ECE/ISE 7202*– Reinforcement Learning

Autumn 2022

1 Course Overview

Mode of delivery: Lectures in person. Office hours in-person on Tuesdays, and online on Thursdays.

Lectures: Tuesdays and Thursdays, 9:35am-10:55am. Location TBD (classroom is likely changing, so please check Buckeyelink).

Instructor: Parinaz Naghizadeh
Email: naghizadeh.1@osu.edu
Office hours: Tuesdays 11:30am-12:30pm (in person, 238 Baker systems) and Thursdays 4-5pm (Zoom), or by appointment (please contact me at least 2 days in advance).

Catalog description: This course explores the fundamentals of Markov decision processes and reinforcement learning algorithms. Topics to be covered include the framework of Markov decision processes, exact dynamic programming, value and/or policy space approximations, and RL algorithms including Q-learning, policy gradient, actor-critic, and temporal differences methods. Prerequisite: graduate standing, or permission by instructor. Knowledge of probability, optimization, and advanced calculus recommended. Units: 3 credit hours.

Informal description: This course focuses on the fundamental ideas behind different reinforcement learning methods by starting at their roots in dynamic programming. We will start with the framework of Markov decision processes, introduce exact dynamic programming, and subsequently study potential approximations in value and/or policy space, exploring RL algorithms such as Q-learning, policy gradient, and actor-critic methods. Time-permitting, the course will also cover some advanced topics. The broad areas to be covered are:

1. Sequential decision making and multi-armed bandits
2. Markov decision processes
3. (Exact) dynamic programming
4. Tabular and Temporal difference learning (Q-learning, SARSA, and other TD methods)
5. Approximate solutions methods (in the value space)
6. Policy gradient methods
7. Actor-critic methods
8. A subset of the following topics (time-permitting): Multi agent and distributed RL; Off policy evaluation and optimization; Reward design and reward shaping; Safe and efficient exploration in RL; Open problems (frontiers, RL and the future of AI).

A tentative schedule of topics to be covered in each lecture is available in Table 1.

Prerequisites: Graduate standing in Engineering or Math. While the course will include implementation of different reinforcement learning algorithms, our primary focus will be on the mathematical foundations behind these various RL algorithms. The students taking this course should therefore have a good knowledge of probability, optimization, calculus, and some linear algebra. This will provide the mathematical maturity to follow and appreciate the fundamentals and reasoning behind the algorithms covered in our course. Knowledge of, or willingness to learn, programming in Python, Matlab, or a language with similar capabilities, will be needed for the coding homework.

*The course is cross-listed. You are welcome to enroll under either ECE 7202 or ISE 7202.

2 Textbooks and Readings

Readings and reference books: There is no required course textbook. The course will primarily progress through the topics covered in "An Introduction to Reinforcement Learning", by Sutton and Barto (2018) *(available at* `http://incompleteideas.net/book/the-book-2nd.html` *).*

Other recommended references are listed below:

- "Reinforcement Learning and Optimal Control" by Dimitri P. Bertsekas, 2019.
- "Algorithms for Reinforcement Learning", by Csaba Szepesvari, 2010. *(available online at* `https://sites.ualberta.ca/~szepesva/papers/RLAlgsInMDPs.pdf` *)*
- "Reinforcement Learning and Stochastic Optimization: A unified framework for sequential decisions", by Powell, 2021. *(current draft available online at* `https://castlelab.princeton.edu/RLSO/` *)*
- "Dynamic Programming and Optimal Control" (2 volumes), by Bertsekas, 2017.
- "Neuro-Dynamic Programming", by Bertsekas and Tsitsiklis, 1996.

We will refer to some (research) papers throughout the course, especially when covering advanced topics, and also for the homework. I will make either the pdf files or the links to the online resources available on the course page.

Copyright disclaimer: The materials used in connection with this course may be subject to copyright protection and are only for the use of students officially enrolled in the course for the educational purposes associated with the course. Copyright law must be considered before copying, retaining, or disseminating materials outside of the course.

3 Grading

Grading: 40% Homework, and 60% Final Project.

Homework will be assigned roughly on a biweekly basis and will be due in one week. You are given *4 late homework days*, to invoke and use as you wish throughout the semester. Late homework beyond this will be accepted only if prior arrangements are made with me, or due to a documented unexpected reason such as a serious illness. We will most likely have 5 homeworks. One of the questions on all homeworks will ask for a critical reading of one or two relevant research papers, and there will typically be a programming question too. Homework is typically due on Fridays, and will be submitted through the CarmenCanvas page. Grading will be in the form of spot-grading or peer-grading (details to be shared in class).

Final Project: This can take on a few forms: (1) the formulation of a problem in the student's field of active research or general interest, and either an analytical evaluation of an RL algorithm on the problem, a computational implementation of a solution using one of the reinforcement learning methods (covered in the course or beyond), or a mixture of both, and (2) an *in depth* literature review on a *recent*, advanced topic in RL. Attempts to extend an existing research paper or an existing RL algorithm will also be allowed, but there should be a prior meeting to ensure there can be substantial extensions over the existing work. You are encouraged to pick a topic well in advance and discuss it in an office hour or separate meeting to make sure it fits the project requirements.

Group projects with 2-3 members will be encouraged, depending on the number of enrolled students. Each student/group will prepare a recording presentation of their work (tentatively, one 3min "lightning version" and one 15min "conference" style), and turn in a final paper based on their work and findings. More details on the format of the final report and interim deadlines for forming teams and/or picking topics will be given separately. The grade for the final project will be assessed based on completion of intermediate steps (topic selection, interim report), recorded presentations, and final report.

4 Course Outline

A tentative schedule of topics to be covered in each lecture, and tentative deadlines for homeworks and the project (in **bold**), are given in Table 1.

Table 1: Lecture outline (subject to change)

Lecture	Date	Topics covered
1	8/23	Introduction to reinforcement learning and its applications, course logistics
2	8/25	An example and introduction to terminology
3	8/30	Sequential decision making: Multi-armed bandits I
	8/31	**(Bonus) Homework 0 due**
4	9/1	Multi-armed bandits II
5	9/6	Markov decision processes I
6	9/8	Markov decision processes II
	9/9	**Homework 1 due**
7	9/13	Markov decision processes III
8	9/15	Dynamic programming I
9	9/20	Dynamic programming II
10	9/22	Dynamic programming III
	9/23	**Homework 2 due**
11	9/27	Dynamic programming wrap up, introduction to Monte Carlo methods
12	9/29	Monte Carlo methods continued
13	10/4	Temporal difference learning: TD methods
14	10/6	Temporal difference learning continued: SARSA, Q-Learning
	10/7	**Homework 3 due**
15	10/11	Wrap up and summary of tabular methods
	10/13	*Fall break. No class.*
16	10/18	Approximate solution methods: RL with function approximation I
17	10/20	RL with function approximation II
	10/21	**Project: interim report due**
18	10/25	RL with function approximation III
19	10/27	RL with function approximation, wrap up
20	11/1	Policy gradient methods I
21	11/3	Policy gradient methods II
	11/4	**Homework 4 due**
22	11/8	Actor-critic methods
23	11/10	Wrap up and summary of function approximation methods
24	11/15	TBD: advanced topics I
25	11/17	TBD: advanced topics II
	11/18	**Homework 5 due**
26	11/22	TBD: advanced topics III
	11/24	*Thanksgiving Day. No class.*
27	11/29	TBD: advanced topics IV
28	12/1	Overview of other advanced topics. Concluding thoughts.
29	12/6	Final project lightning presentations (tentative)
	12/12	**Project: final reports due**

5 Other Course Policies

Health and safety requirements: All students, faculty and staff are required to comply with and stay up to date on all university safety and health guidance (`https://safeandhealthy.osu.edu`), which includes wearing a face mask in indoor spaces (at the time of writing of this syllabus). Non-compliance will be warned first and disciplinary actions will be taken for repeated offenses.

Academic misconduct statement: The Ohio State University and the Committee on Academic Misconduct expect that all students have read and understand the University's Code of Student Conduct, and that all students will complete all academic and scholarly assignments with fairness and honesty. Students must recognize that failure to follow the rules and guidelines established in the University's Code of Student Conduct and this syllabus may constitute "Academic Misconduct," as defined in the University's Code of Student Conduct (Section 3335–23–04). Any student found to have engaged in academic misconduct will be subject to disciplinary action by the university. Please contact me if you have any questions about what might constitutes academic misconduct in this course.

Statement on Title IX: All students and employees at Ohio State have the right to work and learn in an environment free from harassment and discrimination based on sex or gender, and the university can arrange interim measures, provide support resources, and explain investigation options, including referral to confidential resources.

If you or someone you know has been harassed or discriminated against based on your sex or gender, including sexual harassment, sexual assault, relationship violence, stalking, or sexual exploitation, you may find information about your rights and options at `titleix.osu.edu` or by contacting the Ohio State Title IX Coordinator at titleix@osu.edu. Title IX is part of the Office of Institutional Equity (OIE) at Ohio State, which responds to all bias-motivated incidents of harassment and discrimination, such as race, religion, national origin, and disability. For more information on OIE, visit `equity.osu.edu` or email equity@osu.edu.

Your mental health: As a student you may experience a range of issues that can cause barriers to learning, such as strained relationships, increased anxiety, alcohol/drug problems, feeling down, difficulty concentrating and/or lack of motivation. The Ohio State University offers services to assist you with addressing these and other concerns you may be experiencing. If you find yourself feeling isolated, anxious or overwhelmed, please know that there are resources to help: `ccs.osu.edu`. You can reach an on-call counselor when CCS is closed at (614) 292-5766 and 24 hour emergency help is also available through the 24/7 National Prevention Hotline at 1-(800)-273-TALK or at `suicidepreventionlifeline.org`. The Ohio State Wellness app is also a great resource available at `go.osu.edu/wellnessapp`.

Accessibility accommodations for students with disabilities: The University strives to make all learning experiences as accessible as possible. In light of the current pandemic, students seeking to request COVID-related accommodations may do so through the university's request process, managed by Student Life Disability Services. If you anticipate or experience academic barriers based on your disability (including mental health, chronic or temporary medical conditions), please let me know immediately (within the first two weeks of the course) so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion. SLDS contact information: slds@osu.edu; 614-292-3307; `slds.osu.edu`; 098 Baker Hall, 113 W. 12th Avenue.

Accessibility of course technology: This course requires use of Carmen (Ohio State's learning management system) and other online communication and multimedia tools. If you need additional services to use these technologies, please request accommodations with your instructor. Information about accessibility within CarmenCanvas and CarmenZoom is avilable in the following links: Carmen Canvas and CarmenZoom.