
15-889: Advanced Topics in Decision-Making and Reinforcement Learning, Fall 2024

Instructor: Aviral Kumar; Teaching Assistant: Gokul Swamy

Time and Location: 3:30-4:55 pm, Mondays;

Instructor Contact: Email is preferred. aviralku@andrew.cmu.edu; gswamy@andrew.cmu.edu

Office Hours: Aviral (right after class, 4:55 - 5:55 pm); Gokul (Monday, 1:00-2:00 PM, EDSH 119)

Canvas Page: <https://canvas.cmu.edu/courses/41593>

Course Description

Reinforcement learning (RL) provides a general paradigm to convert prediction machines into effective decision-making agents, that can act intelligently in the world when deployed. Concretely, RL frames the problem of sequential decision-making as the problem of optimizing a scalar reward over the course of interaction. This framing is especially critical for obtaining intelligent behaviors in domains where either expert data is hard to collect (e.g., language models), the model must learn on its own via trial-and-error autonomously, or the actual objective is hard to specify (e.g., protein design). The goal of this class is to study advanced topics in this space of decision-making and reinforcement learning. In particular we will focus on algorithms, applications, as well as upcoming topics in the field to make students ready for research in this area. Students will practice essential research skills including reviewing papers, writing research project proposals, executing on research ideas, and technical communication.

Prerequisites

The course is open to all SCS graduate students without strict prerequisites. Familiarity with machine learning, basics of AI, probability, statistics is strongly encouraged. Interested undergraduate students with a strong background or PhD students from other departments may seek approval from the instructor.

Grading

- Participation (5%): Regular attendance and engagement
- Paper Summaries (10%): 1-2 paragraph summary of reading on Canvas
- Paper Presentations (25%): Role-playing in-class presentations on readings
- Midterm Project Report (20%) Literature survey and preliminary results (~2 pages)
- Final Project (40%): Final project report (~6-8 pages, single column) and a final presentation

Learning Objective

At the end of this class, you will be able to: (1) formalize decision-making problems in the terminology and language of reinforcement learning, (2) understand and identify the pros and cons of various approaches to tackle these problems, (3) code up and run the right approach on your problem. From a research perspective, you should be able to: (1) come up with new algorithmic ideas in the area, (2) plan a research project and take the first steps, (3) prepare a scientific presentation or talk and critique research papers.

Paper Summaries & Presentations

There will be several paper discussion days during which you will be assigned research papers to read. You are expected to complete all assigned readings before class and come prepared with comments and questions

to discuss with the group. You will share 1–2 paragraphs with your takeaways or questions on each reading on Canvas, by 10am ET of the day the reading will be discussed. **Please note that the use of LLMs to generate these summaries is entirely forbidden and will result in failing the class.**

The paper discussions will involve role-playing student seminars inspired by Alec Jacobson, Aditi Raghunathan, and Colin Raffel. We will be adopting the following roles inspired by Aditi Raghunathan's class.

- *Positive reviewer*: who advocates for the paper to be accepted at a conference (e.g., NeurIPS)
- *Negative reviewer*: who advocates for the paper to be rejected at a conference (e.g., NeurIPS)
- *Archaeologist*: who determines where this paper sits in the context of previous and subsequent work. They must find and report on at least one older paper cited within the current paper that substantially influenced the current paper and at least one newer paper that cites this current paper. Keep an eye out for follow-up work that contradicts the takeaways in the current paper
- *Academic researcher*: who proposes potential follow-up projects not just based on the current paper but also only possible due to the existence and success of the current paper

Tentative Topics

- **Basics of RL** (MDPs, POMDPs, supervised learning vs RL, value functions, policy gradient)
- **RL algorithms** (model-based, model-free, hybrid algorithms; value function learning, policy learning, online vs offline, challenges in value learning, policy learning, model learning, representation learning)
- **Exploration** (bonus-based algorithms, uncertainty estimation for MDPs, POMDPs with structure, meta reinforcement learning)
- **Offline RL** (conservative RL, behavior constraints, uncertainty estimation, online RL fine-tuning of offline RL policies, offline RL vs imitation learning, stitching and policy improvement)
- **RL with foundation models**:
 - RL fine-tuning of language models (RLHF, RL for self-improvement, multi-turn RL, model-based planning and search, meta learning)
 - RL with transformers (challenges, algorithms, re-evaluating pros and cons of existing approaches)
- **Scaling up RL** (ingredients for scaling up RL; challenges with loss functions / implicit regularization, compute vs data ratio; relative tradeoffs imposed by different algorithms; case studies)
- **Applications of RL on real-world problems** (chip design, recommendation systems, biology)

Class Project

Your class project can be either a thorough literature review (~25 relevant papers, organized so that it identifies gaps in the state of the art) or an exploration of an original research idea. You can choose to work individually or in groups of up to three.

Project proposal | October 4, 2024 11:59pm ET. This is a brief (1 page) summary of your final project. Think of this as an extended abstract: you want to motivate the topic you have chosen and the technical questions that you want to investigate. By this stage, you should have decided if you are doing a project on your own or in a group. This required proposal will allow the instructors to give you early feedback and help you refine the project scope.

Mid-term report | November 10, 2024, 11:59 pm ET. This is intended as a checkpoint to ensure that you are making progress towards your final project. The report length should be a typical robotics workshop paper (2 pages, single-column). You can either submit a preliminary literature review, or preliminary exploration of your research idea.

Project presentation | December 2, 2024, in class. Presentations will be 20 minutes long with 5 minutes for questions. A good presentation will identify and contextualize the problem in the context of prior work and state of the art, state the key idea of the project, and include early results if applicable. All presentation slides should be submitted by the day before presentations start.

Final project reports | December 6, 2024, 11:59 pm ET. The final report should present your final findings in a research paper format. Target length should be a typical ML conference paper (8 pages, single column).

All project deliverables will be submitted through Canvas under the corresponding Assignment. Only one submission per team is required as long as all team members are clearly identified.

Attendance

Class attendance and participation are key for both your and your peer's success in this class. You are expected to attend class in person during the scheduled time, including the final presentations. We understand that occasionally you may have challenges attending (e.g., illness, religious observance, etc.). However, if you anticipate having a challenge regularly attending class, please contact the instructors.

Academic Integrity

Honesty and transparency are critically important. On the flip side, plagiarism and cheating are serious academic offenses with serious consequences. If you are discovered engaging in either behavior in this course, you will earn a failing grade on the assignment in question, and further disciplinary action may be taken. You are encouraged to work together on projects and homework assignments and to make use of campus resources like Student Academic Success Center (SASC) to assist you in your pursuit of academic excellence. However, please note that in accord with the university's policy you must acknowledge any collaboration or assistance that you receive on work that is to be graded, either from a person, reference, or an LLM. **The weekly paper summaries must be entirely your work with no assistance from another person or LLM.**

Late Policy

All homeworks and assignments are assigned due dates and should be submitted through the relevant Canvas portal. If you cannot submit an assignment on time, my default will be to reduce the grade by 10% for each 24 hour period, up to three days, that the assignment is late. This will be automatically applied; you do not have to request it. After three days, the assignment will receive a zero. If you experience an unforeseeable emergency and would like me to consider waiving the late penalty, please email the instructors as early as possible to discuss this request. The 10per day deduction does not apply to unexcused late presentations, which will receive a zero immediately, because they will affect our ability to hold class. Re-scheduling presentations will be based on schedule availability and the instructor's discretion.

Acknowledgements. Policies and format of this class is heavily inspired from prior offerings of 15-884 (Prof. Aditi Raghunathan) and 16-886 (Prof. Andrea Bajcsy).