V.S.B COLLEGE OF ENGINEERING TECHNICAL CAMPUS

DEPARTMENT OF CSE

QUESTION BANK

PART A

UNIT I

Artificial Intelligence

1. **Define Artificial Intelligence (AI).**

   The study of how to make computers do things at which at the moment, people are better.
   Systems that think like humans
   Systems that act like humans
   Systems that think rationally
   Systems that act rationally.

2. **Define Artificial Intelligence formulated by Haugeland.**

   The exciting new effort to make computers think machines with minds in the full and literal sense.

3. **Define Artificial Intelligence in terms of human performance.**

   The art of creating machines that performs functions that require intelligence when performed by
   people.

4. **Define Artificial Intelligence in terms of rational acting.**

   A field of study that seeks to explain and emulate intelligent behaviors in terms of computational processes-Schalkoff. The branch of computer science that is concerned with the automation of intelligent behavior-Luger & Stubblefield.

5. **Define Artificial in terms of rational thinking.**

   The study of mental faculties through the use of computational models-Charniak& McDermott.The study of the computations that make it possible to perceive, reason and act-Winston.

6. **What is meant by Turing test?**

   To conduct this test we need two people and one machine. One person will be an interrogator(i.e.) questioner, will be asking questions to one person and one machine.

Three of them will bein a separate room.Interrogator knows them just as A and B. so it has to identify which is the person and machine.The goal of the machine is to make Interrogator believe that it is the person's answer. If machine succeeds by fooling Interrogator, the machine acts like a human. Programming a computer to pass Turing test is very difficult.

7. **What is called materialism?**

An alternative to dualism is materialism, which holds that the entire world operate according to physical law. Mental process and consciousness are therefore part of physical world, but inherently unknowable they are beyond rational understanding.

8. **What are the capabilities, computer should posses to pass Turing test?**

Natural Language Processing
Knowledge representation

Automated Reasoning

Machine Learning.

9. **Define Total Turing Test?**

The test which includes a video signals so that the interrogator can test the perceptual abilities of the machine.

10. **What are the capabilities computers needs to pass total Turing test?**

Computer Vision
Robotics

11. **Define Rational Agent.**

It is one that acts, so as to achieve the best outcome (or) when there is uncertainty, the best expected outcome.

12. **Define Agent.**

An Agent is anything that can be viewed as perceiving (i.e.) understanding its environment
through sensors and acting upon that environment through actuators.

13. **Define an Omniscient agent.**

An omniscient agent knows the actual outcome of its action and can act accordingly; but omniscience is impossible in reality.

14. **What are the factors that a rational agent should depend on at any given time?**

1. The performance measure that defines degree of success.
2. Ever thing that the agent has perceived so far. We will call this complete perceptual history the percept sequence.
3. When the agent knows about the environment.
4. The action that the agent can perform.

15. **Define Architecture.**

The action program will run on some sort of computing device which is called as Architecture

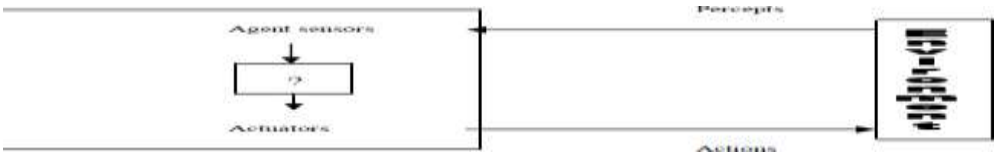
16. List the various type of agent program.

Simple reflex agent program.
Agent that keep track of the world.
Goal based agent program.
Utility based agent program

17. Give the structure of agent in an environment?

Agent interacts with environment through sensors and actuators.

An Agent is anything that can be viewed as perceiving (i.e.) understanding its environment through sensors and acting upon that environment through actuators.



18. Define Percept Sequence.

An agent's choice of action at any given instant can depend on the entire percept sequence observed to elate.

19. Define Agent Function.

It is a mathematical description which deals with the agent's behavior that maps the given percept sequence into an action.

20. Define Agent Program.

Agent function for an agent will be implemented by agent program.

21. How agent should act?

Agent should act as a rational agent. Rational agent is one that does the right thing, (i.e.) right
actions will cause the agent to be most successful in the environment.

22. How to measure the performance of an agent?

Performance measure of an agent is got by analyzing two tasks. They are How and When actions.

23. Define performance measures.

Performance measure embodies the criterion for success of an agent's behavior.

24. Define Ideal Rational Agent.

For each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given the evidence provided by the percept sequence and whatever built in knowledge the agent has.

25. Define Omniscience.

An Omniscience agent knows the actual outcome of its actions and can act accordingly.

26. Define Information Gathering.

Doing actions in order to modify future percepts sometimes called information gathering.

27. What is autonomy?

A rational agent should be autonomous. It should learn what it can do to compensate for partial (or) in correct prior knowledge.

28. What is important for task environment?

PEAS → P- Performance measure

E - Environment
A- Actuators
S – Sensors
Example
Interactive English tutor performance measure maximize student's score on test.
Environment
Set of students testing Agency
Actuators
Display exercises suggestions, corrections.
Sensors
Keyboard entry

29. What is environment program?

It defines the relationship between agents and environments.

30. List the properties of environments.

o Fully Observable Vs Partially Observable
o Deterministic Vs Stochastic
o Episodic Vs Sequential
o Static Vs Dynamic
o Discrete Vs Continuous
o Single Agent Vs Multi agent
a. Competitive Multi agent
b.Co – operative Multi agent

31. What is Environment Class (EC) and Environment Generator (EG)?

EC – It is defined as a group of environment.
EG – It selects the environment from environment class in which the agent has to Run.

32. What is the structure of intelligent Agent?

Intelligent Agent = Architecture + Agent Program

33. Define problem solving agent.

Problem solving agent is one kind of goal based agent, where the agent
Should select one action from sequence of actions which lead to desirable states.

34. List the steps involved in simple problem solving technique.

i. Goal formulation
ii. Problem formulation
iii. Search
iv. Solution
v. Execution phase

35. What are the different types of problem?

Single state problem, multiple state problems, Contingency problem, Exploration problem

36. What are the components of a problem?

There are four components. They are
i. initial state
ii. Successor function
iii. Goal test
iv. Path cost
v. Operator
vi. state space

vii. path

37. Define State Space.

The set of all possible states reachable from the initial state by any sequence of action is called state space.

38. Define Path.

A path in the state space is a sequence of state connected by sequence of actions.

39. Define Path Cost.

A function that assigns a numeric cost to each path, which is the sum of the cost of the each action along the path.

40. Give example problems for Artificial Intelligence.

i. Toy problems
ii. Real world problems

41. Give example for real world end toy problems.

Real world problem examples:
i. Airline travel problem.
ii. Touring problem.
iii. Traveling salesman problem.
iv. VLSI Layout problem
v. Robot navigation
vi. Automatic Assembly
vii. Internet searching
Toy problem Examples:
Vacuum world problem.
8 – Queen problem
8 – Puzzle problem

42. Define search tree.

The tree which is constructed for the search process over the state space is called search tree.

43. Define search node.

The root of the search tree that is the initial state of the problem is called search node.

44. Define fringe.

The collection of nodes that have been generated but not yet expanded, this collection is called fringe or frontier.

45. List the performance measures of search strategies.

i. Completeness
ii. Optimality
iii. Time complexity
iv. Space complexity

46. Define branching factor (b).

The number of nodes which is connected to each of the node in search tree is called Branching factor.

47. Differentiate Blind Search and Heuristic Search.

i) No information about the path cost from the current state to goal state.
ii) Problem is solved with the information we which we know.
iii) Example
a) Breadth first search
b) Uniform cost search
c) Depth first Search
d) Depth limited search
e) Iterative deepening search
f) Bi – Directional Search
Backtracking search.

i) We have some information like minimum path caused to move
ii) Problem can be solved by the information which is already given.
iii) Example
a) Best first search
b) Greedy search
c) A* search

48. Define Uniform cost search.

Uniform cost search expands the node 'n' with the lowest path cost instead of expanding the Shallowest node.

49. Define Depth first search.

It expands the deepest node in the current fringe of the search tree.

50. Define depth limited search.

The problem of unbounded tress can be avoided by supplying depth limit 1(i.e.) nodes at depth 1 are treated as if they have no successors. This is called Depth Limited search.

PART – B

1. Describe briefly the various problem characteristics?

Heuristic search is a very general method applicable to a large class of problem. It includes a variety of techniques. In order to choose an appropriate method, it is necessary to analyze the problem with respect to the following considerations.

A very large and composite problem can be easily solved if it can be broken into smaller problems and recursion could be used. Suppose we want to solve.

Ex:- ∫ x2 + 3x+sin2x cos 2x dx

This can be done by breaking it into three smaller problems and solving each by applying specific rules. Adding the results the complete solution is obtained.

1. Ignorable problems Ex:- theorem proving
· In which solution steps can be ignored.
2. Recoverable problems Ex:- 8 puzzle
· In which solution steps can be undone
3. Irrecoverable problems Ex:- Chess
· In which solution steps can't be undone good solution absolute or relative

There are two categories of problems. In one, like the water jug and 8 puzzle problems, we are satisfied with the solution, unmindful of the solution path taken, whereas in the other category not just any solution is acceptable.

Problem Classification

Actual problems are examined from the point of view , the task here is examine an input and decide which of a set of known classes.

Ex:- Problems such as medical diagnosis , engineering design.

2. Explain in detail about Uninformed Search and Informed Search Strategies.

Uninformed Search Strategies have no additional information about states beyond that provided in the problem definition.

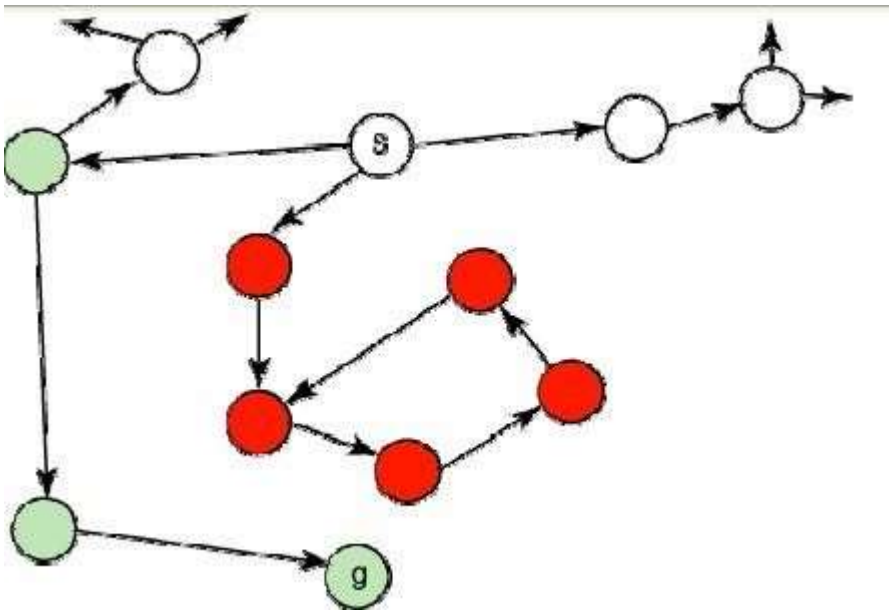
Strategies that know whether one non goal state is ―more promising than another are called informed search or heuristic search strategies.
There are five uninformed search str
ategies as given below.
o Breadth-first search
o Uniform-cost search
o Depth-first search
o Depth-limited search
o Iterative deepening search

3. Explain in detail about Depth first search Algorithm?

The *h* function can be extended to be applicable to (non-empty) paths. The heuristic value of a path is the heuristic value of the node at the end of the path. That is:

h(⟨no,...,nk⟩)=h(nk)

A simple use of a heuristic function is to order the neighbors that are added to the stack representing the frontier in depth-first search. The neighbors can be added to the frontier so that the



best neighbor is selected first. This is known as heuristic depth first search. This search chooses the locally best path, but it explores all paths from the selected path before it selects another path. Although it is often used, it suffers from the problems of depth-fist search.
Another way to use a heuristic function is to always select a path on the frontier with the lowest heuristic value. This is called best-first search. It usually does not work very well; it can follow paths that look promising because they are close to the goal, but the costs of the paths may keep increasing.

4. Discuss in detail about Constraints Satisfaction Problem.

Constraint satisfaction is the process of finding a solution to a set of constraints that impose conditions that the variables must satisfy. A solution is therefore a set of values for the variables that satisfies all constraints—that is, a point in the feasible region.
The techniques used in constraint satisfaction depend on the kind of constraints being considered. Often used are constraints on a finite domain, to the point that constraint satisfaction

problems are typically identified with problems based on constraints on a finite domain. Such problems are usually solved via search, in particular a form of backtracking or local search.
Constraint propagation are other methods used on such problems; most of them are incomplete in general, that is, they
may solve the problem or prove it unsatisfiable, but not always.Constraint propagation methods are also used in conjunction with search to make a given problem simpler to solve. Other considered kinds of constraints are on real or rational numbers; solving problems on these constraints is done via variable elimination or the simplex algorithm.

Complexity

Solving a constraint satisfaction problem on a finite domain is an NP complete problem with respect to the domain size. Research has shown a number of tractable subcases, some limiting the allowed constraint relations, some requiring the scopes of constraints to form a tree, possibly in a reformulated version of the problem. Research has also established relationship of the constraint satisfaction problem with problems in other areas such as finite model theory.

5. Explain in details about Production System?

Types of Production Systems.

A Knowledge representation formalism consists of collections of condition-action rules(Production Rules or Operators), a database which is modified in accordance with the rules, and a Production System Interpreter which controls the operation of the rules i.e
The 'control mechanism' of a Production System, determining the order in which Production Rules are fired. A system that uses this form of knowledge representation is called a production system.
A production system consists of rules and factors. Knowledge is encoded in a declarative from which comprises of a set of rules of the form
Situation ---------- Action
SITUATION that implies ACTION.
Example:-
IF the initial state is a goal state THEN quit.
The major components of an AI production system are
i. A global database
ii. A set of production rules and
iii. A control system
The goal database is the central data structure used by an AI production system. The production system. The production rules operate on the global database. Each rule has a precondition that is either satisfied or not by the database. If the precondition is satisfied, the rule can be applied.
Application of the rule changes the database. The control system chooses which applicable rule should be applied and ceases computation when a termination condition on the database is satisfied. If several rules are to fire at the same time, the control system resolves the conflicts.
Four classes of production systems:-
1. A monotonic production system
2. A non monotonic production system
3. A partially commutative production system
4. A commutative production system.

EXTRAS (FOR PRACTICE)

1. Define Heuristic function.
2. Name the elements of an agent and list down the characteristics of intelligent agent.

3.How would you rank Production system?
4.How would you quote PEAS description?
5.Apply problem solving algorithm to measure performance.
6.Explain the theme of Backtracking search for CSP.
7.Illustrate the categories of production systems.
8.List the types of constraints
9. Point out some of the uninformed search techniques.
10. How would you formulate Constraint Satisfaction Problem?11.Express what is ridge?
12.What do you infer from hill-climbing search algorithm?
13.Generalize your opinion about admissible heuristic.
14.Define problem solving agents and list its algorithms.
15.Why problem formulation must follow goal Formulation?
16.Summarize the factors that make up rationality.
17. What do you infer from the word Agent?
18. How much knowledge would be required by a perfect program for the problem of playing chess? Assume that unlimited computing power is available.
19. Will you state or interpret in your own words PEAS description for a Vacuum cleaner?
20.Show what would happen if problem is decomposed.

UNIT II KNOWLEDGE REPRESENTATION AND REASONING

1. What is informed search?

One that uses problem – specific knowledge beyond the definition of the problem itself and it can find solutions more efficiently than an uninformed strategy.

2. What is the use of QUEUING_FN?

QUEUING_FN inserts asset of elements into the queue. Different varieties of queuing fn produce different varieties of the search algorithm.

3. Mention the criteria for the evaluation of search strategy.

There are 4 criteria: Completeness, time complexity, space complexity, optimality

4. List the various search strategies.

a. BFS
b. Uniform cost search
c. DFS
d. Depth limited search
e. Iterative deepening search
f. Bidirectional search

5. List the various informed search strategy.

Best first search –greedy search ,A* search
Memory bounded search-Iterative deepening A*search -simplified memory bounded A*search -Iterative improvement search –hill climbing -simulated annealing

6. What is Best First Search?

Best First Search is an instance of the general TREE SEARCH or GRAPH SEARCH algorithm in which a node is selected for expansion based on an evaluation function, f(n).

7. Define Evaluation function, f(n).

A node with the lowest evaluation is selected for expansion, because evaluation measures distance to the goal.

8. Define Heuristic function, h (n).

h (n) is defined as the estimated cost of the cheapest path from node n to a goal node.

9. Define Greedy Best First Search.

It expands the node that is closest to the goal (i.e.) to reach solution in a quicker way. It is done by using the heuristic function: f(n) = h(n).

10. Define A* search.

A* search evaluates nodes by combining g(n), the cost to reach the node and h(n), the cost to get from the node to the goal.

f(n) = g(n) + h(n)

11. Define Admissible heuristic h (n).

In A* search, if it is optimal then, h(n) is an admissible heuristic which means h(n) never overestimates the cost to reach the goal.

12. What is triangle inequality?

It states that each side of a triangle cannot be longer than the sum of the other two slides of the triangle.

13. What are the 2 types of memory bounded heuristic algorithms?

i. Recursive Best First Search(RBFS)
ii. Memory bounded A*(MA*)

14. Differentiate BFS & DFS.

BFS means breath wise search. Space complexity is more. Do not give optimal solution Queuing fn is same as that of queue operator

DFS means depth wise search. Space complexity is less Gives optimal solution Queuing fn is somewhat different from queue operator.

15. What is RBFS?

It keeps track of the f-value of the best alternative path available from any ancestor of the current node. RBFS remembers the f-value of the best leaf in the forgotten sub tree and therefore decide whether its worth re expanding the sub tree sometimes later.

16. Define iterative deepening search.

Iterative deepening is a strategy that sidesteps the issue of choosing the best depth limit by trying all possible depth limits: first depth 0, then depth 1,then depth 2& so on.

17. What are the 2 ways to use all available memory?

i. Memory bounded A*(MA*)

ii. Simplified Memory bounded A*(SMA*)

18. What is SMA* search?

SMA* expands the best leaf until memory is full and it drops the oldest worst leaf node and expands the newest best leaf node.

19. What is called as bidirectional search?

The idea behind bidirectional search is to simultaneously search both forward from the initial state & backward from the goal & stop when the two searches meet in the middle.

20. What is metalevel state space?

Each state in a metalevel state space captures the internal state of a program that is searching in an object level state space.

21. What is Manhattan distance, h2?

The sum of the horizontal and vertical distances of the tiles from their goal positions in a 15 puzzle problem is called Manhattan distance (or) city block distance.

22. Give the drawback of DFS.

The drawback of DFS is that it can get stuck going down the wrong path. Many problems have very deep or even infinite search tree. So dfs will never be able to recover from an unlucky choice at one of the nodes near the top of the tree. So DFS should be avoided for search trees with large or infinite maximum depths

23. Define Branching factor b*.

Uniform tree of depth d would have to be in order to contain N+1 nodes is called branching factor.

24. Write the time & space complexity associated with depth limited search.

Time complexity =O (bd) ,

b-branching factor,

d-depth of tree

Space complexity=o (bl)

25. What is Released problems?

A problem with fewer restrictions on the actions is called a relaxed problem.

26. What is a pattern database?

This database is the storage of exact solution costs for every possible sub problem instance.

27. What is a disjoint pattern database?

The sum of the two costs is still a lower bound on the cost of solving the entire problem is

called a disjoint pattern database.

28. What is local search?

It operates using a single current state rather than multiple paths and generally moves only to neighbors of that state.

29. Define Optimization Problems.

The aim of this problem is to find the best state according to an objective function.

30. What are the 2 parts of Landscape?

i. Location defined by the state.
ii. Elevation defined by the value of the heuristic cost function (or) objective function.

31. Define Global minimum.

If elevation corresponds to cost, then the aim is to find the lowest valley is called global minimum.

32. Define Global Maximum.

If elevation corresponds to an objective function, then the aim is to find the highest peak is called global maximum.

33. Define Hill Climbing search.

It is a loop that continually moves in a increasing value direction (i.e.) up hill and terminates when it reaches a "peak" where no neighbor has a higher value.

34. List some drawbacks of hill climbing process.

Local maxima: A local maxima as opposed to a goal maximum is a peak that is lower that the highest peak in the state space. Once a local maxima is reached the algorithm will halt even though the solution may be far from satisfactory.

Plateaux: A plateaux is an area of the state space where the evaluation fn is essentially flat. The search will conduct a random walk.

35. What is the meaning for greedy local search?

It goals (picks) a good neighbor state without thinking ahead about where to go next.

36. Define Local maxima.

A local maximum is a peak that is higher than each of its neighboring states, but lower than the global maximum.

37. What are the variants of hill climbing?

i. Stochastic hill climbing
ii. First choice hill climbing
iii. Simulated annealing search
iv. Local beam search
v. Stochastic beam search

38. Define annealing.

Annealing is the process used to harden metals (or) glass by heating them to a high temperature and then gradually cooling them, thus allowing the material to coalesce into a low energy crystalline state.

39. Define simulated annealing.

This algorithm, instead of picking the best move, it picks a random move. If the move improves the situation, it is always accepted.

40. What is the advantage of memory bounded search techniques?

We can reduce space requirements of A* with memory bounded algorithm such as IDA* & SMA*.

41. Give the procedure of IDA* search.

Minimize f(n)=g(n)+h(n) combines the advantage of uniform cost search + greedy search A* is Complete optimal. Its space complexity is still prohibitive.

Iterative improvement algorithms keep only a single state in memory, but can get stuck on local maxima. In this algorithm each iteration is a dfs just as in regular iterative deepening. The depth first search is modified to use an f-cost limit rather than a depth limit. Thus each iteration expands all nodes inside the contour for the current f-cost.

42. List some properties of SMA* search.

* It will utilize whatever memory is made available to it.
* It avoids repeated states as for as its memory allow.
* It is complete if the available memory is sufficient to store the shallowest path.
* It is optimal if enough memory is available to store the shallowest optimal solution path. Otherwise it returns the best solution that can be reached with the available memory.
*When enough memory is available for entire search tree, the search is optimally efficient.
*Hill climbing.
*Simulated annealing.

43. What is Genetic Algorithms?

Genetic Algorithm is a variant of stochastic beam search in which successor states are generated by combining two parent states, rather than by modifying a single state.

44. Define Online Search agent.

Agent operates by interleaving computation and action (i.e.) first it takes an action, and then it observes the environment and computes the next action.

45. What are the things that agent knows in online search problems?

a. Actions(s)
b. Step cost function C(s, a, s')
c. Goal TEST(s)

46. Define CSP.

Constraint Satisfaction problem (CSP) is defined by a set of variables X1,X2,…Xn and set of constraints C1,C2,…Cm.

47. Define Successor function.

A value can be assigned to any unassigned variable, provided that does not conflict with previously assigned variables.

48. What are the types of constraints?

There are 5 types,
a. Unary constraints relates one variable.
b. A binary constraint relates two variables.
c. Higher order constraints relate more than two variables.
d. Absolute constraints.
e. Preference constraints.

49. Define MRV.

Minimum remaining values heuristic chooses the variable with the fewest "legal" values.

50. Define LCV.

Least constraining value heuristic prefers the value that rules out the fewest choices for the Neighboring variables in the constraint graph.

51. Define Conflict directed back jumping.

A back jumping algorithm that uses conflict sets defined in this way is called Conflict

directed back jumping.

52. Define constraint propagation.

It is the general term for propagating (i.e.) spreading the implications of constraints on the variable on to other variable.

53. Define Cycle cut set.

The process of choosing a subset S from variables [CSP] such that the constraint graph becomes a tree after removal of S. S is called a cycle cut set.

54. Define Tree decomposition.

The constraint graph is divided into a set of connected sub problems. Each sub problem is solved independently and the resulting solutions are then combined. This process is called tree decomposition.

55. Define Alpha beta pruning.

Alpha beta pruning eliminates away branches that cannot possibly influence the final decision

PART – B

1. Describe the Issues in knowledge representation?

Typically, a problem to solve or a task to carry out, as well as what constitutes a solution, is only given informally, such as "deliver parcels promptly when they arrive" or "fix whatever is wrong with the electrical system of the house.

To solve a problem, the designer of a system must flesh out the task and determine what constitutes a solution; represent the problem in a language with which a computer can reason; use the computer to compute an output, which is an answer presented to a user or a sequence of actions to be carried out in the environment; and interpret the output as a solution to the problem. Knowledge is the information about a domain that can be used to solve problems in that domain. To solve many problems requires much knowledge, and this knowledge must be represented in the computer. As part of designing a program to solve problems, we must define how the knowledge will be represented. A representation scheme is the form of the knowledge that is used in an agent. A representation of some piece of knowledge is the internal representation of the knowledge. A representation scheme specifies the form of the knowledge.

A knowledge base is the representation of all of the knowledge that is stored by an agent.

2. Explain in details about first-order logic?

Whereas propositional logic assumes the world contains facts,

• first-order logic (like natural language) assumes the world contains

• Objects: people, houses, numbers, colors, baseball games, wars, …

• Relations: red, round, prime, brother of, bigger than, part of, comes between, …

Atomic sentences

Atomic sentence = *predicate* (*term1*,...,*termn*) or *term1 = term2*

Term = *function* (*term1*,...,*termn*) or *constant* or *variable*

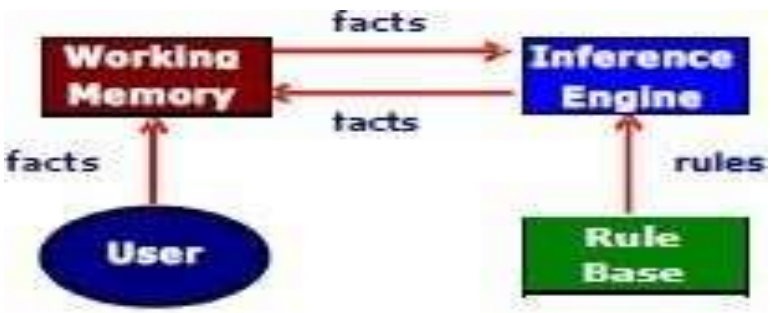
• E.g., *Brother(TaoiseachJohn,RichardTheLionheart) >*
(Length(LeftLegOf(Richard)),
Length(LeftLegOf(TaoiseachJohn)))

Syntax of FOL: Basic elements

- Constants TaoiseachJohn, 2, DIT,...
- Predicates Brother, >,...
- Functions Sqrt, LeftLegOf,...
- Variables x, y, a, b,...
- Connectives ¬, ⇒, ∧, ∨, ⇔
- Equality =
- Quantifiers ∀, ∃

3. Illustrate in detail about forward and backward chaining with suitable example.

Forward Chaining: The Forward chaining system, properties , algorithms, and conflict resolution strategy are illustrated.



‡ facts are held in a working memory

‡ condition-action rules represent actions to be taken when specified facts occur in working memory.

‡ typically, actions involve adding or deleting facts from the working memory.

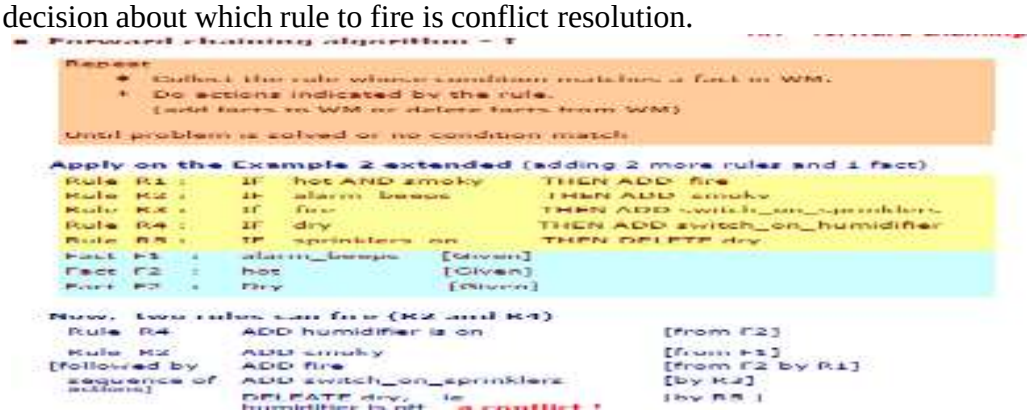
■ Properties of Forward Chaining

‡ all rules which can fire do fire.

‡ can be inefficient - lead to spurious rules firing, unfocused problem solving

‡ set of rules that can fire known as conflict set.

‡ decision about which rule to fire is conflict resolution.



- **Backward Chaining**

Backward chaining system and the algorithm are illustrated.

- **Backward chaining system**
  - Backward chaining means reasoning from goals back to facts. The idea is to focus on the sea h.
  - Rules and facts are processed using backward chaining interpreter.
  - Checks hypothesis, e.g. *"should I switch the sprinklers on?"*
- **Backward chaining algorithm**
  - Prove goal G
    If G is in the initial facts , it is proven.
    Otherwise, find a rule which can be used to conclude G, and try to prove each of that rule's conditions.

Encoding of rules

Rule R1	:	IF hot AND smoky	THEN fire
Rule R2	:	IF alarm_beeps	THEN smoky
Rule R3	:	If fire	THEN switch_on_sprinklers
Fact F1	:	hot	[Given]
Fact F2	:	alarm beeps	[Given]
Goal	:	Should I switch sprinklers on?	

4. Knowledge engineering in FOL

1. Identify the task
2. Assemble the relevant knowledge
3. Decide on a vocabulary of predicates, functions, and constants
4. Encode general knowledge about the domain
5. Encode a description of the specific problem instance
6. Pose queries to the inference procedure and get answers
7. Debug the knowledge base

5. Explain in details about Predicate Calculus?

• An interpretation over D is an assignment of the entities of D to each of the constant, variable, predicate and function symbols of a predicate calculus expression such that:

• 1: Each constant is assigned an element of D

• 2: Each variable is assigned a non-empty subset of D;(these are the allowable substitutions for that variable)

• 3: Each predicate of arity n is defined on n arguments from D and defines a mapping from Dn into {T,F}

• 4: Each function of arity n is defined on n arguments from D and defines a mapping from Dn into D

Truth Value of Predicate Calculus expressions

• Assume an expression E and an interpretation I for E over a non empty domain D. The truth value for E is determined by:

• The value of a constant is the element of D assigned to by I

• The value of a variable is the set of elements assigned to it by I

Similarity with Propositional logic truth values

• The value of the negation of a sentence is F if the value of the sentence is T and F otherwise

UNIT III
UNCERTAINTY

1. Define FOL.

FOL is a first order logic. It is a representational language of knowledge which is powerful than propositional logic (i.e.) Boolean Logic. It is an expressive, declarative, compositional language.

2. Define a knowledge Base:

Knowledge base is the central component of knowledge base agent and it is described as a set of representations of facts about the world.

3. With an example, show objects, properties functions and relations.

Example

"EVIL KING JOHN BROTHER OF RICHARD RULED ENGLAND IN 1200"

Objects : John, Richard, England, 1200

Relation : Ruled

Properties : Evil, King

Functions : BROTHER OF

4. Define a Sentence?

Each individual representation of facts is called a sentence. The sentences are expressed in a language called as knowledge representation language.

5. Define an inference procedure

An inference procedure reports whether or not a sentence is entiled by knowledge base provided a knowledge base and a sentence .An inference procedure 'i' can be described by the sentences that it can derive. If i can derive from knowledge base, we can write. KB --Alpha is derived from KB or i derives alpha from KB.

6. Define Ontological commitment.

The difference between propositional and first order logic is in the ontological commitment. It assumes about the nature of reality.

7. Define Epistemological commitment.

The logic that allows the possible states of knowledge with respect to each fact.

8. Define domain and domain elements.

The set of objects is called domain, sometimes these objects are referred as domain elements.

9. What are the three levels in describing knowledge based agent?

Logical level

Implementation level

Knowledge level or epistemological level

10. Define Syntax?

Syntax is the arrangement of words. Syntax of a knowledge describes the possible configurations that can constitute sentences. Syntax of the language describes how to make sentences.

11. Define Semantics

The semantics of the language defines the truth of each sentence with respect to each possible world. With this semantics, when a particular configuration exists with in an agent, the agent believes the corresponding sentence.

12. Define Logic

Logic is one which consist of i. A formal system for describing states of affairs, consisting of a) Syntax b)Semantics.
ii. Proof Theory – a set of rules for deducing the entailment of a set sentences.

13. What are the 3 types of symbol which is used to indicate objects, relations and functions?
i) Constant symbols for objects
ii) Predicate symbols for relations
iii) Function symbols for functions

14. Define terms.
A term is a logical expression that refers to an object. We use 3 symbols to build a term.

15. Define Atomic sentence.
Atomic sentence is formed by both objects and relations.
Example
Brother (William, Richard)
i. Logical Constants (True, False)

16 .Define Modus Ponen's rule in Propositional logic?
The standard patterns of inference that can be applied to derive chains of conclusions that lead to the desired goal is said to be Modus Ponen's rule.

17 .Define AND –Elimination rule in propositional logic
AND elimination rule states that from a given conjunction it is possible to inference any of the conjuncts. OR-Introduction rule states that from, a sentence, we can infer its disjunction with anything.

18. Define Unification.
Lifted Inference rule require finding substitutions that make different logical expressions look identical (same). This is called Unification.

19. Define Occur check.
When matching a variable in 2 expressions against a complex term, one must check whether the variable itself occurs inside the term, If it does the match fails. This is called occur check.

20. Define pattern matching.
The inner loop of an algorithm involves finding all the possible unifiers with facts in the KB. This is called pattern matching.

21. Explain the function of Rete Algorithm?
This algorithm preprocess the set of rules in KB to constant a sort of data flow network in which each node is a literals from rule a premise.

22. Define magic set.
To rewrite the rule set, using information from the goal, so that only relevant variable bindings called magic set.

23. Define backward chaining.
This algorithm works backward from the goal, chaining through rules to find known facts that support the proof.

24. Define Prolog program.
It is a set of definite clauses written in a notation somewhat different from standard FOL.

25. What are the divisions of knowledge in OTTER theorem?
i. Set of Support (SOS)
ii. Usable axioms

iii. Rewrites (or) Demodulators
iv. A set of parameters and sentences

26. What are the 2 types of frame problem?

i. Representational Frame Problem
ii. Inferential Frame Problem

27. What are the 2 types of processes?

i. Discrete events – it have definite structure
ii. Liquid events - Categories of events with process.

28. Define fluent calculus.

Discard Situation Calculus and invent a new formalism for writing axioms is Called Fluent Calculus.

29. What is important for agent?

Time (i.e.) intervals is important for agent to take an action.
There are 2 kinds;
i. Moments
ii. Extended Intervals

30. Define runtime variables.

Plans to gather and use information are represented using short hand
Notation called runtime variables (n).
Example
[Look up (Agent, "Phone number (Divya)".N), Dial (n)]

PART – B

1. Describe about types of Knowledge representation?

There are four types of Knowledge representation :
Relational, Inheritable, Inferential, and Declarative/Procedural.

◊ **Relational Knowledge :**

provides a framework to compare two objects based on equivalent attributes.
any instance in which two different objects are compared is a relational type of knowledge.

◊ **Inheritable Knowledge**

− is obtained from associated objects.
− it prescribes a structure in which new objects are created which may inherit all or a subset of attributes from existing objects.

◊ **Inferential Knowledge**

− is inferred from objects through relations among objects.
− e.g., a word alone is a simple syntax, but with the help of other
words in phrase the reader may infer more from a word; this inference within linguistic is called semantics.

◊ **Declarative Knowledge**

− a statement in which knowledge is specified, but the use to which that knowledge is to be put is not given.
− e.g. laws, people's name; these are facts which can stand alone, not
dependent on other knowledge;

Procedural Knowledge

− a representation in which the control information, to use the knowledge, is embedded in the knowledge itself.

− e.g. computer programs, directions, and recipes; these indicate specific use or implementation.
− a representation in which the control information, to use the knowledge, is embedded in the knowledge itself.
− e.g. computer programs, directions, and recipes; these indicate specific use or implementation.

2. Discuss in details about inferential knowledge?
This knowledge generates new information from the given information.
This new information does not require further data gathering form soue, but does require analysis of the given information to generate new knowledge.
− given a set of relations and values, one may infer other values or relations.
− a predicate logic (a mathematical deduction) is used to infer from a set of attributes.
− inference through predicate logic uses a set of logical operations to relate individual data.
− the symbols used for the logic operations are :
" → " (implication), " ¬ " (not), " V " (or), " Λ " (and),
" ∀ " (for all), " ∃ " (there exists).
Examples of predicate logic statements :
1. *"Wonder"* is a name of a dog : **dog (wonder)**
2. All dogs belong to the class of animals : **∀ x : dog (x) → animal(x)**
3.All animals either live on land or in **∀ x : animal(x) → live (x,**
water : **land) V live (x, water)**
From these three statements we can infer that :
" Wonder* lives either on land or on water.*"
3. *Discuss about* Predicate Logic ?
Predicate Logic
The propositional logic, is not powerful enough for all types of assertions;
Example : The assertion **"x > 1"**, where **x** is a variable, is not a proposition
because it is neither true nor false unless value of **x** is defined.
For **x > 1** to be a proposition ,
− either we substitute a specific number for **x** ;
− or change it to something like
"There is a number x for which x > 1 holds";
− or **"For every number x, x > 1 holds"**.
Consider example :
“ All men are mortal.
Socrates is a man.
Then Socrates is mortal” ,
These cannot be expressed in propositional logic as a finite and logically valid argument (formula).
We need languages : that allow us to describe properties (*predicates*) of objects, or a relationship among objects represented by the variables .
Predicate logic satisfies the requirements of a language.
− *Predicate logic* is powerful enough for expression and reasoning.
− *Predicate logic* is built upon the ideas of *propositional logic.*
Predicate :
Every complete "sentence" contains two parts : a "subject" and a "predicate".
The *subject* is what (or whom) the sentence is about.

The *predicate* tells something about the subject;
Example :
A *sentence* **"Judy {runs}".**
The subject is **Judy** and the predicate is **runs** .
Predicate, always includes verb, tells something about the subject.
Predicate is a verb phrase template that describes a property of objects, or a relation among objects represented by the variables.
Example:
"The car Tom is driving *is blue*"; "The sky *is blue*" ;"The cover of this book *is blue*"
Predicate is **"is blue"** , describes property.
Predicates are given names; Let **„B"** is name for predicate **"is_blue"**.
Sentence is represented as **"B(x)"** , read as **"x is blue"**;
Symbol **"x"** represents an arbitrary Object .

4. Comparison between procedural and declarative language?

Comparison between Procedural and Declarative Language :

Procedural Language	Declarative Language
• Basic, C++, Cobol, etc.	• SQL
• Most work is done by interpreter of the languages	• Most work done by Data Engine within the DBMS
• For one task many lines of code	• For one task one SQL statement
• Programmer must be skilled in translating the objective into lines of procedural code	• Programmer must be skilled in clearly stating the objective as a SQL statement
• Requires minimum of management around the actual data	• Relies on SQL-enabled DBMS to hold the data and execute the SQL statement .
• Programmer understands and has access to each step of the code	• Programmer has no interaction with the execution of the SQL statement
• Data exposed to programmer during execution of the code	• Programmer receives data at end as an entire set
• More susceptible to failure due to changes in the data structure	• More resistant to changes in the data structure
• Traditionally faster, but that is changing	• Originally slower, but now setting speed records
• Code of procedure tightly linked to front end	• Same SQL statements will work with most front ends Code loosely linked to front end.
• Code tightly integrated with structure of the data store	• Code loosely linked to structure of data; DBMS handles structural issues
• Programmer works with a pointer or cursor	• Programmer not concerned with positioning
• Knowledge of coding tricks applies only to one language	• Knowledge of SQL tricks applies to any language using SQL

UNIT IV LEARNING

1. Why does uncertainty arise ?

Agents almost never have access to the whole truth about their environment.

Agents cannot find a caterorial answer.

Uncertainty can also arise because of incompleteness, incorrectness in agents understanding of properties of environment.

2. State the reason why first order, logic fails to cope with that the mind like medical diagnosis.

Three reasons

a.laziness: o it is hard to lift complete set of antecedents of consequence, needed to ensure and exceptionless rule.

b. Theoritical Ignorance: o medical science has no complete theory for the domain.

Practical ignorance: even if we know all the rules, we may be uncertain about a particular item needed.

3. Define the term utility?

The term utility is used in the sense of "the quality of being useful .", utility of a state is relative to

the agents, whose preferences the utility function is supposed to represent.

4. What is the need for probability theory in uncertainty ?

Probability provides the way of summarizing the uncertainty that comes from our laziness and ignorance . Probability statements do not have quite the same kind of semantics known as evidences.

5. What is the need for utility theory in uncertainty?

Utility theory says that every state has a degree of usefulness, or utility to in agent, and that the agent will prefer states with higher utility. The use utility theory to represent and reason with preferences.

6. What is called as principle of maximum expected utility ?

The basic idea is that an agent is rational if and only if it chooses the action that yields the highest expected utility, averaged over all the possible outcomes of the action. This is known as MEU.

7. What Is Called As Decision Theory ?

Preferences As Expressed by Utilities Are Combined with Probabilities in the General Theory of Rational Decisions Called Decision Theory.

Decision Theory = Probability Theory + Utility Theory.

8. Define Prior Probability?

p(a) for the Unconditional or Prior Probability Is That the Proposition A is True. It is important to remember that p(a) can only be used when there is no other information.

9. Define conditional probability?

Once the agents has obtained some evidence concerning the previously unknown propositions making up the domain conditional or posterior probabilities with the notation p(A/B) is used. This is important that p(A/B) can only be used when all be is known.

10. Define probability distribution:

Eg. P(weather) = (0.7,0.2,0.08,0.02). This type of notations simplifies many equations.

11. What is an atomic event?

An atomic event is an assignment of particular values to all variables, in other words, the complete specifications of the state of domain.

12. Define joint probability distribution

This completely specifies an agent's probability assignments to all propositions in the domain. The joint probability distribution p(x1,x2,--------xn) assigns probabilities to all possible atomic events;

where

X1,X2 ----- Xn 10 =variables.

13. Give the Baye's rule equation

W.K.T P(A ^ B) = P(A/B) P(B)---------- 1

P(A ^ B) = P(B/A) P(A)---------- 2

DIVIDING BY P(A) ;

WE GET

P(B/A) = P(A/B) P(B) ----------P(A)

14. What is meant by belief network?

A belief network is a graph in which the following holds

A set of random variables

A set of directive links or arrows connects pairs of nodes.

The conditional probability table for each node

The graph has no directed cycles.

15. What are the ways in which one can understand the semantics of a belief network?

There are two ways to see the network as a representation of the joint probability distribution to view it as an encoding of collection of conditional independence statements.

16. What is the basic task of a probabilistic inference?

The basic task is to reason in terms of prior probabilities of conjunctions, but for the most part, we will use conditional probabilities as a vehicle for probabilistic inference.

17. What are called as Poly trees?

The algorithm that works only on singly connected networks known as Poly trees. Here at most one undirected path between any two nodes is present.

18. Define casual support

E+X is the casual support for X- the evidence variables "above" X that are connected to X through its parent.

19. Define evidential support

E-X is the evidential support for X- the evidence variables "below" X that are connected to X through its children.

20. What is called as multiple connected graph?

A multiple connected graph is one in which two nodes are connected by more than one path.

21. What is the purpose of learning?

The idea behind learning is that percepts should be used not only for acting but also for improving the agent’s ability to act in the future.

22. What are issues in learning element?

i. Component

ii. Feedback

iii. Representation

23. What are the types of machine learning?

i. Supervised Learning

ii. Unsupervised Learning
iii. Reinforcement Learning

24. Differentiate Supervised and Unsupervised Learning

It involves learning a function from examples of its inputs And outputs	It involves learning patterns in the input when no specific output values are supplied.
Example: Applying Brake on the wet road, we can even skip on the road is a result.	Example: Day by day agent will learn about "Good traffic days" and "Bad traffic days" without any advice.

25. Define Reinforcement Learning.

This Learning is rather than being told what to do by teacher, a reinforcement learning agent must learn from occasional rewards.

Example

If taxi driver does not get a tip at the end of journey, it gives him a indication that his behavior is undesirable.

26. Define Inductive Learning.

An algorithm for supervised learning is given as input the correct value of the unknown function for particular inputs and it must try to recover the unknown function.

27. Define Classification Learning.

Learning a discrete valued function is called is called classification learning.

28. Define Regression learning.

Learning a continuous valued function is called regression learning.

29. What is parity and majority function?

Parity Function : It Returns 1 if and only if an even number of inputs are 1.

Majority function : It Returns 1 if more than half of its inputs are 1.

30. What is training set?

The complete set of examples is called the training set.

Example

Restaurant problem

Goal predicate "will wait"

31. Define Information gain.

Information gain from the attribute test is the difference between the original information requirement and the new requirement.

Gain (A) = I(p/(p+n)), n/ (p+n)) – Remainder(A)

32. What is test set?

Prediction is good if it turns out to be true, so can assess quality of hypotheses by

Checking its predictions against the correct classification once we know it. We do this on a set of examples is known as Test Set.

33. What is over fitting?

Whenever there is a large set of possible hypotheses, one has to be careful not to use the resulting freedom to find meaningless "regularity" in the data. This problem is called over fitting.

34. What is the purpose of cross validation?

It reduces over fitting. It can be applied to any learning algorithm, not just decision tree learning. The basic idea is to estimate how well each hypotheses will predict unseen data.

35. Mention the exercises which broaden the applications of decision trees.

i. Missing data
ii. Multivalued attributes
iii. Continuous and integer valued input attributes
iv. Continuous valued output attributes.

36. What is ensemble learning?

The idea of this learning is to select a whole collection or ensemble, of hypotheses from the hypotheses space and combine their predictions.

37. Define PAC – Learning Algorithm.

An learning algorithm that return hypotheses that are approximately correct is called PAC learning algorithm.

38. Define Decision list.

It is a logical expression of a restricted form, It consists of a series of tests, each of which conjunction of literals. If test succeeds, value is returned. If test fails, processing continues with the next test in the list.

39. What is the purpose of current best hypotheses search?

This search is to maintain a single hypothes

40. Differentiate generalization and specialization.

The extension of the hypotheses must be increased to include it. This is called generalization. The extension of the hypotheses must be decreased to exclude the example. This is called specialization.

41. Define Boundary set.

Each boundary will not be a point but rather a set of hypotheses called a Boundary set.

42. What are the two boundary sets?

i. G Set : a most general boundary set.
ii. S Set : a most specific boundary set.

43. Show the relationship of an entailment constraint.

Hypothesis ^ Descriptions |= classifications

44. Define EBL.

Explanation based learning, from the prior knowledge (or) information; we can infer a general rule. This kind of generalization process called explanation based learning (or) EBL.

45. What is the entailment constraints satisfied by EBL?

Hypothesis ^ Description |= classification
Background |= Hypothesis

46. Define RBL.

Relevance based Learning; the prior knowledge background concerns the relevance of a set of features to the goal predicate. This knowledge together with the observations,
Allows the agent to infer a new, general rule that explains the observations.
Hypothesis ^ Description |= classifications,
Background ^ Description ^ classifications |= Hypothesis.

47. Define knowledge based Inductive learning.

KBIL algorithm finds inductive hypotheses that explain sets of observations with the help of background knowledge.

48. Define Inductive Logic Programming (ILP).

ILP techniques perform KBIL on knowledge that is expressed in first order logic. ILP methods can learn relational knowledge that is not expressible in attribute based systems.

49. What is the purpose of memorization?

Memorization used in computer science to speed up programs by saving the results of computation. The basic idea of memo function is to accumulate a database of input and output pairs, when the function is called; it first checks the database to see whether it can avoid solving the problem.

50. What is the basic EBL process step?

i. Construct a proof using the available background knowledge.
ii. Construct a generalized proof tree for the variabilized goal using the same inference steps as in the original proof.
iii. Construct a new rule where LHS consists of the leaves of the proof tree and R.H.S is the variabilized goal.
iv. Drop any conditions that are true.

51. Define constructive induction algorithm.

Algorithms that can generate new predicates are called constructive induction algorithms.

52. What are the two main subroutines used for generating literals?

i. NEW – LITERALS
ii. CHOOSE - LITERALS

53. What are the 3 kinds of literals that can be added?

i. Literals using Predicate
ii. Equality and inequality literals
iii. Arithmetic comparisons

54. Define Bayesian Learning.

It calculates the probability of each hypotheses, given the data and makes predictions on that basis,
(i.e.) predictions are made by using all the hypotheses, weighted by their probabilities rather than by using just single "best" hypotheses.

55. Define MAP.

Maximum A Posteriori. A very common approximation is to make predictions based on single most probable hypotheses. This is MAP.

56. Define MDL.

The MDL (Maximum Description Length), is a learning method which attempts to minimize the size of the hypotheses and data encodings rather than work with probabilities.

57. What is Maximum – Likelihood hypotheses?

ML – it is reasonable approach when there is no reason to prefer one hypotheses over another a prior.

58. What are the methods for maximum likelihood parameter learning?

i. Write down an expression for the likelihood of the data as a function of the parameter.
ii. Write down the derivative of the log likelihood with respect to each parameter.
iii. Find the parameter values such that the derivatives are zero.

59. Define Naïve Bayes model.

In this model, the "class" variable C is the root and the "attribute" variable XI are the leaves. This model assumes that the attributes are conditionally independent of each other, given the class.

60. Define sum of squared errors.

The difference between the actual value yj and the predicated value ($\theta 1$ xj + $\theta 2$) so E is the sum of squared errors.

61. Define EM.

Expectation Maximization: the idea of EM is to pretend that we know the parameters of the model and then to infer the probability that each data point belongs to each component. After that we refit the components to the data, where each component is fitted to the entire data set with each point weighted by the probability.

62. What are the 2 steps in mixture model parameters?

i. E – Step
ii. M – Step

63. Define Neural Networks.

It consists of nodes or units connected by directed links. A link propagates the activation. Each link has a numeric weight which determines the strength and sign of the connection.

64. Give Activation function.

$$a_i = g(in_i) = g(\sum_{j=0}^{n} W_{j,i}\, a_j)$$

65. What are the two functions in Neural network's Activation functions?

i. Threshold function
ii. Sigmoid function

66. What are the categories of neural network structures?

i. Acyclic (or) Feed – forward networks
ii. Cyclic (or) Recurrent Networks

67. What is single layer feed forward neural network?

A network with all the inputs connected directly to the outputs is called a single layer neural network or a perceptron networks.

68. What is multilayer feed forward neural networks?

It consists of many hidden units. Each hidden unit act as a perceptron that represents a soft threshold functions in the input space. Output unit act as a soft threshold linear combination of several such functions.

69. Define Passive learning.

The agent's policy is fixed and the task is to learn the utilities of states, this could also involve learning a model of the environment.

70. Define Active Learning.

The agent must learn what to do. An agent must experience as much as possible of its environment in order to learn how to behave in it.

71. Define TD.

Temporal Difference learning: The key of TD is to use the observed transitions to adjust the values of the observed states so that they agree with the constraint equations.

Part - B

1. What are the components of planning system and explain in detail.

Reasoning is the act of deriving a conclusion from certain premises using a given methodology.

- Reasoning is a process of thinking; reasoning is logically arguing; reasoning is drawing inference.
- When a system is required to do something, that it has not been explicitly told how to do, it must reason. It must figure out what it needs to know from what it already knows.
- Many types of Reasoning have long been identified and recognized, but many questions regarding their logical and computational properties still remain controversial.

• The popular methods of Reasoning include abduction, induction, model based, explanation and confirmation. All of them are intimately related to problems of belief revision and theory development, knowledge assimilation, discovery and learning.
List the Machine learning algorithms and explain in detail
A formal language may be viewed as being analogous to of words or a collection of sentences.
AI – Reasoning: a collection
In computer science, a formal language is defined by precise mathematical or machine process able formulas.
‡ A formal langu;age **L** is characterized as a set **F** of finite-length sequences of elements drawn from a specified finite set **A** of symbols.
‡ Mathematically, it is an unordered pair **L = { A, F }**
‡ If **A** is words then the set **A** is called alphabet of **L**, and the elements of **F** are called words.
‡ If **A** is sentence then the set **A** is called the lexicon or vocabulary of **F**, and the elements of **F** are then called sentences.
‡ The mathematical theory that treats formal languages in general is known as **formal language theory**.
2. List out the different Methods of Reasoning?
Mostly three kinds of logical reasoning: Deduction, Induction, Abduction.
■ **Deduction**
‡ Example: "When it rains, the grass gets wet. It rains. Thus, the grass is wet."
This means in determining the conclusion; it is using rule and its precondition to make a conclusion.
‡ Applying a general principle to a special case.
‡ Using theory to make predictions
‡ Usage: Inference engines, Theorem provers, Planning.
■ **Induction**
‡ Example: "The grass has been wet every time it has rained. Thus, when it rains, the grass gets wet."
This means in determining the rule; it is learning the rule after numerous examples of conclusion following the precondition.
‡ Deriving a general principle from special cases
‡ From observations to generalizations to knowledge
‡ Usage: Neural nets, Bayesian nets, Pattern recognition
■
AI - Reasoning
Abduction
‡ Example: "When it rains, the grass gets wet. The grass is wet, it must have rained."
Means determining the precondition; it is using the conclusion and the rule to support that the precondition could explain the conclusion.
‡ Guessing that some general principle can relate a given pattern of cases
‡ Extract hypotheses to form a tentative theory
‡ Usage: Knowledge discovery, Statistical methods, Data mining.
■ **Analogy**
‡ Example: "An atom, with its nucleus and electrons, is like the sola system, with its sun and planets."

Means analogous; it is illustration of an idea by means of a more familiar idea that is similar to it in some significant features. And thus said to be analogous to it.
‡ finding a common pattern in different cases
‡ usage: Matching labels, Matching sub-graphs, Matching transformations.
Note: Deductive reasoning and Inductive reasoning are the two most commonly used explicit methods of reasoning to reach a conclusion.

3. Describe Bayes' Theorem?

- Bayesian view of probability is related to degree of belief.
- It is a measure of the plausibility of an event given incomplete knowledge.
- Bayes' theorem is also known as Bayes' rule or Bayes' law, or called Bayesian reasoning.
- The probability of an event A conditional on another event B ie P(A|B) is generally different from probability of B conditional on A ie P(B|A).
- There is a definite relationship between the two, P(A|B) and P(B|A), and Bayes' theorem is the statement of that relationship.
- Bayes theorem is a way to calculate P(A|B) from a knowledge of P(B|A).
- Bayes' Theorem is a result that allows new information to be used toupdate the conditional probability of an event.

Bayes' Theorem

Let S be a sample space.

Let $A_1, A_2, \ldots, A_n$ be a set of mutually exclusive events from S.

Let B be any event from the same S, such that $P(B) > 0$.

Then Bayes' Theorem describes following two probabilities :

$$P(A_k|B) = \frac{P(A_k \cap B)}{P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B)}$$ and

by invoking the fact $P(A_k \cap B) = P(A_k).P(B|A_k)$ the probability

$$P(A_k|B) = \frac{P(A_k).P(B|A_k)}{P(A_1).P(B|A_1) + P(A_2).P(B|A_2) + \cdots + P(A_n).P(B|A_n)}$$

Applying Bayes' Theorem :

Bayes' theorem is applied while following conditions exist.

‡ the sample space S is partitioned into a set of mutually exclusive events $\{A_1, A_2, \ldots, A_n\}$.

‡ within S, there exists an event B, for which $P(B) > 0$.

‡ the goal is to compute a conditional probability of the form : $P(A_k|B)$.

‡ you know at least one of the two sets of probabilities described below

◇ $P(A_k \cap B)$ for each A_k

4. Discuss in details about Rule Based Systems

A rule is an expression of the form **"if A then B"** where
Γ **A** is an assertion and **B** can be either an action or another assertion.
Γ Example : Trouble shooting of water pumps
Γ 1. If pump failure then the pressure is low
Γ 2. If pump failure then check oil level
Γ 3. If power failure then pump failure
Γ ■ Rule based systemnsists of a library of such rules.
Γ ■ Rules reflect essential relationships within the domain.
Γ ■ Rules reflect ways to reason about the domain.
Γ ■ Rules draw conclusions and points to actions, when specific information about the domain comes in. This is called inference.
Γ ■ The inference is a kind of chain reaction like : If there is a power failure then (see rules 1, 2, 3 mentioned above)
Γ Rule 3 states that there is a pump failure, and Rule 1 tells that the pressure is low, and
Γ Rule 2 gives a (useless) recommendation to check the oil level.

It is very difficult to control such a mixture of inference back and forth in the same session and resolve such uncertainties.

5. Explain in details about Bayesian Networks and Certainty Factors

A Bayesian network (or a belief network) is a probabilistic graphical model that represents a set of variables and their probabilistic independencies.
For example, a Bayesian network could represent the probabilistic relationships between diseases and symptoms. Given symptoms, the network can be used to compute the probabilities of the presence of various diseases.
Bayesian Networks are also called : Bayes nets, Bayesian Belief Networks (BBNs) or simply Belief Networks. Causal Probabilistic Networks (CPNs).
A Bayesian network consists of : a set of nodes and a set of directed edges between nodes. the edges reflect cause-effect relations within the domain. The effects are not completely deterministic (e.g. disease -> symptom) the strength of an effect is modeled as a probability.
More Complicated Bayesian Networks: The previous network was simple contained three nodes. Let us look at a slightly more complicated one in the context of heart disease.
Given the following facts about heart disease.
Γ Either smoking or bad diet or both can make heart disease more likely.
Γ Heart disease can produce either or both of the following two symptoms:

‡ high blood pressure
‡ an abnormal electrocardiogram
■ Here smoking and bad diet are regarded as causes of heart disease.
The heart disease in turn is a cause of high blood pressure and an abnormal electrocardiogram.

- **Bayesian Networks**

We have applied Bayesian probability theory, in earlier three examples (example 1, 2, and 3) , to relate two or more events. But this can be used to relate many events by tying them together in a network.

Consider the previous example 3 - Clinic trial

The trial says, the probability of the patients having HIV virus is 0.15.

A blood test done on patients :

If patient has virus, the test is +ve with probability 0.95.

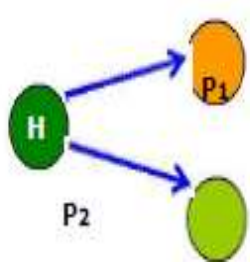
If the patient does not have the virus, the test is +ve with probability 0.02.

This means given : $P(H) = 0.15$; $P(P|H) = 0.95$; $P(P|\neg H) = 0.02$

Imagine, the patient is given a second test independently of the first; means the second test is done at a later date by a different person using different equipment. So, the error on the first test does not affect the probability of an error on the second test.

In other words the two tests are independent. This is depicted using the diagram below :

A simple example of a Bayesian Network.



Event H is the cause of the two events P_1 and P_2.

The arrows represent the fact that H is driving P_1 and P_2.

The network contained 3 nodes.

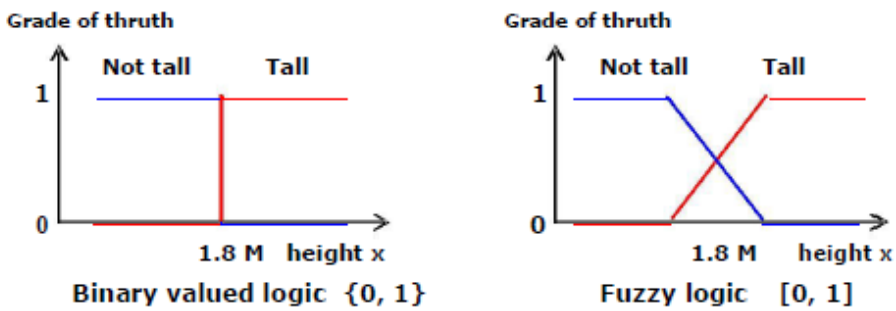
6. Describe in details about Fuzzy Logic

We have discussed only binary valued logic and classical set theory like :
A person belongs to a set of all human beings, and if given a specific subset, say all males, then one can say whether or not the particular person belongs to this set.
This is ok since it is the way human reason. e.g.,
IF person is male AND a parent THEN person is a father. The rules are formed using operators.
Here, it is intersection operator "AND" which manipulates the sets.
However, not everything can be described using binary valued sets.
The grouping of persons into "male" or "female" is easy, but as "tall" or "not tall" is problematic.
A set of "tall" people is difficult to define, because there is no distinct cut-off point at which tall begins.
Fuzzy logic was suggested by Zadeh as a method for mimicking the ability of human reasoning using a small number of rules and still producing a smooth output via a process of interpolation.

- **Description of Fuzzy Logic**

With fuzzy logic an element could partially belong to a set represented by the **set membership**. Example, a person of height **1.79 m** would belong to both tall and not tall sets with a particular degree of membership.

Difference between binary logic and fuzzy logic



A fuzzy logic system is one that has at least one system component that uses fuzzy logic for its internal knowledge representation.

Fuzzy system communicate information using fuzzy sets.

Fuzzy logic is used purely for internal knowledge representation and externally it can be considered as any other system component.

UNIT V INTELLIGENCE AND APPLICATIONS

1. Define planning.

Planning can be viewed as a type of problem solving in which the agent uses beliefs about actions and their consequences to search for a solution.

2. What are the features of an ideal planner?

i. The planner should be able to represent the states, goals and actions.
ii. The planner should be able to add new actions at any time.
iii. The planner should be able to use Divide and Conquer method for solving very big problems.

3. What are the components that are needed for representing an action?

The components that are needed for representing an action are: Action description. Precondition. Effect.

4. What are the components that are needed for representing a plan?

The components that are needed for representing a plan are: A set of plans steps. A set of ordering constraints.

5. What are the different types of planning?

The different types of planning are as follows:
i. Situation space planning. ii. Progressive planning. iii. Regressive planning.
iv. Partial order planning. v. Fully instantiated planning.

6. What are the ways in which incomplete and incorrect information's can be handled in planning?

They can be handled with the help of two planning agents namely,
i. Conditional planning agent.
ii. Replanning agent.

7. Define a solution.

A solution is defined as a plan that an agent can execute and that guarantees the achievement of goal.

8. Define a complete plan.

A complete plan is one in which every precondition of every step is achieved by some other step.

9. Define a consistent plan.

A consistent plan is one in which there are no contradictions in the ordering or binding constraints.

10. Define conditional planning.

Conditional planning is a way in which the incompleteness of information is incorporated in terms of adding a conditional step, which involves if – then rules.

11. Give the classification of learning process.

The learning process can be classified as:
Process which is based on coupling new information to previously acquired knowledge
a. Learning by analyzing differences.
b. Learning by managing models.
c. Learning by correcting mistakes.
d. Learning by explaining experience.
Process which is based on digging useful regularity out of data, usually called as Data base mining:
a. Learning by recording cases.
b. Learning by building identification trees.

c. Learning by training neural networks.

Descriptions from positive and negative examples.

12. What are the different types of induction heuristics?
There are two different types of induction heuristics. They are:
i. Require-link heuristics.
ii. Forbid-link heuristics.
13. What are the principles that are followed by any learning procedure?
i. The wait and see principle.
ii. The no altering principle.
iii. Martin's law.
14. State the wait and see principle.
The law states that, "When there is doubt about what to do, do nothing"
15. State the no altering principle.
The law states that, " When an object or situation known to be an example, fails to match a general model, create a special case exception model".
16. State Martin's law.
The law states that, " You cannot learn anything unless you almost know it already".
17. Define Similarity nets.
Similarity net is an approach for arranging models. Similarity net is a representation in which nodes denotes models, links connect similar models and links are tied to different descriptions.
18. Define Reification.
The process of treating something abstract and difficult to talk about as though it were concrete and easy to talk about is called as reification.
19. What is reified link?
The elevation of a link to the status of a describable node is a kind of reification. When a link is so elevated then it is said to be a reified link.
20. Define Communication.
Communication is the international exchange of information brought about by the production and perception of signs drawn from a shared system of conventional signs.
21. Define Language.
Language enables us to communicate most of what we have observed about the environment.
22. Define Formal Language.
A formal language is defined as a set of strings of terminal symbols. It is called as words.
23. What are the processes in communication?
i. Intention ii. Generation iii. Synthesis
iv. Perception v. Analysis vi Disambiguation vii. Incorporation
24. What are the parts in analyze?
i. Syntactic Interpretation
ii. Semantic Interpretation
iii. Pragmatic Interpretation
25. Define Semantic Interpretation.

The process of extracting the meaning of an utterance an expression in some representation language.
26. What are the processes of Relative Clause.

i. Generating Good English sentences
ii. Over generation
iii. Under generation

27. Define Parsing.

Parsing is the process of finding a parse tree for a given input string. It is also known as syntactic analysis.

28. What are the types of parsing?

i. Top down parsing
ii. Bottom up parsing

29. Define Top down parsing.

It starts with root node S and search for a tree that has the words as it leaves.

30. Define Bottom up parsing.

We start from the leaf nodes (i.e.) with the words and search for a tree with root S.

31. What are the algorithms to have efficient parsing?

i. Left to right parsing algorithm
ii. Chart Parsing algorithm.
iii. Left corner parsing

32. Define Augmentation.

The process of adding the existing rules of a grammar instead of introducing new rules. It is called Augmentation.

33. Define DCG.

The method of rewriting the existing rules in the grammar by the method of augmentation is called as DCG (Define Clause Grammar).

34. Define Sub categorization.

E 2 eliminates VP by mentioning which phrases can allow which verbs which are known as sub categorization.

35. Define Ambiguity.

The sentence that does not provide exact meaning are called ambiguous sentence.

36. What are the types of Ambiguity?

i. Lexical Ambiguity
ii. Syntactic Ambiguity
iii. Semantic Ambiguity

37. Define Disambiguation.

The speaker’s aim is to communicate some words in utterance and hearer work is to get back the meaning of the world from the knowledge of situation.

38. Define Discourse understanding.

A discourage is any string of language usually one that is more than one sentence long.

39. What are the two sub problems in discourse understanding?

The structure of coherent discourse.

40. What are the tasks in probabilistic language model?

i. Information retrieval
ii. Information Extraction
iii. Machine Translation

41. What are the types of smoothing?

i. Add one smoothing.
ii. Linear Interpolation Smoothing

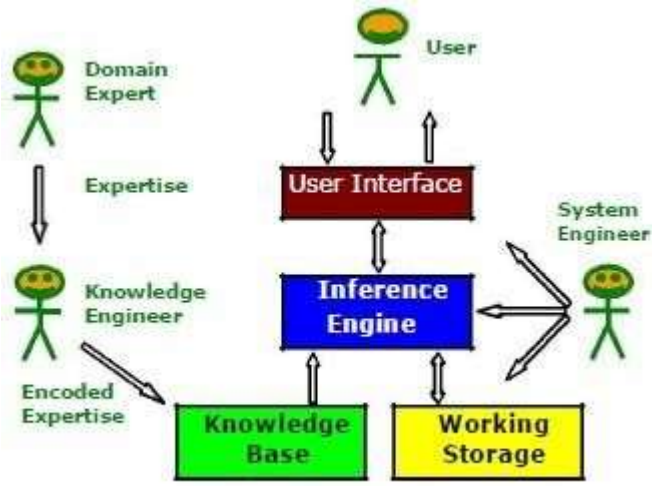
42. Define Segmentation?
The process of finding the words boundaries in a text with no spaces.
43. Define Information Retrieval (IR).
IR is the task of finding documents that are relevant to user's need for information.
44. What are the characteristics of IR?
i. A document collection.
ii. A query posed in a query language.
iii. A result set
iv. A presentation of the result set.
45. What are the types of evaluation of IR systems?
i. Recall
ii. Precision
46. What are the methods to do IR Refinements?
i. Case folding ii. Stemming iii. Recognize synonyms iv. Spelling correction
v. Meta data
47. What are the 3 mechanism to achieve performance improvement?
i. Relevance feedback
ii. Document classification
iii. Document clustering
48. What are the types of clustering technique?
i. Agglomerative clustering
ii. K-means clustering
49. What are the two data structures that make IR systems efficient?
i. Lexicon
ii. Inverted Index
50. Define Information Extraction.
It is a process of creating database entries by skimming a text and looking for occurrences of a particular class of object.
51. What are the types of information extraction systems?
i. Attribute based system
ii. Relational based system
52. What are the stages in Relational based systems?
i. Tokenization ii. Complex word handling iii. Basic groups iv. Complex phrases v. Merges structures
i. Reference Resolution

53. What are the types of machine translation?
i. Rough translation
ii. Restricted source translation
iii. Pre edited translation
iv. Literacy translation
54. How to use the parameters for machine translation?
i. Segment into sentences
ii. Estimate the French language model
iii. Align sentences
iv. Estimate the initial fertility model ,v. Estimate the initial word choice model
vi. Estimate the initial offset model ,vii. Improve all the estimates.

PART - B

1. What is an expert system shell

Expert systems which embody some non-algorithmic expertise for solving certain types of problems.
Expert systems have a number of major system components and interface with individuals who interact with the system in various roles. These are illustrated below.



Components of Expert System

Components and Interfaces
‡ Knowledge base : A declarative representation of the expertise; often in IF THEN rules ;
‡ Working storage : The data which is specific to a problem being solved;
‡ Inference engine : The code at the core of the system which derives recommendations from the knowledge base and problem-specific data in working storage;
‡ User interface : The code that controls the dialog between the user and the system.
■ Roles of Individuals who interact with the system
‡ Domain expert : The individuals who currently are experts in
solving the problems; here the system is intended to solve;
‡ Knowledge engineer : The individual who encodes the expert's
knowledge in a declarative form that can be used by the expert system;
‡ User : The individual who will be consulting with the system to get advice which would have been provided by the expert.

2. Explain in detail about Expert system shells?

Expert System Shells Many expert systems are built with products called expert system shells. A shell is a piece of software which contains the user interface, a format for declarative knowledge in the knowledge base, and an inference engine. The knowledge and system engineers uses these shells in making expert systems.
‡ Knowledge engineer : uses the shell to build a system for a particular problem domain.
‡ System engineer : builds the user interface, designs the declarative format of the knowledge base, and implements the inference engine.

Depending on the size of the system, the knowledge engineer and the system engineer might be the same person.

3. Explain in details about Expert System Features

The features which commonly exist in expert systems are :

■ Goal Driven Reasoning or Backward Chaining

An inference technique which uses IF-THEN rules to repetitively break a goal into smaller sub-goals which are easier to prove;

■ Coping with Uncertainty

The ability of the system to reason with rules and data which are not precisely known;

■ Data Driven Reasoning or Forward Chaining

An inference technique which uses IF-THEN rules to deduce a problem solution from initial data;

■Data Representation

The way in which the problem specific data in the system is stored and accessed;

■User Interface

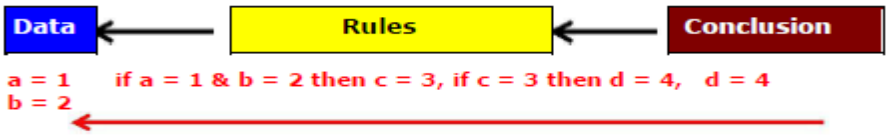
That portion of the code which creates an easy to use system;

■Explanations

The ability of the system to explain the reasoning process that it used to reach a recommendation.

4. Explain in details about Goal Driven Reasoning

Goal-driven reasoning, or backward chaining, is an efficient way to solve problems. The algorithm proceeds from the desired goal, adding new assertions found.



The knowledge is structured in rules which describe how each of the possibilities might be selected.

The rule breaks the problem into sub-problems. Example :

KB contains Rule set :

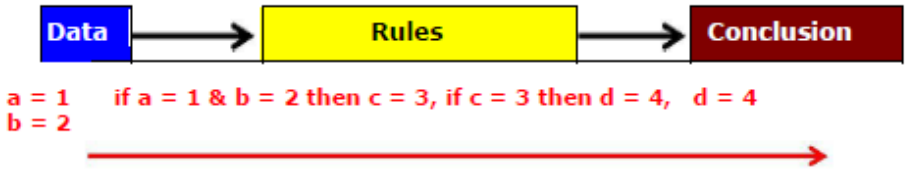
Rule 1:	If A and C	Then	F
Rule 2:	If A and E	Then	G
Rule 3:	If B	Then	E
Rule 4:	If G	Then	D

Problem : prove

If A and B true Then D is true

5. Explain in details about Data Driven Reasoning

The data driven approach, or Forward chaining, uses rules similar to those used for backward chaining. However, the inference process is different. The system keeps track of the current state of problem solution and looks for rules which will move that state closer to a final solution. The Algorithm proceeds from a given situation to a desired goal, adding new assertions found.



The knowledge is structured in rules which describe how each of the possibilities might be selected. The rule breaks the problem into sub-problems.

Example :

KB contains Rule set :

Rule 1:	If A and C	Then	F
Rule 2:	If A and E	Then	G
Rule 3:	If B	Then	E
Rule 4:	If G	Then	D

Problem : prove

If A and B true Then D is true

OBJECT ORIENTED ANALYSIS
DESIGN 2 MARKS

UNIT I - UNIFIED PROCESS AND USE CASE DIAGRAMS

1. What is an object? Give an example.

   A class represents a collection of **objects** having same characteristic properties that exhibit common behavior. It gives the blueprint or description of the **objects** that can be created from it. Creation of an **object** as a member of a class is called instantiation. Thus, **object** is an instance of a class.

2. What is the main advantage of Object Oriented Development?

   Some of the **advantages of object-oriented** programming include: 1. Improved software-**development** productivity: **Object-oriented** programming is modular, as it provides separation of duties in **object**-based program **development**. It is also extensible, as **objects** can be extended to include new attributes and behaviors.

3. Distinguish between method and messages in object.

   A **message** is the name of a selector, and the parameters for that selector. A selector is a symbol. A **method** is a piece of code **in a** class identified by a selector.

4. What is Analysis and Design?

   Object-oriented analysis is a process that groups items that interact with one another, typically by class, data or behavior, to create a model that accurately represents the intended purpose of the system

   Object-oriented design is the process of planning a system of interacting **objects for** the purpose of solving a software problem. It is one approach to software **design**

5. Define Use Case.

   A **use case** is a methodology **used** in system analysis to identify, clarify, and organize system requirements. The **use case** is made up of a set of possible sequences of interactions between systems and users in a particular environment and related to a particular goal

6. Give the different formats of Use cases.

   There are basically two **types of use cases** analysts can draw from: Business **Use Cases** and System **Use Cases**. Business **Use Cases** are more about what a user expects from a system while System **Use Cases** are more about what the system does. Both **use case types** can be represented by diagrams or text.

7. Define OOAD?

   Object-oriented analysis and design (**OOAD**) is a technical approach for analyzing and designing an application, system, or business by applying **object-oriented** programming, as well as using visual modeling throughout the software development process to guide stakeholder communication and product quality.

8. What is UML? List out the UML Diagrams.

   UML (Unified Modeling Language) is a general-purpose modeling language used to represent the structure of complex software in a visual form, and employed in software engineering. **UML** diagrams are also efficient for documenting complex computer systems and software

   Structure **Diagrams:**. Class **Diagram**. Component **Diagram**. Deployment **Diagram**.

41

Object **Diagram**. Package **Diagram**. Profile **Diagram**. Composite Structure **Diagram**.

Behavioral **Diagrams**: **Use Case Diagram**. Activity **Diagram**. State Machine **Diagram**. Sequence **Diagram**. Communication **Diagram**. Interaction Overview **Diagram**.

9. Classify the Three kinds of actors in use case.

An **actor** specifies a role played by a user or any other system that interacts with the subject. .Supporting **Actors**: A supporting **actor** in a **use case** in an external **actor** that provides a service to the system under design.

10. Define Unified Process (UP). List the 4 phases in UP.

The **Unified Process** (UP), or **Unified** Software Development **Process**, is a iterative and incremental software development framework from which a customized **process** can be defined.

Inception - The idea for the project is stated. The development team determines if the project is worth pursuing and what resources will be needed.

Elaboration - The project's architecture and required resources are further evaluated. Developers consider possible applications of the software and costs associated with the development.

Construction - The project is developed and completed. The software is designed, written, and tested.

Transition - The software is released to the public. Final adjustments or updates are made based on feedback from end users.

11. Illustrate the concepts of Generalization Relationship.

Generalization. In the **generalization** process, the common characteristics of classes are combined to form a class in a higher level of hierarchy, i.e., subclasses are combined to form a generalized super-class. It represents an “is – a – kind – of” relationship.

12. Comparison between Include and Extend use case relationships.

Extend is used when a use case conditionally adds steps to another first class use case. Include is used to extract use case fragments that are duplicated in multiple use cases.

13. Describe POS system?

A point of sale **system**, or **POS**, **is the** place where your customer makes a payment for products or services at your store. Simply put, every time a customer makes a purchase at your store, they're completing a point of sale transaction

14. Describe the Primary goals in the Design of UML

The primary goals in the design of the UML were:

Provide users with a ready-to-use, expressive visual modeling language so they can develop and exchange meaningful models.

Provide extensibility and specialization mechanisms to extend the core concepts.

15. Illustrate the relationship used in Use case.

The include **relationship** supports the reuse of functionality in a **use-case** model.

In **UML** modeling, you can **use** an extend **relationship** to specify that one **use case** (extension) extends the behavior of another **use case** (base)

16. What are the three ways and perspectives to Apply UML?

    UML as programming language **Perspectives**-Conceptual **perspective**, Specification (software) **perspective**, Implementation (Software) **perspective**.

17. Generalize the concepts of Use case Modeling and list the advantages of Use case Modeling.

    Use case diagrams are **used** to gather the requirements of a system including internal and external influences. These requirements are mostly design requirements. Hence, when a system is analyzed to gather its functionalities, **use cases** are prepared and actors areidentified.

18. When to use Use cases? Evaluate it?

    Use case diagrams are a way to capture the system's functionality and requirements in UML **diagrams**. It captures the dynamic behavior of a live system. A **use case diagram** consists of a **use case** and an actor

19. Generalize your views about inception in Use case.

    Inception. The primary goal of the **Inception** phase is to establish the case for the viability of the proposed system. The tasks that a project team performs during **Inception** include the following: Defining the scope of the system

20. Evaluate and Name the UML diagrams used for the following:

    a)Modeling Requirements
    b) Modeling Workflows
    c) Modeling behavior of an object.
    d)Interaction between groups of objects.

UNIT II - STATIC UML DIAGRAMS

1. Define Class Diagram?

   A **class diagram** in the Unified Modeling Language (UML) is a type of static structure **diagram** that describes the structure of a system by showing the system's **classes**, their attributes, operations (or methods), and the relationships among objects

2. Define attribute?

   In domain modeling **class diagrams**, an **attribute** represents a data definition for an instance of a classifier **Attributes** that are defined in the scope of the **class**, that is static, are shown as underlined. The visibility styles of **attributes** can be represented as text symbols

3. Express the meaning of Elaboration and What are the tasks.

   Elaboration is the initial series of iterations during which the team does serious investigation, implements (programs and tests) the core architecture, clarifies most requirements, and tackles the high-risk issues. In the UP, "risk" includes business value

4. Define Conceptual class.

   The domain model illustrates **conceptual classes** or vocabulary in the domain. o Informally: a **conceptual class** is an idea, thing, or object. o Formally: a **conceptual class** may be

considered in terms of its symbol, intension, and extension. · Symbol—words or images representing a **conceptual class**.

5. Express why we call a domain model a“VisualDictionary".
6. Illustrate the Relationships used in class diagram

   Generalization **relationships** are used in **class**, component, deployment, and use-case **diagrams** to indicate that the child receives all of the attributes, operations, and **relationships** that are defined in the parent.
7. Define Domain Model.

   Domain modeling is a technique used to understand the project problem description and to translate the requirements of that project into software components of a solution. . A **domain model** contains conceptual classes, associations between conceptual classes, and attributes of a conceptual **class**.
8. Compare Aggregation and Composition.

   In both **aggregation** and **composition** object of one class "owns" object of another class. But there is a subtle **difference**: **Aggregation** implies a relationship where the child can exist independently of the parent..**Composition** implies a relationship where the child cannot exist independent of the parent
9. Illustrate the usage of Description class.

   A **class** represents a collection of objects having same characteristic properties that exhibit common behavior. It gives the blueprint or **description** of the objects that can be created from it. Creation of an object as a member of a **class** is called instantiation. Thus, object is an instance of a **class**.
10. Give the meaning of abstract conceptual class.

    It is useful to identify abstract classes in the domain model because they constrain what classes it is possible to have concrete instances of, thus clarifying the rules of the problem domain. If every

member of a class C must also be a member of a subclass, then class C is called an abstract conceptual class.

11. Comparison between sequence diagram and Use case diagram.

Sequence diagrams only specify the ordering of events and not the exact timings of events. An activation box represents the period during which an operation is executed. Shows interactions **between** objects in visual and chronological (time) order. Refines **use cases** with more details.

Use case diagram depicts how actors (a person or a system) interact with a solution to accomplish one or more of the person's or systems goals. It visually depicts the scope of the solution, the actors involved, and the use cases.

12. Analyze the concepts of Association.

The **concept of association** is used in describing relationships between entities in an entity relationship model or classes in an object model. The basic **concept** is the same regardless of the type of modeling technique being used.

13. Generalize the use of Sequence Diagram.

A **sequence diagram** is a type of interaction **diagram** because it describes how—and in what order—a group of objects works together. These **diagrams** are used by software developers and business professionals to understand requirements for a new system or to document an existing process

14. Differentiate Class diagram and Interaction diagram.

A **class diagram** shows a set of **classes**, interfaces and their relationships and illustrates the static design view of a system, while a **sequence diagram** shows the **sequence** of actions that occurs **in a** system and illustrates the dynamic view of a system.

15. Analyze the concepts of Noun Phrase Identification from use cases

Using this method, you have to read through the Use cases, interviews, and requirements specification carefully, looking for noun phrases.

16. Summarize the strategies to find Conceptual classes.

- Reuse or modify existing models. This is the first, best, and usually easiest approach, and where I will start if I can.
- Use a category list.
- **Identify** noun phrases

17. Define Use case Diagram

Use case diagrams consists of actors, **use cases** and their relationships. The **diagram** is **used** to model the system/subsystem of an **application**. A single **use case diagram** captures a particular functionality of a system. Hence to model the entire system, a number of **use case diagrams** are **used**

18. Illustrate the concepts of Class Hierarchy.

A **class hierarchy** or **inheritance** tree in computer science is a classification of object types, denoting objects as the instantiations of **classes** (**class** is like a blueprint, the object is what is built from that blueprint) inter-relating the various **classes** by relationships such as "inherits", "extends".

19. When to use class diagram?

The **purpose of class diagram** is to model the static view of an application. **Class diagrams** are the only **diagrams** which can be directly mapped with object-oriented languages and thus widely used at the time of construction.

UNIT III - DYNAMIC AND IMPLEMENTATION UML DIAGRAMS

1. Define System sequence diagram.

A **system sequence diagram** (SSD) is a **sequence diagram** that shows, for a particular scenario of a use case, the events that external actors generate, their order, and possible inter-**system** events

2. What are the Common Notations in UML Interaction Diagram?

Interaction diagrams are models that describe how a group of objects collaborate in some behavior - typically a single use-case. The **diagrams** show a number of example objects and the messages that are passed between these objects within the use-case **Interaction**
diagrams come in two forms, both present in the **UML**.

3. Illustrate the concepts and uses of Communication Diagram.

Interaction diagrams are models that describe how a group of objects collaborate in some behavior - typically a single use-case. The **diagrams** show a number of example objects and the messages that are passed between these objects within the use-case **Interaction**
diagrams come in two forms, both present in the **UML**.

4. Compare Activity and state chart diagram? Mention the Elements of an Activity Diagram.
Activity diagrams describe **activities and state charts** describe states. So those models are orthogonal - you might imagine there is an **activity** between two states (something that occurs during the **transition**) and you might say there is a **state** between two **activities**

5. List out the Types of Interactions diagram.
The UML includes **interaction diagrams** to illustrate how objects **interact** via messages. They are used for dynamic object modeling. There are two common **types**: **sequence** and communication **interaction diagrams**

6. Differentiate the strengths and weaknesses of Sequence and Communication Diagram

A Sequence diagram is an interaction diagram that details about the operation that is carried out. Sequence diagram captures the interaction between the objects in the context of collaboration. Sequence diagrams are time focused and they show the order of the interaction visually by using the vertical axis of the diagram to represent time.

Collaboration Diagram represents the interaction of the objects to perform the behavior of a particular use case or a part of use case. The designers use the Sequence diagram and Collaboration Diagrams to define and clarify the roles of the objects that perform a particular flow of events of a use case.

7. Interpret the meaning of event, state and Transition.

Event—occurrence that is relevant to an object or application. **State**—the **state** of an object is determined by the value of some of its attributes and the presence or absences of links with other objects. **Transition**—the movement from one **state** to another, triggered by an **event**

8. Define State Chart Diagram? When to use State Diagram?

Statechart diagram is used to describe the **states** of different objects in its life cycle. Emphasis is placed on the **state** changes upon some internal or external events. These **states** of objects are important to analyze and implement them accurately. **Statechart diagrams** are very important for describing the **states**.

9. Explain how Synchronous and asynchronous messages are depicted in communication diagram.

Synchronous Message. A **synchronous message** requires a response before the interaction can continue. It's usually drawn using a line with a solid arrowhead pointing from one object to another.

10. Differentiate Component and deployment diagram.

Component diagram shows you how different elements of your system have been grouped together (into assemblies / dlls etc) - and the link between these **components**. A **Deployment diagram** takes you one step further and describes on which hardware elements do these **components** reside

11. Analyze the use of UML Package Diagram

Package diagrams are **used** to structure high level system elements. **Packages** are **used** for organizing large system which contains **diagrams**, documents and other key deliverables. **Package Diagram** can be **used** to simplify complex class **diagrams**, it can group classes into **packages**

12. Define Package. Draw UML notation for Package.

Package diagram is **UML** structure **diagram** which shows **packages** and dependencies between the **packages**. Model **diagrams** allow to show different views of a system, for example, as multi-layered (aka multi-tiered) application - multi-layered application model

13. When to use Deployment diagram? Analyze it.

Deployment diagrams are mainly used by system engineers. These **diagrams** are used to describe the physical components (hardware), their distribution, and association. **Deployment diagrams** can be visualized as the hardware components/nodes on which the software components reside.

14. Design the notation of Component and Node

Node instances are based on existing **nodes**. While a **node represents** a generic type of computational device, a **node** instance **represents** a specific and defined **node** in your system environment. You can use **node** instances in **deployment diagrams** to **represent** resources that exist at run time..

15. Describe the basic elements of a Deployment Diagram.

A **deployment diagram** consists of nodes. Nodes are nothing but physical hardware used to **deploy** the application. The application is assumed to be a web-based application, which is **deployed** in a clustered environment using server 1, server 2, and server 3. The user connects to the application using the Internet.

16. What is package diagram? Classify the three layers of package diagram.

UML **Package Diagrams** Notation. **Package diagram** is UML structure **diagram** which shows **packages** and dependencies between the **packages**. Model **diagrams** allow to show different views of a system, for **example**, as multi-layered (aka multi-tiered) application - multi-layered application model.

17. When to use Component Diagram? Analyze it.

Component diagrams are used to display various **components** of a software system as well as subsystems of a single system. They are used to represent physical things or **components** of a system. It generally visualizes the structure and an organization of a system.

18. Generalize the concepts of Node and Component.

Deployment diagrams are mainly used by system engineers. These **diagrams** are used to describe the physical **components** (hardware), their distribution, and association. **Deployment diagrams** can be visualized as the hardware **components**/nodes on which the software **components** reside.

The three-dimensional boxes, known as **nodes**, represent the basic software or hardware elements, or **nodes**, in the system. Lines from **node** to **node** indicate relationships, and the smaller shapes contained within the boxes represent the software artifacts that are deployed.

19. List the two types of deployment diagram node.

Artifacts and **nodes** are **the** common factors **used** for **deployment diagram. The** two **types of nodes** are Device and execution environment.

UNIT IV - DESIGN PATTERNS

1. Define Design Pattern.

A **design pattern** provides a general reusable solution for the common problems occurs in software **design**. The **patterns** typically show relationships and interactions between classes or objects. That **means** a **design pattern** represents an idea, not a particular implementation

2. Illustrate the concepts of GRASP.

The different **patterns** and **principles** used in **GRASP** are controller, creator, indirection, information expert, high cohesion, low coupling, polymorphism, protected variations, and pure fabrication

3. "A system must be loosely coupled and highly cohesive"-Justify.

A system must be loosely coupled and highly cohesive, because highly cohesive means the class does one job well. Low cohesion indicates that there are elements in the class which have little to do with each other.

4. Discover the Limitations of Factory Pattern.

Factory design pattern provides approach to code for interface rather than implementation.Factory pattern removes the instantiation of actual implementation classes from client code. Factory pattern provides abstraction between implementation and client classes through inheritance.

5. Define modular design.

Modular design, or **modularity** in **design**, is an approach (**design** theory and practice) that subdivides a system into smaller parts called modules (such as **modular** process skids), which can be independently created, modified, replaced or exchanged between different systems.

6. Analyze the situation to use Factory method pattern and its advantages.

Factory Method **Pattern** allows the sub-classes to choose the type of objects to create. It promotes the loose-coupling by eliminating the need to bind application-specific classes into the code

7. Generalize your view on creator

Creator is a **GRASP** Pattern which helps to decide which class should be responsible

for creating a new instance of a class

8. Summarize the list of structural patterns used during design phase of software development

structural design **patterns** are design **patterns** that ease the design by identifying a simple way to realize relationships among entities. Examples of **Structural Patterns** include: Adapter **pattern**: 'adapts' one interface for a class into one that a client expects..

9. Analyze the concepts of Coupling and Low coupling

Low coupling refers to a relationship in which one module interacts with another module through a simple and stable interface and does not need to be concerned with the other module's internal implementation.

coupling is the degree of interdependence between software modules; a measure of how closely connected two routines or modules are; the strength of the relationships between modules.

10. Interpret the need of Information Expert.

Information expert (also **expert** or the **expert** principle) is a principle used to determine where to delegate responsibilities such as methods, computed fields, and so on.

11. Distinguish between coupling and cohesion.

Coupling is **the** indication of **the** relationships **between** modules. **Cohesion** is **the** indication of **the** relationship within module. **Cohesion** shows **the** module's relative functional strength. **Cohesion** is a degree (quality) to which a component / module focuses on **the** single thing.

12. When to use Patterns?

In object-oriented programming, a **pattern** can contain the description of certain objects and object classes to be used, along with their attributes and dependencies, and the general approach to how to solve the problem

13. Analyze the benefits of controller.

Controller. The **controller** pattern assigns the responsibility of dealing with system events to a non-UI class that represents the overall system or a use case scenario. A **controller** object is a non-user interface object responsible for receiving or handling a system event.

14. Define Refactoring.

Refactoring is "the process of changing a software system in such a way that it does not alter the external behavior of the code yet improves its internal structure,"

15. Generalize the concepts of responsibility. What are the various types of responsibilities?

object is that these two separate entities are encapsulated together to form a single, atomic entity.....A straightforward **definition** for **object-responsibility** is this: An **object** must contain the data (attributes) and code (methods) necessary to perform any and all services that are required by the **object**.

16. Discuss the benefits and the types of adapter pattern.

A **design pattern** provides a general reusable solution for the common problems occurs in software **design**. The **patterns** typically show relationships and interactions between classes or objects. The idea is to speed up the development process by providing well tested, proven development/**design** paradigm

17. Define Observer Pattern.

The **observer pattern** is a software design **pattern** in which an object, called the subject, maintains a list of its dependents, called observers, and notifies them automatically of any state changes, usually by calling one of their methods

18. Define TDD. Mention the advantages of TDD.

TDD helps to create better modularized, extensible and flexible code. **Test Driven**

Development approach drives the Agile team to plan, develop and test the small units to be integrated at advanced stage. Under this approach, the concerned member delivers and performs better because of being more focused on smaller unit.

19. Give the GOF design patterns.

    Design patterns are solutions to software **design** problems **you** find again and again in real-world application development. **Patterns are** about reusable designs and interactions of objects.

20. Illustrate the benefits of bridge pattern.
    - It enables the separation of implementation from the interface.
    - It improves the extensibility.
    - It allows the hiding of implementation details from the client.

UNIT V- TESTING

1. Describe the term SQA.

   Software quality assurance (**SQA**) is a process that ensures that developed software meets and complies with **defined** or standardized quality specifications. **SQA** is an ongoing process within the software development life cycle (SDLC) that routinely checks the developed software to ensure it meets desired quality measures

2. Give the main techniques of Quality Assurance.

   Data Conversion Testing. When a company migrates data to a new software, it becomes vulnerable. ...
   Regression Testing. In order for a software company to grow, the product needs to do the same. ...
   Mobile Testing. Mobile devices are inherently different than traditional computer environments.
   ...
   User Acceptance Testing.

3. Illustrate the impact object orientation in testing.

   Issues in Object oriented testing. Traditional **testing** methods are not directly applicable to OO programs as they involve OO concepts including encapsulation, inheritance, and polymorphism. These concepts lead to **issues**, which are yet to be resolved.

4. Define the term Object interoperability.

   Interoperability is one of the key aspects related to the construction of large **object**-oriented systems, and can be defined as the ability of two or more en- tities to communicate and cooperate despite differences in the implementation language, the execution environment or the model abstraction.

5. Summarize the basic activities are performed in using debugging tools.

   Discover. Identify a problem.
   Locate. Determine where in the code the problem occurs.
   Inspect. Examine the control flow and data structures of the running code to find the cause of the problem.
   Fix. ...
   Confirm.

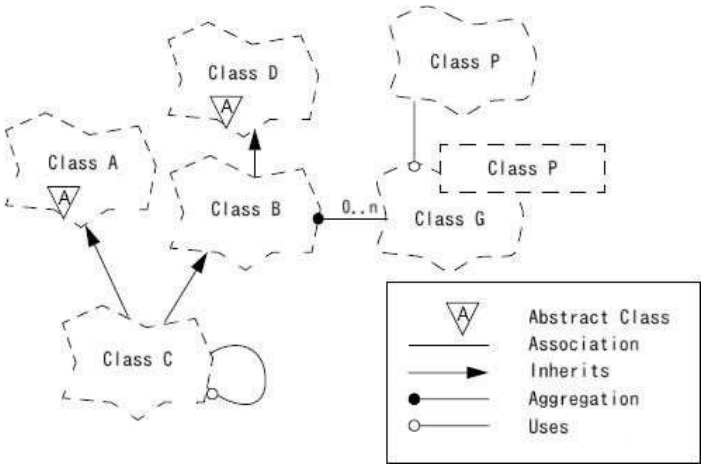
6. Define test plan? What are its components?

   A **test plan** is a detailed document that outlines the **test** strategy, **Testing** objectives, resources (manpower, software, hardware) required for **testing**, **test** schedule, **Test** Estimation and **test** deliverables.

7. Why quality assurance is needed? Summarize it.

Proactiveness in terms of error detection helps organization eliminate the possibilities of a failure, thus, enabling production of a high-performing, intuitive, and stable software. **Quality assurance** activities help identifying and eradicating processes that produce waste or do not meet the standards.

8. Give the Booch methodology diagrams.



9. Define block box testing?

Black-box testing is a method of software **testing** that examines the functionality of an application without peering into its internal structures or workings. This method of **test** can be applied virtually to every level of software **testing**: unit, integration, system and acceptance

10. Illustrate the different kinds of errors you might encounter when you run your program?

Logical Errors. We will divide our errors into three classes: logical, syntax, and semantic. ...

Syntax Errors. Syntax errors in computer programming differ from logical errors in that they don't follow a correct sequence in the computer language. ...

Semantic Errors. Semantic errors are improper uses of “program statements.”

11. List out the Testing strategies.

Different Types Of Software Testing

- Unit Testing.
- Integration Testing.
- System Testing.
- Sanity Testing.
- Smoke Testing.
- Interface Testing.
- Regression Testing.
- Beta/Acceptance Testing.

12. Analyze as to which object oriented methodology is well suited for (i). Design (ii). Analysis (iii).Full life cycle(iv). Real time systems

.

13. Analyze the Booch system development process.

<u>Macro development process:</u>

It consists of following steps:

1. Design or create the system architecture.
2. Evolution or implementation.
3. Conceptualization
4. Maintenance.
5. Analysis and development of the model.

Micro development process:

It represents the daily activities of small or large group of software developers. It describes the day-to-day activities.

The micro development process consists of the following steps.

1. Identify classes and objects.
2. Identify classes and object semantics.
3. Identify classes and object relationships.
4. Identify classes and object interfaces and implementation.

14. What are the steps needed to create a test plan? Illustrate it.

How to write a Test Plan

1. Analyze the product.
2. Design the Test Strategy.
3. Define the Test Objectives.
4. Define Test Criteria.
5. Resource Planning.
6. Plan Test Environment.
7. Schedule & Estimation.
8. Determine Test Deliverables

15. Generalize the concepts of implication of Inheritance.

Generalization is a mechanism for combining similar classes of objects into a single, more general class. **Generalization** is a bottom-up process. **Generalization and inheritance** are powerful abstractions for sharing similarities among classes while preserving their differences.

16. Give the four phases of object oriented modeling Techniques(OMT).

The entire **OMT** software development process has **four phases**: analysis, system design, **object** design, and implementation of the software. Most of the **modeling** is performed in the analysis **phase**. In this **phase**, three basic **models - Object** Model, Dynamic Model and Functional Model are developed.

17. Comparison between patterns and frameworks.

A **framework** is a set **of** related classes **to** perform a certain task. Those classes may or may not implement a certain design **pattern**. A design **pattern** is a well-established design for tackling a problem. A **framework** is an actual package **of** code you use **to** make building applications easier.

18. Generalize the impact of an object orientation on testing.

Issues in Object oriented testing. Traditional **testing** methods are not directly applicable to OO programs as they involve OO concepts including encapsulation, inheritance, and polymorphism. These concepts lead to **issues**, which are yet to be resolved.

19. What is test cases? List the guidelines for developing quality assurance test cases.

- Use a Strong Title. ...
- Include a Strong Description. ...

- Include Assumptions and Preconditions. ...

- Keep the Test Steps Clear and Concise. ...
- Include the Expected result. ...
- Make it Reusable. .

IT8602 Mobile Communication
2 Mark Questions and Answers

UNIT I

1. What is mobile computing?

Mobile computing is a technology that allows transmission of data, via a computer, without having to be connected to a fixed physical link.

2. Difference between FDMA and TDMA.

FDMA	TDMA
Frequency Division Multiplexing Access is a method of dividing a single band into 30 discrete channels.	Time Division Multiplexing Access that also allows more subscribers to use the same frequency band.
Each channel would then be capable of handling separate traffic, whether it's a call or a data transfer.	TDMA divides a single channel into 3 discrete time partitions.
FDMA is also used in the Total Access Communication System (TACS).	TDMA is utilized by Digital-Advanced Mobile Phone System (D-AMPS) and Global System for Mobile communications (GSM).

3. Why do Hidden and Exposed terminal problem arise.

Hidden terminal problem is due to the fact that a node (say A) transmitting to another node (say B) cannot hear transmissions from another node C, which might also be transmitting to B, and might interfere with the A-to-B transmissions.

Exposed node problem occurs when a node is prevented from sending packets to other nodes because of a neighboring transmitter.

4. Prove that Barker code has good auto correlation.

When the receiver attempts to correlate the received coded symbols with respect to any of the codes which it internally generates, it is not able to correlate even when it uses exactly the same code as the one used for transmission.

Reasons for no correlation:

- Propagation delay
- Inappropriate code

5. Write the characteristics of mobile computing.

- Ubiquity
- Location awareness
- Adaptation
- Broadcast
- Personalization

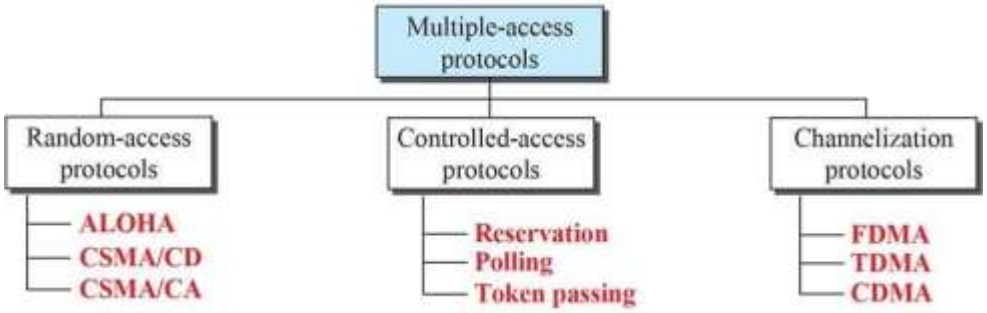
6. Draw the 3 Tier structure of Mobile computing application.

Presentation(Tire-1)
Application(Tier-2)
Data(Tier-3)

7. State the issue in wireless MAC protocol.

- Hidden Terminal Problem.
- Exposed Terminal Problem.

8. Draw the Taxonomy of MAC protocol.



9. Write the different Random assignment scheme in MAC.

- ALOHA
- Slotted ALOHA
- CSMA
- CSMA/CD
- CSMA/CA

10. Difference between Mobile Computing vs Wireless Networking.

Mobile Computing	Wireless Networking
Mobile Computing denotes accessing information and remote computational services.	Wireless Networking provides the basic communication infrastructure.
It is based on wireless networking.	Wireless Networking is an important ingredient of mobile computing.

11. What are applications of Mobile Computing?

- Vehicles
- Emergencies
- Business
- Replacement of wired networks
- Infotainment
- Location dependent services
- Mobile and wireless devices

12. What are the obstacles in mobile communications?

- Interference
- Regulations and spectrum
- Low Bandwidth
- High delays, large delay variation

- Lower security, simpler to attack
- Shared Medium
- Adhoc-networks

13. What are the Advantages of wireless LAN?

- Flexibility
- Planning
- Design
- Robustness

14. Mention some of the disadvantages of WLANS?

- Quality of service
- Proprietary solutions.
- Restrictions
- Safety and Security

15. Ubiquitous computing

Ubiquitous computing enhances computer use by making many computers available throughout the physical environment, while making them effectively invisible to users.

16. What is the aim of ubiquitous computing?

The aim of ubiquitous computing is to design computing infrastructures in such a manner that they integrate seamlessly with the environment and become almost invisible.

It integrates computation into the environment, rather than having computers which are distinct objects

17. Define Mobile computing.
18. Write a short note on mobile networking.
19. List any two mobile computing application.
20. Give two characteristics of mobile computing.
21. Comment on MAC protocol.
22. Write a short note on fixed assignment schemes.
23. What is virtual home environment?
24. What are the three tier setup in mobile computing?
25. Mention three applications of mobile computing.
26. List the advantages of mobile computing. (JUNE 16)
27. Explain hidden and exposed terminal problems in infrastructure – less network. (JUNE 16)
28. What are the different random assignment scheme in MAC? (DEC 16)
29. List out the differences between mobile computing and wireless networking. (MAY 17, DEC 17,MAY 18)
30. List some random assignment scheme. (MAY 17)
31. "MAC protocol designed for infrastructure based wireless network may not work satisfactory in infrastructure- less environment" Justify. (DEC 17)
32. List the issues of wireless MAC. (MAY 18)

UNIT-II

1. Define GSM?

The global system for mobile communication (GSM) was developed by Groupe Speciale Mobile (GSM) which was founded in Europe in 1992. The GSM is a standard for mobile telecommunication through a cellular network at data rates if up to 14.4 kbps. Nowadays, it consists of a set of standards and protocols for mobile telecommunication.

2. Define GPRS?

General Packet Radio Service (GPRS) is a packet oriented service for mobile devices data communication which utilizes the unused channels in TDMA mode in a GSM network and also sends and receives packet of data through the internet.

3. What are subsystems in GSM system?

Radio subsystem (RSS)

Network & Switching subsystem (NSS) Operation subsystem (OSS)

4. What are the control channel groups in GSM?

The control channel groups in GSM are:

Broadcast control channel (BCCH)
Common control channel (CCCH)
Dedicated control channel (DCCH)

5. What are the four types of handover available in GSM?

- Intra cell Handover
- Inter cell Intra BSC Handover
- Inter BSC Intra MSC handover
- Inter MSC Handover

6. What is the frequency range of uplink and downlink in GSM network?

The frequency range of uplink in GSM network is 890-960 MHz The frequency range of downlink in GSM network is 935-960 MHz

7. What are the security services offered by GSM?

The security services offered by GSM are:

- Access control and authentication.
- Confidentiality.
- Anonymity.

8. What are the reasons for delays in GSM for packet data traffic?

Collisions only are possible in GSM with a connection establishment. A slotted ALOHA mechanism is used to get access to the control channel by which the base station is told about the connection establishment attempt. After connection establishment, a designated channel is installed for the transmission.

9. What is meant by beacon?

A beacon contains a timestamp and other management information used for power management and roaming. e.g., identification of the base station subsystem (BSS)

10. List out the numbers needed to locate an MS and to address the MS.

The numbers needed to locate an MS and to address the MS are: Mobile station international ISDN number (MSISDN) International mobile subscriber identity (IMSI)

Temporary mobile subscriber identity (TMSI) Mobile station roaming number (MSRN)

11. What is meant by GPRS?

The General Packet Radio Service provides packet mode transfer for applications that exhibit

traffic patterns such as frequent transmission of small volumes.

12. What is meant by GGSN?

GGSN is Gateway GPRS Support Node. It is the inter-working unit between the GPRS network and external packet data networks. The GGSN is connected to external networks via the GI interface and transfers packets to the SGSN via an IP based GPRS backbone network.

13. What is meant by SGSN?

SGSN is Serving GPRS Support Node. It supports the MS via the GB interface. The GSN is connected to a BSC via frame relay.

14. What is meant by BSSGP?

BSSGP is Base Station Subsystem GPRS Protocol. It is used to convey routing and QoS-related information between the BSS and SGSN.BSSGP does not perform error correction and works on top of a frame relay network.

15. Expand GSM, GPRS and UMTS.

- Global System for Mobile Communication (GSM)
- General Packet Radio Service (GPRS)
- Universal Mobile Telecommunication System (UMTS)

16. Write the GSM services.

- Supplementary Services
- Bearer Services
- Tele Services

17. What is RSS?

- RSS stands for Radio subsystem (RSS)
- RSS comprises all radio specific entities

18. Name the entities of RSS.

- Base Station Subsystem (BSS)
- Base Transceiver Station (BTS)
- Base Station Controller (BSC)
- Mobile Station (MS)

19. Mention the advantages of GSM.

- Communication
- Total mobility
- Worldwide connectivity
- High capacity
- High transmission quality.
- Security functions

20. What does SIM card contain?

- a personal identity number (PIN)
- a PIN unblocking key (PUK)
- an authentication key Ki
- the international mobile subscriber identity (IMSI)

20. Mention the disadvantages of GSM.

- No end-to-end encryption of user data
- Reduced concentration while moving
- Electromagnetic radiation

- High complexity of system
- Several incompatibilities within the GSM standards
- Card-type
- Serial number
- A list of subscribed services

Additional Questions

1. Define GSM.
2. Write a note on network and switching subsystem.
3. What are GSM channels?
4. Write a note on handover procedure in GSM.
5. What is GPRS?
6. Draw the GPRS architecture.
7. What is LLC, RLC, and MAC in GPRS?
8. Define UMTS.
9. List two advantages I in third generation wireless networks.
10. What is PDP context procedure?
11. List the 3 important features of GSM security. **(JUNE 16)**
12. What are the main elements of UMTS? **(JUNE 16)**
13. Name the Teleservices provided by GSM**.(JUNE 17)**
14. Define Handoff. What are its types? **(DEC 17)**
15. List the services of GPRS**.(DEC 17)**
16. What is the frequency range of uplink and downlink in GSM network**? (MAY 18)**

UNIT III

1. Define a network and a wireless network.

A network consists of two or more computers that are linked in order to share resources (such as printers and CDs), exchange files, or allow electronic communications. The computers on a network may be linked through cables, telephone lines, radio waves, satellites, or infrared light beams.

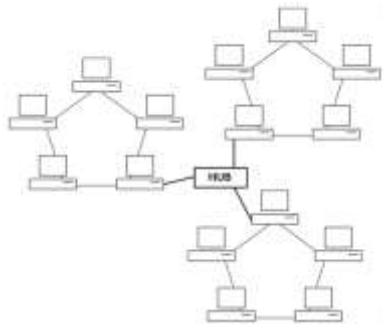
2. Mention three parameters of an efficient network.
   Performance, reliability, and security
3. Compare LAN,MAN and WAN Networks

LAN stands for Local Area Network is a group of network devices which allow the communication between connected devices. The private ownership has the control over the local area network rather than public. LAN has short propagation delay than MAN as well as WAN. It covers the smallest area such as: College, School Hospital and so on.

MAN stands for Metropolitan Area Network. It covers the largest area than LAN such as: Small towns, City etc. MAN connects 2 or a lot of computers that area unit apart however resides within the same or completely different cities. MAN is expensive and should or might not be owned by one organization.

WAN stands for Wide Area Network. It covers the large area than LAN as well as MAN such as: Country/Continent etc. WAN is expensive and should or might not be owned by one organization. PSTN or Satellite medium are used for wide area network.

4. Draw an internetwork arrangement.



Three applications of WLAN. Internet Access. ...
Voice over Wireless. ...
Inventory Control. ...
Health Care. ...
Education. ...
Real Estate. .

5. What is 802.11 standard?

802.11 and *802.11x* refers to a family of specifications developed by the IEEE for **w***ireless* **LAN** (WLAN) technology. 802.11 specifies an over-the-air interface between a wireless client and a base station or between two wireless clients

6. What are the advantages of 802.11 standards?

802.11a uses 5GHz frequency band which is less crowded and hence has relatively smaller interference problem. 802.11a supports up to 54Mbps of bandwidth, which is much faster than the 11Mbps bandwidth provided by 802.11b standard devices 802.11a offers as many as 12 non-overlapping channels.

7. What is known as DLC? Data Link Control

DLC. (1) Short for Data Link Control, the second lowest layer in the OSI Reference Model. Every network interface card (NIC) has a DLC address or DLC identifier (DLCI) that uniquely identifies the node on the network. Some network protocols, such as Ethernet and Token-Ring use the DLC addresses exclusively.

8. What is known as LLC? Logical Link Control

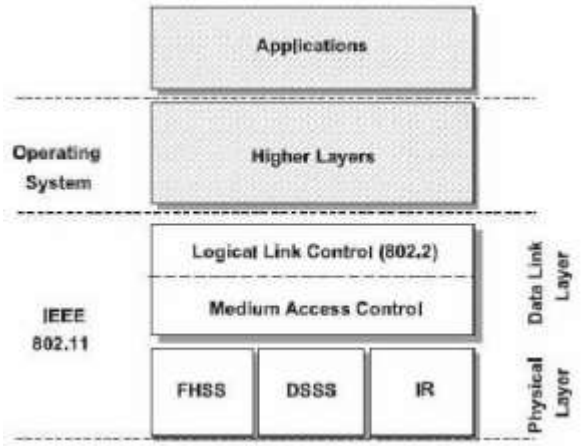
In the IEEE 802 reference model of computer networking, the logical link control (LLC) data communication protocol layer is the upper sublayer of the data link layer (layer 2) of the seven-layer OSI model. The LLC sublayer acts as an interface between the media access control (MAC) sublayer and the network layer.

9. Define medium access control.

Media access control (MAC) is a sublayer of the data link layer (DLL) in the seven-layer OSI network reference model. MAC is responsible for the transmission of data packets to and from the network-interface card, and to and from another remotely shared channel.

10. Differentiate DLC and PHY under MAC.

11. Draw the protocol stack of IEEE 802.11 standard.



Additional Questions

1. Command on authentication facilities in 802.11 standard.
2. What are the functional layers available in 802.11 standard?
3. What are the services defined by IEEE 802.11 standard.

4. Write short note on Bluetooth technology.

Bluetooth is a wireless LAN technology designed to connect devices of different functions such as telephones, notebooks, computers, cameras, printers, coffee maker and so on.

5. Write a short note on Dipole antenna.
6. What are the antenna parameters that can affect the efficiency of Bluetoothtechnology?
7. Define “Return Loss” and “Polarization”.
8. Write a short note on Ad hoc network.
9. What is known as piconet connections?
10. Define scatternet.
11. Define a point to multipoint connection setup.
12. Give an example for scatternet diagram.
13. What is known as Bluetooth device address.
14. What are the four possible modes a Bluetooth can select in a piconet arrangement?
15. Define active and hard modes.
16. What is known as a sniff modes?
17. List some of advantages of Bluetooth technology.
18. Draw the Bluetooth protocol stack architecture.
19. What are the different topologies of Bluetooth modules?
20. What are the three classes of Bluetooth protocol?
21. Compare Bluetooth and wireless technology.
22. What are the two different profiles supported by Bluetooth 1.1 version?
23. List the 9 different applications profiles of Bluetooth 1.1 version.
24. What are the advantages of a Bluetooth radio?
25. Explain Wi-Fi in Detail.
26. Write a short note on WiMAX.

UNIT IV

1. Define mobile IP.

Mobile IP (or MIP) is an Internet Engineering Task Force (IETF) standard communications protocol that is designed to allow mobile device users to move from one network to another while maintaining a permanent IP address.

2. What is Tunneling process?

The encapsulation process creates a logical construct called a tunnel between the device that encapsulates and the one that decapsulates.

3. Define COA or how the mobile nodes identify the current location? Explain.

The mobile nodes identify the current location with the COA. A care-of address (usually referred to as CoA) is a temporary IP address for a mobile device. This allows a home agent to forward messages to the mobile device. A separate address is required because the IP address of the device that is used as host identification is topologically incorrect - it does not match the

network of attachment. The care-of address splits the dual nature of an IP address, that is, its use is to identify the host and the location within the global IP network.

4. State the features of mobile IP. / What are the requirements of mobile IP?

- Transparency
- Compatibility
- Security
- Efficiency and Scalability

5. State the key mechanism used in Mobile IP.

- Discovering the care-of-address
- Registering the care-of-address
- Tunneling to the care-of-address

6. Mention the different entities in a mobile IP.

- Mobile Node
- Correspondent Node
- Home Network
- Foreign Network
- Foreign Agent
- Home Agent
- Care-Of address
- Foreign agent COA
- Co-located COA

7. Define Mobile node.

A mobile node is an end-system or router that can change its point of attachment to the Internet using mobile IP. The MN keeps its IP address and can continuously with any other system in the Internet as long as link layer connectivity is given.

8. Explain Cellular IP.

Cellular IP provides local handovers without renewed registration by installing a single cellular IP gateway for each domain, which acts to the outside world as a foreign agent.

9. What do you mean by mobility binding?

The Mobile Node sends its registration request to the Home Agent. The HA now sets up a mobility binding containing the mobile node's home IP address and the current COA.

10. Define COA.

The COA (care of address) defines the current location of the MN from an IP point of view. All IP packets sent to the MN are delivered to the COA, not directly to the IP address of the MN. Packet delivery toward the MN is done using the tunnel. DHCP is a good candidate for supporting the acquisition of Care Of Addresses.

11. Define a tunnel.

A tunnel establishes a virtual pipe for data packets between a tunnel entry and a tunnel endpoint. Packets entering a tunnel are forwarded inside the tunnel and leave the tunnel unchanged.

12. What is encapsulation?

Encapsulation is the mechanism of taking a packet consisting of packet header and data putting it into the data part of a new packet.

13. What is decapsulation?

The reverse operation, taking a packet out of the data part of another packet, is called decapsulation.

14. Define Snooping TCP.

A protocol in which an agent buffers the packets from the fixed connection layer for transmission to the mobile node on a wireless transceiver; the agent also buffers the packets on the wireless transceiver from the node for transmitting to a layer at the fixed line. The agent snoops at the transmission and reception in place of acknowledgement-or-timeout-based TCP method in the mobile part of the network.

25. What is Ad-hoc network?

The meaning for Ad-hoc is "For the purpose". The network that formed for the particular purpose without infrastructure is called as Ad-hoc network.

26. List the categories of Ad-hoc networks.

- Mobile Ad-hoc Network
- Vehicular Ad-hoc Network
- Sensor Network

27. What are the characteristics of MANET?

- Lack of fixed Infrastructure
- Dynamic topologies
- Bandwidth constrained, variable capacity link
- Energy constrained operations
- Increased vulnerability

28. What are the applications of MANET?

- Communication among portable computers
- Environmental monitoring
- Military
- Emergency applications

29. What are the design issues of MANET?

- Network size and node density
- Connectivity
- Network topology
- User traffic
- Operational environment

- Energy constraint

30. State the essential routing protocols.

- Link state protocols
- Distance vector protocols

31. List the types of communication in MANET.

- Unicast
- Multicast
- Broadcast

32. Classification of unicast routing protocols.

- Proactive protocols
- Reactive Protocols
- Hybrid protocols

33. List the popular routing protocols in MANET.

- Destination sequence distance vector routing protocol
- Dynamic source routing protocol
- Ad-hoc on-demand distance vector
- Zone routing protocol

34. Mention the classification of multicast routing protocols.

- Tree based protocol
- Mesh based protocol

35. Define VANET.

A vehicular Ad-hoc Network is a special type of MANET in which moving automobiles form the nodes of the network.

36. What are the security issues in MANET?

- Lack of physical boundary
- Low power RF transmissions
- Limited computational capabilities
- Limited power supply

37. What are the characteristics of secure ad-hoc networks?

- Availability
- Confidentiality
- Integrity
- Authentication
- Non-repudiation

38. What are the types of attack in MANET?

- Active attack
- Passive attack

39. Define Active attack. Give example.

The attack is destructive and disturbs the normal functionality of the network.
Example: wormhole, black hole

40. Define Passive attack. Give example.

The attack that monitors and steals the data exchanged the network, without disturbing the network operation.
Example: Snooping, Eavesdropping

41. Differentiate cellular and Ad-Hoc Networks:

Cellular Network	Ad-Hoc Network
Infrastructure Networks	Infrastructure less Networks.
Fixed, pre-located cell sites and base stations	No base station, and rapid deployment
Static backbone network topology	Highly dynamic network topologies
Relatively caring environment and stable connectivity	Hostile environment and irregular connectivity
Detailed planning before base station can be installed	Ad-Hoc network automatically forms and adapts to changes
High setup costs	Cost-effective
Large setup time	Less setup time

42. Compare MANET Vs VANET

MANET	VANET
MANET - Mobile Adhoc NETwork	VANET- Vehicular Adhoc NETwork
Nodes moves randomly	Nodes moves regularly
Mobility is low	Mobility is high
Reliability is medium	Reliability is high
Node lifetime depends on power source	Node lifetime depends on vehicle life time
Network topology is sluggish and slow	Network topology is frequent and fast

43. List any five attacks in MANET.

- Snooping,
- Eavesdropping
- wormhole,
- Black hole
- Monitoring

44. Write the type of network used in moving automobile.

The type of network used in moving automobile is VANET.

VANET- The Vehicular Ad-Hoc Network.

VANET, is a technology that uses moves cars as nodes in a network to create a mobile network. VANET turns every participating car into a wireless router or node, allowing cars approximately 100 to 300 meters of each other to connect and, in turn, create a network with a wide range.

45. List the Types of Communications.

- Unicast
  Message is sent to a single destination node
- Multicast
  Message is sent to a selected subset of network nodes
- Broadcast
  Broadcasting is a special case of multicasting
  Message is sent to all the nodes in the network

46. What are the functions of each node in MANET?

Forward the packet to the next hop

Before forwarding, Sender has to ensure that:

- The packet moves towards its destination
- The number of hops (path length) to destination is minimum
- Delay is minimized
- Packet loss is minimum through the path
- Path does not have a loop

47. Discuss the Proactive (table-driven) protocols.

- It is also known as table-driven routing protocols
- Each node in the routing table maintains information about routes to every other node in the network.
- F
- Tables are updates frequently due to
  - Changes in network topology
  - Node Movements
- Nodes shutting down
- Nodes can determine the best route to a destination Generates a large number of control messages to keep the routing tables up-to-date
- Generates overhead which consumes large part of available bandwidth

48. Define Reactive protocols.

- It is also called as On-demand routing protocol
- Nodes do not maintain up-to-date routing information
  New routes are discovered only when required
- Uses flooding technique to determine the route
  Flooding technique is used when the node does not have routing knowledge

<u>Additional questions</u>

1. Define mobile IP.

2. What are the goals of mobile IP?
3. Define node, home agent, and home address in mobile IP.
4. Mention the basic capabilities of mobile IP.
5. Write a note on tunneling in mobile IP.
6. What is MANET?
7. What are the advantages of MANET?
8. List any three characteristics of ad hoc networks.
9. Mention any four advantages of MANET.
10. Write two properties of MANET.
11. What are the types of MANET routing protocols?
12. What are the fundamental steps in routing?
13. Write a note on VANET.
14. List any two security issues in MANET.
15. What is traditional routing?
16. List the characteristics of MANETs**. (JUNE 16)**
17. Compare MANET Vs VANET (**JUNE 16**)
18. What is multicasting? (**DEC 16**)
19. Compare and contrast MANET Vs VANET. **(DEC 16)**
20. List the application of MANETs (**MAY 17**)
21. Distinguish proactive and reactive protocols **(MAY 17)**
22. What is the purpose of DHCP? **(MAY 18)**
23. Differentiate cellular with ad hoc network. **(MAY 18)**
24. Illustrate the use of BOOTP protocol**. (DEC 16)**

UNIT V

1. Define Mobile TCP.

A method of splitting the TCP layer into two TCP sub-layers using a mechanism that reduces window size to zero. The split is asymmetric; the window is set to zero to prevent the transmission from the TCP transport layer at the mobile node (MN) or at the fixed node when disconnection is noticed. The window opens again on getting the packet, there is no slow start by the base transceiver and it is presumed that packet loss is due to disconnection and not due to congestion or interference.

2. Write the concept of "Fast Retransmit/ Fast Recovery Transmission".

A method in which there are four or more phases of fast retransmit and fast recovery –first phase as slow start and beginning (exponential), then fast retransmit/recovery phase 1 (FRR1) on three duplicate acknowledgements, fast retransmit/fast recovery phase 2 (FRR2), and wait (Constant time out and window size).

3. Define T-TCP.

A protocol which is efficient and is used in situations where short messages are to be sent in sequence and a packet is delivered after the SYN and SYN_ACK packet exchanges and the

connection closes after the packet exchanges of FIN, FIN_ACK, and CLOSING.

4. List the features of TCP.

The main features of TCP are:

1) Transmission as data Streams
2) Buffering and retransmission
3) Session-start, data transfer, and session-finish fully acknowledged end to end. 4) In-order delivery
5) Congestion Control and avoidance

5. What is explicit notification?

A method of congestion control by explicit notification of congestion, for example, when a base transceiver at the receiver end is not able to transmit a packet to the mobile node then it sends an ESBN (explicit bad state notification) to the sender (on fixed line) at the other end.

6. What is selective retransmission?

A method in which there is an additional acknowledgement, known as selective acknowledgement; a timeout is set at transmitting end for receiving SACKs. Only the lost packet corresponding to a SACK needs to be retransmitted.

7. What is TCP header?

A header used in the TCP protocol; it consists of fields in five 32-bit words followed by words for the option fields and padding.

8. Methods of Congestion Control.

The methods of congestion control:

1) Slow start and congestion avoidance
2) Fast recovery after packet loss
3) Fast retransmit and fast recovery
4) Selective acknowledgement
5) Explicit congestion notification

9. What are Advantage and Disadvantage of Mobile TCP?

Advantages:

i. M-TCP maintains the TCP end-to-end semantic. The SH does not send any ACK itself but forwards the ACKs from the MH.

ii. If the MH is disconnected, M_TCP avoids useless retransmissions, slow starts or breaking connections by simply shrinking the sender's window to 0;

iii. Since M-TCP does not buffer data in the SH as I-TCP does, it is not necessary to forward buffers to a new SH. Lost packets will be automatically retransmitted to the new SH.

Disadvantages:

i. As the SH does not act as proxy as in I-TCP, packet loss on the wireless link due to bit errors is propagated to the sender. M-TCP assumes low bit error rates, which is not always a valid assumption.

ii. A modified TCP on the wireless link not only requires modification to the MH protocol software but also new network elements like the bandwidth manager

10. What led to the development of Indirect TCP?

TCP performs poorly together with wireless links

TCP within the fixed network cannot be changed. This led to the development of I-TCP which segments a TCP connection into a fixed part and a wireless part.

11. What is the goal of M-TCP?

The goal of M-TCP is to prevent the sender window from shrinking if bit errors or disconnection but not congestion cause current problems. It wants

- To provide overall throughput
- To lower the delay
- To maintain end-to-end semantics of TCP
- To provide a more efficient handover.

12. What do you mean by persistent mode?

Persistent mode is the state of the sender will not change no matter how long the receiver is disconnected. This means that the sender will not try to retransmit the data.

13. Write the congestion avoidance algorithm.

- TCP Tahoe and Reno
- TCP Vegas
- TCP New Reno

14. State the mechanism used in improvement of TCP Performance.

- Slow start
- Mobile TCP
- Indirect TCP
- Snooping TCP

<u>Additional questions</u>

1. Explain the goals of TCP/IP.
2. What are the advantages of TCP/IP over OSI model?
3. List few application protocols.
4. Explain the TCP/IP operation in detail.
5. What are the layers of TCP/IP?
6. What are the applications of TCP/IP?
7. Define WAP.
8. Define WAE.
9. What are the elements of WAE?
10. Write short note on “Wireless Telephony Application”.
11. Explain Wireless Session protocol in detail.

12. Explain protocol stack of WAP.
13. Write short note on WAP 2.0 version.
14. Explain the programming model of WAP.
15. What are the advantages of WAP 2.0 version?
16. Explain the advantages of TCP/IP.

BIG DATA ANALYTICS

2 MARKS WITH ANSWER

UNIT-I

1. What is big data approach? (NOV/DEC 2014)

Many It tools are available for big data projects. Organizations whose data workloads are constant and predictable are better served by traditional database whereas organizations challenged by increasing data demands will need to take advantage of Hadoop's scalable infrastructure.

2. List out the applications of big data analytics. (APR/MAY 2019)

- Marketing
- Finance
- Government
- Healthcare
- Insurance
- Retail

3. List the types of cloud environment.

- Public cloud | Private cloud

4. What is reporting?

It is the process of organizing data into informational summaries in order to monitor how different areas of a business are performing.

5. What is analysis?

It is the process of exploring data and reports in order to extract meaningful insights which can be used to better understand and improve business performance.

6. List out the cross validation technique.

- Simple cross validation
- Double cross validation

- Multicross validation

7. Write short note on MapReduce?

MapReduce provides a data parallel programming model for clusters of commodity machines. It is pioneered by google which process 20PB of data per day. MapReduce is popularized by Apache Hadoop project and used by Yahoo, Facebook, Amazon and others.

8. What is cloud computing?

Cloud computing is internet-based computing. It relies on sharing computing resources on-demand rather than having local servers or PCS and other devices. It is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort.

9. Describe the drawbacks of cloud computing?

In cloud computing, cheap nodes fail, especially when you have many of them. Mean time between failures(MTBF) for 1 node = 3 years – MTBF for 1000 nodes = 1 day and commodity network has low bandwidth.

10. List out the four major types of resampling.

- Randomized exact test
- Cross-validation
- Jackknife
- Bootstrap

11. What are the uses of statistics in data analytics?

Statistics is used to

- to estimate the complexity of a data mining problem;
- suggest which data mining techniques are most likely to be successful; and
- identify data fields that contain the most "surface information".

12. What are the factors to be considered while selecting the sample in statistics?

The sample should be

- Small enough to be manageable.
- Accessible to the sampler.
- Free of bias.

13. Name some advanced database systems.

Object-oriented databases, Object-relational databases.

14. Name some specific application oriented databases.

- Spatial databases,
- Time-series databases.
- Text databases and multimedia databases.

15. What are the basic differences between relational database and HDFS? (NOV/DEC 2014)

Here are the key differences between HDFS and relational database:

RDBMS vs. Hadoop		
	RDBMS	**Hadoop**
Data Types	RDBMS relies on the structured data and the schema of the data is always known.	Any kind of data can be stored into Hadoop i.e. Be it structured, unstructured or semi-structured.
Processing	RDBMS provides limited or no processing capabilities.	Hadoop allows us to process the data which is distributed across the cluster in a parallel fashion.
Schema on Read Vs. Write	RDBMS is based on 'schema on write' where schema validation is done before loading the data.	On the contrary, Hadoop follows the schema on read policy.
Read/Write Speed	In RDBMS, reads are fast because the schema of the data is already known.	The writes are fast in HDFS because no schema validation happens during HDFS write.
Cost	Licensed software, therefore, I have to pay for the software.	Hadoop is an open source framework. So, I don't need to pay for the software.

Best Fit Use Case	RDBMS is used for OLTP (Online Trasanctional Processing) system.	Hadoop is used for Data discovery, data analytics or OLAP system.

16. What is Hadoop and its components.

When "Big Data" emerged as a problem, Apache Hadoop evolved as a solution to it. Apache Hadoop is a framework which provides us various services or tools to store and process Big Data. It helps in analyzing Big Data and making business decisions out of it, which can't be done efficiently and effectively using traditional systems.

- ***Storage unit***– HDFS (NameNode, DataNode)
- ***Processing framework***– YARN (ResourceManager, NodeManager)

17. What are HDFS and YARN? **(APR/MAY 2014)**

HDFS (Hadoop Distributed File System) is the storage unit of Hadoop. It is responsible for storing different kinds of data as blocks in a distributed environment. It follows master and slave topology.

♣ *Tip: It is recommended to explain the HDFS components too i.e.*

- ***NameNode*:** NameNode is the master node in the distributed environment and it maintains the metadata information for the blocks of data stored in HDFS like block location, replication factors etc.
- ***DataNode*:** DataNodes are the slave nodes, which are responsible for storing data in the HDFS. NameNode manages all the DataNodes.

YARN (Yet Another Resource Negotiator) is the processing framework in Hadoop, which manages resources and provides an execution environment to the processes.

18. Write about the various Hadoop daemons and their roles in a Hadoop cluster.

Generally approach this question by first explaining the HDFS daemons i.e. NameNode, DataNode and Secondary NameNode, and then moving on to the YARN daemons i.e. ResorceManager and NodeManager, and lastly explaining the JobHistoryServer.

- **NameNode:** It is the master node which is responsible for storing the metadata of allthe files and directories. It has information about blocks, that make a file, and where those blocks are located in the cluster.
- **Datanode:** It is the slave node that contains the actualdata.
- **Secondary NameNode:** It periodically merges the changes (edit log) with the FsImage (Filesystem Image), present in the NameNode. It stores the modified FsImage into persistent storage, which can be used in case of failure of NameNode.

- **ResourceManager:** It is the central authority that manages resources and schedule applications running on top of YARN.
- **NodeManager:** It runs on slave machines, and is responsible for launching the application's containers (where applications execute their part), monitoring their resource usage (CPU, memory, disk, network) and reporting these to the ResourceManager.
- **JobHistoryServer:** It maintains information about MapReduce jobs after the Application Master terminates.

19. What are active and passive "NameNodes"?

In HA (High Availability) architecture, we have two NameNodes – Active "NameNode" and Passive "NameNode".

- Active "NameNode" is the "NameNode" which works and runs in the cluster.
- Passive "NameNode" is a standby "NameNode", which has similar data as active "NameNode".

When the active "NameNode" fails, the passive "NameNode" replaces the active "NameNode" in the cluster. Hence, the cluster is never without a "NameNode" and so it never fails.

20. Name the three modes in which Hadoop can run.

The three modes in which Hadoop can run are as follows:

1. ***Standalone (local) mode***: This is the default mode if we don't configure anything. In this mode, all the components of Hadoop, such NameNode, DataNode, ResourceManager, and NodeManager, run as a single Java process. This uses the local filesystem.
2. ***Pseudo-distributed mode***: A single-node Hadoop deployment is considered as running Hadoop system in pseudo-distributed mode. In this mode, all the Hadoop services, including both the master and the slave services, were executed on a single compute node.
3. ***Fully distributed mode***: A Hadoop deployments in which the Hadoop master and slave services run on separate nodes, are stated as fully distributed mode.

UNIT – II

1. What are the three stages of IDA process?

oData preparation oData mining and rule finding oResult validation and interpretation

2. What is linear regression? (APR/MAY 2019)

Linear regressionis an approach for modeling the relationship between a scalar dependent variable y and one or more explanatory variables (or independent variables) denoted X. The case of one explanatory variable is called simple **linear regression**.

3. Explain Bayesian Inference ?

Bayesian inference is a method of statistical **inference** in which **Bayes**' theorem is used to update the probability for a hypothesis as more evidence or information becomes available. **Bayesian inference** is an important technique in statistics, and especially in mathematical statistics.

4. What is meant by rule induction?

Rule induction is an area of machine learningin which formal rules are extracted from a set of observations. The rules extracted may represent a full scientific modelof the data, or merely represent local patternsin the data.

5. What are the two strategies in Learn-One-Rule Function.

oGeneral to specific
oSpecific to general

6. Write down the topologies of Neural Network.

- Single layer
- Multi layer
- Recurrent
- Self-organized

7. What is meant by fuzzy logic.

More than data mining tasks such as prediction, classification, etc., fuzzy models can give insight to the underlying system and an be automatically derived from system's dataset. For achieving this, the technique used is grid based rule set.

8. Write short note on fuzzy qualitative modeling. (APR/MAY 2019)

The fuzzy modeling can be interpreted as a qualitative modeling scheme by which the system behavior is qualitatively described using a natural language. A fuzzy qualitative model is a generalized fuzzy model consisting of linguistic explanations about system behavior in the framework of fuzzy logic instead of mathematical equations with numerical values or conventional logical formula with logical symbols.

8. What are the steps for Bayesian data analysis.

- Setting up the prior distribution
- Setting up the posterior distribution
- Evaluating the fit of the model

9. Write short notes on time series model.

A time series is a sequential set of data points, measured typically at successive times. It is mathematically defined as a set of vectors x(t), t=0,1,2,… where t represents the time elapsed. The Variable x9t0 is treated as a randomvariable.

10. What is Visualization?

Visualization is for depiction of data and to gain intuition about data being observed. It Assists the analysts in selecting display formats, viewer perspectives and data representation schema

11. What is clustering?

Clustering is the process of grouping the data into classes or clusters so that objects within a cluster have high similarity in comparison to one another, but are very dissimilar to objects in other clusters.

12. What are the requirements of clustering?

- Scalability
- Ability to deal with different types of attributes
- Ability to deal with noisy data

-

- Minimal requirements for domain knowledge to determine input parameters
- Constraint based clustering

- Interpretability and usability

13. State the categories of clustering methods.

Partitioning methods

Hierarchical methods

Density based methods

Grid based methods

Model based methods

14. What is linear regression?

In linear regression data are modeled using a straight line. Linear regression is the simplest form of regression. Bivariate linear regression models a random variable Y called response variable as a linear function of another random variable X, called a predictor variable.

$$Y = a + b X$$

15. State the types of linear model and state its use.

Generalized linear model represent the theoretical foundation on which linear regression can be applied to the modeling of categorical response variables. The types of generalized linear model are

Logistic regression

Poisson regression

16. Write the preprocessing steps that may be applied to the data for classification and prediction.

a. Data Cleaning
b. Relevance Analysis
c. Data Transformation

17. Define Data Classification.

It is a two-step process. In the first step, a model is built describing a pre-determined set of data classes or concepts. The model is constructed by analyzing database tuples described by attributes. In the second step the model is used for classification.

18. What is a "decision tree"?

It is a flow-chart like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and leaf nodes represent classes or cass distributions. Decision tree is a predictive model. Each branch of the tree is a classification question and leaves of the tree are partition of the dataset with their classification.

19. Where are decision trees mainly used?

Used for exploration of dataset and business problems Data preprocessing for other predictive analysis Statisticians use decision trees for exploratory analysis

20. What is Association rule?

Association rule finds interesting association or correlation relationships among a large set of data items, which is used for decision-making processes. Association rules analyzes buying patterns that are frequently associated or purchased together.

21. Define support.

Support is the ratio of the number of transactions that include all items in the antecedent and consequent parts of the rule to the total number of transactions. Support is an association rule interestingness measure.

22. Define Confidence.

Confidence is the ratio of the number of transactions that include all items in the consequent as well as antecedent to the number of transactions that include all items in antecedent. Confidence

is an association rule interestingness measure.

UNIT - III

1. What is data stream model?

A data stream is a real-time, continuous and ordered sequence of items. It is not possible to control the order in which the items arrive, nor it is feasible to locally store a stream in its entirety in any memory device.

2. Define Data Stream Mining.

Data Stream Mining is the process of extracting useful knowledge from continuous, rapid data streams. Many traditional data mining algorithms can be recast to work with larger datasets, but they cannot address the problem of a continuous supply of data.

3. Write short note about sensor networks.

Sensor networks are a huge source of data occurring in streams. They are used in numerous situations that require constant monitoring of several variables, based on which important decisions are made. in many cases, alerts and alarms may be generated as a response to the information received from a series of sensors.

4. what is meant by one-time queries?

One-Time queries are queries that are evaluated once over a point-in-time snapshot of the data set, with the answer returned to the user.

Eg: A stock price checker may alert the user when a stock price crosses a particular price point.

5. Define biased reservoir sampling.

Biased reservoir sampling is defined as bias function to regulate the sampling from the stream. The bias gives a higher probability of selecting data points from recent parts of the stream as compared to distant past.

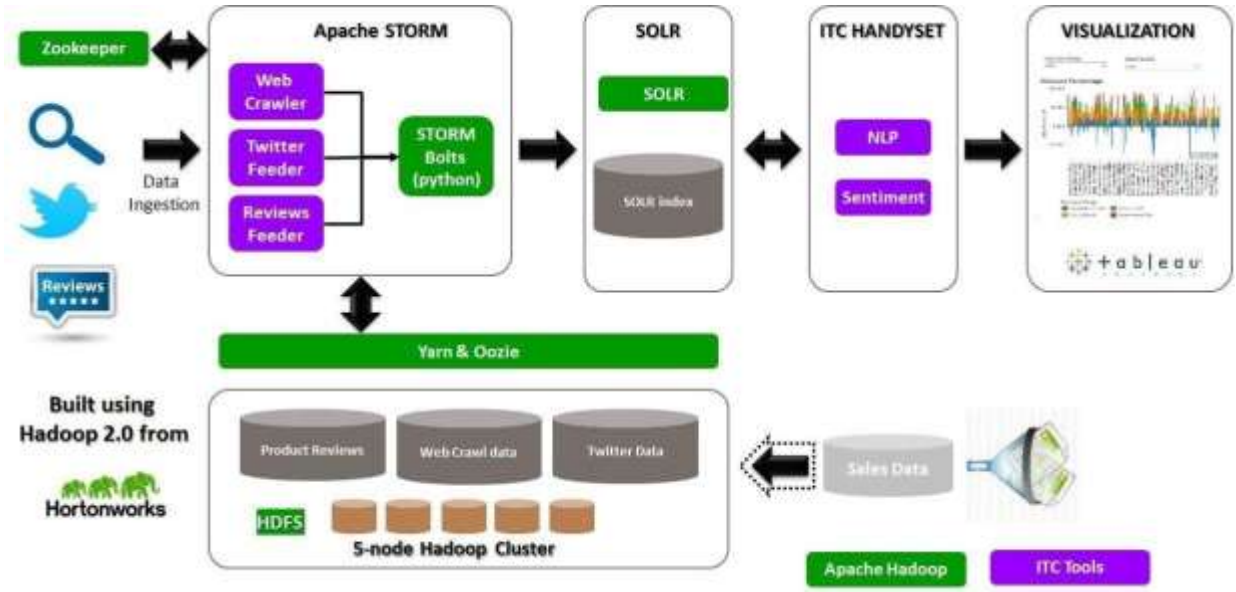
6. What is Bloom Filter? (APR/MAY 2019)

A Bloom Filter is a space-efficient probabilistic data structure, conceived by Burton Howard Bloom in 1970, that is used to test whether an element is a member of set. False Positive matches are possible but false negative are not, thus a Bloom filter has a 100% recall rate.

7. List out the applications of RTAP.

- Financial services oGovernment oE-Commerce sites

8. Draw a High-Level architecture for RADAR.



9. What are the three layers of Lambda architecture.

- Batch Layer- for batch processing of all data.
- Speed Layer- for real-time processing of streaming data.
  oServing Layer- for responding to queries.

10. What is RTSA?

Real-Time Sentiment analysis (also known as opinion mining) refers to the use of natural language processing text analysis and computational linguistics to identify and extract subjective information in source materials.

11. Define Spatial Visualization.

Spatial visualization depicts actual members of the population in their feature space

14. What is Data Generalization?

It is process that abstracts a large set of task-relevant data in a database from a relatively low conceptual to higher conceptual levels 2 approaches for Generalization

a. Data cube approach
b. Attribute-oriented induction approach

15. Define Attribute Oriented Induction. (APR/MAY 2019)

These method collets the task-relevant data using a relational database query and then perform generalization based on the examination in the relevant set of data.

16. What is bootstrap?

An interpretation of the jack knife is that the construction of pseudo value is based on Repeatedly and systematically sampling with out replacement from the data at hand. This lead to generalized concept to repeated sampling with replacement called bootstrap.

17. Explain the view of statistical approach.

Statistical method is interested in interpreting the model. It may sacrifice some performance to be able to extract meaning from the model structure. If accuracy is acceptable then the reason that a model can be decomposed in to revealing parts is often more useful than a 'black box' system, especially during early stages of investigation and design cycle.

18. Define Deterministic models.

Deterministic models, which takes no account of random variables, but gives precise, fixed reproducible output.

19. Define Systems and Models.

System is a collection of interrelated objects and Model is a description of a system. Models are

-

abstract, and conceptually simple.

20. How do you choose the best model?

All things being equal, the smallest model that explains the observations and fits the objectives that should be accepted. In reality, the smallest means the model should optimizes a certain scoring function (e.g. Least nodes, most robust, least assumptions)

Unit-IV

1.What is Association Rule Mining? (APR/MAY 2019)

The Association Rule Mining is main purpose to discovering frequent itemsets from a large dataset is to discover a set of if-then rules called Association rules. The form of an association rules is $I \rightarrow j$, where I is a set of items(products) and j is a particular item.

2. List any two algorithms for Finding Frequent Itemset.

oApriori Algorithm oFP-
Growth Algorithm
oSON algorithm oPCY
algorithm

3. What is meant by curse of dimensionality?

Points in high-dimensional Euclidean spaces, as well as points in non-Euclidean spaces often behave unintuitively. Two unexpected properties of these spaces are that the random points are almost always at about the same distance, and random vectors are almost always orthogonal.

4. Write an algorithm of Park-Chen-Yu.

```
FOR(each basket):
    FOR(each item in basket):
        add 1 to item's count;
    FOR(each pair of items):
```

{hash the pair to a bucket;

add 1 to the count for that bucket:}

5. Define Toivonen's Algorithm

Toivonen's algorithm makes only one full pass over the database. The algorithm thus produces exact association rules in one full pass over the database. The algorithm will give neither false negatives nor positives, but there is a small yet non-zero probability that it will fail to produce any answer at all. Toivonen's algorithm begins by selecting a small sample of the input dataset and finding from it the candidate frequent itemsets.

6. List out some applications of clustering.

- Collaborative filtering
- Customer segmentation
- Data summarization
- Dynamic trend detection
- Multimedia data analysis
- Biological data analysis
- Social network analysis

7. What are the types of Hierarchical Clustering Methods. (APR/MAY 2019)

- Single-link clustering
- Complete-link clustering
- Average-link clustering
- Centroid link clustering

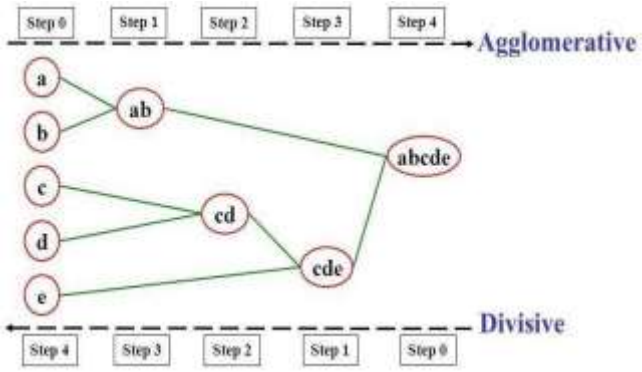
8. Define CLIQUE

CLIQUE is a subspace clustering algorithm that automatically finds subspaces with highdensity clustering in high dimensional attribute spaces. CLIQUE is a simple grid-based method for finding density-based clusters in subspaces. The procedure for this grid-baased clustering is relatively simple.

9. What is meant by k-means algorithm?

The family of algorithms is of the point-assignment type and assumes a Euclidean space. It is assumed that there are exactly *k* clusters for some known *k*. After picking *k* initial cluster centroids, the points are considered one at a time and assigned to the closest centroid.

10. Draw the diagram for Hierarchical Clustering. (APR/MAY 2019)



11. Explain End User Data Access tool.

End User Data Access tool is a client of the data warehouse. In a relational data warehouse, such a client maintains a session with the presentation server, sending a stream of separate SQL requests to the server. Eventually the end user data access tool is done with the SQL session and turns around to present a screen of data or a report, a graph, or some other higher form of analysis to the user. An end user data access tool can be as simple as an Ad Hoc query tool or can be complex as a sophisticated data mining or modelingapplication.

12. Explain Ad Hoc query tool.

A specific kind of end user data access tool that invites the user to form their own queries by directly manipulating relational tables and their joins. Ad Hoc query tools, as powerful as they are, can only be effectively used and understood by about 10% of all the potential end users of a data warehouse.

13. Name some of the data extraction applications.

Data extraction for Biomedical and DNA data analysis

Data mining for Financial data analysis

Data mining for the Retail industry

Data mining for the Telecommunication industry

14. Name some of the data mining applications.

- Data mining for Biomedical and DNA data analysis
- Data mining for Financial data analysis
- Data mining for the Retail industry
- Data mining for the Telecommunication industry

15. What is the difference between "supervised" and unsupervised" learning scheme?

In data mining during classification the class label of each training sample is provided, this type of training is called supervised learning (i.e.) the learning of the model is supervised in that it is told to which class each training sample belongs. Eg. Classification In unsupervised learning the class label of each training sample is not known and the member or set of classes to be learned may not be known in advance. Eg.Clustering

16. Explain the various OLAP operations.

a) Roll-up: The roll-up operation performs aggregation on a data cube, either by Climbing up a concept hierarchy for a dimension.

b) Drill-down: It is the reverse of roll-up. It navigates from less detailed data to more Detailed data.

c) Slice: Performs a selection on one dimension of the given cube, resulting in a Sub cube.

17. Why is data quality so important in a data warehouse environment?

Data quality is important in a data warehouse environment to facilitate decision-making. In order to support decision-making, the stored data should provide information from a historical perspective and in a summarized manner.

18. How can data visualization help in decision-making?

Data visualization helps the analyst gain intuition about the data being observed. Visualization applications frequently assist the analyst in selecting display formats and in viewer perspective.

19. How is association rules mined from large databases?

Association rule mining is a two-step process.

Find all frequent itemsets.

Generate strong association rules from the frequent itemsets.

20. What is the classification of association rules based on various criteria?

1. Based on the types of values handled in the rule.
   Boolean Association rule.

   Quantitative Association rule.

2. Based on the dimensions of data involved in the rule.
   a. Single Dimensional Association rule.
   b. Multi Dimensional Association rule.
3. Based on the levels of abstractions involved in the rule.
   Single level Association rule. Multi level
   Association rule.

4. Based on various extensions to association mining.
   Maxpatterns.

   Frequent closed itemsets.

UNIT-V

1. What are the main goals of Hadoop? (APR/MAY 2019)

- Saclable oFault tolerance oEconomical
- Handle hardware failures.

I

2. What is hive?

Hive provides a warehouse structure for other Hadoop input sources and SQL-Like access for data in HDFS. Hive's query language, HiveQL, compiles to MapReduce and also allows user-defined functions(UDFS).

3. What are the responsibilities of MapReduce Framework?

- Provides overall coordination of execution.
- Selects nodes for running mappers.
- Starts and monitors mapper's execution. oSorts and shuffles output of mappers. oChooses locations for reducer's execution.
- Delivers the output of mapper to reducers node. oStarts and monitors reducers's execution.

4. What is a Key-Value store?

The key-value store uses a key to access a value. The key-value store has a schema-less format. The key can be artificially generated or auto-generated while the value can be string, JSON, BLOB, etc. the key-value uses a hash table with a unique key and a pointer to a particular item of data.

5. What is visualization? What are the three major goals in visualization. (APR/MAY 2019)

Visual Visualization is the presentation or communication of data using interactive interfaces. It has three major goals:

- Communicating/presenting the analysis results efficiently and effectively.
- As a tool for confirmatory analysis that is to examine the hypothesis, analyze and confirm.
- Exploratory data analysis as an interactive and mostly undirected search for finding structures and trends.

6. What is sharding?

Horizontal partitioning of a large database leads to partitioning of rows of the database. Each partition forms part of a shard, meaning small part of the whole. Each part can be located on a separate database server or any physical location.

7. Write about data import in R language (get solved code examples for hands-on experience)?

R Commander is used to import data in R language. To start the R commander GUI, the user must type in the command Rcmdr into the console. There are 3 different ways in which data can be imported in R language-

- Users can select the data set in the dialog box or enter the name of the data set (if they know).

- Data can also be entered directly using the editor of R Commander via Data->New Data Set. However, this works well when the data set is not too large.

- Data can also be imported from a URL or from a plain text file (ASCII), from any other statistical package or from the clipboard.

8) Two vectors X and Y are defined as follows – X <- c(3, 2, 4) and Y <- c(1, 2). What will be output of vector Z that is defined as Z <- X*Y.

In R language when the vectors have different lengths, the multiplication begins with the smaller vector and continues till all the elements in the larger vector have been multiplied.

The output of the above code will be –

Z <- (3, 4, 4)

9) How missing values and impossible values are represented in R language?

NaN (Not a Number) is used to represent impossible values whereas NA (Not Available) is used to represent missing values. The best way to answer this question would be to mention that deleting missing values is not a good idea because the probable cause for missing value could be some problem with data collection or programming or the query. It is good to find the root cause of the missing values and then take necessary steps handle them.

10) R language has several packages for solving a particular problem. How do you make a decision on which one is the best to use?

CRAN package ecosystem has more than 6000 packages. The best way for beginners to answer this question is to mention that they would look for a package that follows good software

development principles. The next thing would be to look for user reviews and find out if other data scientists or analysts have been able to solve a similar problem.

11. What is Hive Metastore? (APR/MAY 2019)

It is the central repository of Apache Hive metadata. It stores metadata for Hive tables (like their schema and location) and partitions in a relational database. It also provides client access to this information with the help of metastore service API

12. What is the default database provided by Apache Hive?

Hive offers an embedded Derby database instance backed by the local disk for the metastore, by default. To this concept what we call embedded metastore configuration.

12. What are the different HIVE operators?

- Arithmetic Operators
- Relational Operators
- Logical Operators
- String Operators
- Operators on Complex Types

14. What is bucketing in HIVE?

For decomposing table data sets into more manageable parts, Apache Hive offers another technique. That technique is called as a Bucketing in Hive.

In Hive Tables or partition are subdivided into buckets based on the hash function of a column in the table to give extra structure to the data that may be used for more efficient queries.

15. What is Internal and External Table in hive?

Internal Table (Managed table): Managed table is also Known as Internal table. This is the default table in Hive. When user create a table in Hive without specifying it as external, by default we will get a Managed table.

If we create a table as a managed table, the table will be created in a specific location in HDFS.

By default, the table data will be created in /usr/hive/warehouse directory of HDFS.

External Tables: External table is mostly created for external use as when the data is used outside Hive. Whenever we want to delete the table's metadata and we want to keep the table's data as it is, we use External table. External table only deletes the schema of the table.

16. Name Key Components of HIVE architecture.

- User Interface
- Compiler
- Metastore
- Driver
- Execute Engine

17. What is SerDe in Apache hive?

A SerDe is a short name for a Serializer Deserializer. Hive uses SerDe to read and write data from tables. An important concept behind Hive is that it DOES NOT own the Hadoop File System format that data is stored in. Users are able to write files to HDFS with whatever tools/mechanism takes their fancy("CREATE EXTERNAL TABLE" or "LOAD DATA INPATH,") and use Hive to correctly "parse" that file format in a way that can be used by Hive. A SerDe is a powerful (and customizable) mechanism that Hive uses to "parse" data stored in HDFS to be used by Hive.

18. Mention some HIVE DDLCommands.

1. Create Database Statement
2. Hive Show Database
3. Drop database
4. Creating Hive Tables
5. Browse the table
6. Altering and Dropping Tables
7. Hive Select Data from Table
8. Hive Load Data

19. What is HCatalog?

HCatalog is a table and storage management layer for Hadoop that enables users with different data processing tools — Pig, MapReduce — to more easily read and write data on the grid.

Hcatalog can be used to share data structures with external systems. Hcatalog provides access to hive metastore to users of other tools on Hadoop so that they can read and write data to hive data warehouse.

20. Why HIVE does not store meta data information in HDFS.

Hive stores metadata information in the metastore using RDBMS instead of HDFS. The main reason for choosing RDBMS is to achieve low latency because HDFS read/write operations are time consuming processes.

21. What are the HIVE data types? (NOV/DEC 2015)

The primitive data types supported by Hive are listed below:

Numeric Types. **TINYINT** (1-byte signed **integer**, from -128 to 127) **SMALLINT** (2-byte signed **integer**, from -32,768 to 32,767) ...
Date/Time Types.
TIMESTAMP. DATE. String
Types. STRING. **VARCHAR**.
...
Misc Types. BOOLEAN. ...
Complex Types. arrays: ARRAY<data_type>

COMPUTER GRAPHICS AND MULTIMEDIA

UNIT - I - ILLUMINATION AND COLOR MODELS
PART A

1. What are subtractive colors?

RGB model is an additive system, the Cyan-Magenta-Yellow (CMY) model is a subtractive color
model. In a subtractive model, the more that an element is added, the more that it subtracts from white. So, if none of these are present the result is white, and when all are fully present the result is black.

2. What do you mean by shading of objects?

A shading model dictates how light is scattered or reflected from a surface. The shading models described here focuses on achromatic light. Achromatic light has brightness and no color; it is a shade of gray so it is described by a single value its intensity. A shading model uses two types of light source to illuminate the objects in a scene : point light sources and ambient light.

3. What is texture?

The realism of an image is greatly enhanced by adding surface texture to various faces of a mesh object. The basic technique begins with some texture function, **texture(s,t)** in **texture space** , which has two parameters s and t. The function texture(s,t) produces a color or intensity value for each value of s and t between 0(dark)and 1(light).

4. What are the types of reflection of incident light?

There are two different types of reflection of incident light

- Diffuse scattering.
- Specular reflections.

5. Define rendering.

Rendering is the process of generating an image from a model (or models in what collectively could be called a *scene*file), by means of computer programs. Also, the results of such a model can be called a rendering.

6. Differentiate flat and smooth shading.

The main distinction is between a shading method that accentuates the individual polygons (flat shading) and a method that blends the faces to de-emphasize the edges between them (smooth shading).

7. Define shading.

Shading is a process used in drawing for depicting levels of darkness on paper by applying media more densely or with a darker shade for darker areas, and less densely or with a lighter shade for lighter areas.

8. What is a shadow?

Shadows make an image more realistic. The way one object casts a shadow on another object gives important visual clues as to how the two objects are positioned with respect to each other. Shadows conveys lot of information as such, you are getting a second look at the object from the view point of the light source.

9. What are the two common sources of textures?

Shadows as Texture. Creating shadows with the use of a shadow buffer.

10. Define intensity of light.

Intensity is the radiant energy emitted per unit time, per unit solid angle, and per unit projected area of source.

11. What is hue?
The perceived light has a dominant frequency (or dominant wavelength). The dominant frequency is also called as hue or simply as color.

12. What is purity of light?
Purity describes how washed out or how “pure” the c olor of the light appears. pastels and pale colors are described as less pure.

13. Define the term chromacity.
The term chromacity is used to refer collectively to the two properties describing color characteristics: purity and dominant frequency.

14. Define complementary colors.
If the two color sources combine to produce white light, they are referred to as 'complementary colors. Examples of complementary color pairs are red and cyan, green and magenta, and blue and yellow.

15. Define primary colors.
The two or three colors used to produce other colors in a color model are referred to as primary colors.

16. State the use of chromaticity diagram.
Comparing color gamuts for different sets of primaries. Identifying complementary colors. Determining dominant wavelength and purity of a given color.

17. Define Color model.
A Color model is a method for explaining the properties or behavior of color within some particular context.

18. What are the uses of chromaticity diagram?
The chromaticity diagram is useful for the following:
• Comparing color gamuts for different sets of primaries.
• Identifying complementary colors.
• Determining dominant wavelength and purity of a given color.

19. Give the transformation matrix for conversion of RGB to YIQ.

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.144 \\ 0.596 & 10.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

20. What is HSV model?
The HSV(Hue,Saturation,Value) model is a color model which uses color descriptions that have a more intuitive appeal to a user. To give a color specification, a user selects a spectral color and the amounts of white and black that are to be added to obtain different shades, tint, and tones.

21. What for CMY color model used?
A color model defined with the primary colors cyan, magenta, and yellow is useful for describing color output to hard-copy devices.

22. What are the parameters in the HLS color model?

Hue, Lightness and Saturation.

23. What is Output Primitive?

Basic geometric structures that describe a scene are referred to as Output Primitives. Points and straight lines segments are the simplest geometric components of pictures. Additional output primitives that can be used to construct a picture include circles and other conic sections, quadric surfaces, spline curves and surfaces, polygon color areas, and character strings.

24. What is DDA?

The Digital Differential Analyzer is a scan-conversion line algorithm based on calculating either difference in y-coordinate (dy) or difference in x-coordinate. We sample the line at unit intervals in one coordinate and determine corresponding integer values nearest the line path for the other coordinate.

25. What are the disadvantages of DDA algorithm?

- Round-off error in successive additions of the floating-point increment can cause the calculated pixel positions to drift away from the true line path for long line segments.
- Rounding operations and floating-point arithmetic in procedure are still time consuming.

26. What is meant by aliasing?

The distortion of information due to low frequency sampling (Under sampling) is called aliasing. We can improve the appearance of displayed raster lines by applying antialiasing methods that compensate for the under sampling process.

PART – B & C

1. Discuss about the properties of light.
2. Describe about the Bresenham's ellipse drawing algorithm.
3. Explain the basic concepts of Midpoint ellipse algorithm. Apply the decision parameter for the algorithm and write down the algorithm steps.
4. Using midpoint circle algorithm calculate the pixel that will be put ON for an origin center at (4,5) of a circle with radius 4.
5. Consider the line from (0,0) to (-8,-4).Use general Bresenham's line algorithm to rasterize this line. Evaluate and tabulate all the steps involved
6. Depict and discuss about the following

- RGB Colour Model
- YIQ Colour Model
- CMY Colour Model
- HSV Colour Model
- HLS Colour Model

UNIT - II - TWO-DIMENSIONAL GRAPHICS
PART A

1. Define Translation.

A translation is applied to an object by repositioning it along a straight line path from one coordinate location to another. We translate a two-dimensional point by adding translation distances, tx and ty, to original coordinate position (x, y) to move the point to a new position (x', y'). x' = x + tx, y' = y + ty. The translation distance pair (tx, ty) is called a translation vector or shift vector.

2. Define Rotation.

A 2-D rotation is applied to an object by repositioning it along a circular path in the xy plane.

3. Define Scaling.

A scaling transformation alters the size of an object. This operation can be carried out for polygons by multiplying the coordinate values (x,y) of each vertex by scaling factors sx and sy to produce the transformed coordinates (x', y'). x' = x. sx, y' = y. sy.

4. Define Reflection.

A Reflection is a transformation that produces a mirror image of an object. The mirror image for a 2D reflection is generated relative to an axis of reflection by rotating the object 180 degree about the reflection axis.

5. Define Shear.

A transformation that distorts the shape of an object such that the transformed shape appears as if the object were composed of internal layers that had been caused to slide over each other is called a shear.

6. What is composite transformation?

A number of transformations or sequence of transformations can be combined into single one called as composition. The resulting matrix is called as composite matrix. The process of combining is called as concatenation.

7. Define Clipping.

Any procedure that identifies those portions of a picture that are either inside or outside of a specified region of space is referred to as a clipping algorithm or simply clipping. The region against which an object is clipped is called a clip window.

8. Define clipping and types of clipping.

Clipping is the method of cutting a graphics display to neatly fit a predefined graphics region or the view port.

- Point clipping
- Line clipping
- Area clipping
- Curve clipping
- Text clipping

9. What is the need of homogeneous coordinates?

To perform more than one transformation at a time, use homogeneous coordinates or matrixes.

They reduce unwanted calculations intermediate steps saves time and memory and produce a sequence of transformations.

10. What is fixed point scaling?

The location of a scaled object can be controlled by a position called the fixed point that is to remain unchanged after the scaling transformation.

11. Define Affine transformation?

A coordinate transformation of the form X= axxx +axyy+bx, y ‟ayxx+ayy y+by is called a two dimensional affine transformation. Each of the transformed coordinates x „and y „is a linear function of the original coordinates x and y, and parameters aij and bk are constants determined by the transformation type.

12. List out the various Text clipping.

- All-or-none string clipping -if all of the string is inside a clip window, keep it otherwise discards.
- All-or-none character clipping – discard only those characters that are not completely inside

the window. Any character that either overlaps or is outside a window boundary is clipped.

13. How will you clip a point?

Assuming that the clip window is a rectangle in standard position, we save a point P=(x,y) for display if the following inequalities are satisfied:

xwmin ≤ x≤ xwmax ywmin ≤ y≤ ywmax

where the edges of the clip window (xwmin ,xwmax, ywmin, ywmax) can be either the world coordinate window boundaries or viewport boundaries. If any one of these inequalities is not satisfied, the points are clipped (not saved for display).

14. Define Cohen–Sutherland algorithm

The Cohen–Sutherland algorithm is a computer-graphics algorithm used for line clipping. The algorithm divides a two-dimensional space into 9 regions and then efficiently determines the lines and portions of lines that are visible in the central region of interest (the viewport).

15. Define Nicholl-Lee-Nicholl Clipping.

The Nicholl–Lee–Nicholl algorithm is a fast line-clipping algorithm that reduces the chances of clipping a single line segment multiple times, as may happen in the Cohen–Sutherland algorithm. The clipping window is divided into a number of different areas, depending on the position of the initial point of the line to be clipped.

16. Define Liang-Barsky Clipping.

The Liang–Barsky algorithm uses the parametric equation of a line and inequalities describing the range of the clipping box to determine the intersections between the line and the clipping box. With these intersections it knows which portion of the line should be drawn. This algorithm is significantly more efficient than Cohen–Sutherland, but Cohen–Sutherland does trivial accepts and rejects much faster, so it should be considered instead if most of the lines you need to clip would be completely in or out of the clip window.

17. Define viewing transformation.

The mapping of a part of world coordinate scene to device coordinates are called viewing transformation. Two dimensional viewing transformations is simply referred to as window to viewport transformation or the windowing transformation.

18. Define Window.

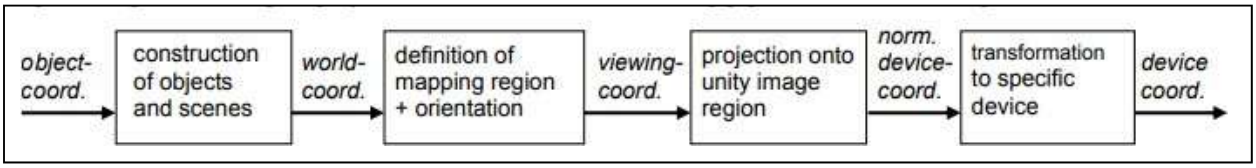
A world-coordinate area selected for display is called a window.

19. Define view port.

An area on a display device to which a window is mapped is called a view port.

20. Outline viewing pipeline

The term Viewing Pipeline describes a series of transformations, which are passed by geometry data to end up as image data being displayed on a device. The 2D viewing pipeline describes this process for 2D data:



PART – B & C

1. Describe 2D geometric transformations.
2. Show that two successive reflections about the coordinate axes is equivalent to a single rotation about
the coordinate origin
3. Illustrate with example the available two dimensional geometric transformations.
4. Explain the brief notes on pivot point rotation of an object.
5. Give a brief note on two dimensional viewing transformation pipeline.
6. Show the different types of clipping operations with neat diagram.
7. Explain Cohen-Sutherland line clipping with example.
8. Examine the Sutherland Hodgeman polygon clipping algorithm with example.
9. Explain about Composite transformation in general and Explain the following with matrix representation:

- Two Successive Translation..
- Two Successive Rotations.
- Two Successive Scaling.
- General Pivot Point Rotation.
- General Fixed Point Scaling.

UNIT - III - THREE-DIMENSIONAL GRAPHICS
PART A

1. What are the various representation schemes used in three dimensional objects?

Boundary representation (B-res) – describe the 3 dimensional object as a set of surfaces that separate the object interior from the environment. Space-portioning representation – describe interior properties, by partitioning the spatial region containing an object into a set of small, no overlapping, contiguous solids.

2. What is Boundary representation?

It describes a 3D object as a set of surfaces that separate the object interior from the environment. e.g. polygon facets and spline patches.

3. What space-partitioning representation?

This is used to describe interior properties, by partitioning the spatial region containing an object in to a set of small, non-overlapping, contiguous solids. e.g.octree.

4. What is parallel projection?

In a parallel projection, coordinate positions are transformed to the view plane along parallel lines.

5. What is Perspective projection?

For a perspective projection object positions are transformed to the view plane along lines that converge to a point called the projection reference point.

6. Define Projection.

The process of displaying 3D into a 2D display unit is known as projection. The projection transforms 3D objects into a 2D projection plane. The process of converting the description of objects from world coordinates to viewing coordinates is known as projection.

7. What is Polygon mesh?

Polygon mesh is a method to represent the polygon, when the object surfaces are tiled, it is more convenient tospecify the surface facets with a mesh function. The various meshes are

(i) Triangle strip – (n-2) connected triangles
(ii) Quadrilateral mesh – generates (n-1)(m-1) Quadrilateral

8. Define B-Spline curve.

A B-Spline curve is a set of piecewise(usually cubic) polynomial segments that pass close to a set of control points. However the curve does not pass through these control points, it only passes close to them.

9. What is a spline?

To produce a smooth curve through a designed set of points, a flexible strip called spline is used. Such a spline curve can be mathematically described with a piecewise cubic polynomial function whose first and second derivatives are continuous across various curve section.

10. What is the use of control points?

Spline curve can be specified by giving a set of coordinate positions called control points, which indicates the general shape of the curve, can specify spline curve.

11. What are the different ways of specifying spline curve?

- Using a set of boundary conditions that are imposed on the spline.
- Using the state matrix that characteristics the spline
- Using a set of blending functions that calculate the positions along the curve path by specifying combination of geometric constraints on the curve.

12. What are the important properties of Bezier Curve?.

It needs only four control points

• It always passes through the first and last control points

• The curve lies entirely within the convex half formed by four control points.

13. Differentiate between interpolation spline and approximation spline.

When the spline curve passes through all the control points then it is called interpolate. When the curve is not passing through all the control points then that curve is called approximation spline.

14. What is a Blobby object?

Some objects do not maintain a fixed shape, but change their surface characteristics in certain motions or when in proximity to other objects. That is known as blobby objects. Example – molecular structures, water droplets.

14. Define Octrees.

Hierarchical tree structures called octrees, are used to represent solid objects in some graphics systems. Medical imaging and other applications that require displays of object cross sections commonly use octree representation.

15. What do you mean by view plane?

A view plane is nothing but the film plane in camera which is positioned and oriented for a particular shot of the scene.

16. What is view-plane normal vector?

This normal vector is the direction perpendicular to the view plane.

17. Explain about axonometric projection.

Orthogonal projections that display more than one face of an object are axonometric projection.

18. Explain about isometric projection.

Isometric projection is obtained by aligning the projection plane so that it intersects each coordinate axis in which the object is defined at the same distance from the origin.

PART – B & C

1. Write short notes and analyze on the list given below

1) Polygon surfaces
2) Curved line surfaces.
2) Quadric surfaces.

2. Explain 3D Transformations.

a. Translation
b. Rotation
c. Scaling
d. Reflect
e. Shearing

3. Describe B spline and Bezier surfaces.
4. Show the general characteristics of B spline curves.
5. Demonstrate uniform and cubic periodic B splines curves.
6. Explain in detail about Visible surface detection methods

UNIT - IV - MULTIMEDIA SYSTEM DESIGN & MULTIMEDIA FILE HANDLING
PART A

1. Give some Multimedia applications.

- Document imaging
- Image processing and Image recognition
- Full-Motion Digital Video Applications
- Electronic Messaging

2. What are the multimedia elements?

Facsimile, Document images, Photographic images, Geographical information system maps, Voice commands and voice synthesis, Audio messages, Video messages, Full-motion stored and live video, Holographic images, Fractuals.

3. What is Holography?

It is defined as the means of creating a unique photographic image without the use of a lens.

4. What is hologram?

The photographic recoding of the image is called a hologram, which appears to be an unrecognizable pattern of stripes and whorls but which when illuminate by coherent light as by a laser beam, organizes the light in to 3D representation of the original object.

5. What are the important processes in image processing?

Image recognition, image enhancement, image synthesis, and image reconstruction.

6. What are hypermedia documents?

In hypermedia documents in addition to text, embedded or linked multimedia objects such as image, audio, hologram, or full -motion video.

7. What are the sub-systems in DSP?

Memory management, hardware-interrupt handling, Multitasking, Inter task synchronization and communication, Multiple timer services, Device-independent I/O.

8. What are the types of images based on multimedia?

Visible images, non-visible images, abstract images.

9. What does a non-visible image refer?

Non-visible images are those that are not stored as images but are displayed as images e.g. pressure gauges, temperature gauges.

10. What are abstract images?

Abstract images are really not images that ever existed as real-world objects or representations. Rather they are computer-generated images based on some arithmetic calculations. e.g. fractals.

11. What is DVI?

The Digital Video Interface (DVI) standard was defined to provide a processor independent specification for a video interface that could accommodate most compression algorithms for fast multimedia displays.

12. What is MIDI?

This is the interface standard for file transfer between a computer and a musical instrument such as a digital piano.

13. What is Apple's Quick time?

The QuickTime standard, developed by Apple Computer, is designed to support multimedia applications. Apple's QuickTime is viewed as a multimedia interface that is evolving to become a standard part of the Apple as well as MS-Windows based systems.

14. What is JPEG?

The Joint Photographic Experts Group, formed as a joint ISO and CCITT working committee, is focused exclusively on still-image compression.

15. What is called Asymmetrical compression based on Compression?

are applications that need to be compressed once but are read many times.

16. What are the considerations in Multimedia storage?

Massive storage volumes, large object sizes, multiple related objects, temporal requirements for retrieval.

17. Define Fractals?

Fractals are regular objects with a high degree of irregular shapes. It is a lossy Compression technique but it doesn't change the shape of the image. Fractals are decompressed images that result from a compression format

18. Explain Fractal Compression?

Fractal Compression is based on image content i.e., it is based on similarity of patterns within an image. The steps in Fractal compression are

(i) A digitized image is broken into segments
(ii) The individual segments are checked against a library of fractals
(iii) The library contains a compact set of numbers called iterated function system codes.
(iv) These system codes will reproduce the corresponding fractal

PART – B & C

1, What are the evolving technologies for multimedia systems? Explain them.
2. Describe in detail about multimedia database.
3. Illustrate short notes on multimedia data interface standards.

4. Derive the issues involved in multimedia storage and retrieval.
5. What are the types of compression available in multimedia? Explain any two types of compression technology
6. Contrast TIFF fill structure and RIFF fill structure.
7. Discuss CCITT group of compression standards in detaiI.
8. Explain how does RAID technologies help in efficient storage and retrieval of multimedia data?
9. Give a detailed summary of MIDI.
10. Write in detail about JPEG & MPEG Compression standard.

UNIT - V - HYERMEDIA
PART A

1. What are the design issues for multimedia authoring?

Display resolution, Data formats for captured data, Compression algorithms,Network interfaces, and Storage formats.

2. What are the types of Multimedia authoring Systems?

Dedicated Authoring system, Timeline-Based Authoring, Structured Multimedia Authoring, Programmable Authoring Systems, Multisource Multi-User Authoring systems, Telephone Authoring Systems.

3. Classify the User interface development tools?

Media editors, An authoring application, Hypermedia object creation, Multimedia object locator and browser.

4. What is panning?

Panning implies that the image window is unable to display the full image at the selected resolution for display. In that case the image can be panned left to right or right to left as well as top to bottom or bottom to top. Panning is useful for finding detail that is not visible in the full image.

5. What are the steps needed for Hypermedia report generation?

Planning, Creating each component, Integrating components.

6. Define mail message.

Mail message is a message of a well-defined type that must include a message header and may include note parts, attachments, and other application-defined components. Note parts may include text, bitmaps, pictures, sound, and video components.

7. What are the characteristics of Document store?

Primary document storage, Linked object storage, Linked object management.

8. What are key issues in data organization for multimedia systems?

Data independence, Common Distributed Database Architecture, Multiple Data servers.

9. What are the functions performed by object request broker?

Object recompilation, Playback control, Format conversion.

10. What are the issues in database replication techniques?

Sharing of all data objects by all users on the networks, Providing acceptable performance to all users, allowing all users to update the database depending on the tasks being performed by them.

11. What are the types of database replication?

Round-robin replication, manual replication, scheduled replication, immediate replication, replication-on-demand, predictive replication, replicating references, no replication.

12. What are the primary n/w topologies used for multimedia?

Traditional LANS, extended LANS, High-speed LANS, WANS.

13. What are the components of a distributed Multimedia system?

Application s/w, Container object store, Image and still video store, Audio and video component store, Object directory service agent, Component service agent, User interface service agent, Networks.

14. What are the key elements in object server architecture of multimedia applications?

Multimedia application, Common object management, Object request broker, Object name server, Object directory manager, Object server, Object manager, Network manager, Object data store.

15. Give the primary goal of MAPI.

Separate client applications from the underlying messaging services, Make basic mail-enabling a standard feature for all applications, Support messaging-reliant workgroup applications.

16. What is the purpose of MIME?

Multipurpose Internet Mail Extension specification defines mechanisms for generalizing the message content to include multiple body parts and multiple data types.

17. What are the characteristics of image and still video stores ?

Compressed information, Multi-image documents, Related annotations, Large volumes, Migration b/w high-volume media such as an optical disk library and high speed media such as magnetic cache storage, shared access.

18. What are the services provided by a directory service agent?

Directory service, Object assignment, Object status management, Directory service domains, Directory service server elements, n/w access.

19. What are the services provided by User Interface Agent?

Window management, Object creation and capture, Object display and play back, Services on workstations, Using display s/w.

PART – B & C

1. Explain the types of multimedia authoring systems and list the main attribute, benefits and drawbacks of authoring systems.
2. Explain about the Hypermedia Message Components.
3. Explain Integrated multimedia message standards.
4. Explain the MAPI Architecture in detail.
5. Illustrate about distributed multimedia systems.
6. Describe Hypermedia messaging.
7. Develop a model with basic shapes, shading and texturing using BLENDER.

IT8076 SOFTWARE TESTING

UNIT – 1

1. Define process.

Process, in the software engineering domain, is the set of methods, practices, standards, documents, activities, policies, and procedures that software engineers use to develop and maintain a software system42and its associated artifacts, such as project and test plans, design documents, code and manuals.

2. Define Validation.

Validation is the process of evaluating a software system or component during, or at the end of, the development cycle in order to determine whether it satisfies specified requirements.

3. Define Verification.

Verification is the process of evaluating a software system or component to determine whether the product of a given development phase satisfy the conditions imposed at the start of that phase.

4. Define Testing.

Testing is generally described as a group of procedures carried out to evaluate some aspects of a piece of software.

(OR)

Testing can be described as a process used for revealing defects in software, and for establishing that the software has attained a specified degree of quality with respect to selected attributes.

5. Define Debugging.

Debugging, or fault localization is the process of

- Locating the fault or defect.
- Repairing the codes.
- Retesting the code.

6. List out the levels of the testing maturity model..

- Level 1: Initial.
- Level 2: Phase Definition.
- Level 3: Integration.
- Level 4: Management and Measurement.
- Level 5: Optimization/Defect prevention and quality control.

7. Define Errors.

An error is a mistake, misconception, or misunderstanding on the part of a software developer.

8. Define Faults.

A fault (defect) is introduced into the software as the result of an error. It is an anomaly in the software that may cause it to behave incorrectly, and not according to its specification.

9. Define Failures.

A failure is the inability of a software system or component to perform its required functions within specified performance requirements.

10. Define Test case.

A test case in the practical sense is a test- related item which contain the following information:

- A set of test inputs.
- Execution conditions.
- Expected outputs.

11. Define Test.

A test is a group of related test cases, or a group of related test cases and test procedures.

12. Define Test Oracle.

A test oracle is a document, or piece of software that allows tester to determine whether a test has been passed or failed.

13. Define Test Bed.

A test bed is an environment that contains all the hardware and software needed to test a software component or a software system.

14. Define Quality.

Two concise definitions for quality.

- Quality relates to the degree to which a system, system component, or process meets specified requirements.
- Quality relates to the degree to which a system, system component, or process meets customer or user needs, or expectations.

15. Define Metric.

A metric is a quantitative measure of the degree to which a system, system component, or process possesses a given attribute.

16. Define Quality Metric.

A quality metric is a quantitative measure of the degree to which an item possesses a quality attribute.

17. Define SQA.

The software quality assurance (SQA) group is a team of people with the necessary training and skills to ensure that all necessary actions are taken during the development process so that the resulting software conforms to established technical requirements.

18. Define Review.

A review is a group meeting whose purpose is to evaluate artifact or a set of software artifacts.

19. Define Precondition.

A precondition is a condition that must be true in order for a software component to operate properly.

Eg; number_of_coins > =0

20. Define Postcondition.

A post condition is a condition that must be true when a software component completes its operation properly.

Eg; number_of_dollars, number_of_cents > = 0

UNIT – 2

1. List the two basic Testing strategies.

- Black box testing.
- White box testing.

2. What are the knowledge sources for Black box testing?

- Requirements
- Document specification
- Domain knowledge
- Defect analysis data

✓

3. What are the knowledge sources for White box testing?

- High level design
- Detailed design
- Control flow graphs
- Cyclomatic complexity

4. List the methods of Black box testing.

- Equivalence class partitioning
- Boundary value analysis
- State transition testing
- Cause and effect graphing
- Error guessing

5. List the methods of White box testing.

- Statement testing
- Branch testing

- Path testing
- Data flow testing
- Mutation testing
- Loop testing

6. Define Random testing.

Each software system or module has an input domain from which test input data is selected. If a tester randomly selects input from the domain, this is called Random testing.

7. Define Equivalence class partitioning.

If a tester is viewing the software-under-test as a black box with well defined inputs and outputs, a good approach to selecting test inputs is to use a method called Equivalence class partitioning.

8. List the advantages of Equivalence class partitioning.

- It eliminates the need for exhaustive testing, which is not feasible.
- It guides a tester in selecting a subset of test inputs with a high probability of detecting a defect.
- It allows a tester to cover a larger domain of inputs/outputs with a smaller subset selected from an Equivalence class.

9. Define Cause effect graphing.

It is a technique that can be used to combine conditions and derive an effective set of test cases that may disclose inconsistencies in a specification.

10. Define State.

A state is an internal configuration of a system or component. It is defined in terms of values assumed at a particular time for the variables that characterize the system or component.

11. Define Finite-state machine.

It is an abstract machine that can be represented by a state graph having a finite number of states and a finite number of transitions between states.

12. Define Usage profiles.

Usage profiles are characterizations of the population of intended uses of the software in its intended environment.

13. What is Certification

Certification refers to third-party assurance that a product, process, or service meets a specific set of requirements.

14. What is Test data set?

A test data set is statement, or branch, adequate if a test set T for program P causes all the statements, or branches, to be executed respectively.

15. Define Path.

A path is a sequence of control flow nodes usually beginning from the entry

node of a graph through to the exit node.

16. Define Variable

Variable is defined in a statement when its value is assigned or changed.

(OR) Variable is defined in a statement when its value is utilized in a statement. The value of the variable is not changed.

17. List the two major assumptions in Mutation testing.

- The component programmer hypothesis
- The coupling effect

18. Define Test set.

A test set T is said to be mutation adequate for program P provided that for every in equivalent mutant Pi of P there is an element t in T such that Pi[t]is not equal to P[t].

19. Define Error guessing.

Error guessing approach is based on the testers/developers past experience with code similar to code-under-test, and their intuition as to where defects may lurk in the code.

20. What is the goal of smart tester?

The goal of the smart tester is to understand the functionality, input/output domain, and the environment of use for the code being tested.

UNIT – 3

1. List the different levels of testing.

- Unit test
- Integration test
- System test
- Acceptance test.

2. Define Unit Testing

A unit is the smallest possible testable software component that can be characterized in several ways.

3. List the components suitable for unit test.

- Procedures and functions
- Classes/objects and methods
- Procedure-sized reusable components.

4. List the phases in the unit test planning.

- Phase 1: Describe unit test approach and risks.
- Phase 2: Identify unit features to be tested.
- Phase 3: Add levels of detailed to the plan.

5. List the issues in the unit test.

- Issue 1: Adequately testing classes.
- Issue 2: Observation of objects states and state changes.
- Issue 3: The retesting of classes-I

6. What is Test harness?

The auxiliary code developed to support to testing of units and components is called a test harness. The harness consists of drivers that call the target code and stubs that represent modules it calls.

7. List the major goals of Integration test.

- To detect defects that occurs on the interfaces of units.
- To assemble the individual units into working subsystems and the finally a complete system that is ready for system test

8. What is the advantage of Bottom up integration?

Bottom-up integration has the advantage that the lower-level modules are usually well tested early in the integration process. This is important if these modules are candidates for reuse.

9. What is a cluster?

A cluster consists of classes that are related, for example, they may work together to support a required functionality for the complete system.

10. List the several types of system tests.

- Functional testing
- Performance testing
- Stress testing
- Configuration testing
- Security testing
- Recovery testing

11. Define Load.

A load is a series of inputs that simulates a group of transactions.

12. List the two major requirements of Performance testing.

- Functional requirements
- Quality requirements.

13. What is meant by Stress testing?

When a system is tested with a load that causes it to allocate its resources in maximum amounts, this is called stress testing.

14. Give the examples of security testing.

- Password checking
- Legal and illegal entry with password
- Password Expiration
- Encryption
- Browsing
- Trap doors
- Viruses.

15. Define Recovery testing.

Recovery testing subjects a system5t1o losses of resources in order to determine if it can recover properly from these losses.

16. List the areas covered during recovery testing.

- Restart
- Switchover.

17. Define Use case.

A use case is a pattern, scenario, or exemplar of usage. It describes a typical interaction between the software system under development and a user.

18. Define Regression testing.

Regression testing is not a level of testing, but it is the retesting of the software that occurs when the changes are made to ensure that the new version of the software has retained the capabilities of the old version and that has no defect have been introduced due to the changes.

19. List the objectives of configuration testing.

- ✓ Show that all the configuration changing commands and menus work properly
- Show that all interchangeable devices are really interchangeable, and that they each enter the proper states for the specified conditions
- ✓ Show that the system's performance level is maintained when devices are interchanged, or when they fail.

20. List the effect of security breaches.

- Loss of information
- Corruption of information
- Misinformation
- Privacy violations
- Denial of service.

UNIT – 4

1. Define Goal in testing

A Goal can be described as

A statement of intent

A statement of a accomplishment that an individual

2. What are the three types of goals in testing

- Business Goal
- Technical Goal
- Political Goal

3. Define the term policy

A policy can be defined as a high-level statement of principle or course of action that is used to govern a set of activities in an organization.

4. Define Test Plan

A Plan is a document that provides a frame work or approach for achieving a set of goals.

5. List the various Test Plan components

- Test Plan identifier
- Introduction
- Items to be tested
- Features to be tested
- Pass/Fail criteria
- Suspension & Resumption criteria
- Testing tasks
- Test environment
- Risks & Contingencies
- Testing costs
- Approvals

6. Define Features

Features may be described as distinguishing characteristics of a software component or system.

7. Define the term Pass / Fail Criteria

Given a test item and a test case, the tester must have a set of criteria to decide on whether the test has been passed or failed upon execution.

8. Define Suspension & Resumption criteria.

The criteria to suspend and resume testing are described in the simplest of cases testing is suspended at the end of a working day and resumed the following

morning.

9. Define Work Breakdown Structure (WBS)

A Work Break Down structure is a hierarchical or tree like representation of all the tasks that are required to complete a project.

10. Define Risks & Contingencies.

Every testing effort has risks associated with it. Testing software with a high degree of critically, complexity, or a tight delivery deadline all impose risks that may have negative impacts on project goals.

11. Define Cost Driver

A Cost Driver can be described as a process or product factor that has an impact on overall project costs.

12. Explain the simple COCOMO equation

$E = a (\text{size in KLOC})^b$

13. What are the various components of the test plan

- Test Design Specification
- Test Case Specification
- Test Procedures specifications

14. Define Test Procedure

A Procedure in general as a sequence of steps required to carry out a specific task

15. Define Test Summary Report

This report is prepared when testing is complete. It is summary of the results of the testing efforts. It also becomes a part of the projects historical database and provides a basis for lessons learned as applied to futureprojects.

16. List the skills needed by a Test specialist

- Organizational and planning skills
- The ability to keep track of and pay attention to details
- The determination to discover and solve problems
- The ability to mentor and train others
- The ability to work with users and clients
- The ability to think creatively

17. What are the steps in forming the test group.

- Upper management support for test function
- Establish test group organization, career paths
- Define education and skill levels
- Develop job description
- Interview candidates
- Select Test group members

18. Explain the Test team hierarchy

- The Test Manager
- The Test Lead
- The Test Engineer
- The Junior Test Engineer

19. What is the use of V-model in testing

The V-model is model that illustrates how testing activities can be integrated in to each phase of the standard software life cycle.

20. What are the various approaches to test cost estimation

- COCOMO Model
- Use of test cost drivers
- Test Tasks
- Testers / Developers ratio
- Expert judgment

UNIT – 5

1. Define the term Project monitoring.

Project Monitoring refers to the activities and tasks managers engage in to periodically check the status of each project. Reports are prepared that compare the actual work done to the work that was planned.

2. Define the term Project controlling.

Project Controlling consists of developing and applying a set of corrective actions to get a project on track when monitoring shows a deviation from what was planned.

3. Define Milestones

Milestones are tangible events that are expected to occur at a certain time in the project's lifetime. Managers use them to determine project status.

4. List some examples of testing Milestones

- Completion of the Master test plan
- Completion of branch coverage for all units
- Execution of all planned system test
- Completion of the test summary report.

5. List various Measurements for monitoring testing status.

- ✓ **Coverage Measures**
- Test Case Development
- Test Execution
- Test Harness Development

6. List the types of testing measurements

- Coverage
- Test Case Development
- Test Execution
- Test Harness

7. What are the various Severity level hierarchy

- Catastrophic
- Critical
- Marginal
- Minor or Annoying

8. What are the four major activities associated with Configuration management.

- Identification of the Configuration items
- Change Control
- Configuration status reporting
- Configuration audits

9. Define Change Control Board (CCB).

There are 2 aspects of change control – one is tool based, the other term based. The team involved is called CCB.

10. Define the term Review.

A review is a group meeting whose purpose is to evaluate a software artifact or a set of software artifact.

11. Explain the benefits of review program

- ✓ Higher – quality software
- Increased productivity
- Closer adherence to project schedule
- Increased awareness of quality issues

12. List the types of reviews.

There are two major types of technical reviews

- Inspections
- Walkthrough

13. What are the various steps in the inspection process

- Entry Criteria
- Initiation
- Preparation
- Inspection Meeting
- Reporting results
- Rework & follow up

14. Define Walkthrough

Walkthrough are a type of technical review where the producer of the reviewed material serves as the review leader and actually guides the progression of the review. Walkthrough have traditionally been applied to design and code.

15. What are the advantages of review approach.

There are two pass approach for detect detection.

- Pass 1 has individuals first reading reviewed item
- Pass 2 has the item read by the group as a whole.

16. What are the various components of review plans.

- Review Goals
- Preconditions and items to be reviewed
- Roles, Participants, Team Size and time requirements
- Review Procedures
- Review Training
- Review Checklist

17. What are the various roles in review program

- Review Leader Review
- Recorder Reader
- Reviewer

18. List the various review team membership constituency

Review Team Members

- SQA Staff Testers
- Developers Users / Clients
- Specialists

19. What are the various different types of software artifacts.

- Requirement Reviews
- Design Reviews Code
- Reviews Test Plan reviews

20. Define Defect Removal Leverage (DRL).

This is a ratio of the defect detection rates from two review or test phases and can be expressed as

Defects / hour (review or test phase X) DRL = ---

Defects / hour (review or test phase Y)