ECE 457C - Reinforcement Learning

Course Description

Introduction to Reinforcement Learning (RL) theory and algorithms for learning decision-making policies in situations with uncertainty and limited information. Topics include Markov decision processes, classic exact/approximate RL algorithms such as value/policy iteration, Q-learning, State-action-reward-state-action (SARSA), Temporal Difference (TD) methods, policy gradients, actor-critic, and Deep RL such as Deep Q-Learning (DQN), Asynchronous Advantage Actor Critic (A3C), and Deep Deterministic Policy Gradient (DDPG).

Note: This course was formerly known as [ECE 493 Topic 42 - Reinforcement Learning](#)

Links to More Information and Resources

[**Full Weekly Schedule and Deadlines (S22 version)**](#)	[**Piazza Discussions**](#)
[**Textbook**](#)	[Course YouTube Channel](#) (from previous terms)
[Course Topics](#)	[Assignments and Tests](#) (weights and deadlines)
[Gingko List of Resources and Links](#)	[S22 Website (OLD)](#)
[Getting Help](#)	[Academic Policies](#)
[Health and Wellness](#)	[COVID-19](#)

Required Background

The course will use concepts from ECE 203 and ECE 307 on Bayesian Probability and Statistics, these will be reviewed but familiarity will help significantly. All other concepts needed for the course will be introduced directly. Examples, assignments and projects will depend on programming ability in Python.

Spring 2023 - Course Staff

Instructor: Prof. Mark Crowley		TA: Martin Ethier
Contact: mcrowley@uwaterloo.ca (but for better results, use piazza or come to office hours)		**Contact:** methier@uwaterloo.ca
Office Hour: One-on-one [bookable on outlook](#) by (E5 4114 (or online))		

Lecture Locations and Times

Course Dates: May 8, 2023 to Aug 1, 2022

Type	Times	Room
Lecture	MF 1:00 pm - 1:00pm	E7 4043
Tutorial *(Mostly for review sessions, will determine based on need.)*	M 7:00pm - 7:50pm	E7 4053
Office Hours	Mondays 3:30-4:30pm	E5 4114 or online via Teams (use [bookable calendar here](#))

Course Learning Objectives and Topic Details

Learning Objectives

This course complements other AI courses in ECE by focussing on the methods for representation and reasoning about uncertain knowledge for the purposes of analysis and decision making. At each stage of the course we will look at relevant applications of the methods being discussed.

For example, in 2016 the AI program "AlphaGO" defeated human world class players of the game Go for the first time. This system requires many different methods to enable reasoning, probabilistic inference, planning and decision optimization. In this course we will build up the fundamental knowledge about these components and how they combine together to make such systems possible.

1. Identify and Explain the component theoretical concepts of Reinforcement Learning systems.
2. Implement or instantiate using a library any of the core Reinforcement Learning algorithms on a variety of domains.

3. Evaluate the performance of a particular RL system on a given domain through proper experimental design, statistical analysis and visualization.

Minor Objectives

1. Learn to write mathematical notation electronically using **LaTex** (also useful for Word/Google Docs/Markdown).

2. Practice and feedback on reviewing, summarizing and utilizing theory and results from academic papers.

Topics

1. Motivation and Context

   - Importance of reasoning and decision making about uncertainty.

   - Connection to Artificial Intelligence and Machine Learning.

   - Probability review.

2. Decision making under uncertainty:

   - Multi-Armed Bandit (MAB) problems, Thompson Sampling.

   - Markov Decision Processes (MDPs)

   - Influence Diagram representation

3. Solving MDPs

   - Theory, Bellman equations

   - Relation to Control Theory

   - Value Iteration, Policy Iteration

4. The Reinforcement Learning Problem

   - Approximately solving MDPs by interacting with the environment

   - SARSA algorithm

   - Q-learning algorithm

   - Other variants of these

5. Temporal Difference Learning

   - Eligibility Traces

   - TD(λ)

6. State Representation

   - Value Function Approximation

   - Stochastic Gradient Descent

7. Basics of Neural Networks (review or refer to ECE657A content)

   - fully connected, multi-layer perceptrons

   - supervised training, back-propagation

- CNNs, LSTMS
- regularization methods

8. Deep Reinforcement Learning
   - Deep Q- Networks (DQN)
   - Experience replay buffers, mini-batch training and other methods of training and architecture

9. Direct Policy Search
   - Policy Gradients methods
   - Actor-Critic methods
   - Discussion of New Algorithms:
     - A3C, A2C, DPG, DDPG, TRPO, PPO, SAC

10. Other ways to solve (PO)MDPs

- Monte-Carlo Tree Search, Explaining AlphaGo, and beyond

11. Other Challenges (brief)
    - Hierarchical RL
    - Training Spectrum from Supervised to Curriculum to Self-supervised Learning
    - Partially Observable MDPs (POMDPs) (skipped in S22)
    - Multi-Agent RL (MARL)
    - Curiosity based learning
    - Soft-Actor Critic

12. Wrap-up and Review

Assessment

Assignments can be done alone or in pairs.

Item	Weight Towards Final Grade	Released (12:00 am)	Due (11:59pm)	Days	Notes
Assignment 1	10.0%	Monday, May 15, 2023	Friday, June 2, 2023	16	**Writing/LaTeX: Probability, MAB, MDPs**
Assignment 2	15.0%	Wednesday, May 31, 2023	Friday, June 23, 2023	23	**Python: Tabular RL on Maze**
Midterm Week (no midterm, regular classes)	0.0%	Monday, June 19, 2023	Friday, June 23, 2023		
Assignment 3	15.0%	Friday, June 23, 2023	Friday, July 14, 2023	21	**OpenAI Gym: Basic Deep RL On Maze, Literature Review**
Assignment 4	20.0%	Friday, June 23, 2023	Monday, July 31, 2023	38	**Deep RL on Bigger Problem, Choice of additional domain or algorithm to use**
Final Exam	40.0%	TBD	TBD		
Total	100%				

Course Tools

I know there has been app/feature/tool creep in courses as they the pandemic has worn on, we're trying to minimize that while still not holding ourselves back when a new tool does something better than an old one.

- **Course Website** :https://markcrowley.ca/rlcourse/
  - Course News, Outline, Learning Goals
  - Schedule of lectures, assignments and tests
  - Links to all resources
- **Learn** : *Log in to learn.uwaterloo.ca*
  - Online Course management system for UWaterloo.

- Your grades will be managed here, up until the final grade submission phase of the course.

    - Links, announcements and course materials will all be made available here as well.

    - Only registered students can access learn.

- **Piazza** : ECE457C Discussions - TBD

    - Online, threaded discussion forum with at the ability for students to construct an answer in addition to the answer provided by course staff.

- **Crowdmark** : *(links will be made available as needed)*

    - A visual grading tool for pdfs submissions of tests and assignments, allows limited online test with mark-down text entry and multiple choice questions.

    - Used by the course staff for grading your tests.

    - Some assignments and tests might be made available online for submission using this tool as well.

Getting Help:

- **Discussion board:**

    - *Piazza* will be the main place for detailed discussion and questions. Students can post anonymously (from students only), post a collaborative answer and course staff can confirm these, post their own or run Live Q&A events.

    - Go there there and sign up with your UWaterloo email now!

- **Pre-recorded Video Lectures:** These will be made available on the [course youtube channel](http://uwaterloo.ca/accessability-services), and links from within Learn

- **LEARN Website:** The main course content, announcements, grade tracking and materials will be made available on Learn. All registered students should see this in their LEARN courses.

- **Email the Teaching Assistant and Instructor:** Office Hours will be arranged once term starts as needed.

- **AccessAbility Services :** http://uwaterloo.ca/accessability-services

    - If you need any accommodation, assistance with exams, learning environment, assignments, then contact them for help setting you up as securely and anonymously as possible.

Discussion Group Protocols

- Posts on Piazza can be public or anonymous to your classmates, but they will *never* be anonymous to the TAs and Instructor.

- Be kind. Assume the best, not the worst. Think before you hit enter.

- Posts which are considered offensive, abusive, bullying, discriminatory to any group or person, will be made private or deleted and followed up with private discussion.

- If you feel there is inappropriate, hurtful behaviour occurring on the discussion forum, please notify the professor, TAs or department staff as you feel appropriate.

Recipe for success:

- Ask questions.

- Connect with your classmates.

- Do the assignments.

- *Ask questions!*

- Most of all, *have fun*! ...yes really.

Course Resources

Primary Reference for Course

[SuttonBarto2018] - Reinforcement Learning: An Introduction. Book, free pdf of draft available.
http://incompleteideas.net/book/the-book-2nd.html

Other Useful Texts

[Dimitrakakis2019] - Decision Making Under Uncertainty and Reinforcement Learning

http://www.cse.chalmers.se/~chrdimi/downloads/book.pdf

[Ghavamzadeh2016] - Bayesian Reinforcement Learning: A Survey. Ghavamzadeh et al. 2016.
https://arxiv.org/abs/1609.04436

- More probability notes online: https://compthinking.github.io/RLCourseNotes/

Open AI Reference Website

This website is a great resource. It lays out concepts from start to finish. Once you get through the first half of our course, many of the concepts on this site will be familiar to you.

- Key Papers in Deep RL List - https://spinningup.openai.com/en/latest/spinningup/keypapers.html

- Fundamental RL Concepts Overview - The fundamentals of RL are briefly covered here. We will go into all this and more in detail in our course.
https://spinningup.openai.com/en/latest/spinningup/rl_intro.html

- Family Tree of Algorithms - Here, a list of algorithms at the cutting edge of RL as of 1 year ago to so, so it's a good place to find out more. But in a fast growing field, it may be a bit out of date about the latest now.
  https://spinningup.openai.com/en/latest/spinningup/rl_intro2.html

COVID-19 Policies

Health and Safety

- see these slides on university policy for COVID-19 safety once in-person classes begin again.

- **Attendance: Students are to be instructed to attend only the section for which they are registered.** If you wish to attend a different section (less people are registered for section 2) you should transfer to that section using official means.

- **Absence:** Students **_shall not attend class if they are experiencing influenza-like illness_**, have been in close contact with someone who is ill, or have travelled outside of Canada within the past 14 days. You will be able to engage with the course content online while reducing the risk of others becoming ill.

- **Face coverings:** Wearing of face-covering/mask is a **requirement** in all common areas on campus, including all indoor instructional spaces.

  - Students who will not wear masks will be asked to leave the classroom. If the student has a medical reason why they cannot wear a mask they should contact the professor electronically and provide proof of this.

  - As such, no food is allowed to be consumed in instructional space. Beverages are allowed if a straw is used or if the mask is lowered only for a brief period.

  - When a student asks or answers a question it may be difficult for them to be heard while wearing a mask. A student may briefly lower their mask to ask/answer the question and then the mask must be replaced.

- **Hand hygiene:** Students are expected to practice frequent hand hygiene (handwashing with soap and water or use of hand sanitizer), including immediately before coming into an instructional space

- **Seating:** Students are permitted to sit where they wish. Students are encouraged to sit with one seat left empty between them and other students when possible.

- **Student illness:** In the event of absence due to influenza-like illness or required self-isolation, students shall submit an Illness Self-declaration. Students can find the Illness Self-declaration form in the Personal Information section of Quest. A doctor's note for accommodation is not required.

Fair Contingencies for Emergency Remote Teaching

We are facing unusual and challenging times. The course outline presents the instructor's intentions for course assessments, their weights, and due dates in Winter 2022. **As best as possible, we will keep to the specified assessments, weights, and dates**. To provide contingency for unforeseen circumstances, the instructor reserves the right to modify course topics and/or assessments and/or weight and/or deadlines with **due and fair notice to students**. In the event of such challenges, **the instructor will work with the Department/Faculty to find reasonable and fair solutions that respect rights and workloads of students, staff, and faculty**.

Wellness Support and Contact Information.

University can be a challenging environment and it is normal to need support from time-to-time. Campus Wellness services are available to students through counselling and health services. If you are struggling or need someone to talk to you, please reach out.

To book an appointment or learn more about the services, call 519-888-4567 x 32655 or explore www.uwaterloo.ca/campus-wellness.

If you're experiencing a crisis and feel unable to cope and Campus Wellness is closed, contact any of these after-hours supports: EmpowerMe (1-833-628-5589), Good2Talk (1-866-925-5454) or Here 24/7 (1-844-437-3247). They are available at any time of the day or night to help.

General University of Waterloo Guidelines:

Academic Integrity: In order to maintain a culture of academic integrity, members of the University of Waterloo community are expected to promote honesty, trust, fairness, respect and responsibility. Check http://www.uwaterloo.ca/academicintegrity/ for more information.

Grievance: A student who believes that a decision affecting some aspect of his/her university life has been unfair or unreasonable may have grounds for initiating a grievance. Read Policy 70, Student Petitions and Grievances, Section 4, http://www.adm.uwaterloo.ca/infosec/ Policies/policy70.htm. When in doubt please be certain to contact the departments administrative assistant who will provide further assistance.

Discipline: A student is expected to know what constitutes academic integrity---check http: //www.uwaterloo.ca/academicintegrity/ to avoid committing an academic offence, and to take responsibility for his/her actions. A student who is unsure whether an action constitutes an offence, or who needs help in learning how to avoid offences (e.g., plagiarism, cheating) or about rules for group work/collaboration should seek guidance from the course instructor, academic advisor, or the

undergraduate Associate Dean. For information on categories of offences and types of penalties, students should refer to Policy 71, Student Discipline, http://www.adm.uwaterloo.ca/infosec/Policies/policy71.htm. For typical penalties check Guidelines for the Assessment of Penalties, http://www.adm.uwaterloo.ca/infosec/guidelines/penaltyguidelines.htm.

Appeals: A decision made or penalty imposed under Policy 70 (Student Petitions and Grievances) (other than a petition) or Policy 71 (Student Discipline) may be appealed if there is a ground. A student who believes he/she has a ground for an appeal should refer to Policy 72 (Student Appeals) http://www.adm.uwaterloo.ca/infosec/Policies/policy72.htm.

Note for Students with Disabilities: The Office for Persons with Disabilities (OPD), located in Needles Hall, Room 1132, collaborates with all academic departments to arrange appropriate accommodations for students with disabilities without compromising the academic integrity of the curriculum. If you require academic accommodations to lessen the impact of your disability, please register with the OPD at the beginning of each academic term.