Western Science

Department of Statistical and Actuarial Sciences
Reinforcement Learning || 9670/9671/9170

Course outline for Winter 2022

THRIVE together

Although this academic year might be different, Western University is committed to a **thriving campus**. We encourage you to check out the Digital Student Experience website to manage your academics and well-being. Additionally, the following link provides available resources to support students on and off campus: https://www.uwo.ca/health/.

Technical Requirements and Important Dates:

 Stable internet connection

 Laptop or computer

 Working microphone

 Working webcam

Classes Start	Reading Week	Classes End	Study day(s)	Exam Period
January 10	February 19 - 27	April 8	April 9	April 10 - 30

1. Course Information

Course information

- Reinforcement Learning (CS 9670/9671/9170), Winter 2022
- **Asynchronous online course.** Videos will be of lectures and programming demonstrations.
- **In-Person Lecture** sometimes. The pandemic situation sometimes changes swiftly, and maybe asynchronous works better.

List of Prerequisite(s)

- **Concepts:** Probability, linear algebra, expectation, partial derivatives (e.g. for gradient descent), search.
- **Programming:** If you do not know Python, you must be aware that you will spend extra time learning this (very beginner-friendly) language as all programming assignments are in Python.

Unless you have either the requisites for this course or written special permission from your Dean to enroll in it, you may be removed from this course and it will be deleted from your record. This

decision may not be appealed. You will receive no adjustment to your fees in the event that you are dropped from a course for failing to have the necessary prerequisites.

2. Instructor Information

Course Instructor
Scott Leith

Contact Information
sleith6@uwo.ca

Office Hours
On request (email me for a zoom meeting)

Students must use their Western (@uwo.ca) email addresses when contacting their instructors.

3. Course Syllabus, Schedule, and Delivery Mode

This course will provide a broad introduction to the foundational concepts and algorithms of reinforcement learning, one of the largest and most active areas in machine learning. The main focus will be on fundamental algorithms and their applications, and will end with an introduction to deep reinforcement learning. Knowledge of probability theory, logic, expectation, and basic machine learning principles (e.g., gradient descent) will be very helpful. The course will be taught in Python, and where appropriate certain Python features and basic elements of software design will be taught.

Type	Mode	Dates	Time	Frequency
Lecture	Asynchronous online NCB 284	Async throughout, in-person sometimes (depends on situation)	1 hour	Tues, Thurs 8:30am - 9:30am

☒ Asynchronous pre-work must be completed [2 days] prior to synchronous sessions
☒ Missed work should be completed within 48 hours
☒ Closed captioning will be provided on audio or video recordings

Learning Outcomes

Upon successful completion of this course, students will be able to:

- Define the core features of reinforcement learning, and explain how RL differs from other artificial intelligence / machine learning approaches.
- Determine if a given problem should be approached as a reinforcement learning problem. Compare different algorithms to select the most appropriate for a particular application/problem space.
- Implement (in code) various common/classic reinforcement algorithms from scratch in Python.
- List and define the various criteria for evaluating reinforcement learning algorithms (e.g., regret, sample efficiency, sample complexity, …).
- List and define the major weaknesses and complications that come with a reinforcement learning approach (e.g., sample *in*efficiency).

Table of Contents and Schedule

Week	Dates	Topic(s)	Reading	Notes
1	Jan 11, 13	Problem definition and key properties, the bandit problem.	Chapter 1, Chapter 2	
2	Jan 18, 20	Markov Decision Processes	3.1 - 3.5	
3	Jan 25, 27	Value Functions and Bellman Equations, Planning Via Dynamic Programming	3.6 - 3.7 Chapter 4	
4	Feb 1, 3	Sample-Based Learning, Temporal Difference Learning	5.1 - 5.4 6.1 - 6.3	
5	Feb 8, 10	Temporal Difference Learning, SARSA, Q-Learning	6.4 - 6.7	
6	Feb 15, 17	MIDTERM		The 15th will be an optional study period where I answer questions
6	Feb 19-27	READING WEEK		Reading Week
7	Mar 1, 3	Planning, Learning, and Acting	8.1 - 8.5	
8	Mar 8, 10	Value Function Approximation	9.1, 9.2, 9.4	No Monday class (Family Day)
9	Mar 15, 17	Feature Construction, Neural Networks	9.5.0, 9.5.3, 9.5.4, 9.6, 9.7	
10	Mar 22, 24	Control with Function Approximation	10.1, 10.3	
11	Mar 29, 31	Policy Gradient Methods	Chapter 13	
12	Mar 31, Apr 2	Experience Replay, Monte Carlo Tree Search	8.11 Chapter 14 Papers TBA	May change to support student final projects
13	Apr 5, 7	Deep RL		Topics guided by student projects / preferences

Online Participation and Engagement

☒ Students are expected to participate and engage with content as much as possible

☒ Students can post on OWL after watching the recording, or alternatively send an email with their thoughts.

☒ Students can also participate by interacting in the forums with their peers and instructors

4. Course Materials

Sutton, R. S., & Barto, A. G. (2011). Reinforcement learning: An introduction.

This is *the* text on reinforcement learning, written by Richard Sutton and Andrew Barto from the University of Alberta and made available for free online. It is roughly split into two parts: Part 1 covers the fundamentals of reinforcement learning, namely core concepts, multi-armed bandits, dynamic programming, temporal difference (TD) learning, Q-learning, Monte-Carlo techniques, and the classic SARSA and Dyna algorithms. Part 2 delves into function approximation, eligibility traces, and policy gradient methods, while Part 3 covers the conceptual cross-field foundations of reinforcement learning with chapters on psychology, neuroscience, and classic case studies.

Additional readings will be provided through OWL as they arise. A selection of possible readings are listed below:

Smith, L., & Gasser, M. (2005). The development of embodied cognition: Six lessons from babies. *Artificial life*, *11*(1-2), 13-29.

Agrawal, S., & Goyal, N. (2013, February). Thompson sampling for contextual bandits with linear payoffs. In *International Conference on Machine Learning* (pp. 127-135).

Liu, F., Tang, R., Li, X., Zhang, W., Ye, Y., Chen, H., ... & Zhang, Y. (2018). Deep reinforcement learning based recommendation with explicit user-item interactions modeling. *arXiv preprint arXiv:1810.12027*.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., ... & Chen, Y. (2017). Mastering the game of go without human knowledge. *Nature*, *550*(7676), 354.

Browne, C. B., Powley, E., Whitehouse, D., Lucas, S. M., Cowling, P. I., Rohlfshagen, P., ... & Colton, S. (2012). A survey of monte carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games*, *4*(1), 1-43.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

Ye, D., Chen, G., Zhang, W., Chen, S., Yuan, B., Liu, B., ... & Yin, Y. (2020). Towards Playing Full MOBA Games with Deep Reinforcement Learning. *Advances in Neural Information Processing Systems*, *33*.

Dulac-Arnold, G., Mankowitz, D., & Hester, T. (2019). Challenges of real-world reinforcement learning. *arXiv preprint arXiv:1904.12901*.

Students should check OWL (http://owl.uwo.ca) on a regular basis for news and updates. This is the primary method by which information will be disseminated to all students in the class. Students are responsible for checking OWL on a regular basis.

All course material will be posted to OWL: http://owl.uwo.ca. Any changes will be indicated on the OWL site and discussed with the class.

If students need assistance, they can seek support on the OWL Help page. Alternatively, they can contact the Western Technology Services Helpdesk. They can be contacted by phone at 519-661-3800 or ext. 83800.

Google Chrome or Mozilla Firefox are the preferred browsers to optimally use OWL; update your browsers frequently. Students interested in evaluating their internet speed, please click here.

5. Methods of Evaluation

The overall course grade will be calculated as listed below:

Programming Assignments	55%
Written Assignments	25%
Midterm Exam	20%

NOTE: NO FINAL EXAM

Accommodated Evaluations

- Late assessments without illness self-reports will be subject to a late penalty 5%/day
- Late assessments with illness self-reports should be submitted within 24 hours of submission of the last illness self-report
- An assessment cannot be submitted after it has been returned to the class; an alternate assessment will be assigned.
- A make-up test will be offered
- If a make-up assessment is missed, the student will receive an INC and complete the task the next time the course is offered
- If permission to waive the requirement that students receive evaluation on work totaling 15% of their final grade at least three days prior to the deadline for withdrawal without academic penalty has been obtained from the Dean's Office, a statement to this effect must be made.

Rounding of Marks Statement

Across the Sciences Undergraduate Education programs, we strive to maintain high standards that reflect the effort that both students and faculty put into the teaching and learning experience during this course. All students will be treated equally and evaluated based only on their actual achievement. ***Final grades*** on this course, irrespective of the number of decimal places used in marking individual assignments and tests, will be calculated to one decimal place and rounded to the nearest integer, e.g., 74.4 becomes 74, and 74.5 becomes 75. Marks WILL NOT be bumped to the next grade or GPA, e.g. a 79 will NOT be bumped up to an 80, an 84 WILL NOT be bumped up to an 85, etc. The mark attained is the mark you achieved, and the mark assigned; requests for mark "bumping" will be denied.

6. Accommodation and Accessibility

Accommodation Policies

Students with disabilities work with Accessible Education (formerly SSD) which provides recommendations for accommodation based on medical documentation or psychological and cognitive testing. The Academic Accommodation for Students with Disabilities policy can be found at:
https://www.uwo.ca/univsec/pdf/academic_policies/appeals/Academic Accommodation_disabilities.pdf

Academic Consideration for Student Absence

Students will have up to two (2) opportunities during the regular academic year to use an on-line portal to self-report an absence during the semester, provided the following conditions are met: the absence is no more than 48 hours in duration, and the assessment for which consideration is being sought is worth 30% or less of the student's final grade. Students are expected to contact their instructors within 24 hours of the end of the period of the self-reported absence, unless noted on the syllabus. Students are not able to use the self-reporting option in the following circumstances:

- for exams scheduled by the Office of the Registrar (e.g., December and April exams)
- absence of a duration greater than 48 hours,
- assessments worth more than 30% of the student's final grade,
- if a student has already used the self-reporting portal twice during the academic year

If the conditions for a Self-Reported Absence are *not* met, students will need to provide a Student Medical Certificate if the absence is medical, or provide appropriate documentation if there are compassionate grounds for the absence in question. Students are encouraged to contact their Faculty academic counselling office to obtain more information about the relevant documentation.

Students should also note that individual instructors are not permitted to receive documentation directly from a student, whether in support of an application for consideration on medical grounds, or for other reasons. **All documentation required for absences that are not covered by the Self-Reported Absence Policy must be submitted to the Academic Counselling office of a student's Home Faculty.**

For policy on Academic Consideration for Student Absences - Undergraduate Students in First Entry Programs, see:
https://www.uwo.ca/univsec/pdf/academic_policies/appeals/Academic_Consideration_for_absences.pdf
and for the Student Medical Certificate (SMC), see:
http://www.uwo.ca/univsec/pdf/academic_policies/appeals/medicalform.pdf

Religious Accommodation

Students should consult the University's list of recognized religious holidays, and should give reasonable notice in writing, prior to the holiday, to the Instructor and an Academic Counsellor if their course requirements will be affected by a religious observance. Additional information is given in the Western Multicultural Calendar:
https://multiculturalcalendar.com/ecal/index.php?s=c-univwo

You may also be eligible to write the Special Exam if you are in a "Multiple Exam Situation" (see http://www.registrar.uwo.ca/examinations/exam_schedule.html).

If a student fails to write a scheduled Special Examination, the date of the next Special Examination (if granted) normally will be the scheduled date for the final exam the next time

this course is offered. The maximum course load for that term will be reduced by the credit of the course(s) for which the final examination has been deferred. See Academic Calendar for details (under Special Examinations).

7. Academic Policies

The website for Registrarial Services is http://www.registrar.uwo.ca.

In accordance with policy, http://www.uwo.ca/its/identity/activatenonstudent.html, the centrally administered e-mail account provided to students will be considered the individual's official university e-mail address. It is the responsibility of the account holder to ensure that e-mail received from the University at his/her official university address is attended to in a timely manner.

All of the remote learning sessions for this course will be recorded.
The data captured during these recordings may include your image, voice recordings, chat logs and personal identifiers (name displayed on the screen). The recordings will be used for educational purposes related to this course, including evaluations. The recordings may be disclosed to other individuals participating in the course for their private or group study purposes. Please contact the instructor if you have any concerns related to session recordings.

Participants in this course are not permitted to record the sessions, except where recording is an approved accommodation, or the participant has the prior written permission of the instructor.

Scholastic offences are taken seriously and students are directed to read the appropriate policy, specifically, the definition of what constitutes a Scholastic Offence, at the following Web site:
http://www.uwo.ca/univsec/pdf/academic_policies/appeals/scholastic_discipline_undergrad.pdf.

All required papers may be subject to submission for textual similarity review to the commercial plagiarism detection software under license to the University for the detection of plagiarism. All papers submitted for such checking will be included as source documents in the reference database for the purpose of detecting plagiarism of papers subsequently submitted to the system. Use of the service is subject to the licensing agreement, currently between The University of Western Ontario and Turnitin.com (http://www.turnitin.com).

Computer-marked multiple-choice tests and exams may be subject to submission for similarity review by software that will check for unusual coincidences in answer patterns that may indicate cheating.

Completion of this course will require you to have a reliable internet connection and a device that meets the technical requirements for this service. Information about the technical requirements are available at the following link:
https://www.proctortrack.com/tech-requirements/

Tests and examinations in this course will be conducted using Zoom. You will be required to keep your camera on for the entire session, hold up your student card for identification purposes, and share your screen with the invigilator if asked to do so at any time during the exam. The exam session will **not** be recorded.*

More information about the use of Zoom for exam invigilation is available in the Online Proctoring Guidelines at the following link:
https://www.uwo.ca/univsec/pdf/onlineproctorguidelines.pdf

* Please note that Zoom servers are located outside Canada. If you would prefer to use only your first name or a nickname to login to Zoom, please provide this information to the instructor in advance of the test or examination.

Professionalism & Privacy

Western students are expected to follow the Student Code of Conduct. Additionally, the following expectations and professional conduct apply to this course:

☒ Students are expected to follow online etiquette expectations provided on OWL

☒ All course materials created by the instructor(s) are copyrighted and cannot be sold/shared

☒ Recordings are not permitted (audio or video) without explicit permission

☒ Permitted recordings are not to be distributed

☒ Students will be expected to take an academic integrity pledge before some assessments

☒ All recorded sessions will remain within the course site or unlisted if streamed

Copyright Statement

Please be aware that all course materials created by the instructor(s) are copyrighted and cannot be **sold/shared**. Those include materials used in tests/quizzes, midterms, and finals. Any posting/sharing of such materials in part or whole without owner's consent is considered as violation of the Copyright Act and will be considered as a scholastic offence.

In addition, online services such as Chegg are actively monitored. Any questions that are coming out during midterms and finals and are posted to an online service will be searched. Such an activity will be considered as a scholastic offence and will result in academic penalty.

8. Support Services

Please visit the Science & Basic Medical Sciences Academic Counselling webpage for information on add/drop courses, academic considerations for absences, appeals, exam conflicts, and many other academic related matters: https://www.uwo.ca/sci/counselling/

Please contact the course instructor if you require lecture or printed material in an alternate format or if any other arrangements can make this course more accessible to you. You may also wish to contact Student Accessibility Services (SAS) at (519) 661-2147 if you have any questions regarding accommodations.

Western University is committed to a thriving campus as we deliver our courses in the mixed model of both virtual and face-to-face formats. We encourage you to check out the Digital Student Experience website to manage your academics and well-being: https://www.uwo.ca/se/digital/

Students who are in emotional/mental distress should refer to Mental Health@Western (http://www.health.uwo.ca/mental_health) for a complete list of options about how to obtain help.

Additional student-run support services are offered by the USC, http://westernusc.ca/services.