

## COS435 / ECE433: Intro RL Final Project (S25)

Unbolded deadlines are subject to change! (peer feedback and experimental figure)

This document describes the final project. During this time, homework assignments will be lighter. Students will work in groups of size 3.<sup>1</sup> The learning objectives are for students to gain hands-on experience working with practical RL algorithms and environments. Each group will choose one of the following options for their project:

**Option 1:** Reproduce a recent reinforcement learning paper. Students will identify a recent (published in 2024 or 2025) RL publication (from NeurIPS, ICML, ICLR, or similar). They will reimplement the method, and reproduce at least some of the experiments in the paper. The aim is to gain familiarity with a recent method, and also to think critically about the experiments performed in the paper: are they statistically rigorous? Would you recommend any changes to the authors? In your writeup and presentation, you should focus on analyzing the claims in that paper.

Some papers that could be good to look at:

- BRO
- SIMBA
- SGCRL
- Maximum Entropy Inverse Reinforcement Learning of Diffusion Models with Energy-Based Models
- Mr.Q
- (we'll keep adding more papers here)

**Option 2:** Applying RL to a new problem. Students may use an off-the-shelf RL library to tackle a decision making problem that (to the best of their knowledge) has not been tackled with RL before. Questions to consider include: what are prior approaches to this problem? Why does it make sense to treat this problem as an RL problem (e.g., as opposed to a bandit problem)?

Some domains you could start looking at:

- Chemical plant control
- Glucose / blood sugar control
- HIV/AIDS treatment
- Something at Princeton. E.g., lights/heating in your dorm, some of the many Princeton apps
- SustainGym
- Air traffic control, BlueSky
- processing minerals / mineral refinement using reinforcement learning to control reaction/timing/process-step
- (we'll keep adding more papers here)

---

<sup>1</sup> At most one group can be of size < 3, if numStudents % 3 != 0

**Option 3:** Try to win one of the RL competitions. E.g.,

- MyoChallenge 2023: Towards Human-Level Dexterity and Agility
- MyoChallenge 2024: Physiological Dexterity and Agility in Bionic Humans
- Large-Scale Auction Challenge: Learning Decision-Making in Uncertain and Competitive Games
- Lux AI Challenge Season 2 NeurIPS Edition
- Lux AI Season 3: Multi-Agent Meta Learning at Scale
- The NeurIPS 2023 Neural MMO Challenge: Multi-Task Reinforcement Learning and Curriculum Generation
- Melting Pot Contest
- The CityLearn Challenge 2023
- The HomeRobot Open Vocabulary Mobile Manipulation Challenge
- The Robot Air Hockey Challenge: Robust, Reliable, and Safe Learning Techniques for Real-world Robotics

## Deliverables

**1a/ Project proposal: 11:59pm on March 24, 2025**, posted publicly on Ed.

The project proposal is an intermediate milestone. The deliverable is a 1 page PDF submitted on Ed. Use [this LaTeX template](#); you do not have to use OverLeaf.

**1b/ 1:1 meeting with instructors: 11:59pm on Mar 31** (late deadline<sup>2</sup>: April 7, 2025).

During any of the office hours, have a 10 min conversation with a TA / instructor about your project. The TA/instructor will register that you've had this conversation. Note that office hours immediately before this deadline will be crowded, so we'd highly encourage students to have this meeting as soon as you've submitted your proposal.

**1c/ Peer feedback on proposal: 11:59pm on Mar 31, 2025**, submitted on Ed.

Post about a paragraph of suggestions on one other group's proposal. This is why we're using Ed for submissions. Each group member should do this independently. Please give feedback on a submission with < 3 comments so far. This means that collectively each proposal should get feedback from 3 students.

**2/ One experimental figure: 11:59pm on April 24, 2025**, added to this [shared slide deck](#).

You will create a high-quality experimental figure, adding it to the shared slide deck alongside a caption that explains (1) what this figure depicts and (2) what the reader is supposed to learn from

---

<sup>2</sup> We expect students to try to have these meetings by Mar 31, but will not penalize students who complete the meeting by April 7.

this figure. A figure should have a clear title, axis labels, a legend (if appropriate), appropriately-sized fonts. It should be high-resolution.

### 3/ Student presentations: 3.5 min, presented during reading period (**May 1, 2025**)

Each group will prepare a 3.5 min presentation about their project. The day and order of presentations will be randomly assigned after the project proposal submissions. The presentation should be prepared in [this shared slide deck](#). There is no limit to the number of slides, but a good rule of thumb is 1 min / slide. Each presentation should make sure to answer the following questions:<sup>3</sup>

- What is the problem?
- Why is it interesting and important?
- Why is it hard? (E.g., why do naive approaches fail?)
- Why hasn't it been solved before? (Or, what's wrong with previous proposed solutions? How does mine differ?)
- What are the key components of my approach and results? Also include any specific limitations.

We will randomly assign groups to one of the two 90-min sessions during reading week.

### 4/ Final written report. **May 11 at 10:30pm**

The final report will be 8 pages (excluding references). [Latex template is here](#); you do not have to use this template. Submit your final report to Gradescope. Make sure to do the following:

- Run a spelling and grammar checker.
- In each figure, make sure to label the axes and include a legend (if applicable). Font sizes should be roughly the same as the surrounding main-body text. The caption should tell the reader (1) what the figure is showing and (2) the main conclusion from the figure.
- Make sure the introduction motivates the problem that your project aims to solve.
  - For reproducibility papers, this is *not* the same as the motivation for the original paper, but rather the motivation for trying to reproduce this paper.
- In the related work section, make sure to explain the connections with *your work*.
- Signposting: each paragraph should start with an intro sentence, each section should start with an intro paragraph. If you just read the first sentence of each paragraph, does the paper make sense?

### 5/Code in final report. **May 11 at 10:30pm.**

Please add a **public GitHub repo link** to your final project report.

Some things to keep in mind:

1. **Please properly cite/comment/document your code for any *nonstandard* components.** i.e. if you are in an application track and use clean-rl's implementation of PPO, please cite this within your code, or write a note at the top of your train file that your code is based off of clean-rl/homework solutions/etc. If you use certain helper

---

<sup>3</sup> Borrowed from this excellent blog post: <https://cs.stanford.edu/people/widom/paper-writing.html#intro>

functions from another repo that are not key to the core algorithm *and* are nonstandard, please add a note/citation about this as well. **We understand that this may not be super clear cut, but please do your best here and err on the side of over vs. under-attribution.**

2. **Please make sure your GitHub repo is public and actually contains code.** The turnaround time for grading is quite tight, so if you do not correctly do so, we may not be able to give credit for code.

## FAQ

- For Track 1 (reproducing), the paper we've selected already provides code. Can we just run their code?
  - A: No. You are welcome to find another RL library and make the changes necessary to implement the proposed method. However, the group is expected to have actually implemented the key lines of code for the proposed method.
- Can I use late days for the project?
  - A: No.

## Grading Rubrics

**1a/ Project proposal (12%).** Does the proposal use the provided template and answer all the questions therein?

**1b/ 1:1 meeting with instructors (3%)**

**1c/ peer feedback (2%)**

**2/ one slide figure (8%)**

**3/ student presentations (15%)**

**4/ final written report (60%)**

### Student presentations

- Does the presentation stick to the allotted time? (5%)
- Does the presentation effectively answer the 5 questions listed in the description?
  - What is the problem? (2%)
  - Why is it interesting and important? (2%)
  - Why is it hard? (E.g., why do naive approaches fail?) (2%)
  - Why hasn't it been solved before? (Or, what's wrong with previous proposed solutions? How does mine differ?) (2%)
  - What are the key components of my approach and results? Also include any specific limitations. (2%)

### Final report

- Technical contributions (60%)
  - Track 1 (reproducibility):

- Does the project make a concerted effort to reproduce the results?
  - Does the project include thoroughly audit the main claims made in paper?
- Track 2 (new applications):
  - Application selection: is this actually a novel application of RL? Is it clear why this application cannot be solved with simpler tools (e.g., bandit algorithms, supervised learning)?
  - Does the project make a concerted effort to get RL to work on the new application?
- Track 3 (competition):
  - Does the project successfully run a baseline method for the competition?
  - Does the project make a concerted effort to beat the baseline and other competitors? This can be, i.e., implementing a new method on the competition or improving upon previous competitor methods.
- Organization + Communication (35%)
  - Organization/Signposting: Does each paragraph have an introduction sentence? Does each section start with an introduction paragraph? Is each section organized around a small set of ideas?
  - Introduction: Does the introduction answer the [5 important questions](#)?
  - Motivation: did you explain the problem? Did you motivate why this problem is important?
  - Relationship with prior work: did you clearly explain how the proposed project is similar to and builds upon prior work?
  - Method: did you clearly explain the method?
  - Limitations: did you mention potential limitations/assumptions of the project?
- Figures (5%)
  - Did you include at least one new figure? Did this figure clarify some part of the paper?
  - You are welcome to additionally provide images/figures from prior work (with proper attribution!), but these do not count towards the requirement that you include one new figure.

*Academic integrity.* See the syllabus for the official course policy, and come to office hours if you have additional questions.