University of Alberta

CMPUT 656: Human-in-the-Loop Reinforcement Learning
Fall 2023

Instructor: Matthew E. Taylor (Matt)
TA: Videh Nema
E-mail: matthew.e.taylor@ualberta.ca (make sure "656" is in the subject)
Web Page: https://sites.google.com/ualberta.ca/F23cmput656/home

Office Hours: Before & after class, and by appointment: https://matttaylor.as.me/schedule.php

Course Format: In person, T/Th 11am-12:20pm. CAB 357

COURSE CONTENT

Course Description

Traditional reinforcement learning (RL) systems assume a single agent interacts with a Markov decision process (MDP). A multi-agent RL setting assumes there are multiple agents. But what happens when one or more of the agents in a system is actually a human? Agents could learn from humans, humans could learn from agents, or agents and humans could learn to accomplish team goals together. This course will provide you with the background and tools needed to conduct research in this emerging area of human-in-the-loop (HitL) AI.

Course Outcomes & Learning Outcomes

- Gain hands-on experience designing, running, and analyzing human subject experiments
- Implement multiple human-in-the-loop RL algorithms, such as learning from demonstration, inverse reinforcement learning, learning from feedback, etc.
- Read and understand a number of classic and cutting edge papers in this area
- Present at least two papers to the class, giving you a chance to go deeper into these papers and work on presentation skills
- Practice giving constructive feedback to others on their presentations and research
- Work on a group project that could form the basis of a future publication or even an MSc thesis

Course Co/Prerequisites

You should have completed one RL class for credit, or be enrolled in RL 1 with Marlos this fall. If you cannot meet these requirements, we should talk about if your non-class experience will be enough for you to:

- Implement basic RL algorithms
- Download and run RL algorithms and RL environments
- Empirically and theoretically evaluate the performance of different RL algorithms
- Understand how to tune hyper parameters in RL algos

Course Schedule

We will discuss the course schedule the first week of the semester and keep it updated on the website.

Course Schedule (subject to change)

	T	Th	Topic	Reading	Deliverables
1	9/5		Introduction / background, Course standards		
		9/7			
2	9/12		Human subject studies		
		9/14			Pick first paper to present
3	9/19				
		9/21			HIPPO Gym exercise
4	9/26				
		9/28			
5	10/3				
		10/5			Group formation and initial proposal
6	10/10				
		10/12			Pick second paper to present
7	10/17				
		10/19			Refined proposal and pilot study
8	10/24				
		10/26			
9	10/31				
		11/2			Pilot study results
10	11/7				
		11/9			First draft of writeup due
12	11/14		*No class: reading week*		
		11/16			
13	11/21				
		11/23			Refined study
14	11/28				
		11/30			
15	12/5		Presentations		Present results and conclusions to class
		12/7	Presentations		Present results and conclusions to class
16	12/10		*No class: course finished*		Final writeup due

Online systems used: We'll mostly use slack for communication. I've been told that "Students MUST NOT use their @ualberta email address or CCID to register into the system and instead should use a non-identifying email address or account". But please do set your slack user name to at least <preferredFirstName LastName>.

Students are expected to join class in person. However, if you are unable to attend class, let Matt or Videh know and we'll provide you a copy of the class recording (if available).

LEARNING RESOURCES

Required Textbook: N/A - everything is freely available online

Course Fees: N/A

Optional Online Learning Resources:
Additional learning resources aimed at facilitating student learning, and perhaps including formative assessment tools, are available from the textbook publisher and may be accessed for a fee paid by the student to the third-party provider (e.g., the textbook company). Students choosing to access and use the online resources should note the following:

1. Registration in the system and any monetary transactions are of your own accord and not the responsibility of the University.
2. Students should be mindful of protecting their personal information and should be aware of how their personal information might be used and/or shared.
3. Students MUST NOT use their @ualberta email address or CCID to register into the system and instead should use a non-identifying email address or account.

Academic Success Centre:
The Academic Success Centre provides professional academic support to help students strengthen their academic skills and achieve their academic goals. Individual advising, appointments, and group workshops are available year round in the areas of Accessibility, Communication, Learning, and Writing Resources. Modest fees apply for some services.

GRADE EVALUATION

Assessment	Weight	Date
Class Participation (interacting during class, doing short reading responses, providing feedback on other's work, etc.)	20%	Term
HIPPO Gym Exercise	5%	9/22
Paper Presentation 1	10%	Term
Paper Presentation 2	10%	Term
Initial proposal	5%	10/6
Refined proposal & pilot study	10%	10/20
Draft writeup	10%	11/10
Final presentation	15%	12/5 or 12/7
Final report	15%	12/10

Note that your report and presentation will be evaluated on their quality, not whether your data proves/disproves your original hypothesis. Put differently, running a pilot study and finding out you should not continue putting time into your idea can still be a perfectly acceptable outcome.

Grades are unofficial until approved by the Department and/or Faculty offering the course.

POLICIES FOR LATE AND MISSED WORK

Missed Term Work:
Graded work will have its grade reduced by 10% every day late. For example, if the HIPPO Gym exercise is submitted 36 hours after the deadline, the earned mark will be multiplied by 0.8.

Missed Presentation:
If the final presentation is not done on time, or the final report is not submitted on time, the student will receive 0% for the work. A student who cannot complete the final presentation due to incapacitating illness, severe domestic affliction or other compelling reasons can apply for an excused absence. To apply for an excused absence, you must contact the instructor within two working days of missing the assessment or as soon as possible. If an excused absence is granted, then you will be given an extension. An excused absence is a privilege and not a right. There is no guarantee that an absence will be excused. Misrepresentation of facts to gain an excused absence is a serious breach of the Code of Student Behaviour. In all cases, instructors may request adequate documentation to substantiate the reason for the absence at their discretion.

Missed Assessments Where the Cause is Religious Belief:
For an excused absence where the cause is religious belief, a student must contact the instructor(s) within two weeks of the start of Fall or Winter classes to request accommodation for the term. Instructors may request adequate documentation to substantiate the student request.

A student who cannot complete their final project due to incapacitating illness, severe domestic affliction or other compelling reasons can apply for an extension.

STUDENT RESPONSIBILITIES

Academic Integrity:
"The University of Alberta is committed to the highest standards of academic integrity and honesty. Students are expected to be familiar with these standards regarding academic honesty and to uphold the policies of the University in this respect. Students are particularly urged to familiarize themselves with the provisions of the Code of Student Behaviour (online at www.governance.ualberta.ca) and avoid any behaviour which could potentially result in suspicions of cheating, plagiarism, misrepresentation of facts and/or participation in an offence. Academic dishonesty is a serious offence and can result in suspension or expulsion from the University."

All forms of dishonesty are unacceptable at the University. Any offence will be reported to the Associate Dean of Science who will determine the disciplinary action to be taken. Cheating, plagiarism and misrepresentation of facts are serious offences. Anyone who engages in these practices will receive at minimum a grade of zero for the exam or paper in question and no opportunity will be given to replace the grade or redistribute the weights. As well, in the Faculty of Science, the sanction for **cheating** on any examination will include **a disciplinary failing grade** (NO EXCEPTIONS) and senior students should expect a period of suspension or expulsion from the University of Alberta.

Students are expected to familiarize themselves with the Academic Integrity resources (covering the topics of cheating, collaboration, plagiarism, and substantial assistance) on the website of the Office of the Dean of Students. Here's a one-page summary of things you shouldn't do.

Appropriate Collaboration: Many activities and work in this class will involve collaboration. Please be sure to note all collaborators and the extent of the collaboration. Students are expected to maintain collegial and respectful interactions online, whether by video, audio, or text.

Students Eligible for Accessibility-Related Accommodations:
Eligible students have both rights and responsibilities with regard to accessibility-related accommodations. Consequently, scheduling exam accommodations in accordance with Accessibility Resources deadlines and procedures is essential. Please note adherence to procedures and deadlines is required for U of A to provide accommodations. Contact Accessibility Resources for further information.

Policy about course outlines can be found in Course Requirements, Evaluations Procedures and Grading of the University Calendar.

Disclaimer:
Any typographical errors in this Course Outline are subject to change and will be announced in class.