CSE4/510: Introduction to Reinforcement Learning
Spring 2020

Instructor: Alina Vereshchaka

Assignment 3 - Policy Gradient & Actor-Critic
Due Date: Sunday, April 19, 11:59pm

1 Assignment Overview

The goal of the assignment is to explore reinforcement learning environments and implement actor-critic algorithms. In the first part of the project we will implement REINFORCE, in the second part we will implement actor-critic algorithm. The purpose of this assignment is to understand the basic policy gradient algorithms. We will train our networks on a reinforcement learning environment among OpenAI Gym or other complex environments.

Part 1 [40 points] - Implement REINFORCE

Implement REINFORCE algorithm. Apply it to solve RL environment. You can choose any environment among OpenAI Gym, Google Football environments or any custom defined **multiagent** environment.

Part 2 [60 points] - Implement Actor-Critic

Implement Actor-critic algorithm. It can be any of your choice: Q Actor-Critic, TD Actot-Critic, Advantage Actor-Critic (A2C), etc. Apply it to solve RL environment, that was used in Part 1.

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 discuss:

- What is REINFORCE?

- Describe actor-critic algorithm, that you choose.

- Describe the environments that you used (e.g. possible actions, states, agent, goal, rewards, etc).

- Show and discuss your results after applying REINFORCE and actor-critic algorithm to an environment (plots may include epsilon decay, reward dynamics, etc). Compare both algorithms in terms of learning speed and overall performance.

2.2   Code

The code of your implementations should be written in Python. You can submit multiple files, but they all need to have a clear name. All project files should be packed in a ZIP file named $YOUR\_UBID\_assignment3.zip$ (e.g. $avereshc\_assignment3.zip$). Your Jupyter notebook should be saved with the results. If you are submitting python scripts, after extracting the ZIP file and executing command python main.py in the first level directory, all the generated results and plots you used in your report should appear printed out in a clear manner.

3   References

- NIPS Styles (docx, tex)

- GYM environments

- Google Research Football

- Richard S. Sutton and Andrew G. Barto, "Reinforcement learning: An introduction", Second Edition, MIT Press, 2019

- Lecture slides

4   Submission

To submit your work, add your pdf, ipynb/python script to the zip file $YOUR\_UBID\_assignment3.zip$ and upload it to UBlearns (Assignments section). After finishing the project, you may be asked to demonstrate it to the instructor if your results and reasoning in the report are not clear enough.

5   Important Information

This assignment 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 for the course. Please refer to the UB Academic Integrity Policy.

6   Important Dates

April 19, Sunday, 11:59pm - Assignment 3 is Due