CSE4/510: Introduction to Reinforcement Learning
Spring 2020

Instructor: Alina Vereshchaka

Project 1 - Building Reinforcement Learning Environment
Due Date: Sunday, March 1, 11:59pm

1   Project Overview

The goal of the project is to explore and get an experience of building reinforcement learning environments, following the OpenAI Gym standards. The project consists of building deterministic and stochastic environments that are based on Markov decision process, and applying a tabular method to solve them.

Part 1 [30 points] - Build a deterministic environment

Define a deterministic environment, where $P(s', r|s, a) = \{0, 1\}$. It has to have more than one state and more than one action.

Environment requirements:

- Min number of states: 4
- Min number of actions: 2
- Min number of rewards: 3

Environment definition should follow OpenAI Gym structure, that includes the following basic methods:

```
def __init__:
    # Initializes the class
    # Define action and observation space

def step:
    # Executes one timestep within the environment
    # Input to the function is an action

def reset:
    # Resets the state of the environment to an initial state

def render:
    # Visualizes the environment
    # Any form like vector representation or visualizing using matplotlib will be sufficient
```

Part 2 [30 points] - Build a stochastic environment

Define a stochastic environment, where $\sum_{s',r} P(s', r|s, a) = 1$. A modified version of the environment defined in Part 1 should be used.

Part 3 [40 points] - Implement tabular method

Apply a tabular method to solve environments, that were built in Part 1 and Part 2.

Tabular methods options:

- Dynamic programming
- Q-learning
- SARSA
- TD(0)
- Monte Carlo

2 Deliverables

There are two parts in your submission:

2.1 Report

Report should be delivered as a pdf file, NIPS template is a suggested report structure to follow.
In your report:

- Describe the deterministic/stochastic environments, that were defined (set of actions/states/rewards, main objective, etc)
- What is the differences between the deterministic/stochastic environments?
- Show your transition-probability matrix for stochastic environment.
- Discuss the main components of the RL environment.
- Show your results after applying an algorithm to solve deterministic and stochastic types of problems, that might include plots and your interpretation of the results.
- Explain tabular method that was used to solve the problems.

2.2 Code

The code of your implementations. Code in Python is the only accepted one for this project. You can submit the code in Jupyter Notebook or Python script. You can submit multiple files, but they all need to have a clear naming. All Python code files should be packed in a ZIP file named $YOUR\_UBID\_project1.zip$ After extracting the ZIP file and executing command python main.py in the first level directory, it should be able to generate all the results and plots you used in your report and print them out in a clear manner.

3 References

- NIPS Styles (docx, tex)
- Overleaf (LaTex based online document generator) - a tool for creating reports
- GYM environments

- Lecture slides
- [Richard S. Sutton and Andrew G. Barto, "Reinforcement learning: An introduction", Second Edition, MIT Press, 2019](#)

4   Submission

To submit your work, add your pdf, ipynb/python script to zip file $YOUR\_UBID\_project1.zip$ and upload it to UBlearns (Assignments section). After finishing the project grading, you may be asked to demonstrate it to the instructor if your results and reasoning in your report are not clear enough.

5   Important Information

This project is done individually. The standing policy of the Department is that all students involved in an academic integrity violation (e.g. plagiarism in any way, shape, or form) will receive an F grade.

6   Important Dates

March 1, Sun, 11:59pm - Assignment 1 is Due