B.Sc. Engg. SWE 6th Semester 11 September 2025

ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)

ORGANISATION OF ISLAMIC COOPERATION (OIC)

Department of Computer Science and Engineering (CSE)

SEMESTER FINAL EXAMINATION **SUMMER SEMESTER, 2023-2024**
DURATION: 3 HOURS **FULL MARKS: 150**

CSE 4617: Artificial Intelligence

Programmable calculators are not allowed. Do not write anything on the question paper.

Answer all 6 (six) questions. Figures in the right margin indicate full marks of questions with corresponding COs and POs in parentheses.

---

1. Carlos is searching for a hidden treasure rumored to be somewhere in a 3-floored tower. He can start in L_1, L_2, or L_3, where L_1 is the lowest floor and L_3 is the highest. Carlos can take the action *climb*, which takes him from L_i to L_{i+1}. If he is already at L_3, he stays there. Similarly, he can take the action *descend*, which takes him from L_i to L_{i-1}. If he is at L_1, he stays there. *Climb* and *Descend* actions always succeed with probability 1.

   At any floor, Carlos can take the action *search*, where he looks for the hidden treasure on that floor. With this action, one of the three outcomes may occur:

   - He finds the hidden treasure, thereby reaching a terminal state named "Found".
   - He triggers a trap and reaches a terminal state named "Trapped".
   - He finds nothing, which keeps him on the same floor.

   Carlos models his predicament as a Markov Decision Process (MDP) where reaching the terminal state "Found" yields +1 reward, reaching the terminal state "Trapped" yields a reward of -1 and all other transitions yield 0 reward. Assume he uses a discount factor of $\gamma = 1$. Carlos creates the Table 1 to represent the transition probabilities of the *search* action in the MDP for different states s and corresponding future states s'.

   Table 1: The transition probability table for Question 1.

	s' ="Found"	$s' = s$	s' ="Trapped"
$s = L_3$	0.3	0.4	0.3
$s = L_2$	0.2	0.5	0.3
$s = L_1$	0.1	0.7	0.2

   a) Carlos plans to go to any specific floor and only take the *search* action until he reaches a terminal state. Which floor should he choose to get the maximum expected reward? What is the maximum expected reward? **10 + 3 (CO3) (PO2)**

   b) Assume that the reward for moving into the "Found" state got changed to +10 and the reward for moving into the "Trapped" state got changed to -10. All other transitions now yield a reward of -1. Everything else about the MDP stays unchanged. **6 + 6 (CO1) (PO1)**
   Perform two iterations of the *Value Iteration* algorithm to find the values of $V_0(s)$ and $V_1(s)$, where $s \in \{L_1, L_2, L_3\}$.

2. Peter, Roy, and Alice are in an $M \times N$ grid with each of them having a tin can telephone that is connected to others using a string. Thus, there are three strings connecting all three people. At each step, they can choose to move *up, down, left, right,* or *stay*. The strings will always be in a straight line and no two person can occupy the same grid. Peter, Roy and Alice want to move the least amount of steps possible until the lengths of all the strings are exactly k. The length of a string is calculated as the straight line distance between the two people it is attached to. The strings can stretch and contract based on the position of the people. All three people can chose an action at each timestep.

a) Compute the size of the state space in terms of M and N. Compute the maximum branching factor for this problem. $5+5$ (CO1) (PO1)

b) Design a non-trivial admissible heuristic for this problem using the positions of each person on the grid. 5 (CO3) (PO2)

c) Suppose G number of ghosts are introduced in the given scenario where, $G < M \times N$. Ghosts can move exactly like humans and if they are in a same grid as another human, the strings connecting that human to others get cut and the human gets removed from the grid. In this modified setting, the goal is the same as before while keeping all three strings intact. In this modified setting, only one person or ghost can make a move each turn. First the three humans each will take turns choosing their actions and then the ghosts each will take turns until one ply is completed. $5+5$ (CO3) (PO2)

   i. With proper justification, recommend an algorithm for finding the optimal move in this scenario.

   ii. "The game will never end if both humans and ghosts play perfectly." - Do you agree with the statement? Justify your position.

3. Sam has a boat that she uses to visit different islands. Currently, she wants to visit the island X. So she formulates the different islands as states. She can choose to take *left*, *right*, *up*, or *down* actions to move from one island to another. But due to rough weather, Sam's boat does not always move in the direction that she wants it to. Despite this setback, she follows some policy π and starts collecting samples in the form of (current state, action, next state, reward) tuples as seen in Table 2.

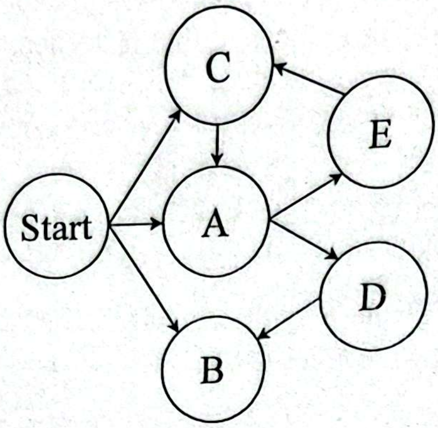
Table 2: The observed samples table for Question 3.

State (s)	Action (a)	Next State (s′)	Reward (r)
F	Right	G	+20
D	Right	G	−10
G	Up	T	−30
P	Down	Z	−15
G	Right	D	+30
W	Down	D	−25
G	Right	D	+30
D	Left	G	−5
G	Right	T	−30
W	Down	X	+100

a) What is $\hat{T}(G, Right, D)$ and $\hat{R}(W, Down, D)$ following the model-based approach of reinforcement learning? 5 (CO1) (PO1)

b) Sam now switches to approximate Q-learning. She takes three features, $f_i(s,a)$ and their corresponding weights w_i where $i \in \{1,2,3\}$. Initially, $w_1 = 2$, $w_2 = 5$, and $w_3 = 10$ while $f_1(G, right) = 6$, $f_2(G, right) = 3$, and $f_3(G, right) = 4$. 20 (CO1) (PO1)
Determine the weights (w_1, w_2, w_3) after performing weight updates for going *right* on G following the samples from Table 2. Assume a discount factor $\gamma = 0.5$ and learning rate $\alpha = 0.5$.

4. Assume you are part of a research team investigating the ecological behavior of a critically endangered snail species. Recently, the team has located a single specimen, providing a rare opportunity to study its foraging behavior in a controlled environment. The directed graph presented in Figure 1 represents the network of feeding areas accessible to the snail. In this graph, each node corresponds to a distinct feeding area, and each directed edge denotes a directional pathway that the snail can use to traverse between locations.



C A D B E

Figure 1: Different feeding areas and the possible actions denoted as a graph for Question 4.

a) If the preferred feeding site is E, find the sequence in which each area will be visited when using BFS, DFS and UCS (assuming costs are denoted by the total number of nodes visited prior to visiting a node). Break all ties in alphabetical order. 6×3 (CO1) (PO1)

b) The new criteria for finding a preferred feeding area is visiting a node three times. That means, the first area to be visited three times will be chosen. Find whether BFS or DFS would find the solution in the shortest number of nodes visited. 7 (CO3) (PO2)

5. You are an air traffic controller at an international airport. The airport only has two runways. You have to schedule a time slot and runway for each of the five aircraft (A, B, C, D, E) waiting for landing or takeoff. There are four time slots $\{1, 2, 3, 4\}$ for each runway, during which a plane can takeoff or land. No two aircraft can reserve the same time slot for the same runway. But there are some constraints such as:

- Plane B is in an emergency and must land within the first time slot.
- Plane D is running late and will arrive at the airport after the second time slot.
- Plane A has been delayed significantly and can not afford to takeoff any later than the second time slot.
- Plane D must land before plane C to transfer some passengers.

a) Formulate the scenario as a Constraint Satisfaction Problem (CSP) by identifying the variables, their possible values, and the constraints. 6 (CO3) (PO3)

b) Sketch the constraint graph for the CSP defined in Question 5.a. 5 (CO1) (PO1)

c) Assume that some additional constraints have been added: 14 (CO1) (PO1)

- Planes A, B, C can only use the first runway.

B,~~I~~,E D,E

- Planes $\boxed{C, D}$ can only use the second runway.

Using the Minimum Remaining Values (MRV) heuristic, perform backtracking search on the modified graph, breaking ties by picking lower values and characters first to find a solution to the modified CSP.

6. Blackbeard is in a crew of 9 miners who are preparing for a challenging expedition to collect valuable metal from a remote site. Before the expedition begins, one member must be selected as the team leader. The role of the leader comes with additional responsibilities, requiring an initial investment of 5 coins to procure extra supplies and equipment needed to manage the team effectively. If the expedition is successful, the leader receives one-fifth (1/5th) of the total resources as compensation for their added responsibilities. The remaining portion is evenly distributed among the other 8 team members.

   The utility of the leader is $U_L(l, t) = 5l + (t/5)$ while the utility of a member of the crew is $U_C(l) = l^2 - 4l$, where, l is the individual net coin gained and t is the total coins collected in a raid.

   a) Let X be the probability that a raid fails and the crew gets 0 coins. If $t = 100$, for what value of X should Blackbeard, a rational person, choose to be the leader of the crew? 15 (CO1) (PO1)

   b) State whether you agree with each of the following statements with brief explanations: 5×2 (CO1) (PO1)

      i. Local search has no optimality guarantees, so we always prefer to use A^* search over local search.

      ii. Reflex agents can be rational.

A B C D E

1,2