

Implementation of RL Algorithms in OpenAI Gym

Douglas Trajano, Sirleno Vidaletti
Master's Degree in Computer Science
Pontifical Catholic University of Rio Grande do Sul - PUCRS

Problem Statement

We want to do a practical project that implements some reinforcement learning algorithms that we had learned in class and some variants of them.

The following is a list of the algorithms we will implement:

- Q-Learning
- Deep Q-Network (DQN)
 - Experience Replay
 - Prioritized Experience Replay

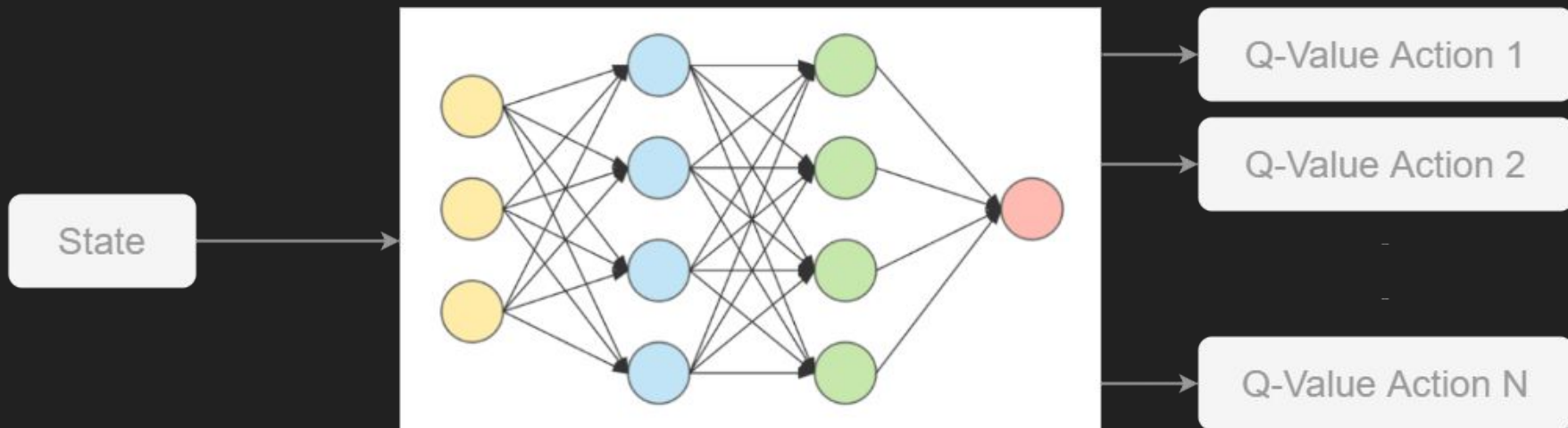


Deep Q-Network (DQN)

The **Deep Q-Network (DQN)** combines Q-Learning with deep neural networks to let RL work for complex, high-dimensional environments, like video games, or robotics.

A critical component of DQN-style algorithms is memory buffer.

In this project we will develop two approaches: **Experience Replay** and **Prioritized Experience Replay**.



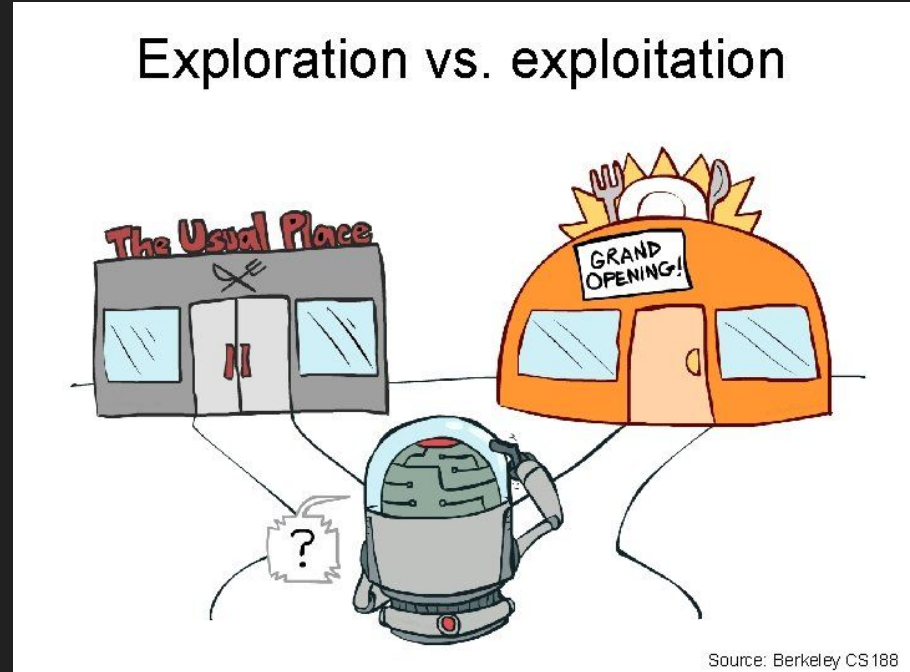
Exploration vs Exploitation

The exploration-exploitation trade-off is a well-known problem in decision-making systems.

Exploration: Select random action.

Exploitation: Select best known action.

In this project we will implement the ϵ -greedy policy in our algorithms.

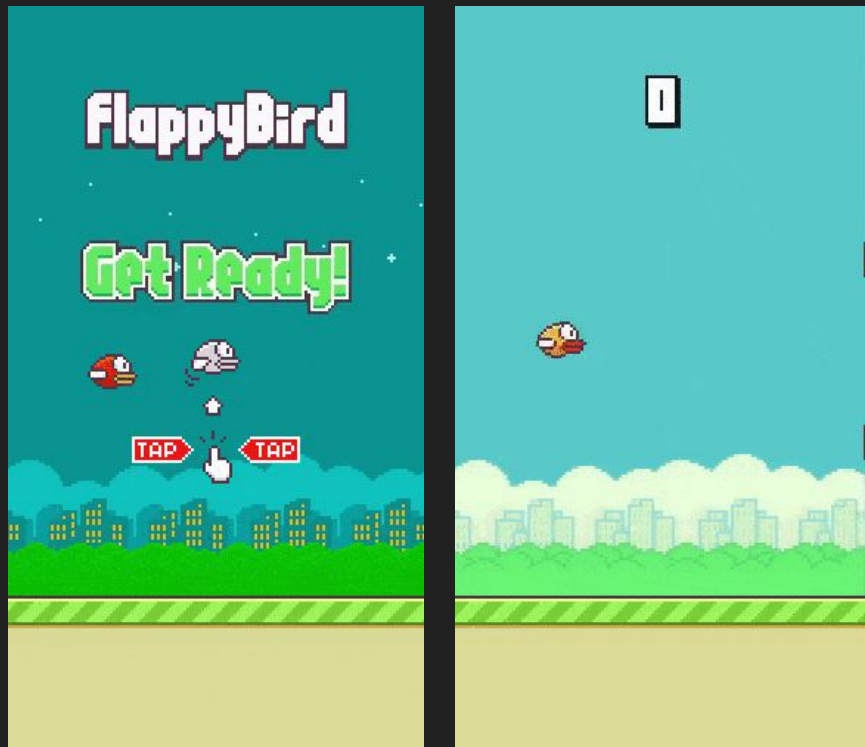


Environment

State observation is composed by RGB-arrays (images) representing the game's screen.

Available actions:

- do nothing
- jump

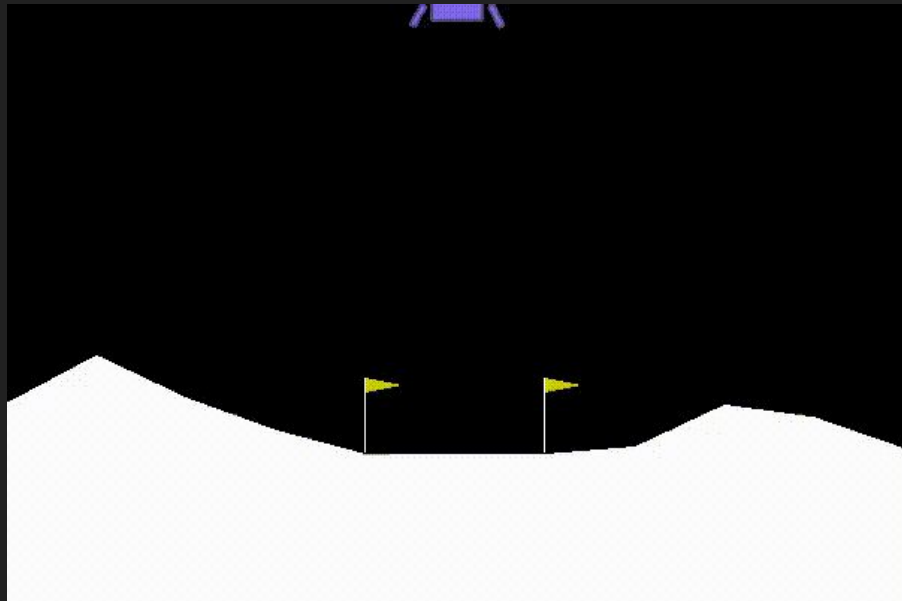


Environment

State observation is composed by a vector with 8 elements.

Available actions:

- do nothing
- fire left orientation engine
- fire main engine
- fire right orientation engine



Experiments

We hope to test the reinforcement learning algorithms in the Lunar Lander OpenAI Gym Environment as explained earlier.

At the end of the experiment, we will provide a report with the results of the experiments and the performance of each algorithm.

Experiment #1

Q-Learning

Experiment #2

Deep Q-Network (Experience Replay)

Experiment #3

Deep Q-Network (Prioritized Experience Replay)

Questions?