Roll No: | | | | | | | | | | | |

PRESIDENCY UNIVERSITY

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

Presidency University Act, 2013 of the Karnataka Act No. 41 of 2013 | Established under Section 2(f) of UGC Act, 1956
Approved by AICTE, New Delhi | Approved By BCI
Bengaluru

Mid Term Examinations - October 2026

Date: 09/10/2026 **Time:** 11:45 AM - 01:15 PM

Course Code: CSA4708 **Course Name:** Reinforcement Learning

Semester: Third Semester **Max. Marks:** 50 **Weightage:** 50%

CO - Levels	CO1	CO2	CO3
Marks	50	46	84

PART A: Answer Following Questions. 10 M

Q.N.	Questions	M	CO	BT
1	What is meant by Markov Decision Process? State the Equation for Solving MDP.	2	CO1	BT1
2	Define Policy? Explain the two types of policy.	2	CO1	BT2
3	Illustrate the agent-environment interactions in MDP.	2	CO2	BT2
4	List the two types of Monte Carlo Prediction.	2	CO3	BT1
5	List the differences between on-policy and off-policy methods.	2	CO3	BT1

PART B : Answer Any 1 Following Questions. 10 M

Q.N.	Questions	M	CO	BT
6	Consider a robot learning to navigate through a grid-based environment to reach a designated destination. Using this scenario, identify and illustrate the different elements involved in Reinforcement Learning and explain how the interactions among these elements enables the robot to learn an optimal behaviour.	10	CO1	BT3
7	Demonstrate Value iteration Algorithm.	10	CO2	BT3

PART C: Answer Any 1 Following Questions. 10 M

Q.N.	Questions	M	CO	BT
8	**Identify and classify** the type of RL environment for each given scenario with suitable examples.	10	CO1	BT3

9	**Consider the following 4 × 4 Frozen Lake environment and solve the problem .** **a) Using formulae to find with discount facor.** **b)Using formulae to find without discount facor.** **C1 C2 C3 C4** **R1** S F F F **R2** F H F H **R3** F F F H **R4** H F F G • Where: **S** = Start State **F** = Frozen/Safe State • **H** = Hole **G** = Goal State The agent can move **Up, Down, Left, or Right**. • Assume the following: Reward for reaching the **Goal (G)** = **+10** • Reward for falling into a **Hole (H)** = **−5** • Reward for every normal movement = **−1**	10	CO2	BT3

PART D: Answer Any 1 Following Questions. 10 M				
Q.N.	**Questions**	**M**	**CO**	**BT**
10	Implement First-Visit Monte Carlo Prediction Algorithm	10	CO3	BT3
11	Demonstrate on-policy MC control Agorithm.	10	CO3	BT3

PART E: Answer Any 1 Following Questions. 10 M				
Q.N.	**Questions**	**M**	**CO**	**BT**
12	Apply first-visit and every-visit Monte Carlo prediction methods for any sequence of episodes and calculate the value of a specified state.	10	CO3	BT3
13	Apply the Monte Carlo method to a given reinforcement learning scenario and demonstrate its limitations in estimating state values with a suitable example.	10	CO3	BT3