Code No.: AD851PE	R20	H.T.No.			8	R					

CMR ENGINEERING COLLEGE: : HYDERABAD
UGC AUTONOMOUS
IV–B.TECH–II–Semester End Examinations (Regular) – April - 2025
REINFORCEMENT LEARNING
(AI&DS)

[Time: 3 Hours] **[Max. Marks: 70]**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 20 marks. Answer all questions in Part A.
Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART-A **(20 Marks)**

1.	a)	What is regret in the context of multi-armed bandits?	[2M]
	b)	What is Thompson Sampling?	[2M]
	c)	What is Bellman's optimality equation?	[2M]
	d)	What is the total reward model and when is it used?	[2M]
	e)	What are the methods used for policy evaluation?	[2M]
	f)	What are the key components of a reinforcement learning system?	[2M]
	g)	What is the goal of Q-learning?	[2M]
	h)	How does SARSA differ from Q-learning?	[2M]
	i)	What happens when $\lambda=0$ in TD(λ)?	[2M]
	j)	What are tilings in tile coding?	[2M]

PART-B **(50 Marks)**

2.	What is a multi-armed bandit problem, and how does it relate to reinforcement learning?	[10M]
	OR	
3.	Describe UCB algorithm to decide which arm to pull in multi-armed bandit scenario.	[10M]
4.	What is value iteration, and how does it differ from policy iteration?	[10M]
	OR	
5.	Differentiate Episodic and Continuing tasks in context of RL.	[10M]
6.	Explain with an example scenario where Monto Carlo control might be applied	[10M]
	OR	
7.	What is Monte Carlo policy evaluation? How does online Monte Carlo differ from batch Monte Carlo learning?	[10M]
8.	Explain the concept of bootstrapping in reinforcement learning. How does it differ from traditional Monte Carlo methods and what are its disadvantageous	[10M]
	OR	
9.	How does fitted Q-Learning leverage the concept of experience reply?	[10M]
10.	Discuss the concept of Eligibility Traces and their role in TD(λ) and n-step Truncated λ-return methods.	[10M]
	OR	
11.	Compare the advantageous and disadvantageous of eligibility traces and linear function approximation in Reinforcement learning.	[10M]

****