Course Code: 23CSHT04

BONAM VENKATA CHALAMAYYA INSTITUTE OF TECHNOLOGY & SCIENCE

(AUTONOMOUS)

III-B. TechII-Semester Regular Examinations (BR23), April- 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-AandPart-B***

PART–A (10X2 = 20M)

		Marks	CO	BL
1 a)	Discuss the scope and limitations of Reinforcement Learning in real-world applications.	(2M)	CO1	BL2
b)	Trace the history and evolution of Reinforcement Learning.	(2M)	CO1	BL2
c)	Illustrate incremental implementation of action-value methods	(2M)	CO2	BL3
d)	Discuss optimistic initial values.	(2M)	CO2	BL2
e)	Describe optimal value functions	(2M)	CO3	BL3
f)	Define returns and discount factor.	(2M)	CO3	BL2
g)	Discuss exploring starts assumption.	(2M)	CO4	BL2
h)	Explain incremental implementation of Monte Carlo algorithms.	(2M)	CO4	BL4
i)	Discuss dynamic channel allocation using RL in communication networks.	(2M)	CO5	BL2
j)	Explain job-shop scheduling using reinforcement learning techniques.	(2M)	CO5	BL4

PART–B (5X10 = 50M)

2a.	Explain the concept of Reinforcement Learning. Compare it with supervised and unsupervised learning with suitable examples.	10(M)	CO1	BL2
	(OR)			
3a.	Explain the Tic-Tac-Toe example and show how learning occurs through interaction.	10(M)	CO1	BL2

4a.	Define the n-armed bandit problem. Explain the exploration–exploitation dilemma.	5(M)	CO2	BL3
b.	Analyze the Upper Confidence Bound (UCB) action selection method and explain its advantages.	5(M)	CO2	BL4
	(OR)			
5a.	Describe Gradient Bandit algorithms.	5(M)	CO2	BL3
b	Compare stationary and non-stationary bandit problems.	5(M)	CO2	BL3

6a.	Describe Markov Decision Processes formally with states, actions, rewards, and transitions	10(M)	CO3	BL3
	(OR)			

7a.	Define Markov Property. Why is it important in Reinforcement Learning?	5(M)	CO3	BL3
b.	Explain the Agent–Environment interface in MDPs with diagram.	5(M)	CO3	BL3

8a.	Explain Monte Carlo prediction methods with first-visit and every-visit approaches.	10(M)	CO4	BL4
	(OR)			
9a.	Explain Monte Carlo control for policy improvement.	10(M)	CO4	BL2

10a.	Explain the working of TD-Gammon and how RL was applied successfully.	5(M)	CO5	BL2
b	Describe the Acrobot problem and learning control strategies.	5(M)	CO5	BL3
	(OR)			
11a.	Discuss Samuel's Checkers Player and its contribution to early Reinforcement Learning research.	10(M)	CO5	BL2

****

Staff

HOD