

BE Sem 8, Choice based, R2019-20, C-Scheme, May 2025
 Date :- 15/5/2025, Branch- INFT

Time: 03 Hours

Marks: 80

- Note: 1) Q1 is compulsory
 2) Attempt any three questions from the remaining questions
 3) Figures to the right indicate full marks

Q 1

- a) Explain elements of Reinforcement Learning. (5)
 b) Suppose $\gamma = 0.5$ and the following sequence of rewards is received $R_1 = 1, R_2 = 2, R_3 = 6, R_4 = 3,$ and $R_5 = 2,$ with $T = 5$. What are G_0, G_1, \dots, G_5 ? (5)
 c) What is policy iteration? Explain policy iteration algorithm. (5)
 d) Explain in brief Markov Decision Process (MDP). (5)

Q 2

- a) Explain Monte Carlo Estimation of action value with suitable example. (10)
 b) Explain n-Armed Bandit Problem. (10)

Q 3

- a) Explain incremental implementation of estimate of action values. (10)
 b) Explain Q-learning algorithm to learn optimal action value function. (10)

Q 4

- a) Describe the application of reinforcement learning to the real world problem of elevator dispatching. (10)
 b) Give similarities and differences between Q-learning and SARSA algorithms. (10)

Q 5

- a) Explain Agent-Environment interface. (10)
 b) Discuss various methods to handle non-stationary bandit problems. (10)

Q 6

- a) Describe asynchronous dynamic programming with an example. (10)
 b) Describe TD control using Q-learning (10)