University of Alberta

CMPUT 628: Deep Reinforcement Learning
LEC B1
Winter 2026

Instructor: Marlos C. Machado
Teaching Assistant: Hon Tik (Rick) Tse

Office: UCOMM 7-241
E-mail: machado@ualberta.ca
Web Page: https://canvas.ualberta.ca/courses/30281

Office hours: Slack and eClass: asynchronously. In-person meetings might be avail. upon request.

Lecture room & time: GSB 7-11, Tuesday and Thursday 11:00 - 12:20
Attendance isn't mandatory, although it is strongly encouraged.

Slack invitation link: We will use Slack as an optional alternative to eClass for communication and question-answering. The invitation link will be provided to the students on eClass.

COURSE CONTENT

Course Description: This course provides an introduction to the fundamentals of deep reinforcement learning, a subfield of reinforcement learning that has been distinguished by the types of problems and solution methods it has focused on. The solution methods in deep reinforcement learning involve using neural networks to parameterize artifacts such as value functions, policies, or models. In this course, we will take the view that deep reinforcement learning studies the design of algorithms to explicitly allow for modern deep learning techniques. This is in contrast to being agnostic to the function class used for function approximation and simply seeing reinforcement learning as a consumer of supervised learning techniques.

The course provides an introduction to the general concepts underlying most of the traditional deep reinforcement learning algorithms that have been used to learn to achieve goals in problems with high-dimensional observation space. The course will cover the design choices that underlie model-free deep reinforcement learning algorithms and the different algorithms that arise from particular choices. Time permitting, it will then discuss some high-profile deep model-based

reinforcement learning algorithms and key policy gradient methods, with the major bulk of the course being model-free methods. The course will take a concepts-based approach instead of focusing on particular algorithms, and it will also touch on empirical design and current reporting practices in the field. A final component of the course will be a set of seminars presented by the students on more recent results in the field.

Course Prerequisites: This course does not have specific courses as prerequisites. Nevertheless, students are expected to have been exposed to the basic ideas of reinforcement learning (as one would acquire by having taken, for example, CMPUT 365 or CMPUT 655) and machine learning in general (as one would acquire by having taken, for example, CMPUT 466/566).

The course will also have a programming component, which should be done in Python.

We will use elementary ideas of probability, calculus, and linear algebra, such as expectations of random variables, conditional expectations, partial derivatives, vectors and matrices. Students should either be familiar with these topics or be ready to pick them up quickly as needed by consulting outside resources.

Course Objectives and Expected Learning Outcomes: By the end of the course, students should have a solid grasp of the main concepts underlying deep reinforcement learning algorithms, which is a leading approach to statistical decision-making. Students will also develop a much better understanding of how to interpret empirical results and how to understand them critically. At the end of this course, students should become better able to assess claims made by others concerning new algorithms, software products, and also be able to appreciate some new research results. Ideally, students will also develop a general understanding of how to implement some of the most traditional deep reinforcement learning algorithms discussed in class.

With a heavy focus on AI as the design of *single-stream* neural-network-powered agents learning *solely* from experience in order to predict and control complex environments, the topics covered in the course will include:

- Model-free deep reinforcement learning concepts (e.g., experience replay buffers, target networks, auxiliary inputs, auxiliary tasks, loss functions) and the algorithms created as different instantiations of these concepts (e.g., DQN, Double DQN, Rainbow, Retrace).
- High-profile model-based deep reinforcement learning methods (e.g., MuZero and Dreamer).
- Key policy gradient methods that are often used with neural networks as function approximators (e.g., PPO, SAC, DDPG, and TD3).

LEARNING RESOURCES

Required Textbook and/or Other Major Course Materials: This course will have no textbook. Lecture notes will be provided for some of the content covered in the course. The course will also have recommended readings of scientific literature published in the last decade.

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.

Faculty of Science Student Services: The Faculty of Science Student Services office is located on the main floor of the Centennial Centre for Interdisciplinary Sciences (CCIS). This office can assist with the planning of Your Academics and provide information related to Student Life & Engagement, Internship & Careers, and Study Abroad opportunities. Please visit Advising for more information about what Faculty Academic Advisors in the Student Services Office can assist you with.

GRADE EVALUATION

Assessment	Weight	Date
Assignment 1	10%	January 24, 2026
Assignment 2	10%	February 7, 2026
Assignment 3	15%	February 28, 2026
Assignment 4	15%	March 14, 2026
Midterm exam	20 %	March 19, 2026
Paper review	15%	April 9, 2026
Seminar	15%	March 24, 2026 – Apr 9, 2026

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

Statement of Expectations for AI Use: The primary goal of this course is to foster *individual* critical, creative thinking, and problem-solving skills related to deep reinforcement learning and, more broadly, machine learning. Therefore, the use of advanced AI tools based on large-language models such as ChatGPT or Gemini is discouraged as they might prevent students from engaging more critically with the content. As stated in the university's AI-Squared - Artificial Intelligence and Academic Integrity webpage, "learning is not only about the product; learning is also about the process of acquiring new knowledge or learning ways to think and reason."

Re-evaluation of midterm exams: Students will have access to their midterm exam during an exam viewing period. A student who has concerns about how specific questions of their midterm exam were marked can submit a request to the instructor via email within two weeks of the date they received their marked exam. The request should specify (1) which question is to be re-evaluated, (2) the rationale for such a request, and (3) the proposed marks. Importantly, once a request for re-evaluation is submitted, it is up to the instructor's discretion to adjust the marks. *Students won't be allowed to take their midterm exams with them, nor to take pictures of them*, so in case of concerns, the student is advised to take notes during the exam viewing period. The TAs are not authorized to weigh in on the midterm exams, this is something only the instructor can do. *Notice marks can also go down once a question is re-evaluated.*

Past or Representative Evaluative Material: Representative material will be discussed in the classroom. This is a new course, and there will be relatively fewer examples of representative evaluative material.

Format of Exams: The midterm exam will be held in GSB 7-11 from 11:00 - 12:20 (80 minutes). *The exam will be closed book and written in real-time.*

Format of Assignments: You are allowed to consult any other material that is available on the internet when solving the assignments. You are also allowed to discuss the assignments with your classmates. Note, however, that you are not allowed to exchange any written text or code or to give and/or receive detailed step-by-step instructions on how to solve the proposed problems.

Grade Evaluation: At the end of the term your percentage grade will be converted into a letter grade by following the table below, extracted from the Department Course Policies.

Letter	Descriptor	Interpretation
A+ A A-	Excellent	Consistently original thinking that extends the material, demonstrated depth and breadth in the material, ability to integrate material with other subjects, ability to analyze and synthesize material at various levels of abstraction.
B+ B	Good	Like an A, but not consistent over time, or weak in a specific area.
B- C+	Satisfactory	Understand the core material but not its subtleties, can apply it to simple situations on own and to more complex situations with hints, evidence that the material has changed the way of thinking.
C C- D+ D F	Failure	Little evidence of understanding of even the surface issues, poor analysis and synthesis, inability to apply the material.

The conversion from numerical grades to letter grades will be determined at the end of the term based on overall class performance. However, the table below provides guaranteed minimum letter grades for the corresponding numerical scores.

Points	97	90	85	80	75	70	60
Grade	A+	A	A-	B+	B	B-	C+

Grade adjustment: *Grades will not be rounded at the end, and no extra marks will be given, with no exceptions.* As much as the instructor understands grades can impact one's eligibility for scholarships and other aspects of a graduate student's career, this will not be considered when assigning grades to students.

Exam Conduct:

- Your student photo I.D. is required at exams to verify your identity.
- Students will not be allowed to begin an examination after 30 minutes have passed. Students must remain in the exam room until at least 30 minutes have elapsed.
- All cell phones must be turned off and stored in your bags.

POLICIES FOR LATE AND MISSED WORK

Late Policies: The late submission penalty is based on the following logistic function, which determines the percentage deducted from your final grade:

$$1 - \frac{1}{1+e^{1.5(t-5)}}$$

where *t* is the number of days late. Note that the penalty is almost negligible for the first 2 to 3 days, with less than a 5% deduction. However, after 5 days, the penalty increases significantly, and by 7 days, the deduction is over 95%. It's important to submit as soon as possible if you're running late, as the early days have minimal impact on your grade. Such a policy was designed to minimize requests for exemptions due to minor issues; as such, minor issues will lead to a delay of a few days with negligible impact on one's grades.

Missed Term Work or Final Exam Due to Non-medical Protected Grounds (e.g., religious beliefs): When a term assessment or final exam presents a conflict based on non-medical protected grounds, students must apply to the Academic Success Centre for accommodations via their Register for Accommodations website. Students can review their eligibility and choose the application process specific for *Accommodations Based on Non-medical Protected Grounds*.

It is imperative that students review the dates of all course assessments upon receipt of the course syllabus and apply *AS SOON AS POSSIBLE* to ensure the timely application of the accommodation. Students who apply later in the term may experience unavoidable delays in the processing of the application, which can affect the accommodation.

Missed midterm exam: A student who cannot complete a midterm exam 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 the student will be able to take another exam at a date to be determined by the instructor. 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 Student Academic Integrity Policy.

REMOTE DELIVERY CONSIDERATIONS

Technology for Remote Learning: To successfully participate in remote learning in this course, it is recommended that students have access to a computer with an internet connection that can support the tools and technologies the University uses to deliver content, engage with instructors, TAs, and fellow students, and facilitate assessment and examinations. Please refer to Technology for Remote Learning - For Students for details. If you encounter difficulty meeting the technology recommendations, please email the Dean of Students Office (dosdean@ualberta.ca) directly to explore options and support.

Please contact the instructor by the add/drop deadline of January 17 if you do not have access to the minimum technology recommended. The instructor will make arrangements for accommodating students who contact the instructor before this date. Failure to do so may result in a zero in any assessment that depends on the minimum technology.

Student Resources for Remote Learning: Online learning may be new to you. Check out tips for success and find out more about online learning on the Campus Life page, specifically on the Academic Skills Online & Remote Delivery Resources page.

STUDENT RESPONSIBILITIES

Academic Integrity and Student Conduct: The University of Alberta is committed to the highest standards of academic integrity and honesty, as well as maintaining a learning environment that fosters the safety, security, and inherent dignity of each member of the community, ensuring students conduct themselves accordingly. Students are expected to be familiar with the standards of academic honesty and appropriate student conduct, and to uphold the policies of the University in this respect.

Students are particularly urged to familiarize themselves with the provisions of the Student Academic Integrity Policy and the Student Conduct Policy, and avoid any behaviour that could potentially result in suspicions of academic misconduct (e.g., cheating, plagiarism, misrepresentation of facts, participation in an offence) and non-academic misconduct (e.g., discrimination, harassment, physical assault). Academic and non-academic misconduct are taken very seriously and can result in suspension or expulsion from the University.

All students are expected to consult the Student Academic Integrity Policy for clarification on the various academic offences. All forms of academic dishonesty are unacceptable at the University. Unfamiliarity of the rules, procrastination or personal pressures are not acceptable excuses for

committing an offence. Listen to your instructor, be a good person, ask for help when you need it, and do your own work -- this will lead you toward a path to success. Any academic integrity concern in this course will be reported to the College of Natural and Applied Sciences.

Suspected cases of non-academic misconduct will be reported to the Office of Student Success and Experience. The College, the Faculty, and the Dean of Students are committed to student rights and responsibilities, and adhere to due process and administrative fairness, as outlined in the Student Academic Integrity Policy and the Student Conduct Policy. Please refer to the policy websites for details on inappropriate behaviours and possible sanctions.

The College of Natural and Applied Sciences (CNAS) has created an Academic Integrity for CNAS Students eClass site. Students can self enroll and review the various resources provided, including the importance of academic integrity, examples of academic misconduct and possible sanctions, and the academic misconduct and appeal process. They can also complete assessments to test their knowledge and earn a completion certificate.

"*Integrity is doing the right thing, even when no one is watching*" -- C.S. Lewis

Contract Cheating and Misuse of University Academic Materials or Other Assets: Contract cheating describes the form of academic dishonesty where students get academic work completed on their behalf, which they then submit for academic credit as if they had created it themselves.

Contract cheating may or may not involve the payment of a fee to a third party, who then creates the work for the student.

Examples include:
1) Getting someone to write an essay or research paper for you.
2) Getting someone to complete your assignment or exam for you.
3) Posting an essay, assignment or exam question to a tutorial or study website; the question is answered by a “content expert”, then you copy it and submit it as your own answer.
4) Posting your solutions to a tutorial/study website, public server or group chat and/or copying solutions that were posted to a tutorial/study website public server or group chat.
5) Sharing your login credentials to the course management system (e.g. Canvas) and allowing someone else to complete your assignment or exam remotely.
6) Using an artificial intelligence bot or text generator tool to complete your essay, research paper, assignment or exam solutions for you.

Contract cheating companies thrive on making students believe that they cannot succeed without their help; they attempt to convince students that cheating is the only way to succeed.

Uploading the instructor's teaching materials (e.g., course outlines, lecture slides, assignment, or exam questions, etc.) to tutorial, study, or note-sharing websites or public servers is a copyright

infringement and constitutes the misuse of University academic materials or other assets. Receiving assignment solutions or answers to exam questions from an unauthorized source puts you at risk of receiving inaccurate information.

Appropriate Collaboration: Students need to be able to recognize when they have crossed the line between appropriate collaboration and inappropriate collaboration. If students are unsure, they need to ask instructors to clarify what is allowed and what is not allowed.

Here are some tips to avoid copying on assessments:

1. Do not write down something that you cannot explain to your instructor.
2. When you are helping other students, avoid showing them your work directly. Instead, explain your solution verbally. Allowing your work to be copied is also considered inappropriate collaboration.
3. It is also possible that verbally discussing the solution in too much detail may result in written responses that are too similar. Try to keep discussions at a general or higher level.
4. If you find yourself reading another student's solution, do not write anything down. Once you understand how to solve the problem, remove the other person's work from your sight and then write up the solution to the question yourself. Looking back and forth between someone else's paper and your own paper is almost certainly copying and considered inappropriate collaboration.
5. If the instructor or TA writes down part of a solution in order to help explain it to you or the class, you cannot copy it and hand it in for credit. Treat it the same way you would treat another student's work with respect to copying, that is, remove the explanation from your sight and then write up the solution yourself.
6. There is often more than one way to solve a problem. Choose the method that makes the most sense to you rather than the method that other students happen to use. If none of the ideas in your solution are your own, there is a good chance it will be flagged as copying.

Cell Phones: Cell phones are to be turned off during lectures, labs and seminars.

Accommodations for Students:
In accordance with the University of Alberta's Discrimination, Harassment, and Duty to Accommodate policy, accommodation support is available to eligible students who encounter limitations or restrictions to their ability to perform the daily activities necessary to pursue studies at a post-secondary level due to medical conditions and/or non-medical protected grounds. Accommodations are coordinated through the Academic Success Centre, and students can learn more about eligibility on the Register for Accommodations website.

It is recommended that students apply *AS SOON AS POSSIBLE* in order to ensure sufficient time to complete accommodation registration and coordination. Students are advised to review and adhere to published deadlines for accommodation approval and for specific accommodation requests (e.g., exam registration submission deadlines). Students who request accommodations

less than a month in advance of the academic term for which they require accommodations may experience unavoidable delays or consequences in their academic programs, and may need to consider alternative academic schedules.

Recording and/or Distribution of Course Materials: Audio or video recording, digital or otherwise, of lectures, labs, seminars or any other teaching environment by students is allowed only with the prior written consent of the instructor or as a part of an approved accommodation plan. Student or instructor content, digital or otherwise, created and/or used within the context of the course is to be used solely for personal study, and is not to be used or distributed for any other purpose without prior written consent from the content author(s).

STUDENT SUPPORTS

Faculty of Science Student Services:
The Faculty of Science Student Services office is located on the main floor of the Centennial Centre for Interdisciplinary Sciences (CCIS). This office can assist with the planning of Your Academics, and provide information related to Student Life & Engagement, Internship and Careers, and Study Abroad opportunities. Please visit Advising for more information about what Faculty Academic Advisors can assist you with.

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 may apply for some services.

Feeling Stressed, Anxious, or Upset?
It's normal for us to have different mental health experiences throughout the year. Know that there are people who want to help. You can reach out to your friends and access a variety of supports available on and off campus at the Need Help Now webpage or by calling the 24-hour Distress Line: 780-482-4357 (HELP).

Learning and Working Environment:
The Faculty of Science is committed to ensuring that all students, faculty and staff are able to work and study in an environment that is safe and free from discrimination, harassment, and violence of any kind. It does not tolerate behaviour that undermines that environment. This includes virtual environments and platforms.

If you are experiencing harassment, discrimination, fraud, theft or any other issue and would like to get confidential advice, please contact any of these campus services:

- Office of Safe Disclosure & Human Rights*: A safe, neutral and confidential space to disclose concerns about how the University of Alberta policies, procedures or ethical standards are being applied. They provide strategic advice and referral on matters such as discrimination, harassment, duty to accommodate and wrong-doings. Disclosures can be made in person or online using the Online Reporting Tool*.
- University of Alberta Protective Services: *Peace officers dedicated to ensuring the safety and security of U of A campuses and community. Staff or students can contact UAPS to make a report if they feel unsafe, threatened, or targeted on campus or by another member of the university community.*
- Office of the Student Ombuds*: A confidential and free service that strives to ensure that university processes related to students operate as fairly as possible. They offer information, advice, and support to students, faculty, and staff as they deal with academic, discipline, interpersonal, and financial issues related to student programs.*
- Office of Student Success and Experience*: They can assist students in navigating services to ensure they receive appropriate and timely resources. For students who are unsure of the support they may need, are concerned about how to access services on campus, or feel like they may need interim support while they wait to access a service, this office is there to help.*

Course Outlines:
Policy about course outlines can be found in the Academic Regulations, Evaluation Procedures and Grading section of the University Calendar.

Disclaimer:
Any typographical errors in this syllabus are subject to change and will be announced in class and/or posted on the course website. The date of final examinations is set by the Registrar and takes precedence over the final examination date reported in the syllabus.

Course Schedule (Tentative) & Assigned Readings

Week	Date	Topic	Deadlines (all due at 23:59:59)	Readings (tentative)
1	Tue, Jan 6	Course Overview Intro to Deep RL		Chapter 1
1	Thu, Jan 8	Background review: Neural networks		Chapter 3
2	Tue, Jan 13	Background review: Neural networks		Chapter 3
2	Thu, Jan 15	Background review: Reinforcement learning		Chapter 2
3	Tue, Jan 20	Value-based Model-free Methods A Generic Framework and DQN		Chapter 4, Sections 4.1 – 4.3
3	Thu, Jan 22	Value-based Model-free Methods The many details and design choices in DQN		Chapter 4, Sections 4.1 – 4.3
Saturday, Jan 24		**Assignment 1: A Reinforcement Learning Learning Setup**		
4	Tue, Jan 27	Value-based Model-free Methods Objective Functions: Double Learning		Chapter 4, Section 4.4.1
4	Thu, Jan 29	Value-based Model-free Methods Objective Functions: Multi-step Methods		Chapter 4, Section 4.4.2
5	Tue, Feb 3	Value-based Model-free Methods Distributional Reinforcement Learning		Chapter 4, Sections 4.4.3 – 4.4.4
5	Thu, Feb 5	Value-based Model-free Methods Auxiliary Objectives		Chapter 4, Section 4.5
Saturday, Feb 7		**Assignment 2: Designing Features for Linear Function Approximation**		
6	Tue, Feb 10	Value-based Model-free Methods Auxiliary Inputs & Experience Replay Buffers		Chapter 4, Section 4.6 – 4.7
6	Thu, Feb 12	Value-based Model-free Methods Neural Network Architectures		Chapter 4, Section 4.8

Feb 16 - Nov 20		Reading week		
7	Tue, Feb 24	Policy Gradient Methods DDPG & TD3		Chapter 6
7	Thu, Feb 26	Policy Gradient Methods DDPG & TD3		Chapter 6
Saturday, Feb 28		**Assignment 3: DQN**		
8	Tue, Mar 3	Policy Gradient Methods TRPO, PPO, & SAC		Chapter 6
8	Thu, Mar 5	Policy Gradient Methods TRPO, PPO, & SAC		Chapter 6
9	Tue, Mar 10	Model-based Methods AlphaGo, AlphaZero, MuZero		Chapter 5
9	Thu, Mar 12	Model-based Methods Dreamer architectures		Chapter 5
Friday, Mar 14		**Assignment 4: Policy Gradient Methods**		
10	Tue, Mar 17	Discussion Day / Buffer		
Thursday, Mar 19		**Midterm Exam**		
11	Tue, Mar 24	Seminar: Day 1 (2 groups, 30 + 10 min each)		
11	Thu, Mar 26	Seminar: Day 2 (2 groups, 30 + 10 min each)		
12	Tue, Mar 31	Seminar: Day 3 (2 groups, 30 + 10 min each)		
12	Thu, Apr 2	Seminar: Day 4 (2 groups, 30 + 10 min each)		
13	Tue, Apr 7	Seminar: Day 5 (2 groups, 30 + 10 min each)		
13	Thu, Apr 9	Seminar: Day 6 (2 groups, 30 + 10 min each)		
Thursday, Apr 9		**Paper Review**		