Advanced Deep Learning

Zeham Management Technologies BootCamp
by SDAIA

September 2nd, 2024

الهيئة السعودية للبيانات
والذكاء الاصطناعي
Saudi Data & AI Authority

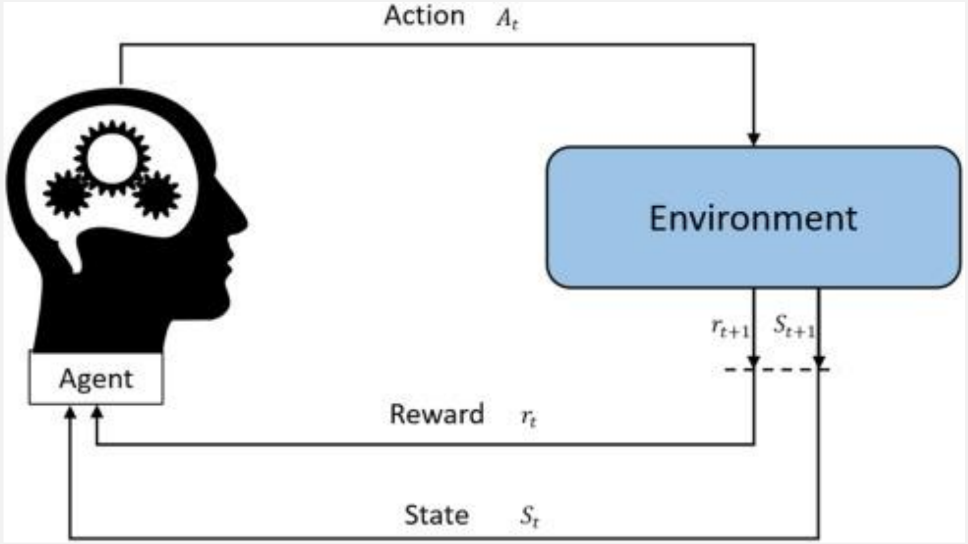
Agenda

- Introduction to Markov Decision Processes
- MDP Example

Introduction to Markov Decision Processes

Introduction to Markov Decision Processes

MDP is a framework for modeling decision-making where outcomes are partly random and partly under control of a decision-maker.



Introduction to Markov Decision Processes

Consists of states, actions, transitions, Policy and rewards.

Why are MDPs Important?

- Widely used in fields such as robotics, economics, and artificial intelligence.
- Helps solve problems where decisions lead to uncertain outcomes over time.

MDPs Components

Components:

- **States (S):** Different configurations or situations (e.g., traffic conditions at an intersection).
- **Actions (A):** The set of all possible decisions or actions available (e.g., changing traffic light timing).
- **Transition probabilities (P):** Probabilities of moving from one state to another, given an action (e.g., probability of traffic becoming congested after a red light).
- **Rewards (R):** A numerical value representing the benefit of an action in a specific state (e.g., smooth traffic flow, reduced congestion).
- **Policy (π):** A strategy that tells the agent which action to take in each state.

MDPs Mathematical Representation

Mathematical Representation of MDP:

- MDP as a Tuple: (S, A, P, R, γ)

Where γ is the discount factor.

- Bellman Equation:

Used to compute the optimal policy and value function.

Key Concepts in MDPs

- **Value Function (V):**

Measures the expected cumulative reward that can be achieved starting from a particular state and following a certain policy.

In traffic management, the value function represents the long-term benefit of a specific traffic control strategy given the current traffic situation.

- **Bellman Equation:**

A foundational equation in MDPs that relates the value of a state to the values of subsequent states, weighted by transition probabilities and rewards.

In a traffic context, it calculates the expected value of future traffic conditions based on current decisions about signal timings.

Key Concepts in MDPs

- **Optimal Policy (π):**

The goal of an MDP is to find the optimal policy, π^*, which maximizes the total expected reward over time (e.g., minimizing traffic congestion).

Techniques such as Value Iteration or Policy Iteration are commonly used to compute the optimal policy.

MDP Example

MDP Example

Autonomous Vehicles: In autonomous driving, MDPs are used to make decisions about route planning, such as whether to take a high-traffic route or wait at a signal. The system evaluates possible outcomes (e.g., traffic congestion or smooth flow) and decides the best action based on expected rewards (minimizing travel time).

MDP Example

Supply Chain Optimization: MDPs are used in supply chain management to decide how much inventory to order at each stage, considering the uncertainty in demand and supply disruptions. The goal is to maximize long-term profitability while minimizing stockouts.

MDP Example in Traffic Management

Traffic Signal Control: *(Detailed Example)*

States: Traffic conditions at an intersection (e.g., low, medium, high congestion).

Actions: Timing changes for red, yellow, and green lights.

Transitions: Probability of congestion changing after adjusting signal timing.

Rewards: Minimizing congestion or maximizing smooth flow of traffic.

MDP Example in Traffic Management

Traffic Signal Control:

Consider a simple traffic light system at an intersection.

The **Bellman equation** computes the value of each state by estimating the expected future congestion based on the current decision.

Tutorial

10-Advanced Deep Learning/LAB/ Markov_Decision_Processes.ipynb

Conclusion

- MDP is a powerful tool for modeling sequential decision-making.
- It provides a structured way to handle uncertainty and optimize outcomes.

Thank you!

الهيئة السعودية للبيانات والذكاء الاصطناعي
Saudi Data & AI Authority