**EE422/CS421 Introduction to Robotics**

**Spring 2025**

Group Members

Fadime Yaren Durmuş 200201003

Mustafa Kağan Aytaç 200202026

**Report – Presentation 3**

**Reinforcement Learning**

Reinforcement learning (RL) is a branch of machine learning where an agent learns how to make decisions by interacting with an environment. Instead of being told what the correct actions are, the agent learns from the consequences of its actions through trial-and-error. This method mimics how humans and animals learn: by experimenting, receiving feedback, and improving based on outcomes.

At the heart of reinforcement learning lies a loop of continuous learning: the agent takes an action, the environment responds with a new state and a reward, and the agent uses this feedback to adjust its future actions. Over time, the agent's behavior becomes optimized to achieve long-term rewards.

**Core Concepts of RL**

* **Agent**: The learner or decision-maker.
* **Environment**: The external system the agent interacts with.
* **State**: A snapshot of the current situation.
* **Action**: A move or decision the agent can make.
* **Reward**: Feedback from the environment based on the action taken.

The agent's goal is to learn a **policy** that maps each state to the best possible action, aiming to maximize the total cumulative reward.

**How RL Works**

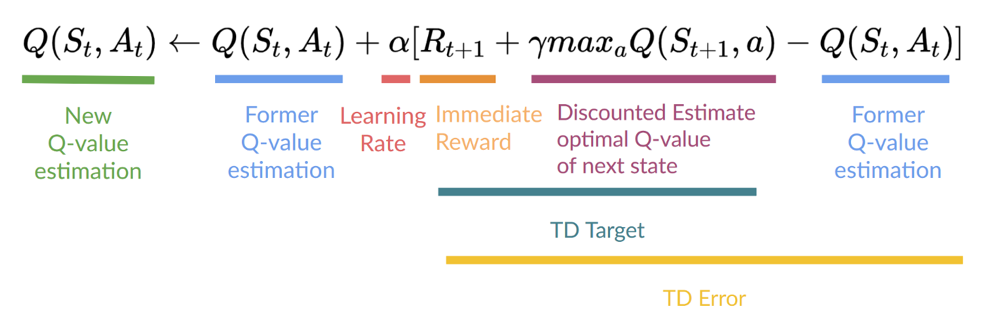
1. The agent starts in an initial state.
2. It selects an action based on a policy (initially random or exploratory).
3. The environment returns a new state and a reward.
4. The agent updates its knowledge.
5. Steps 2–4 repeat over many episodes.

A diagram of a function

AI-generated content may be incorrect.

**Q-Learning**

Q-learning is one of the most widely used RL algorithms. It helps the agent estimate the value of an action in a specific state using a table called the Q-table. This table gets updated with each experience using the following formula:



Where:

* ss: current state
* aa: action taken
* rr: reward received
* s′s': next state
* α\alpha: learning rate
* γ\gamma: discount factor for future rewards

The agent balances **exploration** (trying new actions) and **exploitation** (using known best actions) with an epsilon-greedy strategy.

**Real-Time Simulation: FrozenLake**

To demonstrate these core ideas, we implemented Q-learning in the FrozenLake environment from OpenAI Gym. In this simulation, the agent navigates a frozen grid to reach a goal while avoiding holes. We used a 4x4 grid with non-slippery surfaces for clearer learning behavior.

During training, the Q-table was visualized in real time using a heatmap. This allowed us to see which actions the agent was learning to prefer in each state. Over time, the table showed distinct patterns where the agent consistently learned the most rewarding paths.

After training, the agent was tested in a rendered environment to observe how well it could apply the learned policy. It was able to reach the goal successfully by choosing actions based on its updated Q-table.

**Conclusion**

Reinforcement learning enables agents to learn optimal behaviors through interaction with their environment. By focusing on trial-and-error learning, rewards, and long-term planning, RL stands as a powerful method in AI. While there are more advanced methods for complex tasks, understanding Q-learning provides a solid foundation for exploring the broader world of intelligent decision-making systems.