Unique Paper Code : 32341401
Name of the Course : B. Sc. (Hons.) Computer Science – CBCS (old course)
Name of the Paper : Design and Analysis of Algorithms
Semester : IV
Year of admission : 2015,2016,2017 and 2018
Duration : 3 Hours
Maximum Marks : 75

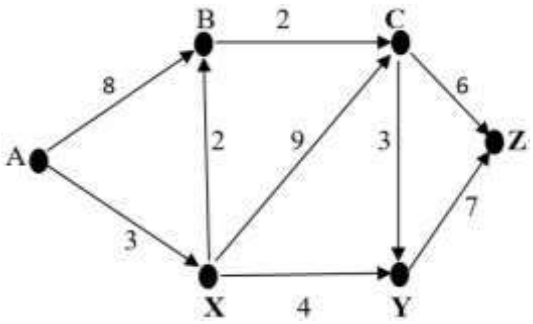
Instructions for Candidates:
Attempt Any Four questions. All Questions carry equal marks.

1. A student named Rohan wishes to enhance his knowledge by studying a few online courses. Every course has a fee that one has to pay to register. Also, each course gives some credit on its successful completion. He has some amount of money saved with him. He needs to choose among the available courses such that he gets maximum sum of credits by spending the money he has. Write an algorithm to find the optimal solution for Rohan's problem.
   Using the algorithm given by you, find the courses he should study for the given instance, illustrating each step clearly.

   Rohan has **$70**.

S. No.	Course Name	Fee (in $)	Credit
1.	C1	40	12
2.	C2	20	8
3.	C3	30	16
4.	C4	10	9
5.	C5	30	5

2. A person named **Albert (denoted by A)** starts a food delivery service. The people who wish to avail this service register themselves with him by providing their address. He hires delivery boys to deliver food and assigns a different delivery boy to each registered person. To save time, he asks the delivery boys to use Scooters to deliver the food. The following diagram shows the time that a delivery boy takes to reach from one place to another. Consider that **B, C, X, Y** and **Z** are the registered people. The number on each edge represents the time (in minutes). Compute the **minimum** time to reach them all individually.



   Now, suppose all Scooters need to be serviced and are unavailable. The delivery boys need to use bicycles to reach the registered people. The amount of time taken to reach the destination using a bicycle is the square of the time taken to reach the destination using a

Scooter. Will the delivery boys be reaching the respective destination using same route as computed above? Does this imply all such routes will never change in any such problem instance where the new edge weights are square of original edge weights?

3. Suppose there are n bottles and each bottle i has some capacity c_i. Write the most efficient algorithm to arrange the bottles in the order of their capacities. Also, explain the running time of your algorithm.
   Next, consider that every bottle has an expiry date before which its content should be consumed. Each bottle is assigned a priority such that a higher value indicates a higher priority. The bottle having a closer expiry date has a higher priority assigned to it. Give an efficient algorithm to arrange the bottles in non-increasing order of their priorities. Does your algorithm involve any comparisons? If yes, count them. Explain the running time of your algorithm.

4. Consider a set of jobs to be scheduled on a processor. Each job has an execution time associated with it. A job-scheduler allocates the processor to one of the given jobs at a time. The scheduler has fixed a *maximum execution time*: the maximum duration for which a job can continue processing once scheduled. If the execution time of any job is greater than this value then it is pre-empted and scheduled again later with the remaining execution time.
   Consider the following variant of **shortest-job-first** scheduling algorithm which schedules the jobs in the ascending order of their execution times. Suppose the data structure used for this scheduling algorithm is a red-black tree (called as scheduling tree) where the jobs are the nodes in the tree.
   Algorithm:
   - The job having least execution time will be chosen and sent for execution and is removed from the tree.
   - If the job reaches its *maximum execution time* while being processed, it is pre-empted and needs to be scheduled later to complete its processing. This job is reinserted into the scheduling tree based on its remaining execution time.
   - Again, the job with the least execution time is selected, repeating the above steps until all the jobs are processed completely.

   Consider a set of jobs whose execution times (in nanoseconds) are as follows: **4,10,7,23,5,30**. Construct the red-black tree corresponding to the given input and then schedule the jobs using the above stated algorithm thereby showing the red-black tree after every step. Assume **the maximum execution time is 20 nanoseconds**.

5. A party is going on in a hotel. When the dinner starts, a waiter is assigned the task of distributing plates to the guests. He creates a pile of plates on a table by adding one plate at a time. The waiter can pick up one or more plates at a time from the pile to distribute. The waiter may add more plates to the pile one by one at any time. Suppose he performs a total of n operations wherein an operation can be putting one plate onto the pile or removing one or more plates from the pile. Analyse the average cost per operation over a sequence of n operations.

Later, after the dinner gets over, a game is played to find out the most social person in the party. Suppose there are ***m*** guests in the party. Each guest considers a set of guests as his friends and gives a flower to each of his friends. The guest who receives the maximum number of flowers wins the game. Suggest an efficient divide-and-conquer algorithm, having least possible worst-case running time, to find a winner. Also, find all the winners, if more than one winner exists. Explain the running of the algorithm given by you.

6. A chemical laboratory has ***n*** chemicals stored in it. There is a risk that some chemicals may react with some others. For safety purpose, the lab assistant decides that for each pair of reactive chemicals, one is kept in a yellow-coloured bottle and the other is kept in a green-coloured bottle. Give an efficient algorithm that he can use to find out whether such a colouring is possible or not. If yes, the algorithm should determine the bottle colour for each of the ***n*** chemicals. If not, the algorithm should report it. Give an instance having at least ***6*** chemicals in which there are a minimum of ***7*** reactive pairs for which such a colouring is possible.

Unique Paper Code	:	**32341601**
Name of the Paper	:	**Artificial Intelligence**
Name of the Course	:	**B.Sc. (H) Computer Science**
Semester	:	**VI**
Duration of Examination	:	**Three Hours**
Maximum Marks	:	**75**

(For students admitted in 2015, 2016, 2017 & 2018)

Instructions for Candidates:

1. Attempt any **FOUR** questions.
2. Each question carries equal marks.

1. Give the Performance Measure, Environment, Actuators, and Sensors (PEAS) description for the *Automated Taxi Driver* environment. Differentiate between the following:

   - Fully observable vs. partially observable
   - Deterministic vs. stochastic
   - Episodic vs. sequential
   - Static vs. dynamic
   - Model-based agent and Goal-based agent
   - Goal-based agent and Utility-based agent

2. What are the differences between Recursive Transition Network (RTN) and Augmented Transition Network (ATN)? Draw the RTN to implement the grammar given below. Show the derivation of the sentence "**Mary slept on the sofa**" and also develop a parse tree using the following grammar:

S	→	NP VP
NP	→	N | DET N
VP	→	V | V PP
PP	→	PREP NP
N	→	Mary | sofa
V	→	slept
DET	→	the
PREP	→	on

3. Using the constraint satisfaction algorithm, solve the following cryptarithmetic problem:

$$\begin{array}{r} \quad B\ A\ S\ E \\ +\ B\ A\ L\ L \\ \hline G\ A\ M\ E\ S \\ \hline \end{array}$$

   Based on the solution of the above cryptarithmetic problem, find the value of B+L+A+M+E. Write a PROLOG program to implement GCD of two numbers.

4. Consider the following axioms:

   A1: Rajesh likes all kind of food.
   A2: Banana and Orange are food.
   A3: Anything anyone eats and not killed is food.
   A4: Madhav eats cashews and is still alive.
   A5: Anyone who is killed, is not alive.
   A6: Pankaj eats everything Madhav eats.

   Express the above axioms into First Order Predicate Logic (FOPL) statements and convert them into clausal form. Using resolution principle, prove that the statement "Rajesh likes cashews" is true.

   Transform the sentence $(\neg A \ \&\ B)\ V\ (A\ \&\ \neg B)\ \&\ C$ into Conjunctive Normal Form.

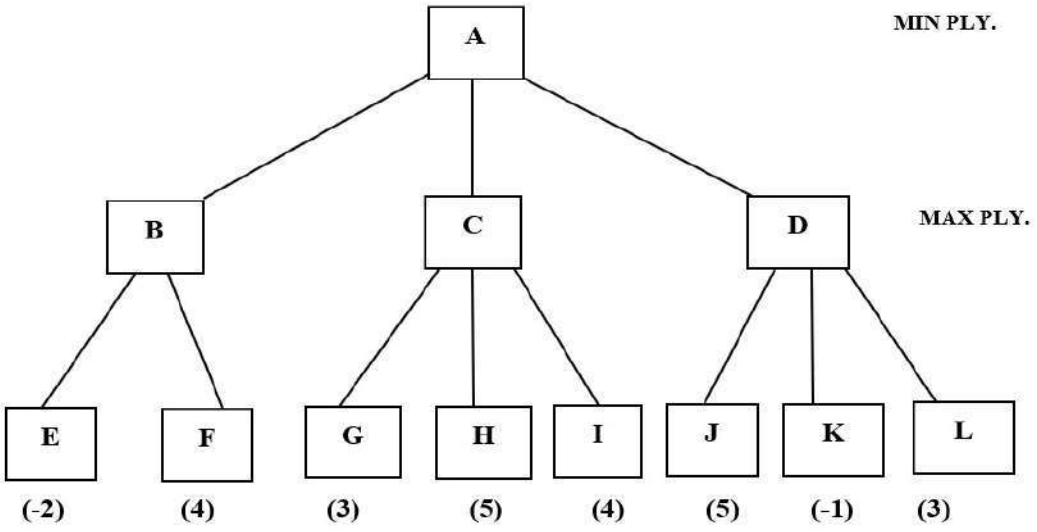
5. Differentiate between the monotonic reasoning and nonmonotonic reasoning. Give one example each of the monotonic and nonmonotonic reasoning.

   From experiments, it has been determined that $P(B|A) = 0.84$, $P(A) = 0.2$, and $P(B) = 0.34$. Find the probability $P(A|B)$ of the event A when it is known that some event B has already occurred. Describe how will you compute $P(A|\sim B)$ given only $P(A)$, $P(B|A)$, and $P(B)$?

   Draw the bayesian belief network for the given joint probability:
   $P(x_1, x_2, \ldots, x_7) = P(x_7|\, x_5, x_6)\, P(x_6|\, x_3, x_4)\, P(x_5|\, x_4)\, P(x_4|\, x_2)\, P(x_3|\, x_2)\, P(x_2|\, x_1)\, P(x_1)$

6. Give the similarities and differences between Best First Search and A* algorithm. Under what conditions A* algorithm provide an optimal solution?

   Consider the following game tree with ply depth 2 where the indicated scores are from the MIN player's point of view. Which move A should choose and why? Which nodes will be pruned according to the α-β pruning procedure? Give justifications of each.



Name of the Course : **B.Sc. (Hons.) Computer Science**

Semester : **VI**

Name of the paper : **Data Mining**

Unique Paper Code : **32347611**

Year of Admission : **2015, 2016, 2017 & 2018**

Duration: 3 Hours **Maximum Marks: 75**

Instructions for Candidates

1. Attempt any FOUR out of SIX questions.
2. All questions carry *equal* marks.
3. Upload single PDF file for each question.

1. Consider a sample dataset of patients visiting a clinic for consultation:

Patient ID	Patient Name	Blood Pressure	Blood Pressure date	Chest Pain	Age	Exercise	Heart Condition
101	Sarita	120	14-10-1995	1	20	High	Good
102	Maddy	110	16-1-2018	0	40	High	Good
103	Rohit	140	18-10-2018	3	?	Low	Bad
104	Gauhar	172	4-6-2018	3	39	Medium	Bad
105	Himani	150	8-6-2018	2	35	Low	Bad
106	Shubham	110	4-7-2018	1	40	Medium	Good
107	Suresh	120	5-3-2018	0	26	High	Good
108	Anmol	110	5-5-2018	1	27	Medium	Good

- Identify the type of each attribute and give justification.
- Perform min-max normalisation on "Blood Pressure" attribute.
- Identify the data quality issues for each of the following fields. Can these issues be resolved? If yes, how?
    - Blood Pressure date for patient 101.
    - The blood pressure meter is miscalibrated and adds 1 mmHg to each reading.
    - Age of the patient 103.

2. A binary classification problem with class labels: 'Yes' and 'No' that denotes the access to the elevator, has the following set of attributes and attribute values:

   Status = {Faculty, Student}
   Floor = {First, Second, Third, Fourth, Fifth}
   Health = {healthy, unhealthy}

   Consider the following set of records for the above classification problem:

Transaction Number	**Status**	**Floor**	**Health**	**Accessible**
1	Faculty	First	Healthy	Yes
2	Student	First	Healthy	No
3	Student	Third	Healthy	No
4	Faculty	Second	Unhealthy	Yes
5	Student	Fifth	Healthy	Yes
6	Faculty	First	Healthy	No
7	Student	Fifth	Unhealthy	No

   A rule-based classifier produces the following rule set:

   R1: Status = Faculty, Floor = Second → Accessible = Yes
   R2: Status = Student, Floor = Second → Accessible = No
   R3: Floor = First → Accessible = No
   R4: Health = unhealthy → Accessible = Yes
   R5: Status = Student, Floor = Fifth, Health = healthy → Accessible = Yes

   - Are the rules in the above rule set mutually exclusive? Justify.
   - Is the rule set exhaustive? Justify.
   - Is ordering needed for this set of rules? Justify
   - Do you need a default class for the rule set? Justify
   - Compute the coverage and accuracy of rules R1 and R5. Which one do you think is a better rule? Why?

3. Consider the training examples shown in the following table for a classification problem.

Student ID	Admission Category	Admission List	Gender	Predicted Class	Actual Class
1	Sports	First	M	C0	C0
2	Arts	Second	F	C0	C1
3	Arts	Third	F	C0	C0
4	Sports	Fourth	M	C0	C1
5	Academics	First	F	C0	C0
6	Arts	Second	F	C1	C1
7	Arts	Third	M	C1	C1
8	Sports	Fourth	M	C1	C0
9	Sports	Third	F	C1	C1
10	Arts	First	F	C1	C0
11	Sports	Third	M	C1	C1
12	Sports	Second	F	C1	C1
13	Sports	First	M	C1	C0

- Compute the Information Gain for the Student ID attribute.
- Compute the Gini Index for the Admission Category Attribute and Admission List attribute
- Which is a better attribute for split based on the Gini Index: Admission Category or Admission List? Why?
- Create a confusion matrix for the above data set and compute False positive rate, accuracy, recall and precision.

4. Consider the following set of points:

{44, 28, 48, 26, 32, 14, 52, 50}

Assuming that k=2, and initial cluster centres for k-means clustering are 5 and 38, compute the sum of squared errors (SSE) and cluster assignment for each iteration.

5. Consider the dataset given below:

Age	Income	Employed	Credit-rating	Buys Car
young	high	yes	fair	yes
young	high	no	good	no
middle	high	no	fair	yes
old	medium	no	fair	yes
old	medium	no	fair	yes
old	low	yes	good	no
old	medium	no	good	no
middle	high	yes	fair	yes

Compute all class conditional and class prior probabilities. Use Naïve Bayes classifier to predict the class of the following tuple:

X = (age = young, income = medium, employed = yes, credit rating = good)

6. Consider the market basket transactions shown in the following table. Use Apriori algorithm to answer the questions that follow.

TID	Item bought
1	oregano, chocolate, milk, cheese, french fries
2	milk, french fries, cheese, ketchup
3	chocolate, cheese, oregano, ketchup
4	chocolate, cheese, french fries
5	french fries, cheese, oregano, chocolate
6	chocolate, ketchup
7	oregano, french fries, ketchup
8	oregano, french fries, chocolate
9	ketchup, oregano, milk
10	french fries, chocolate

- Assuming the minimum support threshold is fixed at 40%, list the set of frequent 1-itemsets (L_1) and with their respective supports.
- List the itemsets in the set of candidate 2-itemsets (C_2) and calculate their supports.
- Generate all association rules from the itemsets in L_2 and also compute the confidence of these rules.