Team Project in Reinforcement Learning

INF581 – Advanced Topics in AI

February 6, 2020

1 Introduction

You work in a team to develop a project on reinforcement learning. The project is to be presented in the form of a report/research paper. You are open to explore different themes of reinforcement learning that have been presented in class (or that can be easily connected to such themes), under the constraints that your results are unique (not found elsewhere or realized by others) and are easily reproducible under Python in AiGym, including an environment which implements the `step` function.

2 Choosing a topic

Some example projects, with some indication of expected scope:

- You build a simple game (e.g., using the PyGame library) and add a `step` function making it compatible with AiGym agents. You select and deploy an off-the-shelf agent, and document the results.
- You choose an existing game and convert it to an environment; proceeding similarly to above.
- You design a simple simulation to emulate some real-world problem (e.g., supply-chain management, energy management, …) as an AiGym environment. You select and deploy an agent for the environment; you discuss the results.
- You take an interest in a particular type of RL learning method. You analyse some of its properties, attempt to improve it, and extensively document and discuss its performance on many off-the-shelf environments.
- You take a particular environment and agent from AIGym. Using extensive parametrization and empirical simulations, you show and discuss the performance obtained, providing insight and reflections.

Note that for this project, 'reinforcement learning' encompasses cases where the environment is fully known to the agent (i.e., MDPs), including also search methods on deterministic environments – if so desired.

3 Writing the report

You write a report detailing your work and discussing the results. Detailed instructions/hints on writing the report are given in `report.pdf` which is itself a template for the report. Of the grade allocated to the report, at *minimum* the grade distribution will consist of:

- 25% Introduction/background
- 25% Results/discussion

and the remainder allocated depending on your particular project (you will be graded on the context you present in the report, as long as it is in scope).

It is essential that you provide sufficient introduction and background material to your project in the report. This means that your report should be understandable/self contained for someone who has taken the course (not requiring further study, outside of what is given in the report). More information/suggestions in `report.pdf`.

Your final report will be a *maximum of 5 pages* of double column as per the format indicated in `report.pdf`, not including references. In addition, you may include an appendix of additional figures/results that you wish to report (e.g., additional results that you obtained, not essential to the discussion).

4 Peer-Review Process

We will run the project in the style of a double-blind[1] peer-review system. It means that your team initially submits an anonymous version of your report. Then you (as an individual) will receive 2-3 reports to comment/provide feedback on. After receiving the feedback from the anonymous reviews, your team will have an opportunity to improve your own report in reflection prior to grading by the INF581 teaching team.

It is important to read the information carefully, and also to respect the timeline. In particular, it is important to bear in mind:

- The feedback/review step is *not* about grading the work of other students. Rather, it is *your* reviews that are being graded by the teaching team, as a reflection of your understanding of the course material.
- The feedback/review phase is an *individual* component, *not* related to the composition of your team or your team's project. It means also that you should not share information with your team (or anyone else) on the reviews you make – and anyway there could be need/benefit to doing so.
- All reviews must be anonymous, positive, and constructive. It is not allowed to write negative reviews (this incurs a low grade *for the reviewer*).
- The *teaching team* will grade the report of all teams *independently* of the student feedback/reviews. In fact the team grade may not correlate at all to initial feedback your team received.

The advantages of this process is

- Become familiar with the peer-review process
- Receive diverse and constructive feedback regarding your work
- A chance to improve and fine-tune/debug your work before grading is carried out
- Encouragement to follow/study all topics covered in the course, rather than over-specialize in material relevant to your project

The process will be managed in the EasyChair system: `https://easychair.org/conferences/?conf=inf5812020`

5 Grading

In addition to the 25% for the class assignment, the project consists of the following components:

5.1 Team Component: The Project Report – 55% of final grade

Although the final grade is carried out by the teaching team, there are minimum requirements for the initial submission (to ensure that everyone has something to provide feedback on). Namely:

1. Your submission must describe at least the environment(s) and agent(s) that you investigate.
2. Your submission must contain a link to your code, which must be easily executable

Despite these minimum requirements, it is highly recommended to submit already a finished report, to make most advantage out of the process.

Your *final* submission is provided on Moodle (see Section 6 for the Timeline).

5.2 Individual Component: Feedback/Reviewing – 20% of final grade

This phase will be carried out after initial submission of the report (to EasyChair), and before final submission for grading (to Moodle).

You will be assigned 2 or 3 *anonymous* reports. You will fill out review forms on EasyChair, using approximately 300–500 words. In this review, you should make

- Positive comments (e.g., *Section 2. Background and Related work gives an excellent overview of the required background knowledge, in particular the outline of Q-learning*)

[1] Authors don't know who reviewers are, and reviewers don't know who authors are

- Constructive comments (e.g., *Time complexity of the agent's learning method could perhaps be reduced by using A* search instead of BFS …*; or *I had trouble understanding Section 4., adding more details about the agent's learning process would help*; or *What value is used for α in the results of Figure 4.8?*)
- Curiosity/speculative comments (e.g., *It would be interesting to observe what happens if you change the reward signal of the environment such that …*)

Your reviews will be graded on account of how well you demonstrate knowledge and understanding related to the concepts covered in the course. It is better to be specific, clear, and detailed on two or three main points than to try list all the small mistakes that you can find. In *no case* should you attempt to act as if grading the report. If you find clear evidence of intentional plagiarism/malpractice, you should mention your doubts by email to a teaching coordinator as you would in any other context. On the other hand, if you believe you have found related work the authors may simply not be aware of – you could mention this to their benefit in the feedback/review.

6 Timeline

The dates here are the *final* day by which to complete the tasks. It is important to adhere to deadlines; in some cases, a delay may affect your grade. Any changes to these deadlines will be announced on Moodle.

04/03/20 (Individual) Create an account on EasyChair[2]

04/03/20 (Team) One member of the team submits your report to the "INF581 conference" in EasyChair.
N.B. The `pdf` file of your report *should be completely anonymized* – but all authors (team members) should be listed *in the submission process.*

06/03/20 (Individual) Accept invitation to be a reviewer of the "INF581 conference"

11/03/20 (Individual) Submit the reviews of papers assigned to you

18/03/20 (Team) Submit the final version of your report in Moodle

[2] `https://easychair.org/conferences/?conf=inf5812020`

7 FAQ - Frequently Asked Questions

WE HAVE AN IDEA OF AN AGENT/ENVIRONMENT FOR THE PROJECT - IS IT SUITABLE? You can discuss with one of the teaching coordinators during the labs. The main points to bear in mind: If you are developing an environment from scratch, be careful not to make it too difficult. On the other hand, if you are reimplementing results from a paper (i.e., an existing environment) you will be expected to perform sufficient experimentation (not restricted to the same results given in the paper). Results must be reproducible in Python and relevant libraries. Should be related to topics covered in the course, or based on them.

DO WE HAVE TO SUBMIT OUR CODE? You *must* include a link to your code in the report, and it must work. Be careful about using links that expire after a number of days.

ARE WE GRADED ON OUR CODE? You are not graded on the quality of your code. However, you are graded on the originality (or, referencing of) and reproducibility of your results: your code must be available and working, and executable using any of the Python libraries from the labs.

It is highly recommended to include a demo script to run at least a part of your code with the least effort possible.

You can use as much code from elsewhere as you wish but *you must* always acknowledge this, and clearly differentiate your original contribution.

Obviously if you 'borrowed' a lot of code you would be expected to make a relatively more sophisticated analysis/simulation/parametrization to compensate.

HOW CAN WE MAKE CODE AVAILABLE ANONYMOUSLY? There are several ways to create an anonymous link to code. If you don't find one which suits your needs, simply send a zip file to one of the teaching team with your code and we can make an anonymous link for you to add to your report.

I DON'T WANT OTHERS TO SEE MY (TEAM'S) WORK – CAN I OPT OUT? It is typical to block certain people from reviewing your work. This is usually known as a 'conflict list'. There is nothing negative associated about it (in fact usually it usually lists close colleagues/groups). You may tell a teaching coordinator if you want a conflict list to avoid people seeing your work (we will not ask or care the reason).

If you want to completely opt out – you still participate in the team project and will be graded in exactly the same way as everybody else. The only difference is that you don't get feedback prior to grading, and you will be given a selection of reinforcement-learning papers from the academic literature to review instead of your anonymized peers' work. In this case, *make sure your team agrees with you!*

THE PAPERS I'M ASKED TO REVIEW IS OF POOR QUALITY OR INCOMPLETE Since grading is done on the final submission (that you do not see) some teams may submit a partially completed initial report, intending to finish later. However, due to the minimum requirements on the initial submission, the papers you receive should describe at least an agent and an environment. In all cases, recall: It is not your task to decide how good or complete a submission is. If there is not much material/results yet – you can instead provide feedback/demonstrate your knowledge by instead suggesting some relevant development/experiments the authors might want to try.

I DIDN'T RECEIVE THE SAME PAPER TO REVIEW AS MY TEAMMATE(S) This is normal! Recall the reviewing process is an individual component! You should not share with your teammates information about the papers you received to review, and there is certainly no need to do so.