Winter – SEMESTER
Course Code: MCSE606L
Course-Title:- Cognitive Robotics
DIGITAL ASSIGNMENT - 1

Name: Nidhi Singh
Reg. No:22MAI0015

Slot- F2/TF2

Faculty: Dr. GOUTAM MAJUMDER - SCOPE

---

1. How do agents balance exploration and exploitation in reinforcement learning, and what are some of the strategies used to address this challenge?

① How do agents balance exploration and exploitation in reinforcement learning, and what are some of the strategies used to address this challenge?

Ans

In reinfoncement learning, agents need to balance exploration and exploitation to achieve optimal performance. Exploration refers to the process of discovering new information about the envirnment by taking actions that the agents has not yet tried, while exploitation refers to the process of maximizing rewards by taking actions that the agent already knows will yield high rewards.

The exploration - exploitation trade-off is a fundamental challenge in reinforcement learning because agents must explore enough to discover new actions that may yield higher rewards, but they also need to exploit actions that have already proven to be successful to maximize their rewards.

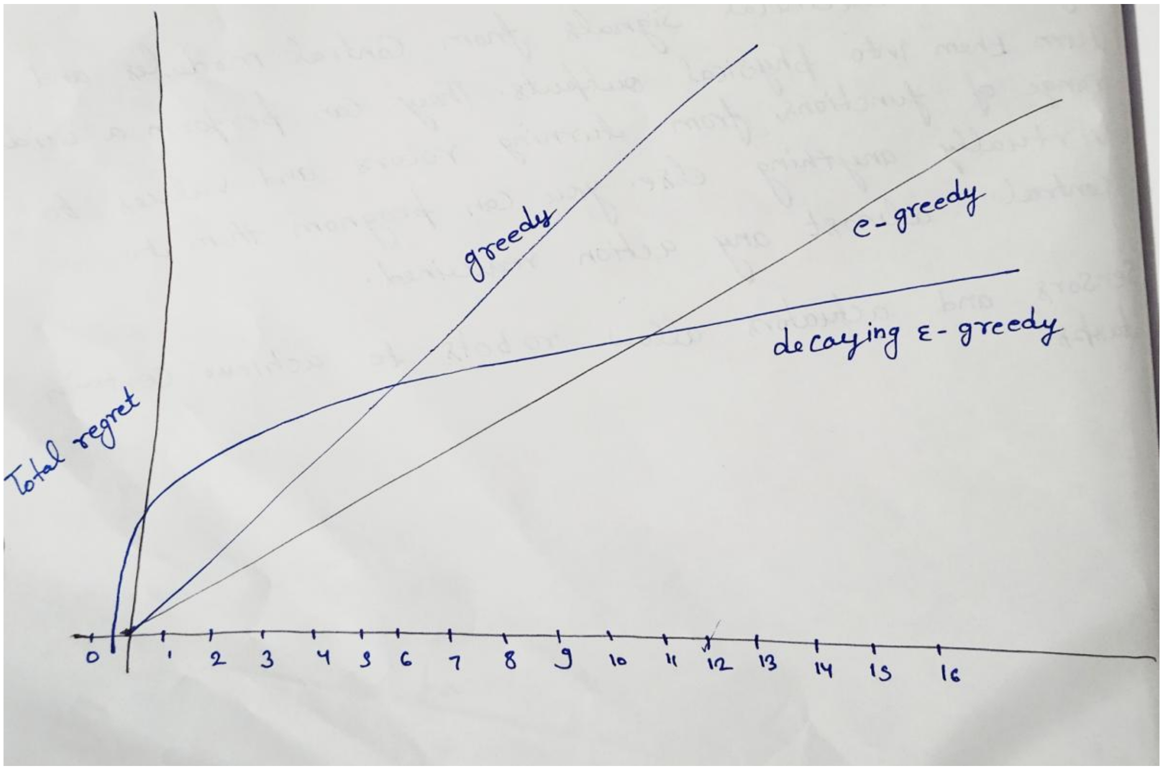
There are several strategies that agents can use to address this challenge:

Classic Exploration Strategies

There are a Simple and Commonly Used strategies where the agent chooses the random action with probability to address this challenge:

① Epsilon - greedy

The agent does random exploration occasionally with probability ϵ and takes the optimal action most of the time with probability $1-\epsilon$.
Epsilon greedy policy is defined as a technique to maintain a balance between exploitation and exploration. However, to choose between exploration and exploitation, a very Simple method is to Select randomly. This can be done by choosing exploitation most of the time with a little exploration.



This exploration rate defines the probability of exploring the environment by the agent rather than exploiting it. It also ensures that the agent will start by exploring the environment with $\varepsilon = 1$.

A machine with the highest Current average payout is Selected with

$$\boxed{\text{Probability} = (1 - \text{epsilon}) + (\text{epsilon} / K)}$$

And, machines that don't have the highest Current payout average are Selected with

$$\boxed{\text{Probability} = \text{epsilon} / K}$$

Over time, the best paying machine will be played more and more often.

② upper Confidence Bound

upper Confidence bound $Q^{\wedge}t(a)+U^{\wedge}t(a)$, where $Q^{\wedge}t(a)$ is the average rewards associated with action a up to time t and $U^{\wedge}t(a)$ is a function reversely proportional to how many times action a has been taken. maximising ~~time~~ this upper Confidence bound is a strategy employed by the agents to move towards the goals. The popular AlphaGo Zero program of DeepMind used Monte Carlo Tree Search, which, in turn, uses a neural network to guide the simulations. Each Simultation in the Go game iteratively selects moves that maximise the upper Confidence bound.

Thompson Sampling :-

This is a Bayesian appraach to reinforcement learning that Samples actions from a posterior distribution over the values of each action. The agent Selects the action with the highest probability of being the best action, which balances exploration and exploitation by choosing actions that have high expected values and high uncertainty.

Thompson Sampling is an algorithm for online decision problems where actions are taken sequentially in a manner that must balance between exploiting what is known to maximize immediate performance and investing to accumulate new information that may improve future performance. The agent keeps track of the probability of optimal actions and samples from this distribution.

Softmax Action Selection :-

This strategy selects actions based on their estimated values but also introduces a temperature parameter that controls the degree of exploration. Higher temperatures lead to more exploration, while lower temperatures lead to more exploitation

Overall, the exploration - exploitation trade-off is a crucial challenge in reinforcement learning, and there are several strategies that agents can use to balance these competing demands and achieve optimal performance.

Exploitation and exploration techniques in reinforcement machine learning have enhanced various type of parameters Such as improved performance, increased learning rate, better decision making etc.

2. How do sensors and actuators work together in robotics to enable robots to interact with their environment and perform tasks?

② How do sensors and actuators work together in robotics to enables robots to interact with their envirnment and perform tasks?

Ans

Sensors and actuators are two essential components that work together in Robotics to enable robots to interact with their envirnment & perform various tasks, Sensors detect the envirnmental conditions and provide feedback to the robot's Control System, while actuators convert this feed back into actions to Control the robot's movements.

For example, Sensors Can detect the presence of obstacles or changes in light, temperature, or Sound, and Send this information to the robot's Control System. The Control System then processes this information and Sends Signals to the appropriate actuators, such as motors or hydraulic Cylinders, to move the robot's joints or manipulate its tools to respond to the detected changes.

In industrial applications, robots are often equipped with a variety of Sensors, Such as proximity Sensors, force sensors, and vision Sensors, to detect and respond to their environment accurately. These Sensors Can detect the position and orientation of objects, measure forces and torques, and identify visual features of the environment.

Actuators, on the other hand, provide the physical motion needed to perform tasks Such as lifting and Carrying objects, welding, painting, or assembling parts. The actuators used in robotics can range from Simple electric motors to more complex hydraulic or pneumatic systems, depending on the application and the reauired precision and force.

overall, Sensors and actuators work together in robotics to enable robots to penceive and respond to their envirnment, allowing them to perform various tasks autonomously and accurately.

Sensors

A Sensor is an electrical instrument that monitors and measures physical aspects of an environment and sends an electrical signal to a Control Center when certain pre-determined Conditions are detected. Sensors turn physical inputs into electrical signals that are output of the Control center.

Actuators

They take electrical signals from Control modules and turn them into physical outputs. They can perform a wide range of functions, from turning rotors and values to virtually anything else. you can prognam them to Control almost any action reauired.

Sensors and actuators allow robots to achieve certain tasks.

Actuators and Sensors together

For instance, a gas furnace will have a thermocouple that monitors the heat from the pilot light. As long as the pilot light keeps burning, the thermocouple generates a current. The greater the heat, the higher the voltage will be.

In furnaces, the gas shutoff value is connected to the thermocouple. As long as current stops such as if the pilot light goes out. the valve shuts, preventing gas from accumulating in the furnace and reducing the risk of an explosion.

In this case, the sensor provides both the energy and the signal for the actuator. In other system, the setup may be more complex with multiple sensors and actuators working in tandem to perform a given task. However the basic principle remains the sensor provides a signal and the actuator adjusts based on that signal.