Professors

Amy Zhang
amy.zhang@austin.utexas.edu

EER 6.878

Peter Stone
pstone@cs.utexas.edu

GDC 3.508

Teaching Assistants

Caroline Wang
caroline.l.wang@utexas.edu
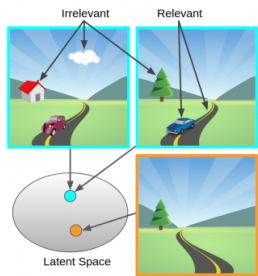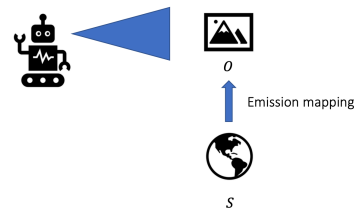
Haoran Xu
haoran.xu@utexas.edu

Shuozhe Li
shuozhe@cs.utexas.edu

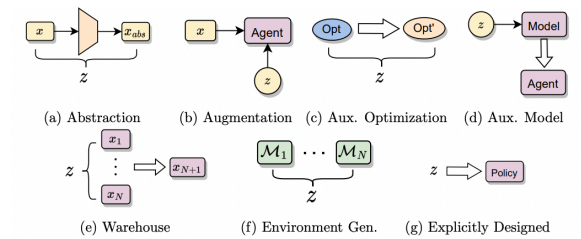Siddhant Agarwal
siddhant@cs.utexas.edu

A bit about me

Reinforcement learning

Structure

Generalization

State abstractions
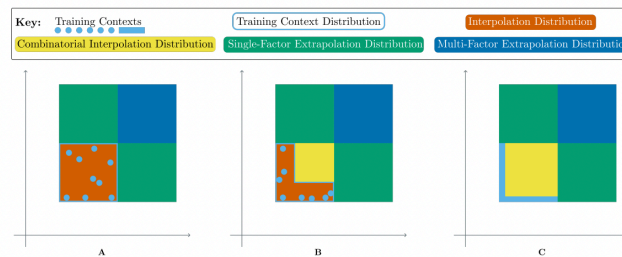
Course Information

Communication:

- Course info on main website
- Grades on edX

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/394Rspring24/

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!

---

hw1_search_q4_a*_graph_search

🔖 Bookmark this page

Q4: A* Graph Search
8.0 points possible (graded)

Consider A* *graph* search on the graph below. Arcs are labeled with action costs and states are labeled with heuristic values. Assume that ties are broken alphabetically (so a partial plan S->X->A would be expanded before S->X->B and S->A->Z would be expanded before S->B->A.
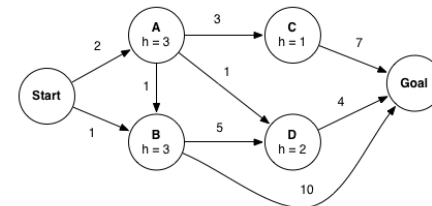
In what order are states expanded by A* graph search? You may find it helpful to execute the search on scratch paper.

○ Start, A, B, C, D, Goal

○ Start, A, C, Goal

○ Start, B, A, D, C, Goal

○ Start, A, D, Goal

○ Start, A, B, Goal

○ Start, B, A, D, B, C, Goal

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%)
  - This includes attendance! We will be checking with live polls.
- edX exercises (15%)
- Programming assignments (25%)
- Midterm (20%)
- Final project (30%)

Academic Honesty

READ THE STATEMENT IN THE SYLLABUS

- 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

Academic Honesty

LLMs

- Please remember that the words in the reading response must be your own. No portion may be copied from the textbook, online, or any generative AI tool without appearing in quotation marks and with full attribution.

- The ideas presented should be your own, and the tool is used for polishing only.

- If you do use a tool: In the response, below the polished version, include the initial (raw) version of your response and acknowledge explicitly what tool(s) you used to improve your language and how. This gives us a trail we can verify if there is ambiguity w.r.t. originality.

- Make sure you completely stand behind the ultimate submission. If you don't agree with or understand the output of the tool, then you should edit it.

Important This Week

- Important this week:

  - **Create an 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