**UNDERSTANDING REINFORCEMENT ALGORITHMS THROUGH OPEN AI GYM (**[**CartPole-v0**](https://gym.openai.com/envs/CartPole-v0/)**)**

**Team:** Abhijith Satyakumar & Ronn George Jacob

Analysis and implementation of different reinforcement learning algorithms in order to solve the Open AI gym CartPole-v0 problem and finding a suitable algorithm to solve the problem. The CartPole-v0 problem is as follows – there exists a pole which would be attached to a cart moving along a track that is frictionless. The pole moves depending on the forces applied to the cart which would either +1 or -1. The pendulum starts upright, and the goal is to prevent the pole from tipping over more than 15 degrees from it’s vertical position. An episode ends when the pole has tipped more than 15 degrees or the cart has moved a certain distance from the center. A state for the problem consists of the information pertaining to the cart position, cart velocity, pole angle and pole velocity when it is on the cart as it tips back and forth. The actions for the problem involve pushing the card to the left and moving the cart to the right. There would be a reward 1 for every step taken until the episode terminates when the pole has tipped more than 15 degrees or when the cart has finally moved past a particular distance from the center. The inputs to the algorithms we intend on delving into would take in the state, action and would then try to maximize the reward accordingly. The output from the algorithm would be consequent values of the cart position, cart velocity, pole angle and pole velocity such that the reward for the episode is maximized. The problem is said to be solved when the average reward is greater than or equal to 195.0 over 100 consecutive episodes. Our aim would be to maximize this average as much as we can. The expectation from the project is to have a better understanding of different reinforcement algorithms and how various parameters such as noise, rewards, discount factor, etc play a role in a specific reinforcement algorithm. We intend to fulfill this expectation by solving the CartPole-v0 problem by implementing different RL agents based on our understanding of these algorithms and observe how the algorithms work out in the real world. We expect to use Q-Learning and Deep Q-Learning Network. The literature survey for the same would carried out with a detailed analysis of the algorithm playing out in the CartPole-v0 environment.

The libraries that we’re not familiar with as of now would be PyTorch, numpy, matplotlib and Tensorflow. The references for learning the same would be from the documentation of each of these libraries as provided below.

* [PyTorch Docs](https://pytorch.org/docs/stable/index.html)
* <https://docs.scipy.org/doc/numpy-1.15.1/reference/>
* <https://matplotlib.org/tutorials/index.html>
* [Tensorflow Docs](https://www.tensorflow.org/api_docs/)

**Milestones**

Milestone 1: Construct a Q-Learning agent to solve the CartPole-v0 problem and acquire necessary knowledge about DQN through available literature and understanding the algorithm.

Milestone 2: Construct a DQN agent to solve the CartPole-v0 problem and explore ways to improve the agent in order to improve the agent’s performance with respect to the problem.

**Week Plan**

Week 1: Getting familiarized with the environment and start constructing the Q-Learning agent for the problem.

Week 2: Complete the Q-Learning agent implementation for the problem and start exploring DQN and understand the algorithm.

Week 3: Learn the necessary tools specified above to implement DQN and get started with implementing DQN agent for the CartPole-v0 problem.

Week 4: Construction of DQN agent for the CartPole-v0 problem.

Week 5: Creation of a better agent to solve the CartPole-v0 problem.

There isn’t a definite or clear delegation of tasks as we will be spending time doing literature survey together and thinking about the implementations together as well.