L-Università ta' Malta

Advanced Reinforcement Learning (ARI3212)

Individual Module Assignment

In this assignment you are expected to show an understanding of advanced concepts of Reinforcement Learning and hands-on skills in solving control problems.

The assignment requires the submission of a paper, with the sections defined below. Each section highlights the questions that are expected to be addressed. In addition, the methodology section highlights the experiments that you are expected to conduct. The results of these experiments are to be presented and discussed in the Results and Discussion section.

The paper should be between 8-10 pages long, including references. The format should follow the ACM Paper latex format available here: ACM SIGPLAN Proceedings Template

The marks will be allocated as follows:

50% - Paper write-up

50% - Experiments results and discussion

- 25% first experiment
- 25% second experiment

The submission has to be done as a one zip file. The zip file should include:

1. Paper (PDF)
2. Python Code or Jupyter notebooks supporting the experiments presented.

Paper Structure and Experiments:

Title – Solving Control Problems using Reinforcement Learning

Introduction

- What is Reinforcement Learning?
- How does it differ from other ML approaches?
- Briefly introduce and explain differences between Value Based, Policy Based and Actor Critic models.

Background

This section is intended for the student to show a good grasp of advanced Reinforcement Learning Concepts and Algorithms. You are expected to present relevant equations or references for better explanations.

- Explain Value Based Methods
- Explain Deep Q-Networks algorithm

- Explain Policy Based Methods, also highlighting the differences when compared to Value Based Methods.
- Explain Actor-Critic Methods, also highlighting the differences between stochastic and deterministic policy approaches.
- Explain two examples of Actor-Critic based algorithms. Your selected algorithms should be those that you apply for solving Experiment 2 below.

Your explanation of Algorithms should include:

1. Presentation of key model training update equation/s

2. Excerpt of your specific code clearly mapping the presented equation/s (include **only** the relevant line/s that specifically transpose the equation/s to code. If more than one line of code is presented, in your explanation refer to the specific line).

Methodology

Your assignment has to cover two experiments:

Experiment 1: In this experiment the objective is to try to solve the Open AI Gym LunarLander (Discrete Actions) problem ("LunarLanderv2") using:

1. Standard DQN
2. Apply two improvements to the standard DQN (apply separately) and compare/discuss against the standard DQN.
3. Combine the selected improvements into one DQN model and compare the new model against the previous models.

You should consider the problem solved when you reach an average score >= 195 over the last 50 episodes.

Experiment 2: In this experiment the objective is to try to solve the Open AI Gym LunarLander (Continuous Actions) problem ("LunarLanderContinuous-v2") using TWO Actor-Critic RL algorithms to solve the continuous case, using:

1. A stochastic policy approach.
2. A deterministic policy approach.

For both experiment 1 and 2, the students are expected to briefly:

- Explain the problem being addressed (environment, actions, rewards, etc).
- The parameters / configurations evaluated and the reasoning behind selection.
- The measures selected to validate the experiment.
- Libraries used.

In both experiments you are asked to make use of the code made available in class and adapt/modify the code accordingly. Alternatively, you can utilize a Reinforcement Learning Library.

Results and Discussion

For each of the above experiments:

1. Using appropriate tables and/or plots (e.g. return vs trained episodes), show the performance of the specific algorithms.
2. Discuss / provide conclusions from your results.

References