

Course Syllabus

18-813B: *Foundations of Reinforcement Learning*
Spring 2023

Instructor: Yuejie Chi
Office Location: Porter Hall B25
Email Address: yuejiechi@cmu.edu
Office Hours: TBD

Teaching Assistant: Jiin Woo
Email Address: jiinw@andrew.cmu.edu
Office Hours: TBD

Course Support: Academic Services Center
Office Location: HH 1113
Website: <https://www.ece.cmu.edu/academics/academic-services-center.html>

Course Description:

Reinforcement learning (RL), which is modeled as sequential decision making in the face of uncertainty, has garnered growing interest in recent years due to its remarkable success in practice. However, the explosion of complexity in emerging applications and the presence of nonconvexity exacerbate the challenge of achieving efficient RL in resource-constrained situations, where data collection and computation is expensive, time-consuming, or even high-stakes (e.g., in clinical trials, autonomous systems, and online advertising). Despite decades-long research efforts, however, the theoretical underpinnings of RL remain far from mature, especially when it comes to understanding and enhancing the sample and computational efficiencies of RL algorithms. An explosion of research has been conducted over the past few years towards advancing the frontiers of these topics, which leverage toolkits that sit at the intersection of multiple fields, including but not limited to control, optimization, statistics and learning. This course aims to present a coherent framework that covers important algorithmic developments in modern RL, highlighting the connections between new ideas and classical topics. Employing Markov Decision Processes (MDPs) as the central mathematical framework, we will cover multiple important scenarios including but not limited to the simulator setting, online RL, offline RL, and multi-agent RL, gravitating our discussions around issues such as sample complexity, computational efficiency, function approximation, distributional robustness, as well as information-theoretic and algorithmic-dependent lower bounds.

Number of Units: 12

Pre-requisites: None. Students should be familiar with basic probability, optimization and linear algebra.

Graduate Area: Artificial Intelligence

Class Lectures: Tuesdays and Thursdays, 3:30pm-4:50pm, WEH 5302

Required Textbook: None

Suggested Reading:

- Reinforcement Learning: Theory and Algorithms (draft), by Alekh Agarwal, Nan Jiang, Sham M. Kakade, Wen Sun
- Reinforcement learning: An introduction, by Richard S. Sutton, Andrew G. Barto
- Bandit Algorithms, by Tor Lattimore, Csaba Szepesvari

Reference Materials:

Brief List of Topics Covered:

- Introduction
- Multi-armed bandits
- Markov decision processes (MDPs)
- Dynamic programming
- Model-based approach (a.k.a. the plug-in approach)
- Model-free approach (e.g., temporal difference (TD) learning, synchronous Q-learning, asynchronous Q-learning, variance-reduced Q-learning)
- Information-theoretic lower bounds
- Online RL and the optimism principle
- Linear function approximation
- Offline RL and the pessimism principle
- Policy gradient methods
- Natural policy gradient methods
- Regularized RL via policy mirror descent
- Multi-agent RL (MARL): Markov games, Nash equilibrium, and coarse correlated equilibrium
- Representation learning in RL

Course Canvas:

Canvas login page: <https://cmu.instructure.com/>. You should check the course Canvas daily for announcements and handouts.

Course Website:

<https://users.ece.cmu.edu/~yuejiec/ece18813B.html>

Grading Algorithm:

20%	Homework
25%	Midterm Paper Presentations
55%	Final Projects (Proposal, Presentation, Final Report)

- **Homework problems (20%).** There will be two homework assignments that involve theory and/or programming components. Your homework must be turned in on Gradescope. One late day is allowed for all the homework; namely, one and only one of the homework can be turned in one day later than the marked deadline. You are encouraged to use LaTeX to typeset your homework.
- **Midterm Paper Presentation (25%).** An in-class presentation on a selected paper (either self-choice upon approval of the instructor or selected from a given pool) is arranged in lieu of the midterm. Students should provide a critical review and in-depth discussion of the selected paper and offer a presentation of 15-20 minutes. It is expected that the students should discuss at least one result from the paper in depths, by providing details of the proof or offering numerical simulations.
- **Final Projects (55%).** A significant component of this class is the final project, which can be either a literature review or original research:
 - (i) *Literature review.* We will provide a list of related papers not covered in the lectures, and the literature review should involve in-depth summaries and exposition of the specific topic (it is expected it is review of the “literature”, not a single paper).
 - (ii) *Original research.* It can be either theoretic or experimental (ideally a mix of the two). If you choose this option, you can do it either individually or in groups of 2. You are encouraged to combine your current research with your term project.

There are 3 milestones / deliverables to help you through the process.

- (i) Proposal. Submit a short report (NeurIPS format, up to 2 pages) stating the topic you plan to survey or the research problems that you plan to work on. Describe why they are important or interesting, and provide some appropriate references.
- (ii) In-class presentation. Prepare an oral presentation with slides. Introduce your topic to outsiders, focus on high-level ideas/insights/key results, and leave technical details to your report.

- (iii) A written report. You are expected to submit a final project report up to 5 pages (NeurIPS format) that summarizes your key findings, plus unlimited supplementary materials.

Tentative Course Calendar:

Date	Day	Class Activity	Assignment	Due Date
January				
17	Tue	Classes begin: introduction		
19	Thu	Multi-arm bandit		
24	Tue	Multi-arm bandit		
26	Thu	Multi-arm bandit	HW1	
31	Tue	MDP and dynamic programming		
February				
2	Thu	MDP and dynamic programming		
7	Tue	MDP and dynamic programming	HW2	
9	Thu	RL with a generative model		
14	Tue	RL with a generative model		
16	Thu	RL with a generative model		
21	Tue	Online RL		
23	Thu	Online RL		
28	Tue	Offline RL		
March				
2	Thu	Offline RL		
7	Tue	Spring Break – No Classes		
9	Thu	Spring Break – No Classes		
14	Tue	Midterm paper presentation		
16	Thu	Midterm paper presentation		
21	Tue	Policy optimization		
23	Thu	Policy optimization		
28	Tue	Actor-critic		
30	Thu	Linear function approximation		
April				
4	Tue	Linear function approximation		
6	Thu	Representation learning in RL		
11	Tue	Multi-agent RL		
13	Thu	Spring Carnival – No Classes		
18	Tue	Multi-agent RL		
20	Thu	POMDP		
25	Tue	Final presentation		
27	Thu	Final presentation		

Education Objectives (Relationship of Course to Program Outcomes)

The ECE department is accredited by ABET to ensure the quality of your education. ABET defines 7 Educational Objectives that are fulfilled by the sum total of all the courses you take. The following list describes which objectives are fulfilled by this course and in what manner they are fulfilled. The objectives are numbered from “1” through “7” in the standard ABET parlance. Those objectives not fulfilled by this course have been omitted from the following list:

1. an ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics
2. an ability to apply engineering design to produce solutions that meet specified needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors
3. an ability to communicate effectively with a range of audiences
4. an ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts
5. an ability to function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives
6. an ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions
7. an ability to acquire and apply new knowledge as needed, using appropriate learning strategies.

ECE Academic Integrity Policy

(<http://www.ece.cmu.edu/programs-admissions/masters/academic-integrity.html>):

The Department of Electrical and Computer Engineering adheres to the academic integrity policies set forth by Carnegie Mellon University and by the College of Engineering. ECE students should review fully and carefully Carnegie Mellon University's policies regarding Cheating and Plagiarism; Undergraduate Academic Discipline; and Graduate Academic Discipline. ECE graduate student should further review the Penalties for Graduate Student Academic Integrity Violations in CIT outlined in the CIT Policy on Graduate Student Academic Integrity Violations. In addition to the above university and college-level policies, it is ECE's policy that an ECE graduate student may not drop a course in which a disciplinary action is assessed or pending without the course instructor's explicit approval. Further, an ECE course instructor may set his/her own course-specific academic integrity policies that do not conflict with university and college-level policies; course-specific policies should be made available to the students in writing in the first week of class.

This policy applies, in all respects, to this course.

CMU Academic Integrity Policy (<http://www.cmu.edu/academic-integrity/index.html>):

In the midst of self-exploration, the high demands of a challenging academic environment can create situations where some students have difficulty exercising good judgment. Academic challenges can provide many opportunities for high standards to evolve if students actively reflect on these challenges and if the community supports discussions to aid in this process. It is the responsibility of the entire community to establish and maintain the integrity of our university.

This site is offered as a comprehensive and accessible resource compiling and organizing the multitude of information pertaining to academic integrity that is available from across

the university. These pages include practical information concerning policies, protocols and best practices as well as articulations of the institutional values from which the policies and protocols grew. The Carnegie Mellon Code, while not formally an honor code, serves as the foundation of these values and frames the expectations of our community with regard to personal integrity.

This policy applies, in all respects, to this course.

The Carnegie Mellon Code

Students at Carnegie Mellon, because they are members of an academic community dedicated to the achievement of excellence, are expected to meet the highest standards of personal, ethical and moral conduct possible.

These standards require personal integrity, a commitment to honesty without compromise, as well as truth without equivocation and a willingness to place the good of the community above the good of the self. Obligations once undertaken must be met, commitments kept.

As members of the Carnegie Mellon community, individuals are expected to uphold the standards of the community in addition to holding others accountable for said standards. It is rare that the life of a student in an academic community can be so private that it will not affect the community as a whole or that the above standards do not apply.

The discovery, advancement and communication of knowledge are not possible without a commitment to these standards. Creativity cannot exist without acknowledgment of the creativity of others. New knowledge cannot be developed without credit for prior knowledge. Without the ability to trust that these principles will be observed, an academic community cannot exist.

The commitment of its faculty, staff and students to these standards contributes to the high respect in which the Carnegie Mellon degree is held. Students must not destroy that respect by their failure to meet these standards. Students who cannot meet them should voluntarily withdraw from the university.

This policy applies, in all respects, to this course.

Carnegie Mellon University's Policy on Cheating

(<http://www.cmu.edu/academic-integrity/cheating/index.html>) states the following:

According to the University Policy on Academic Integrity, cheating "occurs when a student avails her/himself of an unfair or disallowed advantage which includes but is not limited to:

- Theft of or unauthorized access to an exam, answer key or other graded work from previous course offerings.
- Use of an alternate, stand-in or proxy during an examination.

- Copying from the examination or work of another person or source.
- Submission or use of falsified data.
- Using false statements to obtain additional time or other accommodation.
- Falsification of academic credentials.”

This policy applies, in all respects, to this course.

Carnegie Mellon University's Policy on Plagiarism

(<http://www.cmu.edu/academic-integrity/plagiarism/index.html>) states the following:

According to the University Policy on Academic Integrity, plagiarism "is defined as the use of work or concepts contributed by other individuals without proper attribution or citation. Unique ideas or materials taken from another source for either written or oral use must be fully acknowledged in academic work to be graded. Examples of sources expected to be referenced include but are not limited to:

- Text, either written or spoken, quoted directly or paraphrased.
- Graphic elements.
- Passages of music, existing either as sound or as notation.
- Mathematical proofs.
- Scientific data.
- Concepts or material derived from the work, published or unpublished, of another person."

This policy applies, in all respects, to this course.

Carnegie Mellon University's Policy on Unauthorized Assistance

(<http://www.cmu.edu/academic-integrity/collaboration/index.html>) states the following:

According to the University Policy on Academic Integrity, unauthorized assistance "refers to the use of sources of support that have not been specifically authorized in this policy statement or by the course instructor(s) in the completion of academic work to be graded. Such sources of support may include but are not limited to advice or help provided by another individual, published or unpublished written sources, and electronic sources. Examples of unauthorized assistance include but are not limited to:

- Collaboration on any assignment beyond the standards authorized by this policy statement and the course instructor(s).
- Submission of work completed or edited in whole or in part by another person.
- Supplying or communicating unauthorized information or materials, including graded work and answer keys from previous course offerings, in any way to another student.
- Use of unauthorized information or materials, including graded work and answer keys from previous course offerings.
- Use of unauthorized devices.

- Submission for credit of previously completed graded work in a second course without first obtaining permission from the instructor(s) of the second course. In the case of concurrent courses, permission to submit the same work for credit in two courses must be obtained from the instructors of both courses."

This policy applies, in all respects, to this course.

Carnegie Mellon University's Policy on Research Misconduct

(<http://www.cmu.edu/academic-integrity/research/index.html>) states the following:

According to the University Policy for Handling Alleged Misconduct in Research, "Carnegie Mellon University is responsible for the integrity of research conducted at the university. As a community of scholars, in which truth and integrity are fundamental, the university must establish procedures for the investigation of allegations of misconduct of research with due care to protect the rights of those accused, those making the allegations, and the university. Furthermore, federal regulations require the university to have explicit procedures for addressing incidents in which there are allegations of misconduct in research."

The policy goes on to note that "misconduct means:

- fabrication, falsification, plagiarism, or other serious deviation from accepted practices in proposing, carrying out, or reporting results from research;
- material failure to comply with Federal requirements for the protection of researchers, human subjects, or the public or for ensuring the welfare of laboratory animals; or
- failure to meet other material legal requirements governing research."

"To be deemed misconduct for the purposes of this policy, a 'material failure to comply with Federal requirements' or a 'failure to meet other material legal requirements' must be intentional or grossly negligent."

To become familiar with the expectations around the responsible conduct of research, please review the guidelines for Research Ethics published by the Office of Research Integrity and Compliance.

This policy applies, in all respects, to this course.

Take care of yourself. Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. You are not alone. There are many helpful resources available on campus and an important part of the college

experience is learning how to ask for help. Asking for support sooner rather than later is often helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call 412-268-2922 and visit their website at <http://www.cmu.edu/counseling/>. Consider reaching out to a friend, faculty or family member you trust for help getting connected to the support that can help.

If you have questions about this or your coursework, please let me know.

Every individual must be treated with respect. The ways we are diverse are many and are critical to excellence and an inclusive community. They include but are not limited to: race, color, national origin, sex, disability, age, sexual orientation, gender identity, religion, creed, ancestry, belief, veteran status, or genetic information. We at CMU, will work to promote diversity, equity and inclusion because it is just and necessary for innovation. Therefore, while we are imperfect, we will work inside and outside of our classrooms, to increase our commitment to build and sustain a community that embraces these values.

It is the responsibility of each of us to create a safer and more inclusive environment. Bias incidents, whether intentional or unintentional in their occurrence, contribute to creating an unwelcoming environment for individuals and groups at the university. If you experience or observe unfair or hostile treatment on the basis of identity, we encourage you to speak out for justice and support in the moment and and/or share your experience anonymously using the following resources:

Center for Student Diversity and Inclusion: csdi@andrew.cmu.edu, (412) 268-2150, www.cmu.edu/student-diversity

Report-It online anonymous reporting platform: www.reportit.net username: *tartans* password: *plaid*

All reports will be acknowledged, documented and a determination will be made regarding a course of action.” All experiences shared will be used to transform the campus climate.