Course Code: 23AM6T03

BONAM VENKATA CHALAMAYYA INSTITUTE OF TECHNOLOGY & SCIENCE

(AUTONOMOUS)

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

REINFORCEMENT LEARNING (AI&ML)

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	(2M)	CO1	BL1
b)	What is the exploration–exploitation tradeoff	(2M)	CO1	BL1
c)	What is an n-armed bandit problem?	(2M)	CO2	BL1
d)	What is a stationary and non-stationary bandit problem?	(2M)	CO2	BL1
e)	Define state in RL.	(2M)	CO3	BL1
f)	What is the Markov property?	(2M)	CO3	BL1
g)	What is Monte Carlo control?	(2M)	CO4	BL1
h)	What is truncated return?	(2M)	CO4	BL1
i)	What is Samuel's checkers player?	(2M)	CO5	BL1
j)	What is job-shop scheduling?	(2M)	CO5	BL1

PART–B (5X10 = 50M)

2.	Explain the elements of Reinforcement Learning in detail.	10(M)	CO1	BL2
	(OR)			
3.	Explain the history of Reinforcement Learning.	10(M)	CO1	BL2

4	Explain Incremental Implementation in Action–Value Methods. Illustrate with an example.	10(M)	CO2	BL2
	(OR)			
5	Describe Action–Value estimation and the Greedy and ε-Greedy methods.	10(M)	CO2	BL2

6	Explain the agent–environment interface in Reinforcement Learning.	10(M)	CO3	BL2
	(OR)			
7	Explain Finite Markov Decision Processes in detail.	10(M)	CO3	BL2

8.	Describe Monte Carlo estimation of action values.	10(M)	CO4	BL2
	(OR)			
9.	Explain incremental implementation in Monte Carlo methods.	10(M)	CO4	BL2

10.	Explain the working of TD-Gammon.	10(M)	CO5	BL2
	(OR)			
11.	Describe Reinforcement Learning application in dynamic channel allocation.	10(M)	CO5	BL2

****