Indian Institute of Technology Jodhpur

CSL7610: Artificial Intelligence, Minor Exam

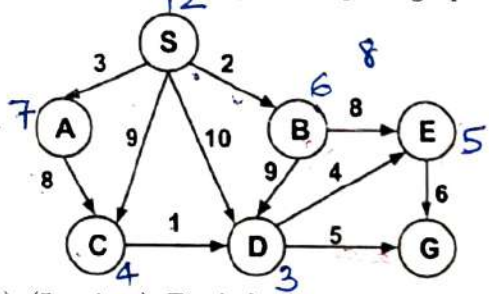
Date: Sep 20, 2024, Max Marks: 60 Max Time: 120 minutes [Pencil/Red Pen not allowed]

1. (a) (2 points) Give the equation that describes the property of a consistent heuristic function.

   (b) (4 points) Prove that if a heuristic function h(n) is consistent then $f(n+1) \geq f(n)$ where n is the immediate predecessor of n+1, g(n) is the cumulative cost of the path from the start node to node n, and $f(n) = g(n) + h(n)$. [This equation has to be proved mathematically and providing textual explanations will not be considered as a proof. Assume non-negative edge costs.]

   (c) (4 points) Now suppose $f(n+1) \geq f(n)$ is True for a particular heuristic function, prove that $h(n) \leq g(G) - g(n)$ where g(n) is the cumulative cost of the path from the start node to node n. [This equation has to be proved mathematically and providing textual explanations will not be considered as a proof. Assume non-negative edge costs.]

2. Consider the following state space graph of a search problem.



	S	A	B	C	D	E
h1	12	7	6	4	3	5
h2	3	7	8	2	2	1

   (a) (5 points) Find the lowest cost path using A* based graph search algorithm while using heuristics function h1. Draw the final search tree when the goal state is achieved. While inserting children nodes from the same expanded node into the fringe, follow an alphabetical order. Also, follow an alphabetical order if two partial plans have equal eligibility for expansion and were inserted into the fringe at the same time. Write the final fringe list without removing the deleted (or expanded) partial plans. The partial plans that were successfully expanded should be shown as crossed out. The partial plans that you tried to expand but failed due to some reason should be shown with an "X" symbol on the left and will be considered as deleted.

   (b) (5 points) Repeat (a) for heuristics function h2.

   (c) (2 points) How many partial plans were successfully expanded in (a) and in (b)

   (d) (1 point) Specify with reason which heuristic function dominates the other.

   (e) (2 points) Specify with reason which among h1 and h2 is/are admissible?

   (f) (5 points) Create a better heuristic function h3 using h1, h2 and any trivial heuristic function if needed, that requires a lesser number of partial plans to be expanded and repeat (a) with h3.

3. (a) (1 point) Explain which one among expectimax or expectiminimax search helps a player play the snakes and ladder game. – It is chance based game.

   (b) (4 points) Write down the Upper Confidence Bound heuristics (UCB1) equation for Monte Carlo Tree Search (MCTS). Consider an MCTS game tree of 2 levels (i.e., a single Max player node at level 0 and 3 Min player nodes at level 1 corresponding to the actions m1, m2, and m3 actions that the Max player can take). Say the Max player has won 3 out of 3 games using m1, 1 out of 3 games using m2 and 0 out 2 games using m3. Draw the MCTS game tree with all the required values and find out which move at level 1 will be selected next using UCB1. Consider all logarithms to be base 2 and the multiplication factor $C=\sqrt{2}$. Take $\sqrt{3} = 1.7$ and $\sqrt{2} = 1.4$.

4. (10 points) Solve the cryptarithmetic puzzle AB + BC = DE using Constraint Satisfaction Problem (CSP): **backtracking + arc consistency + MRV**. Each variable can take values from the set 1,2,3,4,5,6 **with no two variables having the same value**. **Another constraint is that B>C. Obviously AB + BC = DE is also a constraint. Fill the Table shown below. Stop at the first Failure** due to any assignment and **create a new Table** that has all the values copied from the previous table for all the previously assigned variables. e.g., if you had allotted a number to A, B, C without any failure and got a failure assigning the first available value to D, then create a new table and copy the version of the previous Table before assigning a new value to D. Now continue the assignment in the new Table. **Always assign values in the increasing order of the available values.** If multiple variables/nodes have the same importance as per MRV, use alphabetical ordering on top of MRV to decide the next variable. Order column should mention the order in which the variables were assigned values **in that Table only**, even if that resulted in a failure. **Start the assignment from A. In order to reduce the number of failures, do not assign the same value to multiple variables**

	1	2	3	4	5	6	Order
A							
B							
C							
D							
E							

5. (a) (5 points) In Propositional Logic, prove the Modus Ponens Inference Rule, when the knowledge base contains 2 sentences: (1) P⇒Q, (2) P, using proof by enumeration. Follow the same process of assigning values to propositional symbols as discussed in class otherwise no marks wi iven.

(b) (5 points) In each question **verify whether the query can be inferred or not** from given KB showing all the steps and rules using only PL Natural Deduction using Sound Inference Rules (0 marks will be given if you use resolution or truth values or statements like if A is True then B will also be True).

i)
KB
S∨(Q∧S)
(P∨R) ⇒¬S
Query: ¬R

ii)
KB
P⇒Q ⇔ R⇒ ¬Q
Q
Query: ¬R

(c) (5 points) In each question **verify whether the query can be inferred or not** from the given KB showing all the steps and rules using only Resolution Refutation (0 marks will be given if you use natural deduction or truth values or statements like if A is True then B will also be True).

i)
KB
P⇒Q
R⇒S
R∨S
¬Q
Query: S

ii)
KB
P⇒(Q∧R)
(Q∨S)⇒T
¬T
Query: P

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Indian Institute of Technology Jodhpur

Advanced Data Structures and Algorithms (CSL 7560)

Data Structures and Algorithmic Techniques (CSL 7561)

Instructor: Pallavi Jain

Thursday 19th September, 2024

Time: 2 hour | *Minor Examination* | Maximum Marks: 20

Instructions: You are allowed to use the data structures and the operations discussed in the class as a black box (no need to describe them). You only need to attempt questions worth 20 marks. You are also allowed to attempt only subparts of a question.

1. For the following problems, mark **True** or **False**. **The wrong answer is worth -0.5.**
   (a) Consider a B-Tree in which every node contains at least t values. The running time of SEARCH operation (find a given value x in the tree) is $O(\log_t(n))$, where n is the number of values in the tree. [1]
   (b) A red-black tree must have at least one red node. [1]

2. How do you determine whether a graph is connected? What is the complexity of your algorithm? [2]

3. Give an example to show that INSERTION in Red Black Tree is $\Omega(\log n)$. [2]

4. Determine the amortized cost per operation for a sequence of operations on a data structure in which the ith operation costs i if it is an exact power of 2, and 1 otherwise. [3]

5. Let $A[1,\ldots,n]$ be an array of n distinct numbers. If $i < j$ and $A[i] > A[j]$, then the pair (i,j) is called a *bad pair* of A.
   (a) List the five bad pairs of the array $\langle 2,3,8,6,1\rangle$. [1]
   (b) What array with elements from the set $\{1,2,\ldots,n\}$ has the most bad pairs? How many does it have? [2]
   (c) Design an algorithm to count the number of bad pairs in an array of size n in time $O(n\log n)$? [Hint: you can try to use order-statistic tree.] [3]

6. Let A be a set of n distinct numbers. The *gap* in A is the absolute value of the difference of the two closest numbers in A. For example, if $A = \{2,3,8,6\}$, then gap is 1 as 2 and 3 are closest numbers. Augment one of the data structures discussed in the class to find gap in A in $O(\log n)$ time. All the other operations INSERT, DELETE, SEARCH, MAX, MIN should still run in $O(\log n)$ time. Argue the correctness of your procedure. [4]

7. Consider an Interval Tree data structure. Let T be an Interval Tree. Given an interval $[a,b]$, design a procedure to check whether there exists a node x in T that stores interval $[a,b]$. We call this procedure EXACTSEARCH. You are allowed to store additional information in the tree, if required. However, all the operations INSERT, DELETE, MAX, MIN, EXACTSEARCH should run in $O(\log n)$ time. Argue the correctness of your procedure. [Hint: It is similar to the search operation discussed in the class.] [4]

Department of Computer Science and Engineering
Indian Institute of Technology Jodhpur
CSL 7620: Machine Learning
Minor Examination

Total marks: 40 **Time: 2 hour**

Instructions:

1. No queries will be answered during the exam. If you find anything unclear/incorrect in any question, make a reasonable assumption and proceed.
2. Total marks for each question are indicated beside or below the question inside the brackets.
3. All segments of one question must be answered in one place.
4. Write brief and to-the-point answers.
5. Write in clear and legible handwriting.
6. All the terms used to describe equations must be properly defined.
7. Do not write anything on the question paper. Use the last page of the answer sheet for rough work.

Answer all the questions.

C1 C2

1. You are building a binary classification model to determine whether an email is spam or not. The training dataset consists of 30% spam emails, while the remaining 70% are non-spam. Each email is described by two key features: whether it contains the word "sale" and whether it has more than 2 links. 60% of spam emails contain the word "sale," and 15% of non-spam emails contain the word "sale." Additionally, 80% of spam emails contain more than 2 links, while 20% of non-spam emails contain more than 2 links. Suppose a new email contains more than 2 links and the word "sale". What is the probability that this new email is spam? **[4 marks]**

2. Explain gradient descent in the context of linear regression. Compare batch gradient descent, stochastic gradient descent, and mini-batch gradient descent. **[3+3=6 marks]**

3. Explain how reinforcement learning differs from supervised and unsupervised learning. **[2 marks]**

4. What is inductive bias? Explain the inductive bias in the context of linear regression. Discuss the bias-variance tradeoff in machine learning. **[2+2=4 marks]**

5. "DBSCAN is less affected by outliers compared to k-means clustering". Do you agree with the statement? Explain by mentioning the steps of both DBSCAN and k-means clustering algorithms. **[4 marks]**

6. Define VC-dimension using the explanation of shattering. How is VC-dimension related to test error? **[3+2=5 marks]**

Department of Computer Science and Engineering
Indian Institute of Technology Jodhpur
CSL 7620: Machine Learning
Minor Examination

Total marks: 40 **Time: 2 hour**

7. "Maximum Likelihood Estimation (MLE) can be used for performing classification". Do you agree with the statement? Explain. Also, explain the steps of MLE. **[1+2=3 marks]**

8. Why do we need standardization of input data? Explain with suitable equations. **[2 marks]**

9. What is conditional independence? Explain a Bayesian classification approach that uses conditional independence. **[1+2=3 marks]**

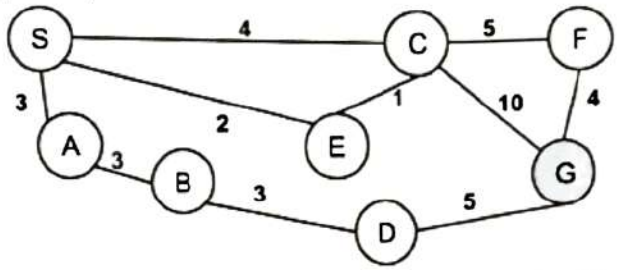
10. "Overfitting is more likely when a model becomes too complex". Do you agree? Explain. Suggest a method to prevent overfitting. **[2+1=3 marks]**

11. How is minimum error classification achieved using the Bayes Decision Rule? Provide relevant equations. Explain the notion of risk in the context of Bayesian classification. **[2+2=4 marks]**

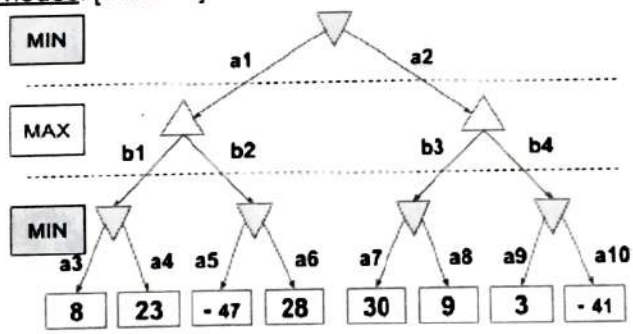
CSL7610 AI Quiz 1 (2024)

Total Marks: 30 **Duration:** 40 mins **Note: Answers written in Pencil will get 0 Marks**

1. Consider the following state space graph of a search problem.
   a. Find the lowest cost path from S to G using the graph-based UCS algorithm and show the resulting path. [2 Marks]
   b. Draw the final search tree when the goal state is achieved. While inserting children nodes from the same expanded node into the fringe, follow an alphabetical order. Also, follow an alphabetical order if two partial plans have equal eligibility for expansion and were inserted into the fringe at the same time. [3 Marks]
   c. Write the final fringe list without removing the deleted (or expanded) partial plans. **The deleted (or expanded) partial plans should be shown as crossed out.** [3 Marks]
   d. How many partial plans were expanded, including the partial plan only containing the start state and the partial plan containing the goal state that is identified as the solution? [2 Marks]

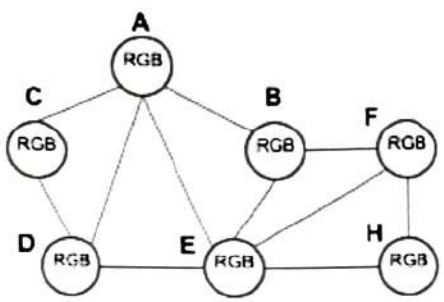


2. Perform MiniMax+alpha-beta pruning on the game tree to select the next move for the MIN Player.
   a. What should be the next move/action for MIN? [3 Marks]
   b. Copy the game tree and mention the final values at each min and max nodes, the final alpha and beta values for all nodes in the game tree, the number of pruned nodes and also cross the pruned nodes. [7 Marks]



min. remaining value

3. In the following graph coloring problem, use **backtracking + arc consistency + MRV** (for ordering the variables/nodes), to color all the nodes such that **no neighbors (having a direct edge between them) have the same color**. Colors available are **Red, Green, Blue**. **If multiple variables/nodes have the same importance as per MRV, use alphabetical ordering on top of MRV to decide the next variable.**
   a. If **node A is colored as Red**, **which variable/node** will you assign a color to **next**? [5 Marks]
   b. How many more nodes after A and your answer in part (a), will you assign colors to **before none of the remaining nodes have more than 1 color left**? [5 Marks]



Indian Institute of Technology Jodhpur

Statistics for Data Science - MAL 7060

First Semester (2024-25) Minor Exam (25%)

Note - Q. 1 to 6 are compulsary for all the students. Q. 7 is for Fractal 1 (Only) students. Q. 8 is for students who have opted both the fractals.

1. If a random variable has the normal distribution with $\mu = 16.2$ and $\sigma = 1.562$. find the probabilities that it will take on a value [3]

   (a) greater than 16.8;

   (b) less than 14.9;

   (c) between 13.6 and 18.8.

2. Assuming $\sigma = 20$, how large a random sample be taken to assert with probability 0.95 that the sample mean will not differ from the true mean by more than 2 points? [3]

3. If the service life of electron tubes in a particular application is normally distributed, and if 92.5 % of the tubes have lives greater than 2,160 hours, while 3.92 % have lives greater than 17,040 hours, what are the mean and S.D. of the service life? [4]

4. Let random variables X_1, X_2, X_3 be independent and distributed according to $N(1;1)$; $N(3;1)$, and $N(5;2)$, respectively. Determine probability $P(X_1 + X_2 + X_3 > 7)$. [4]

5 Consider the following sample for fat content of $n = 10$ randomly selected hot dogs: [2+4]

25 20 22 17 30 21 25 16 21 20

Assuming that these were selected from normal population, obtain a point estimator and a 95% CI for population mean.

6. The following measurements of the diameters of Indian mounds at a place were gathered [4+3]

22 24 24 30 22 20 28 30 24 34 36 15 37

   (a) Do these data substantiate the conjecture that the population mean diameter is larger than 21 feet? Test at $\alpha = 0.01$.

   (b) Determine a 90% confidence interval for the population mean diameter of Indian mounds.

7. **(For fractal I (only) students)** Let p denote the probability that, for a particular tennis player, the first serve is good. Since $p = 0.40$, this player decided to take lessons in order to increase p. When the lessons are completed, the hypothesis $H_0 : p = 0.40$ will be tested against $H_1 : p > 0.40$ based on $n = 25$ trials. Let Y equal the number of first serves that are good, and let the critical region be defined by $C = \{Y : Y \geq 13\}$. [4+4]

   (a) Determine the significance level α of the test.

   (b) Find the probability of Type-II error when $p = 0.60$.

8. **(For fractal I and II students)** In a small village of Rajasthan, a survey was conducted as follows. [4+2+2]

Visual Acuity Test(A)

Balance Test (B)	Left eyed	Ambiocular	Right-eyed
Left-handed	48	25	52
Ambidextrous	32	13	25
Right-handed	94	33	91

Using the Chi-square test at 5% significance level, determine whether there is any association between A and B. If association exists, find the degree of association by using the Karl Pearson's coefficient of contingency and the Tschuprow's coefficient.

Total Marks: 40

Duration: 120 mins

The questions have all the required information available.
Use of pencil is NOT allowed. If pencil is used to answer any question, that answer will not be checked.

1. **[8 marks]** Consider a Bayesian network representing the relationships between weather, traffic, and being late for work. The conditional probability tables are as follows:

W	P(W)
sunny	0.7
rainy	0.3

W	T	P(T|W)
sunny	+t	0.9
sunny	-t	0.1
rainy	+t	0.2
rainy	-t	0.8

W	T	L	P(L|T,W)
sunny	+t	+l	0.9
sunny	+t	-l	0.1
sunny	-t	+l	0.1
sunny	-t	-l	0.9
rainy	+t	+l	0.7
rainy	+t	-l	0.3
rainy	-t	+l	0.3
rainy	-t	-l	0.7

The nodes in the network are Weather (W), Traffic (T), and Late (L). Solve the room allotment problem with the following constraints:

a. Draw the Bayesian Network [2 Marks]
b. What is the probability that it is rainy, there is traffic (+t), and you are late for work (+l) [3 Marks]
c. What is the probability distribution over L with W=sunny, i.e. P(L,W=sunny) [3 Marks]

2. **[5 marks]** Let's consider a Bayesian network (BN) involving five variables: Health (H), Exercise (E), Diet (D), Stress (S), and Heart Disease (A). Suppose, all the conditional probability distributions have been provided to you, namely: P(H), P(E|H), P(D|H), P(S|D,E), P(A|S).

a. Draw the Bayesian Network [2 Marks]
b. What is the name of the algorithm used for checking for conditional independence in a BN [1 Mark]
c. If A=+a is known, check if E & D are conditionally independent using the above method. [2 Marks]

3. **[5 Marks]** Solve the following MDP (described in the table below) using Value Iteration:

Start State	Action	End State	Transition Prob	Reward
A	a	B	0.9	1
A	a	A	0.1	1
A	b	C	0.9	2
A	b	A	0.1	1
B	a	A	0.8	1
B	a	C	0.2	1.5
B	b	B	1	1
C	a	D	0.9	10
C	a	A	0.1	2
C	b	B	1	-1
D	a	D	0	0
D	b	D	1	0

There are 4 states A, B, C, D and 2 actions a and b. The discounting factor is 0.1. Repeat the iterations for either a maximum of 5 iteration steps or as long as the maximum difference between the previous and current expected value of any state is >0.1, whichever comes earlier. **Round off the expected V value obtained after every iteration to 1 decimal position.**

a. Draw the MDP state space with regular and Q nodes [2 Marks]
b. What is the final policy obtained? Show all the steps, and intermediate values. [3 Marks]

4. **[4 Marks]** Solve the following MDP. There are 5 states, A, B, C, D and two actions a and b. The policy is fixed with action a being recommended for all states. The transition probability and reward functions are not available, but the effects of 1 episode of the agent performing actions in the environment have been recorded. In this episode, A goes to B through action a, B goes to C through action a and C goes to D through action a. The rewards obtained are R1, R2, R3, i.e. Episode 1: (A,a,B,R1), (B,a,C,R2), (C,a,D,R3). No transition sample is available from D. Show that finding V(A), V(B), V(C), using Temporal Difference (TD) learning for this single episode is the same as using Direct Evaluation. Use discounting factor=0.1. Also in TD, at each timestep within an episode, the V(A) calculations only looks at the immediate next state after transition which is B in this case. Therefore, repeat the state-wise V calculations in TD till the previous calculation matches the current one. In TD, V(.) is initially 0 for all states. For Direct evaluation note that t starts from 0.

sampling are both same as prior sampling.

6. **[3 Marks]** In the following CSP variable value assignment table, **complete** the assignment of values (V1-9) to variables (R1-9) using backtracking, cycle consistency and MRV.

	V1	V2	V3	V4	V5	V6	V7	V8	V9	Order
R1	V1	x	x	x	x	x	x	x	x	1
R2	x	x	x	V4	x	x	x	x	x	2
R3	x	x	x	x	x	x	V7	x	x	3
R4	x	x		x	x		x	x	V9	7
R5	x	V2	x	x	x	x	x	x	x	4
R6	x	x	x	x	V5	x	x	x	x	5
R7	x	x	x	x	x	x	x	V8	x	6
R8	x	x	V3	x	x		x	x		8
R9	x	x		x	x	V1	x	x		9

In case of a tie between values follow the order V1>V2>V3>V4>V5>V6>V7>V8 and in case of tie between variables follow the order R1>R2>R3>R4>R5>R6>R7>R8>R9. **Some variables are already allotted. Don't change any existing entries.** Don't miss out on the Order column. x denotes values that cannot be assigned to current row (variable). The only constraint that you should look out for is that, adjacent variables should not have adjacent values, e.g., if R2 has V2 then R1 and R3 cannot have V1 or V3. Please Note that R1 and R9 are not adjacent variables and V1 and V9 are not adjacent values. One variable can have only one value and vice versa

7. **[4 Marks]** Verify if the query can be inferred from KB showing **all the steps and rules** using Resolution Refutation. (For inference only Resolution inference rule can be used, Don't use truth values or statements like if A is True then B will also be True) **[Solve any one of the two]**

a) $A \rightarrow (B \rightarrow C)$
$(A \vee C) \rightarrow D$
Query: C

b) $A \vee C$
$A \rightarrow B$
$C \rightarrow D$
$\neg B$
Query: D

8. **[4 Marks]** Verify if the query can be inferred from KB showing **all the steps and rules** using only PL Natural Deduction (no resolution or truth value). (Don't use truth values or statements like if A is True then B will also be True) **[Solve any one of the two]**

a) $P \vee (Q \wedge R)$
$(P \vee R) \rightarrow S$
Query: S

b) $A \vee (A \wedge B)$
$\neg C \rightarrow \neg A$
$C \rightarrow D$
Query: D

9. **[5 Marks]** Suppose you have an array of size N=5 and you have to search for a number D. The starting state is the middle position (3). You can start checking from the starting state using DFS with the graph formulation. Each state can only move 1 step to the left or right direction. Once DFS has been completed, use the closed set position of each state to get the heuristic value, i.e., if the final closed set after DFS is S,A,C,B,D, then $h(S)=5-1$. The cost of moving left or right is 1. Now solve this problem using A* search. **Draw "only" the last tree and "other values" for both methods. Donot miss out nodes that cannot be expanded. Also, write down the final solution for DFS and A*.** Note for this question: In cases where multiple children of a node have to be inserted into the fringe simultaneously and they have the same overall cost depending on the algorithm, insert them in such a way that when removing them from the fringe, the removal is automatically in alphabetical order.

1	2	3	4	5
B	A	S	C	D

END

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

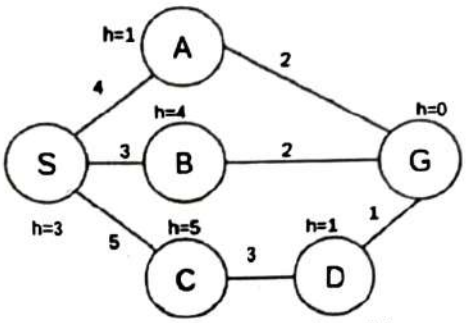
Indian Institute of Technology Jodhpur

September 9, 2023

Artificial Intelligence (CSL7610/CS323)

Time: 1 hour — *Minor 1 Exam* — Maximum Marks: 30

1. Convert the city to city traversal problem given in the figure below to a search problem such that you can apply any uninformed/informed search. Here S is the start city, and G is the target city.



   a) Define all the components of a search problem for this case. [3 Marks]

   b) Apply BFS, DFS and UCS tree search to find the solution. Incase of ties follow alphabetical order. For each approach, show all the intermediate search trees, fringe list and other values as the nodes get expanded and also write the solution sequence of states [6 Marks]

   c) Apply Greedy and A* tree search to find the solution. Incase of ties follow alphabetical order. For each approach, show all the intermediate search trees, fringe list and other values as the nodes get expanded and also write the solution sequence of states. [4 Marks]

   d) Is the solution provided by A* for the above problem optimal? Explain in detail what causes the optimality or non-optimality in this case. [2 Marks]

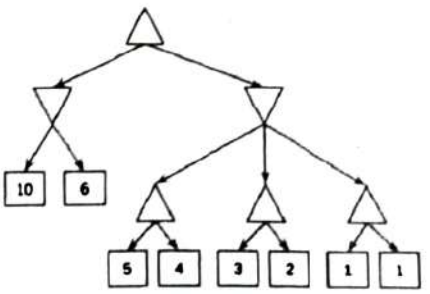
2. a) Difference between an admissible and a consistent heuristic function? [1 Mark]

   b) Suppose A* uses f(x)=max(g(x),p(x)), where g(x) follows the usual definition. If the path to x from start state s follows s,c,d,x, then p(x)=(h(c)-h(s))+(h(d)-h(c))+(h(x)-h(d)), where h(.) is a consistent and admissible heuristic function. Suppose there is an optimal goal state A and a suboptimal goal state B. Prove that f(A)<f(B). [Donot prove using examples] [2 Marks]

   c) The current algorithm is the same as which other search algorithm and why? [1 Mark]

3. What is the shoulder problem in hill-climbing and how enforced hill-climbing avoids this. [2 Marks]

4. a) Where is alpha-beta pruning used and why? Apply alpha-beta pruning below. Show all the intermediate states and checks involved. Also show its negative effect here [6 Marks]



5. "Expectimax gives the best solution for the MAX player given an optimal MIN player." Explain why or why not? Modify an expectiminimax game tree to get a minimax solution without changing any of the nodes or actions? Show an example game before and after the modification. [3 Marks]

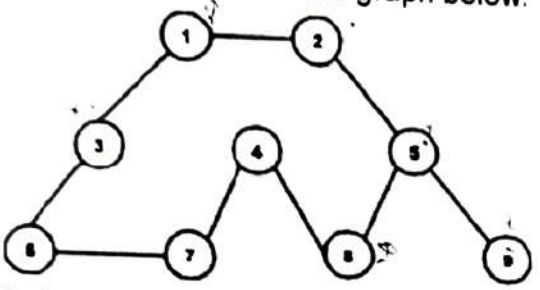
End of exam

CSL7610 AI Minor 2

Total Marks: 30 **Duration:** 60 mins

Use of pencil is NOT allowed. If pencil is used to answer any question, that answer will not be checked.

1. **[8 marks]** Consider the following problem: 8 cricketers arrive at a hotel after their match and have to be allotted rooms for their stay. The cricketers include 3 batsmen: B1, B2, B3, 3 bowlers: W1, W2, W3 and 2 all rounders: A1, A2. There are 9 rooms in the hotel that are arranged as shown in the graph below.



	B1	B2	B3	W1	W2	W3	A1	A2	Order
1	B1	X	X	X	X	X	X	X	1
2	X								
3	X								
4	X								
5	X								
6	X								
7	X								
8	X								
9	X								

Solve the room allotment problem with the following constraints:
C1: The path between the room/nodes of the same types of cricketers should contain more than 1 node, excluding the start and end nodes, e.g., if W2 gets room 2, then no other bowler can be allotted rooms 1, 3, 5, 8, 9
C2: Room number of B2 > Room number of A1
C3: Room number of B3 > Room number of W2

Solve using CSP with backtracking **a) with only forward checking** and **b) with forward checking, arc consistency, and MRV.** In simple forward checking, assign cricketers to rooms in increasing order of the room/node number (1,2,..). In the table, add X to denote that the current room (row) cannot be allotted to this cricketer (column). If multiple cricketers are available to be allotted to a room, then try to allot that room in the order B1>B2>B3>W1>W2>W3>A1>A2. Fill the above table for both of the methods. If more than one room is equally eligible to be allotted next, follow the increasing order of the room number to allot first. **Stop at the first Failure and mention FAILURE**. Failure can be caused by not having any rooms for a specific cricketer or having more than one room that cannot be allotted to anyone. Room 1 is allotted to B1 by default. **A cricketer can be allotted only one room and one room can have only one cricketer. Order column should mention the order in which you tried to allot the room, even if you were not able to assign any cricketer to that room.**

2. **[7 Marks]** Consider a problem where you have to make two dishes D1 and D2. D1 requires the ingredients T1 and T2, whereas D2 requires the ingredient T3. Initially, none of the ingredients and dishes are available. Model this problem as a planning problem using the STRIPS formulation. The actions available are getT1, getT2, getT3, makeD1, makeD2. getT1 can get the ingredient T1 but this cannot be done if T3 or G1 is already present. getT2 can get the ingredient T2 but can only be performed if T1 is already present. getT3 can get the ingredient T3 but this cannot be done if T1 or G2 is already present. makeD1 requires T1 and T2 to be present and it consumes/removes T1 and T2 to produce D1. makeD2 requires T3 to be present, and it consumes/removes T3 to produce D3.

Continued next page...

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Indian Institute of Technology Jodhpur

Advanced Data Structures and Algorithms (CSL 7560)

Data Structures and Algorithmic Techniques (CSL 7561) Instructor: Pallavi Jain

Tuesday 17th October, 2023

Time: 1 hour | *Minor 2* | Maximum Marks: 15

1. For the following problems, mention whether they are solvable in polynomial time or NP-complete. No proofs are necessary. For problems where there is an additional parameter k, it is an integer, and it is part of the input, and could be as large as n. **The wrong answer is worth -0.5.**

   (a) Given an undirected graph on n vertices and two designated vertices s and t, is there a path of length *at most* k between s and t? [1]

   (b) Given an undirected graph on n vertices and two designated vertices s and t, is there a path of length *at least* k between s and t? [1]

   (c) Given a graph G, a set $S \subseteq V(G)$ is called a *vertex cover* of G if for every edge $uv \in E(G)$, either u or v is in S. Given an undirected graph on n vertices, does it have a vertex cover on 15 vertices? [1]

   (d) Given a graph G, a set $S \subseteq V(G)$ is called a *clique* if G has an edge between every pair of vertices in S. Given a bipartite graph, does it have a clique on k vertices? [1]

   (e) Given a graph G, partition the vertex set into sets X and Y such that X and Y are independent sets in G. [1]

   (f) Find the smallest independent set in a graph. [1]

2. Show ONE of the following problems is NP-complete.

   (a) Given an undirected graph and an integer k, does it have at most k vertices that cover all cycles in the graph? I.e. the removal of the k vertices makes the graph acyclic. [4]

   (b) Given a set S of positive integers, is there a way to partition S into two subsets S_1 and S_2 that have the same sum? [4]

   You can assume that the following two problems are NP-Complete.

   - VERTEX COVER: Given a graph G and an integer k, find a vertex cover of G of size at most k.
   - SUBSET SUM: Given a set S of positive integers and a target integer T, is there a subset of S whose sum is T?

3. Consider the flow network D given in Figure 1.

   (a) Draw the residual network R of D [1]

   (b) Show an augmenting path in the residual network R. [1]

   (c) Show an augmented flow for D. [1]

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Indian Institute of Technology Jodhpur

Advanced Data Structures and Algorithms (CSL 7560)

Instructor: Pallavi Jain

Tuesday 5th December, 2023

Time: 2 hours — *End Semester Examination* — Maximum Marks: 35

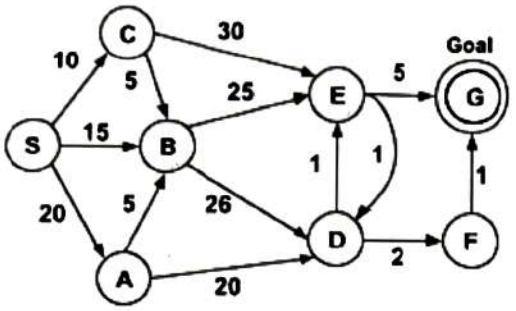
1. For the following problems, mark **True** or **False**. No proofs are necessary. For problems where there is an additional parameter k, it is an integer, and it is part of the input, and could be as large as n. **The wrong answer is worth -0.5.**
   (a) In a bipartite graph $G = (V, E)$, a clique on k vertices can be found in $O(1)$ time, if it exists. F [1]
   (b) Given an undirected graph on n vertices and m edges, we can test in $O(n + m)$ time if it is connected? [1]
   (c) The recurrence $F(n) = n + 2\sqrt{n} \cdot F(\sqrt{n})$ has the solution $\Theta(n \log n)$. [1]
   (d) A polynomial-time reduction from X to 3-SAT proves that X is NP-hard. [1]
   (e) k-CENTER does not admit 1.35 approximation unless P=NP. [1]
   (f) SAT does not have a polynomial time algorithm unless P=NP. T [1]
   (g) Searching a number in a sorted array of numbers can be done in $O(\log n)$ time. T [1]
   (h) Given an undirected graph on n vertices and two designated vertices s and t, is there a simple path of length at most k between s and t? F [1]
   (i) Given a list of $3n$ numbers, we can find the n-th largest element in $O(n)$ time. [1]
   (j) We can find a minimum sized cut of a graph G in $O(n^2)$ time deterministically. [1]
2. Show the red-black trees that result after successively inserting the keys 41, 38, 31, 12, 19, 8 into an initially empty red-black tree. [3]
3. Suppose you are given a directed graph $G = (V, E)$, two vertices s and t, a capacity function $c: E \to \mathbf{N}$, and a second function $f: E \to \mathbb{Z}_{\geq}0$. Describe an algorithm to determine whether f is a maximum (s, t)-flow in G. [3]
4. Design an algorithm to count the number of bad pairs in an array of size n in time $O(n \log n)$. [3]
5. Design an algorithm for 3-SAT that runs in $O(1.6181^n)$ time. [4]
6. This problem asks you to describe polynomial-time reductions between two closely related problems.
   - SUBSET SUM: Given a set S of positive integers and a target integer T, is there a subset of S whose sum is T?
   - PARTITION: Given a set S of positive integers, is there a way to partition S into two subsets S_1 and S_2 that have the same sum?

   (a) Describe a polynomial-time reduction from SUBSET SUM to PARTITION. [4]
   (b) Describe a polynomial-time reduction from PARTITION to SUBSET SUM. [4]

Total Marks: 55 **AI Major [2024]** **Time:3 hours**

PENCIL not allowed. In the questions with options, **choose only one option and mention it in your answer sheet** and not on the question paper.

Q1.



h(S)	h(A)	h(B)	h(C)	h(D)	h(E)	h(F)	h(G)
45	26	28	34	2	5	1	0

i. If you perform UCS from A, which of the following sequences of the first four expanded nodes ordered from left to right is possible? [2 Marks]
a) SABC b) ABDE c) ABDF d) SABD e) None of these

ii. Which of the following nodes has the heuristics value equal to the cost of the UCS path from that node to the goal node G. [2 Marks]
a) A b) B c) C d) D e) None of these

iii. What heuristic function for A* gives the same shortest path cost as UCS in this graph?
a) $\forall x\ h(x)=0$ b) $\forall x\ h(x)=1$ c) Options a and b d) None of these [1 Mark]

iv. Using graph-based A* search from S, which of the following sequences of the first four expanded nodes ordered from left to right is possible? [2 Marks]
a) SCBB b) SCBA c) ABDF d) SBDC e) None of these

v. What is the cost of the lowest-cost path found using graph-based A* search from S?
a) 44 b) 43 c) 41 d) 40 e) None of these [2 Marks]

Q2. Consider a Nims game with 5 marbles with a Max player and a Min player. In this game, based on the turn, one player picks either 1 or 2 marbles, and then the other player picks either 1 or 2 marbles. If after a Min/Max node picks some marbles and no marbles are left, then add a terminal node/state. The value of the terminal node/state is +10 if its parent is a Max node and is -10 if its parent is a Min node, i.e., if Max player picks 1 or 2 marbles and then if no marbles are left then Max is the winner and we get +10 reward. If at any Max/Min node only one marble is left, then the player can only pick one marble.

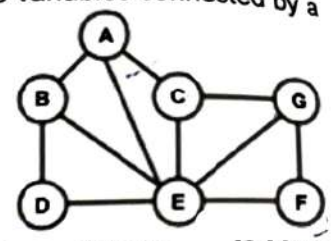
i. If Max player makes the first move, then what is the maximum and minimum depth of the game tree, including the terminal state/node, assuming a tree containing only the first Max player node is of depth 0. [1 Mark]
a) min=0, max=3 b) min=1, max=3 c) min=3, max=5 d) min=0, max=5 e) None of these

ii. Apply mini-max to find how many marbles should the Max player pick in his/her first move?
a) 0 b) 1 c) 2 d) 3 e) None of these [2 Marks]

iii. If you apply alpha-beta pruning to this game tree, how many Max/Min/Terminal nodes get deleted? [2 Marks]
a) 0 b) 2 c) 9 d) 10 e) None of these

iv. Basic alpha-beta pruning can be applied to the Expectimax game tree without requiring additional information. [1 Mark]
a) True b) False

Q3. Consider the below CSP with 5 variables with available values R,G,B. Two variables connected by a direct edge cannot have the same color.



i. List out all the possible solutions to the CSP? [4 Marks]
ii. What is the minimal cycle cutset for this CSP? [2 Marks]
iii. What is the number of residual CSPs to be solved due to cutset conditioning of the above minimal cutset? [2 Marks]

Q4. Is $(A \Rightarrow B \Rightarrow C \Rightarrow D \Rightarrow E) \wedge \neg (B \Rightarrow \neg C) \wedge A \wedge D \wedge E$ a tautology? a) True b) False [2 Marks]

Q5. i. Is $A(x) \vee B(x) \vee C(x)$ a definite clause in First Order Logic? a) True b) False [2 Marks]

ii. Is $B(x) \Rightarrow C(x)$ a definite clause in First Order Logic? a) True b) False [2 Marks]

iii. Knowledge Base

$\forall x,y\ (A(x) \wedge B(y) \rightarrow C(x,y))$
$\forall z,w\ (C(z,w) \wedge D(w) \rightarrow E(w))$
$\forall v\ (E(v) \rightarrow G(v))$
A(Ram)
B(Shyam)

Check if you can derive the query G(Shyam) from the Knowledge Base using Forward Chaining without assuming any other sentence. [3 Marks]

iv. Consider $\forall x,y\ (A(x) \wedge B(y) \rightarrow C(x,y))$. When you convert this sentence to propositional logic how many unique propositional symbols will you need? Assuming that Ram and Shyam are the only possible values for instantiation. Do not make any other assumptions. [3 Marks]

Q6. Is the following is satisfiable or not $(A \vee \neg B) \wedge (B \vee \neg C) \wedge (\neg A \vee \neg C) \wedge (A \vee \neg D) \wedge (\neg D \vee E) \wedge (\neg E \vee X) \wedge (\neg X \vee \neg G) \wedge (G \vee H) \wedge (\neg H \vee I) \wedge (I \vee \neg J) \wedge (\neg A \vee \neg B \vee \neg C) \wedge (A \vee B \vee C \vee D \vee E)$.

a) True b) False [3 Marks]

Q7. Bag A has 3 blue and 5 red balls, Bag B contains 4 blue and 6 red balls. 1 ball is picked at random from the bags and discovered to be red. Determine the probability that it was taken from bag B. Show the answer as a **fraction only**. [3 Marks]

Q8. Consider the joint distribution table with binary random variables A,B,C,D.

A	B	C	D	P(A,B,C,D)
-a	-b	-c	-d	0.30
-a	-b	-c	+d	0.05
-a	-b	+c	-d	0.12
-a	-b	+c	+d	0.03
-a	+b	-c	-d	0.03
-a	+b	-c	+d	0.05
-a	+b	+c	-d	0.05
-a	+b	+c	+d	0.05
+a	-b	-c	-d	0.10
+a	-b	-c	+d	0.02
+a	-b	+c	-d	0.05
+a	-b	+c	+d	0.02
+a	+b	-c	-d	0.03
+a	+b	-c	+d	0.05
+a	+b	+c	-d	0.02
+a	+b	+c	+d	0.03

i. What is the value of P(-a|-d). Show the answer as a **fraction only**.[2 Marks]
ii. What is the value of P(-c|-d). Show the answer as a **fraction only**.[2 Marks]
iii. What is the value of P(+a). Show the answer as a **fraction only**. [1 Mark]
iv. Assume the Bayesian Network has the following conditional probability tables: P(A|B,D), P(C|B,D), P(D), P(B). Now suppose you use likelihood weighted sampling to solve P(A|B,+d) and get a sample +a,+b,-c,+d. What is the weight of this sample? [2 Marks]

Q9. Consider a Markov Decision Process with states A, B, C, actions a1 and a2, with discount factor 0.1 and the following transition T and reward R details: [(A,a1,B) T=0.5, R=10], [(A,a1,A) T=0.5, R=2], [(A,a2,C) T=0.5, R=2], [(A,a2,B) T=0.5, R=4], [(B,a1,B) T=0.5, R=-1], [(B,a1,A) T=0.5, R=3], [(B,a2,A) T=0.6, R=1], [(B,a2,B) T=0.4, R=-1], [(C,a1,C) T=1, R=1], [(C,a2,C) T=0.5, R=1], [(C,a2,B) T=0.5, R=4].

Initial policy π, i.e., $\pi(A)=a2$, $\pi(B)=a1$, $\pi(C)=a2$ and you are using **Policy Iteration**

i. Using the initial policy, what are the values of $V_1^{\pi}(A), V_1^{\pi}(B), V_1^{\pi}(C)$ rounded to 2 decimal places? [3 Marks]

ii. Find the updated policy, using the values of $V_1^{\pi}(A), V_1^{\pi}(B), V_1^{\pi}(C)$ rounded to 2 decimal places. [4 Marks]

Name: Roll Number:

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
Indian Institute of Technology Jodhpur
CSL 7620: Machine Learning
Major Examination

Total marks: 60 **Set B** **Time: 3 hour**

Read the instructions carefully:

- You need to attempt all the questions. The marks allocated to each question is mentioned in brackets beside them.
- During the exam, invigilators will not address any queries. If you encounter anything unclear or incorrect in a question, make a reasonable assumption and proceed.
- Writing on the question paper is not allowed. Use the designated rough work page in the answer sheet for any calculations.
- Each question with options may have multiple correct options. You have to mark all the correct options.
- For the MCQs/ MSQs, mark the correct choices with a pen. No correction is allowed in the marking. You must not write anything inside the rectangular boxes in the OMR sheet.

1. Given the symmetric matrix $A = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 0 \\ 1 & 2 & 2 \end{bmatrix}$. What is the sum of its eigenvalues? (2)

A. 3 B. 4 C. 7 D. None

2. Given an input image of size $(224, 256)$, a filter of size $(4, 2)$ is applied with a stride of $(2, 3)$ and padding of $(2, 1)$. Which one of the following is the correct dimensions of the output feature map? (2)

A. $(113, 86)$ B. $(112, 85)$ C. $(112, 86)$ D. $(111, 85)$

3. A dataset contains 1000 emails labeled as "Spam" or "Not Spam". The emails are further categorized based on whether the word "money" is present. The data summary is as follows: (2)

	Spam	**Not Spam**	**Total**
"Money" Present	300	150	450
"Money" Absent	200	350	550
Total	500	500	1000

What is the probability that an email is "Spam" given that the word "money" is present?

A. 0.50 B. 0.55 C. 0.66 D. 0.75

4. A company has two servers, S_1 and S_2, processing customer requests. $P(S_1) = 0.4$ and $P(S_2) = 0.6$. The probability of a request failing on S_1, $P(F|S_1)$, is 0.03. The probability of a request failing on S_2, $P(F|S_2)$, is 0.01. Given that a request failed, what is the probability it came from server S_2, $P(S_2|F)$? (2)

A. 0.18 B. 0.33 C. 0.67 D. 0.81

5. A three-class classification problem has the following prior probabilities: (2)

$$P(C_1) = 0.1, \quad P(C_2) = 0.4, \quad P(C_3) = 0.5.$$

The Gaussian distribution formula gives the class-conditional densities for class C:

$$P(X = x|C) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}},$$

where μ is the mean and σ is the standard deviation of the distribution.

For class C_1, $\mu = 1$ and $\sigma = 2$, for class C_2, $\mu = 2$ and $\sigma = 1$, and for class C_3, $\mu = 1$ and $\sigma = 3$. Given $X = 0.5$, we conclude that the test sample belongs to _____?

A. C_1 B. C_2 C. C_3 D. None

6. Consider a set of 2D data points given by: (3)

$$(1,2), (2,9), (5,3), (5,9), (8,10), (11,11), (13,12), (15,5), (16,6), (17,7).$$

We want to apply k-means clustering with $k = 3$ clusters using Euclidean distance. The initial centroids are chosen as: $C_1 = (3.25, 5.75)$, $C_2 = (11.5, 7.5)$, $C_3 = (14.25, 9)$ After the first iteration, the new centroids are computed. Which of the following represents the correct positions of the new centroids?

A. $C_1 = (2.67, 4.67)$, $C_2 = (10, 8)$, $C_3 = (14.25, 9)$

B. $C_1 = (2.67, 4.67)$, $C_2 = (7.5, 9.5)$, $C_3 = (14.4, 8.2)$

C. $C_1 = (3.25, 5.75)$, $C_2 = (11.5, 7.5)$, $C_3 = (14.25, 9)$

D. $C_1 = (3.25, 5.75)$, $C_2 = (11.67, 9)$, $C_3 = (15, 8)$

7. The following table gives a data set for deciding whether to play or cancel a ball game, depending on the weather conditions. (3)

Outlook	Temp (F)	Humidity (%)	Windy	Class
sunny	75	65	true	Play
sunny	80	90	true	Don't Play
sunny	85	75	false	Don't Play
sunny	72	95	false	Don't Play
sunny	69	70	false	Play
overcast	72	90	true	Play
overcast	83	78	false	Play
overcast	64	65	true	Play
overcast	81	75	false	Play
rain	71	75	true	Don't Play
rain	65	70	true	Don't Play
rain	75	80	false	Don't Play
rain	68	80	false	Don't Play
rain	70	96	false	Play

Using the above data, we split the Humidity attributes into two child nodes using a threshold of 75 (i.e., binary split: humidity ≤ 75 and humidity > 75). Which of the following represents the correct entropy values of the child nodes (use logarithm with base 2 for the calculation of entropy)?

A. (0.954, 0.918) B. (1, 0.918) C. (0.954, 1) D. (0.985, 0.985)

8. Consider a dataset with three training examples. Adaboost is used to build a boosting ensemble. Each data point is assigned equal weights at the start i.e. $w_1 = w_2 = w_3 = \frac{1}{3}$ (w_1, w_2, and w_3 are the weights of training examples 1, 2, and 3, respectively). After the first iteration, the first stump makes predictions and misclassifies the second training example only. For growing the second stump, the updated weight (normalized) for the misclassified training example w_2 is _____? (3)

A. 0.269 B. 0.298 C. 0.365 D. 0.403

9. Suppose you have a dataset: $\mathcal{X} = \{1.8, 3.6, 2.4, 4.0, 1.6, 3.2, 3.8, 2.0\}$. This data is generated from a mixture of two Gaussian distributions with unknown parameters. You want to estimate the parameters of the two Gaussian distributions, (μ_1, σ_1^2) and (μ_2, σ_2^2) using the EM algorithm. Start with initial guesses for the parameters: $\mu_1 = 2.0$, $\mu_2 = 3.5$, $\sigma_1^2 = \sigma_2^2 = 1.0$. The initial mixing coefficients for the two Gaussians are $\pi_1 = 0.5, \pi_2 = 0.5$, respectively. What are the updated values of parameters μ_1 and mu_2 after 1 iteration of E-step and M-step? (3)

A. $\mu_1 = 2.277$, $\mu_2 = 3.297$

B. $\mu_1 = 3.297$, $\mu_2 = 2.277$

C. $\mu_1 = 3.703$, $\mu_2 = 1.723$

D. $\mu_1 = 1.723$, $\mu_2 = 3.703$

10. A neural network consists of two input neurons (x_1 and x_2), two hidden layer neurons (h_1 and h_2), and two output neurons (y_1 and y_2). Both the hidden and output layers include a bias term. The network is trained using Stochastic Gradient Descent (SGD) with the following parameters: the learning rate is $\eta = 0.5$, the activation function used in both hidden and output layers is the sigmoid function, and the biases for the hidden and output layers are $b_{\text{hidden}} = 0.35$ and $b_{\text{output}} = 0.6$, respectively. The loss function employed is the Mean Squared Error (MSE). (3)

The network receives inputs $x_1 = 0.05$ and $x_2 = 0.10$, and the target outputs are $y_1 = 0.01$ and $y_2 = 0.99$. Assume that the network is trained with only this training data. Consider the initial weights for the connections between the input layer and hidden layer $W^h_{\text{hidden} \times \text{input}} = \begin{bmatrix} 0.15 & 0.25 \\ 0.20 & 0.30 \end{bmatrix}$, and for the connections between the hidden layer and the output layer $W^o_{\text{output} \times \text{hidden}} = \begin{bmatrix} 0.40 & 0.50 \\ 0.45 & 0.55 \end{bmatrix}$.

Calculate the total net input to each hidden layer neuron. Sum the weighted inputs and add the bias term, followed by applying the activation function to squash the total input. Subsequently, compute the predictions for the output layer using the hidden layer outputs. Perform a backward pass to calculate the gradients and update the weights using gradient descent. Consider the training is complete. Now suppose a test data with $x_1 = 0.2$ and $x_2 = 0.15$ arrives. Compute the network's predictions for this test data. Compare the prediction with the ground truth $y_1 = 0.01$ and $y_2 = 0.99$. What is the error between the prediction and the ground truth if we use the MSE formula? Select the closest.

A. 0.181 B. 0.291 C. 0.345 D. 0.513

11. Fill in the blank with the options to complete the paragraph in the best possible way. (2)

In (a), we need not specify the number of clusters. However, in (b) and in (c), we need to specify the number of clusters. The number of parameters in (d) is two. While swapping of cluster centroid is a vital step in (e), such a step is not present in (f) or (g).

A. (a)k-means; (b)DBSCAN; (c)k-medoids; (d)DBSCAN; (e)k-means; (f)k-medoids; (g)DBSCAN;

B. (a)DBSCAN; (b)k-means; (c)k-medoids; (d)DBSCAN; (e)k-means; (f)k-medoids; (g)DBSCAN;

C. (a)k-means; (b)DBSCAN; (c)DBSCAN; (d)DBSCAN; (e)k-means; (f)k-medoids; (g)DBSCAN;

D. (a)DBSCAN; (b)k-means; (c)k-medoids; (d)DBSCAN; (e)k-medoids; (f)k-means; (g)DBSCAN;

12. Consider the following statements regarding Adaboost. (2)

a. Bootstrap sampling is an essential step in Adaboost

b. Normally, trees in Adaboost can have any number of nodes

c. In Adaboost, all the trees do not play an equal role in decision-making

d. Only a subset of features are considered for finding the best feature at the nodes of the trees in Adaboost

e. Adaboost can not be used for data with continuous feature values

f. Adaboost can be used for regression problems

Now, based on the above statements, find out the best option from the following.

A. Statements (b) and (e) are incorrect; Statements (a), (d), (c), and (f) are correct

B. Statements (a) and (e) are incorrect; Statements (b), (d), (c), and (f) are correct

C. Statements (a), (b), (d), and (e) are incorrect; Statements (c)and (f) are correct

D. Statements (c) and (e) are incorrect; Statements (a), (b), (d), and (f) are correct

13. Select the correct matching from Column A to Column B (2)

Column A	**Column B**
(a) Support Vectors	(i) Controls the penalty associated with crossing the class boundary
(b) Kernel Trick	(ii) A mathematical function that helps in reducing computation when higher-dimensional transformation of data is required
(c) Margin	(iii) Determined by support vectors
(d) C Parameter	(iv) Points closest to the decision boundary that influence the position of the decision boundary

A. $(a) \rightarrow (i), (b) \rightarrow (iv), (c) \rightarrow (iii), d \rightarrow (ii)$
B. $(a) \rightarrow (iii), (b) \rightarrow (i), (c) \rightarrow (ii), d \rightarrow (iv)$
C. $(a) \rightarrow (i), (b) \rightarrow (iii), (c) \rightarrow (ii), d \rightarrow (iv)$
D. $(a) \rightarrow (iv), (b) \rightarrow (ii), (c) \rightarrow (iii), d \rightarrow (i)$

14. Select the correct matching from Column A to Column B (2)

Column A	**Column B**
(a) Activation Function	(i) The process used for updating weights and bias in a neural network by calculating successive gradients.
(b) Backpropagation	(ii) Parameter sharing across the input.
(c) Convolution operation	(iii) The function that introduces non-linearity into the network and helps it learn complex patterns.
(d) Loss Function	(iv) A measure of the model's performance, which is minimized during training.

A. $(a) \rightarrow (ii), (b) \rightarrow (iii), (c) \rightarrow (i), d \rightarrow (iv)$
B. $(a) \rightarrow (ii), (b) \rightarrow (i), (c) \rightarrow (iii), d \rightarrow (iv)$
C. $(a) \rightarrow (i), (b) \rightarrow (ii), (c) \rightarrow (iv), d \rightarrow (iii)$
D. $(a) \rightarrow (iii), (b) \rightarrow (ii), (c) \rightarrow (i), d \rightarrow (iv)$

15. Select the correct matching from Column A to Column B (2)

Column A	Column B
(a) Bagging	(i) The process of putting more importance to misclassified data points
(b) Entropy	(ii) Splitting criterion in tree-based models based on the level of impurity or disorder in the data
(c) Variance	(iii) A technique where multiple models are trained in parallel and their results averaged or voted upon
(d) Boosting	(iv) Used for regression problem in tree-based models

A. $(a) \rightarrow (iv), (b) \rightarrow (ii), (c) \rightarrow (iii), d \rightarrow (i)$
B. $(a) \rightarrow (iii), (b) \rightarrow (ii), (c) \rightarrow (iv), d \rightarrow (i)$
C. $(a) \rightarrow (iv), (b) \rightarrow (ii), (c) \rightarrow (i), d \rightarrow (iii)$
D. $(a) \rightarrow (ii), (b) \rightarrow (i), (c) \rightarrow (iii), d \rightarrow (iv)$

16. Select the correct matching from Column A to Column B (2)

Column A	Column B
(a) Overfitting	(i) Occurs when the model is too simple to capture the underlying pattern in data
(b) Underfitting	(ii) Higher chance of overfitting
(c) Model-Complexity	(iii) May be related to the number of trainable parameters
(d) Small amount of training data	(iv) Model memorizes training data but performs poorly on unseen data

A. $(a) \rightarrow (i), (b) \rightarrow (ii), (c) \rightarrow (iv), d \rightarrow (iii)$
B. $(a) \rightarrow (ii), (b) \rightarrow (iv), (c) \rightarrow (iii), d \rightarrow (i)$
C. $(a) \rightarrow (i), (b) \rightarrow (iv), (c) \rightarrow (iii), d \rightarrow (ii)$
D. $(a) \rightarrow (iv), (b) \rightarrow (i), (c) \rightarrow (iii), d \rightarrow (ii)$

17. Fill in the blank with the options to complete the paragraph in the best possible way. (2)

PCA, LDA, and Autoencoders can be used for dimensionality reduction. Although (a) and (b) typically do not use class labels for dimensionality reduction, (c) uses class labels for the same. Utilizing class labels, important features can also be found using (d). However, after dimensionality reduction using class labels, (e) produces composite features unlike (f).

A. (a)LDA; (b)Autoencoder; (c)PCA; (d)Random forest; (e)PCA; (f)Random forest;
B. (a)PCA; (b)Autoencoder; (c)LDA; (d)Random forest; (e)LDA; (f)Random forest;
C. (a)PCA; (b)LDA; (c)Autoencoders; (d)Random forest; (e)Autoencoders; (f)Random forest;
D. (a)PCA; (b)Autoencoder; (c)LDA; (d)GMM; (e)LDA; (f)GMM;

18. Fill in the blank with the options to complete the paragraph in the best possible way. (1)

Theoretically, activation function $f(x)$ should be (a), or all values of x. However, (b) and (c) functions do not satisfy this criterion.

A. (a)Differentiable; (b)Relu; (c)Leaky Relu;
B. (a)Monotonically decreasing; (b)Sigmoid; (c)Relu;
C. (a)Monotonically increasing; (b)Sigmoid; (c)Tanh;
D. (a)Linear; (b)Sigmoid; (c)Tanh;

19. In the context of hypothesis evaluation, how does the VC dimension influence the generalization error of a model? (1)

A. Lower VC-dimension reduces the capacity to fit complex patterns, lowering generalization error
B. Higher VC-dimension always results in better generalization
C. VC-dimension has no direct impact on the generalization error
D. Higher VC-dimension increases the likelihood of overfitting, causing poor generalization

20. Which of the following is/ are correct about Autoencoders? (1)

A. Overcomplete autoencoders are not useful for any practical applications.
B. Undercomplete autoencoder may just copy the input copy and may not learn any salient features.
C. Overcomplete autoencoders can learn useful features without the need for any regularization.
D. Undercomplete autoencoders are designed to learn compressed representations of data.

21. Which of the following statements about PCA, LDA, and ICA are correct? (1)

A. PCA is primarily used for supervised learning tasks.
B. ICA assumes that observed data is a linear mixture of statistically independent components.
C. LDA and PCA always produce the same projections when applied to the same dataset.
D. LDA is a supervised technique that uses class labels to determine the optimal projection.

22. Which of the following statements about the role of inductive bias in decision trees are correct? (1)

A. The preference for splits that reduce impurity reflects the inductive bias of decision trees.
B. Inductive bias allows decision trees to avoid overfitting by restricting their growth.
C. Inductive bias ensures that a decision tree will always achieve perfect accuracy on the training data.
D. The inductive bias in decision trees always ensures optimal performance on unseen data.

23. Which of the following statements about the data normalization is/are correct? (1)

A. Normalization is preferred when the data has outliers.
B. Standardization and normalization always produce the same results.
C. Normalization ensures that the relative scale of the features is preserved.
D. Both normalization and standardization change the distribution of the data to a uniform distribution.

24. Which of the following is/ are true about the EM algorithm and Gaussian Mixture Model (GMM)? (1)

A. Each component in a GMM is characterized by its own mean, variance, and weight.
B. GMM requires the data to be linearly separable.
C. The E-step in EM calculates the probabilities of latent variables given the current parameters.
D. The EM algorithm is used to find the maximum likelihood estimates of parameters in probabilistic models.

25. Which of the following is/ are true about kernels? (1)

A. A kernel function implicitly defines a feature space and a mapping function.
B. A kernel is a function that computes the inner product of two points in a high-dimensional feature space.
C. The Radial Basis Function (RBF) kernel maps data into an infinite-dimensional space.
D. None of the others.

26. Which of the following is/ are not true about CNN? (1)

A. In CNN models, we never use dense (fully connected) layers

B. CNNs may suffer from vanishing gradient problem

C. CNNs do not suffer from the problem of overfitting

D. Autoencoders can be designed using CNN

27. Explain the working principle of soft Margin SVM with suitable equations and figures. "The kernel trick does not provide any computational advantage." Do you agree with the statement? Explain. (5)

28. "Random forest aims to create diverse decision trees" Do you agree with the statement? Explain. Explain the inference (testing) process in random forest. (3)

29. "Class labels are not useful for any dimensionality reduction technique." Do you agree with the statement? Explain. (2)

30. Compare batch gradient descent and stochastic gradient descent. (2)

Major Examination

Indian Institute of Technology Jodhpur

Matrix Theory (MAL7051)

Marks: 45 Time: 180 Mins

---

1. The set of all $n \times n$ symmetric matrices having trace equal to zero is a subspace W of $M_{n\times n}(F)$. What is the dimension of W? Write down the basis of W when $n = 3$. (2)

2. Let V be an inner product space, and suppose that x and y are orthogonal vectors in V. Prove that (2)

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

3. Let $A = \begin{pmatrix} 3 & 5 \\ 5 & 2 \end{pmatrix}$, then what is the trace of A^{100}? (2)

4. Prove that if $A, B \in M_{n\times n}(\mathbb{F})$ have same eigenvalues, then $\det(A) = \det(B)$. Is converse also true? (2)

5. Calculate the eigenvalues of the following matrices. If matrices are diagonalizable then find the invertible matrix P s.t. $D = P^{-1}AP$, where D is diagonal matrix. (3)

$$\begin{pmatrix} 0 & -2 & -3 \\ -1 & 1 & -1 \\ 2 & 2 & 5 \end{pmatrix}$$

6. Find a singular value decomposition and pseudoinverse for the given matrix (4)

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

7. How to check a norm is induced from an inner-product? If a norm is induced from an inner-product, how to find that inner-product? (4)

8. Let F be a field. Consider $A \in \mathbb{M}(m \times n, F)$ and $B \in \mathbb{M}(n \times q, F)$. Pick the correct relation and justify your choice (4)

   (i) row/column space of $A \subseteq / \supseteq$ row/column space of AB. State when equality will hold.

   (ii) row/column space of $AB \subseteq / \supseteq$ row/column space of B. State when equality will hold.

9. For a 2×3 matrix, write the matrices associated with elementary row operations $E(1,2)$, $E(1,2)^{-1}$, and the elementary column operations $E(1,2)$, $E(1,2)^{-1}$. (4)

10. Consider $V = \{f : \mathbb{R} \to \mathbb{R} : f \text{ is infinitely often differentiable }\}$, which is a real vector space, and let $T : V \to V$ be $T(f) = f'$. Find the eigenvalues and eigenvectors of T. (2)

11. Let V, W be real vector spaces and $T \in L(V, W)$. If $Y \subset V$ is convex, what can be said about $T(Y)$? (2)

12. Suppose that A is a real, $n \times n$ symmetric matrix with $A^3 = A^2 + A - I$. Is A invertible? If so, find its inverse. (2)

13. Let A is a $n \times n$ real matrix such that $\sum_{j=1}^{n} a_{ij} = 1, \forall i = 1, \cdots, n$. Let $B = A^k$ for some $k \in \mathbb{N}$. Compute $\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}$. (2)

14. Let $A = \begin{bmatrix} 1 & -1 & 0 & 1 \\ x & 0 & 1 & x+1 \\ 1 & x-1 & 1 & x+1 \\ x & 0 & x & x \end{bmatrix}$, $x \in \mathbb{R}$. What is the rank of A dependent of $x \in \mathbb{R}$. (2)

(4)

15. Let A and B be $n \times n$ matrices. Prove or disprove each of the following.

(a) If A and B are diagonalizable, then so is $A + B$.

(b) If A and B are diagonalizable, then so is AB.

(c) If $A^2 = A$, then A is diagonalizable.

(d) If A^2 is diagonalizable, then so is A.

(4)

16. Let

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Define

$$\begin{aligned} T : \mathcal{M}_2(\mathbb{R}) &\longrightarrow \mathcal{M}_2(\mathbb{R}) \\ B &\longmapsto AB - BA. \end{aligned}$$

(a) Fix an ordered basis $\mathcal{B}$ of $\mathcal{M}_2(\mathbb{R})$ and find the matrix of T with respect to $\mathcal{B}$.

(b) Find a basis for each of the eigenspaces of T.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Indian Institute of Technology Jodhpur

Advanced Data Structures and Algorithms (CSL 7560)

Instructor: Pallavi Jain

Tuesday 26th November, 2024

Time: 3 hours | *Major Examination* | Maximum Marks: 40

Instructions: You are allowed to use the data structures and the operations discussed in the class as a black box (no need to describe them).

1. For the following problems, mark **True** or **False**. **The wrong answer is worth -0.5.**
   (a) Consider a red-black tree with at least one node. The tree must have at least one black node. [1]
   (b) Consider a red-black tree with at least one node. The tree must have at least one red node. [1]
   (c) Given an interval tree T and an interval i, we can list all intervals in T that overlap with i in $O(\lg n)$ time. [1]
   (d) Given a directed graph and two designated vertices s and t, we can test in polynomial time that there are two edge disjoint paths between s and t? [1]
   (e) Finding smallest independent set in a graph is NP-hard. [1]
   (f) Given a graph G, a set $S \subseteq V(G)$ is called a clique if for every pair of vertices $u, v \in S$, uv is an edge in G. Given an undirected graph on n vertices, can we test in polynomial time if it has a clique on 15 vertices? [1]
   (g) The number of global min cuts in a graph is polynomial. [1]
   (h) A polynomial-time reduction from X to 3-SAT proves that X is NP-hard. [1]
   (i) In a bipartite graph $G = (V, E)$, a clique on k vertices can be found in $O(1)$ time, if it exists. [1]
   (j) Given an undirected graph on n vertices and m edges, we can test in $O(n + m)$ time if it is connected? [1]

2. Suppose you are given a directed graph $G = (V, E)$, two vertices s and t, a capacity function $c\colon E \to \mathbb{N}$, and a second function $f\colon E \to \mathbb{Z}_{\geq}0$. Describe an algorithm to determine whether f is a maximum (s, t)-flow in G. [4]

3. Given an element x in an n-node order statistic tree and a natural number i, design an algorithm that finds ith successor of x in the linear order of tree in $O(\lg n)$ time. [5]

4. In the PRIME FACTORIZATION problem, you are given a positive integer n and the objective is to output all prime divisors of n. Example: If $n = 864$ then $864 = 2^5 \times 3^3$, and output is 2 and 3.
   (a) What is the size of the input to the problem PRIME FACTORIZATION? [1]
   (b) What does polynomial time algorithm would mean for the problem PRIME FACTORIZATION? [2]
   (c) How many distinct prime divisors a number n can have? Example: 864 has 2 distinct prime divisors: 2 and 3. [3]

5. In the VERTEX COVER problem, we are given a graph G and an integer k. The goal is to decide whether there exists a set $S \subseteq V(G)$ of size at most k such that for every $uv \in E(G)$, either u or v is in S. Design an algorithm that solves the problem in $O(1.46^k \cdot n^{O(1)})$ time. [4]

6. Consider Algorithm 1.

```
1: M = ∅
2: while G has at least one edge do
       let uv be any edge in G ;
       add uv to M;
       remove u and v from G;
   end while
3: return M
```

Algorithm 1: Greedy Algorithm

(a) Prove that M is a maximal matching, that is, M is not a subgraph of another matching in G. [2]

(b) Prove that M contains at most twice as many edges as the smallest maximal matching in G. [3]

7. In this problem we will show a problem to be NP-complete.

(a) A problem Π is NP-complete if (a) ___ and (b)___ holds. What are (a) and (b)? [2]

(b) In the CLIQUE problem, we are given as input an undirected graph $G = (V, E)$ and a positive integer k, and the objective is to test whether there exists a set $S \subseteq V$, of size at least k such that S forms a clique (that is, a set of vertices such that every two distinct vertices in the subset are adjacent.). Show that CLIQUE is NP-complete. [4]

Hint: You could either assume that 3-SAT *is NP-complete or* INDEPENDENT SET *is NP-complete.*