



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

# PRESIDENCY UNIVERSITY

## BENGALURU

### Mid - Term Examinations – October 2025

**Date:** 08-10-2025

**Time:** 09.30am to 11.00am

<b>School:</b> SOIS	<b>Program:</b> BCA	
<b>Course Code:</b> CSA3074	<b>Course Name:</b> REINFORCEMENT LEARNING	
<b>Semester:</b> V	<b>Max Marks:</b> 50	<b>Weightage:</b> 25%

<b>CO - Levels</b>	<b>CO1</b>	<b>CO2</b>	<b>CO3</b>	<b>CO4</b>	<b>CO5</b>
<b>Marks</b>	<b>26</b>	<b>24</b>			

**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.**

**5Q x 2M=10M**

<b>1</b>	What is meant by a positive reward and a negative reward?	<b>2 Marks</b>	<b>L2</b>	<b>CO1</b>
<b>2</b>	Differentiate Supervised Learning and Reinforcement Learning.	<b>2 Marks</b>	<b>L2</b>	<b>CO1</b>
<b>3</b>	Define a deterministic environment in Reinforcement Learning and give an example.	<b>2 Marks</b>	<b>L2</b>	<b>CO1</b>
<b>4</b>	Write the incremental mean update formula used in Monte Carlo prediction.	<b>2 Marks</b>	<b>L2</b>	<b>CO2</b>
<b>5</b>	Why is Off-Policy MC control considered more flexible than On-Policy MC control?	<b>2 Marks</b>	<b>L2</b>	<b>CO2</b>

## Part B

### Answer the Questions.

**Total Marks 40M**

<b>6.</b>	<b>a.</b>	Explain Markov Decision Process (MDP). Explain its key components with an example.	<b>10 Marks</b>	<b>L4</b>	<b>CO1</b>
-----------	-----------	--	-----------------	-----------	------------

**Or**

<b>7.</b>	<b>a.</b>	Compare deterministic policy and stochastic policy in terms of their effect on value functions.	<b>10 Marks</b>	<b>L4</b>	<b>CO1</b>
-----------	-----------	---	-----------------	-----------	------------

<b>8.</b>	<b>a.</b>	Discuss the role of policies, value functions, and the Bellman equation in solving MDPs, along with the limitations and scope of RL.	<b>10 Marks</b>	<b>L3</b>	<b>CO1</b>
-----------	-----------	--	-----------------	-----------	------------

**Or**

<b>9.</b>	<b>a.</b>	Differentiate model-based vs. model-free learning with Example.	<b>10 Marks</b>	<b>L3</b>	<b>CO1</b>
-----------	-----------	---	-----------------	-----------	------------

<b>10.</b>	<b>a.</b>	Explain Off-Policy Monte Carlo (MC) Control in reinforcement learning. How does it differ from On-Policy MC Control?	<b>10 Marks</b>	<b>L4</b>	<b>CO2</b>
------------	-----------	--	-----------------	-----------	------------

**Or**

<b>11.</b>	<b>a.</b>	Explain the Monte Carlo control method with an epsilon--greedy policy for reinforcement learning.	<b>10 Marks</b>	<b>L3</b>	<b>CO2</b>
------------	-----------	---	-----------------	-----------	------------

<b>12.</b>	<b>a.</b>	Compute final value of states for the given policy.	<b>10 Marks</b>	<b>L3</b>	<b>CO2</b>
------------	-----------	---	-----------------	-----------	------------

**Or**

<b>13.</b>	<b>a.</b>	Illustrate MC prediction Algorithm with Suitable Example.	<b>10 Marks</b>	<b>L4</b>	<b>CO2</b>
------------	-----------	---	-----------------	-----------	------------