



**SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR
(AUTONOMOUS)**

Siddharth Nagar, Narayanavanam Road – 517583

QUESTION BANK (DESCRIPTIVE)



Subject with Code: REINFORCEMENT LEARNING(23CS1117)

Course & Branch: B.Tech- CAD

Regulation: R23

Year & Sem: IV-B.Tech & I-Sem

UNIT – I

The Reinforcement Learning Problem

1	a)	Define Reinforcement Learning.	[L1][CO1]	[2M]
	b)	List the elements of Reinforcement Learning.	[L1][CO1]	[2M]
	c)	What are the limitations of Reinforcement Learning.	[L1][CO1]	[2M]
	d)	What is meant by the exploration–exploitation trade-off .	[L1][CO1]	[2M]
	e)	Why is Tic-Tac-Toe considered an extended example in Reinforcement Learning.	[L2][CO1]	[2M]
2	a)	Discuss the role of trial-and-error learning in Reinforcement Learning.	[L2][CO1]	[5M]
	b)	Illustrate the elements of Reinforcement Learning with a neat diagram.	[L3][CO1]	[5M]
3		Analyze the Reinforcement Learning framework and discuss its characteristics and applications using a neat diagram.	[L4][CO1]	[10M]
4	a)	Illustrate the characteristics of a Reinforcement Learning problem with suitable examples.	[L2][CO1]	[5M]
	b)	Describe the interaction between an agent and its environment.	[L2][CO1]	[5M]
5	a)	Compare Reinforcement Learning with Supervised and Unsupervised Learning.	[L4][CO1]	[5M]
	b)	Evaluate the suitability of Reinforcement Learning for a given learning problem.	[L2][CO1]	[5M]
6	a)	Explain how rewards influence the learning behavior of an agent.	[L2][CO1]	[5M]
	b)	Discuss the concept of reward in Reinforcement Learning.	[L2][CO1]	[5M]
7	a)	Explain the major components involved in a Reinforcement Learning problem.	[L2][CO1]	[5M]
	b)	Analyze the scope and limitations of Reinforcement Learning.	[L4][CO1]	[5M]
8	a)	Describe the Tic-Tac-Toe example used in Reinforcement Learning.	[L2][CO1]	[5M]
	b)	Discuss the history and evolution of Reinforcement Learning.	[L2][CO1]	[5M]
9		Analyze the complete Reinforcement Learning problem formulation and discuss its importance in Artificial Intelligence.	[L4][CO1]	[10M]
10	a)	Explain the exploration and exploitation concepts with examples.	[L2][CO1]	[5M]
	b)	Explain the importance of Reinforcement Learning in Artificial Intelligence with suitable examples.	[L2][CO1]	[5M]
11		Illustrate the Reinforcement Learning process with the help of a real-world example.	[L3][CO1]	[10M]

UNIT – II

Multi-arm Bandits

1	a)	Define the n-Armed Bandit Problem.	[L1][CO2]	[2M]
	b)	What is an Action-Value Method.	[L1][CO2]	[2M]
	c)	What is meant by Nonstationary Problem.	[L1][CO2]	[2M]
	d)	Define Upper-Confidence-Bound (UCB) Action Selection.	[L1][CO2]	[2M]
	e)	What are Contextual Bandits.	[L1][CO2]	[2M]
2		Illustrate the n-Armed Bandit Problem with a suitable example and discuss its significance.	[L3][CO2]	[10M]
3	a)	Explain the exploration and exploitation dilemma in the n-Armed Bandit Problem.	[L2][CO2]	[5M]
	b)	Describe the role of preferences in Gradient Bandits.	[L2][CO2]	[5M]
4		Describe the Action-Value Methods for solving Multi-arm Bandit problems with mathematical formulation.	[L2][CO2]	[10M]
5	a)	Explain Incremental Implementation with the update equation.	[L2][CO2]	[5M]
	b)	Analyze how a Bandit method can be adapted to a nonstationary environment.	[L4][CO2]	[5M]
6	a)	Examine the role of Learning Rate in updating action-value estimates in nonstationary Bandit problems.	[L4][CO2]	[5M]
	b)	Explain how action values can be updated incrementally.	[L2][CO2]	[5M]
7	a)	Apply the concept of Sample-Average Method for estimating action values in Multi-Armed Bandit problems.	[L4][CO2]	[5M]
	b)	Discuss about the usefulness of Optimistic Initial Values in solving the exploration problem.	[L2][CO2]	[5M]
8	a)	Describe Optimistic Initial Values and their significance.	[L2][CO2]	[5M]
	b)	Explain about the Upper-Confidence-Bound (UCB) Action Selection method.	[L2][CO2]	[5M]
9	a)	Compare ϵ -Greedy, Optimistic Initial Values, and UCB methods.	[L4][CO2]	[5M]
	b)	Explain different exploration strategies used in Multi-arm Bandit problems.	[L2][CO2]	[5M]
10	a)	Describe about the Gradient Bandit Algorithm.	[L2][CO2]	[5M]
	b)	Analyze the problem of selecting actions when reward estimates are uncertain.	[L4][CO2]	[5M]
11		Discuss Associative Search (Contextual Bandits) with practical applications.	[L2][CO2]	[10M]

UNIT – III

Finite Markov Decision Processes

1	a)	Define a Markov Decision Process.	[L1][CO3]	[2M]
	b)	What is the Markov Property.	[L1][CO3]	[2M]
	c)	Define Value Function.	[L1][CO3]	[2M]
	d)	What is meant by Return in Reinforcement Learning.	[L2][CO3]	[2M]
	e)	Define Optimal Value Function.	[L1][CO3]	[2M]
2	a)	Explain the Agent–Environment Interface with a neat diagram.	[L2][CO3]	[5M]
	b)	Discuss Goals and Rewards in Reinforcement Learning.	[L2][CO3]	[5M]
3	a)	Describe the importance of rewards and returns in decision making.	[L2][CO3]	[5M]
	b)	Explain how the Agent–Environment interaction can be represented using an MDP.	[L2][CO3]	[5M]
4	a)	Apply the concept of return to explain how future rewards affect decision making.	[L2][CO3]	[5M]
	b)	Explain the concept of Value Functions.	[L2][CO3]	[5M]
5	Differentiate between Episodic Tasks and Continuing Tasks using the unified notation with suitable examples.		[L4][CO3]	[10M]
6	a)	Analyze why the Markov Property is important for decision making	[L4][CO3]	[5M]
	b)	Explain Optimal Value Functions with suitable examples.	[L2][CO3]	[5M]
7	Examine the components of a Markov Decision Process (MDP) and analyze how they contribute to decision-making in Reinforcement Learning with suitable examples.		[L4][CO3]	[10M]
8	a)	Explain the relationship between value functions and optimal value functions.	[L2][CO3]	[5M]
	b)	Describe the importance of the Markov Decision Process in Reinforcement Learning.	[L2][CO3]	[5M]
9	Illustrate State-Value Functions and Action-Value Functions with mathematical expressions.		[L3][CO3]	[10M]
10	Analyze the role of optimality and approximation in solving finite MDPs.		[L4][CO3]	[10M]
11	Compare finite Markov Decision Processes with Multi-arm Bandit Problems.		[L4][CO3]	[10M]

UNIT-IV

Monte Carlo Methods

1	a)	Define Monte Carlo Prediction.	[L1][CO4]	[2M]
	b)	What is meant by Importance Sampling.	[L1][CO4]	[2M]
	c)	Define Off-Policy Learning.	[L1][CO4]	[2M]
	d)	What do you mean by Monte Carlo Control.	[L1][CO4]	[2M]
	e)	Define Exploring Starts.	[L1][CO4]	[2M]
2	a)	Explain Monte Carlo Prediction with an example.	[L2][CO4]	[5M]
	b)	Discuss Monte Carlo Estimation of Action Values.	[L2][CO4]	[5M]
3	a)	Describe Monte Carlo Control and its working procedure.	[L2][CO4]	[5M]
	b)	Analyze the role of exploring starts in Monte Carlo Control.	[L4][CO4]	[5M]
4	a)	Differentiate Monte Carlo Prediction and Monte Carlo Control.	[L4][CO4]	[5M]
	b)	Analyze the use of Importance Sampling for truncated returns in Monte Carlo methods.	[L4][CO4]	[5M]
5	a)	Explain about the usefulness of Off-Policy Monte Carlo Control.	[L2][CO4]	[5M]
	b)	Discuss about why Exploring Starts are considered in Monte Carlo Control.	[L2][CO4]	[5M]
6	a)	Illustrate the concept of Off-Policy Prediction using Importance Sampling.	[L3][CO4]	[5M]
	b)	Explain the basic principle of Monte Carlo methods in Reinforcement Learning.	[L2][CO4]	[5M]
7	a)	Compare On-Policy and Off-Policy Monte Carlo methods.	[L4][CO4]	[5M]
	b)	Discuss how the action values can be estimated using Monte Carlo methods.	[L2][CO4]	[5M]
8	a)	Compare and evaluate Monte Carlo Prediction and Monte Carlo Control methods, highlighting their effectiveness in different scenarios.	[L4][CO4]	[5M]
	b)	Differentiate between conventional and incremental implementation of Monte Carlo estimation.	[L4][CO4]	[5M]
9		Examine the Incremental Implementation in Monte Carlo Methods through mathematical derivation and analyze its computational advantages.	[L4][CO4]	[10M]
10		Compare and analyze Off-Policy Monte Carlo Control with On-Policy Monte Carlo Control, highlighting its working mechanism.	[L4][CO4]	[10M]
11	a)	Differentiate between Monte Carlo Estimation of Action Values and Monte Carlo Prediction with suitable examples.	[L4][CO4]	[5M]
	b)	Explain the major techniques included under Monte Carlo Methods.	[L2][CO4]	[5M]

UNIT-V

Applications and Case Studies

1	a)	What is TD-Gammon.	[L1][CO5]	[2M]
	b)	Who developed Samuel's Checkers Player.	[L1][CO5]	[2M]
	c)	What is the Acrobot problem.	[L2][CO5]	[2M]
	d)	Define Dynamic Channel Allocation.	[L1][CO6]	[2M]
	e)	What is Job-Shop Scheduling.	[L1][CO6]	[2M]
2	a)	Explain TD-Gammon as a successful application of Reinforcement Learning.	[L2][CO5]	[5M]
	b)	Discuss Samuel's Checkers Player and its significance.	[L2][CO5]	[5M]
3		Analyze the similarities and differences between TD-Gammon and Samuel's Checkers Player in terms of learning strategies and effectiveness.	[L4][CO6]	[10M]
4	a)	Illustrate the Acrobot problem and demonstrate how Reinforcement Learning can be used to solve it.	[L3][CO5]	[5M]
	b)	Describe the application of Reinforcement Learning in Elevator Dispatching.	[L2][CO5]	[5M]
5	a)	Apply the principles of Reinforcement Learning to Dynamic Channel Allocation and discuss its benefits.	[L3][CO5]	[5M]
	b)	. Compare the Acrobot and Elevator Dispatching applications.	[L4][CO5]	[5M]
6		Develop a Reinforcement Learning approach for Elevator Dispatching and evaluate its effectiveness in reducing waiting time and improving service efficiency.	[L6][CO6]	[10M]
7		Examine the application of Reinforcement Learning in Job-Shop Scheduling and analyze its advantages and challenges.	[L4][CO6]	[10M]
8	a)	Differentiate between TD-Gammon and Samuel's Checkers Player.	[L4][CO6]	[5M]
	b)	Explain how Reinforcement Learning is applied in industrial automation.	[L6][CO6]	[5M]
9		Formulate a comprehensive analysis of Reinforcement Learning case studies and propose their significance in advancing modern Artificial Intelligence.	[L6][CO6]	[10M]
10		Construct a Reinforcement Learning model for solving real-world optimization problems and discuss its implementation strategy.	[L6][CO6]	[10M]
11	a)	Analyze the different Reinforcement Learning applications covered in the case studies.	[L4][CO6]	[5M]
	b)	Discuss real-world applications of Reinforcement Learning in gaming and robotics.	[L2][CO6]	[5M]

Prepared By Dr. E. Murali, Dr. N. Babu, Mr. V.M. Vinayagam