Project Instructions for Introduction to Reinforcement Learning (CSC2547) (2021 Spring)

Amir-massoud Farahmand
Department of Computer Science, University of Toronto
Vector Institute, Toronto, Canada

1 Introduction

The goal of the project is to give you an opportunity to gain experience in doing research in reinforcement learning (RL).[1] The range of acceptable research topics is wide. Your work can be theoretical, empirical, algorithmic, applied, or a combination of them. This is discussed more in Section 2.

The project contributes to 30% of your final mark. It has three components. The percentage of each component and the deadline for each of them are (subject to minor change)

- Proposal (5%): March 12th (Section 3)
- Presentation (5%): April 6th (and possibly March 30th, if some teams are available) (Section 4)
- Report (and Source Code) (20%): April 16th (Section 5)

Collaboration. The project should be conducted collaboratively as a team of 2-3 members. Everyone should contribute to the project in a meaningful way, and they should be clear about their contributions.

Some frequently asked questions are answered in Section 6.

2 Types of Projects

You may choose any topic of your choice, as long as it is related to RL. Below I describe a few general directions you may want to pursue, and how you can choose a topic within them. Your project might be a combination of some

You do not need to invent a new algorithm, reach a state of the art performance, or completely solve a new application domain to be successful in this project. I realize that there is not much time in a semester to learn about a new research area well enough to innovate a new idea and execute it completely (though sometimes your fresh perspective might lead to a new idea that others have not thought about it. In that case, you have to cherish that opportunity and pursue it). The goal is to give you a taste of what research in RL might look like. If you feel delighted enough after the end of this course, you have the option to continue working on the project with your team. This being said, you have to spend a good amount of time on this project and you must follow rigorous scientific methodology in pursuing your research.

Before providing specific suggestions for each type of project, I have a general suggestion on how you can start if you do not have any idea already: You can consult the proceedings of top AI/ML venues such as NeurIPS, ICML, COLT, ICLR, AAAI, IJCAI, AISTATS, JMLR, MLJ, and JAIR,

[1] This document may be slightly changed in the coming weeks in order to clarify any questions that you have. This version: 2021 March 1.

where RL papers are often published, to find many interesting papers. One of them may catch your attention. Read it carefully. Afterwards, read some of the papers that are cited within that paper, and backtrack. You often find the same set of papers referred to again and again. It is also very helpful to see what other more recent papers have cited your originally selected paper. This helps you figure out what advances has been done since that original paper.[2]

2.1 Theoretical Analysis

You may decide to work on the theoretical understanding of an RL algorithm or problem. Some example research directions are (by no means comprehensive):

- Understanding the convergence proofs of different RL algorithms (TD with function approximation, policy gradient, actor-critic) under various assumptions (e.g., linear function approximator, nonlinear function approximator, on-policy vs. off-policy sampling distribution, etc.)
- Understanding the sample complexity guarantee for RL problems and algorithms, investigate what the upper and lower bounds are, and whether they match.
- Investigate how the discounted MDP framework can be extended to other settings such as semi-Markov Decision Processes, Partial Observable MDPs, risk minimization instead of maximizing the mean return, etc., and study what this entails on standard algorithms such as VI, PI, LP, and their corresponding RL algorithms.
- Theoretically analyze an algorithm with a good empirical performance and try to analyze its theoretical properties, perhaps in a simplified setting (linear FA or even finite state-action MDP).

You should understand an aspect of theory literature very well and figure out

- What are interesting questions to ask? For example, is the stability of the algorithm the main concern? Or the sample complexity is?
- What are known and unknown about the topic? For example, do we have any upper bound for the sample complexity? What about a lower bound? Do they match?
- What assumptions are required to make the analysis work? Is there any assumption that can be relaxed? For example, do we need an i.i.d. assumption in the proofs? What changes if we relax that assumption? Or is boundedness of some quantities assumed? Is that necessary?
- Is there any part of the theoretical analysis that can be improved? Is there any tool that the authors of the paper use that is known to be improvable? For example, if they use Hoeffding's inequality, can the use of Bernstein's inequality lead to any improvement?

2.2 Algorithm Design

You may decide to design a new algorithm by varying certain components of an already existing algorithm, and effectively search in the space of "adjacent possible". For example, what happens if you use a DNN in a Fitted Q-Iteration framework? (Voila! You have re-invented DQN!). Or what happens if you use various variance reduction techniques in a policy gradient method? Or what happens if instead of using scalar gains in an online algorithm such as TD, you use a matrix gain? And perhaps design an adaptive mechanism to change the gain?

The space of all possible algorithms is combinatorially large. Maybe we can use a computer to automatically explore it. For this course, however, we want you to explore it based on the insight gained in this course as well as your other courses and research experience.

[2] You may also consult the following good set of suggestions by Csaba Szepesvári at https://rltheory.github.io/pages/assignments/, especially if you want to work on theoretical work. In addition to that, I consulted Animesh Garg's course on 3D and Geometric Deep Learning http://www.pair.toronto.edu/csc2547-w21/project/, and Roger Grosse's project instructions https://www.cs.toronto.edu/~rgrosse/courses/csc2541_2021/assignments/project.pdf for his Neural Net Training Dynamics. You may find good advice there too. Note that their evaluation criteria and what is acceptable or not is not the same as this course, so do not rely on that.

You do not want to randomly wander in the space of all algorithms. Any new algorithm should be justified. You do not need to rigorously prove that the algorithm works before trying it empirically, but it is good to have a theoretical insight before your start implementing.

2.3 Empirical Study

Empirically investigating an already existing algorithm is a reasonable project. Think of yourself as an experimentalist who wants to understand the behaviour of an algorithm through careful design and conduct of experiments.

You want to know when the algorithm works and when it does not. The original paper that introduced the algorithm might have not explored all relevant questions. It is likely that their authors only reported successful results. Your goal is to empirically investigate the range of problems and conditions that results in success and failure of an algorithm.

For example, you may want to know the effect of stochasticity of the environment, learning rate, or the particular exploration technique used on the performance of an online algorithm. Or you may choose to see how an algorithm works on a wide range of environments. You may choose a state of the art algorithm and implement it on domains that were not included in the original paper.

2.4 Application

If your main research topic is in an application area, broadly defined, you can investigate whether you can formulate it as an RL problem and solve it using RL algorithms.

For example, if you are working on robotics, computer networks, computer games, healthcare, autonomous vehicles, energy management, scheduling, you can use RL, as many have already done so. Of course, there are many more application domains that can potentially fit well within the RL framework, but have not been explored yet. Discovering this possibility is an exciting area.

3 Proposal

Your proposal should be a maximum of two (2) page summary of your intended research direction (references excluded in the page limit). You need to clearly define the problem, provide a brief summary of prior work, and what you intend to achieve.

The instructor and the TAs will provide you feedback on this. We also provide some office hours before the proposal deadline, in case you want to bounce ideas back and forth before submitting them in the written form.

4 Presentation

You need to prepare a presentation defining the problem, and very briefly summarizing what is known about it, your contributions and results, and some conclusions. Depending on the type of paper you have, the format might be different (we do not expect any figure in a purely theoretical work).

You have about 5min to present your work. The exact amount depends on how many teams we will have. As this is a short time for a presentation, try to be efficient in your communication.

Note that you need to present the material in the class and all the team members should be present to answer questions (unless there is any valid excuse to miss the class, e.g., a medical problem). We also record the presentations and intend to upload them on YouTube to showcase your achievements. If you are against uploading your presentation on YouTube, please contact us well in advance of the presentation day.

5 Report (and Source Code)

You should write a 6-8 page report summarizing your work. We encourage you to use LaTeX.[3] You can use the NeurIPS style file https://nips.cc/Conferences/2020/PaperInformation/StyleFiles, though you are not required to use this specific style.

At a high-level, your report should include

- Problem definition and motivation: Clearly state what problem you are tackling and why we should care about it.
- Summary of prior work: What other attempts have been done in order to address this problem.
- Your contributions: Statement and proof of a new result or summary and critique of prior results (Theoretical); clear description of the algorithm and evidence (theoretical or empirical) supporting how it works (Algorithmic); the description of the algorithm, the experimental design to evaluate them, and the empirical results (Empirical); how you formulated your application, the description of algorithms you have tried, and the performance you achieved in comparison with other (non-RL) baselines (Application).
- Conclusions: What have you learned and what is remained to be done or figured out?

Your report will be evaluated based on its quality of writing and explanations, how well you cover the prior work, precision of your statements, your contributions (which depends on the type of research you have conducted), and following the good scientific methodology.

It is common in research that each co-authors contribute to different aspects of the project. Some may come up with the high-level ideas, some design the algorithm, some study the idea theoretically, some design and conduct the experiments, and some others write the paper. I'd like to acknowledge that this is how modern science works and let you have different contributions. That being said, you need to have a section describing the rule of each team member in the whole project in some detail. The only requirement is that *all team members must be involved in writing the paper*. If you are not good in writing yet, the grad school is the right place to practice it.[4]

If your project has a source code, you should submit it too.

6 FAQ

Q: Is it acceptable to have a project that overlaps with my research project?

A: Yes, and I encourage it. But you should be clear about what part has been done before this project, and what contributions are new. The basic idea is that you should not reuse your prior work for this project; you have to spend a lot of time during this semester to work on this project, but you can use it for your thesis (of course, if your supervisor is OK with it).

Q: Is it acceptable to have a project that overlaps with another course project?

A: Try to avoid it! If there is a good reason to have a project that spans more than one course, that can be discussed. You need to get the permission of all instructors for this. Which means that you need to send an email to me and the other instructor(s) and get a joint permission. Since your research is done within a team and you may have different teams in different courses, this makes the credit assignment difficult, hence the discouragement.

Q: Can I extend the project from a previous course?

A: Yes! You should mention it in your proposal, include the report from the previous project in your submission, and be explicit about the new contributions specific to this course. In other words, be clear about the Δ.

Q: Can I have a team size of 4+ or 1?

[3] If you do not know how to use LaTeX, this is a great opportunity to learn.

[4] If there is any good reason that you cannot write the paper, for example a medical reason, you should discuss it with me beforehand.

A: Teams should consist of two or three members. A solo team is only acceptable only if there is a good reason. You need my permission for doing so. Team with more than 3 members are only acceptable under very specific circumstances, e.g., if you are building a complex system that requires various expertise, I am open to discuss it. I do not expect to have any 4+ team member though.

Q: What if I discover that someone else has done a very similar thing to what I am doing in this project?

A: That is OK. It is a part of science. Make sure to mention those paper(s) in your prior work, and provide a detailed comparison.