COMP 5900J
Reinforcement Learning
Course Outline

Instructor: Junfeng Wen (junfeng.wen [AT] carleton.ca)

Winter 2024
School of Computer Science
Carleton University

Course Information

Instructor: Junfeng Wen

Contact: junfeng.wen [AT] carleton.ca

Classroom: Room location is posted on the public class schedule

Lectures: Mondays & Wednesdays 10:05 am - 11:25 am

Office Hours: By appointment

Brightspace access for University of Ottawa Students; please see information here.

For information about Carleton's academic year, including registration and withdrawal dates, see Carleton's Academic Calendar.

Course Description

Reinforcement learning (RL) focuses on learning by interacting with a complicated environment through trial and error. In this research-oriented course, students will first learn the fundamental concepts of RL, including Markov decision processes, value prediction, and optimal control. Then we will investigate several advanced topics in the recent RL literature, such as imitation learning, batch RL, and distributional RL. Priority will be given to OCICS joint institute students and students in thesis-based research programs.

Recommended Textbooks

- *Reinforcement Learning: An Introduction* (2nd edition), Sutton Barto
- *Algorithms for Reinforcement Learning*, Szepesvari
- *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Puterman

Prerequisites

Students are expected to be familiar with linear algebra, calculus, basic statistics and Python programming.

Topics Covered and Learning Outcomes

This course will cover the following (tentative) topics

- Markov decision process
- Dynamic programming methods
- Monte-Carlo methods
- Temporal difference learning
- Prediction and control with function approximation
- Policy gradient and actor-critic methods
- Common deep RL algorithms
- Offline/Batch RL
- Distributional RL
- Exploration
- RL applications

Upon completion, students should be able to

- Develop a solid understanding of the fundamental concepts and principles in reinforcement learning
- Understand a wide range of reinforcement learning algorithms, their applicability, strengths and weaknesses
- Design and implement reinforcement learning algorithms for real-world problems, and evaluate their performance

Evaluation

Two assignments 40% (20% each)

- Done individually
- For assignments, you have three excused days **throughout the term** (rounded up to the nearest day) to account for any unforeseeable difficulties. After that no late submission will be accepted without proper justifications
- Submissions are handled electronically (i.e., through Brightspace). Technical problems do not exempt you from late policy, so if you wait until the last minute and then have issues with your connection, it will still count as a late submission. Consequently, you are advised to
    1. Periodically upload you progress
    2. Attempt to submit your final submission early (e.g., at least one hour in advance of the due date and time) and
    3. Download the submitted files to make sure they are correct

Paper review 10%

Paper presentation 10%

Project 40%

Intellectual Property

Sharing assignment or quiz specifications or posting them online (to sites like Chegg, CourseHero, OneClass, etc.) is considered academic misconduct. You are never permitted to post, share, or upload course materials without explicit permission from your instructor. Academic integrity offences are reported to the office of the Dean of Science. Penalties for such offences can be found on the ODS webpage.

University Polices

Academic Accommodations. Carleton is committed to providing academic accessibility for all individuals. Please review the academic accommodation available to students here.

Student Academic Integrity Policy. Every student should be familiar with the Carleton University student academic integrity policy. A student found in violation of academic integrity standards may be awarded penalties which range from a reprimand to receiving a grade of F in the course or even being suspended or expelled from the program or University. Examples of punishable offences include: plagiarism and unauthorized co-operation or collaboration. Any such reported offences will be reviewed by the office of the Dean of Science. More information on this policy may be found on the ODS Academic Integrity page.

Plagiarism. As defined by Senate, "plagiarism is presenting, whether intentional or not, the ideas, expression of ideas or work of others as one's own". Such reported offences will be reviewed by the office of the Dean of Science. More information and standard sanction guidelines can be found here.

Unauthorized Collaboration. Senate policy states that "to ensure fairness and equity in assessment of term work, students shall not co-operate or collaborate in the completion of an academic assignment, in whole or in part, when the instructor has indicated that the assignment is to be completed on an individual basis".