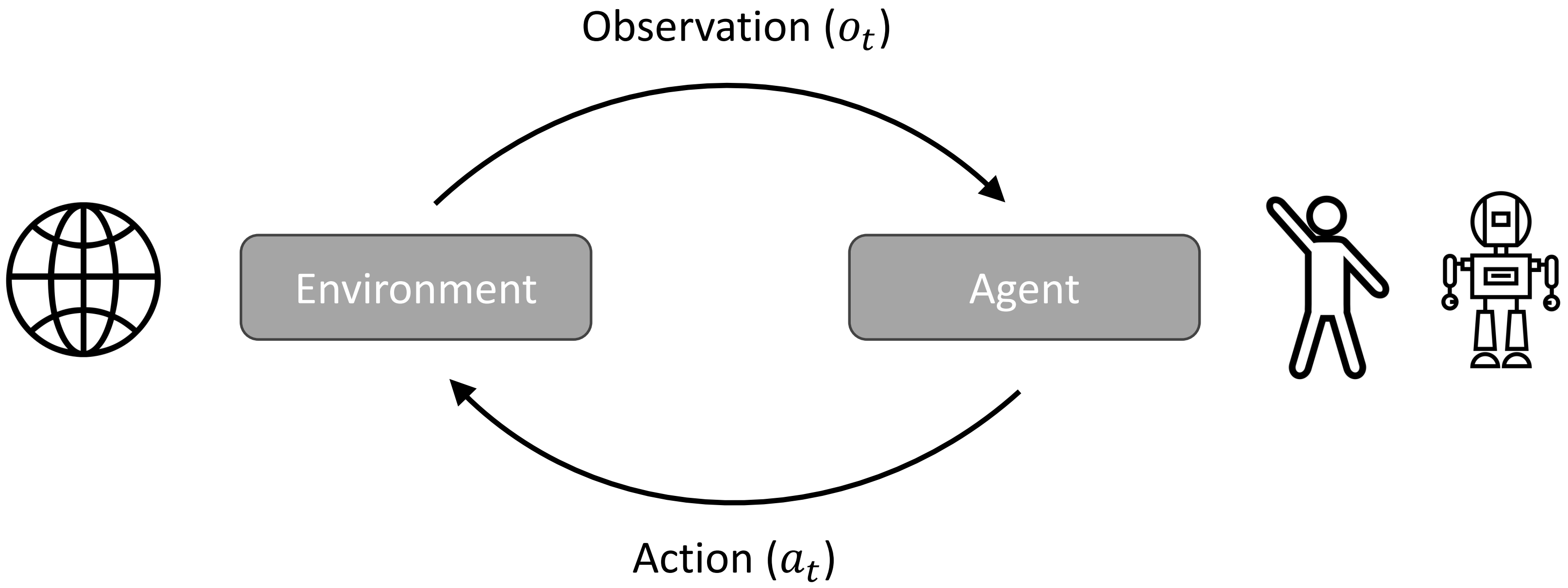
CSE 574 Planning and Learning Methods in AI

Ransalu Senanayake

www.ransalu.com


Observation (o_t)
Environment
Agent
Action (a_t)

Methods for Designing Decision Agents

- Explicit programming
  - Anticipate all the different scenarios the agent might find
  - Explicitly program the agent
- Through supervised learning
  - It might be easier to show what to do than explicitly program E.g., Behavioral cloning
- Optimization
  - Designer specifies the space of possible strategies and a performance metric
  - Dynamic model should be known but not used to guide the search
- Planning
  - Assumes the model of the problem dynamics is known and it is used to help guide the search
- Reinforcement learning
  - We don’t need to know the model

Class Logistics

- Teaching team
- Communication
- Prerequisites and infrastructure
- Learning outcomes
- Assessments and grading
- Schedule

https://lens-ai-lab.github.io/CSE574_Fall23/

LENS Lab

Welcome to the Laboratory for Learning Evaluation of autoNomous Systems (LENS Lab). We are interested in operationalizing machine learning (ML) models in autonomous systems and robots. We delve deeper into critical questions:

- How can we explain black-box neural networks?
- How can we evaluate uncertainty of decisions?
- When and how do models fail?
- Are models fair and unbiased?
- Will a model perform effectively in new or changing environments?

We examine these questions before, during, and after deployment. These efforts not only aid engineers in debugging models but also assist legislative bodies in establishing legal and ethical guidelines.

Twitter
https://twitter.com › ransalu

LinkedIn
https://www.linkedin.com › ransalu

Ransalu Senanayake
(Instructor)

E-mail: ransalu@asu...

Office hours: Fri,

11:00 AM - 12:00 PM

Virtual: Meeting Link

Yancheng Wang
(Teaching Assistant)

E-mail:

ywan1053@asu...

Office hours: Wed,

2:00 PM - 3:00 PM

Virtual: Meeting Link

Office: BYENG 211

Akshara Trichambaram
(Instructional Assistant)

E-mail:

atricham@asu...

Office hours: Mon,

11:00 AM - 12:00 PM

Virtual: Meeting Link

Suresh Kondepudi

(Grader)

E-mail:

nkondepu@asu...

Office hours: Ad-hoc

(post grading)

Harsh Mankodiya

(Grader)

E-mail:

mankodi@asu...

Office hours: Ad-hoc

(post grading)

Communication

- **General Concerns**: For any other concerns related to assignments (such as late or missed submissions), accommodations, or any other issues, contact the IA via email. If you do not receive a response within 48 hours, contact the instructor.

- **Technical Questions**: For any technical questions regarding assignments, use Canvas. If you do not receive an answer within 48 hours, first contact the TA via email and subsequently the instructor.

Note: You get extra credits for participating in Canvas discussions

Prerequisites

Highly recommended

- CSE 471: Introduction to AI
- CSE 475: Foundations of Machine Learning

Computing and Software

No labs. Three programming assignments and a project.

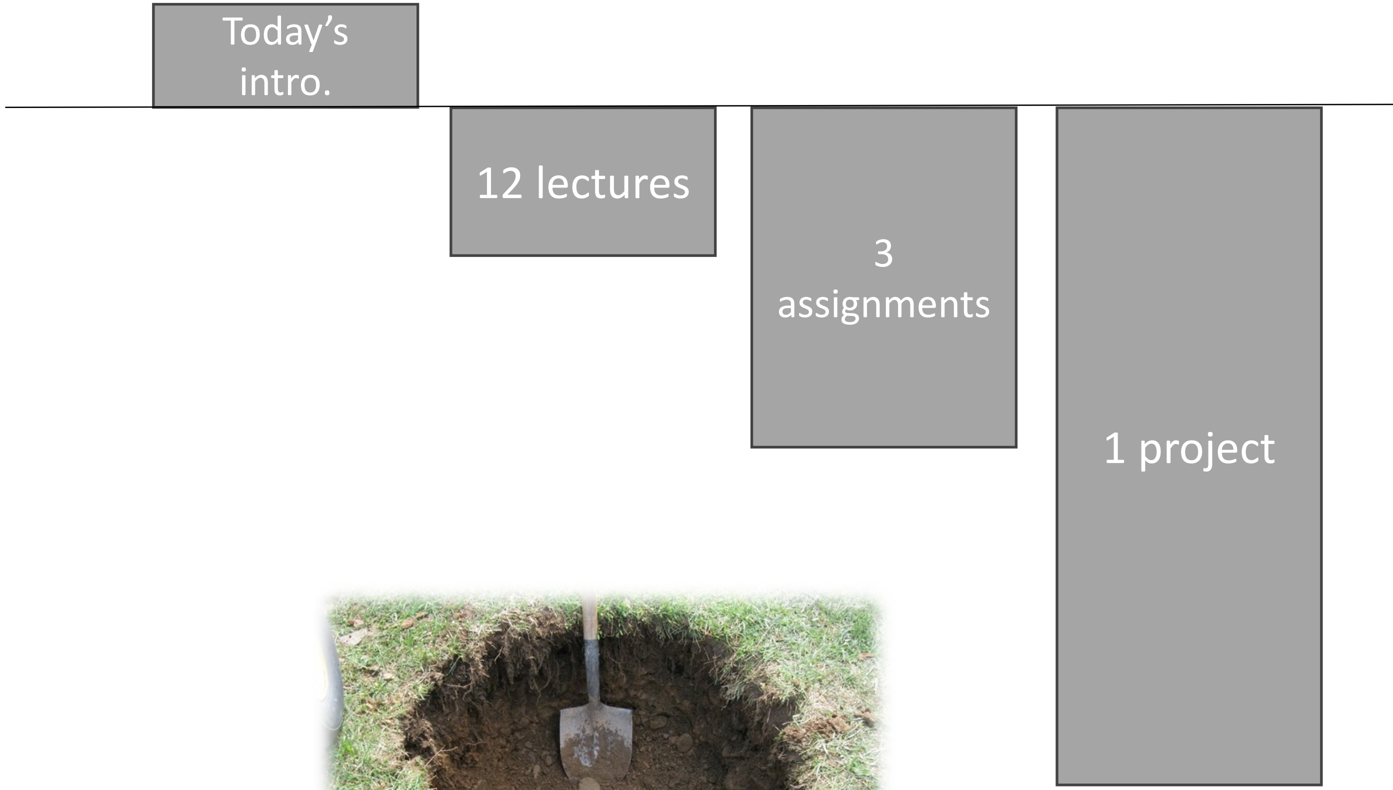
- A computer with Python (PyTorch)
- Sol/Agave supercomputer - we'll send instructions later

Learning Outcomes

- Understand the role of planning and learning in AI-based decision-making systems
- Explain the fundamental concepts and techniques of AI planning and machine learning
- Analyze and compare different AI planning algorithms and approaches, highlighting their strengths and weaknesses
- Design and implement AI planning and learning techniques to develop intelligent systems capable of adaptive decision-making
- Stay up-to-date on recent advancements as well as challenges in AI planning and learning by reading and understanding research papers in the field
- Collaborate effectively in teams to solve AI-related challenges and complete practical projects
- Communicate technical concepts related to AI planning and learning clearly and persuasively, both in written reports and oral presentations

Grades

- Participation (5%)
- Three assignments (50%)
  - Homework component (30%)
  - In-class component (20%)
- Group project (45%)
  - Proposal (5%)
  - Progress presentation (10%)
  - Final presentation (15%)
  - Final written report (15%)
- Extra credits



Grades

- Participation (5%)
- Three assignments (50%)
  - Homework component (30%)
  - In-class component (20%)
- Group project (45%)
  - Proposal (5%)
  - Progress presentation (10%)
  - Final presentation (15%)
  - Final written report (15%)
- Extra credits

- In-person
- Zoom (in emergencies)

  Watch live, AND

  Submit a paragraph via Canvas

  before 11.59pm on Sunday
- Must attend at least 11/14

Grades

- Participation (5%)
- Three assignments (50%)
  - Homework component (30%)
  - In-class component (20%)
- Group project (45%)
  - Proposal (5%)
  - Progress presentation (10%)
  - Final presentation (15%)
  - Final written report (15%)
- Extra credits

- Within 24 hours of the deadline, max grade = 90% of the total grade;
- Within 24 and 72 hours of the deadline, max grade = 50% of the total grade;
- Will not be accepted after 72 hours

Note: If you do submit a late homework, in addition to uploading to Canvas, send an email to both graders and IA with the homework as an attachment

Grades

- Participation (5%)
- Three assignments (50%)
  - Homework component (30%)
  - In-class component (20%)
- Group project (45%)
  - Proposal (5%)
  - Progress presentation (10%)
  - Final presentation (15%)
  - Final written report (15%)
- Extra credits

Final report:

- Within 24 hours of the deadline, max grade = 90% of the total grade;
- Within 24 and 72 hours of the deadline, max grade = 50% of the total grade;
- Will not be accepted after 72 hours

Note: If you do submit the report late, in addition to uploading to Canvas, send an email to both graders and IA with the report as an attachment

Grades

- Participation (5%)
- Three assignments (50%)
  - Homework component (30%)
  - In-class component (20%)
- Group project (45%)
  - Proposal (5%)
  - Progress presentation (10%)
  - Final presentation (15%)
  - Final written report (15%)
- Extra credits

This is a 16-week ongoing project. Start thinking about the project today. It cannot be completed within just a few weeks.

Final report:

- Within 24 hours of the deadline, max grade = 90% of the total grade;
- Within 24 and 72 hours of the deadline, max grade = 50% of the total grade;
- Will not be accepted after 72 hours

Note: If you do submit the report late, in addition to uploading to Canvas, send an email to both graders and IA with the report as an attachment

Grades

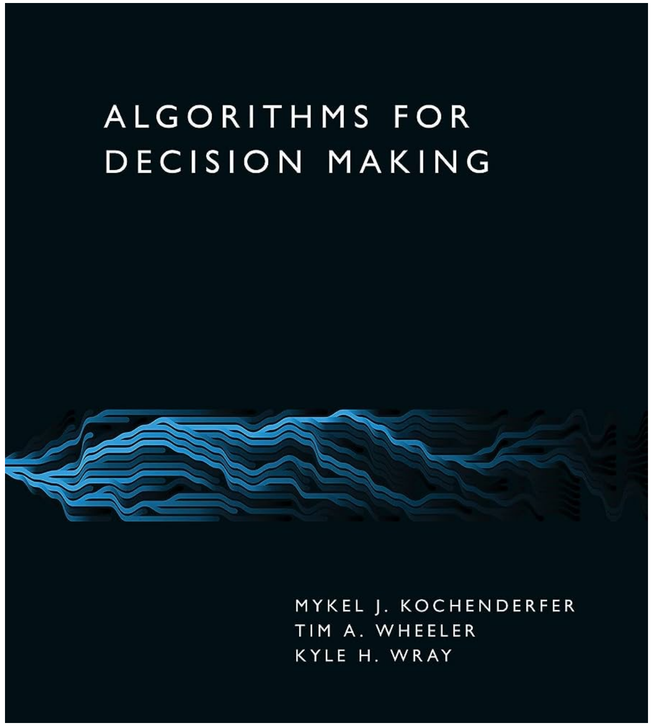
- Participation (5%)
- Three assignments (50%)
  - Homework component (30%)
  - In-class component (20%)
- Group project (45%)
  - Proposal (5%)
  - Progress presentation (10%)
  - Final presentation (15%)
  - Final written report (15%)
- Extra credits

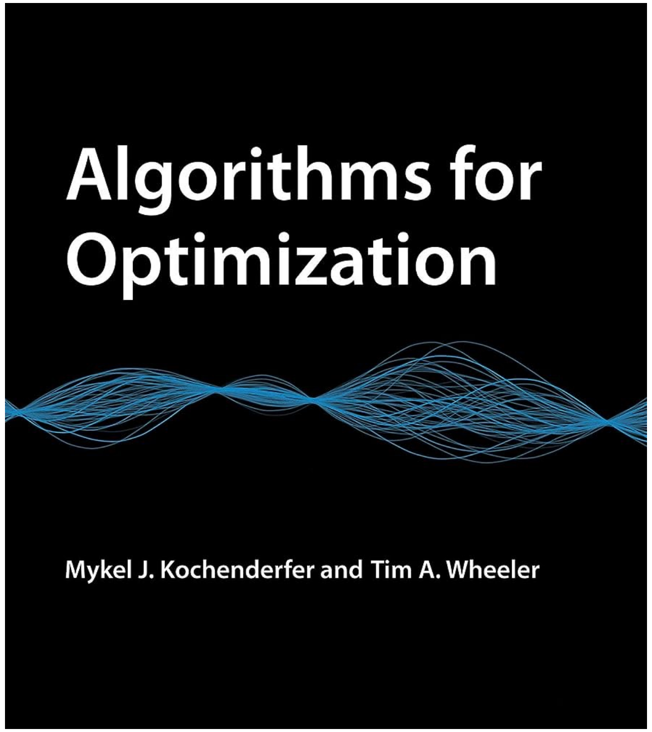
- Active discussion on slack, or
- Remarkable project

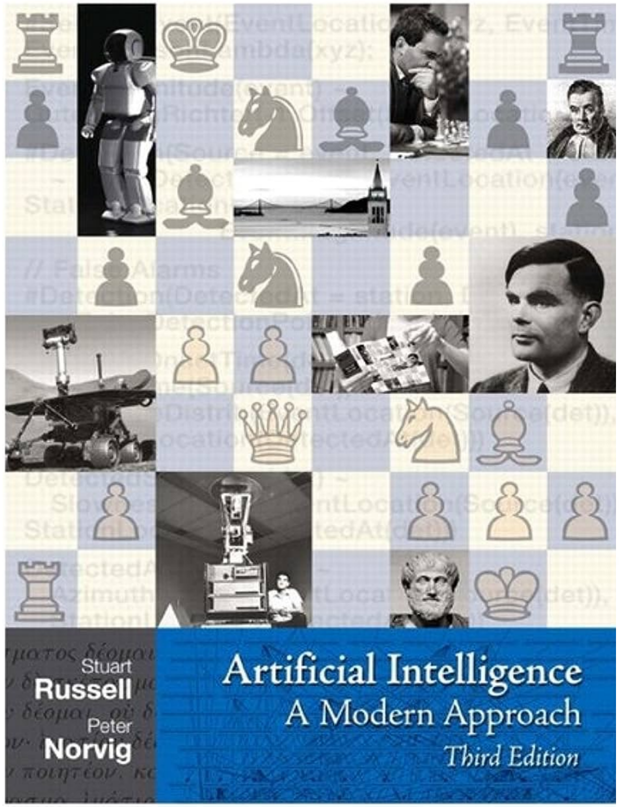
Schedule

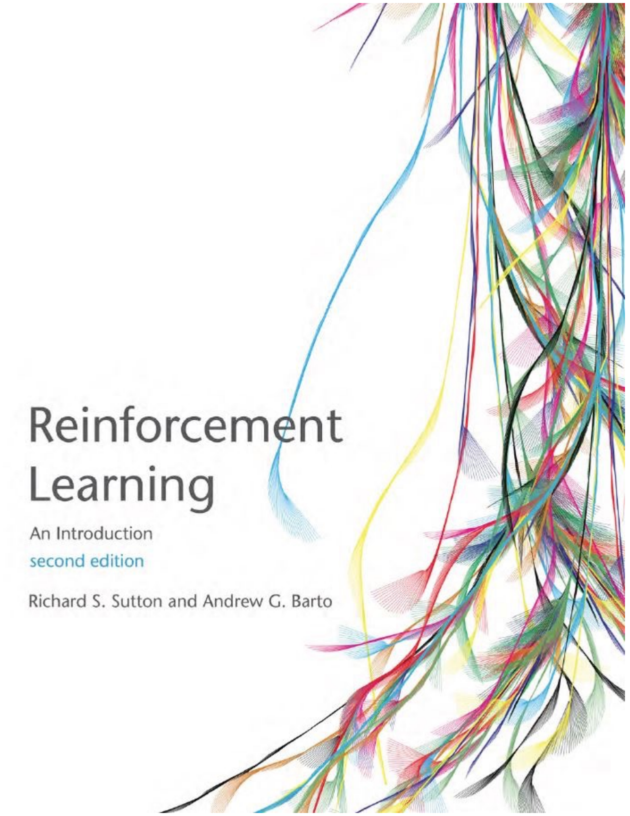
Week	Date	Theory
Week 1	08/18/2023	Introduction to decision-making
Week 2	08/25/2023	Myopic planning
Week 3	09/01/2023	Myopic planning
Week 4	09/08/2023	Non-myopic planning and reinforcement learning
Week 5	09/15/2023	Non-myopic planning and reinforcement learning
Week 6	09/22/2023	Imitation learning for decision-making
Week 7	09/29/2023	Imitation learning for decision-making
Week 8	10/06/2023	Human-in-the-loop planning
Week 9	10/13/2023	Human-in-the-loop planning
Week 10	10/20/2023	Task and motion planning
Week 11	10/27/2023	Task and motion planning
Week 12	11/03/2023	Multiagent planning and decision-making under uncertainty
~~Week 13~~	~~11/10/2023~~	Veterans Day (No class)
Week 14	11/17/2023	Classical planning and hierarchical planning/learning
~~Week 15~~	~~11/24/2023~~	Thanksgiving (No class)
Week 16	12/01/2023	Project presentations

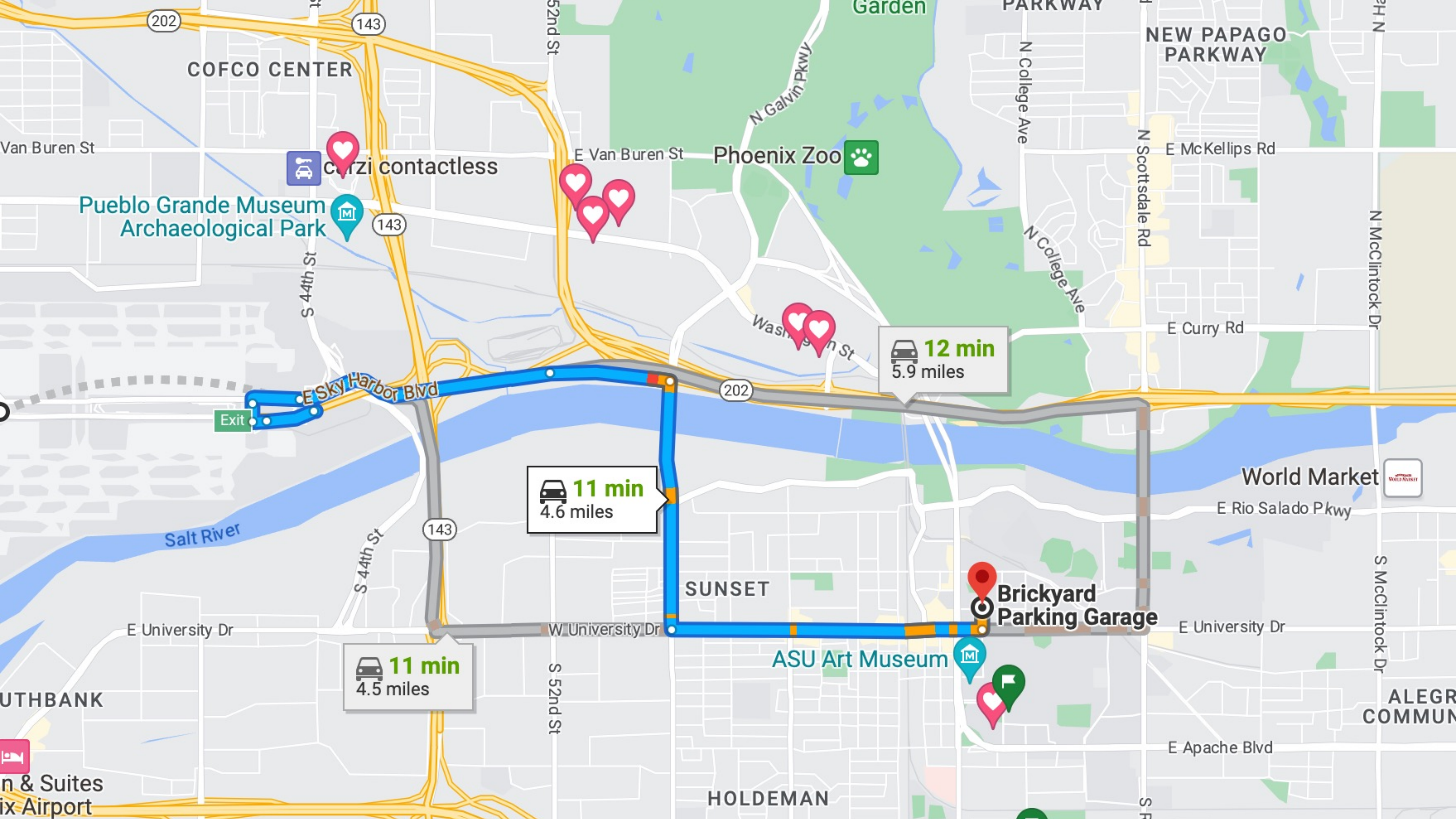
Textbooks

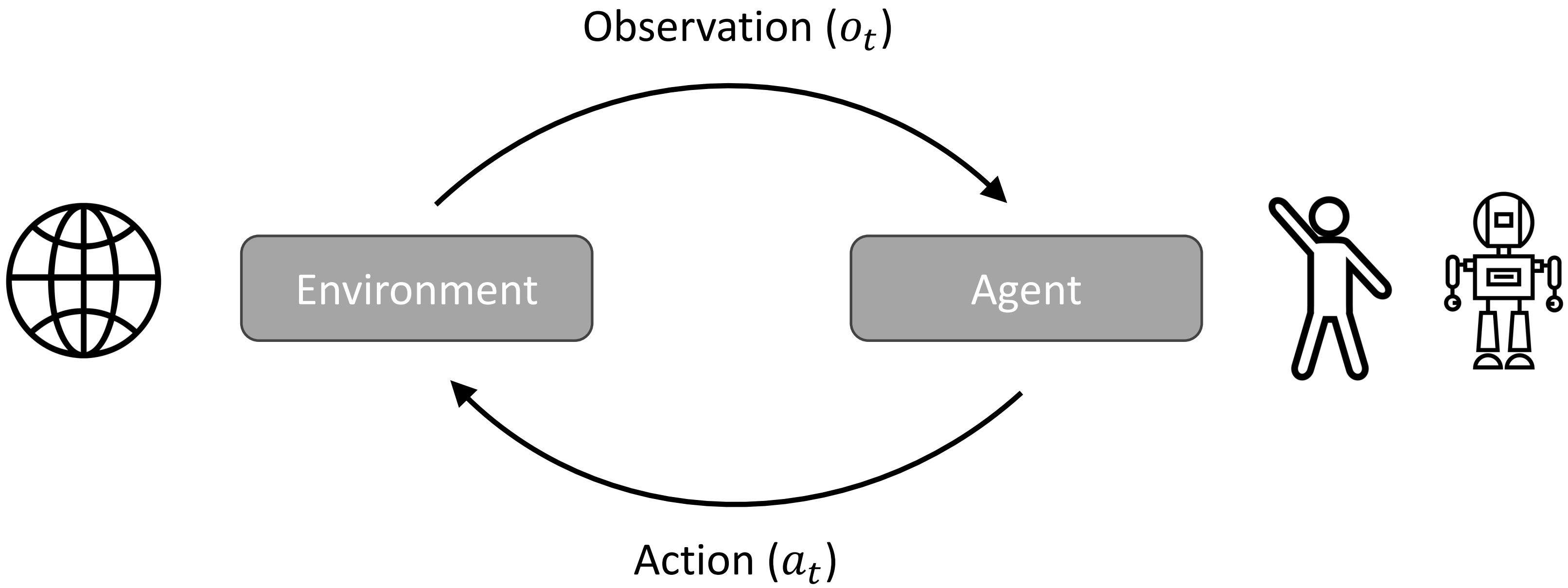

ALGORITHMS FOR
DECISION MAKING
MYKEL J. KOCHENDERFER
TIM A. WHEELER
KYLE H. WRAY


Algorithms for
Optimization
Mykel J. Kochenderfer and Tim A. Wheeler


Stuart
Russell
Peter
Norvig
Artificial Intelligence
A Modern Approach
Third Edition


Reinforcement
Learning
An Introduction
second edition
Richard S. Sutton and Andrew G. Barto


COFCO CENTER
Van Buren St
carzi contactless
Pueblo Grande Museum
Archaeological Park
E Van Buren St
Phoenix Zoo
N Galvin Pkwy
Garden
PARKWAY
NEW PAPAGO
PARKWAY
N College Ave
N Scottsdale Rd
E McKellips Rd
E Curry Rd
N McClintock Dr
S 44th St
52nd St
E Sky Harbor Blvd
Exit
12 min
5.9 miles
11 min
4.6 miles
11 min
4.5 miles
Salt River
World Market
E Rio Salado Pkwy
SUNSET
Brickyard
Parking Garage
E University Dr
W University Dr
ASU Art Museum
S 52nd St
S McClintock Dr
E Apache Blvd
HOLDEMAN
UTHBANK
n & Suites
ix Airport
ALEGR
COMMUN
202
143


Observation (o_t)
Environment
Agent
Action (a_t)


WAYMO
7ZGY053

Different aspects of a plan

- A plan
  - A sequence of actions
  - A policy (state → action mapping)

- Objective/s, constraints, and preferences

- Guarantees

  \- Optimality – based on some metrics (e.g., shortest path, highest expected value)

  \- Robustness – against perturbations to the environment

Terminology

- Input (sensors, symbolic/language inputs)
- Observations (fully (MDP), partially (belief state/POMDP))
- Policy (deterministic $\pi: S \rightarrow A$, stochastic. $\pi: S \times A \rightarrow [0,1]$)
- Multiagent, collaborative (with human - explainable), hierarchical

- Planning, control theory, exploration, learning (supervised, unsupervised, semi-supervised, self-supervised, reinforcement), operations research/optimization, scheduling, constrained satisfaction

Utility Theory

- When we make a decision there are different outcomes $x \in X$
- Each outcome has a probability p
- *Utility* indicates how worth an outcome is (worthiness depends on our *preferences*). A *utility function* maps outcomes to a real number $u: X \rightarrow \mathbb{R}$
- The total utility is $\sum_{x \in X} p_x u(x)$
- If we observe o and take action a, we end up in state s' with the probability $p(s'|o,a)$. Therefore, the expected utility is, $\mathbb{E}[u(a|o)] = \sum_{s'} p(s'|o,a) u(s')$
- A rational agent chose the action that maximizes the expected utility

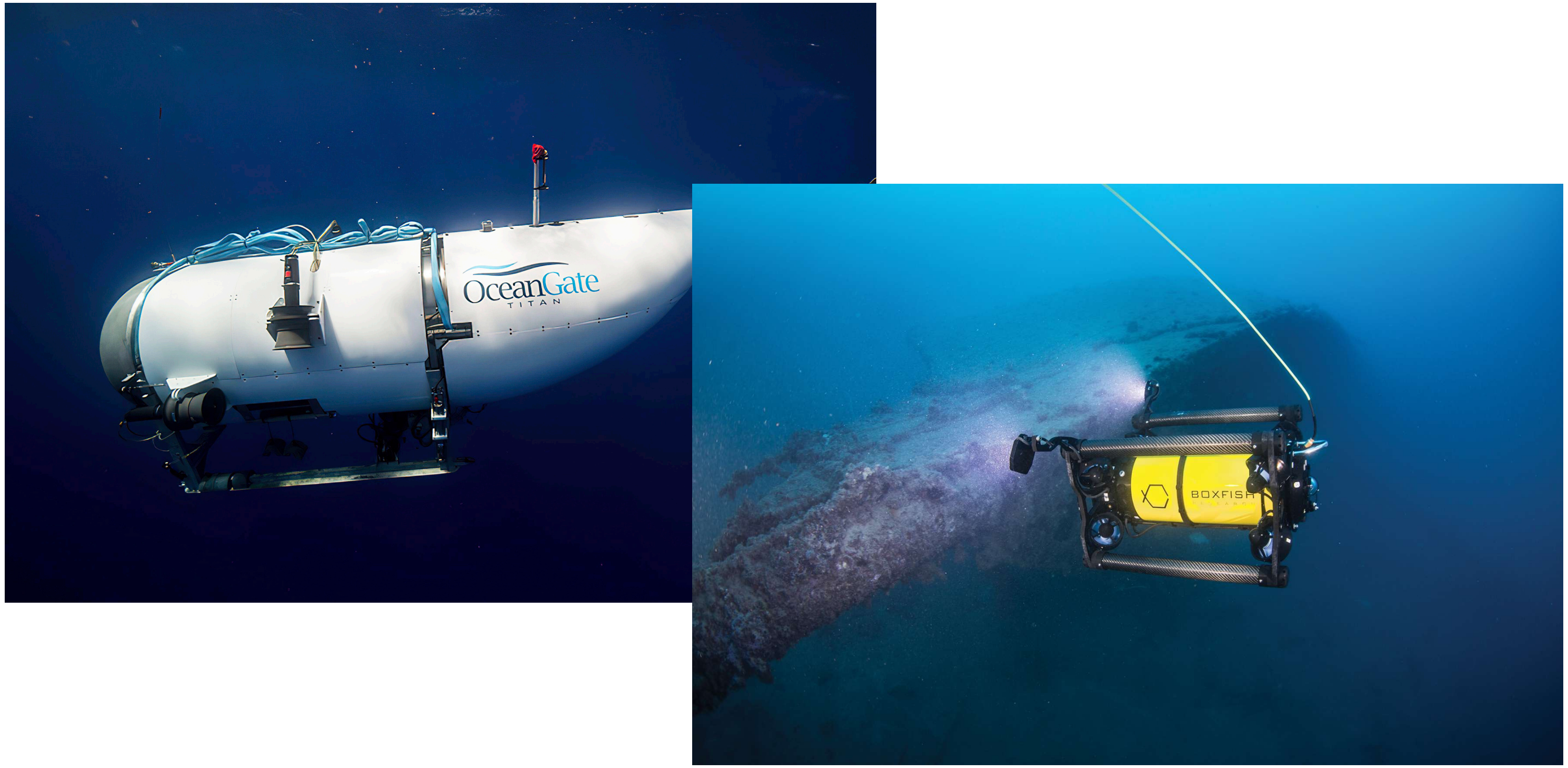
$$a^* = \underset{a}{\operatorname{argmax}}\, \mathbb{E}[u(a|o)]$$

Schedule

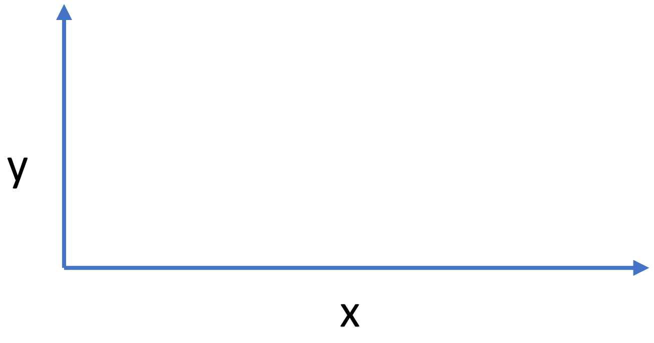
Week	Date	Theory
Week 1	08/18/2023	Introduction to decision-making
Week 2	08/25/2023	Myopic planning
Week 3	09/01/2023	Myopic planning
Week 4	09/08/2023	Non-myopic planning and reinforcement learning
Week 5	09/15/2023	Non-myopic planning and reinforcement learning
Week 6	09/22/2023	Imitation learning for decision-making
Week 7	09/29/2023	Imitation learning for decision-making
Week 8	10/06/2023	Human-in-the-loop planning
Week 9	10/13/2023	Human-in-the-loop planning
Week 10	10/20/2023	Task and motion planning
Week 11	10/27/2023	Task and motion planning
Week 12	11/03/2023	Multiagent planning and decision-making under uncertainty
~~Week 13~~	~~11/10/2023~~	Veterans Day (No class)
Week 14	11/17/2023	Classical planning and hierarchical planning/learning
~~Week 15~~	~~11/24/2023~~	Thanksgiving (No class)
Week 16	12/01/2023	Project presentations

Myopic Planning

- E.g., Bayesian Optimization

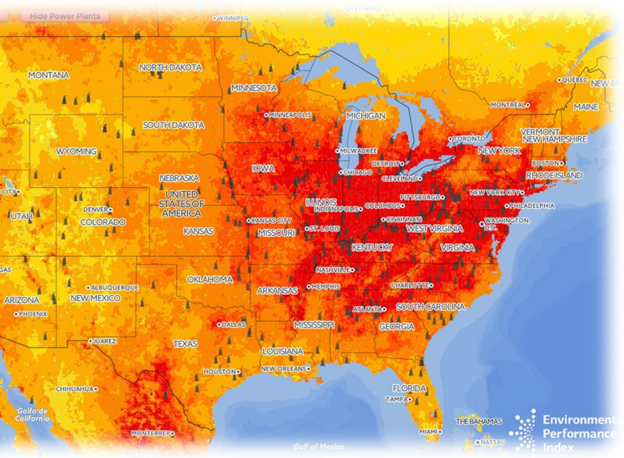


Informative path planning

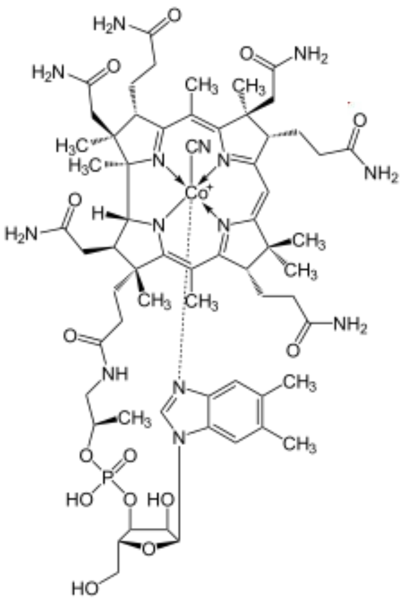


Applications

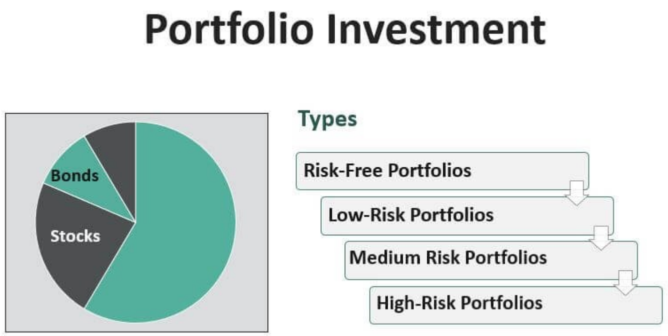
The underlying process is costly to evaluate



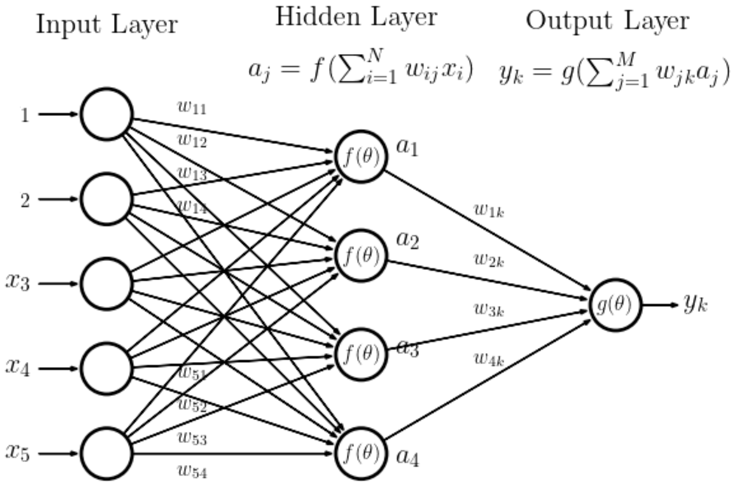
Environmental monitoring



Drug discovery



Financial decisions



Neural architecture search

Schedule

Week	Date		Theory	
Week 1	08/18/2023	Assignment 1	Introduction to decision-making	
Week 2	08/25/2023		Myopic planning	PyTorch tutorial
Week 3	09/01/2023		Myopic planning	

BoTorch

$$f_1(\mathbf{x}) = (1 + g(\mathbf{x}_M)) \cos\left(\frac{\pi}{2} x_1\right)$$

$$f_2(\mathbf{x}) = (1 + g(\mathbf{x}_M)) \sin\left(\frac{\pi}{2} x_1\right)$$

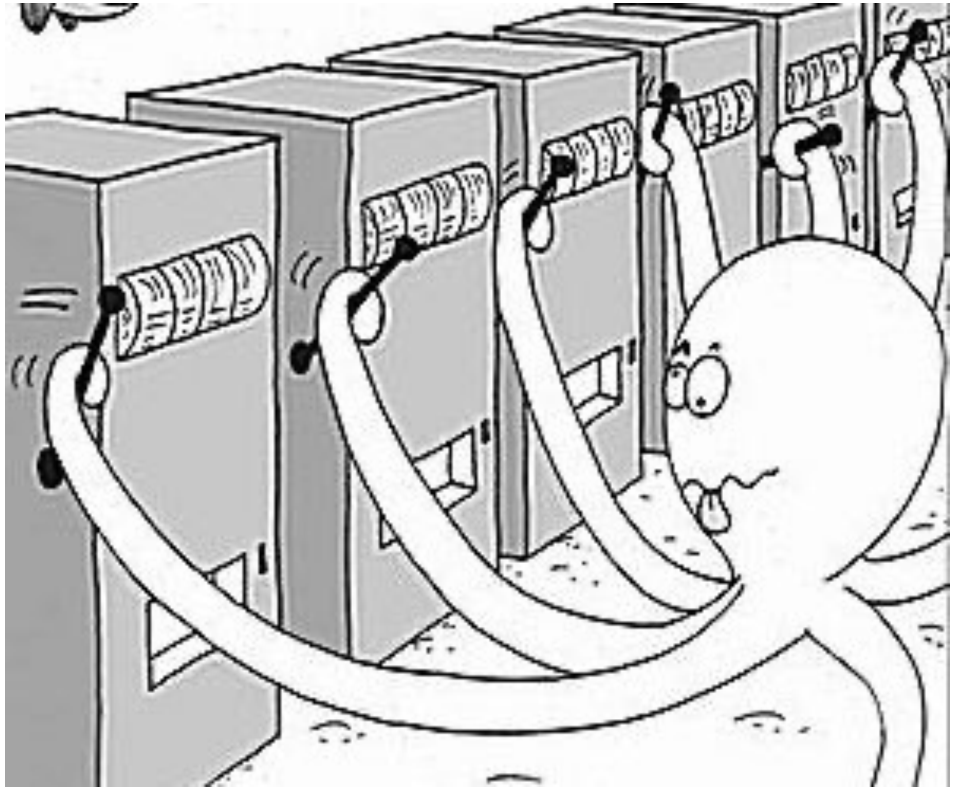
where $g(\mathbf{x}) = \sum_{x_i \in \mathbf{x}_M} (x_i - 0.5)^2, \mathbf{x} \in [0,1]^d$, and $\mathbf{x}_M$ represents the last $d - M + 1$ elements of $\mathbf{x}$. Additionally, the C2-DTLZ2 problem uses the following constraint:

$$c(\mathbf{x}) = -\min\left[\min_{i=1}^{M}\left((f_i(\mathbf{x}) - 1)^2 + \sum_{j=1, j=i}^{M}(f_j^2 - r^2)\right), \left(\sum_{i=1}^{M}\left((f_i(\mathbf{x}) - \frac{1}{\sqrt{M}})^2 - r^2\right)\right)\right] \geq 0$$

where $\mathbf{x} \in [0,1]^d$ and $r = 0.2$.

Exploration vs exploitation

- Multi-arm bandits



Schedule

Week	Date	Theory
Week 1	08/18/2023	Introduction to decision-making
Week 2	08/25/2023	Myopic planning
Week 3	09/01/2023	Myopic planning
Week 4	09/08/2023	Non-myopic planning and reinforcement learning
Week 5	09/15/2023	Non-myopic planning and reinforcement learning
Week 6	09/22/2023	Imitation learning for decision-making
Week 7	09/29/2023	Imitation learning for decision-making
Week 8	10/06/2023	Human-in-the-loop planning
Week 9	10/13/2023	Human-in-the-loop planning
Week 10	10/20/2023	Task and motion planning
Week 11	10/27/2023	Task and motion planning
Week 12	11/03/2023	Multiagent planning and decision-making under uncertainty
~~Week 13~~	~~11/10/2023~~	Veterans Day (No class)
Week 14	11/17/2023	Classical planning and hierarchical planning/learning
~~Week 15~~	~~11/24/2023~~	Thanksgiving (No class)
Week 16	12/01/2023	Project presentations

Reinforcement Learning

- No supervision/labels, just a *reward*
- You see the consequences of your current decision many steps later
- Active data gathering (typically non-iid)
- We'll learn various RL algorithms

OpenAI Gym stable-baselines

Language Models as Zero-Shot Planners: Extracting Actionable Knowledge for Embodied Agents

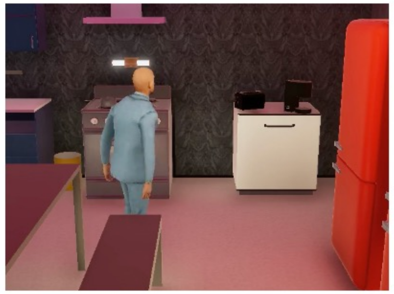
ICML 2022

Wenlong Huang[1], Pieter Abbeel[1], Deepak Pathak[2*], Igor Mordatch[3*] (*equal advising)
[1]University of California, Berkeley, [2]Carnegie Mellon University, [3]Google Brain

arXiv | PDF | Code | Video | Interview


Browse Internet


Empty Dishwasher


Turn off TV


Organize Closet


Wash Face


Take off Shoes

Large Language Models (LLMs) such as GPT-3 and Codex can plan actions for embodied agents, even without any additional training.

Website

Do As I Can, Not As I Say:
Grounding Language in Robotic Affordances

Michael Ahn* Anthony Brohan* Noah Brown* Yevgen Chebotar* Omar Cortes* Byron David* Chelsea Finn*
Chuyuan Fu* Keerthana Gopalakrishnan* Karol Hausman* Alex Herzog* Daniel Ho* Jasmine Hsu* Julian Ibarz*
Brian Ichter* Alex Irpan* Eric Jang* Rosario Jauregui Ruano* Kyle Jeffrey* Sally Jesmonth* Nikhil Joshi*
Ryan Julian* Dmitry Kalashnikov* Yuheng Kuang* Kuang-Huei Lee* Sergey Levine* Yao Lu* Linda Luu* Carolina Parada*
Peter Pastor* Jornell Quiambao* Kanishka Rao* Jarek Rettinghouse* Diego Reyes* Pierre Sermanet* Nicolas Sievers*
Clayton Tan* Alexander Toshev* Vincent Vanhoucke* Fei Xia* Ted Xiao* Peng Xu* Sichun Xu* Mengyuan Yan* Andy Zeng*

Robotics at Google Everyday Robots

* Authors listed in alphabetical order (see paper appendix for contribution statement).

Paper **Video** **Blogpost** **Code** **Demo**

Video

Foundation Models for Decision Making: Problems, Methods, and Opportunities

Sherry Yang*[1,2] Ofir Nachum[1] Yilun Du[3] Jason Wei[1] Pieter Abbeel[2] Dale Schuurmans[1,4]

[1]Google Research, Brain Team, [2]UC Berkeley, [3]MIT, [4]University of Alberta

Foundation models pretrained on diverse data at scale have demonstrated extraordinary capabilities in a wide range of vision and language tasks. When such models are deployed in real world environments, they inevitably interface with other entities and agents. For example, language models are often used to interact with human beings through dialogue, and visual perception models are used to autonomously navigate neighborhood streets. In response to these developments, new paradigms are emerging for training foundation models to interact with other agents and perform long-term reasoning. These paradigms leverage the existence of ever-larger datasets curated for multimodal, multitask, and generalist interaction. Research at the intersection of foundation models and decision making holds tremendous promise for creating powerful new systems that can interact effectively across a diverse range of applications such as dialogue, autonomous driving, healthcare, education, and robotics. In this manuscript, we examine the scope of foundation models for decision making, and provide conceptual tools and technical background for understanding the problem space and exploring new research directions. We review recent approaches that ground foundation models in practical decision making applications through a variety of methods such as prompting, conditional generative modeling, planning, optimal control, and reinforcement learning, and discuss common challenges and open problems in the field.

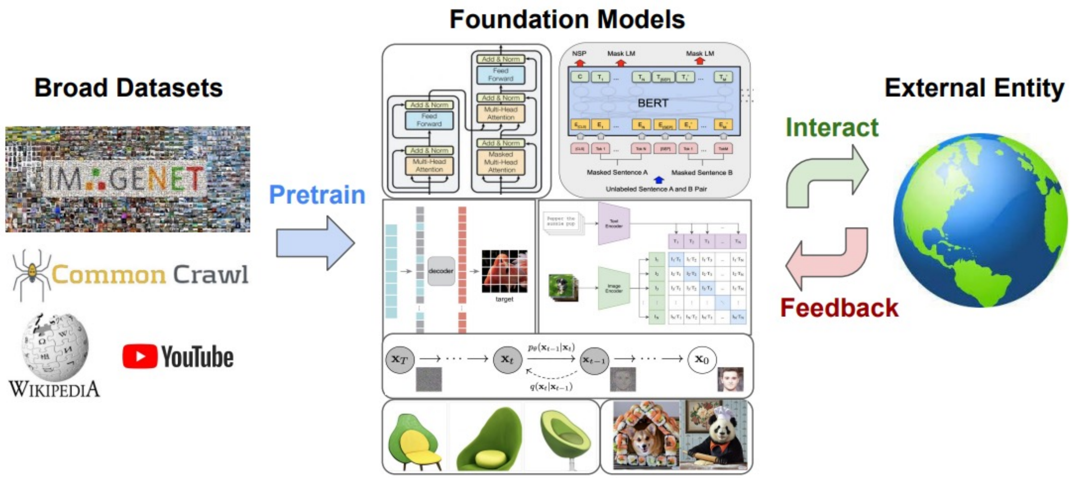


Fig. 1. Overview of foundation models for decision making. Foundation models pretrained on broad data are

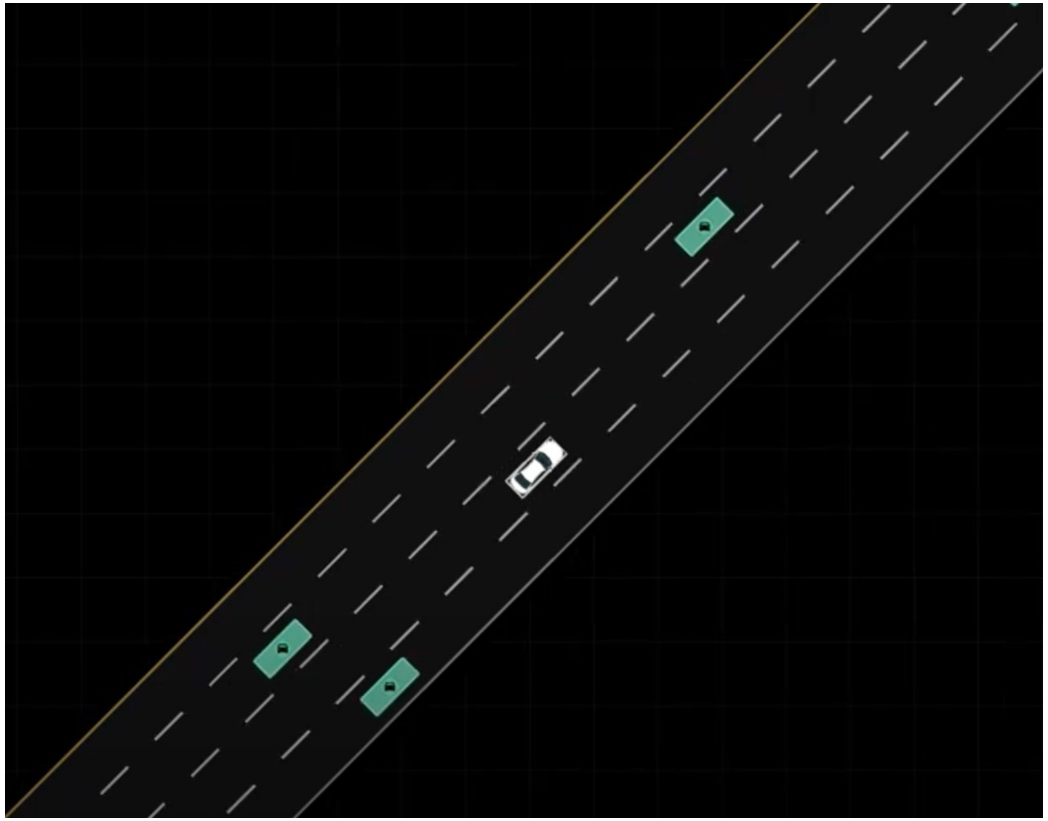
Schedule

Week	Date	Theory
Week 1	08/18/2023	Introduction to decision-making
Week 2	08/25/2023	Myopic planning
Week 3	09/01/2023	Myopic planning Robotics tutorial
Week 4	09/08/2023	Non-myopic planning and reinforcement learning
Week 5	09/15/2023	Non-myopic planning and reinforcement learning
Week 6	09/22/2023	Imitation learning for decision-making
Week 7	09/29/2023	Imitation learning for decision-making
Week 8	10/06/2023	Human-in-the-loop planning
Week 9	10/13/2023	Human-in-the-loop planning
Week 10	10/20/2023	Task and motion planning
Week 11	10/27/2023	Task and motion planning
Week 12	11/03/2023	Multiagent planning and decision-making under uncertainty
~~Week 13~~	~~11/10/2023~~	Veterans Day (No class)
Week 14	11/17/2023	Classical planning and hierarchical planning/learning
~~Week 15~~	~~11/24/2023~~	Thanksgiving (No class)
Week 16	12/01/2023	Project presentations

Schedule

Week	**Date**	**Theory**
Week 1	08/18/2023	Introduction to decision-making
Week 2	08/25/2023	Myopic planning
Week 3	09/01/2023	Myopic planning
Week 4	09/08/2023	Non-myopic planning and reinforcement learning
Week 5	09/15/2023	Non-myopic planning and reinforcement learning
Week 6	09/22/2023	Imitation learning for decision-making
Week 7	09/29/2023	Imitation learning for decision-making
Week 8	10/06/2023	Human-in-the-loop planning
Week 9	10/13/2023	Human-in-the-loop planning
Week 10	10/20/2023	Task and motion planning
Week 11	10/27/2023	Task and motion planning
Week 12	11/03/2023	Multiagent planning and decision-making under uncertainty
~~Week 13~~	~~11/10/2023~~	Veterans Day (No class)
Week 14	11/17/2023	Classical planning and hierarchical planning/learning
~~Week 15~~	~~11/24/2023~~	Thanksgiving (No class)
Week 16	12/01/2023	Project presentations

- E.g., Behavioral cloning, DAgger



Schedule

Week	Date	Theory
Week 1	08/18/2023	Introduction to decision-making
Week 2	08/25/2023	Myopic planning
Week 3	09/01/2023	Myopic planning
Week 4	09/08/2023	Non-myopic planning and reinforcement learning
Week 5	09/15/2023	Non-myopic planning and reinforcement learning
Week 6	09/22/2023	Imitation learning for decision-making
Week 7	09/29/2023	Imitation learning for decision-making
Week 8	10/06/2023	Human-in-the-loop planning
Week 9	10/13/2023	Human-in-the-loop planning
Week 10	10/20/2023	Task and motion planning
Week 11	10/27/2023	Task and motion planning
Week 12	11/03/2023	Multiagent planning and decision-making under uncertainty
~~Week 13~~	~~11/10/2023~~	Veterans Day (No class)
Week 14	11/17/2023	Classical planning and hierarchical planning/learning
~~Week 15~~	~~11/24/2023~~	Thanksgiving (No class)
Week 16	12/01/2023	Project presentations

Default (GPT-3.5)

RA

Is taking a grad level course on planning and learning methods in AI time consuming? Give me a short answer.

Yes, taking a graduate-level course on planning and learning methods in AI is typically time-consuming due to the complex nature of the subject matter and the depth of understanding required.

RL feedback

Training language models to follow instructions with human feedback

Long Ouyang* Jeff Wu* Xu Jiang* Diogo Almeida* Carroll L. Wainwright*

Pamela Mishkin* Chong Zhang Sandhini Agarwal Katarina Slama Alex Ray

John Schulman Jacob Hilton Fraser Kelton Luke Miller Maddie Simens

Amanda Askell† Peter Welinder Paul Christiano*†

Jan Leike* Ryan Lowe*

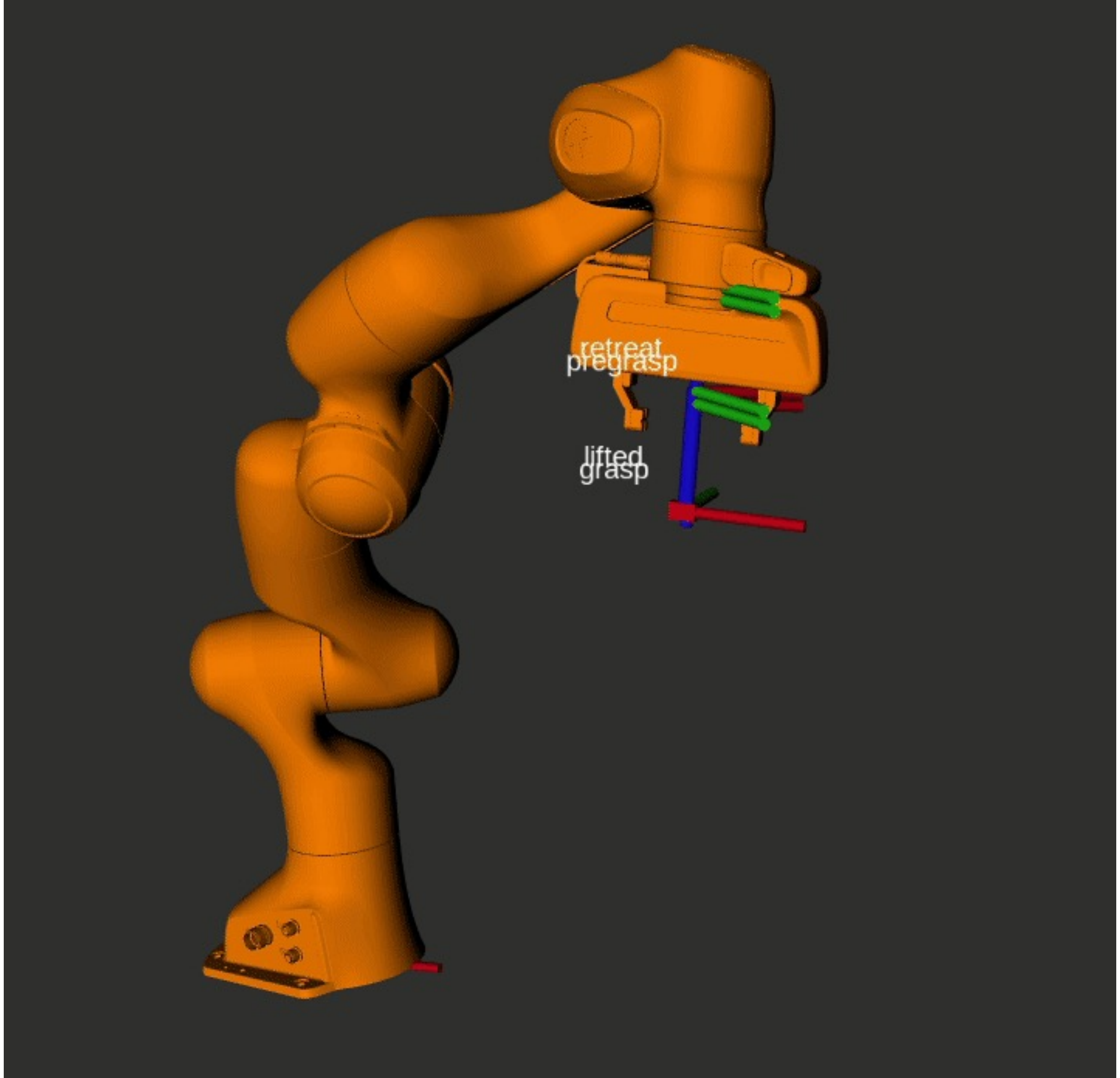
OpenAI

InstructGPT: an instruction-following LLM that uses PPO

[OpenAI Preference Demo]

Schedule

Week	Date	Theory
Week 1	08/18/2023	Introduction to decision-making
Week 2	08/25/2023	Myopic planning
Week 3	09/01/2023	Myopic planning
Week 4	09/08/2023	Non-myopic planning and reinforcement learning
Week 5	09/15/2023	Non-myopic planning and reinforcement learning
Week 6	09/22/2023	Imitation learning for decision-making
Week 7	09/29/2023	Imitation learning for decision-making
Week 8	10/06/2023	Human-in-the-loop planning
Week 9	10/13/2023	Human-in-the-loop planning
Week 10	10/20/2023	Task and motion planning
Week 11	10/27/2023	Task and motion planning
Week 12	11/03/2023	Multiagent planning and decision-making under uncertainty
~~Week 13~~	~~11/10/2023~~	Veterans Day (No class)
Week 14	11/17/2023	Classical planning and hierarchical planning/learning
~~Week 15~~	~~11/24/2023~~	Thanksgiving (No class)
Week 16	12/01/2023	Project presentations


retreat
pregrasp
lifted
grasp

Schedule

Week	Date	Theory
Week 1	08/18/2023	Introduction to decision-making
Week 2	08/25/2023	Myopic planning
Week 3	09/01/2023	Myopic planning
Week 4	09/08/2023	Non-myopic planning and reinforcement learning
Week 5	09/15/2023	Non-myopic planning and reinforcement learning
Week 6	09/22/2023	Imitation learning for decision-making
Week 7	09/29/2023	Imitation learning for decision-making
Week 8	10/06/2023	Human-in-the-loop planning
Week 9	10/13/2023	Human-in-the-loop planning
Week 10	10/20/2023	Task and motion planning
Week 11	10/27/2023	Task and motion planning
Week 12	11/03/2023	Multiagent planning and decision-making under uncertainty
~~Week 13~~	~~11/10/2023~~	Veterans Day (No class)
Week 14	11/17/2023	Classical planning and hierarchical planning/learning
~~Week 15~~	~~11/24/2023~~	Thanksgiving (No class)
Week 16	12/01/2023	Project presentations