

Roll No.																			
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--



**PRESIDENCY  
UNIVERSITY**

**BENGALURU**

**School of Computer Science and Engineering**

**Mid - Term Examinations - November 2024**

**Semester:** V

**Course Code:** CSA3074

**Course Name:** REINFORCEMENT LEARNING

**Program:** BSD

**Date:** 07-11-2024

**Time:** 11.45am to 01.15pm

**Max Marks:** 50

**Weightage:** 25%

**Instructions:**

*(i) Read all questions carefully and answer accordingly.*

*(ii) Do not write anything on the question paper other than roll number.*

---

**Part A**

**Answer ALL the Questions. Each question carries 2marks.**

**5Qx2M=10M**

- |   |   |         |    |     |
|---|---|---------|----|-----|
| 1 | Define a trajectory with an example                   | 2 Marks | L1 | C01 |
| 2 | Define Dynamic Programming and its Disadvantages      | 2 Marks | L1 | C02 |
| 3 | Discuss optimal policy                                | 2 Marks | L2 | C02 |
| 4 | Differentiate the episodic task and non-episodic task | 2 Marks | L2 | C01 |
| 5 | State the goal of Reinforcement learning              | 2 Marks | L1 | C01 |

**Part B**

**Answer ALL Questions. Each question carries 10 marks.**

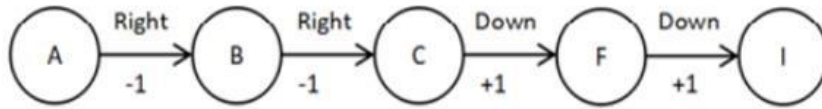
**4QX10M=40M**

- |   |                                    |        |    |     |
|---|------------------------------------|--------|----|-----|
| 6 | a. Discuss the Elements of RL      | 4Marks | L2 | C01 |
|   | b. Illustrate typical RL algorithm | 4Marks | L3 | C02 |
|   | c. Define Policy                   | 2Marks | L1 | C01 |

**Or**

- |   |  |        |    |     |
|---|--|--------|----|-----|
| 7 | a. Compare Discrete Action space and continuous action space | 4Marks | L2 | C01 |
|   | b. Interpret Value function                                  | 4Marks | L3 | C02 |
|   | c. Define model based learning                               | 2Marks | L1 | C01 |

- 8 a. Discuss different types of RL environments with an example each 4Marks L2 C02  
 b. Calculate the return for the following Episode 4Marks L3 C01



- c. Discuss the limitations of Monte carlo method 2Marks L2 C02

Or

- 9 a. Illustrate Monte Carlo prediction algorithm for every visit 4Marks L2 C02  
 b. Using the Model dynamics table of State A compute the optimal policy using Policy Iteration 4Marks L3 C01

State ( $s$ )	Action ( $a$ )	Next State ( $s'$ )	Transition Probability $P(s' s, a)$ or $P_{ss'}^a$	Reward Function $R(s, a, s')$ or $R_{ss'}^a$
A	0	A	0.1	0
A	0	B	0.8	-1
A	0	C	0.1	1
A	1	A	0.1	0
A	1	B	0.0	-1
A	1	C	0.9	0

Table 3.14: Model dynamics of state A

- c. Identify the value of the cards J,4,Q and 'Ace' in the blackjack game? 2Marks L2 C02

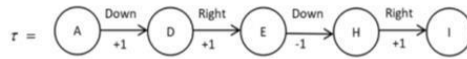
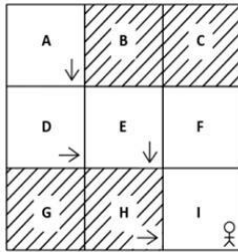
- 10 a. Discuss Markov Decision Process in detail 4Marks L2 C02  
 b. Calculate the return for the following Episode 4Marks L3 C01



- c. Define Q function 2Marks L1 C02

Or

- 11 a. Discuss Markov property with examples 4Marks L2 C02  
 b. For the following grid world Environment Solve the Value function which follows a deterministic policy. 4Marks L3 C01



c. Interpret the goal of a reinforcement learning

2Marks L2 C02

12 a. Differentiate the Discrete action space and continuous action space

4Marks L2 C01

b. Illustrate Monte carlo prediction algorithm

4Marks L3 C02

c. Distinguish between continuous task and episodic task

2Marks L2 C01

Or

13 a. Discuss epsilon greedy technique

4Marks L2 C01

b. Illustrate Monte carlo control algorithm

4Marks L3 C02

c. Interpret exploration-exploitation dilemma

2Marks L2 C02