ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)

ORGANISATION OF ISLAMIC COOPERATION (OIC)

Department of Computer Science and Engineering (CSE)

SEMESTER FINAL EXAMINATION — WINTER SEMESTER, 2022-2023

DURATION: 3 HOURS — FULL MARKS: 150

CSE 4711: Artificial Intelligence

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

Answer **all 5 (five)** questions. Figures in the right margin indicate full marks of questions whereas corresponding CO and PO are written within parentheses.

1. Consider placing a robot in an unfamiliar environment. The robot perceives its current state, s_t, takes an action, a_t, transitions to a new state, s_{t+1}, and receives an instant reward, r_t. Assuming a discount factor of 0.5 and a learning rate of 0.5, the robot explores the environment, resulting in the experiences outlined in Table 1.

Table 1: Experience achieved in 5 iterations for Question 1

t	s_t	a_t	s_{t+1}	r_t
0	X	*Left*	Y	2
1	Y	*Left*	Y	-4
2	Y	*Right*	Y	0
3	Y	*Right*	X	3
4	X	*Right*	X	-1

a) Formulate the scenario as a sample-based Q-learning problem by determining Q-values for each state-action pair based on the experience of the robot. 24 (CO2) (PO2)

b) Determine the optimal policy. 6 (CO3) (PO3)

2. Consider creating a chatbot application where users ask questions, the chatbot replies, and users rate the answers. The rating dictates the utility of the chatbot. If the chatbot is unsure about an answer, it can choose actions as specified in Table 2.

Table 2: Possible chatbot scenarios for Question 2

Action	Potential Outcome	Probability	Utility
Say "I don't know"	User gets bored	1.0	0
Guess the answer	Correct answer	0.6	1
	Wrong answer	0.4	X
Trick the user to click "like"	Tricking succeeds	p	1
	Tricking fails	$1-p$	$2X$

a) Assume $X = -1$ and $p = 0$. With proper justification, recommend the optimal action for a rational chatbot. 10 (CO3) (PO3)

b) Assume that $p = 0.7$. Determine the necessary and sufficient condition in terms of X so that a rational chatbot performs the same action as recommended in Question 2.a). 8 (CO1) (PO1)

3. In Figure 1, a simplified map shows seven cities connected by bidirectional roads, each associated with a positive integer travel time. A robot plans to journey from city S to city T using Graph Search variant of search algorithms to find the path.

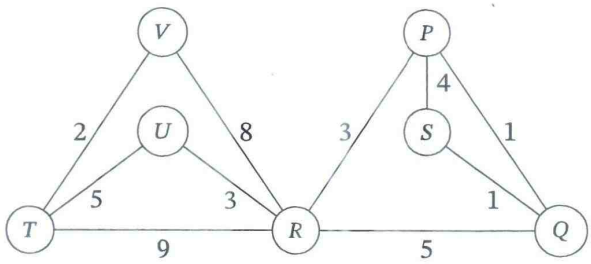

Figure 1: A simplified map of a country for Question 3

Two heuristics h_0 and h_1 are shown in Table 3. Here, h_0 is consistent, but h_1 is not consistent.

Table 3: Heuristic Functions for Question 3

Node	P	Q	R	S	T	U	V
$\mathbf{h_0}$	8	9	7	9.5	0	4	1.5
$\mathbf{h_1}$	10	12	8	10	0	4.5	1

Following is a list of paths which can be possibly be returned by a search algorithm:

- **Path 1:** $S \to P \to R \to T$
- **Path 2:** $S \to Q \to R \to T$
- **Path 3:** $S \to Q \to P \to R \to U \to T$

a) For each of the following algorithms, state the path(s) (among the aforementioned paths) returned: 5×6 (CO1) (PO1)

i. Depth First Search
ii. Breadth First Search
iii. Uniform Cost Search
iv. A* Search with Heuristic h_0
v. A* Search with Heuristic h_1

b) You are designing a new heuristic function h_2 for the given scenario. You have come up with all the values except $h_2(Q)$ as shown in Table 4: 3×10 (CO1) (PO1)

Table 4: Incomplete Heuristic Functions for Question 3.b)

Node	P	Q	R	S	T	U	V
$\mathbf{h_2}$	9	?	7	10	0	4.5	1.5

For each of the following scenarios, write the possible range of values for $h_2(Q)$ with a brief explanation:

i. h_2 will be admissible.
ii. h_2 will be consistent.
iii. A* Graph Search will traverse cities in the order: $S \to P \to Q \to R$.

4. Consider that an agent is moving in the Gridworld shown in Figure 2. The only action available in the cells A and E is `Exit` with rewards 0 and X, respectively. Here, X = The last digit of your student ID + 1. For the other cells, the agent can take the action `Left` or `Right`, which results in the agent moving to the immediate left or right cell, respectively. There is no living reward or penalty. All actions succeed with probability 1. The discount factor is 1.

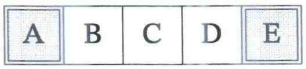

Figure 2: Gridworld for Question 4

Assume that we will run policy iteration to find the optimal action for the agent. In case of ties, we will choose `Left` instead of `Right`. The initial policy, $\pi_0(s)$ = `Left` for $s \in \{B, C, D\}$ and $\pi_0(s)$ = `Exit` for $s \in \{A, E\}$.

a) Formulate the scenario as a Markov Decision Process. 10 (CO2) (PO2)

b) With a brief explanation, determine the resulting policy after one iteration. 19 (CO1) (PO1)

c) Given that we start with the policy π_0, determine the number of iterations of policy iteration required to compute the optimal policy. 2 (CO1) (PO1)

5. After the yearly checkup of a patient, an AI-based agent, Alice said it has a 'bad' news and a 'good' news. The bad news is that the patient tested positive for a serious disease and that the test is 99% accurate. The good news is that this is a rare disease, striking only 1 in 10,000 people of the patient's age. Another AI-based agent, Bob concluded that the patient actually has the disease. 5 + 6 (CO3) (PO3)

a) Provide arguments to support the decision of Alice labeling the good news as 'good'.

b) Criticize the decision made by Bob.