

**BONAM VENKATA CHALAMAYYA INSTITUTE OF TECHNOLOGY & SCIENCE
(AUTONOMOUS)**

III - B.Tech II-Semester (Honor) Regular Examinations (BR23), MAY - 2026

REINFORCEMENT LEARNING (CSE-Honors)

Time: 3 hours

Max. Marks: 70

*Question Paper consists of Part-A and Part-B
Answer **ALL** the question in **Part-A and Part-B***

PART-A (10X2 = 20M)

	Marks	CO	BL
1. a) Define Reinforcement Learning and give one real-world example.	(2M)	CO1	BL2
b) List the key elements of Reinforcement Learning.	(2M)	CO1	BL2
c) What is the n-armed bandit problem?	(2M)	CO2	BL4
d) What is the purpose of Upper Confidence Bound (UCB)?	(2M)	CO2	BL4
e) Define Markov Property.	(2M)	CO3	BL4
f) What is a Value Function in MDP?	(2M)	CO3	BL4
g) What is Monte Carlo Prediction?	(2M)	CO4	BL3
h) Define On-policy and Off-policy learning.	(2M)	CO4	BL3
i) Mention two applications of Reinforcement Learning.	(2M)	CO5	BL2
j) What is TD-Gammon?	(2M)	CO5	BL2

PART-B (5X10 = 50M)

2a. Explain the fundamentals of Reinforcement Learning with a suitable example.	5(M)	CO1	BL2
b. Discuss the scope, limitations, and history of Reinforcement Learning.	5(M)		
(OR)			
3a. Explain the elements of Reinforcement Learning in detail.	5(M)	CO1	BL2
b. Describe the Tic-Tac-Toe example used in RL.	5(M)		
4a. Explain action-value methods with equations and examples.	5(M)	CO2	BL4
b. Compare ϵ -greedy and UCB action selection methods.	5(M)		
(OR)			
5a. Explain gradient bandit algorithms.	5(M)	CO2	BL4
b. Discuss non-stationary bandit problems and their solutions.	5(M)		
6a. Explain Markov Decision Process (MDP) with components.	5(M)	CO3	BL4
b. Discuss return and value functions in RL.	5(M)		
(OR)			
7a. Explain optimal value functions and Bellman equations.	5(M)	CO3	BL4
b. Describe the agent-environment interaction in RL.	5(M)		

- | | | | | |
|------|--|------|-----|-----|
| 8a. | Explain Monte Carlo prediction methods. | 5(M) | CO4 | BL3 |
| b. | Discuss Monte Carlo control with exploring starts. | 5(M) | | |
| | (OR) | | | |
| 9a. | Explain off-policy Monte Carlo methods using importance sampling. | 5(M) | CO4 | BL3 |
| b. | Discuss incremental implementation of Monte Carlo methods. | 5(M) | | |
| 10a. | Explain RL applications in robotics and scheduling. | 5(M) | CO5 | BL2 |
| b. | Discuss case study of Samuel's Checkers Player. | 5(M) | | |
| | (OR) | | | |
| 11a. | Explain TD-Gammon and its significance. | 5(M) | CO5 | BL2 |
| b. | Discuss real-world applications like elevator dispatching and job-shop scheduling. | 5(M) | | |
