BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(MID SEMESTER EXAMINATION)

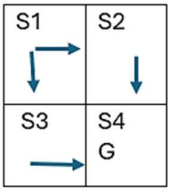
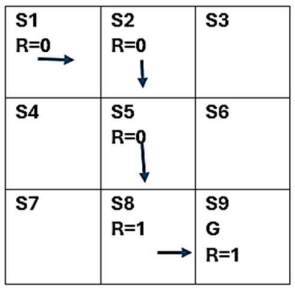
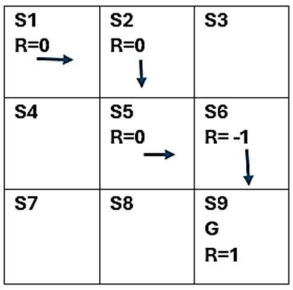
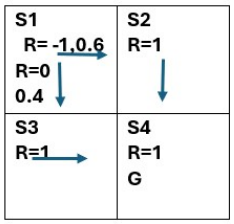
CLASS: BTECH **SEMESTER : VII**
BRANCH: AIML **SESSION : MO/2025**

SUBJECT:AI401 REINFORECEMENT LEARNING

TIME: 02 Hours **FULL MARKS: 25**

INSTRUCTIONS:

1. The question paper contains 5 questions each of 5 marks and total 25 marks.
2. Attempt all questions.
3. The missing data, if any, may be assumed suitably.
***4*. Tables/Data handbook/Graph paper etc., if applicable, will be supplied to the candidates**

			CO	BL
Q.1(a)	Explain with neat block diagram the framework of reinforcement learning	[2]	1	2
Q.1(b)	 Design the tabular representation of a policy for the above grid .the probabilities are 0.5 for a1(R) and a2(D) respectively, the other actions as policies are a3(U),a4(L),a5(still).	[3]	1	3
Q.2(a)	 Fig1(a)  Fig1(b) Determine the best policy by considering the finite length and infinite length trajectory. Assume ¥ as discount factor.	[2]	1	3
Q.2(b)	Explain mathematically why Markov property refers to memoryless property of a stochastic process. What is the difference between Markov Process and Markov Decision Process.	[3]	1	3
Q.3(a)	Is the reward a function of next state ? Justify with a mathematical equation	[2]	1	3
Q.3(b)	Derive Bellman's equation from $G_t=R_{t+1}+YR_{t+2}+Y^2R_{t+3}+$…….and reduce it to $$v_\pi(s) = \sum_{a\in\mathcal{A}} \pi(a\|s) \sum_{s'\in\mathcal{S}} p(s'\|s,a)\,[r(s') + \gamma v_\pi(s')]\,.$$	[3]	2	3
Q.4(a)	What is the relationship between state values and action values?	[2]	2	2
Q.4(b)	 Determine the matrix form of Bellman's equations for the above grid. For actions refer question 1(b)	[3]	2	3
Q.5(a)	Derive action values in terms of state values	[2]	2	1
Q.5(b)	Why do we care about the values of the actions that a given policy cannot select? Referring to the figure given in question 4(b). Determine action values of state s1 in terms of state values .For actions refer question 1(b)	[3]	2	3

:::::18/09/2025:::::M