CS 4180/5180: Reinforcement Learning and Sequential Decision Making (Fall 2020)

1 General Information

Time: Monday, Wednesday 2:50–4:30 Location: 103 Churchill Hall

2 Teaching Staff

- **The preferred platform for asking questions and contacting staff is Piazza.**
- If e-mail contact is necessary (e.g., sending attachments), **the preferred e-mail address that reaches all staff is** `cs41805180-staff@ccs.neu.edu`.
- Only e-mail individual staff if absolutely necessary (e.g., confidential issue), and note that response will typically be slower than contacting all staff via Piazza or the staff mailing list.

Role	Name and E-mail	Office Hours	Location
Instructor	Lawson L.S. Wong `lsw@ccs.neu.edu`	Fri 1–3 and by appointment	Teams
TA	Shuo Jiang `jiang.shuo@northeastern.edu`	Thu 5–7	Teams
TA	Lingzhi Kong `kong.ling@northeastern.edu`	Tue 2–4	Teams
IA	Juhner Yatin Padia `padia.j@northeastern.edu`	N/A	N/A

3 Course Overview

This course will introduce students to the fundamentals of reinforcement learning (RL) and sequential decision making, as well as a selection of recent advances in the field. The course is centered around the framework of Markov decision processes (MDPs), including aspects of modeling, planning, and learning in MDPs.

The content of the course emphasizes the foundational ideas and algorithms of reinforcement learning, and only lightly covers a small portion of the recent advances that have generated much recent excitement in the field. Due to the rapidly developing nature of the field, and partly because we believe it is premature to consider the latest developments, we will omit most algorithms that were introduced in the past five years. Nevertheless, by the end of the course, students will have developed a sufficiently broad set of technical tools, and have gained sufficent depth of understanding in the core concepts, that will enable them to solve real-world problems, self-learn additional techniques, and pursue further specialized content in reinforcement learning. Students are also encouraged to investigate recent methods in their course projects.

The course material will focus on problem formulation, models, and algorithms. Applications will be discussed when relevant, but will not be the focus of the content. However, in the spirit of experiential learning, there will be significant opportunities for implementation and application, through the programming assignments and the project.

4 Textbook and Reference Materials

The main textbook is *Reinforcement Learning: An Introduction* (2nd edition), by Richard S. Sutton and Andrew G. Barto. This textbook gives a broad introductory overview of most topics that are covered in this class. We will follow the book closely for the first two months of the course, covering Part I and sections of Part II in detail.

- *Reinforcement Learning* (2nd edition), by Richard S. Sutton and Andrew G. Barto.
  Website (with free PDF): `http://www.incompleteideas.net/book/the-book-2nd.html`

In the final part of the course (November), we will be covering more recent material drawn from recent papers in reinforcement learning. Links to relevant papers will be provided then.

When learning, sometimes it helps to be exposed to multiple perspectives of the same content. The following material from courses at other institutions is optional, but may be helpful to view when the textbook / lectures are insufficient:

- UCL: Advanced Deep Learning and Reinforcement Learning (2018)
  This is a two-part course with interleaved lectures on deep learning and reinforcement learning. The reinforcement learning portion strand of the course follows our textbook closely, and therefore is aligned with our course.
  Lecture videos:
  `https://www.youtube.com/playlist?list=PLqYmG7hTraZDNJre23vqCGIVpfZ_K2RZs`

- UC Berkeley: CS 285 – Deep Reinforcement Learning (2019)
  This is the latest complete offering of a course on deep RL; there is another offering this semester but most of the material is not present yet. The focus of the course is on recent advances in reinforcement learning, and therefore is less aligned with our course. However, this would be excellent supplementary material for those wishing to learn more about the frontiers of the field (the part we only lightly cover in November).
  Course website: `http://rail.eecs.berkeley.edu/deeprlcourse-fa19/`
  Lecture videos:
  `https://www.youtube.com/playlist?list=PLkFD6_40KJIwhWJpGazJ9VSj9CFMkb79A`

5 Prerequisites

- All programming assignments must be completed in Python 3. In the second half of the course, you will need to install and use OpenAI Gym and PyTorch. A Linux-based OS such as Ubuntu is recommended, but other common operating systems should work as well.

- Familiarity with mathematical concepts, including probability, optimization, linear algebra, and calculus. Familiarity with reading pseudocode and implementing algorithms.

- For implementing deep RL algorithms such as those covered in the final part of the course, it may help to have access to a CUDA-enabled GPU, but it is not strictly necessary.

6 Hybrid NUflex

This is the inaugural semester of Hybrid NUflex; as such, we would greatly appreciate your patience as we explore this new territory together. The course will be conducted in hybrid format, with the instructor teaching in-person (when possible), and students participating both in-person and remotely. To facilitate this, we will use a combination of tools: Canvas, Teams, and Piazza. (We do not plan on using Zoom.)

- Canvas: `https://northeastern.instructure.com/courses/22218`
  Content: Course announcements, syllabus, lecture slides, assignments, grades.
  All official content except lectures and office hours will be posted on Canvas. The Canvas sites for CS 4180 and CS 5180 have been merged; CS 4180 is officially a "section" of CS 5180.

- Teams: `https://teams.northeastern.edu`
  Content: Lecture live-streams and recordings (when possible), office hours.
  All registered students have been added to the "Khoury - CS 4180 1/CS 5180 1 (Fall 2020)" team. We will live-stream lectures in the "Lectures" channel and hold office hours remotely in the "Office hours" channel. If possible, we would like to record lectures for students in the course to refer to later; we will ask for consent to record at the beginning of each lecture. These recordings will be made available only to students enrolled in the course, instructor of record, and any teaching and instructional assistants assigned to the course.

- Piazza: `https://piazza.com/northeastern/fall2020/cs41805180`
  Content: Q&A and discussion. **Preferred platform for contacting course staff.**
  The site also offers an excellent discussion forum, where both instructors and fellow students can answer questions. Everyone is encouraged to participate. Questions/notes can be posted anonymously or with identity, and may also be posted privately only to instructors. Note that posting questions/notes via Piazza will most likely result in faster responses compared to e-mailing individual instructors.
  Piazza is also embedded as an app in Canvas, but you can also use it separately.

7 Coursework

Type	Total weight	Frequency	Due dates
Exercises	40%	$\sim$ Weekly ($\sim$ 8–10 total)	Sunday (11:59 PM)
Exam	20%	1 total	November 9 (Monday)
Project	40%	1 total	See schedule below

- Exercises are based on the previous week of material; we roughly cover one topic per week. Exercises may include both written and programming components. Students may discuss the problems with other students, but must write up their own solutions / code up their own implementations. On each assignment, please also indicate who you discussed with (if any). Lateness: Up to two days late (24-hour period), penalized by 5% per day.
- The exam will be administered remotely during class time (2:50-4:30 PM on Monday, November 9). Details will be forthcoming; we expect the exam to be open book, open computer, and possibly even open Internet. No external help / collaboration is permitted.
- The project offers an opportunity to apply learned techniques on a substantial problem that interests the student. The project should be completed in teams of 1–3 (ideally 2). Further details and (non-exhaustive) topic suggestions will be provided in October. Here is a rough timeline for the project, but is subject to change:
  - October 25: Project proposal due
  - November 15: Milestone 1
  - December 4: Milestone 2
  - December 7/9: Presentation
  - December 11: Draft report
  - December 14/15: Interview / debriefing
  - December 18: Final report
- Students enrolled in CS 4180 will complete shorter versions of selected exercises, and will be graded more leniently overall.

8 Academic Integrity

Cheating and other acts of academic dishonesty will be referred to OSCCR (office of student conduct and conflict resolution) and the Khoury College of Computer Sciences.

9 Schedule (subject to change; version 20201115)

Date	Lec #	Topic	Reference	Assignments due (Fri 11:59 PM)
9/9	1	Course overview	Ch. 1	
9/14	2	Multi-armed bandits	Ch. 2	
9/16	3	Multi-armed bandits	Ch. 2	
9/18				Ex. 0: Introduction
9/21	4	Markov decision processes	Ch. 3	
9/23	5	Markov decision processes	Ch. 3	
9/25				Ex. 1: Bandits
9/28	6	Dynamic programming	Ch. 4	
9/30	7	Dynamic programming	Ch. 4	
10/4				Ex. 2: MDPs
10/5	8	Monte-Carlo methods	Ch. 5	
10/7	9	Monte-Carlo methods	Ch. 5	
10/11				Ex. 3: Dynamic programming
10/12		Columbus Day (no class)		
10/14	10	Monte-Carlo methods	Ch. 5	
10/19	11	Temporal-difference learning	Ch. 6	
10/21	12	Temporal-difference learning	Ch. 6	
10/21				Ex. 4: Monte-Carlo
10/25				*Project proposal due 10/25*
10/26	13	n-step bootstrapping	Ch. 7	
		Eligibility traces	Ch. 12	
10/28	14	Planning	Ch. 8	
11/1				Ex. 5: Temporal-difference
11/2	15	Function approximation	Ch. 9	
11/4	16	Function approximation	Ch. 9	
11/9		~~Exam (remote)~~ Cancelled		
11/11		Veterans' Day (no class)		Ex. 6: Planning
11/15				*Project milestone 1 due 11/15*
11/16	17	Function approximation	Ch. 10–11	
11/18	18	Deep Q-Networks	Ch. 16.1, 16.5	
11/22				
11/23	19	AlphaGo	Ch. 8.8–8.11, 16.6	Ex. 7: Function approximation
11/25		Thanksgiving (no class)		
11/30	20	Policy-gradient methods	Ch. 13	
12/2	21	Further topics in RL	TBD	
12/4				*Project milestone 2 due 12/4*
12/7		*Project presentations*		
12/9		*Project presentations*		
12/11				*Draft report due 12/11*
				Final report due 12/18