INF581: Class Project

1 Introduction

You work in a team to develop a project on reinforcement learning.

- You design and implement an environment
- You design and deploy an agent for the environment
- You adapt and apply your agent to another environment[1]
- And 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. We will run the project in the style of a double-blind[2] peer-review system. This means that you submit your report and will review/comment on an *anonymized* version of other's reports, and then have an opportunity to respond and improve your own report in reflection (or justify why not) before the grading (by the teaching team). **Important note**: this does *not* mean that you grade each others work. The teaching team will review the report of all teams and the grade assigned may or may not correlate at all to the reviews that your received. The advantages of this process is

- Become familiar with the peer-review process
- You will receive diverse and constructive feedback regarding your work (any reviews that are written negatively or unfairly about your work will induce a penalty to the reviewer – see Section 2 and Section 4)
- You have a chance to improve your work before the final grading
- You have the chance to see other work and you are encouraged to follow the lecture material in general rather than over-specialising in your project; so that you may comment appropriately on it and demonstrate your own understanding of the material

2 Grading

There is 80% of your course grade related to the project. In summary, your grade will be based upon how well you

- Understand and present a selection of the lecture material
- Understand and approach a reinforcement-learning problem, and adapt to the challenges faced
- Implement using libraries seen in the labs, and is reproducible
- Discuss and present the results (including tables, figures, text, and code, . . .)

There will be two components to the grade

70% Your team's written report (see `report.pdf`) – the final version, including team's responses/reactions to reviews.

10% Your reviews on other reports (see `review.txt`)

3 Possibly-Asked-Questions

Q1a. *I feel that I would recognise* and *not be able to give a fair appraisal of someone's [team's] work in particular* (for whatever reason, usually positive reason, such as being friends with that person out of class).

[1] The *Artificial Life/Bug World* environment: `https://github.com/jmread/alife`

[2] Authors don't know who reviewers are, and reviewers don't know who authors are; the teaching team are the *program chairs*.

Q1b. *I feel that somebody in particular would recognise* and *not be able to give a fair appraisal of my [team's] work* (for whatever reason, usually positive reason, such as being friends with that person out of class).

As in a normal peer-review conference, we will have 'conflict' lists. You simply add the person to the list. There is nothing negative associated with this list, it usually means you have collaborated closely in the recent past or cannot provide a fair review for any other reason. It is not revealed to anyone and has no affect on grading/reviewing at all. We will not even ask what the reason is.

Q2 *Actually I am uncomfortable with* anyone *in this class reading my work at all.*

In this case your conflict list is everybody. You still participate and will be graded in exactly the same way as everybody else. The only difference is that all comments will come from the teaching team and you will be given a selection of reinforcement-learning papers from the academic literature to review instead of your anonymized peer's work. *Of course, make sure your team agrees with you!*

4 Reviewing (10%)

Your reviewing is worth 10%. You will be assigned 2 *anonymous* reports. You are to fill out the review form based on `review.txt` for each report, using approximately 1–2 pages. In this, 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*)
- 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 Algorithm 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 ...*)

It is better to be specific, clear and justified on two or three points than to try to list as many mistakes as you can find (the review does *not* need to be exhaustive). Recall that you are only rewarded for sharing knowledge/insight that will potentially improve the report, not for finding mistakes.

5 Rebuttal and Revision

You (and your team) read the reviews received. You update your report (you can add 1 extra page if necessary for new material) and also respond to the comments made, such as

- Detailing changes, e.g., *We have rewritten Section 3 to be more clear; and have now specified alpha=0.3*, or *Interesting comment regarding the reward signal. We have re-run a few simulations with the new function suggested and added a new figure with the results (Figure 3.4b). It shows that ....*
- Justifying non-changes, e.g., *Actually in our case, we did not apply A* because ...*, or *Interesting comment regarding the reward signal. Unfortunately we did not have time to run a full simulation with this change, but we could speculate that the agent would take much longer to learn the task in that case, since ....*

This component is included in your report grade. If there are too few comments regarding extra material, you can follow your own suggestions that you mentioned in 'future work'. Further comments may be appended by the teaching team (the "*program chairs*").

6 Timeline

The dates here are the *final* day by which to complete the tasks. The dates marked with an asterisk* are compulsory (failure to fulfil will affect your grade – except † which is an optional step)! The first two dates are recommended.

25 Jan Create a group of 2–3 people; discuss topics

31 Jan Start work on the project

28 Feb* Create an account on EasyChair[3] and *one* of your team submits your abstract to the "INF581 conference"[4] – listing the others as authors *in the submission system* (*not in the* `pdf!`)

28 Feb*† Send your conflict list, if any

07 Mar* Accept invitation to be a reviewer/on programme committee of the "INF581 conference"

07 Mar* Submit your report to the "INF581 conference" in EasyChair as a `pdf` file

14 Mar* Submit the reviews of papers assigned to you

21 Mar* Submit the final version[5] in EasyChair, *and* also to Moodle, along with a copy of your rebuttal

Your grade is then assigned by members of the teaching team (i.e., the *program chairs* of the conference) *after* the final submission based on careful consideration of your team's paper, and the reviews that *you* wrote. Note that this grade may not correspond with reviews received (*unlike* a real conference, your grade is assigned conditionally independent of the reviews received – also keep in mind that reviewers see only a small subset making it difficult to calibrate their reviews).

[3] `https://easychair.org/conferences/?conf=inf581`
[4] This step is only to give us time to see how many teams there are and organize the reviewing – nevertheless it is important
[5] aka the "camera-ready" version