

# COGS 182 - Project 2

Deadline: Mar 15 at 11:59pm

In this project you will apply reinforcement learning methods to **a problem of your choice**. Any problem or domain that interests you is valid. This could be, for example, a domain in which you have been working for research or in some other class. You may look online for project ideas and implementations (though see below the rules regarding attribution and plagiarism). Alternatively, you can try to do something that has never been done before. It's up to you!

Your solution to the problem of choice should include at least two RL algorithms. Each of the two algorithms must come from a different chapter of the textbook. Examples of algorithm pairings are: MC and TD; TD(0) and n-step TD; TD(0) and TD( $\lambda$ ); Q-learning and Dyna-Q; Tabular TD and TD with function approximation.

You may work in teams of 2 if you wish, or you may work independently. Working in a group means you can work on the same problem and discuss your approach with each other. However, each person must turn in their own submission, including writing your own code. If you work in a group, each person must also include a brief note summarizing each person's contribution to the final product.

You may build on existing work for this project (as long as you write the code yourself), but you must give proper attribution to all existing work that you build on and make clear what your new contribution is. Any unattributed or uncited work that you use will be considered a breach of academic honesty and dealt with according to the course policy in the syllabus. Furthermore, you may not claim your own existing work as a new contribution. You may build on your own work, but it must be clearly cited as existing work and you must do new work for the class project.

This project will have 3 submission checkpoints:

## Checkpoint 1: 20% of grade

*Due: Feb 28 at 11:59pm*

Write one paragraph describing what you plan to do and how you plan to do it, and a set of bullet points describing the high-level aspects of your project. You will receive feedback on your proposal.

The paragraph should be written with the goal of convincing the reader that what you are proposing to do is interesting and non-trivial (though not necessarily completely original — see below). It is ok if your proposal is based on something you read about, as long as you are going to write the code yourself. Just make sure to acknowledge any ideas that you borrow and code that you use as inspiration, and be sure to clearly identify what you are going to do that has not been done previously.

Below the paragraph, you should provide answers to the following questions:

1. What is the agent?
2. What is the environment?
3. What are the states?
4. What are the actions?

5. What reward function are you planning to use?
6. In what sense is this problem sequential?
7. What is the simplest possible first result that you will try to get?
8. What RL algorithms will you use?

## Checkpoint 2: 20% of grade

*Due: Mar 7 at 11:59pm*

You should implement the environment – i.e.  $p(s', r|s, a)$ . Show, in graphs, that your environment is working by displaying in some form that the environment returns the right state ( $s'$ ) and reward ( $r$ ) for a series of state-action pairs ( $s, a$ ). Submit it in .pdf format.

## Checkpoint 3: 60% of grade

*Due: Mar 15 at 11:59pm*

You should implement the two algorithms for solving the problem of your choice.

Submit your final project in .pdf format, including:

- Source code, executable and README . We recommend you using a github repository to hold the source code, executable and a README file that provides a brief guide to run your code. In this case, you just need to provide the github link within your final report. If you want to keep your repository private for any reason, please zip your project folder including the source code, executable and the README file, and include that with your submission.
- A 5-minute mp4 YouTube video summarizing your project and the main results (make sure to choose unlist option so that only people with the URL can access your video). The URL should be included in your final report PDF, at the beginning of your report, along with your github link. The detail instructions for uploading the video are on Piazza.
- A detailed written report describing your project, including its merits, and its deficiencies. As much as possible, you should relate your approach to the readings from throughout the course. View this report as a term paper. The report should be roughly in the style of a conference paper, including introduction, motivation, related work, etc. All writing should be your own – all quotes must be clearly attributed. Remember to proofread and spell-check!

If any parameteres are mentioned in the report, be sure to mention how you arrived at their values. Was it the first thing you tried? Trial and error? Roughly how many trials? etc.

Be very clear about what code you've used from other sources, if any. Clear citations are essential. Failure to credit ideas and code from external sources is cheating.

Each team member should individually (and privately) identify what was your role in the overall project, and what was your partner's role.