Homework 4

ELEN E6885: Introduction to Reinforcement Learning

Due: 11/25/2019

Problem 0 Background and Environment Configuration

Since we will have a DRL part in this assignment, you may use Google Colab to make things easier instead of running locally. It comes with the common libraries installed.

Colab is a free Jupyter notebook environment that runs entirely in the cloud. You can request a virtual machine with powerful computing power on Colab for free. There're some limits such as one can't use a virtual machine on Colab for more than 6 consecutive hours or you can't use Colab to mine bitcoins. You can click this link: Colaboratory for more information.

Problem 1 Basic tasks (20 credits)

In this part, you are supposed to implement some basic tasks by following the notebook. These tasks include:

1. Incremental implementation of average
2. Exploration vs exploitation comparison.

Problem 2 Model-based RL: value iteration and policy iteration (30 credits)

In this part, you will implement model-based methods - value iteration and policy iteration to solve the Taxi game problem by following the notebook.

Problem 3 Model-free RL: Q-learning and SARSA (20 credits)

In this part, you will implement the model-free methods, including Q-learning and Sarsa, to solve the Taxi game problem.

1. Implement Q-learning algorithm, your output should be an array of Q value for each state-action pair.
2. Implement Sarsa algorithm, your output should be an array of Q value for each state-action pair.
3. Plot the learning process of both algorithms for training 1000 episodes. (x-axis: episodes numbers, y-axis: average rewards) As is shown in Figure 1, the left can be a performance of Sarsa, the right one can be a result of Q-learning. (maybe different from your result)

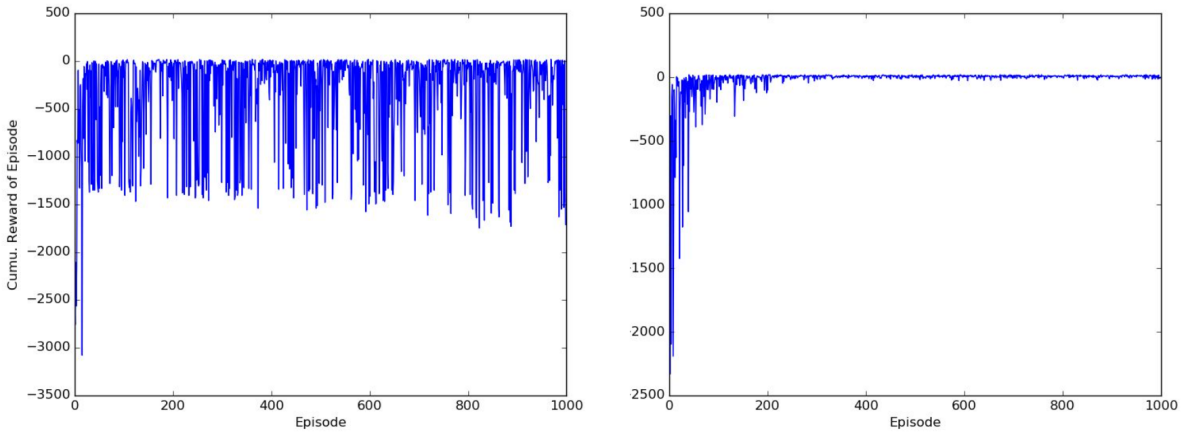

Figure 1: Learning Process of Q-learning and Sarsa

Problem 4 DRL problem (30 credits) In this part, you will implement a DQN agent to play the cart pole game. Please follow the instructions in the notebook of part 4 to finish the code.