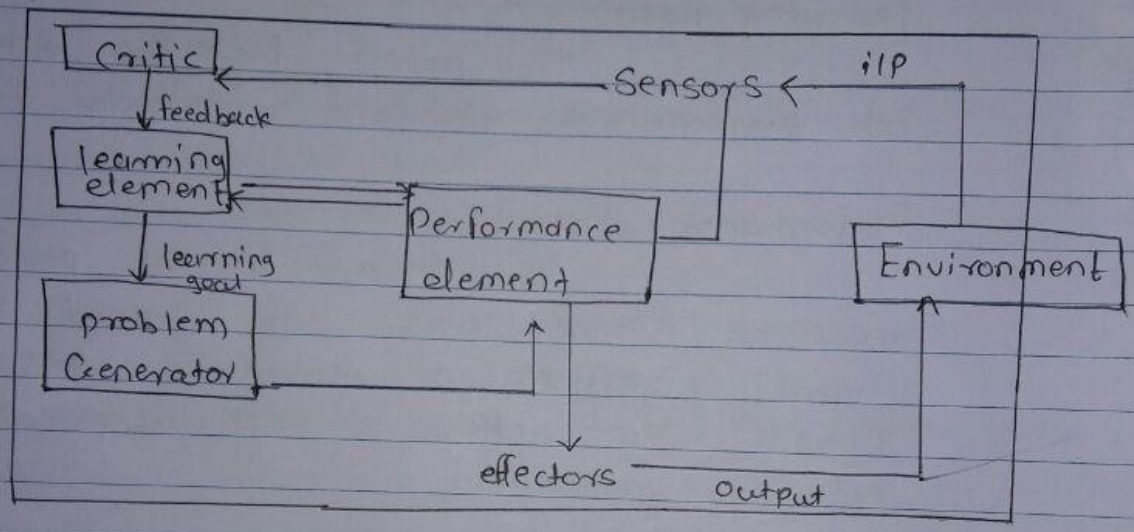
SSDA

Q1 Explain architecture of general learning Agent

performance Agent



- Why we give mock test ? when we get less marks for some questions you come to know that you have made some mistake in our answer . Then we learn the correct answer & when we get that same question is further examinations , you write the correct ans. & avoid the mistake which were made in the mock test.

- learning agent is advantages , with its basic knowledge it can initially operate in an known env. then it can gain knowledge from the env. based on few Parameta. & perform action to give better result

Critic - It is one who compares sensors input with performance standard.

learning Agent -

To learn from the difference bet^n performance standard & feedback from critic

Performance element -

Based on current percept received from sensor & obtained by learning agent, performance element is responsible to choose the action to act upon the external environment.

b) Explain any two local search algorithm.

→ Local search algorithm start with local Solution & proceed ~~proce~~ to iterate the remaining search space.

-Important factor in local search algorithm are,

① Selected initial solution
② Stop criteria

① Hill climbing Search

- Hill climbing search algorithm is a special type of local search algo. which regularly update search tree by refering to goal state.

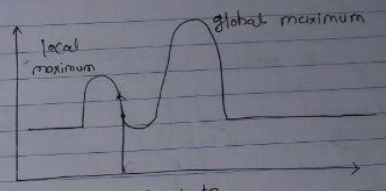


Fig. Hill climbing

c) Explain procedure for FOL to CNF with example

Q2

a) Explain any three functionalities of inteligent agent ?

→ The basic abilities of an intelligent agent are to self guared, responsive, goal oriented.

- Capabilities of an intelligent goal oriented

① Ability to remain atonous
② Responsive
③ Goal oriented.

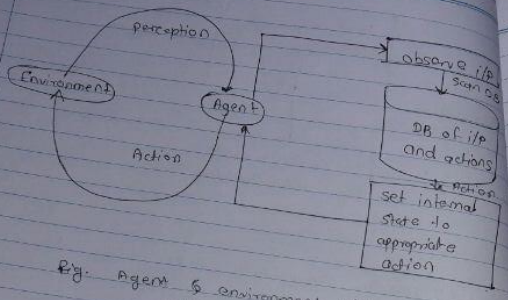


Fig: Agent & environment with agent state.

① Reactivity :-

The reaction to situation in stipulated time frame. An agent can receive the environment & responsed to the situation in particulary time frame.

② Pro-activeness -

It is controlling a situation rather than just responding to it. Itelligent agent show goal oriented behaviour by taking initiative.

② Social Ability -

Intelligent agent can interact with other agents take our driver example, where agent might have to interact with other agent or a human being while driving cars

b) Describe effectiveness of alpha-beta purning.

→ - Alpha-Purning is search algorithm that decreases the no. of node while search proress & evaluation.
- The min-max algorithm serves as base of alpha beta purning. The min-max algorithm, minimax decision is computed from current state
- Let have a look on generalized pseudocode for minimax algo before diving into alpha-beta purning

DCOER
Page No.:

Q.2 b] Describe effectiveness of a alpha-beta pruning.

1) Alpha-beta pruning is a search algorithm that decreases the number of nodes while search process of evaluation.
2) The minimax algorithm serves as the base of Alpha-Beta Pruning algorithm.
3) In Minimax algorithm, minimax decision is computed from current state.
4) Lets have a look on a generalized Pseudocode for minimax algorithm before dividing into alpha-beta pruning.
Following is the generalized algorithm for minimax approach.

```
function minimax (node, depth, maximizingPlayer).
if depth = 0    or node is a terminal node.
        return the heuristic value of node.
  if maximizingPlayer.
  bestValue := -∞
  for each child of node.
            v := minimax (child, depth-1, FALSE)
        bestValue := max (bestValue, v)
        return bestValue.
else (* minimizing player *)
                bestValue := +∞
                for each child of node.
v := minimax (child, depth-1, TRUE)
                bestValue := min (bestValue, v)
                return bestValue.
```

DCOER
Page No.:

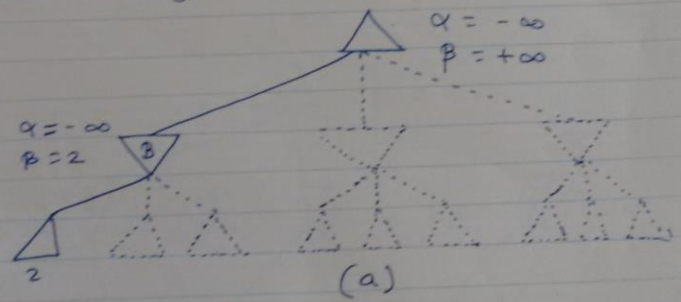
5) Alpha-beta pruning is a kind of adversarial search problem. This algorithm works quite efficiently for gaming problem like chess, tic-tac-toe etc.

6) The Alpha & Beta can be specified as follows:

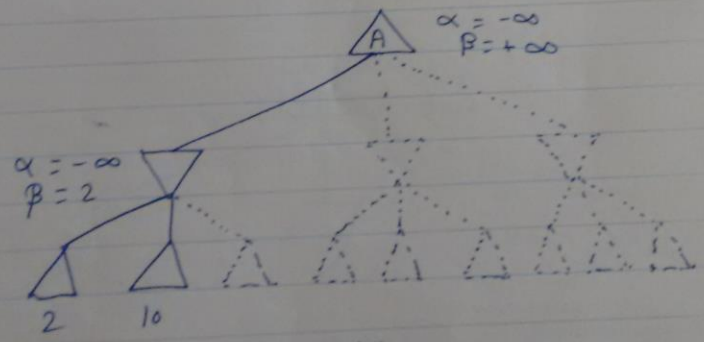
Alpha: Best option already explored along path to the root for maximizer.

Beta: Beta option already explored along path to the root for minimizer.

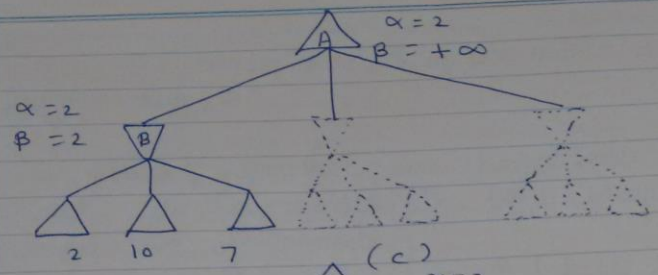
7) Lets take an example of calculating an optimal decision using alpha-beta pruning.



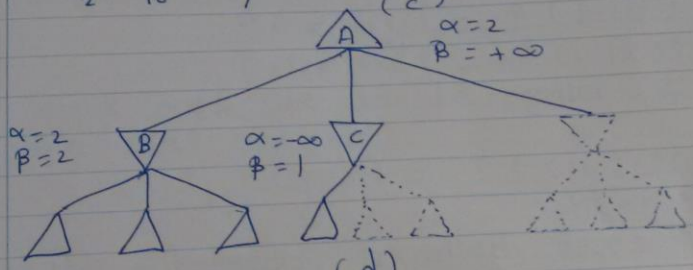
(a)



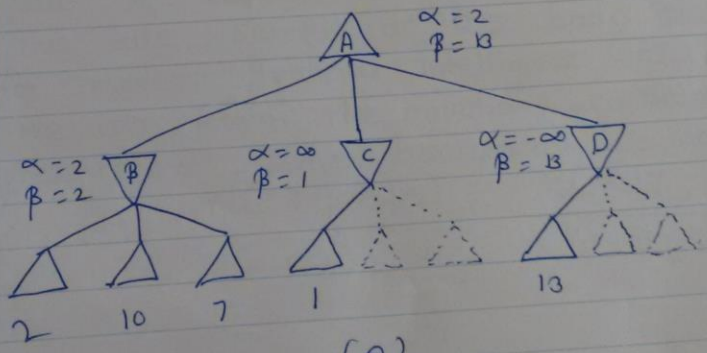
(b)


A
α = 2
β = +∞
α = 2
β = 2
B
2
10
7

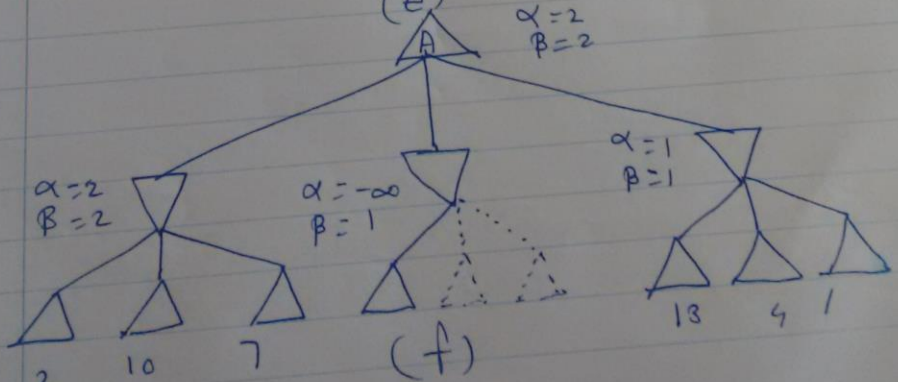
(c)


A
α = 2
β = +∞
α = 2
β = 2
B
α = -∞
β = 1
C

(d)


A
α = 2
β = 13
α = 2
β = 2
B
α = ∞
β = 1
C
α = -∞
β = 13
D
2
10
7
1
13

(e)


A
α = 2
β = 2
α = 2
β = 2
α = -∞
β = 1
α = 1
β = 1
2
10
7
13
4
1

(f)

DCOER
Page No.

Q.3 a] Explain the bay's rule & its use with a suitable example.

→ 1) Baye's Rule =

1) In previous sections we have seen product rule that can be written in two ways due to the commutative of conjunction.

$$P(x \wedge y) = P(x|y)\,P(y).$$
$$P(x \wedge y) = P(y|x)\,P(x).$$

2) Right hand side can be equated & can be devided by $P(x)$ to get.

$$P(y|x) = \frac{P(x|y)\,P(y)}{P(x)}$$

3) This ~~rat~~ equation is called Bayes Rule or Bayes law or Baye's theorem.

4) This rule serves as the basic probabilities inference for modern Artificial intelligence (AI) systems.

5) We can reform the equation conditionalised in presence of some background evidence e.

$$P(Y|X,e) = \frac{P(X|Y,e)\,P(Y|e)}{P(X|e)}.$$

DCOER

Page No.

Q. 4a] Write a note on Hidden Markov Models.

1) Hidden markov model or HMM is used to describe the state of a process by a single discrete random variable.

2) It is a temporal probability model where possible states of the world is represented by possible values of the variable.

3) It is possible however to add more variable in HMM framework but at the end they all state variable are combined into a single variable called "mega variable".

4) Many speech recognition application uses the HMM.

5) Let X_t is the state variable having value represented by integers $1, \ldots, N$, where N is the number of possible states.
So, the transition model $(X_t \mid X_{t-1})$ will become on N x N matrix T where,

6) $T_{ij} = P(X_t = j \mid X_{t-1} = i)$ and T_{ij} is nothing but the probability of transition of state 'i' to state 'j'.

7) Let concider discussed in previous section where the evidence is Noise & state is Traffic. So the transition matrix for "Noise & Traffic" would will be,

$$T = P(X_t \mid X_{t-1}) = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}$$

8) We can form the matrix with sensor model also, the model will specify the probability of evidence e_t to appear. Diagonal matrix O_t can be formed for each step t, &

and diagonal values will given by $P(e_t | X_{t=i})$ & other entries will be 0.

g) So, for example if in the first snapshot of world (Day - 1) Noise is true (N_1 = true) we can specify diagonal matrix

$$O_1 = \begin{pmatrix} 0.9 & 0 \\ 0 & 0.2 \end{pmatrix}.$$

9.4 b] Explain the construction of Dynamic Bayesian Networks with a suitable example.

→ 1) Another kind temporal probability model is dynamic Bayesian Network.

2) The noise & traffic network example & 1 & 2 is an example of dynamic bayesian network.

3) Many kalman filter networks also fall under the category of dynamic bayesian network.

4) Each & every snapshot of a Dynamic Bayesian Network (DBN) contains any number of evidence variable E_t & state variable X_t.

5) For simplicity, it is assumed that the relation or link between the variables are exactly same in each snapshot & can be represented by markov chain.

6) Markov process model put restriction on DBN such that each variable should have parent only in its own snapshot.

7) Hidden markov model can also be represented as DBN. In this case it will have single state variable & single evidence variable.

Also, every discrete variable DBN can be specified as Hidden Markov Model (HMM).

8) So we can conclude that each HMM is a DBMN & each DBN can be represented into a HMM.

9) Every kalman filter with continuous variable & Gaussian distribution can be represented in DBN but not each DBN can be translated into kalman filter.

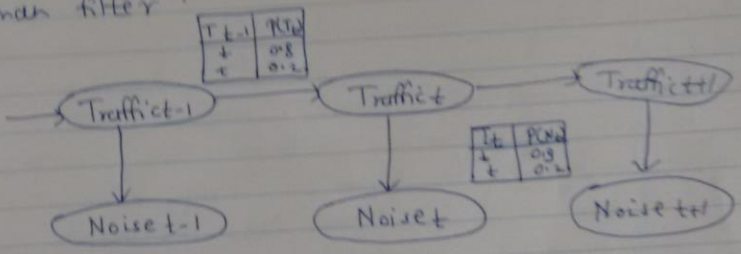


Fig.1] Bayesian network for sensor model.

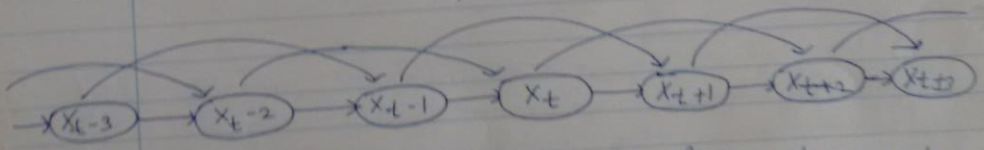


Fig.2] Bayesian network for first order markov process.

DCOER
Page No:

Q.5a) Explain any one supervised learning approach.

1) The main aim of supervised learing is learning a general rule that maps the input to output.
2) In supervised learning suppose 'teacher' gives some input, then computer is presented with example input & also their desired outputs.
3) For example :
Given training set of positive negative examples of a concept. The description is constructed which will accurately classify whether the future examples are positive or negative. That is, learn some good estimate of function of given a training set.
$\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$.
4) Where each y_i is distribution of either + or - or a probability distributed over +/-.

Q.5b) Explain Nonparametric models.
1) Non parametric statistics is mainly a branch of statistics concerned with non parametric statistical models & non parametric statistical tests.
2) Non parametric inferential statistical methods are mathematical procedures for statistical hypothesis testing.
3) Non parametric models are different from parametric models in that model structure is not specified previously but is instead

determined from data.

4) The term non-parametric doesn't imply that such different models completely lack the parameter but number & nature of parameters are flexible & not fixed in advanced.

5) A histogram calculates non-parametric estimate of a probability distribution.

6) Kernel density estimation provides better estimates of density.

7) Non parametric regression & semiparametric regression methods are developed based on splines, kernels & wavelets.

8) KNN mainly classify the unseen instance based on k points in training set which are nearest to it.

9) A support vector machine is mainly a non parametric large-margin classifier.

6a) Write a note Artificial Neural Networks.

→ 1) Artificial neural networks are most powerful learning model. They can have wide range of complex functions which represents multidimensional input-output maps.

2) ANN is generally presented by a set of nodes, arranged in layers & set of weighted directed links connecting them. The nodes are equivalent to neurons while the link denotes synapses. The nodes can be represented as information processing units

and links act as communication media.

3) For e.g :- A neural n/w for handwriting recognition is defined by a set of input neurons which may be activated by pixels or an input image.

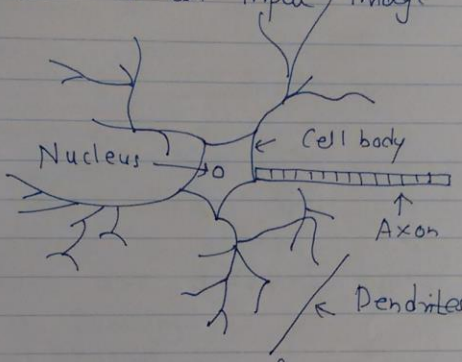


Fig.] Artifitial Neural network (ANN)

Q.6b) Explain Ensemble Learning.

→1) The ensemble learning method is used to select the collection of different hypothesis from hypothesis space & also combine their predictions.

2) For e.g :- At the time of cross validation we might generate twenty different decision trees & then vote on best classification for new example.

3) The motivation of ensemble learning is simple.

4) Cdr an ensemble of K=5 hypothesis & suppose that we combine their predictions using simple majority voting for the ensemble to misclassify a new example, at least three

of five hypothesis have to misclassify it.

5) The hope is that this is much less likely than misclassification by single hypothesis.

6) In statistics & machine learning ensemble methods are multiple learning algorithms to obtain better predictive performance that could be obtained from any of learning algorithm.

7) A machine learning ensemble mainly refers to concrete finite set; where as statistical ensemble is infinite.

8) There are different types of ensemble methods
   1) Boosting.
   2) Bagging.
   3) Bays Optimal classifier.
   4) Boysian model average.

9) The most widely used ensemble method is Boosting.

Q.7 a) What are the information retrival characteristics How to evaluate & refine information retrieval system?

→ Following are the characteristics of information retrieval =

1) A corpus of documents = Each system must decide what it wants to treat as a ~~dot~~ document: a paragraph, a page or a multipage text.

DCOER
Page No.:

2) Queries posed in a query language =
A query specifies what the user wants to know.

3) A result set =
This is the subset of documents that the IR system judges to be relevant to the query.

4) A presentation of the result =
This can be as simple as a ranked list of document titles or as complex as a rotating color map of the result set projected onto a three dimentional space, rendered as a two-dimentional display.

* IR system evaluation =
1) How do we know whether an IR system is performing well?
2) Traditionally, there have been two measures used in the scoring: recall & precision.
3) Precision measures the proportion of documents in the result set that are actually relevant.
4) Recall measures the proportion of all the relevant documents in the collection that are in the result set.
5) All we can do is either estimate recall by sampling or ignore recall completely & just judge precision.
6) In the case of a Web search engine, there may be thousands of documents in the result set, so it makes more sense to measure

precision for several different sizes, such as "p@10" or "p@50," rather than to estimate precision in the entire result set.

- IR refinements =

1) There are many possible refinments to the system described here & indeed Web search engines are continually updating their algorithms as they discover new approaches & as the Web grows & changes.
2) One common refinement is a better model of the effect of document length on relevance.
3) The BM25 scoring function uses a word model that treats all words as completely independent, but we know that some words are correlated: "couch" is closely related to both "couches" & "sofa". Many IR systems attempt to account for these correlations.
4) For example, if the query is [couch.], it would be a shame to exclude from the result set.
5) In german, for e.g., it is not uncommon to see words like "Lebensversicherungsgesellschaftsangestellter".
6) Languages such as Finnish, Turkish, Inuit & Yupik have recursive morphological rules that in principle generate words of unbounded length.

DCOER
Page No.:

4) ~~Language~~ Explain the procedure for Machine translation.

The procedure for machine Translation is as given below =

1) Find parallel texts =
First, gather a parallel bilingual corpus. For e.g., a Hansard HANSARD is a record of parliamentary debate. Canada, Hong Kong, & other counties produce bilingual Hansards, the European Union publishes its official documents in 11 languages & the United Nations publishes multilingual documents.

2) Segments into sentences =
The unit of translation is a sentence, so we will have to break the corpus into sentences. One way to decide if a period ends a sentence is to train a model that takes as features the surrounding words & their parts of speech. This approach achieves about 98% accuracy.

3) Align sentences =
For each sentence in the English version, determine what sentence(s) it corresponds to in the French version. Usually, the next sentence of English corresponds to the next sentence of French in a 1:1 match, but sometimes there is variation the order of two sentences will be swapped, resulting in a 2:2 match.

4) Align phrases =
Within a sentence, phrases can be aligned by a process that is similar to that used for sentence alignment, but requiring iterative improvement. When we start, we have no way of knowing that "qui dort" aligns with "sleeping", but we can arrive at that alignment by a process of aggregation of evidence.

5) Extract distortions =
Once we have an alignment of phrases we can define distortion probabilities.

6) Improve estimates with EM=
Use expectation-maximization to improve the estimates of P(f|e) & P(d) values. We compute the best alignments with the current values of these parameters in the E step, then update the estimates in the M step & iterate the process until convergence.

8a] Define Robotic Perception in brief.
y

Robotic Perception=

1) The most important task that has to be accomplished by an autonomous robot of any kind is to acquire the knowledge about its environment & its where abouts.

2) The system can use various sensors to take the measurements & then extract important information about the environment from that measurement to achieve the knowledge about its environments.

3) The we can say that the process in which the robot maps the sensor measurements into internal representations of the environment is called as perception.

4) The problem with perception is that it is quite difficult since the measurement source i.e. the sensors itself are noisy, & the environment is not totally observable, & the parameters of the environment keeps changing & is unpredictable.

5) The representation must have following 3 properties:-
1) Information contained there in must be sufficient enough that the robot is able to make good decisions.
2) The representations are structured so that efficient updating is possible.
3) The internal state variables must correspond to natural state variables in the real world.

Q.9 a] Basis of Utility Theory =

→ 1) There are certain constraints that are used to derive the Maximum Expected Utility (MEU) principle.

2) These constraints are applied on preferences that rational agent considers while making decision.

3) Following notations are used to express agent's preference =

i) $X > Y$ X is preferred to Y

ii) $X \sim Y$ The agent is unbiased for both X & Y.

iii) $X \geq Y$ Agent prefers X to Y or unbiased between them.

where X & Y are world's state observable to agent.

4) Utility functions are of 2 types =
   1) Cardinal Utility Functions.
   2) Ordinal Utility Functions.

① Cardinal Utility Functions =

→ In cardinal utility, the magnitude of utility is considered while deciding preferences.

② Ordinal Utility =

→ In ordinal utility, the amount of utility or the magnitude of utility has no meaning, while making preferences.

9b] How to Evaluate & Choose the Best Hypothesis:

→ 1) We want to learn a hypothesis that fits the future data best. To make that easy & precise we need to define "Future data" & "best". We make stationary assumption;

1) That there is a probability distribution over examples that remain stationary over time.

2) While evaluating & choosing best hypothesis there points should be considered:

(1) Each observed value is sampled from that distribution of probability & is independent of previous examples.

(2) Examples have identical prior probability distribution.

(3) Examples that satisfy the assumptions are called as independent & identically distributed (i.i.d).

(4) The error rate of hypothesis h is the proportion of mistakes it makes.

(5) Just because a hypothesis h has low error rate on training set, does not mean that it will generalize well.

2) The simplest approach is to randomly split the available data into a training set & from which learning algorithm produces h & a test set on which accuracy of h is evaluated. This method is called as holdout cross-validation.

3) The disadvantage of holdout cross-validation is fails to use all the variable data. If we use half data for test set, then we are only training on half the data & because of that we may get a poor hypothesis.

Q.10a) How to represent & evaluate decision problem with a decision network.

→ Decision Network =

1) Decision network is a directed graph that specifies agent's preference considering uncertainty in world & possible decisions.

2) Decision Networks are also known as influence diagram & it represents three important things :=
   i) Current state of agent's world.
   ii) Set of possible actions.
   iii) Set of possible outcomes.

3) A decision network consists of =
   i) Circle or chance nodes = It represent uncertain variable.
   ii) Rectangle or decision nodes = It represents decisions & attached with chance nodes.
   iii) Diamond shapes or utility nodes = It represents the utility of agent that tells the usefulness of state with respect to decision.

4) An example of simple decision n/w is shown below :=

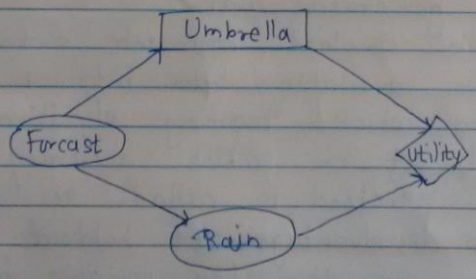


Fig.1 Decision Network.

Q.10b) Explain any four prime application domains of robotics technology.

→ 1) Agriculture
2) Social & personal services.
3) Industries.
4) Transportation.

1) Agriculture =

Robots can assist farmers to relieve them from their laborious task like plowing, seeding, spraying pesticides, watering, maintaining crops & harvesting. The robots can also be deployed for transporting the materials & packaging.

2) Social & personal services =

1) Robots can be used to help humans in their daily chores like domestic work that includes cleaning, washing, cooking & other similar activities, help children in their studies or take them to school or any other places safely.
2) Work in hotels, restaurants & shops etc. Robotic information kiosks can be set up in malls, exhibition, trade fairs etc., in museums or places of tourist attraction as tour guides.
3) Enquiry counters like railway, bus or taxi stations & airports. They can be used as security systems for various homes & offices.

3) Industries =

1) Robots find infinite applications in the industries. Work that requires high dexterity, difficult human labor, or areas that are hazardous for humans to work, robots can be deployed for accomplishing such tasks.
2) One of the most popular applications of robots in industry is the assembly line, where the robots perform the tasks such as material handling, transport, part placements, welding, grinding, milling & painting repeatedly without breaks.
3) Robots can perform these tasks efficiently & more accurately than humans with less errors & prove much cost effective than human labour.

4) Transportation =

1) Robotic cars = Cars are the most convenient mode of transport for most of us. Due to carelessness, negligence, distraction of any kinds or lack of skills cause lot of road accidents & more than millions of people lose their lives. Robotic cars can reduce the risks to quite a great extent.
2) An automatic or driverless car or self-driving cars are the other names for robotic cars. The Mercedes Benz's & EUREKA's Prometheus,

ZEAL COER
Page No.:
www.zealeducation.com

DARPA Grand Challenge Boss and STANLEY are some exeumples of robotic cars.

3) The advantages of these cars are =

1) Fewer traffic collisions.
2) Increased roadway capacity.
3) Higher speed limits.
4) Safe.
5) Smoother ride.
6) Relief from driving & navigation.
7) Reduction in redundant passenger, etc.