IE 598 Special Topics

Reinforcement Learning and Learning-based Control

Instructor:	Yingying Li yl101@illinois.edu
Time & Place:	Tuesday/Thursday 2:00pm–3:20pm 137 Loomis Laboratory
Office Hours: **Location:**	Wednesday 4:30pm–5:30pm or by appointment. Coordinated Science Lab, 347.
Textbook:	Dimitri Bertsekas. Dynamic Programming and Optimal Control: Volume I and II
Reference books:	Dimitri Bertsekas and John N. Tsitsiklis. Neuro-Dynamic Programming. Andrew Barto and Richard S. Sutton. Reinforcement Learning: An Introduction
Grading:	**Assignments: 30%. Midterm exam: 30%. Final project: 40%.**
Homework:	We will have **3 assignments**. Assignments are due on Thursdays before 11:59pm. Tentative due dates: Sep 18, Oct 2, Nov 7. Assignments will be posted on Canvas. The first assignment will be posted on Sep 1. Later assignments will be posted on the due date of the previous assignment. Students are expected to submit ther answers in pdf via **GradeScope**. Assignments may be discussed with other students. However, students are expected to submit **independent write-ups** and **independent code** for computer assignments. Assignments may be checked for similarity. **Late submission:** every student has **one chance** of late submission if they couldn't submit their homework in time, as long as they inform the instructor **before the due date via email**. Late answers must be submitted within **4 days** of the due date, i.e., **on the next Monday, before 11:59pm**.
Midterm:	In-class midterm, Date TBD. Open book: students can bring any print-outs, notes, papers, etc. The use of any electronic device will **not** be allowed during examinations.
Modifications:	The instructor reserves the right to modify any or all parts of this syllabus throughout the semester. Any modifications will be announced to the students and posted on the course website.
Final project:	There will not be a final exam. Instead, there will be a final project, which involves a project proposal, at least one meeting with the instructor about the project, a presentation in the class, and a final report. More details are provided later.

Course Description

Reinforcement learning (RL) and learning-based control (LBC) are two closely connected fields, both of which focus on decision-making and control of uncertain dynamical systems, also known as Markov decision processes. This course will discuss the similarities, differences, and interconnections of RL and LBC from the aspects of algorithm design, theoretical tools, and real-world applications. In particular, the course will start with dynamic programming as a common tool for RL and LBC, then diving into different algorithm design philosophies of RL and LBC, e.g., model-free approaches such as Q learning, actor-critic, policy gradient for RL, and model-based approaches such as system identification, certainty equivalence, and model predictive control for LBC. The course will mostly focus on tabular cases and unconstrained cases. Function approximations and safety RL/LBC will also be discussed if time permits.

Prerequisites

MATH 257, MATH 415, or equivalent course on linear algebra and MATH 362, MATH 461, or equivalent course on probability. Previous knowledge of random processes, machine learning, and/or control systems is strongly recommended.

Tentative Schedule

Part 1: When and How to Model Problems as Markov Decision Processes (MDP) and Optimal Control (OC) Weeks 1-2

In this part, we will introduce the basic formulation of MDP and Optimal Control and discuss how to model problems in these forms.

Part 2: When Models are Known. Weeks 2-6.

In this part, we will learn an important and commonly used tool to handle both MDP and OC: Dynamic Programming (DP).

We will start with the simplest setting: finite-horizon formulation, then cover the infinite-horizon formulation, and briefly discuss the partial observation case.

We will also introduce popular algorithms inspired by DP, such as value iteration, policy iteration, rollout algorithms, and model predictive control.

Part 3: When Models are Unknown. Weeks 7-10.

We will discuss major directions of algorithm designs for RL and LBC: including Q learning, policy gradient, actor-critic, certainty equivalence, etc. We will also introduce an important tool for RL in both algorithmic and theoretical senses: multi-armed bandit.

Part 4: When Models are Too Large: Approximations and Neural Networks. Week 10-12

This part will introduce state/action/function approximations for RL and LBC to tackle the curse-of-dimensionality challenge.

Part 5: State-of-the-art Research Topics. Weeks 11-13.

This part will introduce some advanced research topics on RL and LBC. Some sessions will have invited speakers discussing state-of-the-art research progress. Some tentative topics include:

- Safe and Robust RL.
- Online-learning and online control.
- Multi-agent MDP and RL.
- RL theory.
- Applications in e.g. robotics.

Tentative speakers include

- Prof. Laixi Shi, Johns Hopkins
- Prof. Nan Jiang, UIUC
- Prof. Kaiqing Zhang, U of Maryland
- Prof. Guannan Qu, CMU
- Prof. Vasileios Tzoumas, U of Michigan

More Information on the Final Project

40% of your final grade will come from the final project. The goal of the project is for you to write a NeurIPS-style paper, and give a cogent talk on the selected topic. The project has two intended outcomes: (i) enable you to critique, extend and apply published results on dynamic programming and reinforcement learning, and (ii) allow you to walk through the paper-writing process. The scope of the project is determined by the student and approved by the instructor, following the criteria outlined below.

1. Project Scope You have two overlapping options. You can single out one of the tracks below for your project scope, or propose a project that combines both tracks.

1). *Review and critique one existing paper on reinforcement learning (RL), and/or learning-based control (LBC).* I will upload a set of potential papers to Canvas before the midterm, but you can suggest your own. The review **should not** merely summarize the results of the paper: it should contextualize the work with existing literature (both prior to and after publication), explain proof techniques used, present a cogent review of the results and suggest potential extensions, applications or alternative proofs. Your analysis should build on concepts seen in class.

2). Apply what we learned from this course to different research fields. You can apply the tools learned in class to your own research field, or other fields you may be interested in. If you choose to present a review of existing applications of reinforcement learning and/or learning-based control to a field, a cohesive literature survey should be presented, and the criteria described in item (1) should be satisfied. If you are trying out something new, the report must show a clear effort (successful or not) in using reinforcement learning and/or learning-based control to tackle a specific problem and/or model in the chosen field. You will not be measured by your success, but by showing clear effort in formulating an interesting problem, and analyzing it using techniques from class. **Successful projects that lead to new theorems, proof techniques, extensions of the existing theorems and proofs, and/or new numerical results will get extra points**.

2. Project Dynamics and Evaluation The project can be done in groups of **up to two students**. Group size will be taken into account when evaluating both the final write-up and the in-class presentations. The project will have four components that will be used to determine the final score.

1). *Project proposal.* We ask that you put together a project proposal (1 page, 11pt font, single column) to be discussed with the instructor. The proposal is due on **Nov 3** 2025. The project proposal will count towards **10% of your final grade**.

2). *Proposal discussion meeting.* Before **Nov 10**, it is your responsibility to coordinate with the instructor to schedule a **15 minute meeting** to discuss your proposal and progress. You should be prepared to answer questions about your project choice. The meeting should start with a **1 minute elevator pitch** about your proposed project. Additional meetings may be scheduled at the instructor's discretion.

3). *In-class presentation.* In December, we will have in-class presentations of the final projects. The exact duration of the presentation will be determined during the semester. This presentation should be treated as a **short conference talk**, with one exception: it should also start with a **1 minute elevator pitch about the idea**. The presentation must have a story arch, even if the ending is "the work is still in progress." The in-class presentation together with the proposal meeting will count towards **15% of your final grade**.

4). *Final paper.* The final paper is due on Dec 15 2025. The paper should be typeset in the NeurIPS conference format (style files will be uploaded to the course website), and should be at most 5 pages long except for references. The following sections (or equivalent) must be included in the final paper: abstract, introduction, notation, future work/conclusions, and bibliographical references. Other sections are at the students discretion. It is expected that the final paper is well-written in proper English, and reviewed for typos. The final paper will account for **15% of your final grade**.

Campus Policies

Emergency Response Recommendations: Emergency response recommendations can be found at the following website:

http://police.illinois.edu/emergency-preparedness/

You are encouraged to review this website and the campus building floor plans website within the first 10 days of class:

http://police.illinois.edu/emergency-preparedness/building-emergency-action-plans/

Academic Integrity: Students are expected to adhere to the Student Code:

http://studentcode.illinois.edu

and in particular, *Article 1, Part 4: Academic Integrity*.

Academic dishonesty will result in a sanction proportionate to the severity of the infraction, with possible sanctions described in 1-404 of the Student Code (https://studentcode.illinois.edu/article1/part4/1-404/). Every student is expected to review and abide by the Academic Integrity Policy as defined in the Student Code: https://studentcode.illinois.edu/article1/part4/1-401/. As a student it is your responsibility to refrain from infractions of academic integrity and from conduct that aids others in such infractions. A short

guide to academic integrity issues may be found at https://provost.illinois.edu/policies/policies/academic-integrity/students-quick-reference-guide-to-academic-integrity/. Ignorance of these policies is not an excuse for any academic dishonesty. It is your responsibility to read this policy to avoid any misunderstanding. Do not hesitate to ask the instructor(s) if you are ever in doubt about what constitutes plagiarism, cheating, or any other breach of academic integrity.

In this course you are expected to produce your own work in all assignments. Written assignments will be submitted through SafeAssign, a software tool that compares your writing against a large database as well as to the work of your current classmates and previously submitted assignments. Assignments with close matches to other work will be flagged and investigated.

In this course the use of calculators or electronic devices (cell phones or others) will **not** be allowed during examinations. If you are found using one, it will be investigated as potential cheating.

Religious Observances: In the case of missed exams or project presentations for religious observations, students should complete the *Request for Accommodation for Religious Observances* form:

https://cm.maxient.com/reportingform.php?UnivofIllinois&layout_id=19

Please make requests for absence letters as early as possible in the semester.

Privacy & FERPA: Any student who has suppressed their directory information pursuant to Family Educational Rights and Privacy Act (FERPA) should self-identify to the instructor to ensure protection of the privacy of their attendance in this course. See

https://registrar.illinois.edu/academic-records/ferpa/

for more information on FERPA.

Accommodations: To obtain disability-related academic adjustments and/or auxiliary aids, students with disabilities must contact the course instructor and the as soon as possible. To ensure that disability-related concerns are properly addressed from the beginning, students with disabilities who require assistance to participate in this class should contact Disability Resources and Educational Services (DRES) and see the instructor as soon as possible. If you need accommodations for any sort of disability, please speak to me after class, or make an appointment to see me or see me during my office hours. DRES provides students with academic accommodations, access, and support services. To contact DRES you may visit 1207 S. Oak St., Champaign, call 333-4603 (V/TDD), or e-mail disability@illinois.edu. For more information, please visit:

http://www.disability.illinois.edu/

Misconduct Reporting: The University of Illinois is committed to combating sexual misconduct. Faculty and staff members are required to report any instances of sexual misconduct to the University's Title IX and Disability Office. In turn, an individual with the Title IX and Disability Office will provide information about rights and options, including accommodations, support services, the campus disciplinary process, and law enforcement options.

A list of the designated University employees who, as counselors, confidential advisors, and medical professionals, do not have this reporting responsibility and can maintain confidentiality, can be found here:

http://wecare.illinois.edu/resources/students/#confidential

Other information about resources and reporting is available here:

http://wecare.illinois.edu