CS260R Reinforcement Learning
Winter 2026

Course Logistics

Instructor: Professor - Bolei Zhou <bolei@cs.ucla.edu>

TAs:
- Haoyuan Cai <haoyuan@cs.ucla.edu >
- Matthew Leng <matthewleng@cs.ucla.edu >

Lecture Time: Tuesday/Thursday 10:00 am - 11:40 am

Lecture Location
- **(in-person): ENG VI Mong Lecture hall (first floor of ENG VI)**

TA's Discussion Session and Office Hours (you can go to either one as long as space allows):
- DIS 1A, Friday 12 pm - 1:50 pm, DODD 170, Matthew Leng
- DIS 1B, Friday 2 pm - 3:50 pm, BOLTER 2444, Haoyuan Cai

Prof. Zhou's Office Hours: Eng VI, 295D, Thursday 5:00 pm - 5:45 pm

Piazza Forum Link: https://piazza.com/ucla/winter2026/cs260r
TA's expected response time on Piazza: less than 36 hours

Bruinlearn course site: https://bruinlearn.ucla.edu/courses/221115

Prerequisites: Familiarity with **Python programming**, **Linear Algebra**, **Calculus**, and **Probability**. Students should have taken **at least one course relevant** to machine learning or pattern recognition or data mining.

Grade Structure: Letter grades are assigned based on the following:

Attendance	10%
Assignments	50%
Final exam	40%

Textbooks: No textbook is required; the slides are self-contained. The following optional textbooks are excellent references (all PDF available):

- Sutton and Barton: http://incompleteideas.net/book/the-book-2nd.html
- Kevin P. Murphyhttps://arxiv.org/pdf/2412.05265
- Dimitri P. Bertsekas https://web.mit.edu/dimitrib/www/RLbook.html

Details on Homework Assignments: There will be three assignments, released on the following schedule. Each problem set is allotted 2 weeks for completion. However, you will be allotted 3 late days for assignments of your choice. Beyond this 3-day grace period, late assignments will not be accepted without a valid explanation. The grace period is calculated on a per-day basis. **Any portion of a 24-hour period counts as one full day.** For example, if an assignment is submitted 25 hours after the deadline, it will be considered two late days.

Assignments are released here at
https://github.com/ucla-rlcourse/cs260r-assignment-2026winter/

Assignment 0	OUT: now	DUE: End of Week 1
Assignment 1	OUT: End of Week 1	DUE: End of Week 3
Assignment 2	OUT: End of Week 3	DUE: End of Week 5
Assignment 3	OUT: End of Week 5	DUE: End of Week 7
Assignment 4	OUT: End of Week 7	DUE: End of Week 9
Mini project*	OUT: End of week 6	DUE: End of Week 10

* Mini project is an open-ended tournament. Tutorials will be provided in discussion. More information is yet to come. The mini-project will be a one-person project.
** End means 23:59 Sunday of that week.

Final exam: Tuesday, March 17, 2026, 3:00 PM - 6:00 PM at TBD location

Course Description

This course covers fundamental topics in reinforcement learning, a computational approach to learning in which an agent seeks to maximize the total reward it receives while interacting with a complex, uncertain environment. Recent progress in deep reinforcement learning and its applications will be discussed.

Learning Objectives

- Students should understand the reinforcement learning foundations such as value-based RL, policy-based RL, and actor-critic.

- Students should understand the standard RL algorithms such as PPO, SAC, and apply them to some applications.
- Students should understand the strengths and limitations of RL and its real-world applications.

Video recordings

The lectures are recorded. You can access the recorded video in the Assignment tab in BruinLearn. The recordings are only for officially enrolled students.

Schedule (to be updated)

Course slides are released here

Week 1: Overview
Lecture 1: Course introduction
Lecture 2: RL basics and coding with RL
Action: Assignment 0 and Assignment 1 out, Assignment 0 Due
Discussion Session: N/A

Week 2: MDP and policy/value iteration
Lecture 1: Markov decision process
Lecture 2: Policy iteration and value iteration
Discussion Session: Policy Evaluation/Iteration Example

Week 3: Tabular methods
Lecture 1: Model-free prediction
Lecture 2: Model-free control
Action: Assignment 1 due, Assignment 2 out
Discussion Session: Value Iteration

Week 4: Value function approximation and deep Q learning
Lecture 1: value function approximation
Lecture 2: Deep Q Learning
Discussion Session: Assignment 1 Review, Value Function Approximation

Week 5: Policy-based RL: basics
Lecture 1: Policy Optimization 1
Lecture 2: Policy Optimization 2
Action: Assignment 2 due, Assignment 3 out
Discussion Session: Policy Gradient

Week 6: Policy-based RL: advanced methods
Lecture 1: Policy Optimization 3
Lecture 2: Policy Optimization 4
Action: Mini project out
Discussion Session: Assignment 2 Review

Week 7: Model-based RL
Lecture 1: Model-based RL
Lecture 2: Connection to optimal control
Action: Assignment 3 due, Assignment 4 out
Discussion: Tutorial on course project

Week 8: Advanced topics
Lecture 1: Imitation learning
Lecture 2: Human-in-the-loop RL (PVP)
Discussion: Assignment 3 Review

Week 9: Advanced topics
Lecture 1: Distributed computing and RL system design
Lecture 2: Course summary
Action: Assignment 4 due, Final Exam Sample Out
Discussion: Q&A

Week 10: Advanced topics
Lecture 1: RL for LLM (TBD)
Lecture 2: Final Exam Sample Review + Q&A (by TAs)
Discussion: Assignment 4 Review

Week 11: Final

Tuesday: Final Exam, 3:00 pm - 6:00 pm at TBD

Thursday EOD: Mini Project Due (Hard DDL, no grace period applied)