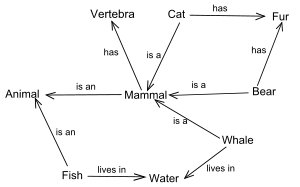
**Lists** - linked lists are used to represent hierarchical knowledge

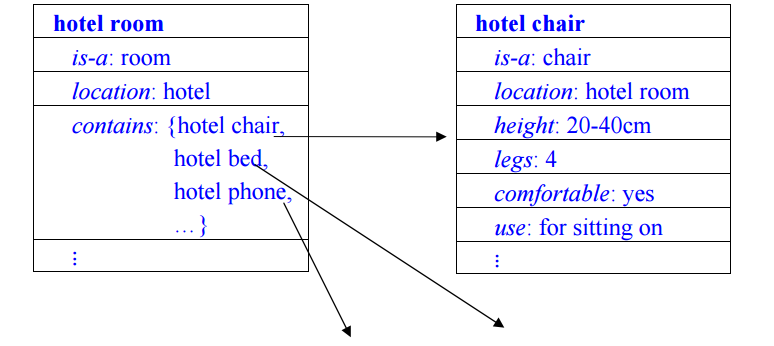
**Trees** - graphs which represent hierarchical knowledge. LISP, the main programming language of AI, was developed to process lists and trees.

**Semantic networks** - nodes and links - stored as propositions. A semantic network or net is a graphic notation for representing knowledge in patterns of interconnected nodes and arcs.



**Schemas** - used to represent commonsense or stereotyped knowledge.

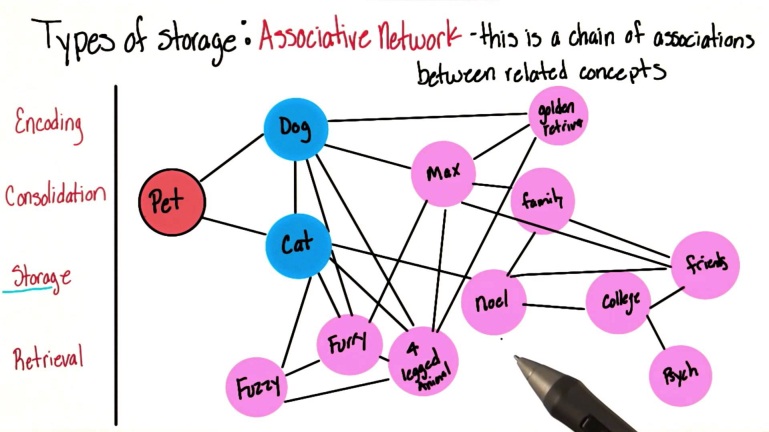
**Frames** - Describe objects. Consist of a cluster of nodes and links manipulated as a whole.Knowledge is organised in slots. Frames are hierarchically organised. Semantic network can be converted into frames. In short Frames are a OOPS representation of semantic net.



**Scripts**  - Describe event rather than objects. Consist of stereotypically ordered causal or temporal chain of events.

**Associative Networks:-**

Pictorial representation of objects, their attributes and relationships



**Conceptual Dependency:-**

Structures and primitives to represent sentences.

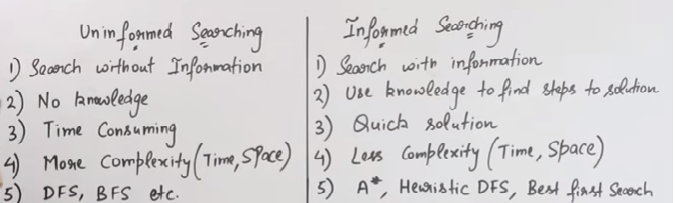
Know the fundamental search strategies and algorithms

**uninformed search :-**

breadth-first, depth-first, uniform-cost, iterative deepening, bidirectional

**informed search :-**

best-first (greedy, A\*), heuristics, memory-bounded



A **Horn clause** is a clause containing at most one positive literal.

A **Definite clause** contains exactly one positive literal.

A Horn clause with no positive literals is sometimes called a **goal clause**.

**Search algorithms:-**

There are several types of search algorithms that are routinely applied in planning. These include the following.

**well-known uninformed search** algorithms like depth-first search, breadth-first search

**systematic heuristic search algorithms** with optimality guarantees, for example A\* and its variants like IDA\*, WA\*

**systematic heuristic search algorithms** without optimality guarantees, for example the standard "best-first" search algorithm which is like A\* but ignores the cost-so-far component of the valuation function

**incomplete unsystematic search algorithms**, most notably stochastic search algorithms

**MEA (Means-Ends Analysis)** is an a approach that puts together aspects of both forward and backward reasoning in that both the condition and action portions of rules are considered when we decide which rules to apply. The logic of the process takes into account the gap between the current situation and the desired goal – where we wish to get to and proposes actions in order to close the gap between the two.

**Expert system shells** are the software containing an interface, an inference engine, and the formatted skeleton of a knowledge base. In essence, an expert system shell is an empty bowl to be filled with the expert knowledge elements that the inference engine may process for users.

Expert systems shells provide the bare bones for imitation of human expert reasoning in rule methods known as forward chaining and backward chaining

**Minimax** is a decision rule used in decision theory, game theory for minimizing the possible loss for a worst case (maximum loss) scenario. Originally formulated for two-player zero-sum game theory, covering both the cases where players take alternate moves and those where they make simultaneous moves, it has also been extended to more complex games and to general decision making in the presence of uncertainty.

**Steepest Ascent Hill Climbing(SAHC):-**

Steepest ascent hill climbing examines all possible next steps and chooses the best one. In this way, it is analogous to the use of best-first search to traverse graphs.Considers all moves from current state and selects best move.

Classic hill climbing takes the first step that it encounters which improves the objective function.

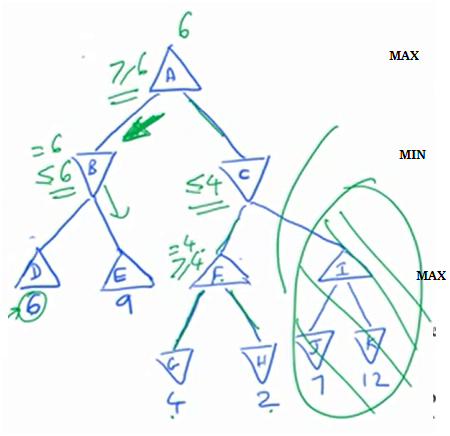
**Branch And Bound :-**

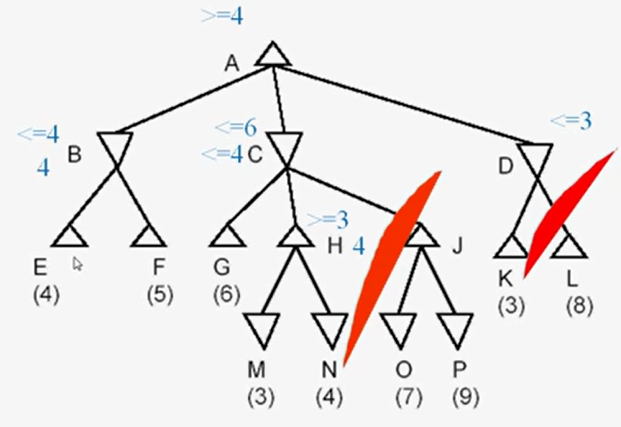
Keeps track of all partial paths which can be candidate for further exploration.

**Alpha-beta pruning**

-is a procedure to reduce the amount of computation and searching during minimax. It seeks to decrease the number of nodes that are evaluated by the minimax algorithm in its search tree. It is an adversarial search algorithm used commonly for machine playing of two-player games (Tic-tac-toe, Chess, Go, etc.). It stops completely evaluating a move when at least one possibility has been found that proves the move to be worse than a previously examined move.

-The strategy used to reduce the number of tree branches and the number of static evaluations applied in case of a game tree is





**best-first search** a search in which the next fringe node to be expanded is the node that is estimated to be the best according to an evaluation function. If the evaluation function f(n) is appropriately chosen, this becomes A\*.

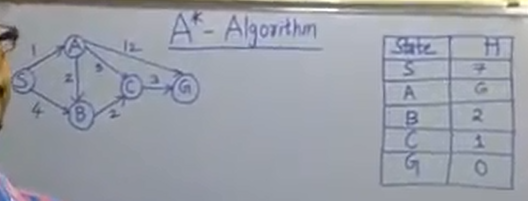
Fringe Node = The outer or less important part of an area, group, or activity.

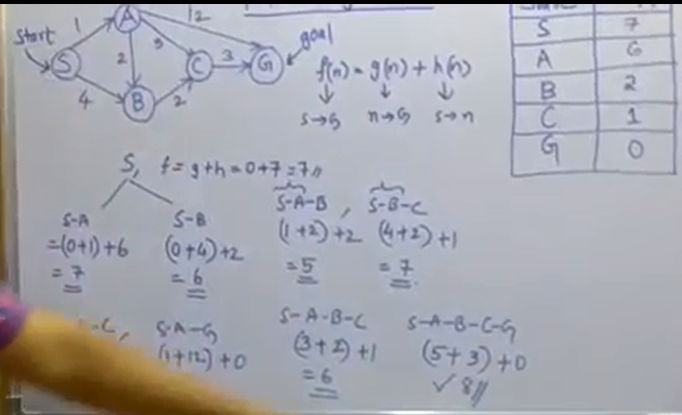
Ref: https://www.youtube.com/watch?v=i4MA\_hFkKDg&ab\_channel=Education4u

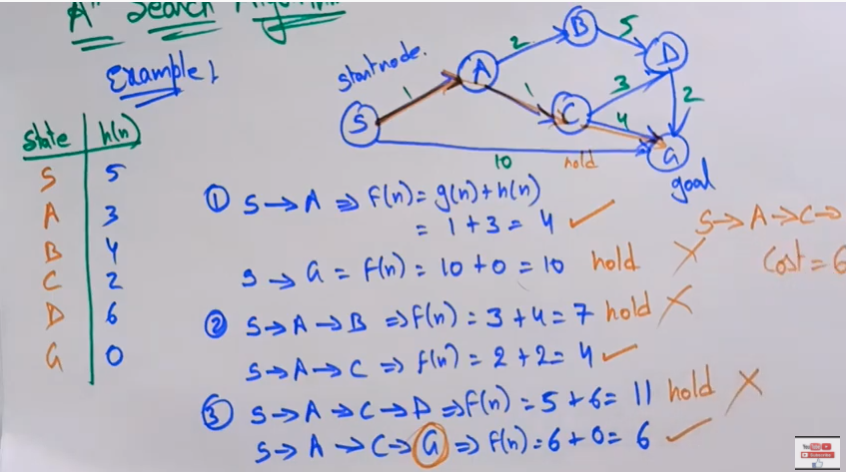
**A\* algorithm** (pronounced “A-star'')

-an algorithm for heuristic search, using an admissible heuristic function f(n) = g(n) + h(n), where g(n) is the known cost to reach node n from the start, and h(n) is the estimated cost from node n to a goal. A\* is guaranteed to find a minimum-cost solution if one exists and examines the fewest possible nodes.

- It uses f' = g + h' to estimate the cost of getting from the initial state to the goal state, where g is a measure of the cost of getting from initial state to the current node and the function h' is an estimate of the cost of getting from the current node to the goal state. To find a path involving the fewest number of steps, we should set g = 1







https://www.youtube.com/watch?v=PzEWHH2v3TE&ab\_channel=Education4u

**AO\* Algorithm :-**

A\* Algorithm follows OR arcs. But AO\* Algorithm follows both OR arcs and AND arcs.

A\* algorithm contains both BFS and DFS feature.

To select best node for generating new successor nodes in AO\* Algorithm not only consider whether that is having less Heuristic Function value also calculate path value if that node is located in AND arc by propagating backword and add 1 unit cost to current f value i.w is shown above diagrams., then U can follow the above AO\* search procedure. nothing difficult.

**AND/OR graph** a graph or tree structure describing the decomposition of a goal in terms of alternative sub goals (OR nodes) or combinations of sub goals that must all be satisfied (AND nodes).

**Alpha and Beta cutoffs** in alpha/beta search, threshold values that allow un-promising avenues of play to be identified and eliminated from the search.

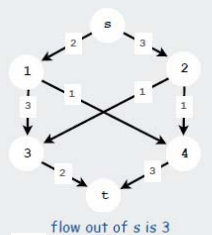
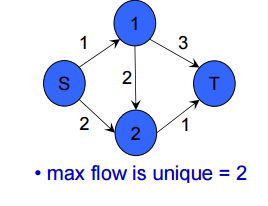
**Alpha/Beta** search a kind of game tree search that uses alpha and beta cutoffs to avoid unnecessary search of un-promising avenues of play.

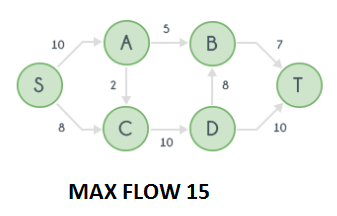
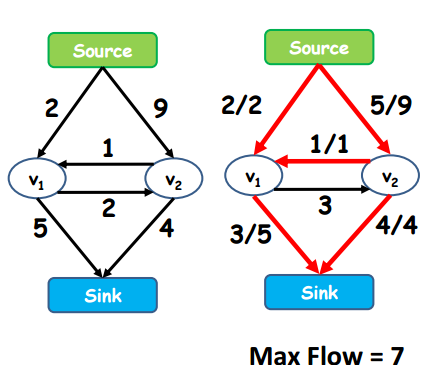
**Predicate calculus** a form of logical representation that includes: quantification of variables, terms (including functions of terms), and predicates whose arguments are terms.

**Greedy Best-First Search :-**

Greedy best-first search tries to expand the node that is closest to the goal, on the grounds that it is likely to lead to a solution quickly. Thus, it evaluates nodes by using just the heuristic function; that is, f(n) = h(n).

**MAX FLOW – MIN CUT :-**

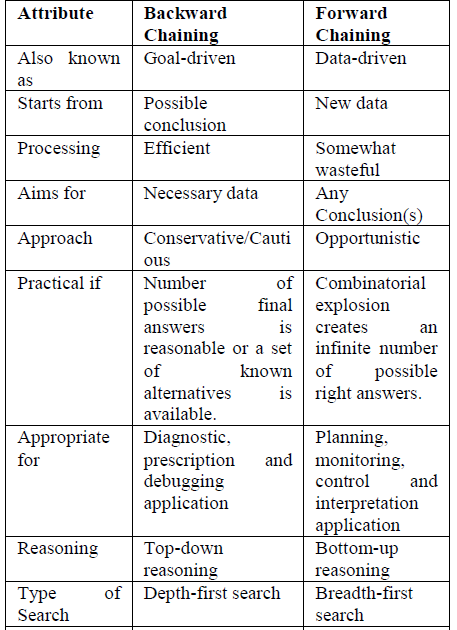




**Hill climbing :-**

In numerical analysis, hill climbing is a mathematical optimization technique which belongs to the family of local search. It is an iterative algorithm that starts with an arbitrary solution to a problem, then attempts to find a better solution by incrementally changing a single element of the solution. If the change produces a better solution, an incremental change is made to the new solution, repeating until no further improvements can be found.

**FORWARD CHAINING AND BACKWARD CHAINING :-**



**FUZZY SET AND FUZZY LOGIC**

**Universe of Discourse**

The Universe of Discourse is the range of all possible values for an input to a fuzzy system.

**Fuzzy Set**

A Fuzzy Set is any set that allows its members to have different grades of membership (membership function) in the interval [0,1].

**Support**

The Support of a fuzzy set F is the crisp set of all points in the Universe of Discourse U such that the membership function of F is non-zero.

**Crossover point**

The Crossover point of a fuzzy set is the element in U at which its membership function is 0.5.

**Membership Functions:-**

A membership function (MF) is a curve that defines how each point in the input space is mapped to a membership value (or degree of membership) between 0 and 1. The input space is sometimes referred to as the universe of discourse. The only condition a membership function must really satisfy is that it must vary between 0 and 1. A classical set might be expressed as A = {x | x > 6}

A fuzzy set is an extension of a classical set. If X is the universe of discourse and its elements are denoted by x, then a fuzzy set A in X is defined as a set of ordered pairs.

A = {x, μA(x) | x belongs to X}

μA(x) is called the membership function (or MF) of x in A. The membership function maps each element of X to a membership value between 0 and 1.

Example Gaussian membership function, sigmoidal membership function.

**alpha-cut:-**

An alpha-cut of the membership function A (denoted aA) is the set of all x such that A(x) is greater than or equal to alpha (a). Similarly, a strong alpha-cut (denoted a+A) is the set of all x such that A(x) is strictly greater than alpha (a).

**Fuzzy Singleton**

A Fuzzy singleton is a fuzzy set whose support is a single point in U with a membership function of one.

**Union :-**

The membership function of the Union of two fuzzy sets A and B is defined as the maximum of the two individual membership functions. This is called the maximum criterion.

C:\Users\antony\Desktop\union.gif

**Intersection:-**

The membership function of the Intersection of two fuzzy sets A and B with membership functions and respectively is defined as the minimum of the two individual membership functions. This is called the minimum criterion.

C:\Users\antony\Desktop\intersect.gif

**Complement:-**

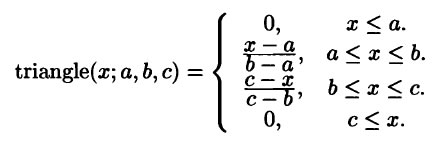
The membership function of the Complement of a Fuzzy set A with membership function is defined as the negation of the specified membership function. This is called the negation criterion.

C:\Users\antony\Desktop\complement.gif

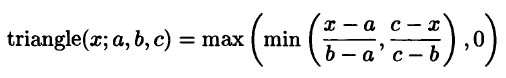
**Membership Functions**

**1 Triangular MFs**

A triangular MF is specified by three parameters {a, b, c} as follows:

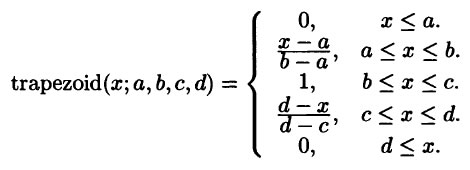


By using min and max, we have an alternative expression for the preceding equation:

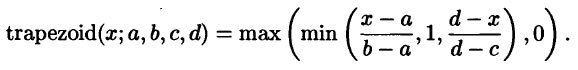


**2 Trapezoidal MFs**

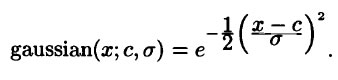
A trapezoidal MF is specified by four parameters {a, b, c, d} as follows:



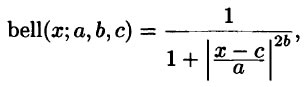
An alternative concise expression using min and max is:



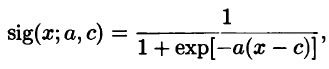
**3 Gaussian MFs**



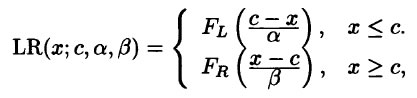
**4 BELL MFs**



**5 Sigmoid MFs**



**6 Left-right MF (L-R MF)**

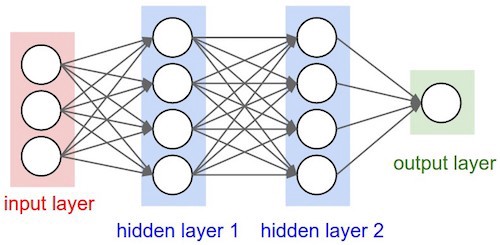


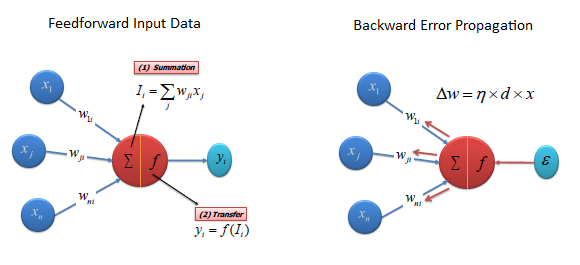
The term fuzzy logic was introduced with the 1965 proposal of fuzzy set theory by LotfiZadeh

Fuzzy logic has been used in numerous applications such as facial pattern recognition, air conditioners, washing machines, vacuum cleaners, antiskid braking systems, transmission systems, control of subway systems and unmanned helicopters, knowledge-based systems for multiobjective optimization of power systems, weather forecasting systems, models for new product pricing or project risk assessment, medical diagnosis and treatment plans, and stock trading. Fuzzy logic has been successfully used in numerous fields such as control systems engineering, image processing, power engineering, industrial automation, robotics, consumer electronics, and optimization. This branch of mathematics has instilled new life into scientific fields that have been dormant for a long time.

**NEURAL NETWORK**

**Artificial Neural Network**





OR PROBLEM:- Can be solved by single layer perceptron

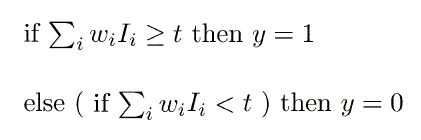
XOR PROBLEM:- Can be solved by radial basis functron

**The perceptron:-**

The most basic form of an activation function is a simple binary function that has only two possible results.

This function returns 1 if the input is positive or zero, and 0 for any negative input. A neuron whose activation function is a function like this is called a perceptron.

If any thresh hold value given, if the output greater than or equal to thresh hold value then output 1 else 0.(see example)



C:\Users\antony\Desktop\threshold.gif

The perceptron is a single layer feed-forward neural network with preprocessing. It is not an autoassociative network because it has no feedback and is not a multiple layer neural network because the preprocessing stage is not made of neurons.

Perceptron can learn AND,OR,NAND,NOR but not XOR.

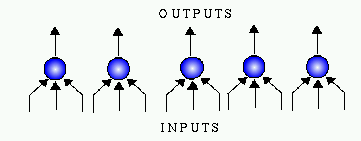
**An auto associative network** is equivalent to a neural network that contains feedback. The number of feedback paths(loops) does not have to be one

**Single-Layer Perceptron Neural Networks**

A single-layer perceptron network consists of one or more artificial neurons in parallel. The neurons may be of the same type we've seen in the Artificial Neuron Applet.

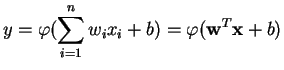
Each neuron in the layer provides one network output, and is usually connected to all of the external (or environmental) inputs.

The applet in this tutorial is an example of a single-neuron, single-layer perceptron network, with just two inputs.



**Multi-Layer Perceptron Neural Networks:-**

An MLP is a network of simple neurons called perceptrons. The perceptron computes a single output from multiple real-valued inputs by forming a linear combination according to its input weights and then possibly putting the output through some nonlinear activation function. Mathematically this can be written as



where w denotes the vector of weights, x is the vector of inputs, b is the bias and phi is the activation function

**Radial Basis Function Network (RBFN):-**

can be used for clustering of data.

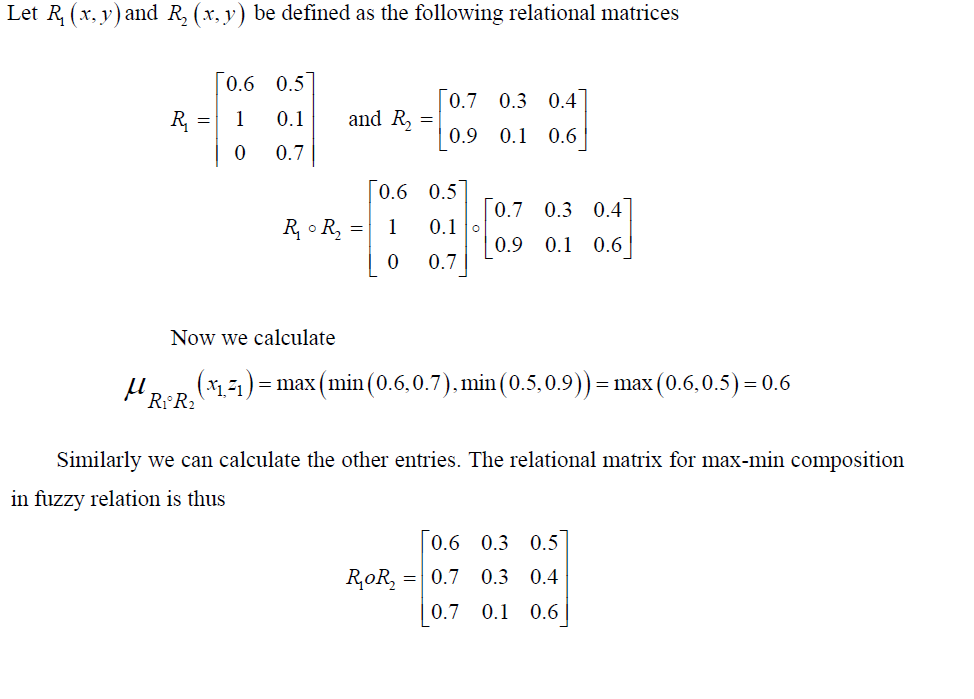
**Self Organization Maps:-**

x

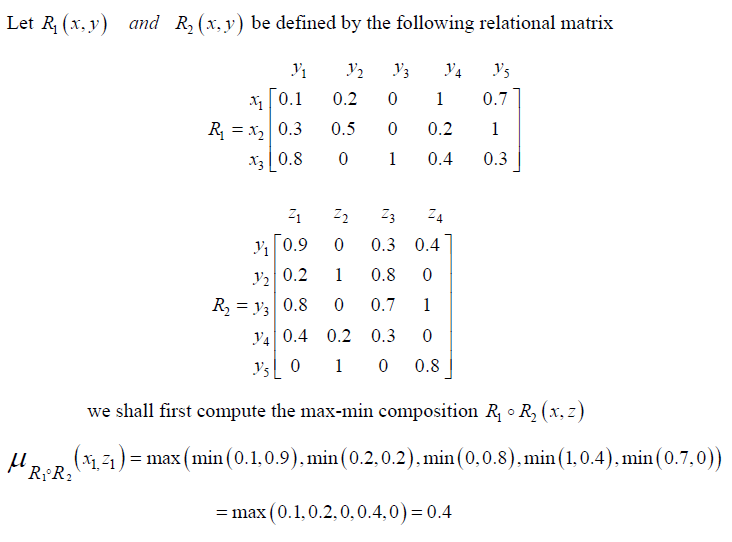
**DELTA RULE:-**

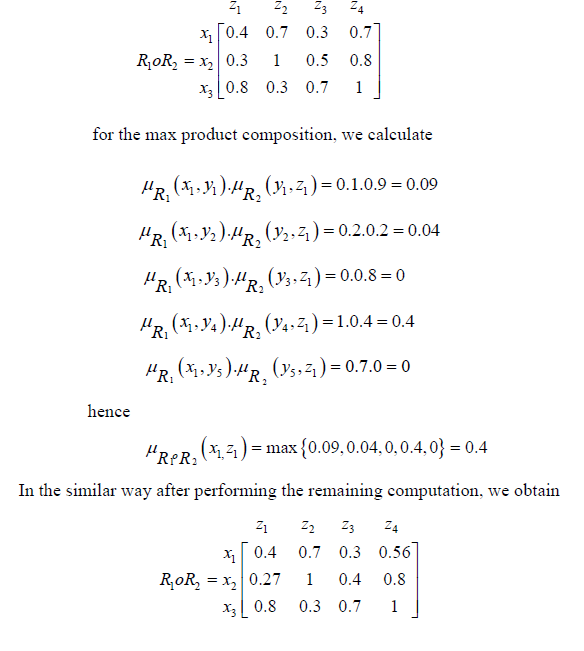
The Perceptron Learning Rule is an algorithm for adjusting the network weights w to minimize the difference between the actual and the desired outputs. the gradient of the function as the current point.

**Max – Min Composition :-**



**Max-Product Composition:-**





It is clear from the example that max-min composition and max product composition of crisp relations will yield the same result, but in fuzzy max-min composition and max product composition have different result.

**Neural Network Learning Rules:-**

We know that, during ANN learning, to change the input/output behavior, we need to adjust the weights. Hence, a method is required with the help of which the weights can be modified. These methods are called Learning rules, which are simply algorithms or equations.

### Perceptron Learning Rule

This rule is an error correcting the supervised learning algorithm of single layer feedforward networks with linear activation function, introduced by Rosenblatt.

### Delta Learning Rule (Widrow-Hoff Rule)

It is introduced by Bernard Widrow and Marcian Hoff, also called Least Mean Square (LMS) method, to minimize the error over all training patterns. It is kind of supervised learning algorithm with having continuous activation function.

**MACHINE LEARNING:-**

—**Inductive Learning:** This type of AI learning model is based on inferring a general rule from datasets of input-output pairs.. Algorithms such as knowledge based inductive learning(KBIL) are a great example of this type of AI learning technique. KBIL focused on finding inductive hypotheses on a dataset with the help of background information.

—**Deductive Learning:** This type of AI learning technique starts with the series of rules and infers new rules that are more efficient in the context of a specific AI algorithm. Explanation-Based Learning(EBL) and Relevance-Based Learning(RBL) are examples of deductive techniques. EBL extracts general rules from examples by “generalizing” the explanation. RBL focuses on identifying attributes and deductive generalizations from simple example.

—**Unsupervised Learning:** Unsupervised models focus on learning a pattern in the input data without any external feedback. Clustering is a classic example of unsupervised learning models. Hopfield network is an unsupervised learning.

**—Supervised Learning:** Supervised learning models use external feedback to learning functions that map inputs to output observations. In those models the external environment acts as a “teacher” of the AI algorithms. Multilayer perceptron and support vector machine networks are example of the supervised learning.

—**Semi-supervised Learning:** Semi-Supervised learning uses a set of curated, labeled data and tries to infer new labels/attributes on new data data sets. Semi-Supervised learning models are a solid middle ground between supervised and unsupervised models.

—**Reinforcement Learning:** Reinforcement learning models use opposite dynamics such as rewards and punishment to “reinforce” different types of knowledge

The main differences between Supervised and Unsupervised learning are given below:

|  |  |
| --- | --- |
| **Supervised Learning** | **Unsupervised Learning** |
| Supervised learning algorithms are trained using labeled data. | Unsupervised learning algorithms are trained using unlabeled data. |
| Supervised learning model takes direct feedback to check if it is predicting correct output or not. | Unsupervised learning model does not take any feedback. |
| Supervised learning model predicts the output. | Unsupervised learning model finds the hidden patterns in data. |
| In supervised learning, input data is provided to the model along with the output. | In unsupervised learning, only input data is provided to the model. |
| The goal of supervised learning is to train the model so that it can predict the output when it is given new data. | The goal of unsupervised learning is to find the hidden patterns and useful insights from the unknown dataset. |
| Supervised learning needs supervision to train the model. | Unsupervised learning does not need any supervision to train the model. |
| Supervised learning can be categorized in **Classification** and **Regression** problems. | Unsupervised Learning can be classified in **Clustering** and **Associations** problems. |
| Supervised learning can be used for those cases where we know the input as well as corresponding outputs. | Unsupervised learning can be used for those cases where we have only input data and no corresponding output data. |
| Supervised learning model produces an accurate result. | Unsupervised learning model may give less accurate result as compared to supervised learning. |
| Supervised learning is not close to true Artificial intelligence as in this, we first train the model for each data, and then only it can predict the correct output. | Unsupervised learning is more close to the true Artificial Intelligence as it learns similarly as a child learns daily routine things by his experiences. |
| It includes various algorithms such as Linear Regression, Logistic Regression, Support Vector Machine, Multi-class Classification, Decision tree, Bayesian Logic, etc. | It includes various algorithms such as Clustering, KNN, and Apriori algorithm. |

**Quest**:- Where is the minimum criterion used ?

**Answer:-** A. When there is an AND operation .

**Explanation:** The minimum criterion takes the least value in the set given. In an AND operation all the values have to be true in order for the whole expression to be true. Hence taking the least of the values ensures that the range of all values are met.

**Ques:-** What are the following sequence of steps taken in designing a fuzzy logic machine?

**Ans:-**Fuzzification->Rule evaluation->Defuzzification

Explanation: When designing a fuzzy logic, we first have to define the fuzzy sets, and make appropriate member functions. Then rule evaluation comes in which matches the sets to its corresponding rules(a series of if-then statements).

**EXERCISES1:-**

A neuron with 4 inputs has the weight vector w = [1 2 3 4]t .The activation function is linear, that is, the activation function is given by f(net) = 2 \* net. If the input vector is X = [5 6 7 8] , then find the output of the neuron?

**ANS** :-Net =1\*5+2\*6+3\*7+4\*8=70

F(net)=2\*net=140

**Questions:-**

A 4-input neuron has weights 1, 2, 3 and 4. The transfer function is linear with the constant of proportionality being equal to 2. The inputs are 4, 10, 5 and 20 respectively. The output will be:

**Explanation:-**

The output is found by multiplying the weights with their respective inputs, summing the results and multiplying with the transfer function. Therefore:

Output = 2 \* (1\*4 + 2\*10 + 3\*5 + 4\*20) = 238

**QUES:-**What is the Fuzzy Appoximation Theorem(FAT)?

**Answer:** A fuzzy sytem can model any continous sytem.

Explanation: The Fuzzy Approximation Theorem(FAT) as stated by Bart Kosko shows a fuzzy sytem can model any contionous system.

Each of the rules acts as a fuzzy patch that the system places so as to resemble the response of the continous system.

**QUES:-**What is the main difference between probability and fuzzy logic

**Answer:-** Probability is ADDITIVE, meaning all its values must add up to one.

**Explanation:**-This is the main difference between fuzzy logic and probability. Although, both probability and fuzzy logic contain values between the ranges 1 and 0, fuzzy logic tells the extent of a specific member function, whereas probability gives the frequency, hence all values of its set must add up to one.

***QUES:- What is the best method to go for the game playing problem?***

(A) Optimal Search (B) Random Search (C) Heuristic Search(yes) (D) Stratified Search

**Solution:** A heuristic function, also called simply a heuristic, is a function that ranks alternatives in search algorithms at each branching step based on available information to decide which branch to follow. For example, it may approximate the exact solution.

***A neuron with inputs has the weight vector [0.2 −0.1 0.1]T and a bias θ=0. If the input vector is X=[ 0.2 0.4 0.2 ]T. Then the total number of input to the neuron is:***

A) 0.20 B)1.0 C)0.02(yes) D)-1.0

f(x)=WT∗X+θ

f(x) = 0.2\*0.2 - 0.1\*0.4 + 0.1\*0.2 = 0.02

***Which of the following can be used for clustering the data?***

A)Single layer perception B)Multilayer perception

C)Self organizing map(yes) D)Radial basis function

**Explanation:-**Self-organizing maps are a data visualization technique which reduce the dimensions of data through the use of self-organizing neural networks. It is a type of unsupervised learning. The goal is to discover some underlying structure of the data

***The strategy used to reduce the number of tree branches and the number of static evaluations applied in case of a game tree is***

(A) Minmax strategy (B) Alpha-beta pruning strategy (yes)

(C) Constraint satisfaction strategy (D) Static max strategy

***Given U = {1, 2, 3, 4, 5, 6, 7} A = {(3, 0.7), (5, 1), (6, 0.8)} then ~A will be : (where ~ ® complement)***

(A) {(4, 0.7), (2, 1), (1, 0.8)} (B) {(4, 0.3), (5, 0), (6, 0.2) }

(C) {(1, 1), (2, 1), (3, 0.3), (4, 1),(6, 0.2), (7, 1)}(yes) (D) {(3, 0.3), (6.0.2)}

**Explanation:-**

The first is (1,1). The first 1 is in U but not in A, so it should be added in the complement. The second 1 is because the membership function is 1- µA(x). 1-0=1.

***Consider a fuzzy set old is defined as below Old = {(20, 0.1), (30, 0.2), (40, 0.4), (50, 0.6), (60, 0.8), (70, 1), (80, 1)} Then the alpha-cut for alpha = 0.4 for the set old will be***

(A) {(40, 0.4)} (B) {50, 60, 70, 80} (C) {(20, 0.1), (30, 0.2)}

(D) {(20, 0), (30, 0), (40, 1),(50, 1), (60, 1), (70, 1), (80, 1)}

**Explanation:-**alpha cut means it is the cut off value . the number having member ship value more than or equal to in the given fuzzy set will be considered as present in the result . now in the set OLD person with age 40,50,60,70,80 are having membership values more than or equal to 0.4 so they are considered as member with value 1 and 20 30 with value less than cut off are not member hence they are shown with value 0 so

ans is D ( P.S. it is one form of showing the result another may be {40,50,60,70,80} but it is not in choice )

***Given U = {1, 2, 3, 4, 5, 6, 7} A = {(3, 0.7), (5, 1), (6, 0.8)} then ~A will be : (where ~ ® complement)***

(A) {(4, 0.7), (2, 1), (1, 0.8)} (B) {(4, 0.3), (5, 0), (6, 0.2) }

(C) {(1, 1), (2, 1), (3, 0.3), (4, 1),(6, 0.2), (7, 1)} (D) {(3, 0.3), (6.0.2)}

Explanation:-

Membership of 3=.7

Membership of 5=1.

Membership of 6=.8

Membership of 3’=.3

Membership of 5’=0,

Membership of 6’=.2.

A’ is universe of discourse set not in Set A.

{(1,1),(2,2),(3,.3),(4,1),(6.2),(7,1)}.

***Consider a fuzzy set A defined on the interval X=[0, 10] of integers by the membership function μA(x) = x / (x+2) Then the α cut corresponding to α = 0.5 will be***

a) {0,1,2,3,4,5,6,7,8,9,10}

b) {1,2,3,4,5,6,7,8,9,10}

c) {2,3,4,5,6,7,8,9,10}

d) {}

**Explanation:-**

putting value of X from 0 to 10 in membership function = x / ( x + 2 ) we get

0/2 , 1/3 , 2/4 ,3/5 ,4/6 ,5/7 , 6 /8 , 7/9 , 8/10 , 9/11 , 10/12 means they are belonging with degree 0, 0.33 , 0.5, 0.6 .........0.83 here alpha cut = 0.5 so first two elements will not be included in the result as their degree of belongingness < 0.5

so answer will be C