# Machine Learning Summer School 2020 - Reinforcement Learning Workshop - Assignment 2

World Description

A pole is attached by an un-actuated joint to a cart which moves along a frictionless track. The system is controlled by applying a force of +1 or -1 to the chart. The pendulum start upright, and the goal is to prevent it from falling over. A reward of +1 is provided from every tilmestep that the pole remains upright. The episode ends when the pole is more then 15 degrees from vertical, or the cart moves more then 2.4 from the center. At every time step, the agent observes its position, velocity, angle and angular velocity. These are the observable states of this world. At any state, the cart only has two possible actions: move to the left (0) and move to the right (+1). In other words, the state-space has four dimensions of continues values and the action space has one dimension of two discrete values.
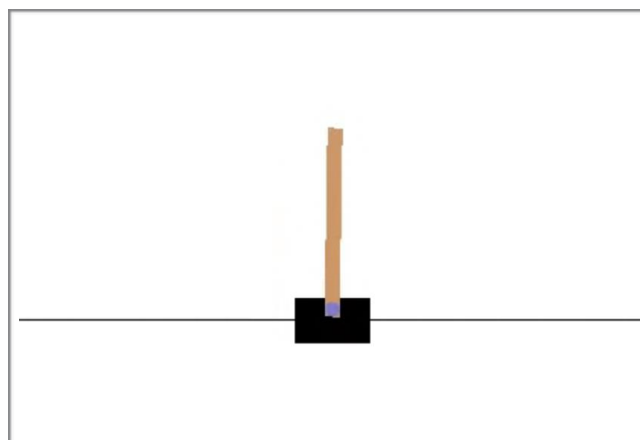
CartPole environment described above is provided in cartpole_env.py with CartPoleEnv class. The class acts as a "wrapper" around OpenAI Gym environment. OpenAI gym is an open-source library for developing and comparing reinforcement learning algorithms.

Questions

1. Implement Deep Q-Learning on the CartPole environment. (deepQLearning_Basic.py).

2. Implement Deep Q-Learning with Experience Reply (deepQLearning_Replay.py) and compare it to the original Deep Q-Learning algorithm.

   - What would be the best way to compare these 2 approaches?

   - Increase the number of episodes 2 times for both approaches. Which method has lower variance and provides more stable results? (Tip: You can train both approaches in parallel)