

Team Project in Reinforcement Learning

INF581 – Advanced Machine Learning and Autonomous Agents

1 Introduction

You work in a team of 3-5 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 any topic(s) in reinforcement learning that have been presented in class, or that can be easily connected to such themes.

2 Choosing a topic

Constraints You may choose any topic under the following constraints:

1. The methods used have been presented in class, or is based on the material presented in class
2. Your project demonstrates some kind of novel/unique aspect. This can be as simple as a unique modification to an environment, a novel set of experiments on an existing environment, different parametrization of a method, etc. You are not expected to make groundbreaking research; but you should aim to satisfy your curiosity on some aspect related to RL.
3. The code required to reproduce your results should be easily reproducible under Python with the AIGYM library
4. All aspects of the project which are not original contributions by the team members in the context of this course are clearly acknowledged/cited.

Scope and Examples Some example recipes for 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 and discuss 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 method. You analyse some of its properties, and document and discuss its performance on many off-the-shelf environments, and improve it or suggest ways to improve it based on your analysis (even if these improvements are specific).
- You take a particular environment and agent from AIGYM. Using extensive parametrization and empirical simulations, you demonstrate and discuss the performance obtained, providing insight and reflections regarding reinforcement learning.

3 Writing the report (50% of your final grade) – Team Component

You write a report/paper detailing your work and discussing your results. Instructions and hints for writing are given in `report.pdf` which is itself a template for the report.

- 25% Introduction and Background
- 25% Results and Discussion

- 50% The methodology of your particular project (you will be graded on the context you present in the report, as long as it is in scope).

Page limit: Your final report will be a **maximum of 5 pages** of double column as per the format indicated in `report.pdf`, *not including* references (i.e., references may go over the limit).

Submission of the report You submit the report twice. This is in order to receive feedback via a peer-review process in order to improve your report (an important step, because there is no oral presentation to clear up misunderstandings); see Section 4 for more details on the peer-review process, and Section 5 for important dates. Your first submission is not graded, but it must meet minimum requirements such that the final submission will be graded. Namely, it must

1. Have an introduction
2. Describe the environment(s) and agent(s) that you investigate.
3. Contain a link to the code which must be easily executable (max 5 minutes ‘setup’¹) using tools from the lab

Only final reports that have met these minimum requirements for the first submission will be graded. Beyond these minimum requirements it is highly recommended to submit already a finished report to make the most advantage out of the process.

After taking into account any feedback to improve your report, your **final submission** (with your names) should be submitted **on Moodle** by one member of your team. This version will be graded by the teaching team (provided you met the minimum requirements earlier).

4 Peer-Review Process (10% of final grade) – Individual Component

We will run the project in the style of a double-blind² peer-review system. It means that **as a team** you initially submit an **anonymous** version of your report. Then you **as an individual** will receive one anonymous report to comment/provide feedback on. After receiving the anonymous feedback in the form of ‘reviews’, your team will have an opportunity to improve your own report in reflection, prior to grading by the INF581 teaching team.

Remarks:

- The feedback/review step is *not* about grading the work of other students. Rather, it is *your* reviews that will be graded by the teaching team (10% of your final grade), as a demonstration 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.
- You are expected only to have the knowledge of what has been presented in the course (lectures and labs) when performing reviews (you are not expected to research new methods presented in a report; is it up to the writers to explain them in a way that someone who has followed the course can understand).
- All reviews must be anonymous, positive, and constructive. Negative/discouraging reviews incur a low grade **for the reviewer**.
- The *teaching team* will grade the report of 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 (to be able to provide good reviews)

¹Running in its most basic configuration. Hint: have a `README.md` with brief instructions; use a Jupyter Notebook, or include a `run_demo.py` script. Of course you should provide a minimum of documentation for trying different/more complex configurations

²Authors don’t know who reviewers are, and reviewers don’t know who authors are

How to provide feedback You will be assigned an **anonymous** reports. You will fill out review forms on EasyChair (links below), 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*)
- Constructive comments (e.g., *I think SARSA might be better suited to learning in this environment because ...; or I had trouble understanding agent's parametrization, what is λ ? (it is not specified); or Figure 4.8 is difficult to interpret – what is on the vertical axis?*)
- Curiosity/speculative comments (e.g., *It would be interesting to observe what happens if you change the reward function of the environment such that ...*)

Your review 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. You should indicate if there are problems but in no case should you attempt to suggest a grade. If you find clear evidence of intentional plagiarism/malpractice, you should mention your doubts *to a teacher* but provide a review anyway. On the other hand, if the authors have simply missed some related work or forgotten to cite some source – definitely mention this in the feedback/review.

5 Timeline

The dates here are the **final** day by which to complete the tasks. All steps must be completed. There will be a grade penalty for any delay (because this can hold up the process for everyone). Any changes to these deadlines will be announced on Moodle.

March 02 (Team) *One member* of the team submits your report to the “INF581 conference” in EasyChair³.

Important: You submit a **pdf** file of your report **completely anonymized**, but you list all authors (members of your team) along with the same email address as they use on Moodle in the EasyChair submission form (these details will not be visible to reviewers).

March 04 (Individual) Accept the e-mail invitation to be a reviewer for the “INF581 conference”. Didn't receive it? Check your spam folder and make sure that your email address is entered correctly in the previous step!

Shortly after (an announcement will be made) you will find an anonymous report has been assigned to you.

March 11 (Individual) Submit your review on EasyChair

March 16 (Team) Submit the final version of your report in Moodle (this time with author names listed)

³<https://easychair.org/conferences/?conf=inf581-rl-2022>

6 FAQ - Frequently Asked Questions

WE HAVE AN IDEA OF AN AGENT/ENVIRONMENT FOR THE PROJECT - IS IT SUITABLE?

Discuss with one of the teaching team. 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 directly on your code. However, you are graded on the originality (or, proper referencing of) and reproducibility of your results; therefore your code must be available and working, and executable using any of the Python libraries from the labs. Include a demo script to run at least a part of your code with the least time/effort possible.

CAN WE TAKE CODE FROM ELSEWHERE? You can use as much code from elsewhere as you wish but *you must always* acknowledge this clearly and unambiguously, and specifically differentiate your original contribution. (Obviously if you take a lot of code from somewhere else 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? At conferences there is what is known as a 'conflict [of interest] list'. There is nothing negative associated about it (in fact usually it includes close colleagues and personal friends). You may tell a teaching coordinator if you want a specific conflict list if some of your peers.

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 your team doesn't get anonymous 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 partially completed work, intending to finish later. It is unfortunate for them, but it will not inconvenience your review; you can instead provide feedback/demonstrate your knowledge by 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.