No.	Crisp sets	Fuzzy sets
1.	A crisp set is a set containing elements that have fixed membership value.	A fuzzy set is a set containing elements that have varying degree of membership value in the set.
2.	Membership value is 1 (true) or 0 (false).	Membership values lies between 0 and 1.
3.	Law of excluded middle and non-contradiction hold : $A \cap \bar{A} = \phi$ $A \cup \bar{A} = \Omega$	Law of excluded middle and non-contradiction do not necessarily hold : $A \wedge \bar{A} \geq \phi$ $A \vee \bar{A} \leq \Omega$
4.	Intersection (AND), union (OR), and negation (NOT) operation are fixed.	Flexibility in choosing the intersection (T-norm), union (S-norm) and negation operations.

Que 3.5. **Write short notes on fuzzy arithmetic.**

Answer

1. Fuzzy arithmetic is generalization of interval arithmetic, where rather than considering intervals at one constant level only, several levels are considered in [0, 1].
2. This is because a fuzzy set allows degree of membership for an element of the universal set.
3. It plays an important role in many applications, such as fuzzy control, decision making, approximate reasoning, optimization and statistics with imprecise probabilities.

Que 3.6. **Explain different crisp sets with example.**

Answer

Different crisp sets are :

1. **Null set / empty set :**
   a. A set is said to be a null set or empty set if it has no member.
   b. A null set is indicated as ϕ or { } and indicates and impossible event.

   For example : The set of all prime ministers who are below 15 years of age.

2. **Singleton set :** A set with a single element is called a singleton set. A singleton set has cardinality of 1.

   For example : If $A = \{a\}$, then $|A| = 1$.

3. **Subset :**
   a. Give sets A and B defined over E the universal set, A is said to be subset of B if A is fully contained in B, that is, every element of A is in B.
   b. It is denoted as $A \subset B$, we say that A is a subset of B, or A is a proper subset of B.

   For example : Let $A = \{3, 4,\}$, $B = \{3, 4, 5\}$

   Here, $A \subset B$

4. **Superset :**
   a. Given sets A and B on E the universal set, A is said to be a superset of B if every element of B is contained in A.
   b. It is denoted as $A \supset B$, we say A is a superset of B or A contains B. If A contains B and is equivalent to B, then we denote it as $A \supseteq B$.

   For example : Let $A = \{3, 4\}$ $B = \{3, 4, 5\}$ and $C = \{4, 5, 6\}$

   Here, $A \subset B$, and $B \supset A$

   $C \subseteq B$, and $B \supseteq C$

5. **Power set :**
   a. A power set of a set A is the set of all possible subsets that are derivable from A including null set.
   b. A power set is indicated as $P(A)$ and has cardinality of $|P(A)| = 2^{|A|}$

   For example :

   Let $A = \{3, 4, 6, 7\}$

   $P(A) = \{\{3\}, \{4\}, \{6\}, \{7\}, \{3, 4\}, \{4, 6\}, \{6, 7\}, \{3, 7\}, \{3, 6\}, \{4, 7\}, \{3, 4, 6\}, \{4, 6, 7\}, \{3, 6, 7\}, \{3, 4, 7\}, \{3, 4, 6, 7\}, \phi\}$

   Here, $|A| = 4$ and $|P(A)| = 2^4 = 16$.

Que 3.7. Explain crisp set operations.

Answer

Operation on crisp set :

1. **Union ($\cup$) :** The union of two set A and B ($A \cup B$) is the set of all elements that belong to A or B or both.

$$A \cup B = \{x/x \in A \text{ or } x \in B\}$$

For example : Given $A = \{a, b, c, 1, 2\}$ and $B = \{1, 2, 3, a, c\}$, we get $A \cup B = \{a, b, c, 1, 2, 3\}$

2. **Intersection ($\cap$) :** The intersection of two set A and B ($A \cap B$) is the set of all elements that belong to A and B.

$$A \cap B = \{x/x \in A \text{ and } x \in B\}$$

For example : Given $A = \{a, b, c, 1, 2\}$ and $B = \{1, 2, 3, a, c\}$, we get $A \cap B = \{a, c, 1, 2\}$

3. **Complement (^{c}) :** The complement of a set A ($\bar{A}$ /A^c) is the set of all elements which are in E but not in A.

$$A^c = \{x/x \notin A \text{ or } x \in E\}$$

For example : Given $X = \{1, 2, 3, 4, 5, 6, 7,\}$ and $A = \{5, 4, 3\}$, we get $A^c = \{1, 2, 6, 7\}$

4. **Difference (–) :** The difference of the set A and B is $A - B$, the set of all elements which are in A but not in B

$$A - B = \{x \mid x \in A \text{ or } x \notin B\}$$

For example : Given $A = \{a, b, c, d, e\}$ and $B = \{b, d\}$, we get $A - B = \{a, c, e\}$

Que 3.8. Write short notes on

1. **Universe of discourse**
2. **Set**
3. **Venn diagram**
4. **Membership**
5. **Cardinality**

Answer

1. **Universe of discourse :** The universe of discourse or universal set is the set which, with reference to a particular context, contains all possible elements having the same characteristics and from which sets can be formed. The universal set is denoted by E.

For example :

i. The universal set of all number in Euclidean space.

ii. The universal set of all students in a university.

2. **Set :**
   a. A set is a well defined collection of object. The object either belongs to or does not belong to the set.
   b. A set in certain contexts may be associated with its universal set from which it is derived. Given a set A whose objects are $a_1, a_2, a_3, ..., a_n$, we write A as $A = \{a_1, a_2, ..., a_n\}$. Here, $a_1, a_2, ..., a_n$ are called the members of the set.

   For example :

   $A = \{x \mid x$ is an odd number$\}$

   $B = \{y \mid y > 0$ and $y \bmod 5 = 0\}$

3. **Venn diagram :** Venn diagrams are pictorial representations to denote a set. Given a set A defined over a universal set E, the Venn diagram for A and E is as shown in Fig. 3.8.1

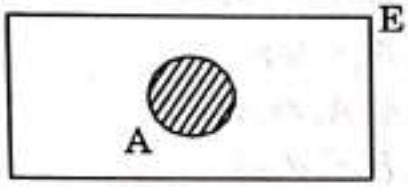


Fig. 3.8.1.

   For example : In Fig. 3.8.1, if E represents the set of university students then A may represent the set of female students.

4. **Membership :**
   a. An element x is said to be a member of a set A if x belongs to the set A.
   b. The membership is indicated by '$\in$' and is pronounced belongs to. Thus $x \in A$ means x belongs to A and $x \notin A$ means x does not belong to A.

   For example : Given $A = \{4, 5, 6, 7, 8, 10\}$ for $x = 3$ and $y = 4$, we have $x \notin A$ and $y \in A$.

5. **Cardinality :** The number of elements in the set is called its cardinality. It is denoted as $n(A)$ or $|A|$.

   For example : If $A = \{4, 5, 6, 7, 8\}$ then $n(A) = 5$.

Que 3.9. **Write note on partition and covering.**

AKTU 2014-15(Sem-IV), Marks 05

Answer

Partition :

1. A partition on set A is defined to be a set of non-empty subsets A_i, each of which is pair-wise disjoint and whose union yields the original set A.
2. Partition on set A is indicated as $\pi(A)$, therefore

i. $A_i \cap A_j = \phi$ for each pair $(i, j) \in I, i \neq j$

ii. $\bigcup_{i \in I} A_i = A$

3. The members A_i of the partition are known as blocks.

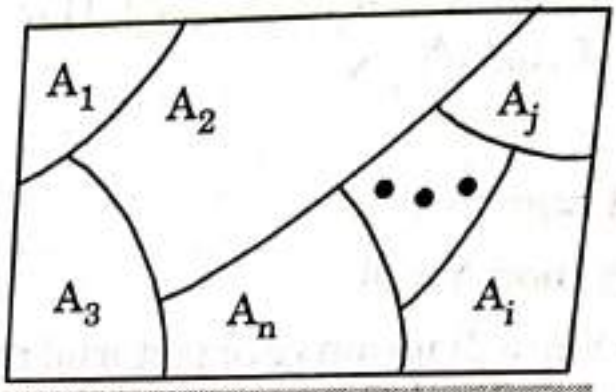


Fig. 3.9.1. Partition of set A.

Example : Given $A = \{a, b, c, d, e\}$

$A_1 = \{a, b\}, A_2 = \{c, d\}, A_3 = \{e\}$

$A_1 \cap A_2 = \phi, A_1 \cap A_3 = \phi, A_2 \cap A_3 = \phi$

$A_1 \cup A_2 \cup A_3 = A = \{a, b, c, d, e\}$

Hence $\{A_1, A_2, A_3\}$ is a partition on set A.

Covering :

1. A covering on set A is defined to be a set of non-empty subsets A_i, whose union yields the original set A.
2. The non-empty subsets need not to be disjoint.

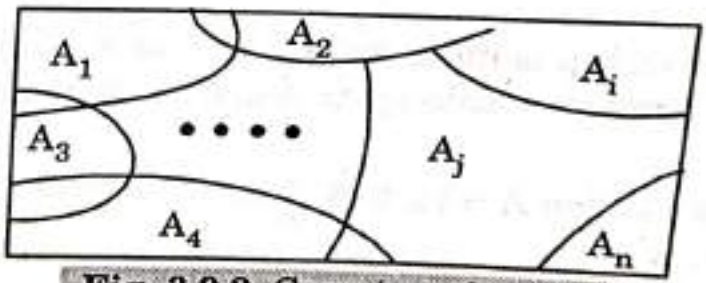


Fig. 3.9.2. Covering of set A.

Example : Given $A = \{a, b, c, d, e\}$

$A_1 = \{a, b\}$

$A_2 = \{b, c, d\}$

$A_3 = \{d, e\}$

$A_1 \cap A_2 = \{b\}$

$A_1 \cap A_3 = \phi$

$A_2 \cap A_3 = \{d\}$

$A_1 \cup A_2 \cup A_3 = A = \{a, b, c, d, e\}$.

Hence $\{A_1, A_2, A_3\}$ is a covering on set A.

PART-2

Fuzzy Set Theory and Operations, Properties of Fuzzy Set, Fuzzy and Crisp Relations, Fuzzy to Crisp Conversion.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 3.10. **What is fuzzy set theory ? Explain different fuzzy sets and its operations.**

AKTU 2013-14(Sem-III), Marks 10

OR

Explain the different types of operation used in Fuzzy set with suitable examples.

AKTU 2016-17(Sem-IV), Marks 10

OR

Discuss the basic fuzzy set operations.

AKTU 2014-15(Sem-IV), Marks 05

Answer

Fuzzy set theory :

1. Fuzzy set theory permits membership function valued in the interval [0, 1].
2. Fuzzy set is characterized by values from zero to one, with 0 representing no membership and 1 representing complete membership in a set *i.e.*, the grade of membership for all its members.

Fuzzy set operations : Let X to be the universal set and $\tilde{A}$ and $\tilde{B}$ are Fuzzy sets with $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$ as their respective membership function, the basic Fuzzy set operations are as follows :

1. **Union :** The union of two Fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a new Fuzzy set $\tilde{A} \cup \tilde{B}$ with a membership function defined as

$$\mu_{\tilde{A} \cup \tilde{B}}(x) = \max\left(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\right)$$

For example : Let $\tilde{A} = \{(x_1, 0.5), (x_2, 0.7), (x_3, 0)\}$

and $\tilde{B} = \{(x_1, 0.8), (x_2, 0.2), (x_3, 1)\}$

$$\tilde{A} \cup \tilde{B} = \{(x_1, 0.8), (x_2, 0.7), (x_3, 1)\}$$

Since, $\mu_{\tilde{A} \cup \tilde{B}}(x_1) = \max(\mu_{\tilde{A}}(x_1), \mu_{\tilde{B}}(x_1))$

$= \max(0.5, 0.8)$

$= 0.8$

$\mu_{\tilde{A} \cup \tilde{B}}(x_2) = \max(\mu_{\tilde{A}}(x_2), \mu_{\tilde{B}}(x_2)) = \max(0.2, 0.7) = 0.7$

$\mu_{\tilde{A} \cup \tilde{B}}(x_3) = \max(\mu_{\tilde{A}}(x_3), \mu_{\tilde{B}}(x_3)) = \max(0.1) = 1$

2. **Intersection :** The intersection of two Fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a new Fuzzy set $\tilde{A} \cap \tilde{B}$ with membership function defined as

$$\mu_{\tilde{A} \cap \tilde{B}}(x) = \min(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x))$$

For example : Let $\tilde{A} = \{(x_1, 0.5), (x_2, 0.7), (x_3, 0)\}$ and $\tilde{B} = \{(x_1, 0.5), (x_2, 0.7), (x_3, 0)\}$

then $\tilde{A} \cap \tilde{B} = \{(x_1, 0.5), (x_2, 0.2), (x_3, 0)\}$

since, $\mu_{\tilde{A} \cap \tilde{B}}(x_1) = \min(\mu_{\tilde{A}}(x_1), \mu_{\tilde{B}}(x_1))$

$= \min(0.5, 0.8) = 0.5$

$\mu_{\tilde{A} \cap \tilde{B}}(x_2) = \min(\mu_{\tilde{A}}(x_2), \mu_{\tilde{B}}(x_2))$

$= \min(0.7, 0.2) = 0.2$

$\mu_{\tilde{A} \cap \tilde{B}}(x_3) = \min(\mu_{\tilde{A}}(x_3), \mu_{\tilde{B}}(x_3))$

$= \min(0, 1) = 0$

3. **Complement :** The complement of a Fuzzy set $\tilde{A}$ is a new Fuzzy set $\tilde{A}$ with a membership function defined as

$$\mu_{\tilde{A}}(x) = 1 - \mu_{\tilde{A}}(x)$$

For example : Let $\tilde{A} = \{(x_1, 0.5), (x_2, 0.7), (x_3, 0)\}$

then $\tilde{A}^c = \{(x_1, 0.5), (x_2, 0.3), (x_3, 1)\}$

since, $\mu_{\tilde{A}^c}(x_1) = 1 - \mu_{\tilde{A}}(x_1) = 1 - 0.5 = 0.5$

$\mu_{\tilde{A}^c}(x_2) = 1 - \mu_{\tilde{A}}(x_2) = 1 - 0.7 = 0.3$

$\mu_{\tilde{A}^c}(x_3) = 1 - \mu_{\tilde{A}}(x_3) = 1 - 0 = 1$

4. **Product of two Fuzzy sets :** The product of two Fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a new Fuzzy set $\tilde{A} \cdot \tilde{B}$ whose membership function is defined as

$$\mu_{\tilde{A} \cdot \tilde{B}}(x) = \mu_{\tilde{A}}(x) \mu_{\tilde{B}}(x)$$

For example : Let $\tilde{A} = \{(x_1, 0.2), (x_2, 0.8), (x_3, 0.4)\}$

$\tilde{B} = \{(x_1, 0.4), (x_2, 0), (x_3, 0.1)\}$

$$\tilde{A} \cdot \tilde{B} = \{(x_1, 0.08), (x_2, 0), (x_3, 0.04)\}$$

Since
$$\mu_{\tilde{A}\cdot\tilde{B}}(x_1) = \mu_{\tilde{A}}(x_1) \cdot \mu_{\tilde{B}}(x_1)$$
$$= 0.2 \cdot 0.4 = 0.08$$
$$\mu_{\tilde{A}\cdot\tilde{B}}(x_2) = \mu_{\tilde{A}}(x_2) \cdot \mu_{\tilde{B}}(x_2)$$
$$= 0.8 \cdot 0 = 0$$
$$\mu_{\tilde{A}\cdot\tilde{B}}(x_3) = \mu_{\tilde{A}}(x_3) \cdot \mu_{\tilde{B}}(x_3)$$
$$= 0.4 \cdot 0.1$$
$$= 0.04$$

5. **Equality :** Two Fuzzy sets $\tilde{A}$ and $\tilde{B}$ are said to be equal ($\tilde{A} = \tilde{B}$) if $\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x)$

For example : Let $\tilde{A} = \{(x_1, 0.2), (x_2, 0.8)\}$

$$\tilde{B} = \{(x_1, 0.6), (x_2, 0.8)\}$$
$$\tilde{C} = \{(x_1, 0.2), (x_2, 0.8)\}$$
$$\tilde{A} \neq \tilde{B}$$

since $\mu_{\tilde{A}}(x_1) = \mu_{\tilde{B}}(x_1)$ although

$$\mu_{\tilde{A}}(x_2) = \mu_{\tilde{B}}(x_2)$$

but $\tilde{A} = \tilde{C}$

since $\mu_{\tilde{A}}(x_1) = \mu_{\tilde{C}}(x_1) = 0.2$

and $\mu_{\tilde{A}}(x_2) = \mu_{\tilde{C}}(x_2) = 0.8$

6. **Difference :** The difference of two Fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a new Fuzzy set $\tilde{A} - \tilde{B}$ defined as

$$\tilde{A} - \tilde{B} = (\tilde{A} \cap \tilde{B}^c)$$

For example : Let $\tilde{A} = \{(x_1, 0.2), (x_2, 0.5), (x_3, 0.6)\}$ and $\tilde{B} = \{(x_1, 0.1), (x_2, 0.4), (x_3, 0.5)\}$

$$\tilde{B}^c = \{(x_1, 0.9), (x_2, 0.6), (x_3, 0.5)\}$$
$$\tilde{A} - \tilde{B} = \tilde{A} \cap \tilde{B}^c$$
$$= \{(x_1, 0.2), (x_2, 0.5) (x_3, 0.5)\}$$

Que 3.11. **Explain the following terms :**

i. **Fuzzy arithmetic**
ii. **Fuzzy to crisp conversion**
iii. **Fuzzy relations**

AKTU 2013-14(Sem-III), Marks 10

OR

Write short notes on the following :

i. **Fuzzy relations**

ii. Fuzzy to crisp conversion **AKTU 2013-14(Sem-IV), Marks 10**

Answer

i. Fuzzy arithmetic : Refer Q. 3.5, Page 3–4F, Unit-3.

ii. Fuzzy to crisp conversion :

1. Defuzzification is the process of conversion of fuzzy set to crisp set.
2. It is a mapping from a space of fuzzy control action defined over an output universe of discourse into a space of non-fuzzy (crisp) control actions. It is most commonly used in fuzzy control systems.
3. These will have a number of rules that transform a number of variables into a fuzzy result, that is, the result is described in terms of membership in fuzzy sets.
4. Defuzzification is interpreting the membership degrees of the fuzzy sets into a specific decision or real value.

iii. Fuzzy relations :

1. Fuzzy relation is a fuzzy set defined on the Cartesian product of crisp sets $X_1, X_2, \ldots\ldots X_n$ where the n-tuples $(x_1, x_2, \ldots\ldots, x_n)$ may have varying degrees of membership within the relation.
2. The membership values indicate the strength of the relation between the tuples.

Example : Let R be the fuzzy relation between two sets X_1 and X_2 where X_1 is the set of diseases and X_2 is the set of symptoms.

X_1 = {typhoid, viral, common cold}

X_2 = {running nose, high temperature, shivering}

The fuzzy relation R may be defined as :

X_1 \ X_2	Running nose	High temperature	Shivering
Typhoid	0.1	0.9	0.8
Viral	0.2	0.9	0.7
Common cold	0.9	0.4	0.6

Que 3.12. **Let R be a crisp relation among the two sets X = {dollar, pound, franc, mark} and Y = {United States, France, Canada, Britain, Germany}, which associates a country with a currency. Represent it as a fuzzy relation using 2-dimensional membership array.**

AKTU 2015-16(Sem-IV), Marks 05

Answer

Let R be a crisp relation among the two sets X = {dollar, pound, franc, mark} and Y = {united states, France, Canada, Britain, Germany}, which associates a country with a currency as follows :

$R(X, Y)$ = {(dollar, United States),(franc, France),(dollar, Canada), (pound, Britain),(mark, Germany)}

This relation can also be represented by the following two-dimensional membership array :

Country / Currency	U.S.	France	Canada	Britain	Germany
dollar	1	0	1	0	0
pound	0	0	0	1	0
franc	0	1	0	0	0
mark	0	0	0	0	1

Que 3.13. **The task is to recognize English alphabetical characters (F, E, X, Y, I, T) in an image processing system. Two fuzzy sets $\tilde{I}$ and $\tilde{F}$ are defined to represent the identification of characters I and F.**

$\tilde{I}$ = {$(F, 0.4), (E, 0.3), (X, 0.1), (Y, 0.1), (I, 0.9), (T, 0.8)$}

$\tilde{F}$ = {$(F, 0.99), (E, 0.8), (X, 0.1), (Y, 0.2), (I, 0.5), (T, 0.5)$}

Find the following :

i. $\tilde{I} \cup \tilde{F}$

ii. $\tilde{I} - \tilde{F}$

iii. $\tilde{F} \cup \tilde{F}^c$

iv. **Verify de Morgan's law**

AKTU 2013-14(Sem-IV), Marks 10

AKTU 2014-15(Sem-IV), Marks 10

Answer

i. $\tilde{I} \cup \tilde{F}$: {$(F, 0.99), (E, 0.8), (X, 0.1), (Y, 0.2), (I, 0.9), (T, 0.8)$}

ii. $\tilde{I} - \tilde{F}$:

$$\tilde{I} - \tilde{F} = (\tilde{I} \cap \tilde{F}^c)$$

$\because$ $\tilde{I}$ = {$(F, 0.4), (E, 0.3), (X, 0.1), (Y, 0.1), (I, 0.9), (T, 0.8)$}

$\because$ $\tilde{F}$ = {$(F, 0.99), (E, 0.8), (X, 0.1), (Y, 0.2), (I, 0.5), (T, 0.5)$}

So, $\tilde{F}^c = \{(F, 0.01), (E, 0.2), (X, 0.9), (Y, 0.8), (I, 0.5), (T, 0.5)\}$

$\tilde{I} - \tilde{F} = \{(F, 0.01), (E, 0.2), (X, 0.1), (Y, 0.1), (I, 0.5), (T, 0.5)\}$

iii. $\tilde{F} \cup \tilde{F}^c$ **:**

$\because \quad \tilde{F} = \{(F, 0.99), (E, 0.8), (X, 0.1), (Y, 0.2), (I, 0.5), (T, 0.5)\}$

$\because \quad \tilde{F}^c = \{(F, .01), (E, 0.2), (X, 0.9), (Y, 0.8), (I, 0.5), (T, 0.5)\}$

$\tilde{F} \cup \tilde{F}^c = \{(F, 0.99), (E, 0.8), (X, 0.9), (Y, 0.8), (I, 0.5), (T, 0.5)\}$

iv. De Morgan's Law : De Morgan's law stated that :

$$(\tilde{I} \cup \tilde{F})^c = (\tilde{I}^c \cap \tilde{F}^c) \text{ or}$$

$$(\tilde{I} \cap \tilde{F})^c = (\tilde{I}^c \cap \tilde{F}^c)$$

So, now

$\because \quad (\tilde{I} \cup \tilde{F}) = \{(F, 0.99), (E, 0.8), (X, 0.1), (Y, 0.2), (I, 0.9), (T, 0.8)\}$

$$(\tilde{I} \cup \tilde{F})^c = 1 - (\tilde{I} \cup \tilde{F})$$

$(\tilde{I} \cup \tilde{F})^c = \{(F, 0.01), (E, 0.2), (X, 0.9), (V, 0.8), (I, 0.1), (T, 0.2)\}$

$(\tilde{I}^c \cap \tilde{F}^c) = \{(F, .01), (E, 0.2), (X, 0.9), (Y, 0.8), (I, 0.1), (T, 0.2)\}$

So, $(\tilde{I} \cup \tilde{F})^c = (\tilde{I}^c \cap \tilde{F}^c)$

Que 3.14. **Consider two fuzzy sets**

$\tilde{A} = \{(2, 1), (4, 0.3), (6, 0.5), (8, 0.2)\}$

$\tilde{B} = \{(2, 0.5), (4, 0.4), (6, 0.1), (8, 1)\}$

Perform union, intersection, difference and complement over fuzzy sets $\tilde{A}$ and $\tilde{B}$.

AKTU 2014-15(Sem-III), Marks 10

Answer

$\widetilde{A} = \{(2, 1), (4, 0.3), (6, 0.5), (8, 0.2)\}$

$\widetilde{B} = \{(2, 0.5), (4, 0.4), (6, 0.1), (8, 1)\}$

1. Union :

$\mu_{\widetilde{A} \cup \widetilde{B}}(2) = \max(\mu_{\widetilde{A}}(2), \mu_{\widetilde{B}}(2)) = \max(1, 0.5) = 1$

$\mu_{\widetilde{A} \cup \widetilde{B}}(4) = \max(\mu_{\widetilde{A}}(4), \mu_{\widetilde{B}}(4)) = \max(0.3, 0.4) = 0.4$

$\mu_{\widetilde{A} \cup \widetilde{B}}(6) = \max(\mu_{\widetilde{A}}(6), \mu_{\widetilde{B}}(6)) = \max(0.5, 0.1) = 0.5$

$\mu_{\widetilde{A} \cup \widetilde{B}}(8) = \max(\mu_{\widetilde{A}}(8), \mu_{\widetilde{B}}(8)) = \max(0.2, 1) = 1$

$\widetilde{A} \cup \widetilde{B} = \{(2, 1), (4, 0.4), (6, 0.5), (8,1)\}$

2. Intersection :

$$\mu_{\tilde{A}\cap\tilde{B}}(2) = \min(\mu_{\tilde{A}}(2), \mu_{\tilde{B}}(2)) = \min(1, 0.5) = 0.5$$
$$\mu_{\tilde{A}\cap\tilde{B}}(4) = \min(\mu_{\tilde{A}}(4), \mu_{\tilde{B}}(4)) = \min(0.3, 0.4) = 0.3$$
$$\mu_{\tilde{A}\cap\tilde{B}}(6) = \min(\mu_{\tilde{A}}(6), \mu_{\tilde{B}}(6)) = \min(0.5, 0.1) = 0.1$$
$$\mu_{\tilde{A}\cap\tilde{B}}(8) = \min(\mu_{\tilde{A}}(8), \mu_{\tilde{B}}(8)) = \min(0.2, 1) = 0.2$$
$$\tilde{A}\cap\tilde{B} = \{(2, 0.5), (4, 0.3), (6, 0.1), (8, 0.2)\}$$

3. Difference :

$$\widetilde{A} - \widetilde{B} = \widetilde{A}\cap\widetilde{B}^C$$
$$\widetilde{B}^C = \{(2, 0.5), (4, 0.6), (6, 0.9), (8, 0)\}$$
$$\mu_{\widetilde{A}\cap\widetilde{B}^C}(2) = \min(\mu_{\widetilde{A}}(2), \mu_{\widetilde{B}^C}(2)) = \min(1, 0.5) = 0.5$$
$$\mu_{\widetilde{A}\cap\widetilde{B}^C}(4) = \min(\mu_{\widetilde{A}}(4), \mu_{\widetilde{B}^C}(4)) = \min(0.3, 0.6) = 0.3$$
$$\mu_{\widetilde{A}\cap\widetilde{B}^C}(6) = \min(\mu_{\widetilde{A}}(6), \mu_{\widetilde{B}^C}(6)) = \min(0.5, 0.9) = 0.5$$
$$\mu_{\widetilde{A}\cap\widetilde{B}^C}(8) = \min(\mu_{\widetilde{A}}(8), \mu_{\widetilde{B}^C}(8)) = \min(0.2, 0) = 0$$
$$\widetilde{A} - \widetilde{B} = \{(2, 0.5), (4, 0.3), (6, 0.5), (8, 0)\}$$

4. Complement :

$$\widetilde{A}^C = 1 - \widetilde{A}$$
$$= \{(2, 0), (4, 0.7), (6, 0.5), (8, 0.8)\}$$
$$\widetilde{B}^C = 1 - \widetilde{B}$$
$$= \{(2, 0.5), (4, 0.6), (6, 0.9), (8. 0)\}$$

Que 3.15. **Consider the fuzzy sets A and B defined on the interval $X = [0, 5]$ of real numbers by the membership grade functions**

$$\mu_{\tilde{A}}(x) = \frac{x}{x+1}, \quad \mu_{\tilde{B}}(x) = 2^{-x}$$

Determine the mathematical formulae and graphs of the membership grade functions of each of the following sets :

a. A^c, B^c b. $A \cup B$

c. $A \cdot B$ d. $(A \cup B)^c$

AKTU 2014-15(Sem-IV), Marks 10

Answer

a. $$\mu_{\tilde{A}^c}(x) = 1 - \mu_{\tilde{A}}(x) = 1 - \frac{x}{x+1}$$
$$= \frac{1}{x+1}$$
$$\mu_{\tilde{B}^c}(x) = 1 - \mu_{\tilde{B}}(x) = 1 - 2^{-x} = \frac{2^x - 1}{2^x}$$

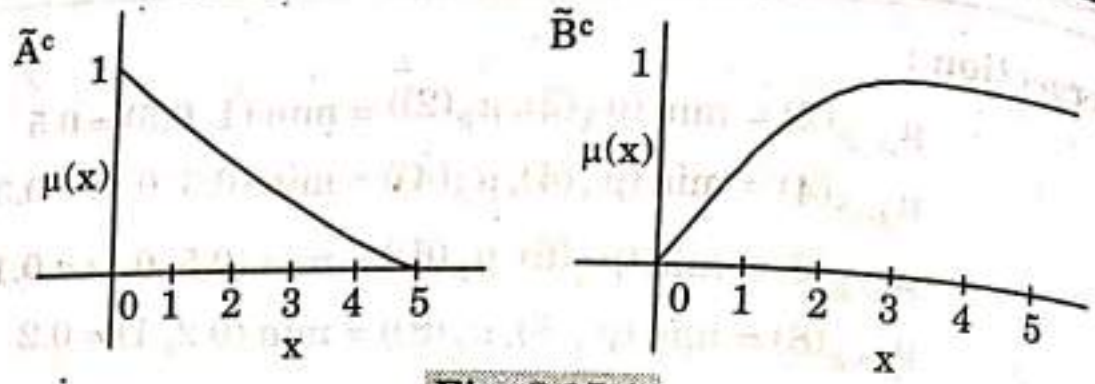


Fig. 3.15.1.

b. $\mu_{\tilde{A} \cup \tilde{B}}(x) = \max(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x))$

$$= \max\left(\frac{x}{x+1}, 2^{-x}\right)$$

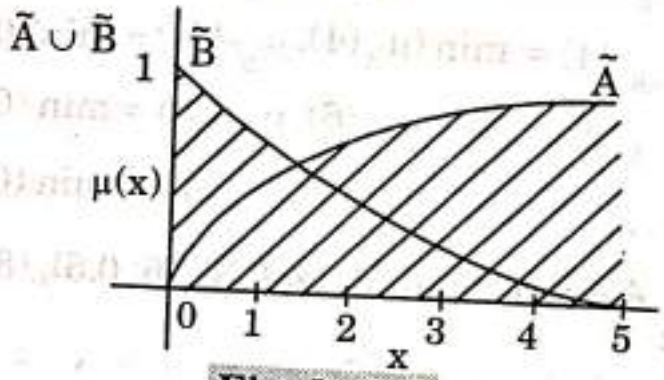


Fig. 3.15.2.

c. $\mu_{\tilde{A}.\tilde{B}}(x) = \frac{x}{x+1} \times 2^{-x} = \frac{x \cdot 2^{-x}}{x+1}$

d. $\mu_{(\tilde{A} \cup \tilde{B})^c}(x) = \mu_{\tilde{A}^c \cap \tilde{B}^c}(x)$ {∵ De Morgan's law}

$$= \min\left[\frac{1}{x+1}, \frac{2^x - 1}{2^x}\right]$$

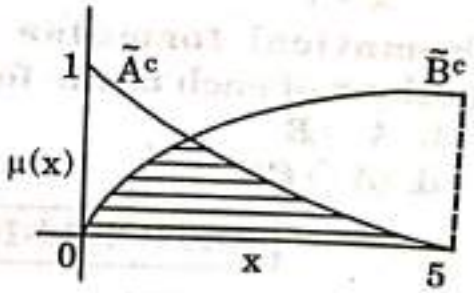


Fig. 3.15.3.

Que 3.16. **Consider three fuzzy sets given by :**

A = {(low, 1), (medium, 0.2), (high, 0.5)}
B = {(positive, 0.9), (zero, 0.4), (negative, 0.9)}
C = {(low, 0.1), (medium, 0.2), (high, 0.7)}

i. Find the fuzzy relation for the Cartesian product of A and B.

ii. Find CoR using max-min composition.

AKTU 2017-18(Sem-IV), Marks 07

Answer

i. $\tilde{R} = \tilde{A} \times \tilde{B}$

	Positive	Zero	Negative
Low	0.9	0.4	0.9
Medium	0.2	0.2	0.2
High	0.5	0.4	0.5

ii. $\tilde{C} \circ \tilde{R} = [0.1 \quad 0.2 \quad 0.7]_{1\times 3} \begin{bmatrix} 0.9 & 0.4 & 0.9 \\ 0.2 & 0.2 & 0.2 \\ 0.5 & 0.4 & 0.5 \end{bmatrix}_{1\times 3}$

$= [0.5 \quad 0.4 \quad 0.5]$

For instance,

$\mu_{\tilde{C}\circ\tilde{R}}(x_1, y_1) = \max[\min(0.1, 0.9), \min(0.2, 0.2), \min(0.7, 0.5)]$

$= \max(0.1, 0.2, 0.5)$

$= 0.5$

$\mu_{\tilde{C}\circ\tilde{R}}(x_2, y_2) = \max[\min(0.1, 0.4), \min(0.2, 0.2), \min(0.7, 0.7)]$

$= \max(0.1, 0.2, 0.4)$

$= 0.4$

$\mu_{\tilde{C}\circ\tilde{R}}(x_3, y_3) = \max[\min(0.1, 0.9), \min(0.2, 0.2), \min(0.7, 0.5)]$

$= \max(0.1, 0.2, 0.5)$

$= 0.5$

Que 3.17. What do you mean by fuzzy relation explain with the help of an example. For a speed control of DC motor the membership functions of series, resistance, armature current and speed are given as follows :

$$\tilde{R}_{se} = \{0.4/30, 0.6/60, 1.0/100, 0.1/120\}$$

$$I_a = \{0.2/20, 0.3/40, 0.6/60, 0.8/80, 1.0/100, 0.2/120\}$$

$$\tilde{N} = \{0.35/500, 0.67/1000, 0.97/1500, 0.25/1800\}$$

Compute relation T for relating series resistance to motor speed that is $\tilde{R}_{se}$ to $\tilde{N}$.

AKTU 2017-18(Sem-IV), Marks 07

Answer

Fuzzy relation : Refer Q. 3.11, Page 3–11F, Unit-3.

The operations needed to develop these relations uses two fuzzy Cartesian products and one composition :

$$\tilde{R} = \tilde{R}_{se} \times \tilde{I}_a$$

$$\tilde{S} = \tilde{I}_a \times \tilde{N}$$

$$\tilde{T} = \tilde{R} \, o \, \tilde{S}$$

$$\mu^{Rse}(\%se) = \frac{0.3}{30} + \frac{0.7}{60} + \frac{1.0}{100} + \frac{0.2}{120}$$

$$\mu^{Ia}(\%a) = \frac{0.2}{20} + \frac{0.4}{40} + \frac{0.6}{60} + \frac{0.8}{80} + \frac{1.0}{100} + \frac{0.1}{120}$$

The membership value for $\tilde{N}$ is given in units of motor speed in rpm as :

$$\mu^{N}(\text{rpm}) = \frac{0.33}{500} + \frac{0.67}{1000} + \frac{1.00}{1500} + \frac{0.15}{1800}$$

The following relation then uses Cartesian product to determine R.

$$\tilde{R} = \begin{array}{c} \\ 30 \\ 60 \\ 100 \\ 120 \end{array} \begin{array}{c} \begin{array}{cccccc} 20 & 40 & 60 & 80 & 100 & 120 \end{array} \\ \begin{bmatrix} 0.2 & 0.3 & 0.3 & 0.3 & 0.3 & 0.1 \\ 0.2 & 0.4 & 0.6 & 0.7 & 0.7 & 0.1 \\ 0.2 & 0.4 & 0.6 & 0.8 & 1 & 0.1 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.1 \end{bmatrix} \end{array}$$

For example, $\mu_{\tilde{R}}(60, 40) = \min(0.7, 0.4) = 0.4$, $\mu_{\tilde{R}}(100, 80) = \min(1.0, 0.8) = 0.8$,

Cartesian product to determine S :

$$\tilde{S} = \begin{array}{c} \\ 20 \\ 40 \\ 60 \\ 80 \\ 100 \\ 120 \end{array} \begin{array}{c} \begin{array}{cccc} 500 & 1000 & 1500 & 1800 \end{array} \\ \begin{bmatrix} 0.2 & 0.2 & 0.2 & 0.15 \\ 0.33 & 0.4 & 0.4 & 0.15 \\ 0.33 & 0.6 & 0.6 & 0.15 \\ 0.33 & 0.67 & 0.8 & 0.15 \\ 0.33 & 0.67 & 1 & 0.15 \\ 0.1 & 0.1 & 0.1 & 0.1 \end{bmatrix} \end{array}$$

$\mu_{\tilde{S}}(80, 1000) = \min(0.8, 0.67) = 0.67.$

The following relation result uses max–min composition for T :

$$\tilde{T} = \tilde{R} \, o \, \tilde{S} = \begin{matrix} \\ 30 \\ 60 \\ 100 \\ 120 \end{matrix} \begin{matrix} \begin{matrix} 500 & 1000 & 1500 & 1800 \end{matrix} \\ \begin{bmatrix} 0.3 & 0.3 & 0.3 & 0.15 \\ 0.33 & 0.67 & 0.7 & 0.15 \\ 0.33 & 0.67 & 1 & 0.15 \\ 0.2 & 0.2 & 0.2 & 0.15 \end{bmatrix} \end{matrix}$$

$\mu_{\tilde{T}}(60, 1500) = \max[\min(0.2, 0.2), \min(0.4, 0.4), \min(0.6, 0.6),$
$\min(0.7, 0.8), \min(0.7, 1.0), \min(0.1, 0.1)].$
$= \max[0.2, 0.4, 0.6, 0.7, 0.7, 0.1] = 0.7$

Que 3.18. Explain fuzzy set properties.

Answer

i. Let fuzzy set $\tilde{A}$ is a subset of the reference set X. Also, the membership of any element belonging to the null set ϕ is 0 and the membership of any element belonging to the reference set is 1. Following are the properties :

1. **Commutativity :** $\tilde{A} \cup \tilde{B} = \tilde{B} \cup \tilde{A}$

   $\tilde{A} \cap \tilde{B} = \tilde{B} \cap \tilde{A}$

2. **Associativity :** $\tilde{A} \cup (\tilde{B} \cup \tilde{C}) = (\tilde{A} \cup \tilde{B}) \cup \tilde{C}$

   $\tilde{A} \cap (\tilde{B} \cap \tilde{C}) = (\tilde{A} \cap \tilde{B}) \cap \tilde{C}$

3. **Distributivity :** $\tilde{A} \cup (\tilde{B} \cap \tilde{C}) = (\tilde{A} \cup \tilde{B}) \cap (\tilde{A} \cup \tilde{C})$

   $\tilde{A} \cap (\tilde{B} \cup \tilde{C}) = (\tilde{A} \cap \tilde{B}) \cup (\tilde{A} \cap \tilde{C})$

4. **Indempotence :** $\tilde{A} \cup \tilde{A} = \tilde{A}$

   $\tilde{A} \cap \tilde{A} = \tilde{A}$

5. **Identity :** $\tilde{A} \cup \varnothing = \tilde{A}$

   $\tilde{A} \cup X = \tilde{A}$

   $\tilde{A} \cap \varnothing = \varnothing$

   $\tilde{A} \cup X = X$

6. **Transitivity :** If $\tilde{A} \subseteq \tilde{B} \subseteq \tilde{C}$, then $\tilde{A} \subseteq \tilde{C}$

7. **Involution :** $(\tilde{A}^c)^c = \tilde{A}$

8. **De Morgan's Laws :** $(\tilde{A} \cap \tilde{B})^c = (\tilde{A}^c \cup \tilde{B}^c)$

   $(\tilde{A} \cup \tilde{B})^c = (\tilde{A}^c \cap \tilde{B}^c)$

Que 3.19. **Define fuzzy sets. Explain why the law of contradiction and the law of exclusive middle are violated in fuzzy set theory under the standard fuzzy sets operations. What is the significance of this ?**

AKTU 2013-14(Sem-IV), Marks 10

Answer

Fuzzy sets : Rcfer Q. 3.2, Page 3–2F, Unit-3.

1. Law of contradiction does not hold good for fuzzy sets. To verify this, we have to show that the equation

$$\min [A(x), \bar{A}(x)] = 0$$

is violated for at least one $x \in X$. The equation is obviously violated for any value $A(x) \in (0, 1)$. Hence, law of contradiction does not hold good in fuzzy sets.

2. Similarly, law of excluded middle does not hold good for fuzzy sets. To verify this, we have to show that the equation

$$\max [A(x), \bar{A}(x)] = 1$$

is violated for at least one $x \in X$. The equation is obviously violated for any value, $A(x) \in (0, 1)$. Hence, law of excluded middle does not hold good for fuzzy sets.

For example :

Let us consider a fuzzy set A on $X = \{1, 2, 3, 4\}$ is

$$A = \frac{0.1}{1} + \frac{0.7}{2} + \frac{0.5}{3} + \frac{1.0}{4}$$

then

$$\bar{A} = \frac{0.9}{1} + \frac{0.3}{2} + \frac{0.5}{3} + \frac{0.0}{4}$$

$$\min [A(x), \bar{A}(x)] = (A \cap \bar{A})(x) = \frac{0.1}{1} + \frac{0.3}{2} + \frac{0.5}{3} + \frac{0.0}{4}$$

Clearly, $(A \cap \bar{A})(x) \neq 0, \forall x \in X$

$$\max [A(x), \bar{A}(x)] = (A \cup \bar{A})(x) = \frac{0.9}{1} + \frac{0.7}{2} + \frac{0.5}{3} + \frac{0.0}{4}$$

Clearly, $(A \cup \bar{A})(x) \neq 1, \forall x \in X$

Therefore, law of contradiction $A \cap \bar{A} = \phi$ and law of exclusive middle $A \cup \bar{A} = X$ are violated.

Significance : Absence of these two laws has profound implication for fuzzy logic and serves to distinguish it from crisp set theory.

Que 3.20. **Verify De Morgan's Law using truth table (For 3 states).**

AKTU 2014-15(Sem-IV), Marks 05

Answer

De Morgan's Law for three states :

i. $(A \cup B \cup C)^c = A^c \cap B^c \cap C^c$

A	B	C	$A \cup B$	$A \cup B \cup C$	$(A \cup B \cup C)^c$	A^c	B^c	C^c	P
F	F	F	F	F	T	T	T	T	T
F	F	T	F	T	F	T	T	F	F
F	T	F	T	T	F	T	F	T	F
F	T	T	T	T	F	T	F	F	F
T	F	F	T	T	F	F	T	T	F
T	F	T	T	T	F	F	T	F	F
T	T	F	T	T	F	F	F	T	F
T	T	T	T	T	F	F	F	F	F

From the table it is clear that

$$(A \cup B \cup C)^C = P$$

Where, $P = A^C \cap B^C \cap C^C$

Hence proved.

ii. $(A \cap B \cap C)^c = A^c \cup B^c \cup C^c$

$A \cap B$	$A \cap B \cap C$	$(A \cap B \cap C)^C$	$A^C \cup B^C$	$A^C \cup B^C \cup C^C$
F	F	T	T	T
F	F	T	T	T
F	F	T	T	T
F	F	T	T	T
F	F	T	T	T
F	F	T	T	T
T	F	T	F	T
T	T	F	F	F

It is clear from table that

$$(A \cap B \cap C)^C = A^C \cup B^C \cup C^C$$

Hence proved.

VERY IMPORTANT QUESTIONS

Following questions are very important. These questions may be asked in your SESSIONALS as well as UNIVERSITY EXAMINATION.

Q. 1. Define fuzzy logic and its importance in our daily life. What is role of crisp sets in fuzzy logic ?

Ans: Refer Q. 3.1.

Q. 2. Define classical set and fuzzy sets. State the importance of fuzzy sets.

Ans: Refer Q. 3.2.

Q. 3. Explain different types of operations used in fuzzy sets with example.

Ans: Refer Q. 3.10.

Q. 4. Explain the following terms :

i. Fuzzy arithmetic

ii. Fuzzy to crisp conversion

iii. Fuzzy relations

Ans: Refer Q. 3.11.

Q. 5. Explain fuzzy set properties.

Ans: Refer Q. 3.18.

Q. 6. Define fuzzy sets. Explain why the law of contradiction and the law of exclusive middle are violated in fuzzy set theory under the standard fuzzy sets operations. What is the significance of this ?

Ans: Refer Q. 3.19.

Q. 7. Consider three fuzzy sets given by :

A = {(low, 1), (medium, 0.2), (high, 0.5)}

B = {(positive, 0.9), (zero, 0.4), (negative, 0.9)}

C = {(low, 0.1), (medium, 0.2), (high, 0.7)}

i. Find the fuzzy relation for the Cartesian product of A and B.

ii. Find CoR using max-min composition.

Ans: Refer Q. 3.16.

4

UNIT

Fuzzy Logic-II (Fuzzy Membership, Rules)

CONTENTS

PART-1

Membership Functions, Inference in Fuzzy Logic.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 4.1. Define the membership function and state its importance in fuzzy logic. Also discuss the features of membership functions.

AKTU 2013-14(Sem-IV), Marks 10

Answer

Membership function :

1. A membership function for a fuzzy set A on the universe of discourse X is defined as $\mu_A : X \rightarrow [0,1]$, where each element of X is mapped to a value between 0 and 1.
2. This value, called membership value or degree of membership, quantifies the grade of membership of the element in X to the fuzzy set A.
3. Membership functions characterize fuzziness (*i.e.*, all the information in fuzzy set), whether the elements in fuzzy sets are discrete or continuous.
4. Membership functions can be defined as a technique to solve practical problems by experience rather than knowledge.
5. Membership functions are represented by graphical forms.

Importance of membership function in fuzzy logic :

1. It allows us to graphically represent a fuzzy set.
2. It helps in finding different fuzzy set operation.

Features of membership function :

1. Core :

a. The core of a membership function for some fuzzy set $\tilde{A}$ is defined as that region of the universe that is characterized by complete and full membership in the set.

b. The core comprises those elements x of the universe such that $\mu_{\tilde{A}}(x) = 1$.

2. **Support :**

a. The support of a membership function for some fuzzy set $\tilde{A}$ is defined as that region of the universe that is characterized by nonzero membership in the set $\tilde{A}$.

b. The support comprises those elements x of the universe such that $\mu_{\tilde{A}}(x) > 0$.

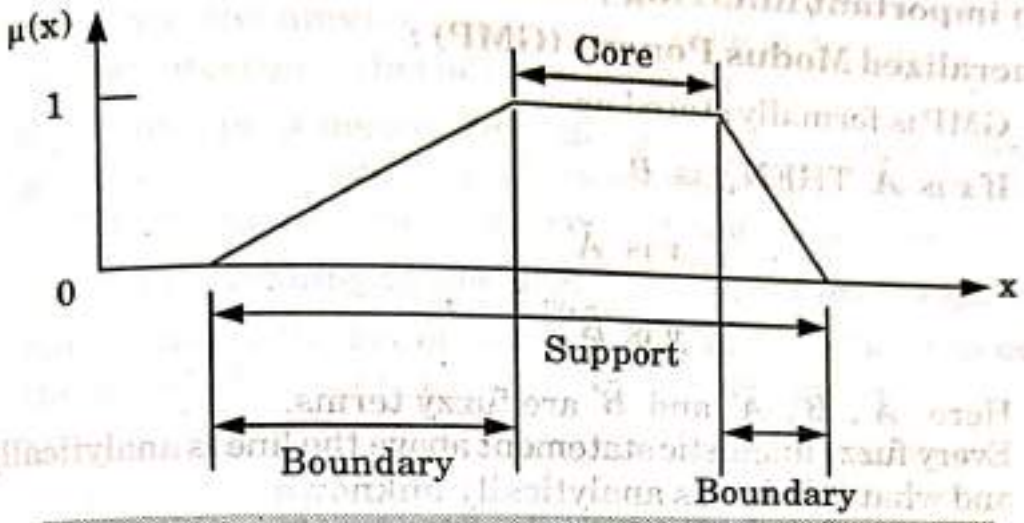


Fig. 4.1.1. Core, support, and boundaries of a fuzzy set.

3. **Boundaries :**

a. The boundaries of a membership function for some fuzzy set $\tilde{A}$ are defined as that region of the universe containing elements that have a non-zero membership but not complete membership.

b. The boundaries comprise those elements x of the universe such that $0 < \mu_{\tilde{A}}(x) < 1$.

Que 4.2. Explain membership function in fuzzy logic. What is the inference in fuzzy logic ? **AKTU 2013-14(Sem-III), Marks 10**

Answer

Membership function : Refer Q. 4.1, Page 4–2F, Unit-4.

Fuzzy Inference :

1. Inferences is a technique where facts, premises $F_1, F_2, \ldots\ldots, F_n$ and a goal G is to be derived from a given set.
2. Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic.
3. The mapping then provides a basis from which decisions can be made.
4. Fuzzy inference (approximate reasoning) refers to computational procedures used for evaluating linguistic (IF-THEN) descriptions.

5. The two important inferring procedures are :
   a. Generalized Modus Ponens (GMP).
   b. Generalized Modus Tollens (GMT).

Que 4.3. Explain two important inferring procedures.

Answer

The two important inferring procedures are :

i. Generalized Modus Ponens (GMP) :

1. GMP is formally stated as

   If x is $\tilde{A}$ THEN y is $\tilde{B}$

   $$\frac{x \text{ is } \tilde{A}'}{y \text{ is } \tilde{B}'}$$

   Here, $\tilde{A}$, $\tilde{B}$, $\tilde{A}'$ and $\tilde{B}'$ are fuzzy terms.
2. Every fuzzy linguistic statement above the line is analytically known and what is below is analytically unknown.

   Here $$\tilde{B}' = \tilde{A}' \circ \tilde{R}(x, y)$$

   where 'o' denotes max-min composition (IF-THEN relation)
3. The membership function is

   $$\mu_{\tilde{B}'}(y) = \max(\min(\mu_{\tilde{A}'}(x), \mu_{\tilde{R}}(x,y)))$$

   where $\mu_{\tilde{B}'}(y)$ is membership function of $\tilde{B}$, $\mu_{\tilde{A}'}(x)$ is membership function of $\tilde{A}'$ and $\mu_R(x,y)$ is the membership function of implication relation.

ii. Generalized Modus Tollens (GMT)

1. GMT is defined as

   If x is $\tilde{A}$. Then y is $\tilde{B}$

   $$\frac{y \text{ is } \tilde{B}'}{x \text{ is } \tilde{A}'}$$

2. The membership of $\tilde{A}'$ is computed as

   $$\tilde{A}' = \tilde{B}' \circ \tilde{R}(x, y)$$
3. In terms of membership function

   $$\mu_{\tilde{A}'}(x) = \max(\min(\mu_{\tilde{B}'}(y), \mu_{\tilde{R}}(x,y)))$$

Que 4.4. Compare and contrast classical logic and fuzzy logic.

Justify the following statement :
"Partial membership is allowed in fuzzy sets."

AKTU 2014-15(Sem-III), Marks 10

Answer

Compare classical logic and fuzzy logic : Refer Q. 3.3, Page 3–3F, Unit-3.

Justification :

1. In classical set theory, an element belongs to a set of all possible elements (the universe of discourse) and given any specific subset we can say precisely whether that element is or is not a member.
2. For example, a person belongs to the set of all human beings, and given a specific subset, such as all males, we can say whether or not each particular element (person) belongs to this set.
3. But not everything can be described using binary valued sets.
4. For example, the set of tall people is more difficult to define because there is no distinct cut-off point at which tall begins.
5. Now, we could simply select a height, say 1.8 m, at which the set tall begins. The output of a reasoning system built using this definition would not be smooth with respect to the height of a person, a person of height 1.799 m would produce a different output than a person of 1.801 m.
6. In fuzzy logic, human reasoning process use a relatively small number of rules and produce a smooth output via a process of interpolation.
7. It form rules based upon multivalued logic and so the concept of set membership is used.
8. In fuzzy logic an element could partially belong to a set and this was represented by the set membership.
9. In the given example, a person of height 1.799 m would belong to both tall and not tall sets with a particular degree of membership. As the height of a person increased the membership grade within the tall set would increase while the membership grade within the not tall set would decrease.
10. Fuzzy logic is simply the extension of conventional logic to the case where partial set membership can exist, rule conditions can be partially Fuzzy logic support a flexible sense of membership of elements to a set *i.e.*, in fuzzy logic many degrees of membership (between 0 and 1) are allowed.

Que 4.5. **Let sets of values of variables X and Y be $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2\}$, respectively. Assume that a proposition "if X is a, then Y is B" is given, where $A = .5/x_1 + 1/x_2 + .6/x_3$ and $B = 1/y_1 + .4/y_2$. Then, given a fact expressed by the proposition "x is A'",**

where $A' = .6/x_1 + .9/x_2 + .7/x_3$, use the generalized modus ponens to derive a conclusion in the form "Y is B".

AKTU 2013-14(Sem-IV), Marks 10

Answer

$$X = \{x_1, x_2, x_3\}$$

$$Y = \{y_1, y_2\}$$

$$A = \frac{0.5}{x_1} + \frac{1}{x_2} + \frac{0.6}{x_3}$$

$$B = \frac{1}{y_1} + \frac{0.4}{y_2}$$

$$A' = \frac{0.6}{x_1} + \frac{0.9}{x_2} + \frac{0.7}{x_3}$$

$$B' = A' \circ R(x, y)$$

$$\mu'_B(y) = \max(\min(\mu_A'(x)), \mu_R(x, y))$$

$$R(x, y) = \max(A \times B, \bar{A} \times Y)$$

$$A \times B = \begin{bmatrix} 0.5 & 0.4 \\ 1 & 0.4 \\ 0.6 & 0.4 \end{bmatrix}$$

$$\bar{A} \times Y = \begin{bmatrix} 0.5 & 0.5 \\ 0 & 0 \\ 0.4 & 0.4 \end{bmatrix}$$

$$R(x, y) = \begin{bmatrix} 0.5 & 0.5 \\ 1 & 0.4 \\ 0.6 & 0.4 \end{bmatrix}$$

$$\mu_B'(y) = \max((0.5, 0.9, 0.6), (0.5, 0.4, 0.4))$$

$$= (0.9, 0.5)$$

$$B' = \frac{0.9}{y_1} + \frac{0.5}{y_2}$$

Que 4.6. Write short notes on inference in propositional logic.

Answer

1. Inference is a technique by which, given a set of facts or postulates or axioms or premises $F_1, F_2, ..., F_n$, a goal G is to be derived.
2. In propositional logic, following three rules are used for inferring facts :

i. Modus ponens (mod pons) :

Given $P \Rightarrow Q$ and P to be true, Q is true.

$$\frac{\begin{array}{c} P \Rightarrow Q \\ P \end{array}}{Q}$$

Here, the formulae above the line are the premises and the one below is the conclusion which can be inferred from the premises.

ii. Modus tollens :

Given $P \Rightarrow Q$ and $\sim Q$ to be true, $\sim P$ is true.

$$\frac{\begin{array}{c} P \Rightarrow Q \\ \sim Q \end{array}}{\sim P}$$

iii. Chain rule :

Given $P \Rightarrow Q$ and $Q \Rightarrow R$ to be true, $P \Rightarrow R$ is true.

$$\frac{\begin{array}{c} P \Rightarrow Q \\ Q \Rightarrow R \end{array}}{P \Rightarrow R}$$

The chain rule is a representation of the transitivity relation with respect to the '$\Rightarrow$' connective.

Que 4.7. **Given :**

i. $C \vee D$ ii. $\sim H \Rightarrow (A \wedge \sim B)$

iii. $(C \vee D) \Rightarrow \sim H$ iv. $(A \wedge \sim B) \Rightarrow (R \vee S)$

Can $(R \vee S)$ be inferred from inference rules ?

AKTU 2014-15(Sem-IV), Marks 10

Answer

From (i) and (iii) using the rule of Modus Ponens, $\sim H$ can be inferred.

$$\frac{\begin{array}{c} (C \vee D) \Rightarrow \sim H \\ C \vee D \end{array}}{\sim H} \quad ...(4.7.1)$$

From (ii) and (iv) using chain rule, $\sim H \Rightarrow (R \vee S)$ can be inferred.

$$\frac{\begin{array}{c}\sim H \Rightarrow (A \wedge \sim B) \\ (A \wedge \sim B) \Rightarrow (R \vee S)\end{array}}{\sim H \Rightarrow (R \vee S)} \quad \ldots(4.7.2)$$

From (4.7.1) and (4.7.2) using Modus Ponens rule $(R \vee S)$ can be inferred.

$$\frac{\begin{array}{c}\sim H \Rightarrow (R \vee S) \\ \sim H\end{array}}{R \vee S}$$

Hence, $(R \vee S)$ be inferred from given inference rules.

Que 4.8. What are different attributes of predicate logic ? Using inference in predicate logic prove following statement :

i. All men are mortal.

ii. Socrates is a man.

Prove : Socrates is mortal. **AKTU 2016-17(Sem-IV), Marks 10**

Answer

Attributes of predicate logic :

1. Variables :

i. Variable are symbols which represent values acquired by the objects as quantified by the quantifier with which they are associated with.

ii. For example, if 'Hx'= 'x is happy', then Hx says only that x, whoever that is, is happy.

2. Constants :

i. Constants represent specific individuals or objects that do not change values.

ii. When a constant follows a predicate, the statement made by the predicate asserts that the individual named has a certain property.

iii. So, if 'Hx' = 'x is happy' and 'Txy' = 'x is taller than y' and 'a' be the individual constant for Alice and 'b' the individual constant for Bob, 'Hb' says that 'Bob is happy' and 'Tba' says that 'Bob is taller than Alice'.

3. Quantifiers :

i. Quantifiers are the symbol which indicates how many objects the predicate is asserted.

ii. If we want to assert a predicate of all objects, we use the universal quantifier,"$(\forall x)$". For example, "$(\forall x)Mx$" says that, for all x, x is mortal.

iii. If we want to assert a predicate of some objects (atleast one), we use the existential quantifier, "$(\exists x)$". For example, "$(\exists x)Mx$" says that, for some (atleast one) x, x is mortal.

iv. Quantifiers bind variables and thus convert expressions containing variables into sentences.

4. **Predicates :**

i. Predicate are representative of associations between objects that are constant or variables and acquire truth values 'True' of 'False'.

ii. A predicate carries a name representing the association followed by its arguments representing the objects it is to associate.

For example :

likes (Ram, tea) : Ram likes tea

plays (Sita, x) : Sita plays anything

Here, likes and plays are predicate names and Ram, tea and Sita, x are the associated objects.

Numerical :

Given :

i. All men are mortal.

ii. Socrates is mortal.

Proof :

Translating the above into predicate logic statement as :

i. $\forall x$ (man (x) $\Rightarrow$ mortal (x))

ii. man (Socrates)

iii. mortal (Socrates)

Since (i) is a tautology qualified by the universal quantifier for x = Socrates, the statement is true, *i.e.*,

man (Socrates) $\Rightarrow$ moral (Socrates)

$\Rightarrow$ ~ man (Socrates) $\vee$ moral (Socrates) [$\because p \to q = \sim p \vee q$]

But from (ii), man (Socrates) is true. [$\therefore F \vee P = P$]

Hence, false $\vee$ mortal (Socrates)

= mortal (Socrates)

Hence, Socrates is mortal has been proved.

Que 4.9. Define fuzziness of fuzzy set and what is a fuzzy function ?

AKTU 2017-18(Sem-III), Marks 10

Answer

1. Fuzziness of a fuzzy set directly comes from the membership function, *i.e.*, height and width of a fuzzy set which correspond to (a) and (b) in Fig. 4.9.1, respectively.

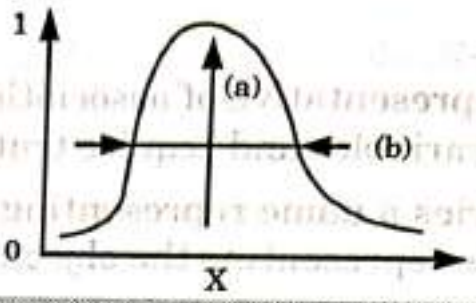


Fig. 4.9.1. Two types of fuzziness.

2. The fuzziness tends to be measured by the following two criteria :
   a. If the height of a fuzzy set is tall, the fuzziness is small.
   b. If the breadth of a fuzzy set is narrow, the fuzziness is small.
3. A fuzzy set should be judged to have small fuzziness even when the membership value becomes close to zero.
4. Since the criterion (a) is not precise enough to represent this fuzziness, the criterion (a) is rewritten as the following criteria (i) and (ii), where (i) is equivalent to (a) and (ii) is newly added.
   i. If the membership value is close to 1, the fuzziness is small.
   ii. If the membership value is close to 0, the fuzziness is small.
5. The breadth that the criterion (b) not suitable for the measure of fuzziness. Although the fuzzy set is often denoted as total ignorance, it is impossible to affirm that the fuzzy set is value. Therefore (b) should be removed from the criteria.
6. Finally, criteria (i) and (ii) help to measure the fuzziness of a fuzzy set.

Fuzzy functions :

1. Fuzzy functions consist of crisp function with fuzzy constraint and fuzzifying function.
2. Fuzzy function can be classified into following three groups :
   a. Crisp function with fuzzy constraint.
   b. Crisp function which propagates the fuzziness of independent variable to dependent variable.
   c. Function that is itself fuzzy.

Que 4.10. **Given a conditional and qualified Fuzzy proposition 'P' of the form *P* : If *x* is *A*, then *y* is *B* is S where 'S' is fuzzy truth qualifier and a fact is in the form "*x* is *A*" We want to make an**

inference in the form "y is B". Develop a method based on the truth-value restrictions for getting the inference.

AKTU 2016-17(Sem-III), Marks 10

Answer

1. A conditional and qualified fuzzy proposition p is the from

$$p : \text{If } x \text{ is } A, \text{ then } y \text{ is } B \text{ is } S \qquad \ldots(4.10.1)$$

where S is a fuzzy truth qualifier, and a fact is in the from "X is A", we want to develop an inference in the form "Y is B".

2. A method developed for this purpose, called a method of truth-value restrictions, is based on a manipulation of linguistic truth values.

3. The method involves the following four steps :

Step 1 : Calculate the relative fuzzy truth value of A' with respect to A, denoted by $RT(A'/A)$, which is a fuzzy set on the unit interval defined by

$$RT(A'/A)(a) = \sup_{x : A(x) \to a} A'(x), \forall\, a \in [0, 1] \qquad \ldots(4.10.2)$$

The relative fuzzy truth value $RT(A'/A)$ expresses the degree to which the fuzzy proposition (1) is true given the available fact "X is A'".

Step 2 : Choose a suitable fuzzy implication J by which the fuzzy proposition "If X is A, then y is B is S" is interpreted. This is similar to the selection of fuzzy implication :

$$J(a, b) = \min(1, 1 - a + b),$$

whose purpose is to express a conditional but unqualified fuzzy proposition as a fuzzy relation.

Step 3 : Calculate the relative fuzzy truth value $RT(B'/B)$ by using the formula :

$$RT(B'/B)(b) = \sup_{a \in [0,1]} \min[RT(A'/A)(a), S(J(a, b))], \forall\, b \in [0, 1]$$

where S is the fuzzy qualifier as in (1). The role of the qualifier S is to modify the truth value of $J(a, b)$, when S is true (*i.e.*, $S(a) = a$) for all $a \in [0,1]$, then

$$S(J(a, b)) = J(a, b).$$

The relative fuzzy truth value $RT(B'/B)$ expresses the degree to which the conclusion of the fuzzy proposition (1) is true.

Step 4 : Calculate the set B' involved in the inference "y is B'" by the equation

$$B'(y) = RT(B'/B)(B(y)), \forall\, y \in Y$$

PART-2

Fuzzy if Then Rules, Fuzzy Implications and Fuzzy Algorithm.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 4.11. Describe the following :

i. Fuzzy set properties

ii. Linguistic variable and membership function

AKTU 2017-18(Sem-IV), Marks 07

Answer

i. Fuzzy set properties : Refer Q. 3.18, Page 3–19F, Unit-3.

ii. Linguistic variable :

1. A linguistic variable is a variable whose values are words or sentences in a natural or artificial language.
2. When the fuzzy numbers represent linguistic concepts, such as very small, small medium, and so on, the resulting constructs are usually called linguistic variables.
3. Each linguistic variable interpreted as specific fuzzy numbers is defined in terms of a base variable, the values of which are real numbers within a specific range.
4. Each linguistic variable is fully characterized by a quintuple (v, T, X, g, m) in which v is the name of the variable, T is the set linguistic terms of v that refer to a base variable whose values range over a universal set X, g is syntactic rule (a grammar) for generating linguistic terms, and m is a semantic rule that assigns to each linguistic term $t \in T$ its meaning, $m(t)$ which is a fuzzy set on X (*i.e.*, $m : T \rightarrow F(X)$).

Membership function : Refer Q. 4.1, Page 4–2F, Unit-4.

Que 4.12. Define the term fuzzy IF-THEN rules.

OR

Define fuzzy IF-THEN rules with suitable examples.

AKTU 2014-15(Sem-IV), Marks 05

Answer

IF-THEN rule :

1. Fuzzy rules take the form :

   IF(conditions) and THEN (actions)

   where conditions and actions are linguistic variables, respectively.
2. These rules encode the knowledge about a system.
3. If THEN rule in the form of fuzzy set is given as

   IF $(x_1$ is $\tilde{A}_1, x_2$ is $\tilde{A}_2, \ldots, x_n$ is $\tilde{A}_n)$

   THEN $(y_1$ is $\tilde{B}_1, y_2$ is $\tilde{B}_2, \ldots, y_n$ is $\tilde{B}_n)$

   where linguistic variable x_i, x_j take the values of fuzzy sets $\tilde{A}_i$ and $\tilde{B}_i$ respectively.

   For example :

   IF temperature is very cold THEN air conditioner is off.

Que 4.13. Write short note on fuzzy implication and fuzzy algorithm.

Answer

Fuzzy implication :

1. A fuzzy implication (IF-THEN rule) assumes the form as :

   IF x is A, THEN y is B.

   Where A and B are linguistic values defined by fuzzy sets on Universe of Discourses (UoD) X and Y, respectively.
2. x is A called antecedent or premise, while y is B is called the consequence or conclusion.
3. Fuzzy implication is an important connective in fuzzy control systems because the control strategies are represented by sets of IF-THEN rules.
4. Fuzzy implication is denoted as

   $$R : A \rightarrow B$$

5. It represents a binary fuzzy relation R on the Cartesian product of $\tilde{A} \times \tilde{B}$.

Fuzzy algorithm : Fuzzy algorithm is a algorithm that includes atleast some fuzzy instructions such as conditional or unconditional action statements.

Que 4.14. **How fuzzy logic differs from crisp logic ? How rules are defined in fuzzy rule base system ?**

AKTU 2015-16(Sem-IV), Marks 10

Answer

Difference : Refer Q. 3.3, Page 3–3F, Unit-3.

Rules in fuzzy rule base system are defined by IF-THEN rule.

IF-THEN rule : Refer Q. 4.12, Page 4–12F, Unit-4.

PART-3

Fuzzifications and Defuzzifications.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 4.15. **Define the following terms :**

i. **Fuzzy algorithm**
ii. **Fuzzification**
iii. **Defuzzification**
iv. **Fuzzy if then rules**

AKTU 2013-14(Sem-III), Marks 10

Answer

i. **Fuzzy algorithm :** Refer Q. 4.13, Page 4–13F, Unit-4.

ii. **Fuzzification :**

1. Fuzzification is the method of transforming a crisp quantity into fuzzy quantity.
2. This can be achieved by identifying the various known crisp and deterministic quantities as completely non-deterministic and quite uncertain in nature.
3. This uncertainty may be emerged because of vagueness and imprecision which then leads the variables to be represented by a membership function.

iii. **Defuzzification :** Refer Q. 3.11, Page 3–11F, Unit-3.

iv. **Fuzzy if then rules :** Refer Q. 4.12, Page 4–12F, Unit-4.

Que 4.16. **Explain different methods of defuzzification.**

OR

Define the defuzzification. What are the different methods of defuzzification process ? Discuss one method in detail.

AKTU 2013-14(Sem-IV), Marks 10

OR

Explain the concept of defuzzification and discuss its method.

AKTU 2014-15(Sem-III), Marks 05

OR

Explain the method of converting fuzzy set into crisp set in detail.

AKTU 2014-15(Sem-III), Marks 10

OR

Define fuzzification and defuzzification with example.

AKTU 2014-15(Sem-IV), Marks 05

OR

Define the defuzzification. Explain methods of defuzzification process.

AKTU 2017-18(Sem-IV), Marks 07

Answer

Defuzzification : Refer Q. 3.11, Page 3–11F, Unit-3.

Different methods of duffuzzification process are :

a. **Center of Area (COA) :** The centroid defuzzification method selects the output crispy value corresponding to the center of gravity of the output membership function which is given by the expression :

$$x^* = \frac{\int w\mu(w)dw}{\int \mu(w)dw}$$

b. **Center of Sums (COS) :**

1. This is the most commonly used defuzzification technique.
2. In this method, the overlapping area is counted twice.
3. COS builds the resultant membership function by taking the algebraic sum of outputs from each of the contributing fuzzy sets $\tilde{A}_1, \tilde{A}_2, \ldots \tilde{A}_n$.
4. The defuzzified valued x^* is defined as :

$$x^* = \frac{\sum_{i=1}^{N} x_i \cdot \sum_{k=1}^{n} \mu_{\tilde{A}_k}(x_i)}{\sum_{i=1}^{N} \sum_{k=1}^{n} \mu_{\tilde{A}_k}(x_i)}$$

Here, n is the number of fuzzy sets, N is the number of fuzzy variable, $\mu_{\tilde{A}_k}(x_i)$ is the membership function for the k^{th} fuzzy set.

c. Height Method (HM) :

i. In the height method, the centroid of output membership function for each rule is evaluated first.

ii. The final output is then calculated as the average of the individual centroids, weighted by their heights (degree of membership) as follows :

$$x^* = \frac{\sum_{j=1}^{n} w_j \mu(w_j)}{\sum_{j=1}^{n} \mu(w_j)}$$

d. Middle of Maxima (MOM) :

i. The MOM strategy generates a control action which represents the mean value of all local control actions whose membership functions reach the maximum and may be expressed as :

$$x^* = \sum_{j=1}^{l} w_j / l$$

e. Center of Largest Area (COLA) :

i. The COLA method is used in the case when universe of discourse W is non-convex, *i.e.*, it consists of at least two convex fuzzy subsets.

ii. Then the method determines the convex fuzzy subset with the largest area and defines the crisp output value x^* to be the center of area of this particular fuzzy subset.

$$x^* = \frac{\int \mu_{\tilde{C}_m}(x)\, x\, dx}{\int \mu_{\tilde{C}_m}(x)\, dx}$$

f. Height Weighted Second Maxima (HWSM) :

i. In this method, the second maximum of each output membership function for each rule is evaluated first.

ii. The final output is calculated as the average of the individual maxima, weighted by their heights (degree of membership) as follows :

$$x^* = \frac{\sum_{j=1}^{n} w_j \mu(w_j)}{\sum_{j=1}^{n} \mu(w_j)}$$

where w_j takes the largest value of the domain with maximal membership degree.

Example of defuzzification :

Center of sum (COS) : The aggregated fuzzy set of two fuzzy sets C_1 and C_2 is shown in Fig. 4.16.1. Let the area of this two fuzzy sets are A_1 and A_2.

$$A_1 = ½ \times [(8 - 1) + (7 - 3)] \times 0.5 = ½ \times 11 \times 0.5$$
$$= 55/20 = 2.75$$

$$A_2 = ½ \times [(9-3)+(8-4)] \times 0.3 = ½ \times 10 \times 0.3$$
$$= 3/2 = 1.5$$

Now the center of area of the fuzzy set C_1 is

$$\bar{x}_1 = (7+3)/2 = 5$$

The center of area of the fuzzy set C_2 is

$$\bar{x}_2 = (8+4)/2 = 6.$$

Now the defuzzified value

$$x^* = \frac{(A_1 \bar{x}_1 + A_2 \bar{x}_2)}{A_1 + A_2} = \frac{(2.75 \times 5 + 1.5 \times 6)}{(2.75 + 1.5)}$$
$$= 22.75 / 4.25 = 5.35$$

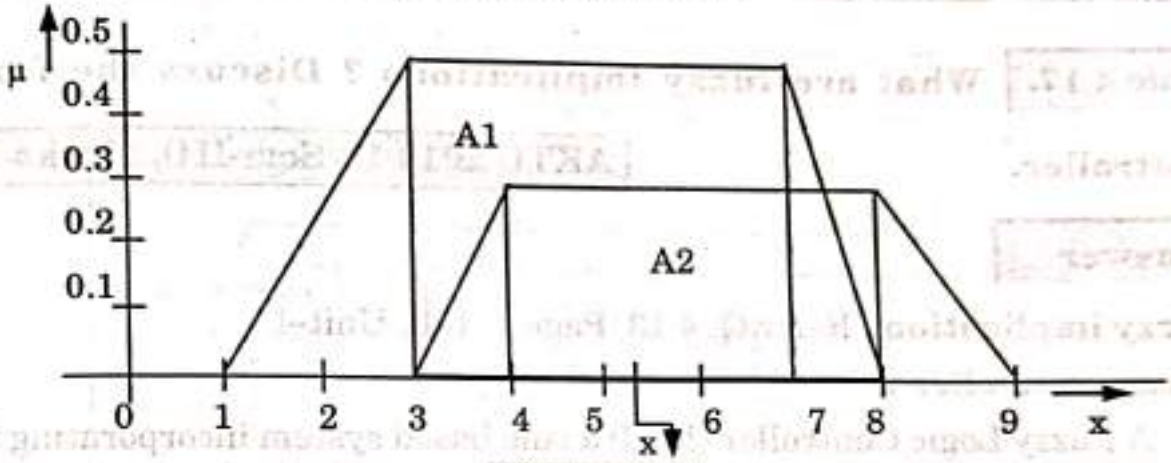


Fig. 4.16.1.

Fuzzification : Refer Q. 4.15, Page 4–14F, Unit-4.

Example of fuzzification :

1. Suppose at some time t, the temperature is 42 degrees and fan speed is 31 rpm. The corresponding membership values and the fuzzy regions are shown in Table 4.16.1.

Table 4.16.1.

S.No.	Parameter	Fuzzy regions	Memberships
1.	Temperature	Warm, Hot	0.142, 0.2
2.	Fan speed	Medium, High	0.25, 0.286

2. From 1, the temperature 42 degrees correspond to two membership values 0.142 and 0.2 that belong to Warm and Hot fuzzy regions respectively.
3. Similarly from 2, the fan speed 31 rpm corresponds to two membership values 0.25 and 0.286 that belong to Medium and High fuzzy regions respectively.
4. From Table 4.16.1, there are four combinations possible :
   a. If temperature is Warm and fan speed is Medium.
   b. If temperature is Warm and fan speed is High.
   c. If temperature is Hot and fan speed is Medium.

d. If temperature is Hot and fan speed is High.

PART-4

Fuzzy Controller, Industrial Applications.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 4.17. What are fuzzy implications ? Discuss the fuzzy controller.

AKTU 2013-14(Sem-III), Marks 10

Answer

Fuzzy implication : Refer Q. 4.13, Page 4–13F, Unit-4.

Fuzzy controller :

1. A Fuzzy Logic Controller (FLC) a rule based system incorporating the flexibility of human decision making is used for fuzzy structural optimization.
2. The fuzzy functions are intended to represent a human expert's conception of the linguistic terms.

Components of Fuzzy Logic Controller (FLC) :

Fuzzy logic controller process is divided into three stages as shown in Fig. 4.17.1.

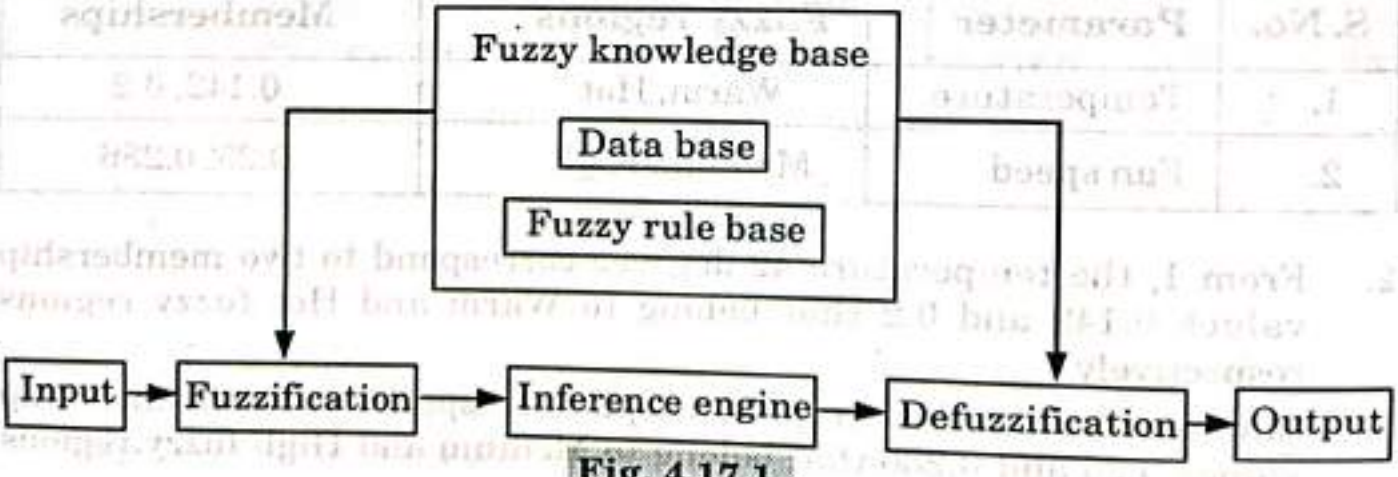


Fig. 4.17.1.

a. **Fuzzification :** It is a process evaluating the input variable with respect to corresponding linguistic terms in the condition side.

b. **Fuzzy inference :** It is a process evaluating the activation strength of every rule base and combine their action sides.

c. **Defuzzification :** It is process of converting the fuzzy output into precise numerical value.

Que 4.18. **Draw the block diagram of a fuzzy logic system, and define membership function.** **AKTU 2016-17(Sem-III), Marks 10** **AKTU 2017-18(Sem-III), Marks 10**

Answer

Fuzzy logic system :

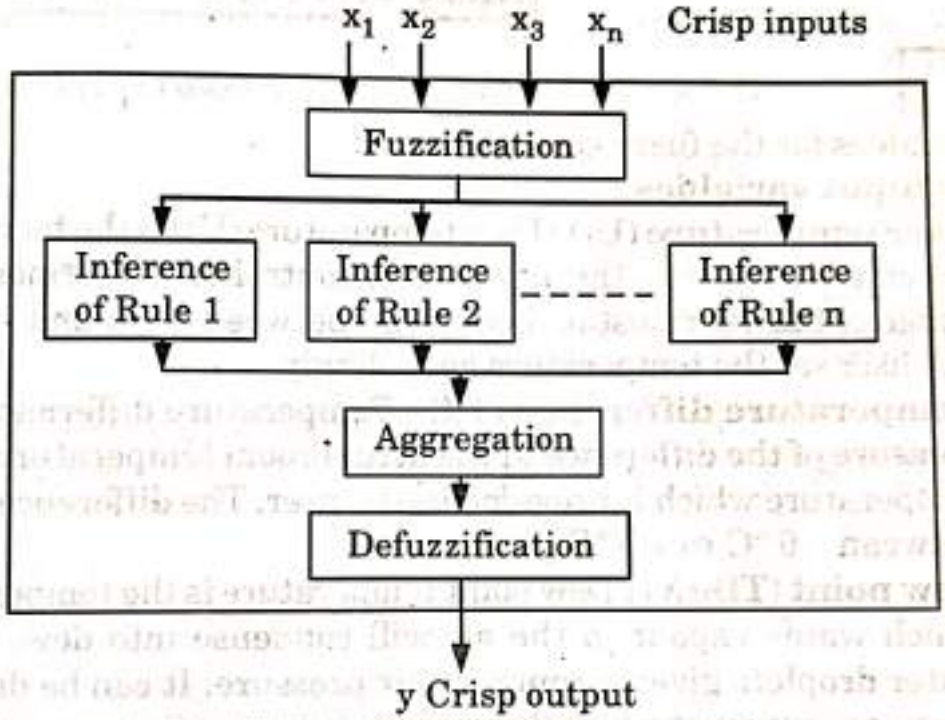


Fig. 4.18.1.

Membership function : Refer Q. 4.1, Page 4–2F, Unit-4.

Que 4.19. **What are the advantages and disadvantages of hybrid fuzzy controller in soft computing ?** **AKTU 2016-17(Sem-III), Marks 10**

Answer

Advantage of hybrid fuzzy controller :

1. Similar to human reasoning.
2. Based on linguistic model.
3. High precision.
4. Rapid operation.
5. Use simple mathematics for non- integrated and c

Disadvantage of hybrid fuzzy controller :

1. Lack of real time response.
2. The lower speed and also longer run time of system.
3. Restricted number of usage of input variables.
4. Not capable to receive feedback for implementation of learning strategy.
5. For more accuracy, needs more fuzzy grades which results to increase the rule exponentially.

Que 4.20. For an air conditioner what will be the input and output in a fuzzy controller ?

AKTU 2016-17(Sem-III), Marks 10

AKTU 2015-16(Sem-III), Marks 05

Answer

Various variables for the fuzzy controller are :

1. **Fuzzy input variables :**
   - i. **User temperature (Ut) :** User temperature (Ut) is the temperature provided by the user through remote controller or thermostat. The range of this thermostat should vary between 18 °C and 30 °C. So; the user set the temperature accordingly.
   - ii. **Temperature difference (Td) :** Temperature difference (Td) is measure of the difference in the actual room temperature and the temperature which is provided by the user. The difference range is between – 6 °C to + 6 °C.
   - iii. **Dew point (TDew) :** Dew point temperature is the temperature at which water vapour in the air will condense into dew, frost, or water droplets given a constant air pressure. It can be defined as the temperature at which the saturation vapour pressure and actual vapor pressure are equal.
   - iv. **Occupancy (Occ) :** Occupancy is the number of people exposed to air conditioner. The range of people will decide the level of occupancy as low, medium or high. In the absence of people the compressor as well as the fan remains off. The ranges can be varied according to various scenarios like indoor stadiums, auditoriums, etc.
   - v. **Time of Day (TDay) :** Time of Day is the period during which the AC would be working. The temperature and dew point values vary significantly during morning/night time with that of afternoon time as per the data provided by IMD. Accordingly the range of requirement can be decided for an optimum cooling and power consumption.
2. **Fuzzy output variables :**
   - i. **Compressor speed (Sc) :** The speed of compressor is varied between 30 to 100 %. Accordingly it will affect the room temperature as per to the given input.
   - ii. **Fan Speed (Sf) :** Fan speed gives the information about the fan running inside the air conditioner. The speed of fan is accordingly varied between 30 to 100 %.

iii. **Mode of Operation (Mo) :** Air conditioning system can act as a cooler as well as dehumidifier. In the cooling state it will regulate the air to release cool air. But as dehumidifier it can absorb the humid content of the air by passing dry air into the room. This process does not increase the temperature of the room. This setting preference is usually not given to the user and is performed implicitly by the AC. This parameter leads to greater efficiency and comfort levels.

iv. **Fin Direction (Fn) :** The fins are the set of blades attached to the air conditioner to ensure a swift flow of air in a particular direction. The direction of these fins will define the flow of air either towards or away from the user. The angle of propagation of blades is set accordingly considering 0° as "towards" and 90° as "away".

Que 4.21. Write fuzzy logic control system models.

AKTU 2015-16(Sem-IV), Marks 10

OR

Explain the Greg Viot's Fuzzy Cruise Controller.

AKTU 2016-17(Sem-IV), Marks 05

OR

Explain the working of any one fuzzy controller.

AKTU 2017-18(Sem-IV), Marks 07

Answer

Fuzzy logic control system models are :

a. **Greg Viot's fuzzy cruise controller :**

1. This controller is used to maintain a vehicle at a desired speed.
2. The system consists of two fuzzy inputs, namely speed difference, acceleration and one fuzzy output, namely throttle control.

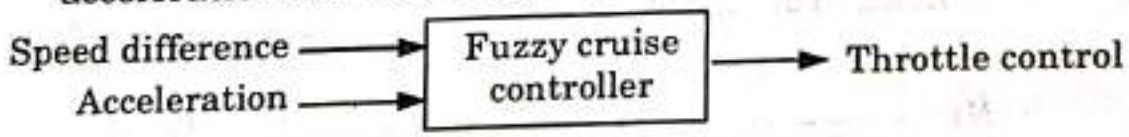


Fig. 4.21.1. Fuzzy cruise controller.

3. A sample fuzzy rule base R governing the cruise control is given as :

Rule 1 : If (speed difference is NL) and (acceleration is ZE) then (throttle control is PL)

Rule 2 : If (speed difference is ZE) and (acceleration is NL) then (throttle control is PL)

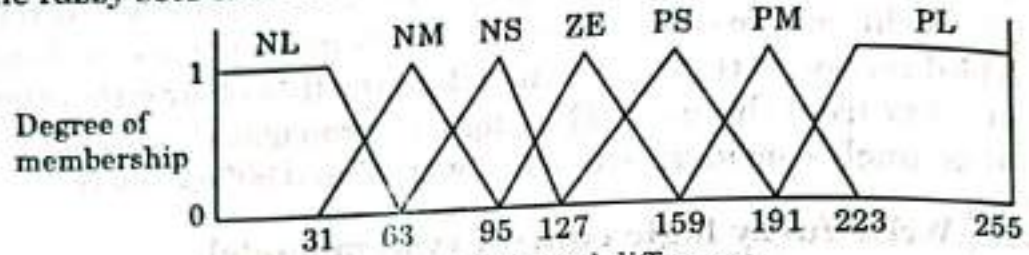
Rule 3 : If (speed difference is NM) and (acceleration is ZE) then (throttle control is PM)

Rule 4 : If (speed difference is NS) and (acceleration is PS) then (throttle control is PS)

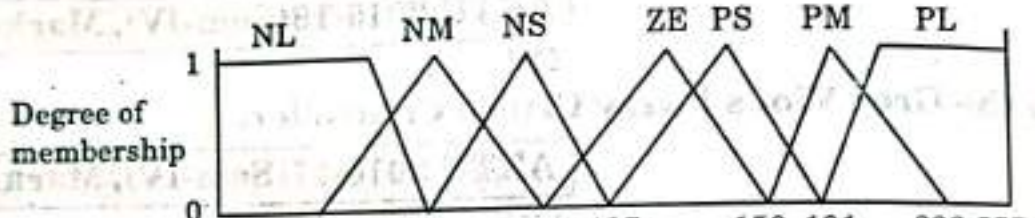
Rule 5 : If (speed difference is PS) and (acceleration is NS) then (throttle control is NS)

Rule 6 : If (speed difference is PL) and (acceleration is ZE) then (throttle control is NL)
Rule 7 : If (speed difference is ZE) and (acceleration is NS) then (throttle control is PS)
Rule 8 : If (speed difference is ZE) and (acceleration is NM) then (throttle control is PM)
Here, NL – Negative Large, ZE – Zero PL – Positive Large, NM – Negative Medium, PM – Positive Medium, PS – Positive Small, NS – Negative Small.

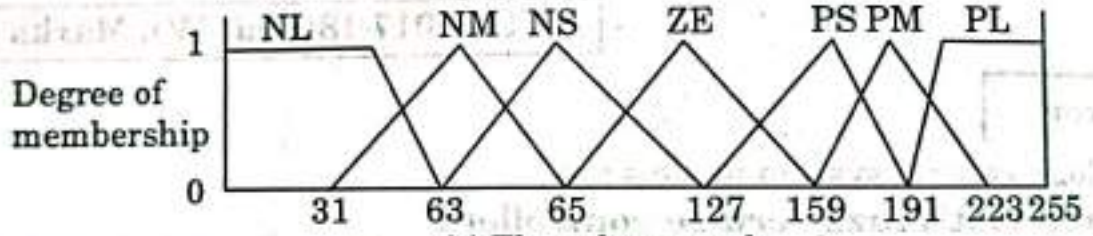
4. The fuzzy sets characterize the inputs and output as :



(a) Speed difference



(b) Acceleration



(c) Throttle control

Fig. 4.21.2.

5. For the fuzzification of inputs,

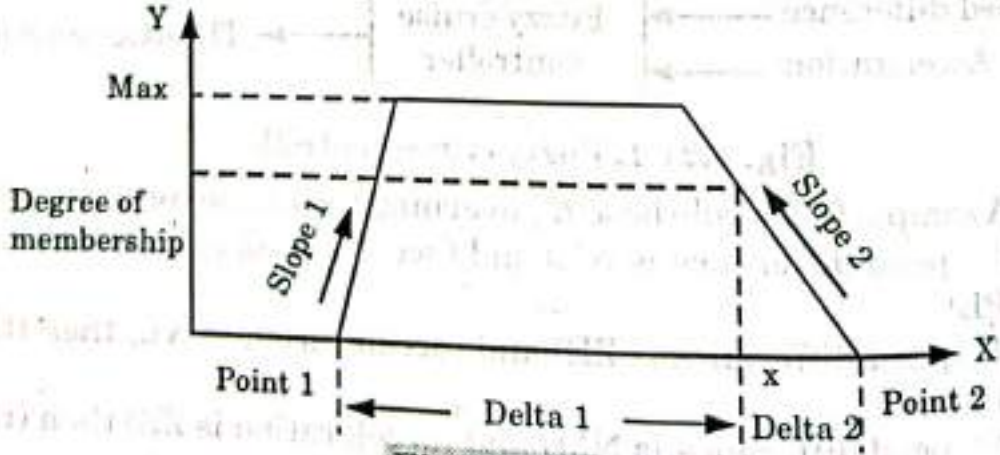


Fig. 4.21.3.

For degree of membership :
Compute Delta 1 = x – point 1
Delta 2 = Point 2 – x
If (Delta 1 ≤ 0) or (Delta 2 ≤ 0) then degree of membership = 0

$$\text{Else degree of membership} = \min \begin{bmatrix} \text{Delta 1} * \text{slope 1} \\ \text{Delta 2} * \text{slope 2} \\ \text{Max} \end{bmatrix}$$

Here, x which is the system input has its membership function values computed for all Fuzzy sets *i.e.*, namely NL, NM, NS, ZE, PS, PM, PL.

b. Yamakawa's air conditioner controller :

1. The system as illustrated in Fig. 6 comprises a dial to control the flow of warm/hot or cool/cold air and a thermometer to measure the room temperature (T °C).

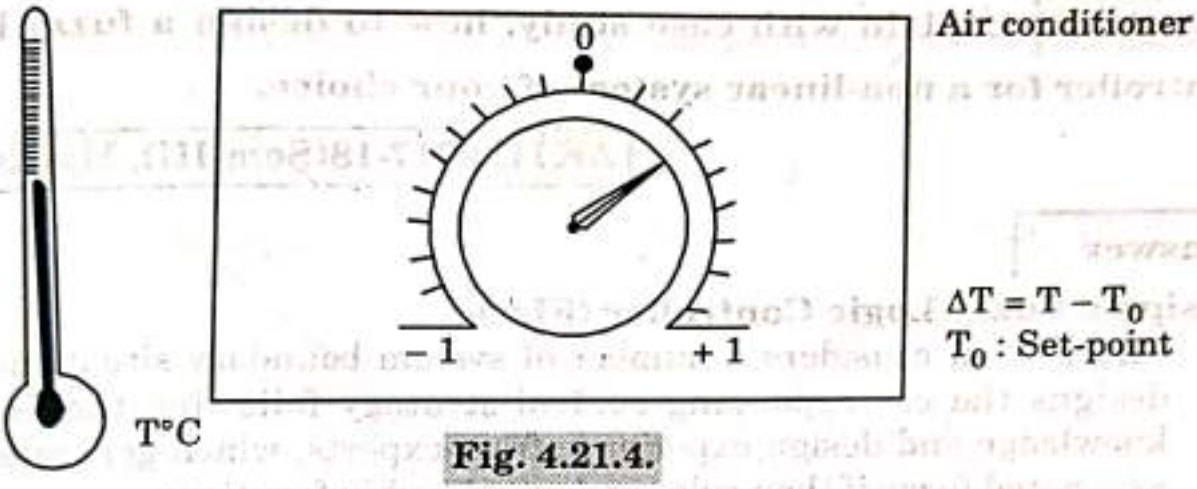


Fig. 4.21.4.

2. When the dial is turned positive, warm/hot air is supplied from the air conditioner and if it is turned negative, cool/cold air is supplied. If set to zero, no air is supplied.
3. Difference in temperature (ΔT °C) between the room temperature (T °C) as measured by the thermometer and the desired temperature (T_0 °C) at which the room is desired to be kept (set-point).
4. The problem now is to determine to what extent the dial should be turned so that the appropriate supply of air (hot/warm/cool/cold) will nullify the change in temperature.

Que 4.22. Sketch the block diagram of fuzzy logic controller for a non-linear process.

AKTU 2017-18(Sem-III), Marks 10

Answer

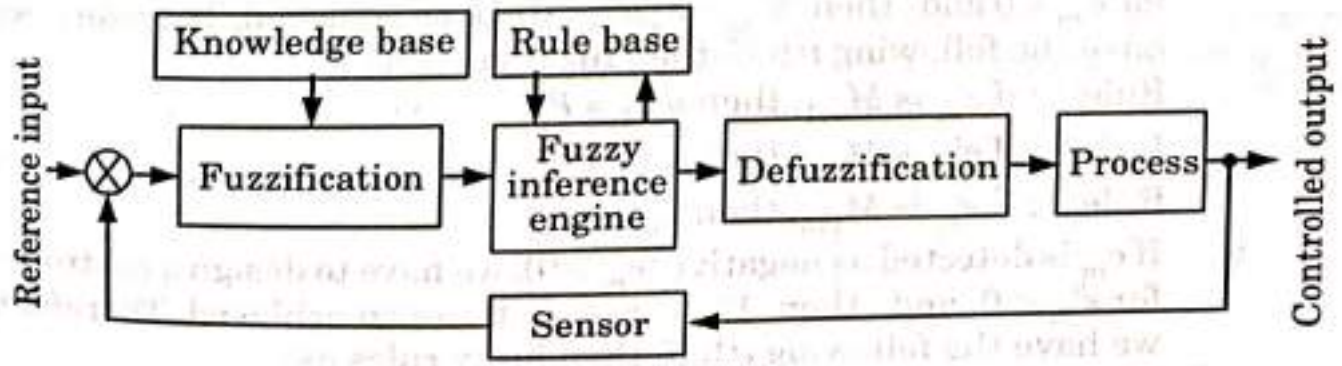


Fig. 4.22.1.

Que 4.23. Explain how fuzzy logic control can be used for process control.

AKTU 2017-18(Sem-III), Marks 10

Answer

Fuzzy Logic Control (FLC) can be considered an intelligent way to control process parameters through the formulation of IF-THEN based rule functions related to process control.

If-Then rule : Refer Q. 4.12, Page 4–12F, Unit-4.

Que 4.24. Explain with case study, how to design a fuzzy logic controller for a non-linear system of your choice.

AKTU 2017-18(Sem-III), Marks 10

Answer

Design a Fuzzy Logic Controller (FLC) :

1. FLC system considers a number of system boundary situations and designs the corresponding control strategy following the domain knowledge and design experience from experts, which generates the associated fuzzy if then rules and membership functions.
2. The fuzzy rules base consists of collection of fuzzy if then rules expressed in the form if a is A and then b is B where "if a is A" is the antecedent part and "b is B" is the consequent part; a and b denote linguistic variables and A and B represent linguistic values which are characterized by membership functions.
3. We use the error derivatives $e'(t) = [e'_1, e'_2, \ldots, e'_m, \ldots, e'_n]^T$ as the antecedent part and the corresponding control input u is considered to be the consequent part of the FLC system in this article.
4. In particular, u is designed as a constant column vector to stabilize the error dynamics.
$$u = [u_1, u_2, \ldots, u_m, \ldots, u_n]^T$$
5. Assuming the upper bound and lower bound of e'_m as Z_m and $-Z_m$, respectively, the Fuzzy logic controllers can be designed step by step as follows :
   i. If e_m is detected as positive ($e_m > 0$), we have to design a controller for $e'_m < 0$ and then, $V'_m = e_m e'_m < 0$ can be achieved. Therefore, we have the following ith if–then fuzzy rules as :
   Rule 1 : if e'_m is M_{m1}, then $u_{m1} = P_m$
   Rule 2 : if e'_m is M_{m2}, then $u_{m2} = 0$
   Rule 3 : if e'_m is M_{m3}, then $u_{m3} = e_m$
   ii. If e_m is detected as negative ($e_m > 0$), we have to design a controller for $e'_m < 0$, and then, $V'_m = e_m e'_m < 0$ can be achieved. Therefore, we have the following ith if–then fuzzy rules as :
   Rule 1 : if e'_m is M_{m1}, then $u_{m1} = -P_m$

Rule 2 : if e'_m is M_{m2}, then $u_{m2} = 0$

Rule 3 : if e'_m is M_{m3}, then $u_{m3} = e_m$

iii. If e_m approaches to zero, then the synchronization is nearly achieved. Therefore, we have the following ith if-then fuzzy rules as :

Rule 1 : if e'_m is M_{m1}, then $u_{m1} = e_m$

Rule 2 : if e'_m is M_{m2}, then $u_{m2} = e_m$

Rule 3 : if e'_m is M_{m3}, then $u_{m3} = e_m$

where $M_{m1} = M_{m2} = |e'_n|/Z_m$ and $M_{m3} = (Z_m - |e'_m|)/Z_m$, $M_{m1} = M_{m2}$ and M_{m3} refer to the membership functions of Positive (P), Negative (N), and Zero (Z) separately, which are presented in Fig. 4.24.1.

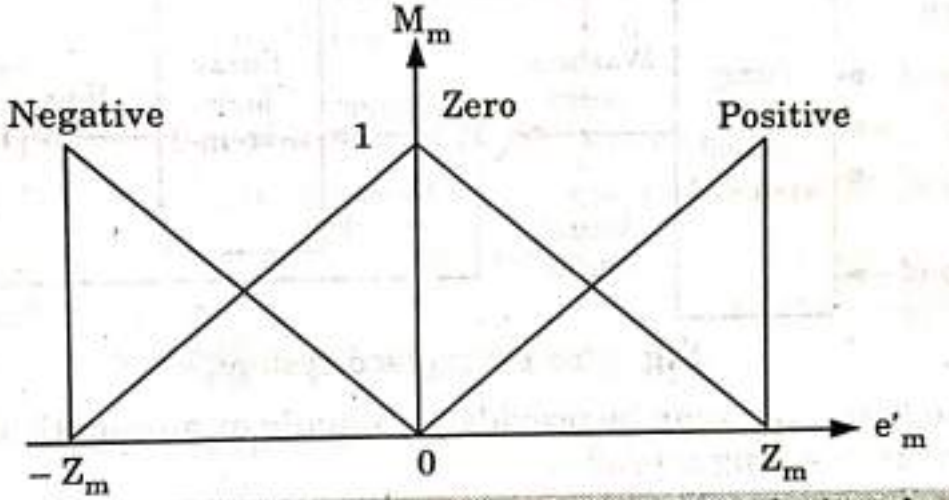


Fig. 4.24.1. Design of membership functions for each situation of e'_m

6. For each case u_{mi}, $i = 1$ to 3 is the ith output of e'_m which is a constant controller. The centroid defuzzifier evaluates the output of all rules as follows :

$$u_m = \frac{\sum_{i=1}^{3} M_{mi} \times u_{mi}}{\sum_{i=1}^{3} M_{mi}}$$

7. The fuzzy rule base is listed in Table 2, in which the input variables in the antecedent part of the rule are e'_m and the output variable in the consequent part is u_{mi}.

Table 4.24.1. Rule table of FLC

Rule	Antecedent part (e'_m)	Consequent part u_{mi}
1.	Negative (N)	u_{m1}
2.	Positive (P)	u_{m2}
3.	Zero (Z)	u_{m3}

Que 4.25. **A modern fully automatic washing machine is fixed with what type of intelligent control ? Define the input and output of the system.** **AKTU 2017-18(Sem-III), Marks 10**

Answer

1. The washing machine is fitted with fuzzy logic intelligent control.
2. The schematic diagram of fully automatic washing machine is shown in Fig. 4.25.1.
3. The fuzzy logic control system -1 is used in determining the washing speed with four inputs.

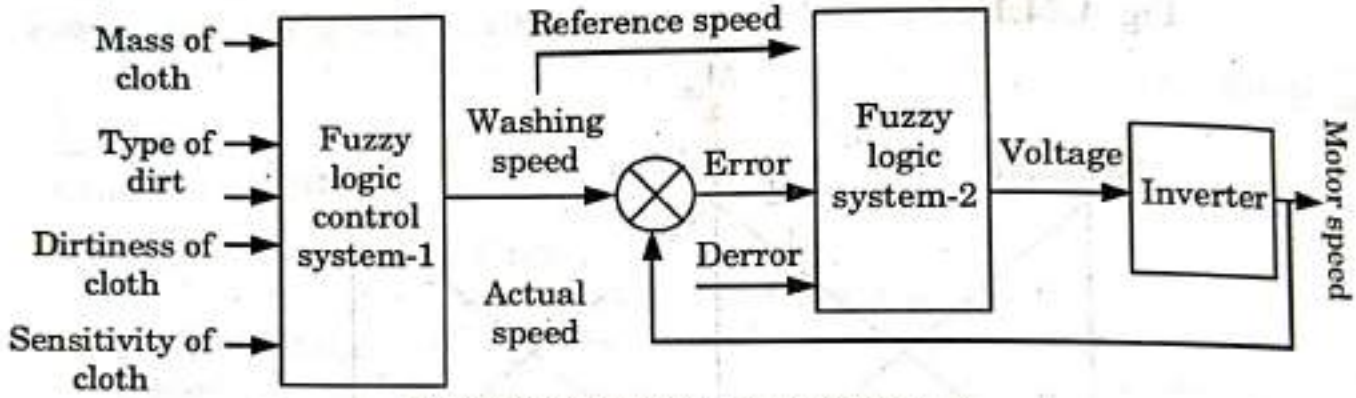


Fig. 4.25.1. Proposed system.

4. These four inputs can be provided manually or automatically by using the corresponding sensor.
5. After the system-1 produces the output the system then performs a process whereby the output on the system-2 is defined as the reference speed at the system-2 input.
6. The system-2 runs the process of regulating motor speed stability with three inputs.
7. Two other inputs include error is the difference between the reference speed with the actual speed and the current error difference (derror) with the previous error.

System-1 :

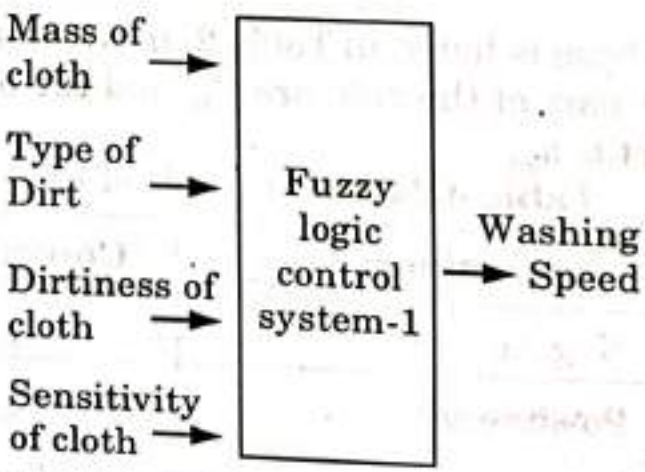


Fig. 4.25.2. Fuzzy logic control for washing machine.

i. The first step in using fuzzy logic control system is to determine the input and output.

ii. In this model the desired output is obviously the washing speed.

System-2 :

i. The input and output for fuzzy logic modeling of induction motor speed can be seen in Fig. 4.25.3.

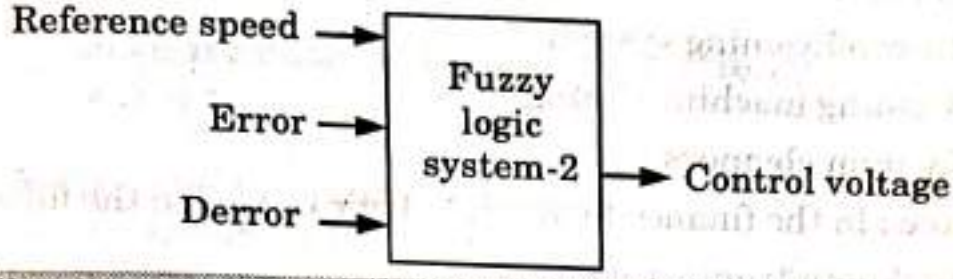


Fig. 4.25.3. Fuzzy logic control for induction motor speed.

ii. In this model the input parameters are reference speed, error, and derror value.

iii. The reference speed is obtained from the previous system output.

iv. Error is the difference between set point speed and actual speed.

v. Derror is the current error increment with the previous error.

vi. While output is a voltage (volts) that will be given to the inverter as a control voltage to run the motor.

Que 4.26. Explain the industrial applications of fuzzy logic.

AKTU 2016-17(Sem-III), Marks 10

Answer

Industrial application of fuzzy logic :

1. **Aerospace :** In aerospace, fuzzy logic is used in the following areas :
   i. Altitude control of spacecraft
   ii. Satellite altitude control
   iii. Flow and mixture regulation in aircraft deicing vehicles
2. **Automotive :** In automotive, fuzzy logic is used in the following areas :
   i. Shift scheduling method for automatic transmission
   ii. Intelligent highway systems
   iii. Traffic control
   iv. Improving efficiency of automatic transmissions
3. **Business :** In business, fuzzy logic is used in the following areas :
   i. Decision-making support systems
   ii. Personnel evaluation in a large company
4. **Defence :** In defence, fuzzy logic is used in the following areas :

i. Underwater target recognition
ii. Automatic target recognition of thermal infrared images
iii. Control of a hypervelocity interceptor

5. **Electronics :** In electronics, fuzzy logic is used in the following areas :
   i. Control of automatic exposure in video cameras
   ii. Air conditioning systems
   iii. Washing machine timing
   iv. Vacuum cleaners
6. **Finance :** In the finance field, fuzzy logic is used in the following areas :
   i. Banknote transfer control
   ii. Fund management
   iii. Stock market predictions
7. **Manufacturing :** In the manufacturing industry, fuzzy logic is used in following areas :
   i. Optimization of cheese production
   ii. Optimization of milk production
8. **Marine :** In the marine field, fuzzy logic is used in the following areas :
   i. Autopilot for ships
   ii. Optimal route selection
9. **Medical :** In the medical field, fuzzy logic is used in the following areas :
   i. Medical diagnostic support system
   ii. Control of arterial pressure during anesthesia
   iii. Radiology diagnoses
10. **Securities :** In securities, fuzzy logic is used in following areas :
   i. Decision systems for securities trading
   ii. Various security appliances

VERY IMPORTANT QUESTIONS

Following questions are very important. These questions may be asked in your SESSIONALS as well as UNIVERSITY EXAMINATION.

Q. 1. Define the membership function and state its importance in fuzzy logic. Also discuss the features of membership functions.

Ans. Refer Q. 4.1.

Q. 2. What are different attributes of predicate logic ? Using inference in predicate logic prove following statement :

i. All men are mortal.

ii. Socrates is a man.

Prove : Socrates is mortal.

Ans. Refer Q. 4.8.

Q. 3. Write short note on fuzzy implication.

Ans. Refer Q. 4.13.

Q. 4. Define the following terms :

i. Fuzzy algorithm

ii. Fuzzification

iii. Defuzzification

iv. Fuzzy if then rules

Ans. Refer Q. 4.15.

Q. 5. What are fuzzy implications ? Discuss the fuzzy controller.

Ans. Refer Q. 4.17.

Q. 6. Draw the block diagram of a fuzzy logic system, and define membership function.

Ans. Refer Q. 4.18.

Q. 7. For an air conditioner what will be the input and output in a fuzzy controller ?

Ans. Refer Q. 4.20.

Q. 8. Explain how fuzzy logic control can be used for process control.

Ans. Refer Q. 4.23.

Q. 9. Explain with case study, how to design a fuzzy logic controller for a non-linear system of your choice.

Ans. Refer Q. 4.24.

Q. 10. A modern fully automatic washing machine is fixed with what type of intelligent control ? Define the input and output of the system.

Ans. Refer Q. 4.25.

☺☺☺

5 UNIT

Genetic Algorithm (GA)

CONTENTS

PART-1

Basic Concepts, Working Principle, Procedures of GA, Flow Chart of GA.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 5.1. **Explain genetic algorithm. Also draw and explain the flow chart of genetic algorithm.**

AKTU 2013-14(Sem-III), Marks 10

OR

What is Genetic Algorithm ? Draw the general flow diagram of genetic algorithm.

AKTU 2016-17(Sem-IV), Marks 10

OR

Define the basic concept of genetic algorithm.

AKTU 2014-15(Sem-IV), Marks 05

Answer

Genetic algorithm (GA) :

1. The genetic algorithm is a method for solving both constrained and unconstrained optimization problems that is based on natural selection.
2. The genetic algorithm repeatedly modifies a population of individual solutions.
3. At each step, the genetic algorithm selects individuals at random from the current population to be parents and uses them to produce the children for the next generation.
4. Over successive generations, the population evolves toward an optimal solution.

Flow chart : The genetic algorithm uses three main types of rules at each step to create the next generation from the current population :

a. **Selection rule :** Selection rules select the individuals, called parents, that contribute to the population at the next generation.

b. **Crossover rule :** Crossover rules combine two parents to form children for the next generation.

c. **Mutation rule :** Mutation rules apply random changes to individual parents to form children.

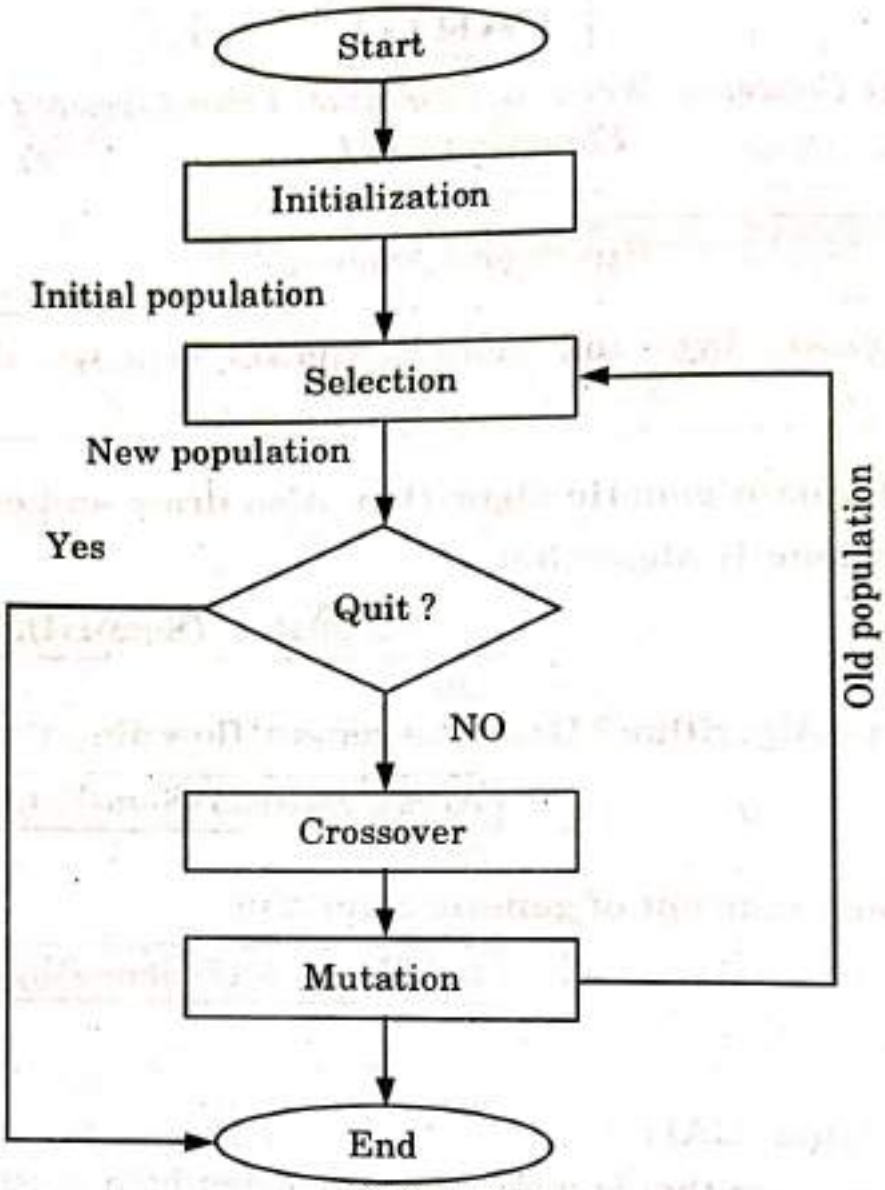


Fig. 5.1.1.

Que 5.2. Draw a flowchart of GA and explain the working principle.

AKTU 2016-17(Sem-III), Marks 10

Answer

Genetic algorithm : Refer Q. 5.1, Page 5–2F, Unit-5.

Working principle :

1. To illustrate the working principle of GA, we consider unconstrained optimization problem.
2. Let us consider the following maximization problem :

   maximize $f(X)$

   $X_i^{(L)} \leq X_i \leq X_i^{(U)}$ for $i = 1, 2 \ldots N$,
3. If we want to minimize $f(X)$, for $f(X) > 0$, then we can write the objective function as :

maximize $\frac{1}{1+f(X)}$

4. If $f(X) < 0$ instead of minimizing $f(X)$, maximize $\{-f(X)\}$. Hence, both maximization and minimization problems can be handled by GA.

Que 5.3. Write short notes on procedures of GA.

AKTU 2013-14(Sem-IV), Marks 05

Answer

1. **Start :** Generate random population of n chromosomes.
2. **Fitness :** Evaluate the fitness $f(x)$ of each chromosome x in the population.
3. **New population :** Create a new population by repeating following steps until the new population is complete.
   a. **Selection :** Select two parent chromosomes from a population according to their fitness.
   b. **Crossover :** With a crossover probability crossover the parents to form new offspring (children). If no crossover was performed, offspring is the exact copy of parents.
   c. **Mutation :** With a mutation probability mutate new offspring at each locus (position in chromosome).
   d. **Accepting :** Place new offspring in the new population.
4. **Replace :** Use new generated population for a further run of the algorithm.
5. **Test :** If the end condition is satisfied, stop, and return the best solution in current population.
6. Go to step 2

Que 5.4. What are the benefits of using GA ? What are its limitations ?

AKTU 2017-18(Sem-IV), Marks 07

OR

Write benefits of genetic algorithm. Explain the back propagation algorithm.

AKTU 2015-16(Sem-III), Marks 10

Answer

Benefits of using GA :

1. It is easy to understand.
2. It is modular and separate from application.
3. It supports multi-objective optimization.

4. It is good for noisy environment.

Limitations of genetic algorithm are :

1. The problem of identifying fitness function.
2. Definition of representation for the problem.
3. Premature convergence occurs.
4. The problem of choosing the various parameters like the size of the population, mutation rate, crossover rate, the selection method and its strength.
5. Cannot use gradients.
6. Cannot easily incorporate problem specific information.
7. Not good at identifying local optima.
8. No effective terminator.
9. Not effective for smooth unimodal functions.
10. Needs to be coupled with a local search technique.

Back propagation algorithm : Refer Q. 2.20, Page 2–20F, Unit-2.

PART-2

Genetic Representations, (Encoding) Initialization and Selection.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 5.5. Write short notes of genetic representations.

AKTU 2013-14(Sem-IV), Marks 05

Answer

1. Genetic representation is a way of representing solutions/individuals in evolutionary computation methods.
2. Genetic representation can encode appearance, behavior, physical qualities of individuals.
3. All the individuals of a population is represented by using binary encoding, permutational encoding, encoding by tree.
4. Genetic algorithms use linear binary representations. The most standard method of representation is an array of bits.
5. These genetic representations are convenient because parts of individual are easily aligned due to their fixed size which makes simple crossover operation.

Que 5.6. **Give the detail of genetic representation (Encoding).**

AKTU 2014-15(Sem-IV), Marks 05

OR

Explain different types of encoding in genetic algorithm.

Answer

Genetic representations :

1. **Encoding :**
   a. Encoding is a process of representing individual genes.
   b. The process can be performed using bits, numbers, trees, arrays, lists or any other objects.
   c. The encoding depends mainly on solving the problem.
2. **Binary encoding :**
   a. Binary encoding is the most commonly used method of genetic representation because GA uses this type of encoding.
   b. In binary encoding, every chromosome is a string of bits, 0 or 1.

Chromosome A	10110010110010101100101
Chromosome B	11111110000011000011111

   c. Binary encoding gives many possible chromosomes.
3. **Octal or Hexadecimal encoding :**
   a. The encoding is done using octal or hexadecimal numbers.

Chromosome	Octal	Hexadecimal
Chromosome A	54545345	B2CAE5
Chromosome B	77406037	FE0C1F

4. **Permutation encoding (real number encoding) :**
   a. Permutation encoding can be used in ordering problems, such as Travelling Salesman Problem (TSP).
   b. In permutation encoding, every chromosome is a string of numbers, which represents number in a sequence.

Chromosome A	1 5 3 2 6 4 7 9 8
Chromosome B	8 5 6 7 2 3 1 4 9

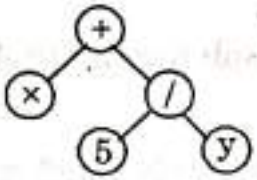
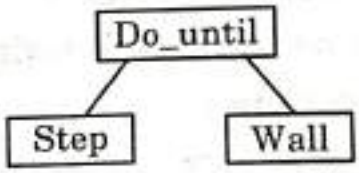
5. **Value encoding :**
   a. Direct value encoding can be used in problems, where some complicated values, such as real numbers, are used.
   b. In value encoding, every chromosome is a string of some values.
   c. Values can be anything connected to problem, real numbers or chars to some complicated objects.

Chromosome A	1.2324 5.3243 0.4556 2.3293 2.4545
Chromosome B	ABDJEIFJDHDIERJFDLDFLFEGT
Chromosome C	(back), (back), (right), (forward), (left)

6. Tree encoding :

a. Tree encoding is used for evolving programs or expressions, for genetic programming.

b. In tree encoding, every chromosome is a tree of some objects, such as functions or commands in programming language.

Chromosome A	Chromosome B
 (+ x (/ 5 y))	 (do_until step wall)

c. Programming language LISP is often used to this, because programs in it are represented in this form and can be easily parsed as a tree, so the cross-over and mutation can be done relatively easily.

Que 5.7. Explain different methods of selection in genetic algorithm in order to select a population for next generation.

AKTU 2014-15(Sem-IV), Marks 05

OR

State the different selection method in GA.

AKTU 2015-16(Sem-IV), Marks 10

Answer

The various methods of selecting chromosomes for parents to cross over are :

a. Roulette-wheel selection :

i. Roulette-wheel selection is the proportionate reproductive method where a string is selected from the mating pool with a probability proportional to the fitness.

ii. Thus, ith string in the population is selected with a probability proportional to F_i where F_i is the fitness value for that string.

iii. Since the population size is usually kept fixed in Genetic Algorithm, the sum of the probabilities of each string being selected for the mating pool must be one.

iv. The probability of the ith selected st is

$$p_i = \frac{F_i}{\sum_{j=1}^{n} F_j}$$

where 'n' is the population size.

v. The average fitness is

$$\bar{F} = \sum_{j=1}^{n} F_j / n \qquad \text{...(5.7.1)}$$

b. Boltzmann selection :

i. Boltzmann selection uses the concept of simulated annealing.

ii. Simulated annealing is a method of functional minimization or maximization.

iii. This method simulates the process of slow cooling of molten metal to achieve the minimum function value in a minimization problem.

iv. The cooling phenomenon is simulated by controlling a temperature so that a system in thermal equilibrium at a temperature T has its energy distributed probabilistically according to

$$P(E) = \exp\left(-\frac{E}{kT}\right) \qquad \text{...(5.7.2)}$$

where 'k' is Boltzmann constant.

v. This expression suggests that a system at a high temperature has almost uniform probability of being at any energy state, but at a low temperature it has a small probability of being at a high energy state.

vi. Therefore, by controlling the temperature T and assuming search process follows Boltzmann probability distribution, the convergence of the algorithm is controlled.

c. Tournament selection :

i. GA uses a strategy to select the individuals from population and insert them into a mating pool.

ii. A selection strategy in GA is a process that favours the selection of better individuals in the population for the mating pool.

iii. There are two important issues in the evolution process of genetic search.

1. **Population diversity :** Population diversity means that the genes from the already discovered good individuals are exploited.
2. **Selective pressure :** Selective pressure is the degree to which the better individuals are favoured.

iv. The higher the selective pressure the better individuals are favoured.

d. Rank selection :

i. Rank selection first ranks the population and takes every chromosome, receives fitness from the ranking.

ii. The worst will have fitness 1, the next 2, ..., and the best will have fitness N (N is the number of chromosomes in the population).

iii. The method can lead to slow convergence because the best chromosome does not differ so much from the other.

e. Steady-state selection :

i. The main idea of the selection is that bigger part of chromosome should survive to next generation.

ii. GA works in the following way :

1. In every generation a few chromosomes are selected for creating new off springs.
2. Then, some chromosomes are removed and new offspring is placed in that place.
3. The rest of population survives a new generation.

Que 5.8. Differentiate between Roulette-wheel based on fitness and Roulette-wheel based on rank with suitable example.

AKTU 2016-17(Sem-IV), Marks 10

Answer

Difference :

S. No.	Roulette-wheel based on fitness	Roulette-wheel based on rank
1.	Population is selected with a probability that is directly proportional to their fitness values.	Probability of a population being selected is based on its fitness rank.
2.	It computes selection probabilities according to their fitness values but do not sort the individual in the population.	It first sort individuals in the population according to their fitness and then computes selection probabilities according to their ranks rather than fitness values.
3.	It gives a chance to all the individuals in the population to be selected.	It selects the individuals with highest rank in the population.
4.	Diversity in the population is preserved.	Diversity in the population is not preserved.

Example :

1. Imagine a Roulette-wheel where all chromosomes in the population are placed, each chromosome has its place accordingly to its fitness function :

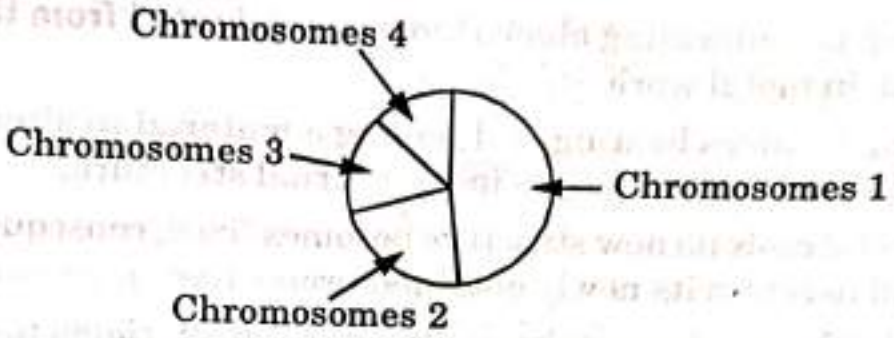


Fig. 2.8.1. Roulette-wheel selection.

2. When the wheel is spun, the wheel will finally stop and pointer attached to it will points to the one of chromosomes with bigger fitness value.
3. The different between roulette-wheel selection based on fitness and rank is shown in Fig. 2.8.1 and Fig. 2.8.3.

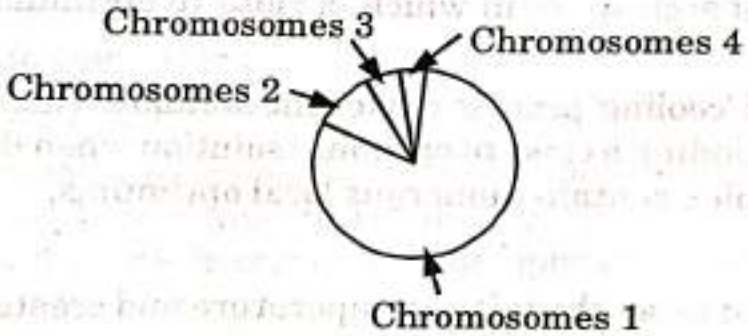


Fig. 2.8.2. Situation before ranking (graph of fitnesses).

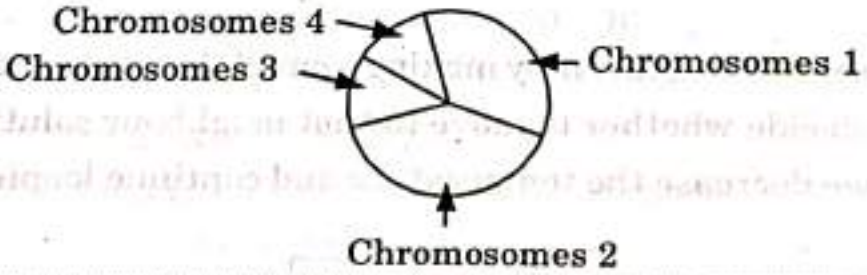


Fig. 2.8.3. Situation after ranking (graph of order numbers).

Que 5.9. What is simulated annealing ? What is the structure of a simulated annealing algorithm ?

AKTU 2015-16(Sem-IV), Marks 10

Answer

1. Stimulated annealing is a mathematical and modeling method that is often used to find a global optimization in a particular function or problem.
2. The simulated annealing algorithm was originated from the process of annealing in metal work.
3. Annealing involves heating and cooling a material to alter its physical properties due to the changes in its internal structure.
4. As the metal cools its new structure becomes fixed, consequently causing the metal to retain its newly obtained properties.
5. In simulated annealing, we keep a temperature variable to simulate this heating process. We initially set it high and then allow it to slowly cool as the algorithm runs.
6. While this temperature variable is high the algorithm will be allowed, with more frequency, to accept solutions that are worse than our current solution.
7. As the temperature is reduced so is the chance of accepting worse solutions, therefore allowing the algorithm to gradually focus in on a area of the search space in which, a close to optimum solution can be found.
8. This gradual 'cooling' process makes the simulated annealing algorithm effective at finding a close to optimum solution when dealing with large problems which contain numerous local optimums.

Algorithm :

1. First we need to set the initial temperature and create a random initial solution.
2. Then we begin looping until stop condition is met. Usually either the system has sufficiently cooled, or a good enough solution has been found.
3. Then we select a neighbour by making a small change to current solution.
4. We then decide whether to move to that neighbour solution.
5. Finally, we decrease the temperature and continue looping.

PART-3

Genetic Operators, Mutation, Generational Cycle, Applications.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 5.10. **What are the genetic operators ? What are the roles of genetic operators in GA ?** **AKTU 2013-14(Sem-III), Marks 10**

OR

Explain two point crossover and uniform crossover in genetic algorithm. **AKTU 2016-17(Sem-III), Marks 10**

AKTU 2015-16(Sem-III), Marks 10

OR

Discuss different genetic operators. **AKTU 2014-15(Sem-IV), Marks 05**

OR

Discuss crossover operation in genetic algorithm and its type. **AKTU 2017-18(Sem-IV), Marks 07**

Answer

Genetic operator : A genetic operator is an operator used in genetic algorithm to guide the algorithm towards a solution to a given problem.

Different genetic operator :

i. **Bit-wise operator :**

1. **One's compliment operator (~) :** It is unary operator that causes the bits of its operand to be inverted (*i.e.*, reversed), so that 1 becomes 0 and 0 becomes 1. This operator always precedes its operand.

2. **Bit-wise AND (&) operator :** A bit-wise AND (&) expression returns 1 if both the bits have a value 1, otherwise it returns a value 0.

   Parent 1a = 1010 1010 → 10 10

   Parent 2b = 1100 0011 → 12 3

   Child a&b = 1000 0010 → 8 2

3. **Bit-wise exclusive-OR value (∧) operator :** A bit-wise exclusive-OR (∧) expression returns a value 1 if one of the bits have a value 1 and the other has a value of 0 otherwise it returns a value 0.

   Parent 1a = 1010 1010 → 10 10

   Parent 2b = 1100 0011 → 12 3

   Child a&b = 0110 1001 → 6 9

4. **Bit-wise OR (|) operator :** A bit-wise (|) expression returns 1 if one or more bits have a value of 1 otherwise it returns a value 0.

   Parent 1a = 1010 1010 → 10 10

Parent 2a = 1100 0011 → 12 3

Child a&b = 1110 1011 → 13 11

ii. Shift operators :

1. **Shift left operator (<<) :** The shift left operator causes all the bits in the first operand to be shifted to the left by the number of positions indicated by the second operand. The leftmost bits (*i.e.*, the overflow bits) in the original bit pattern is lost. The rightmost bit positions that become vacant are to be filled with zeroes.

a = 1010 1010 → 10 6

a << 2 = 1001 1000 → 9 8

2. **Masking :** Masking is a process in which a given bit pattern is transformed into another bit pattern by means of logical bit-wise operation. The original bit pattern is one of the operands in the bit-wise operation. The second operand called mask, is a specially selected bit pattern that brings about the desired transformation.

iii. Mutation operator :

1. Mutation is a genetic operator used to maintain genetic diversity from one generation of a population of genetic algorithm chromosomes to the next.
2. It is analogous to biological mutation.
3. Mutation alters one or more gene values in a chromosome from its initial state.
4. In mutation, the solution may change entirely from the previous solution.
5. Mutation occurs during evolution according to a user-definable mutation probability.

iv. Crossover operator :

a. Crossover operator is applied to the mating pool with a hope that it would create a better string.

b. The aim of the crossover operator is to search the parameter space.

c. In addition, search is to be made in such a way that the information stored in the present string is maximally preserved because these parent strings are instances of good strings selected during reproduction.

1. Single-site Crossover :

i. In a single-site crossover, a cross-site is selected randomly along the length of the mated strings and bits next to the cross-sites are exchanged as shown in Fig. 5.10.1.

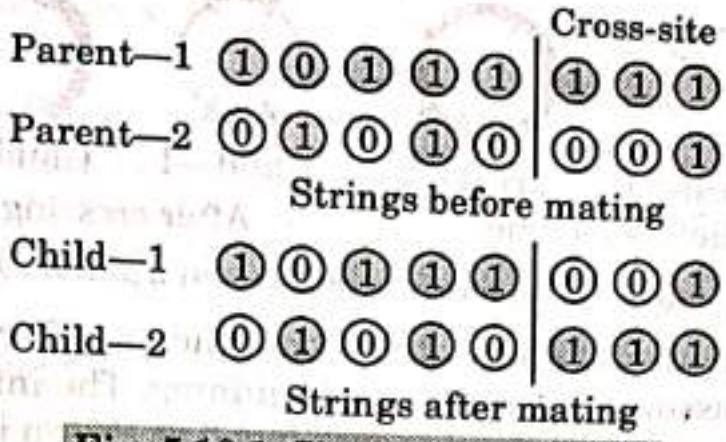


Fig. 5.10.1. Single-site crossover.

ii. If an appropriate site is chosen, better children can be obtained by combining good substances of parents.

iii. Since the knowledge of the appropriate site is not known and it is selected randomly, this random selection of cross-sites may produce enhanced children if the selected site is appropriate. If not, it may severely hamper the string quality.

2. Two-point Crossover :

i. In a two-point crossover operator, two random sites are chosen and the contents bracketted by these sites are exchanged between two mated parents.

ii. If the cross-site 1 is three and cross-site 2 is six, the strings between three and six are exchanged as shown in Fig. 5.10.2.

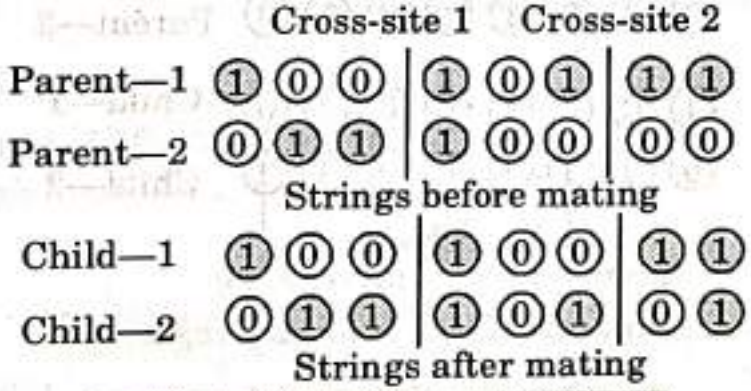


Fig. 5.10.2. Two-point crossover.

3. Multi-point Crossover :

i. In a multi-point crossover, there are two cases.

ii. One is even number of cross-sites and second one is the odd number of cross-sites.

iii. In case of even numbered cross-sites, the string is treated as a ring with no beginning or end.

iv. The cross-sites are selected around the circle uniformly at random.

v. Now the information between alternate pairs of sites is interchanged as shown in Fig. 5.10.3.

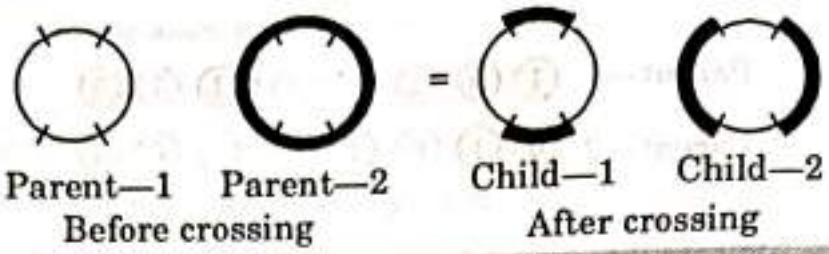


Fig. 5.10.3. Multi-point crossover with even number of cross-sites.

vi. If the number of cross-sites is odd, then a different cross-point is always assumed at the string beginning. The information (genes) between alternate pairs is exchanged as shown in Fig. 5.10.4.

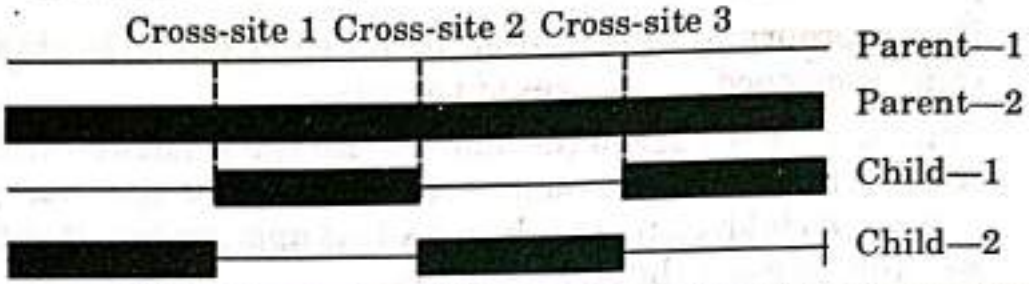


Fig. 5.10.4. Multi-point crossover with odd number of cross-sites.

4. Uniform Crossover :

i. In a uniform crossover operator, each bit from either parent is selected with a probability of 0.5 and then interchanged as shown in Fig. 5.10.5(a).

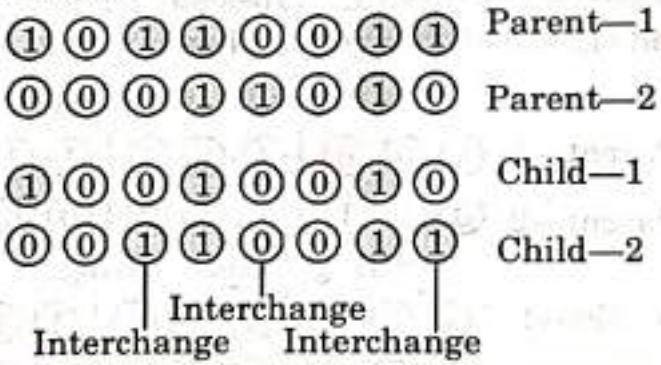


Fig. 5.10.5. (a) Uniform crossover.

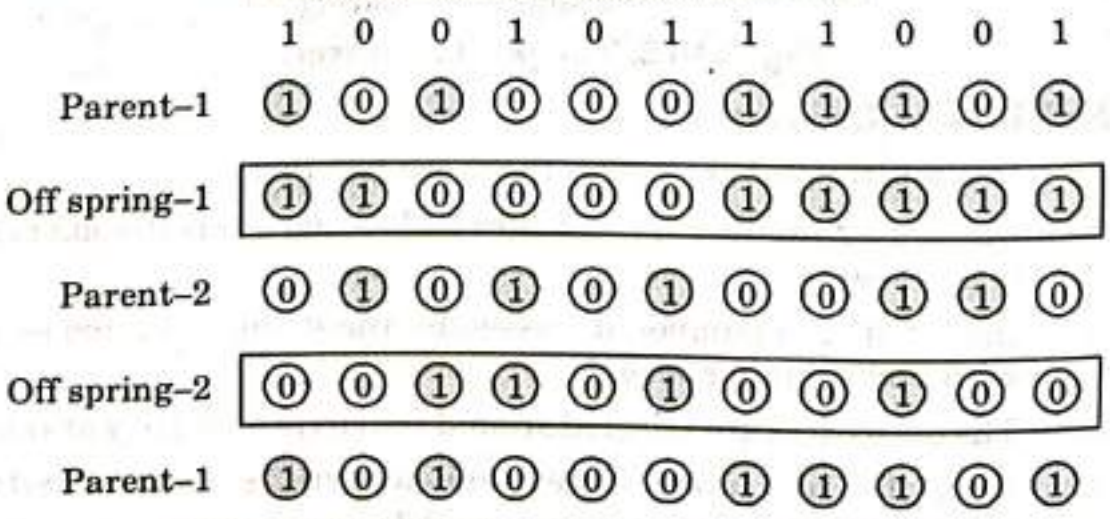


Fig. 5.10.5. (b) Uniform crossover using mask.

ii. Uniform crossover is radically different from one-point crossover.

iii. Sometimes gene in the offspring is created by copying the corresponding gene from one or the other parent chosen according to a randomly generated crossover mask.

iv. When there is 1 in the mask, the gene is copied from the first parent and when there is 0, the gene is copied from second parent as shown in Fig. 17(b).

v. The process is repeated with the parents exchanged to produce the second offspring.

vi. Offspring therefore contains a mixture of genes from each parent. The number of effective crossing points is not fixed but averages to L/2 (where L is chromosome length).

Roles of genetic operators in GA :

1. Genetic algorithm uses these operators to calculate successive generations.
2. It is used to select pairs of individuals from the current population.

Que 5.11. What are the mutation in GA ? Explain the generational cycle in GA. **AKTU 2013-14(Sem-III), Marks 10**

OR

Draw flow chart and genetics cycle for genetic algorithm. **AKTU 2014-15(Sem-IV), Marks 05**

Answer

Mutation : Refer Q. 5.10, Page 5–12F, Unit-5.

Generational cycle of GA :

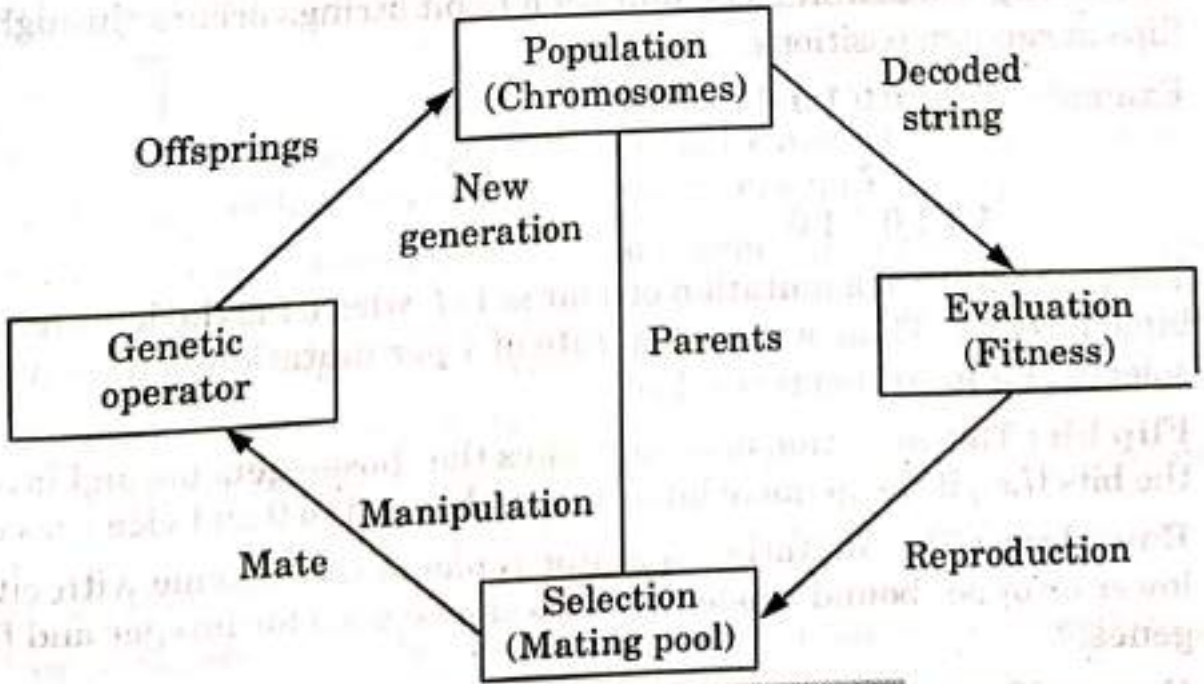


Fig. 5.11.1. The GA cycle.

Components of generational cycle in GA :

1. **Population (Chromosomes) :** A population is collection of individuals. A population consists of a number of individuals being tested, the phenotype parameters defining the individuals and some information about search space.
2. **Evaluation (Fitness) :** A fitness function is a particular type of objective function that quantifies the optimality of a solution (*i.e.*, a chromosome) in a genetic algorithm so that particular chromosome may be ranked against all the other chromosomes.
3. **Selection :** During each successive generation, a proportion of the existing population is selected to breed a new generation. Individual solutions are selected through a fitness-based process.
4. **Generic operator :** A genetic operator is an operator used in genetic algorithm to guide the algorithm towards a solution to a given problem.

Flow chart : Refer Q. 5.1, Page 5–2F, Unit-5.

Que 5.12. Why mutation is done in genetic algorithm ? Explain types of mutation. **AKTU 2015-16(Sem-III), Marks 10**

Answer

Mutation is done in genetic algorithm because :

1. It maintains genetic diversity from one generation of a population of genetic algorithm chromosomes to the next.
2. GA can give better solution of the problem by using mutation.

Types of mutation :

1. **Bit string mutation :** The mutation of bit strings occurs through bit flips at random positions.

   Example : 1 0 1 0 0 1 0

   ↓

   1 0 1 0 1 1 0

   The probability of a mutation of a bit is $1/l$, where l is the length of the binary vector. Thus, a mutation rate of 1 per mutation and individual selected for mutation is reached.
2. **Flip bit :** This mutation operator takes the chosen genome and inverts the bits (*i.e.*, if the genome bit is 1, it is changed to 0 and vice versa).
3. **Boundary :** This mutation operator replaces the genome with either lower or upper bound randomly. This can be used for integer and float genes.
4. **Non-uniform :** The probability that amount of mutation will go to 0 with the next generation is increased by using non-uniform mutation

operator. It keeps the population from stagnating in the early stages of the evolution.

5. **Uniform :** This operator replaces the value of the chosen gene with a uniform random value selected between the user-specified upper and lower bounds for that gene.
6. **Gaussian :** This operator adds a unit Gaussian distributed random value to the chosen gene. If it falls outside of the user-specified lower or upper bounds for that gene, the new gene value is clipped.

Que 5.13. Write short notes on mutation and mutation rate.

AKTU 2013-14(Sem-IV), Marks 05

Answer

Mutation : Refer Q. 5.10, Page 5–12F, Unit-5.

Mutation rate :

1. Mutation rate is the probability of mutation which is used to calculate number of bits to be muted.
2. The mutation operator preserves the diversity among the population which is also very important for the search.
3. Mutation probabilities are smaller in natural populations leading us to conclude that mutation is appropriately considered a secondary mechanism of genetic algorithm adoption.

Que 5.14. What is the main function of crossover operation in genetic algorithm ?

AKTU 2017-18(Sem-III), Marks 10

Answer

1. Crossover is the basic operator of genetic algorithm. Performance of genetic algorithm depends on crossover operator.
2. Type of crossover operator used for a problem depends on the type of encoding used.
3. The basic principle of crossover process is to exchange genetic material of two parents beyond the crossover points.

Function of crossover operation/operator in genetic algorithm :

1. The main function of crossover operator is to introduce diversity in the population.
2. Specific crossover made for a specific problem can improve performance of the genetic algorithm.
3. Crossover combines parental solutions to form offspring with a hope to produce better solutions.

4. Crossover operators are critical in ensuring good mixing of building blocks.
5. Crossover is used to maintain balance between exploitation and exploration. The exploitation and exploration techniques are responsible for the performance of genetic algorithms. Exploitation means to use the already existing information to find out the better solution and exploration is to investigate new and unknown solution in exploration space.

Que 5.15. Write short notes on application of GA.

AKTU 2013-14(Sem-IV), Marks 05

OR

Explain the applications of G.A. in general life.

AKTU 2017-18(Sem-IV), Marks 07

OR

Discuss the different applications of genetic algorithms.

AKTU 2014-15(Sem-IV), Marks 10

Answer

Application of GA :

1. **Optimization :** Genetic Algorithms are most commonly used in optimization problems wherein we have to maximize or minimize a given objective function value under a given set of constraints.
2. **Economics :** GAs are also used to characterize various economic models like the cobweb model, game theory equilibrium resolution, asset pricing, etc.
3. **Neural networks :** GAs are also used to train neural networks, particularly recurrent neural networks.
4. **Parallelization :** GAs also have very good parallel capabilities, and prove to be very effective means in solving certain problems, and also provide a good area for research.
5. **Image processing :** GAs are used for various digital image processing (DIP) tasks as w[illegible]ike dense pixel matching.
6. **Machine learning :** Genetics based machine learning (GBML) is a nice area in machine learning.
7. **Robot trajectory generation :** GAs have been used to plan the path which a robot arm takes by moving from one point to another.

Que 5.16. Explain optimization of travelling salesman problem using genetic algorithm and give a suitable example too.

AKTU 2016-17(Sem-III), Marks 10

Answer

1. The TSP consist a number of cities, where each pair of cities has a corresponding distance.
2. The aim is to visit all the cities such that the total distance travelled will be minimized.
3. Obviously, a solution, and therefore a chromosome which represents that solution to the TSP, can be given as an order, that is, a path, of the cities.
4. The procedure for solving TSP can be viewed as a process flow given in Fig. 5.16.1.

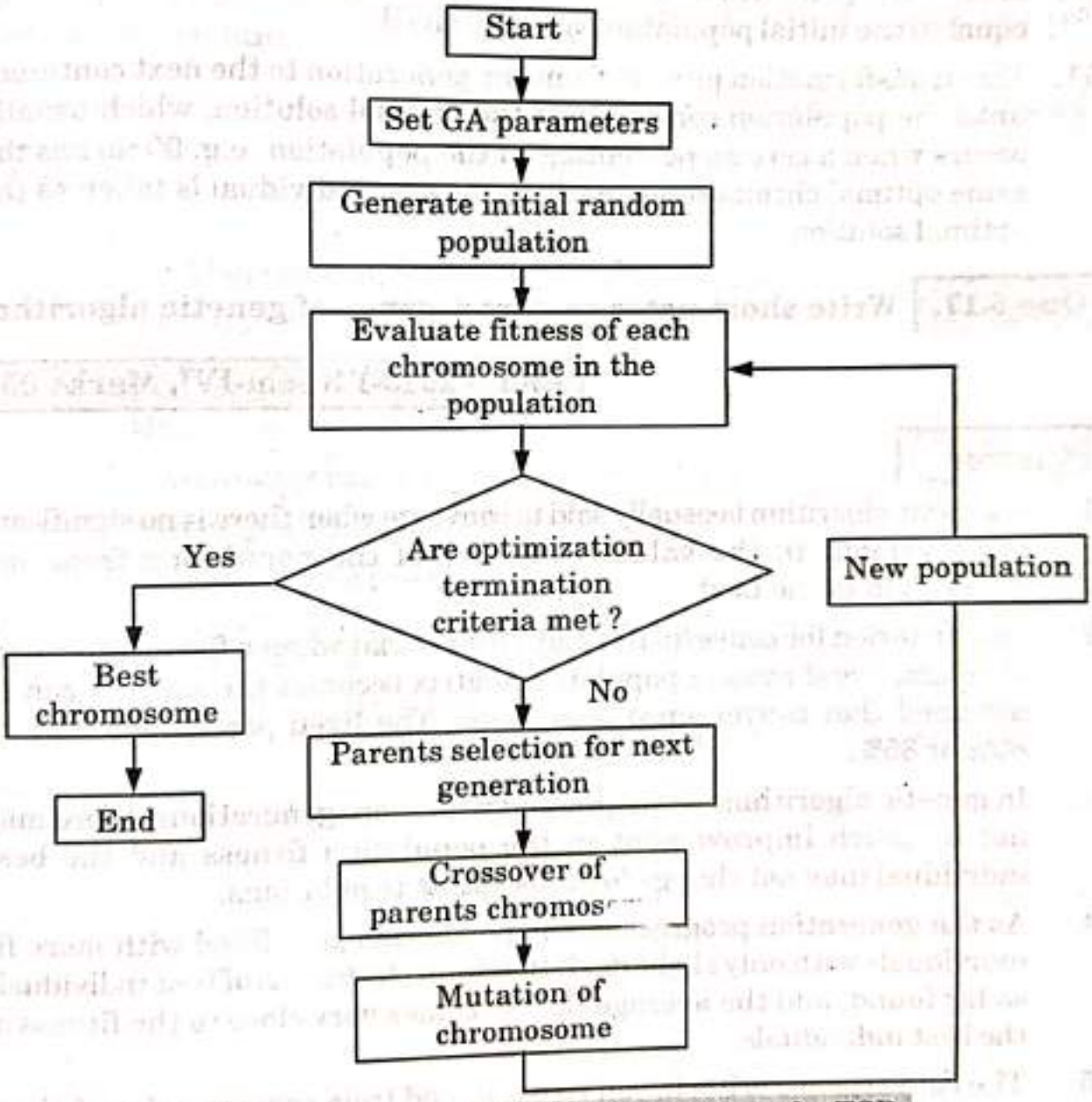


Fig. 5.16.1. Genetic algorithm procedure for TSP.

5. The GA process starts by supplying important information such as location of the city, maximum number of generations, population size, probability of crossover and probability of mutation.

6. An initial random population of chromosomes is generated and the fitness of each chromosome is evaluated.
7. The population is then transformed into a new population (the next generation) using three genetic operators : selection, crossover and mutation.
8. The selection operator is used to choose two parents from the current generation in order to create a new child by crossover and/or mutation.
9. The new generation contains a higher proportion of the characteristics possessed by the good members of the previous generation and in this way good characteristics are spread over the population and mixed with other good characteristics.
10. After each generation, a new set of chromosomes where the size is equal to the initial population size is evolved.
11. This transformation process from one generation to the next continues until the population converges to the optimal solution, which usually occurs when a certain percentage of the population (e.g. 90 %) has the same optimal chromosome in which the best individual is taken as the optimal solution.

Que 5.17. Write short notes on convergence of genetic algorithm

AKTU 2013-14(Sem-IV), Marks 05

Answer

1. A genetic algorithm is usually said to converge when there is no significant improvement in the values of fitness of the population from one generation to the next.
2. One criterion for convergence may be such that when a fixed percentage of columns and rows in population matrix becomes the same, it can be assumed that convergence is attained. The fixed percentage may be 80% or 85%.
3. In genetic algorithms as we proceed with more generations, there may not be much improvement in the population fitness and the best individual may not change for subsequent populations.
4. As the generation progresses, the population gets filled with more fit individuals with only slight deviation from the fitness of best individuals so far found, and the average fitness comes very close to the fitness of the best individuals.
5. The convergence criteria can be explained from schema point of view.
6. A schema is a similarity template describing a subset of strings with similarities at certain positions. A schema represents a subset of all possible strings that have the same bits at certain string positions.

7. Since schema represents a robust of strings, we can associate a fitness value with a schema, *i.e.*, the average fitness of the schema.
8. One can visualize GA's search for the optimal strings as a simultaneous competition among schema increases the number of their instances in the population.

Que 5.18. **Consider four travel packages offered by Thomas Cook, Club Mahindra, World around, and Himalaya Travels. We want to choose one. Their costs are INR 100,000, INR 200,000, INR 150,000 and INR 175,000. Their travel time in hours is 150, 200, 100, and 125 respectively. They are viewed as interesting with degrees 0.4, 0.3, 0.6, 0.5. Define your own fuzzy set of interesting travel packages whose cost and travel times are acceptable and use this set to choose one of your own packages.**

AKTU 2015-16(Sem-IV), Marks 05

Answer

Consider,

Thomas cook = A_1

Club Mahindra = A_2

World around = A_3

Himalaya travels = A_4

Affordable cost = C_1

Acceptable time = C_2

Interesting package = C_3

Based on the available information the affordable cost may be represented by the fuzzy set μ, where

$$\mu(x) = 1, x \leq 150{,}000$$
$$= (210{,}000 - 0.4x)/150{,}000, x > 150{,}000$$

Then,

$$\mu(100{,}000) = \mu(150{,}000) = 1$$
$$\mu(200{,}000) = 0.86$$
$$\mu(175{,}000) = 0.93$$

Similarly, acceptable time may be represented by the fuzzy set υ, where

$$\upsilon(x) = (x - 100)/60 + 1$$

Then,

$$\upsilon(100) = 1$$
$$\upsilon(150) = 1.83$$
$$\upsilon(200) = 2.6$$
$$\upsilon(125) = 1.41$$

Table 5.18.1. Matrix representation of alternatives, criteria and payoff.

	C_1	C_2	...	C_n
A_1	x_{11}	x_{12}	...	x_{1n}
A_2	x_{21}	x_{22}	...	x_{2n}
⋮	⋮	⋮		⋮
A_m	x_{m1}	x_{m2}	...	x_{mn}

Assume that a tourist expert opined that the packages are interesting with the degrees 0.4, 0.3, 0.6, 0.5 respectively.

The criteria satisfied by the alternatives are expressed in a matrix form as shown in Table 5.18.2.

Suppose the decision maker assigns weights of 0.3, 0.2 and 0.5 to the three criteria respectively. If m denotes the basic mass assignment function, then

$$m(C_1) = 0.3,$$
$$m(C_2) = 0.2,$$
$$m(C_3) = 0.5.$$

Calculation of Bel gives the following :

$$\text{Bel}(A_1) = 0.866$$
$$\text{Bel}(A_2) = 0.928$$
$$\text{Bel}(A_3) = 0.8$$
$$\text{Bel}(A_4) = 0.859$$

$\therefore$ From the above analysis one can concluded that the second package, *i.e.*, club Mahindra is the best alternative.

Table 5.18.2. Degree to which criteria are satisfied by the alternatives.

	C_1	C_2	C_3
	(Affordable Cost)	**(Acceptable Time)**	**(Interesting Package)**
A_1	**1.0**	**1.83**	**0.4**
A_2	**0.86**	**2.6**	**0.3**
A_3	**1.0**	**1.0**	**0.6**
A_4	**0.93**	**1.41**	**0.5**

Que 5.19. **Use GA to solve the following non-linear programming problem :**
Minimize $(x-2.5)^2+(y-5)^2$ subject to $5.5x+2y^2-18\le 0, 0\le x, y\ge 5$.

AKTU 2016-17(Sem-IV), Marks 15

Answer

Let $$f(x) = (x-2.5)^2+(y-5)^2$$
$$= x^2+y^2-5x-10y+31.25$$

To find minimum value differentiate partially w.r.t x and y,

$$\frac{\partial}{\partial x}(x^2+y^2-5x-10y+31.25) = 2x-5$$

$$f'(x) = 0 \Rightarrow 2x-5=0 \quad x=5/2 \quad x=2.5\approx 3$$

$$\frac{\partial^2}{\partial x^2}(2x-5) = 2>0$$

Since $f''(x)>0$ so, minimum value of $f(x)$ is 3.

$$\frac{\partial}{\partial y}(x^2+y^2-5x-10y+31.25) = 2y-10$$

$$f'(x) = 0 \quad 2y-10=0 \quad y=5$$

$$\frac{\partial^2}{\partial y^2}(2y-10) = 2>0$$

Since $f''(x)>0$ so, minimum value of $f(x)$ is 5.

Que 5.20. **Let a function $f(x) = x-x^2/16$ be defined on the interval [0, 31]. Illustrate the use of GA for determining the maximum of the given function.**

AKTU 2017-18(Sem-III), Marks 10

Answer

1. Genetic algorithm for determining the maximum of function $f(x)=x-x^2/16$ in the given interval [0, 31], the interval is approximated by 32 integer points *i.e.*, 0, 1, ..., 31, and these points are coded by the corresponding binary numbers. Then, gene pool $G = \{0, 1\}$ and all possible chromosomes are binary integers from 00000 to 11111.
2. Assume $m = 4$ and $P^{(1)} = \{00010, 01001, 10011, 11000\}$ in step 1 as shown in Table 5.20.1(*a*). Using function f as the fitness function, the fitness of each chromosome in $P^{(1)}$ is calculated using step 2.
3. Then, using the deterministic sampling in step 3, the population $P_n^{(1)} = (01001, 10011, 10011, 11000)$ as shown in Table 5.20.1(*b*) is obtained. If given stopping criteria in step 4 are not met, proceed to step 5.
4. Assuming that the condition $P_n^{(k)} = P^{(k)}$ was chosen as the stopping criterion, the algorithm does not stop at this point and proceeds to

step 5. In this step, lets assume that only simple crossovers are used, each of which produces one of the two possible offsprings.

5. For each x in $P_n^{(1)}$, a mate y in $P_n^{(1)}$ and a crossover point are chosen randomly and, then, the offsprings x' is produced as shown in Table 5.20.1(b). Next, in step 6, the old population $P_n^{(1)}$, is replaced with the new population $P^{(2)}$ of offsprings produced in step 5, k is increased by one, and then proceed to step 2. Step 2 and 3 are now repeated for $k = 2$, and the results are shown in Table 5.20.1(c).

6. The stopping criterion in step 4 is again not satisfied; consequently, proceed to step 5. The result of this step is shown in Table 5.20.1(d).

7. In step 6, replace $P^{(2)}$ with $P^{(3)}$, increase k by one, and proceed to step 2. The application of steps 2 and 3 for $k = 3$ results in $P_n^{(3)}$, shown in Table 5.20.1(e). Now, the stopping criterion $P_n^{(3)} = P^{(3)}$ is satisfied in step 4, and the algorithm terminates.

8. The chromosome 10000, which has the highest fitness, represents the solution. This chromosome corresponds to the integer 16 which is, indeed, the point for which the function f reaches its maximum.

Table 5.20.1. Illustration of genetic algorithm

a. $k = 1$: Step 2 and 3 where P : population, $g(x)$ = relative fitness :

Chromosome in $P^{(1)}$	Integers	Fitness	$g(x)$	$4g(x)$ $m = 4$	Number of selected copies
00010	2	13.75	0.360	1.444	0
01001	9	15.94	0.418	1.672	1
10011	19	8.44	0.483	1.932	2
11000	24	0	0	0	1
	$\Sigma f(x) =$	38.127			

b. $k = 1$: Step 5 :

Chromosome in $P_n^{(1)}$	Mate (randomly selected)	Crossover site (randomly selected)	Resulting chromosomes in $P^{(2)}$
01001	10011	3	01011
10011	01001	3	10001
10011	11000	1	11000
11000	10011	1	10011

Similarly for the values of k 2, 3, values are set to calculate the fitness value, mate, and the crossover site.

c. $k = 2$: Step 2 and 3 :

Chromosome in $P^{(2)}$	Integers	Fitness	$g(x)$	$4g(x)$	Number of selected copies
01011	11	15.437	0.443	1.772	2
10001	17	10.937	0.314	1.256	2
11000	24	0	0	0	0
10011	19	8.438	0.242	0.968	1

d. $k = 2$: Step 5 :

Chromosome in $P_n^{(2)}$	Mate (randomly selected)	Crossover site (randomly selected)	Resulting chromosomes in $P^{(2)}$
10001	3	2	10011
10001	4	3	10000
11000	1	2	11001
10011	2	3	10010

e. $k = 3$: Step 2 and 3 :

Chromosome in $P^{(2)}$	Integers	Fitness	$g(x)$	$4g(x)$	Number of selected copies
10011	19	10.5	0.279	1.112	1
10000	16	14.0625	0.374	1.496	1
11001	25	0	0	0	1
10001	17	13	0.346	1.384	1

VERY IMPORTANT QUESTIONS

Following questions are very important. These questions may be asked in your SESSIONALS as well as UNIVERSITY EXAMINATION.

Q. 1. Explain genetic algorithm. Also draw and explain the flow chart of genetic algorithm.

Ans: Refer Q. 5.1.

Q. 2. Draw a flowchart of GA and explain the working principle.
Ans: Refer Q. 5.2.

Q. 3. Write short notes on procedures of GA.
Ans: Refer Q. 5.3.

Q. 4. Write short notes of genetic representation.
Ans: Refer Q. 5.5.

Q. 5. Give the detail of genetic representation (Encoding).
Ans: Refer Q. 5.6.

Q. 6. Explain different methods of selection in genetic algorithm in order to select a population for next generation.
Ans: Refer Q. 5.7.

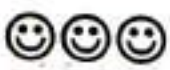
Q. 7. What is simulated annealing ? What is the structure of a simulated annealing algorithm ?
Ans: Refer Q. 5.9.

Q. 8. What are the genetic operators ? What are the roles of genetic operators in GA ?
Ans: Refer Q. 5.10.

Q. 9. What are the mutation in GA ? Explain the generational cycle in GA.
Ans: Refer Q. 5.11.

Q. 10. What is the main function of crossover operation in genetic algorithm ?
Ans: Refer Q. 5.14.

Q. 11. Write short notes on application of GA.
Ans: Refer Q. 5.15.



Neural Network-I (Introduction and Architecture) (2 Marks Questions)

1.1. What are neurons ? **AKTU 2015-16(Sem-III), Marks 02**

Ans: A neuron is a small cell that receives electro-chemical signals from its various sources and in return responds by transmitting electrical impulses to other neurons.

1.2. What is soft computing ?

AKTU 2015-16(Sem-IV), Marks 02

Ans: Soft computing refers to a collection of computational techniques in computer science, machine learning and some engineering disciplines which study, model and analyze very complex phenomena.

1.3. Give the difference between supervised and unsupervised learning in artificial neural network.

AKTU 2015-16(Sem-IV), Marks 02

AKTU 2015-16(Sem-III), Marks 02

Ans:

S. No.	Supervised learning	Unsupervised learning
1.	It uses known and labeled data as input.	It uses unknown data as input.
2.	It uses offline analysis.	It uses real time analysis of data.
3.	Number of classes are known.	Number of classes are not known.
4.	Accurate and reliable results.	Moderate accurate and reliable results.

1.4. Define activation function.

Ans: An activation function is the basic element in neural model. It is used for limiting the amplitude of the output of a neuron. It is also called squashing function.

1.5. Give types of activation function.

OR

What are the different activation functions ? Explain.

AKTU 2017-18(Sem-III), Marks 02

Ans: Types of activation function :

i. Signum function
ii. Sigmoidal function
iii. Identity function
iv. Binary step function
v. Bipolar step function

1.6. Give advantages of neural network.

Ans: Advantages of neural network :

i. A neural network can perform tasks that a linear program cannot.
ii. It can be implemented in any application.
iii. A neural network learns and does not need to be reprogrammed.

1.7. What are disadvantages of neural network (NN) ?

Ans: Disadvantages of neural network :

i. The neural network needs training to operate.
ii. It requires high processing time for large NN.

1.8. List the various types of soft computing techniques and mention some application areas for neural network.

AKTU 2015-16(Sem-IV), Marks 02

Ans: Types of soft computing techniques :

1. Fuzzy logic control
2. Neural network
3. Genetic algorithms
4. Support vector machine

Application areas for neural network :

1. Speech recognition
2. Character recognition
3. Signature verification application
4. Human face recognition

1.9. Draw a biological NN and explain the parts.

AKTU 2015-16(Sem-IV), Marks 02

Ans:

i. Biological neural networks are made up of real biological neurons that are connected in the peripheral nervous system.

ii. In general a biological neural network is composed of a group of chemically connected or functionally associated neurons.

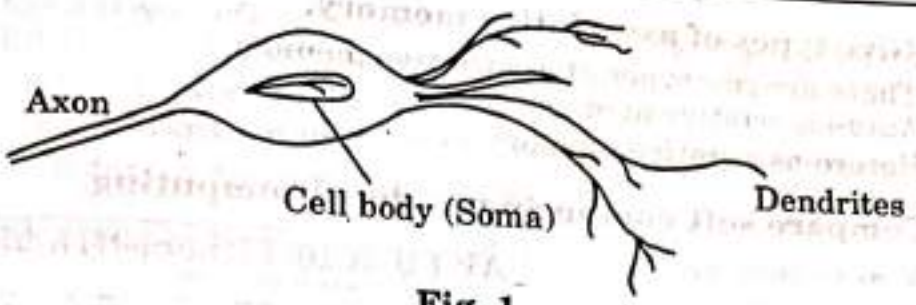


Fig. 1.

A biological neural network has three major parts :

1. **Soma or cell body :** It contains the cell's nucleus and other vital components called organelles which perform specialized tasks.
2. **A set of dendrites :** It forms a tree like structure that spread out from the cell. The neuron receives its input electrical signal along these set of dendrites.
3. **Axon :** It is tabular extension from the cell (Soma) that carries an electrical signal away from Soma to another neuron for processing.

1.10. What is single layer feed forward network ?

Ans. Single layer feed forward network is the simplest form of a layered network where an input layer of source nodes that projects onto an output layer of neurons, but not vice versa.

1.11. Write different applications of neural networks (NN).

AKTU 2016-17(Sem-IV), Marks 02

Ans. **Application of NN are :**

1. Image recognition
2. Data mining
3. Machine translation
4. Spell checking
5. Stock and sport bet prediction
6. Statistical modeling

1.12. Draw an artificial neural network.

AKTU 2015-16(Sem-IV), Marks 02

Ans. An Artificial Neuron Network (ANN) is a computational model based on the structure and functions of biological neural networks.

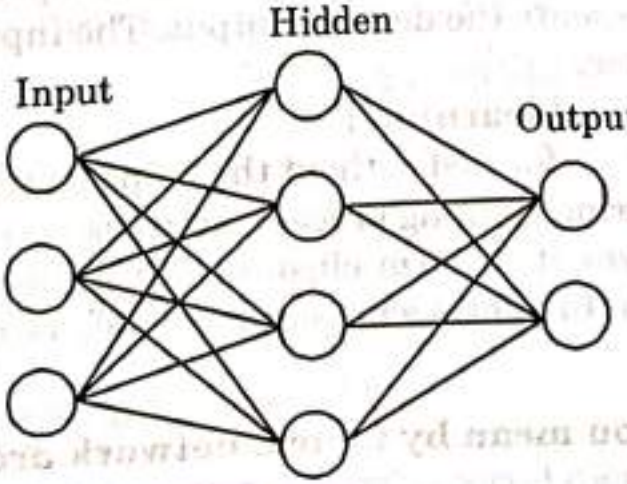


Fig. 2.

1.13. Give types of associative memory.

Ans: There are two types of associative memory :

i. Auto-associative memory

ii. Hetero-associative memory

1.14. Compare soft computing vs. hard computing.

AKTU 2016-17(Sem-III), Marks 02

AKTU 2017-18(Sem-III), Marks 02

OR

How soft computing is different from conventional computing ? **AKTU 2017-18(Sem-IV), Marks 02**

Ans:

S. No.	Soft computing	Hard (Conventional) computing
1.	Soft computing is liberal of imprecision, uncertainty, partial truth and approximation.	Hard computing requires a precisely state analytic model.
2.	Soft computing is based on fuzzy logic, neural sets, and probabilistic reasoning.	Hard computing is based on binary logic, crisp system, numerical analysis and crisp software.
3.	Soft computing will emerge its own programs.	Hard computing requires programs to be written

1.15. Define supervised and unsupervised learning in artificial neural network. **AKTU 2016-17(Sem-III), Marks 02**

Ans:

1. Supervised learning :

i. Learning is performed with the help of a trainer.

ii. In ANN, each input vector requires a corresponding target vector, which represents the desired output. The input vector results in output vector.

2. Unsupervised learning :

i. Learning is performed without the help of a trainer.

ii. In ANN, during training process, network receives input patterns and organizes it to form clusters. No feedback is applied from environment to inform what output should be or whether they are correct.

1.16. What do you mean by neural network architecture ?

AKTU 2016-17(Sem-III), Marks 02

Ans: Neural network architecture refers to the arrangement of neurons into layers and the connection patterns between layers, activation functions, and learning methods. The neural network model and the architecture of a neural network determine how a network transforms its input into an output.

1.17. Name some application of competitive learning network.

AKTU 2016-17(Sem-III), Marks 02

Ans: **Applications of competitive learning networks :**

1. Vector quantization
2. Analyze raw data
3. Image browsing systems
4. Medical diagnosis (visualization of data)
5. Speech recognition
6. Data compression
7. Feature extraction

1.18. Artificial Intelligence can be used in Neural Network or not. Justify your answer.

AKTU 2016-17(Sem-IV), Marks 02

Ans: Artificial intelligence can be used in neural networks because we are able to combine the desirable features of adaptivity, robustness and uniformity offered by neural networks with the representation, inference and universality that are special features of symbolic artificial networks.

1.19. What is reinforcement learning ?

AKTU 2016-17(Sem-IV), Marks 02

Ans: Reinforcement learning is the type of dynamic programming that trains algorithms by using a system of reward and punishment. The agent learns by interacting with its environment to achieve a goal in an uncertain, potentially complex environment.

1.20. What are the various learning rules ? Explain.

AKTU 2017-18(Sem-III), Marks 02

Ans: **Different learning rules in the neural network :**

1. **Hebbian learning rule :** It identifies how to modify the weights of nodes of a network.
2. **Perceptron learning rule :** Network starts its learning by assigning a random value to each weight.
3. **Delta learning rule :** Modification in weight of a node is equal to the multiplication of error and the input.
4. **Correlation learning rule :** The correlation rule is the supervised learning.

5. **Outstar learning rule :** We can use it when nodes or neurons in a network are arranged in a layer.

1.21. What do you mean hetro-associative memory ?

AKTU 2017-18(Sem-III), Marks 02

Ans:

1. In a hetero-associative memory, the retrieved pattern is different from the input pattern not only in content but also different in type and format.
2. In hetero-associative memory, the training input and target output vectors are different.

1.22. Why sigmoid function is so important activation function in neural network ?

AKTU 2017-18(Sem-IV), Marks 02

Ans: A sigmoid function takes any real number as input and returns a value between 0 and 1. Therefore, it is especially used for models where we have to predict the probability as an output. Since probablity of anything exists only between the range of 0 and 1, so sigmoid function is important in neural network.

1.23. Whether the Hebbian learning rule is an example of unsupervised learning or not, justify your answer ?

AKTU 2017-18(Sem-IV), Marks 02

Ans:

1. The Hebbian learning rule states that if two neighbour neurons are activated and deactivated at the same time, then the weight connecting these neurons should increase.
2. For neurons operating in the opposite phase, the weight between them should decrease. If there is no signal correlation, the weight should not change.
3. Response of neurons is not taken into account. Since desired responses of neurons are not used in the learning procedure, this is the unsupervised learning rule. Therefore, Hebbian learning rule is an example of unsupervised learning.

1.24. Consider an auto associative net with the bipolar step function as the activation function and weights set by Hebb rule (outer diagonal) where the main diagonal of the weight matrix is set to zero. Find the weight matrix to store the vector $v_1 = (1\ 1\ 1\ 1\ -1\ -1)$.

AKTU 2015-16(Sem-III), Marks 02

AKTU 2015-16(Sem-IV), Marks 02

Ans: The vectors v_1 is stored with weight matrix

$$w = \sum_{p=1}^{p} S^T(p)\, S(p)$$

$$w = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} [1 \quad 1 \quad 1 \quad 1 \quad -1 \quad -1]$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 \end{bmatrix}$$

Setting diagonal to zero we have

$$= \begin{bmatrix} 0 & 1 & 1 & 1 & -1 & -1 \\ 1 & 0 & 1 & 1 & -1 & -1 \\ 1 & 1 & 0 & 1 & -1 & -1 \\ 1 & 1 & 1 & 0 & -1 & -1 \\ -1 & -1 & -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & -1 & 1 & 0 \end{bmatrix}$$

1.25. Use the Hebb rule to store the vector (1 1 –1 –1) in an auto associative neural network. Find the weight matrix.

AKTU 2015-16(Sem-IV), Marks 02

Ans: By using Hebb rule, the stored vector with weight matrix is,

$$w = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} [1 \;\; 1 \; -1 \; -1] = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$$

1.26. Use the Hebb rule to store the vector [1 1 1 – 1] in an auto-associative neural network.

AKTU 2016-17(Sem-IV), Marks 02

Ans: The vector s is stored with the weight matrix as :

$$w = s^T s$$

$$w = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} [1\ 1\ 1\ -1] = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$$

1.27. Name any three commercial software used for soft computing techniques.

AKTU 2015-16(Sem-III), Marks 02

Ans: **Three commercial software used for soft computing techniques are :**

1. Open CV 2. MATLAB 3. LabVIEW

1.28. What is leaky learning ?

AKTU 2016-17(Sem-III), Marks 02

AKTU 2017-18(Sem-III), Marks 02

Ans: Leaky learning is a variant of the competitive learning algorithm which requires all weights to learn on each time step, rather than just the winning weight.

1.29. Which neural network architecture is used for online spell checking ?

AKTU 2015-16(Sem-III), Marks 02

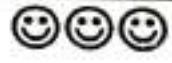
Ans: Personalized Error Correction using Neural Network (PENN) is used for online spell checking.

1.30. What is difference between autoassociative memory and heteroassociative memory ?

AKTU 2015-16(Sem-III), Marks 02

Ans:

S. No.	Autoassociative Memory	Heteroassociative Memory
1.	In autoassociative memory, the associated pattern pair (x, y) is same.	In heteroassociative memory, the associated pattern pair (x, y) is different.
2.	In this model, given a distorted or a partial input pattern x as input, the whole pattern with perfect form y is recalled as output.	This model recalls output y on a given input x or vice-versa.
3.	Useful for image refinement.	Useful for the association of patterns.



2 UNIT

Neural Network-II (Back Propagation Network) (2 Marks Questions)

2.1. Give characteristics of multilayer perceptron.

Ans: Characteristics of multilayer perceptron :

i. In this model, each neuron in the network includes a non-linear activation function.
ii. The network contains hidden neurons that are not a part of input or output of the network.
iii. The network exhibits a high degree of connectivity.

2.2. Give the name of the tuning parameters of back propagation neural network. **AKTU 2017-18(Sem-IV), Marks 02**

Ans: Tuning parameters of back propagation neural network are :

i. Momentum factor
ii. Learning coefficient
iii. Sigmoidal gain
iv. Threshold value

2.3. What do you mean by momentum factor ?

Ans: Momentum factor determines the step size of change in weights or biases. If momentum factor is zero, the smoothening is minimum and entire weight adjustment comes from newly calculated change. If it is one, new adjustment is ignored and previous one is repeated.

2.4. What are the various parameters for selection in BPN ?

Ans: Various parameters for selection in BPN are :

i. Number of hidden nodes
ii. Momentum coefficient (α)
iii. Sigmoidal gain (λ)
iv. Local minima
v. Learning coefficient (η)

2.5. Give applications of back propagation network.

Ans: Applications of back propagation networks are :

i. Optical character recognition.
ii. Non-linear simulation.
iii. Data compression.

2.6. Define ADALINE network.

Ans. The adaptive linear neural element network makes use of supervised learning. In this, there is only one output neuron and the output values are bipolar (– 1 or + 1).

2.7. Write a short note on MADALINE network.

Ans. A MADALINE (many ADALINE) network is created by combining a number of ADALINES. The use of multiple ADALINES helps counter the problem of non-linear separability. In this, each ADALINE unit receives the input bits x_1, x_2 and the bias input $x_0 = 1$ as its inputs.

2.8. If the net input to an output neuron is 0.64, calculate its output when the activation function is binary sigmoidal.

AKTU 2015-16(Sem-III), Marks 02

AKTU 2017-18(Sem-III), Marks 02

Ans. For binary sigmoidal,

$$y = f(\text{net}) = \frac{1}{1 + e^{-\lambda(\text{net})}}$$

$$y = \frac{1}{1 + e^{-0.64}} = 0.6548$$

3 UNIT

Fuzzy Logic-I (Introduction) (2 Marks Questions)

3.1. What is the difference between crisp set and fuzzy set ?

AKTU 2015-16(Sem-III), Marks 02

AKTU 2016-17(Sem-III), Marks 02

Ans:

S.No.	Crisp set	Fuzzy set
1.	In crisp set, an element either belongs to or does not belong to a set.	Fuzzy sets support a flexible sense of membership of elements to a set.
2.	Membership value in crisp set is 0 and 1.	Membership value in fuzzy set is in between 0 and 1.

3.2. Define a fuzzy Cartesian product.

AKTU 2016-17(Sem-III), Marks 02

Ans: Let $\tilde{A}$ be a fuzzy set defined on the universe X and $\tilde{B}$ be a fuzzy set defined on the universe Y. The Cartesian product between that fuzzy sets $\tilde{A}$ and $\tilde{B}$ indicated as $\tilde{A} \times \tilde{B}$ and resulting in a fuzzy relation $\tilde{R}$ is given by

$$\tilde{R} = \tilde{A} \times \tilde{B} \subset X \times Y$$

where $\tilde{R}$ has its membership function given by

$$\mu_{\tilde{R}}(x, y) = \mu_{\tilde{A} \times \tilde{B}}(x, y)$$
$$= \min(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y))$$

3.3. What are the different fuzzy relation operations ?

AKTU 2017-18(Sem-III), Marks 02

Ans. Different fuzzy relation operations :

1. Union
2. Intersection
3. Complement
4. Composition

3.4. Give the industrial applications of fuzzy logic.

AKTU 2017-18(Sem-III), Marks 02

Ans. Industrial applications of fuzzy logic :

1. Aerospace
2. Automotive
3. Defence
4. Finance
5. Medical

3.5. What is set ?

Ans. A set is a well defined collection of objects. A set in certain contexts may be associated with its universal set from which it is derived. For example :

$$A = \{x \mid x \text{ is an odd number}\}$$

3.6. What do you mean by superset ?

Ans. Given sets A and B on E the universal set, A is said to be a superset of B if every element of B is contained in A. It is denoted as $A \supset B$.

3.7. Give operations on crisp sets.

Ans. Operation on crisp sets are :

i. Union ($\cup$)
ii. Intersection ($\cap$)
iii. Complement (c)
iv. Difference ($-$)

3.8. List down properties of crisp set.

Ans. Properties of crisp set are :

i. Commutativity
ii. Associativity
iii. Distributivity
iv. Idempotence
v. Identity
vi. Law of absorption
vii. Law of excluded middle
viii. Law of contradiction

3.9. Define covering.

Ans. A covering on A is defined to be a set of non-empty subsets A_i whose union yields the original set A. The non-empty subsets need not be disjoint.

3.10. Give properties of fuzzy sets.

Ans. Properties of fuzzy sets are :

i. Commutativity
ii. Associativity

iii. Distributivity
iv. Idempotence
v. Identity
vi. Transitivity
vii. Involution
viii. De Morgan's law

3.11. Write a short note on partition.

Ans: A partition on A is defined to be a set of non-empty subsets A_i, each of which is pairwise disjoint and whose union yields the original set A. Partition on A indicated as $\prod(A)$.

3.12. Let $\tilde{A}$ and $\tilde{B}$ be the two fuzzy sets given by : $\tilde{A} = \{(x_1, 0.4), (x_2, 0.7), (x_3, 0.6)\}$ $\tilde{B} = \{(x_1, 0.5), (x_2, 0.3), (x_3, 0.1)\}$. Find the membership value of x_1 and x_2 in $A \cap B$.

AKTU 2017-18(Sem-IV), Marks 02

Ans:

$$\tilde{A} = \{(x_1, 0.4), (x_2, 0.7), (x_3, 0.6)\}$$

$$\tilde{B} = \{(x_1, 0.5), (x_2, 0.3), (x_3, 0.1)\}$$

$$\tilde{A} \cap \tilde{B} = \{(x_1, 0.4), (x_2, 0.3), (x_3, 0.1)\}$$

$$\mu_{\tilde{A} \cap \tilde{B}}(x_1) = \min\ (\mu_{\tilde{A}}(x_1), (\mu_{\tilde{B}}(x_1))$$
$$= \min(0.4, 0.5) = 0.4$$

$$\mu_{\tilde{A} \cap \tilde{B}}(x_2) = \min\ (\mu_{\tilde{A}}(x_2), \mu_{\tilde{B}}(x_1))$$
$$= \min(0.7, 0.3) = 0.3$$

☺☺☺

4 UNIT

Fuzzy Logic-II (Fuzzy Membership, Rules) (2 Marks Questions)

4.1. What are the disadvantages of fuzzy systems ?

AKTU 2016-17(Sem-III), Marks 02

Ans: Disadvantages of fuzzy system :

1. Fuzzy systems do not have the capability of machine learning as-well-as neural network type pattern recognition.
2. Expert input or instructions are needed in order to define fuzzy rules.
3. The process of tuning of the system parameters requires a relatively long time, if there is a high number of fuzzy rules in the system.

4.2. Define the fuzzy inference.

AKTU 2016-17(Sem-IV), Marks 02

Ans: Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. It is a mapping that provides a basis from which decisions can be made. Fuzzy inference systems have been successfully applied in fields such as automatic control, data classification, decision analysis, expert systems, and computer vision.

4.3. What is FLC ?

AKTU 2015-16(Sem-IV), Marks 02

AKTU 2016-17(Sem-IV), Marks 02

Ans: A Fuzzy Logic Controller (FLC) is a rule based system incorporating the flexibility of human decision making which is used for fuzzy structural optimization. It is a control law that is described by a knowledge-based system consisting of IF...THEN rules with vague predicates and a fuzzy logic inference mechanism.

4.4. Give the industrial applications of fuzzy logic.

Ans: Industrial applications of fuzzy logic :

1. Aerospace
2. Automotive

3. Defence
4. Finance
5. Medical

4.5. Draw fuzzy membership function to describe cold, warm and hot water.

AKTU 2017-18(Sem-IV), Marks 02

Ans:

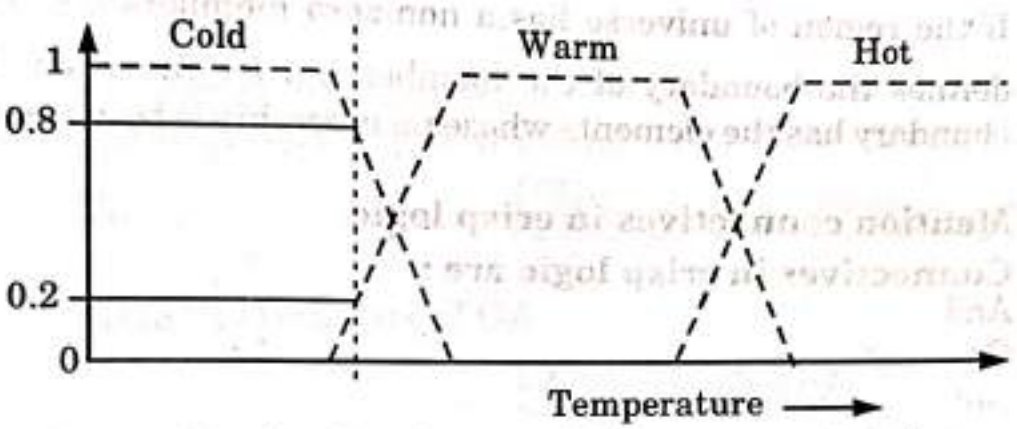


Fig. 1. Membership functions for water temperature.

1. Fig. 2 illustrates the membership functions, and shows the degree of particular temperature, is either cold, warm or hot.
2. The meaning of the expressions cold, warm, and hot is represented by functions mapping a temperature scale.
3. A point on that scale, where the vertical line is drawn, has three memberships.
4. The vertical line in the image represents a particular temperature that is too cold, warm and not hot.
5. Since the 'hot' function is at 0, this temperature may be interpreted as 'not hot'.
6. The value of warm function is about 0.2, so it may be described by 'slightly warm', and the cold function is about 0.8, indicating it is 'fairly cold'.

4.6. Give features of membership functions.

Ans: Features of membership function are :

i. Core
ii. Support
iii. Boundary

4.7. Define core.

Ans: If the region of universe is characterized by full membership (*i.e.*, 1) in the set $\widetilde{A}$, then this gives the core of the membership function of fuzzy at $\widetilde{A}$. The elements which have the membership function as 1, are the elements of core.

4.8. What is support ?

Ans: If the region of universe is characterized by non-zero membership in the set $\widetilde{A}$, then this defined the support of a membership function. The support has the elements whose membership is greater than 0.

4.9. Explain boundary.

Ans: If the region of universe has a non-zero membership, then this defines the boundary of the membership for fuzzy set $\widetilde{A}$. The boundary has the elements whose membership is between 0 and 1.

4.10. Mention connectives in crisp logic.

Ans: Connectives in crisp logic are :

i. And
ii. Or
iii. Not
iv. Implication
v. Equality

4.11. Give connectives in fuzzy logic.

Ans: Connectives in fuzzy logic are :

i. Negation
ii. Disjunction
iii. Conjunction
iv. Implication

4.12. List the methods used for defuzzification.

Ans: Methods used for defuzzification are :

i. Centroid method
ii. Centre of sums (COS) method
iii. Mean of maxima (MOM) defuzzification

4.13. Give two applications of fuzzy systems.

Ans: Two application of fuzzy system are :

i. Greg Viot's fuzzy cruise control system
ii. Yamakawa's air conditioner controller

4.14. What is the significance of fuzzy quantifier ?

AKTU 2016-17(Sem-IV), Marks 02

Ans: Significance of fuzzy quantifier : It provide an approximate idea of the number of elements in a subset fulfilling a certain condition.

5 UNIT

Genetic Algorithm (GA) (2 Marks Questions)

5.1. Define genetic algorithm and write down the advantages of GA.

AKTU 2016-17(Sem-III), Marks 02

OR

Write the benefits of GA.

AKTU 2016-17(Sem-IV), Marks 02

OR

Write benefits of genetic algorithm.

AKTU 2015-16(Sem-IV), Marks 02

OR

What are the advantages of GA over conventional algorithms ?

AKTU 2017-18(Sem-III), Marks 02

Ans. Genetic algorithms are computerized search and optimization algorithm based on mechanics of natural genetics and natural selection. GA converts the design space into genetic space.

Advantages of GA :

1. Easy to understand.
2. Supports multi-objective optimization.
3. Good for noisy environment.
4. Inherently parallel and easily distributed.

5.2. Define mutation.

AKTU 2015-16(Sem-IV), Marks 02

AKTU 2016-17(Sem-III), Marks 02

OR

What is mutation ?

AKTU 2016-17(Sem-IV), Marks 02

Ans. Mutation is the process by which a random deformation of the string after crossover is done with a certain probability to produce a modified generation.

5.3. What is simulated annealing ?

AKTU 2017-18(Sem-III), Marks 02

Ans: Simulated Annealing (SA) is an effective and general form of optimization. It is useful in finding global optima in the presence of large numbers of local optima. "Annealing" refers to an analogy with thermodynamics, specifically with the way that metals cool and anneal. Simulated annealing uses the objective function of an optimization problem instead of the energy of a material.

5.4. Describe in brief tree encoding in GA.

AKTU 2017-18(Sem-IV), Marks 02

Ans: Tree encoding is used for evolving programs or expressions, for genetic programming. In tree encoding, every chromosome is a tree of some objects, such as functions or commands in programming language.

5.5. Explain binary encoding in genetic algorithm.

AKTU 2015-16(Sem-III), Marks 02

Ans: **Binary encoding :**

i. Binary encoding is the most commonly used method of genetic representation because GA uses binary encoding.

ii. In binary encoding, every chromosome is a string of bits, 0 or 1.

For example :

Chromosome A10110010110010101011100101

iii. Binary encoding gives many possible chromosomes.

5.6. What is inheritance in genetic algorithm ?

AKTU 2015-16(Sem-III), Marks 02

Ans: Inheritance in genetic algorithm refers to the process of generating offspring that competes for survival to make up the next generation of population. Some of the genetic inheritance operators are reproduction, crossover, mutation, inversion, segregation etc.

5.7. List down various types of encoding of genetic algorithm.

Ans: **Various types of encoding of GA are :**

i. Binary encoding
ii. Octal encoding (0 – 7)
iii. Hexadecimal encoding (0123456789 ABCDEF)
iv. Permutation encoding
v. Value encoding
vi. Tree encoding

5.8. What are the basic operators of genetic algorithm ?

Ans: **Basic operators of genetic algorithm are :**

i. Selection or Reproduction

ii. Mutation
iii. Crossover

5.9. Give various selection methods in GA.

Ans: Various selection methods are :
i. Roulette-wheel selection
ii. Boltzmann selection
iii. Tournament selection
iv. Steady-state selection
v. Rank selection

5.10. Define genetic algorithm.

Ans: Genetic algorithms are computerized search and optimization algorithm based on mechanics of natural genetics and natural selection. These algorithms use the principle of natural genetics and natural selection to construct search and optimization procedure. GA converts the design space into genetic space.

5.11. Give some aspects of genetic algorithm.

Ans: Three most important aspects of genetic algorithm are :
i. Definition of objective function.
ii. Definition and implementation of genetic representation.
iii. Definition and implementation of genetic operators.

5.12. What is reproduction ?

Ans: Reproduction or selection is the first operator applied on population. Chromosomes are selected from the population to be parents to crossover and produce new population.

5.13. What do you mean by crossover operator ?

Ans: The crossover operator applied to the mating pool to create a better string. It searches a parameter space in a way that the information stored in the present string is maximally preserved. It is of three types :
i. Single-site crossover
ii. Two-point crossover
iii. Multi-point crossover

5.14. What do you mean by convergence of GA ?

AKTU 2016-17(Sem-IV), Marks 02

Ans: A Genetic Algorithm (GA) is said to be converged when there is no significant improvement in the values of fitness of the population from one generation to the next. It causes evolution to halt because every individual in the population is identical. Full convergence might be seen in genetic algorithms using crossover.

☺☺☺

B. Tech.
(SEM. VII) ODD SEMESTER THEORY EXAMINATION, 2019-20
APPLICATION OF SOFT COMPUTING

Time : 3 Hours **Max. Marks : 70**

Note : Attempt **all** sections. If require any missing data; then choose suitably.

Section-A

1. Attempt **all** questions in brief.

a. How AI can be used in neural network? Give proper justification.

Ans. A neural network is either a system software or hardware that works similar to the task performed by neurons of a human brain. Neural network include various technologies like deep learning, and machine learning as a part of AI. Artificial neural network is a tool of machine learning.

b. Differentiate the soft computing and hard computing.

Ans. Refer Q. 1.14, Page SQ–4F, Unit-1, Two Marks Question.

c. Draw biological neural network and explain each part.

Ans. Refer Q. 1.9, Page SQ–2F, Unit-1, Two Marks Question.

d. What is fuzzy quantifier ?

Ans. Fuzzy quantifiers allow us to express fuzzy quantities or proportions in order to provide an approximate idea of the number of elements of a subset fulfilling a certain condition or the proportion of the number in relation to the total number of possible elements.

e. Use the Hebb rule to store the vector [1 –1 1 –1] in an auto-associative neural network.

Ans: By using Hebb rule, the stored vector with weight matrix is,

$$w = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} [1 \ -1 \ \ 1 \ -1] = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$

f. Explain convergence in genetic algorithm,

Ans: Refer Q. 5.14, Page SQ–19F, Unit-5, Two Marks Question.

g. Explain conditional and unconditional fuzzy proposition.

Ans: **Conditional and unconditional fuzzy proposition :**

1. **Conditional and unqualified propositions :** The canonical form of this type of fuzzy proposition is p : if X is A, then Y is B, where, X, Y are variables in universes $U1$ and $U2$ and A, B are fuzzy sets on X, Y.
2. **Conditional and qualified propositions :** The canonical form of this type of fuzzy proposition is p : (if X is A, then Y is B) is S, where, all variables have same meaning.
3. **Unconditional and unqualified propositions :** The canonical form of this type of fuzzy proposition is p : V is F where, V is a variable which takes value V from a universal set U, F is a fuzzy set on U that represents a given inaccurate predicate such as fast, low, tall etc.
4. **Unconditional and qualified propositions :** The canonical form of this type of fuzzy proposition is p : V is F is S where, V and F have the same meaning and S is a fuzzy truth qualifier.

Section-B

2. Attempt any **three** of the following : **(7 × 3 = 21)**

a. Implement MADALINE network to solve XOR problem.

Ans: Refer Q. 2.12, Page 2–12F, Unit-2.

b. Explain generational cycle of GA with diagram.

Ans: Refer Q. 5.11, Page 5–16F, Unit-5.

c. Discuss the selection of various parameters in BPN.

Ans: Refer Q. 2.24, Page 2–25F, Unit-2.

d. Explain the Greg Voit's fuzzy cruise controller.

Ans: Refer Q. 4.21, Page 4–21F, Unit-4.

e. For the given input vectors $S = (S_1, S_2, S_3, S_4)$ and output vector $T = (T_1, T_2)$, find the weight matrix using hetero associative training algorithm.

$S = (S_1, S_2, S_3, S_4)$ $T = (T_1, T_2)$

I	(1, 1, 0, 0)	(0, 1)
II	(1, 0, 0, 1)	(1, 1)
III	(1, 1, 0, 0)	(1, 0)
IV	(0, 1, 0, 1)	(0, 0)

Ans: The weight matrix for S = I is calculated as :

$$S_1 = (1\ 1\ 0\ 0) \text{ and } t = (0\ 1)$$

$$W_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} [0\ 1] = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Similarly, for S = II,

$$S_2 = (1\ 0\ 0\ 1) \text{ and } t = (1\ 1)$$

$$W_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} [1\ 1] = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$$

For S = III, $S_3 = (1\ 1\ 0\ 0)$ and $t = (1\ 0)$

$$W_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} [1\ 0] = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

For S = IV, S_4 = (0 1 0 1) and t = (0 0)

$$W_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} [0\ 0] = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

The weight matrix to store all four pattern pairs is sum of the weight matrices to store each pattern pairs separately, *i.e.*,

$$= W_1 + W_2 + W_3 + W_4$$

$$= \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$$

Section-C

3. Attempt any **one** part of the following : (7 × 1 = 7)

a. What is simulated annealing ? Explain the structure of the simulated annealing algorithm.

Ans: Refer Q. 5.9, Page 5–10F, Unit-5.

b. Consider four destination wedding packages P1, P2, P3 and P4. We want to choose one. Their cost are INR 1,00,000, INR 1,50,000, INR 2,50,000 and INR 3,50,000. Their days for packages are 10, 20, 30 and 35 respectively. They are viewed as interesting with degrees 0.4, 0.3, 0.6, 0.5. Define your own fuzzy set of acceptable days of packages. Then determine the fuzzy set of interesting wedding packages whose cost and days for package are acceptable and use this set to choose one of your package.

Ans: Consider four wedding package P_1, P_2, P_3, P_4 with

Affordable cost = C_1

Acceptable day = C_2, and

Interesting wedding package = C_3

Based on the available information the affordable cost may be represented by the fuzzy set μ, where

$$\mu(x) = 1, x \leq 250{,}000$$

$$= (300{,}000 - 0.5x)/250{,}000, x > 250{,}000$$

Then,

$$\mu(100{,}000) = \mu(150{,}000) = \mu(250{,}000) = 1$$

$$\mu(350{,}000) = 0.5$$

Similarly, acceptable day may be represented by the fuzzy set v, where

$$v(x) = (x - 10)/7 + 1$$

Then, $v(10) = 1$

$v(20) = 2.42$

$v(30) = 3.85$

$v(35) = 4.57$

Table 1. Matrix representation of alternatives, criteria and payoff.

	C_1	C_2	...	C_n
P_1	x_{11}	x_{12}	...	x_{1n}
P_2	x_{21}	x_{22}	...	x_{2n}
.	.	.		.
.	.	.		.
.	.	.		.
P_m	x_{m1}	x_{m2}	...	x_{mn}

Assume that a tourist expert opined that the packages are interesting with the degrees 0.4, 0.3, 0.6, 0.5 respectively.

The criteria satisfied by the alternatives are expressed in a matrix form as shown in Table 2.

Suppose the decision maker assigns weights of 0.3, 0.2 and 0.5 to the three criteria respectively. If m denotes the basic mass assignment function, then

$$m(C_1) = 0.3,$$

$$m(C_2) = 0.2,$$

$$m(C_3) = 0.5.$$

Calculation of Bel gives the following :

$$\text{Bel}(P_1) = 0.7$$

$$\text{Bel}(P_2) = 0.934$$

$$\text{Bel}(P_3) = 1.37$$

$$\text{Bel}(P_4) = 1.314$$

$\therefore$ From the above analysis one can concluded that the third package is the best alternative.

Table 2. Degree to which criteria are satisfied by the alternatives.

	C_1	C_2	C_3
	(Affordable Cost)	(Acceptable Day)	(Interesting Package)
P_1	1	1	0.4
P_2	1	2.42	0.3
P_3	1	3.85	0.6
P_4	0.5	4.57	0.5

4. Attempt any **one** part of the following : $(7 \times 1 = 7)$

a. Write the expression of bipolar continuous and bipolar binary activation function.

Ans: Refer Q. 1.11, Page 1–11F, Unit-1.

b. Discuss the applications of GA in detail.

Ans: Refer Q. 5.15, Page 5–19F, Unit-5.

5. Attempt any **one** part of the following : **(7 × 1 = 7)**

a. Discuss the membership function and state its importance in fuzzy logic. Also discuss the features of membership functions.

Ans: Refer Q. 4.1, Page 4–2F, Unit-4.

b. Write short note on genetic representation.

Ans: Refer Q. 5.5, Page 5–5F, Unit-5.

6. Attempt any **one** part of the following : **(7 × 1 = 7)**

a. Explain back propagation algorithm and factors that may affect the back propagation neural network.

Ans: **Back propagation algorithm :** Refer Q. 2.20, Page 2–20F, Unit-2.

Factors affecting back propagation neural network : Refer Q. 2.23, Page 2–24F, Unit-2.

b. Explain fuzzy relations and fuzzy to crisp conversion in detail.

Ans: Refer Q. 3.11, Page 3–11F, Unit-3.

7. Attempt any **one** part of the following : **(7 × 1 = 7)**

a. Explain mutation and mutation rate with example.

Ans: Refer Q. 5.13, Page 5–18F, Unit-5.

For example : The string 00000100 might be mutated in its second position to yield 01000100. Mutation can occur at each bit position in a string with some probability usually very small (e.g., 0.001).

b. Write short note on the following :

i. Hopfield Network

ii. Supervised learning and Unsupervised learning

Ans.

i. Refer Q. 1.32, Page 1–35F, Unit-1.

ii. Refer Q. 1.19, Page 1–18F, Unit-1.