Professors

Scott Niekum
sniekum@cs.utexas.edu

GDC 3.404

Peter Stone
pstone@cs.utexas.edu
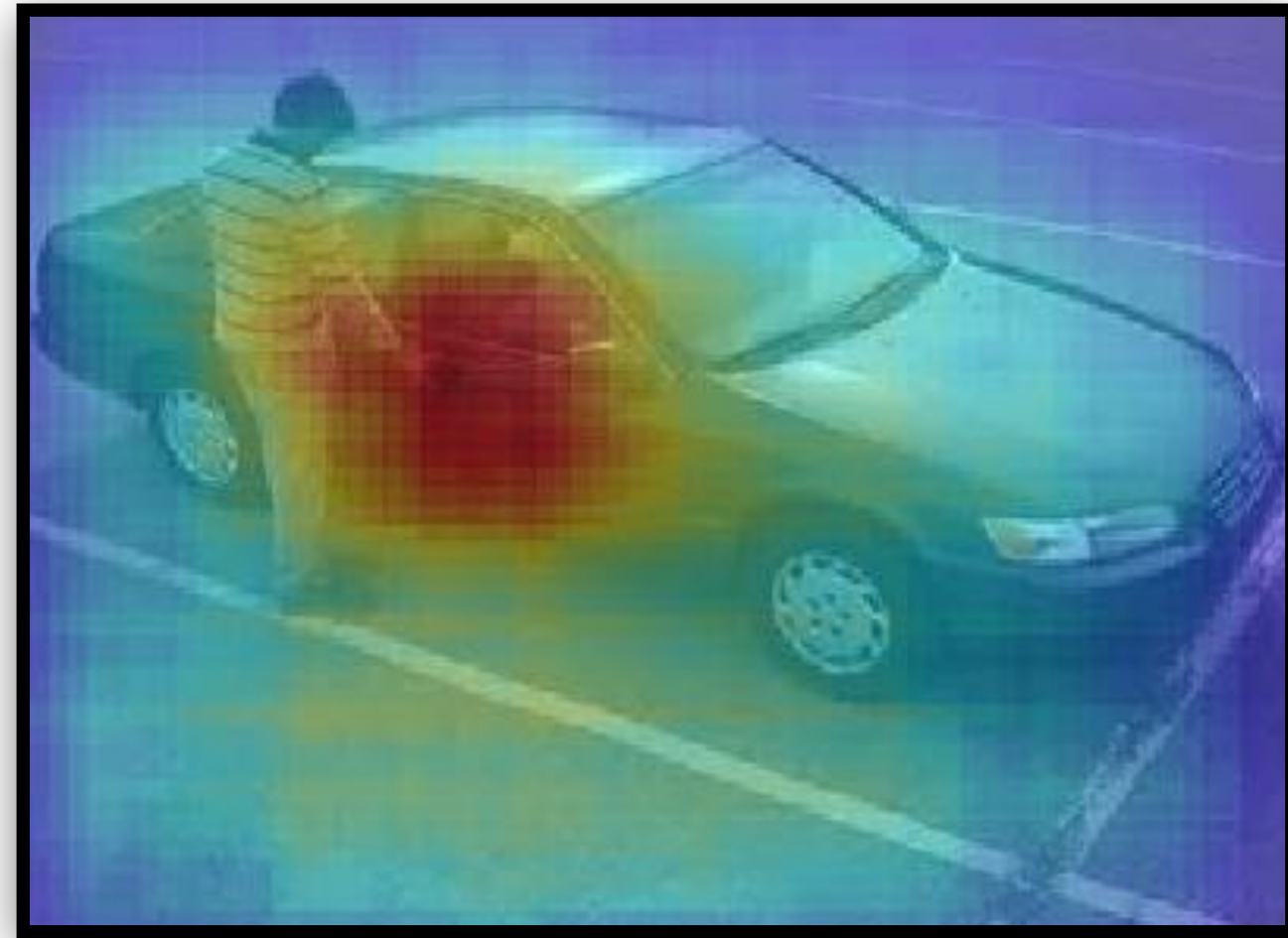
GDC 3.508

Teaching Assistants

Harshit Sikchi
hsikchi@utexas.edu

William Macke
wmacke@cs.utexas.edu

Rohan Nair
rohan.nair@utexas.edu

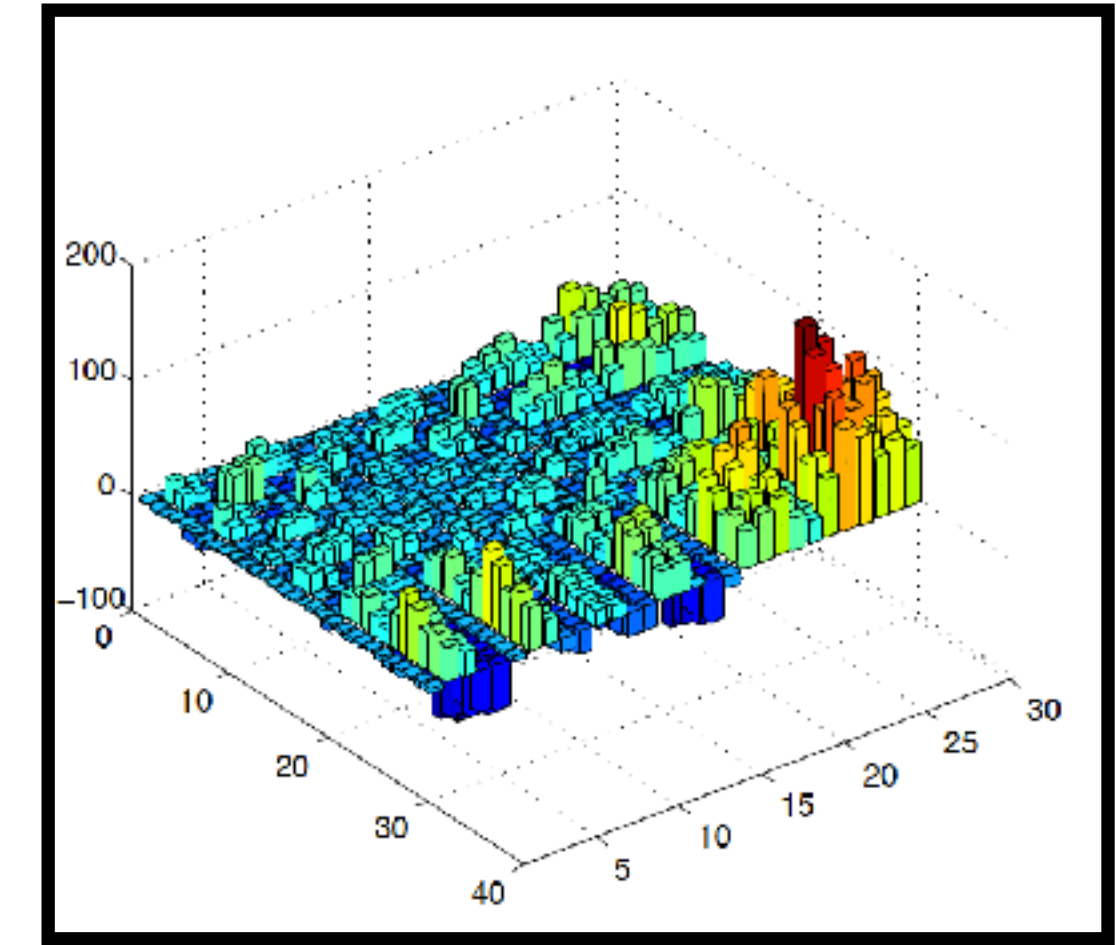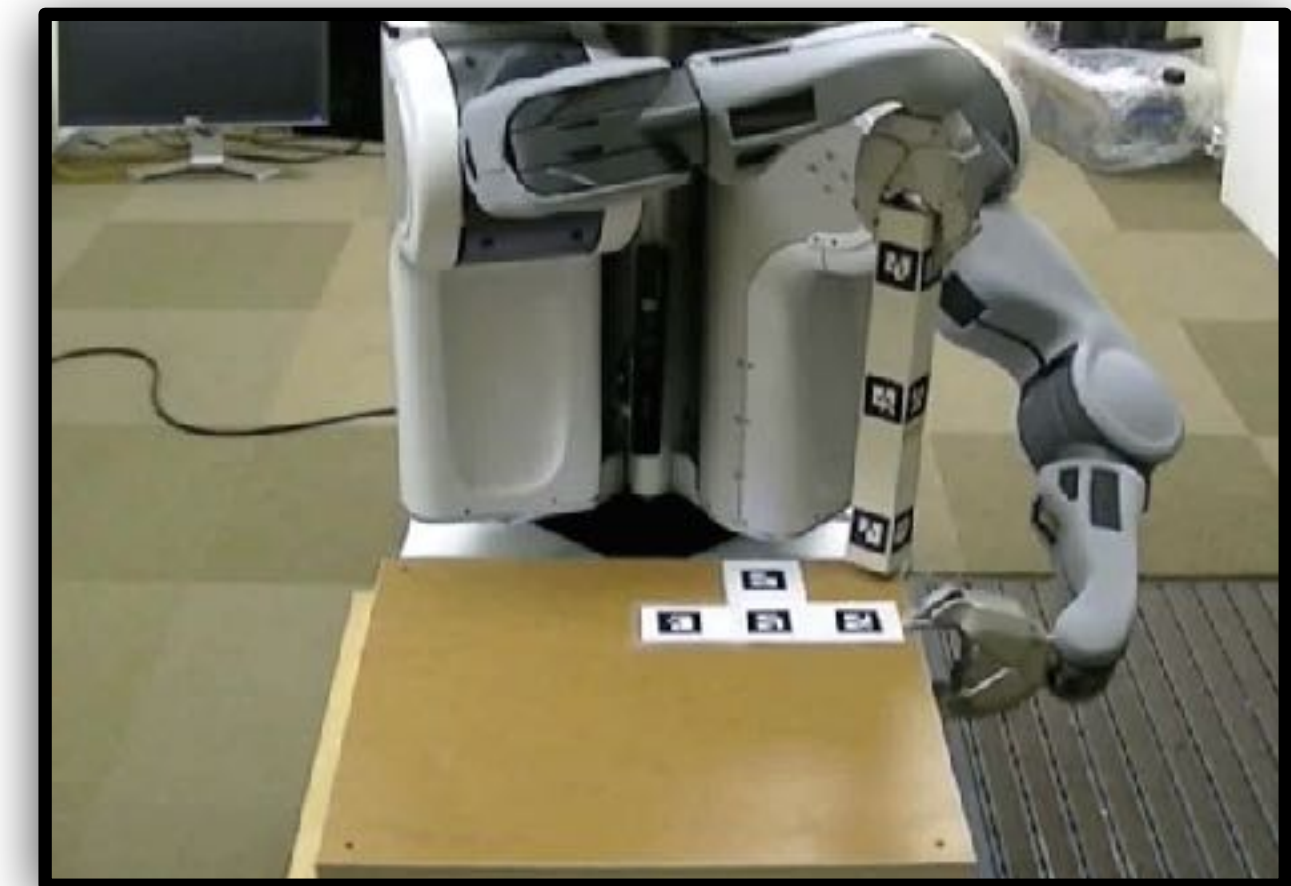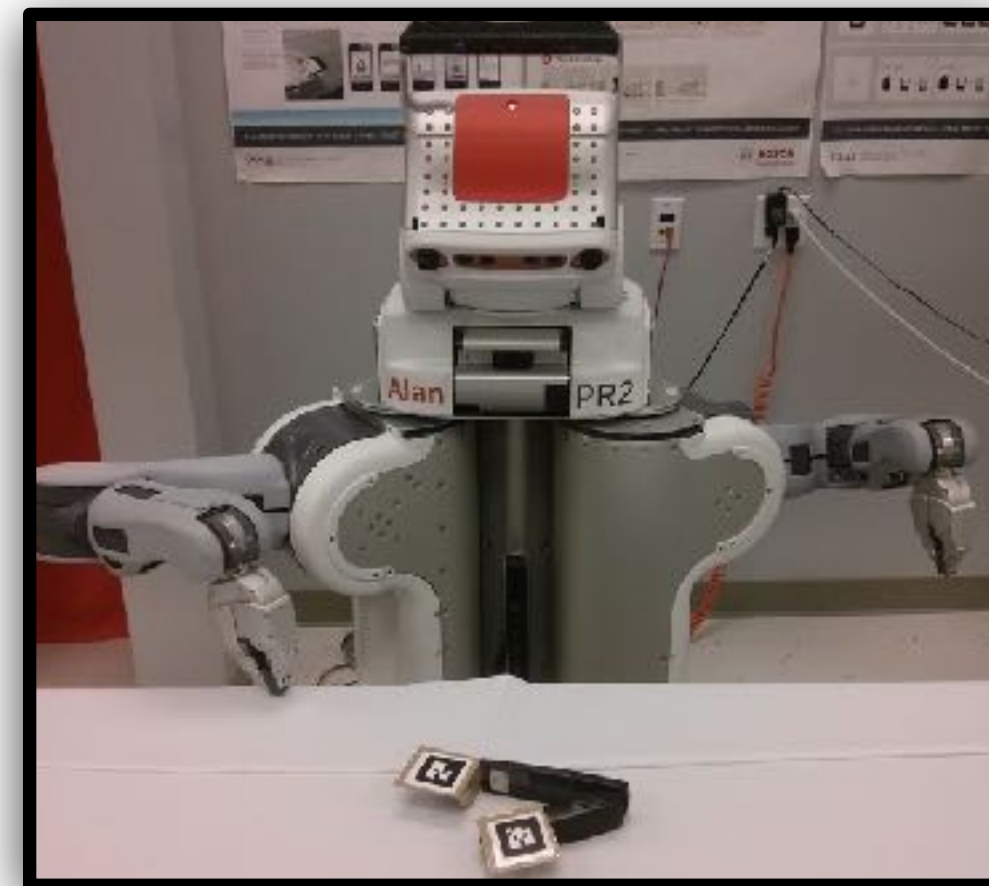Zizhao Wang
zizhao.wang@utexas.edu

A bit about me

Perception

PeARL
Personal Autonomous Robotics Lab

Reinforcement learning

Robotic manipulation, imitation learning, and safety

Course Information

- Communication:

  - Course info on main website

  - Grades on edX

  - Piazza for discussion forums

- edX Edge

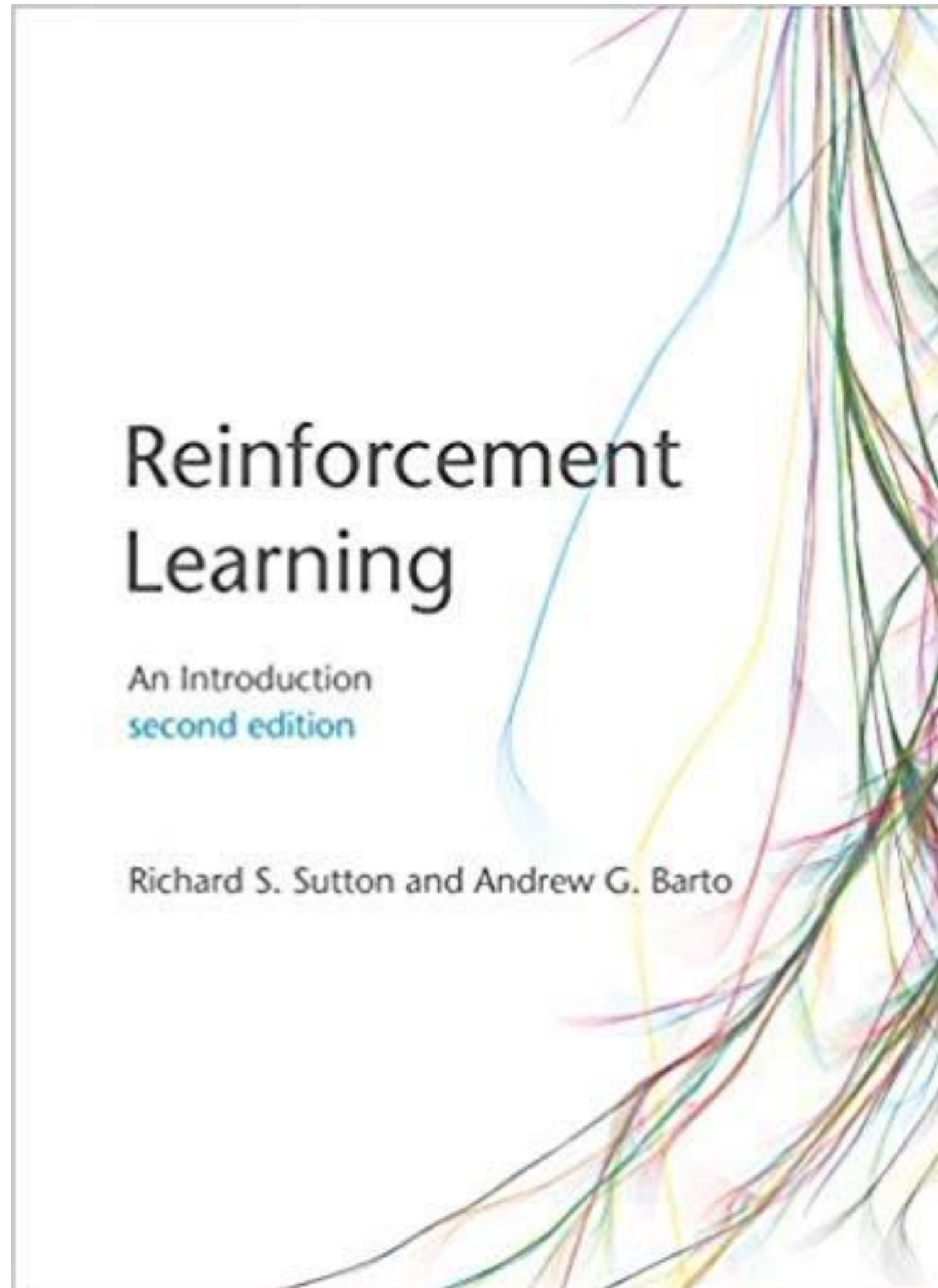 - Interactive homework problems

  - Autograded programming projects

  - Create an edX **Edge** account immediately

  - Use your real full name!

Class website:

https://www.cs.utexas.edu/~pstone/Courses/394Rspring22/

Textbook

**Reinforcement Learning: An Introduction
Second Edition**
Richard S. Sutton and Andrew G. Barto

http://incompleteideas.net/book/the-book-2nd.html

The **free online edition** is better than physical:
Many errors have been corrected!

Readings

Readings are extremely important!

We will not review many of the concepts from the book in class

Instead, they are a **prerequisite** for understanding the class discussion

Readings

Each Monday **by 5pm**, you must submit a written response to that week's readings, which may include:

- Insightful questions
- Clarification questions about ambiguities
- Comments about the relation of the reading to previous readings
- Critiques
- Thoughts on what you would like to learn about in more detail
- Possible extensions or related studies
- Summaries of the most important things you learned

Homework Exercises

- Online on edX

- Autograded text boxes / multiple choice

- Goal: self-assess and prepare for midterm

- Can discuss at high-level, but work alone

- Some problems randomized

- No spoilers on forum discussions!

Programming Assignments

- Projects will be in Python — learn basics ASAP if not familiar

- Roughly one every two weeks, but may adjust

- Submitted and autograded on edX

- Implement core algorithms introduced in the book

- Deeper explorations into some topics outside of the book

Lateness policy

- Reading responses **will not be accepted late**, as they are critical for class discussion

- edX homework and programming assignments will have deadlines posted on the course website.

- 20% penalty for late submissions until end of semester

Exams

- Midterm covering book material

- One page of notes, but not open book

- No final exam; open-ended project instead

Grading

Plus/minus grading - adjustable scale, but no more harsh than:

```
                        A [94-100]    A- [90-94)
        B+ [87-90)      B [84-87)     B- [80-84)
        C+ [77-80)      C [74-77)     C- [70-74)
        D+ [67-70)      D [64-67)     D- [60-64)
        F  [0-60)
```

Grades will be weighted as follows:

- Written reading responses / class participation (10%)
- edX exercises (15%)
- Programming assignments (25%)
- Midterm (25%)
- Final project (25%)

Academic Honesty

- Discuss concepts, but don't share solutions or written work with other students

- Don't look for answers / code online or elsewhere

- Automated tools will be used to discover cheating

- If unsure, check departmental guidelines or ask — ignorance is not an excuse

- We will pursue the harshest penalties available, so please don't cheat!

- To be clear: you will fail the class automatically and be reported to the university

Important This Week

- Important this week:

  - **Create and edX edge account and register** for the class — **use your real full name!**

  - **Catch up on the reading if you missed the first one**

  - **First reading response due tomorrow!**

  - **Get familiar with Python if you aren't already**

- Also:

  - **If you are wait-listed,** you may or may not get in depending on how many students drop. Be patient if possible — many students often drop early in the course.

  - **Office Hours** begin next week

Course Topics

Core RL (book)

Tabular methods

Bandits

MDPs

Dynamic Programming

Monte Carlo

Temporal Difference

Planning

Function approx

Prediction

Control

Eligibility Traces

Off-Policy RL

Policy Gradient

Advanced topics

Applications and case studies

Abstractions and hierarchy

Learning from humans

Exploration

Modern methods