

Introduction to Reinforcement Learning

- In Reinforcement learning, we teach our algorithm or ‘agent’ in an environment that produces a state and reward.
- The agent performs actions/interacts with this environment which results in various responses.

Learning in Reinforcement Learning

- The agent in this environment examines the state and the reward information it receives.
- It chooses an action that maximizes the reward feedback it receives.
- The agent learns by repeated interaction with the environment.

A successful agent needs to

- Learn the interaction between states, actions and their corresponding rewards.
- Determine which action(s) are the best to take.

Q Learning

- Q-learning is a model-free reinforcement learning algorithm.
- The goal of Q-learning is to learn a **policy**, which tells an agent what action to take under what circumstances.
- It does not require a model of the environment, and it can handle problems with stochastic transitions and rewards, without requiring adaptations.

Q Learning State Reward Tables

	Action 1	Action 2
State 1	0	5
State 2	5	0
State 3	0	5
State 4	5	0

- This is a simple State-Action-Reward table where in this ‘game’ we can see at each state the best reward the agent needs to take.
- If an agent randomly explored this game and summed up which actions received the most reward in each of the four states (and stored this in an array), then it would basically learn the functional form of the table above.

	Action 1	Action 2
State 1	0	5
State 2	5	0
State 3	0	5
State 4	5	0

- After extensive trials, an agent will be able to learn that taking Action 2 repeatedly for States 1,2,3 and 4 leads to the greatest reward.

The Q Learning Rule

We can define the Q-Learning update rule as:

- r – Reward
- γ – Discounts reward impact (0 to 1)

$\max_{a'} Q(s', a')$ - This is the maximum Q value possible in the next state. It represents the maximum future reward thus to encourage the agent to aim for the max reward in as little action steps