

Round 3: Reinforcement learning

Introduction: Imagine teaching a dog new tricks. When the dog does something right, you give it a reward; when it makes a mistake, you don't reward it. Over time, the dog learns which actions lead to treats and repeats those behaviors. This is exactly how reinforcement learning works! In reinforcement learning, an AI agent learns by interacting with an environment, just like the dog. When it takes an action, it gets feedback in the form of rewards or penalties.

Question	Options	Type	Information	Hint
Question 1: Which scenario best represents the reward mechanism in a Reinforcement Learning system?	<ol style="list-style-type: none"> A drone learning to maintain stability in varying wind conditions by receiving altitude and orientation feedback signals A factory robot learning assembly patterns from a database of correct assemblies A trading algorithm using historical price patterns to identify market trends A quality control system detecting defects based on images of good and bad products 	UC	- Dynamic responses	"Think about dynamic responses."
		PC	<ul style="list-style-type: none"> - Learning through actual experience, not just past data - Immediate feedback is essential - The agent should learn without being shown correct examples 	"Consider these questions: Which scenario involves learning through actual experience rather than from previous data? Does it get immediate feedback that helps it improve right away? Can it get better during operation without being shown correct examples? "
		NC	<ul style="list-style-type: none"> - Reinforcement Learning involves real-time learning and adaptation - Analogy: Riding a bicycle on a windy day 	"Imagine riding a bicycle on a windy day - you constantly adjust your balance based on wind gusts, your speed, and the road conditions. Each adjustment affects your stability, and you learn from each moment how to stay upright. Which scenario involves similar real-time learning and adaptation to changing conditions?"

Question	Options	Type	Information	Hint
Question 2: What is the difference between an agent and its environment in Reinforcement Learning?	<ol style="list-style-type: none"> The agent takes actions, and the environment provides feedback based on these actions. The agent collects labeled data while the environment provides unlabeled data The agent stores data while the environment processes it The environment learns from the agent's actions, not the other way around. 	UC	<ul style="list-style-type: none"> - Focus on interaction between the agent and the environment. 	"Think about interaction."
		PC	<ul style="list-style-type: none"> - UC Information - The agent takes actions. - The environment responds. 	"The agent makes moves and the environment responds to moves."
		NC	<ul style="list-style-type: none"> - UC + PC Information - The agent learns from the outcomes. - Full explanation with analogy 	"Imagine playing tennis with a wall: the agent decide how to hit the ball, while the wall and court (the environment) determine where the ball goes and if you scored. You make choices and learn, while the environment shows you what happens as a result. Which option describes this kind of relationship?"

Question	Options	Type	Information	Hint
Question 3: A smart traffic system uses Reinforcement Learning to control traffic lights in a city. Which description best matches how it actually works?	<ol style="list-style-type: none"> The traffic light controller looks at how many cars are waiting and how long they've been waiting. It tries different signal timing patterns and learns from seeing how these changes affect the average waiting time. The cameras watch traffic and predict busy periods based on past data. They get rewarded for guessing rush hours correctly. The system records how long lights stay green and stores traffic data. It aims to save storage space efficiently. The analyzer looks at different types of vehicles and groups similar traffic patterns together. 	UC	<ul style="list-style-type: none"> - Focus on control—the system should actively adjust traffic lights rather than just observing. 	“Think about who's in control.”
		PC	<ul style="list-style-type: none"> - UC Information - The system should actively try to improve traffic flow, not just analyze or store data. - Trial-and-error learning 	“Which option shows a system that's actively trying to improve things rather than just analyzing or storing information?”
		NC	<ul style="list-style-type: none"> - UC + PC Information - It sees how traffic responds and learns from it - Full explanation with parallel example 	“Think of it like playing a game: you see the current traffic situation, you adjust the traffic lights, then you see if cars move faster. You keep trying different timings until you find what works best. Which option describes this complete cycle?”