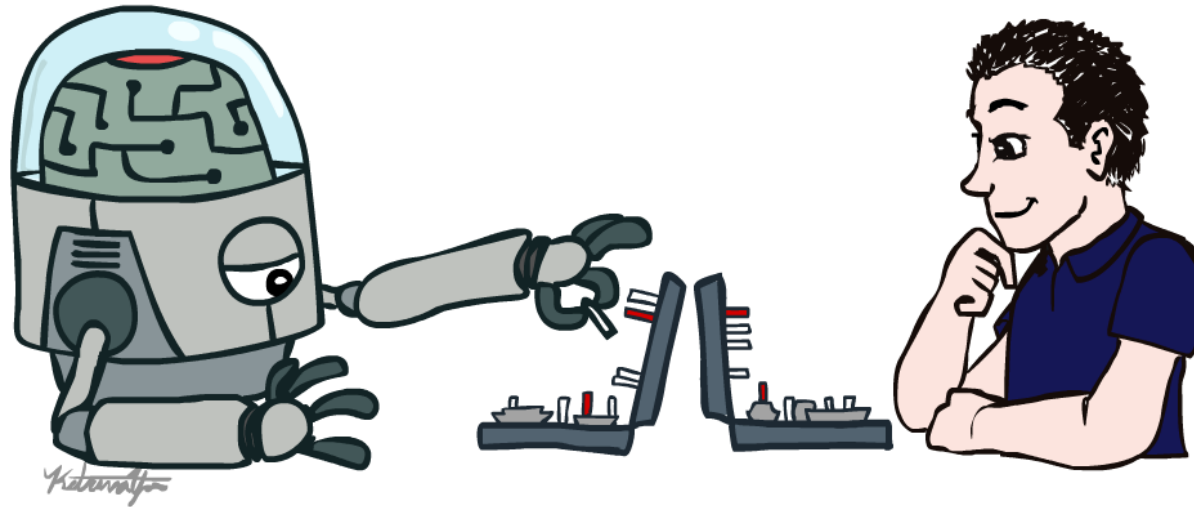


CS 188: Artificial Intelligence

Introduction



Spring 2025

University of California, Berkeley

[These slides were created by Dan Klein and Pieter Abbeel for CS188 Intro to AI at UC Berkeley (ai.berkeley.edu).]

First Half of Today: Intros and Logistics

- Staff introductions: John Canny, Oliver Grillmeyer, and course staff
- Course logistics
 - Lectures, discussions, office hours, and exams
 - Resources and communication platforms
 - Collaboration and academic honesty
 - DSP and extenuating circumstances
 - Stress management and mental health

Staff Introductions: John (he/him)

- Professor at UC Berkeley since 1987
- 1987, Ph.D., Electrical Engineering, MIT
- 1983, M.S., Electrical Engineering, MIT
- 1980, B.E. (Hons), Electrical Engineering, Adelaide University
- 1979, B.Sc., Computer Science and Theoretical Physics, Adelaide University
- Research areas:
 - Artificial Intelligence
 - Control, Intelligent Systems, and Robotics
 - Graphics
 - Human Computer Interaction
 - Security



Staff Introductions: Oliver (he/him)

- Did BA, MS, and Ph.D. at UC Berkeley
 - Masters: Multimedia-based CS textbook
 - Doctorate: Using Animations to Teach Computer Science
- Teaching experience at UC Berkeley and USF
 - Berkeley: intro classes in Pascal, LISP, Scheme, and 61A
 - USF: Intro courses, Project courses, Databases, Grad AI
- Work Experience
 - Expert Systems; CTO for education start-up; Enterprise website development; Disaster management; Deep learning for satellite detection and document classification
 - Consultant for electrochromic window control and HotMath
- Side Hustles
 - Wrote CS textbook; Home remodeling; Real estate investor
- For fun
 - Martial arts; Photography; Travel; Beer brewing; Sailing; Skiing



Our talented course staff!



Noah Adhikari
he/him



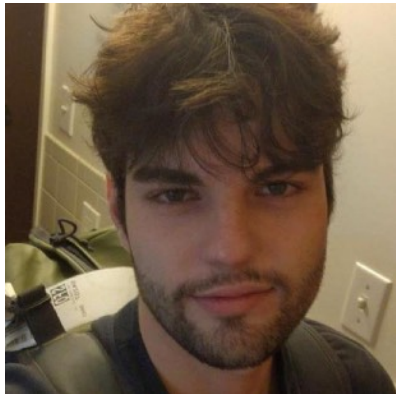
Pranav Muralikrishnan
he/him



Xavier Yin
he/him



Catherine Chu
she/her



Arnaud Fickinger
he/him



Joshua Liao
he/him



Mustafa Mirza
he/him



Suhong Moon
he/him

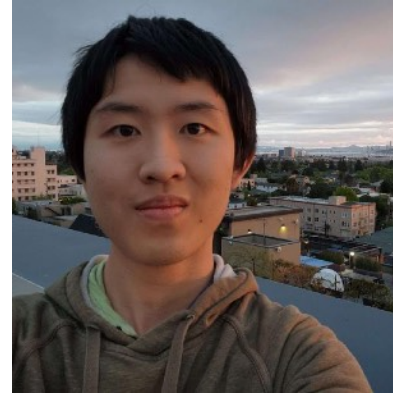
Our talented course staff!



Saathvik Selvan
he/him



Daniel Toktarbayev
he/him



Kenny Wang
he/him



Michael Wu
he/him



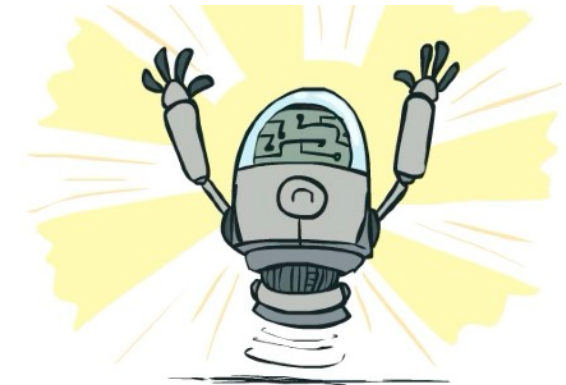
Wesley Zheng
he/him



Advika Bhike
she/her



Andrew Choy
he/him



Gradescope
she/her

Enrollment

- Course staff does not control enrollment; we have to follow department policy
 - Only CS and Data Science majors will be able to enroll this spring
 - More details on the course website
 - Please do not email us if you are a concurrent enrollment student with a pending application; you will be added automatically within 3-4 days of submitting your application.

Course Structure: Lectures

- You are here!
- Tuesday/Thursday, 12:30–2:00 PM PT
- Attendance is not taken, but highly encouraged
- You can attend:
 - In-person in 155 Dwinelle
 - Remotely over Zoom
 - Asynchronously by watching recordings (posted next day on website)

Course Structure: Discussions

- We offer two types of discussions
 - Regular discussions
 - Exam prep discussions
 - See course website for breakdown
 - We'll try to make recordings, but no promises
- Discussion schedule available on website
 - Discussions start next week (January 27)
- You can attend any discussion section you want (no need to enroll in a section)

Course Structure: Office Hours

- Join in-person or remotely to talk to staff about content, ask questions on assignments, or raise any concerns you have
- Schedule and queue available on website
 - Office hours start next week (January 27)

Course Structure: Exams

- Save the dates!
 - Midterm: Wednesday, March 19, 7:00–9:00 pm PT
 - Final exam: Thursday, May 15, 3:00–6:00 pm PT
- If you have a time conflict:
 - We'll offer an in-person-only alternate exam right after the listed time
- More logistics closer to the exam

Resources

- Course website: <https://inst.eecs.berkeley.edu/~cs188/sp25/>
 - All resources (slides, notes, recordings, assignments, etc.) posted here
- Ed: Discussion forum
- Staff email for private concerns: cs188@berkeley.edu
 - Making a private post on Ed is easier/faster
- Gradescope: Submit assignments here

Grading Structure

- **Projects (25%)**
 - Python programming assignments, autograded
 - You can optionally work with a partner
 - No late submissions, unless you have an extension
- **Homeworks (20%)**
 - Electronic homework: Autograded on Gradescope
 - Written homework: One exam-style question per week, with self-assessment due the following week. Graded on completion and quality of self-assessment.
 - Submit individually (cite your collaborators if you discuss with others)
 - No late submissions, unless you have an extension
- **Midterm (20%), Final Exam (35%)**

Extensions and Accommodations

- We'll drop your lowest homework score
- If you ever need an extension, please request one!
 - We're here to support you, and we understand that life happens.
 - Extension form will be posted on the website

DSP

- Disabled Students' Program (DSP)
 - There's a variety of accommodations UC Berkeley can help us set up for you in this class
 - <https://dsp.berkeley.edu/>
- Are you facing barriers in school due to a disability?
 - Apply to DSP!
 - We maintain proper access controls on this information: Only instructors, course managers, head TAs, and logistics TAs can access any DSP-related info
- Our goal is to teach you the material in our course. The more accessible we can make it, the better.

Collaboration and Academic Dishonesty

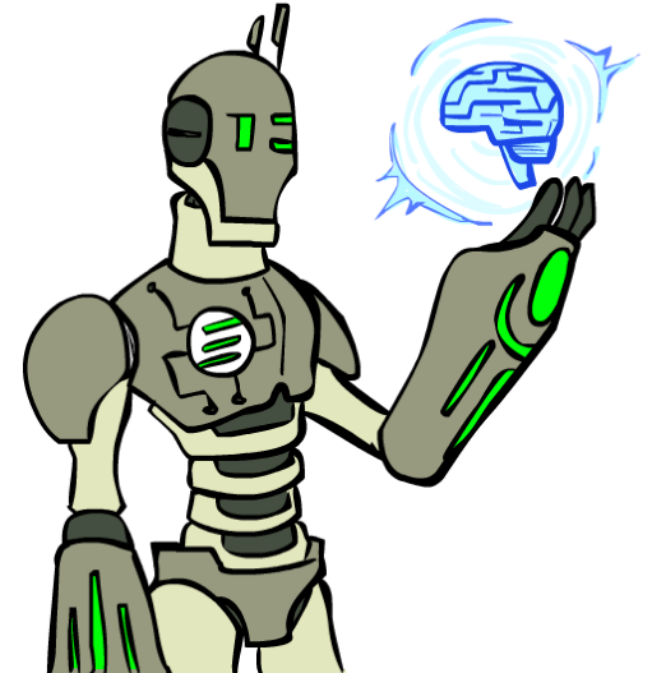
- We're here to help! There are plenty of staff and resources available for you
 - You can always talk to a staff member if you're feeling stressed or tempted to cheat
 - Collaboration on homework is okay, but please cite collaborators
 - Do not post solutions online or share with others!
- Academic dishonesty policies
 - Reported to Center of Student Conduct
 - Negative points on assignments, and/or F in the class

Stress Management and Mental Health

- **Your health is more important than this course**
- **If you feel overwhelmed, there are options**
 - Academically: Ask on Ed, talk to staff in office hours, set up a meeting with staff to make a plan for your success this semester
 - Non-academic:
 - Counselling and Psychological Services (CAPS) has multiple free, confidential services
 - Casual consultations: <https://uhs.berkeley.edu/counseling/lets-talk>
 - Crisis management: <https://uhs.berkeley.edu/counseling/urgent>
 - Check out UHS's resources: <https://uhs.berkeley.edu/health-topics/mental-health>

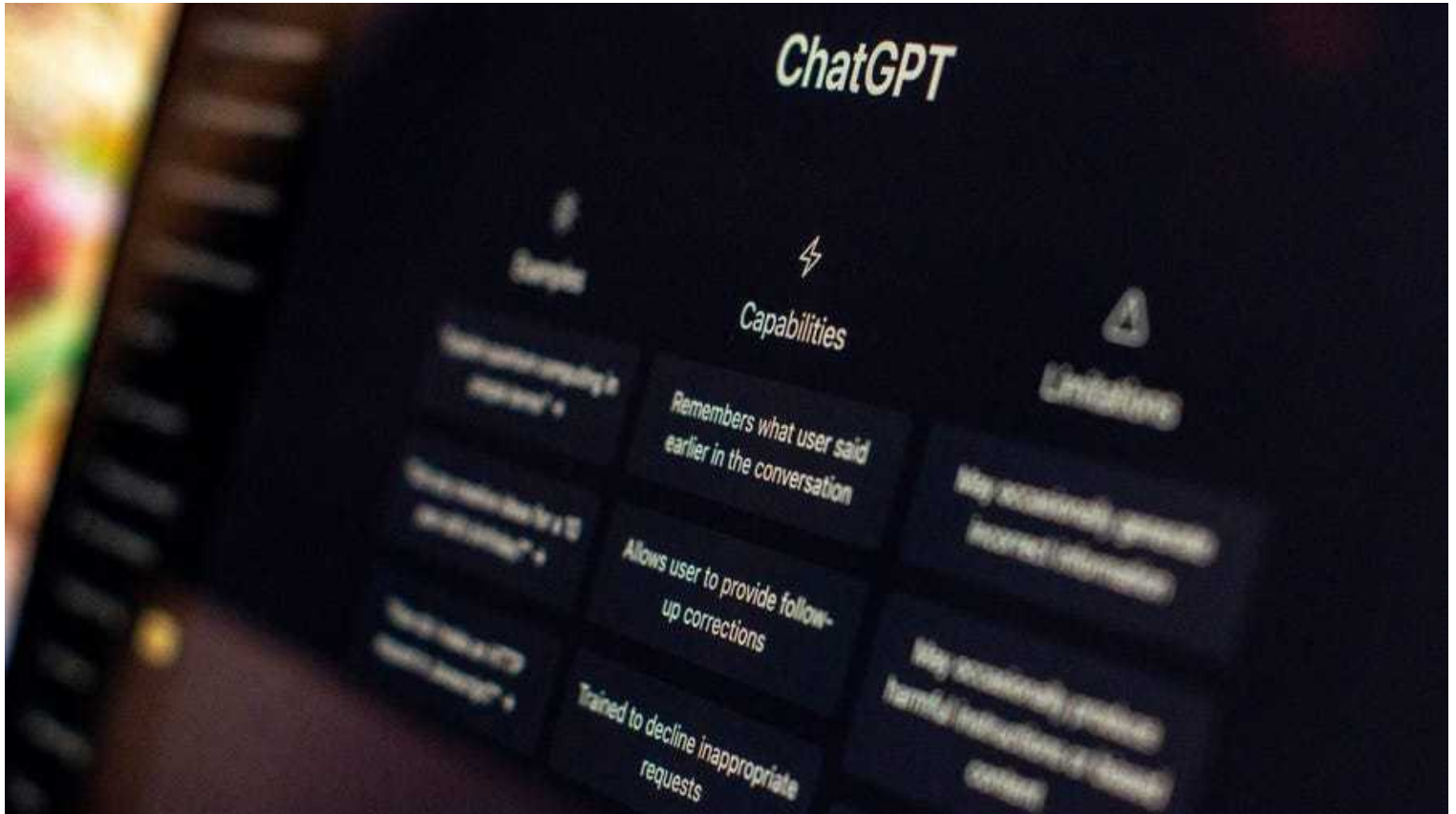
Second Half of Today: What is AI?

- What is artificial intelligence?
- What can AI do?
- What is this course?



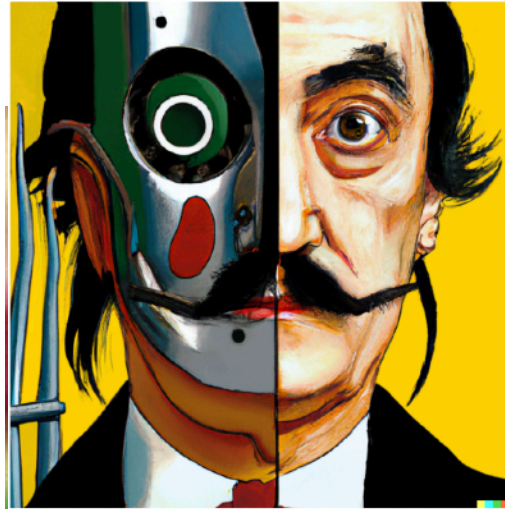
AI is having real-world impact

- Public imagination
 - Text assistants



AI is having real-world impact

- Public imagination
 - Text assistants
 - Image generation



vibrant portrait painting of Salvador Dalí with a robotic half face



a shiba inu wearing a beret and black turtleneck



a close up of a handpalm with leaves growing from it



an espresso machine that makes coffee from human souls, artstation



panda mad scientist mixing sparkling chemicals, artstation



a corgi's head depicted as an explosion of a nebula

AI is having real-world impact



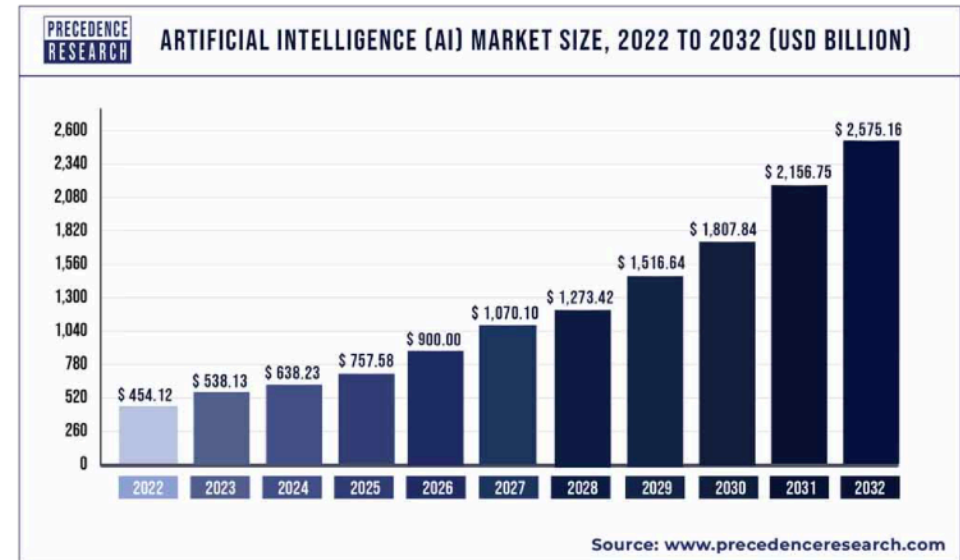
AI is having real-world impact



AI is having real-world impact

- Public imagination
- Economy
 - 454 billion USD globally

The global artificial intelligence (AI) market size was valued at USD 454.12 billion in 2022 and is expected to hit around USD 2,575.16 billion by 2032, progressing with a CAGR of 19% from 2023 to 2032. The North America artificial intelligence market was valued at USD 167.30 billion in 2022.



<https://www.precedenceresearch.com/artificial-intelligence-market>

AI is having real-world impact

- Public imagination
- Economy
- Politics



AI is having real-world impact

- Public imagination
- Economy
- Politics



AI is having real-world impact

- Public imagination
- Economy
- Politics



AI is having real-world impact

- Public imagination
- Economy
- Politics



AI is having real-world impact

- Public imagination
- Economy
- Politics
- Law

Aug. 18, 2023, 12:18 PM; Updated: Aug. 18, 2023, 12:48 PM

AI-Generated Art Lacks Copyright Protection, D.C. Court Says (1)



Riddhi Setty
Reporter



Isaiah Poritz
Legal Reporter



Bloomberg Law, 2023

AI is having real-world impact

- Public imagination
- Economy
- Politics
- Law
- Labor

Finance & economics | Free exchange

New research shows the robots are coming for jobs—but stealthily

Look beneath the aggregate economic numbers, and change is afoot

The Economist, 2021

The Optimist's Guide to Artificial Intelligence and Work

The focus of much discussion is on how it will replace jobs, but nothing is inevitable.

New York Times, 2023

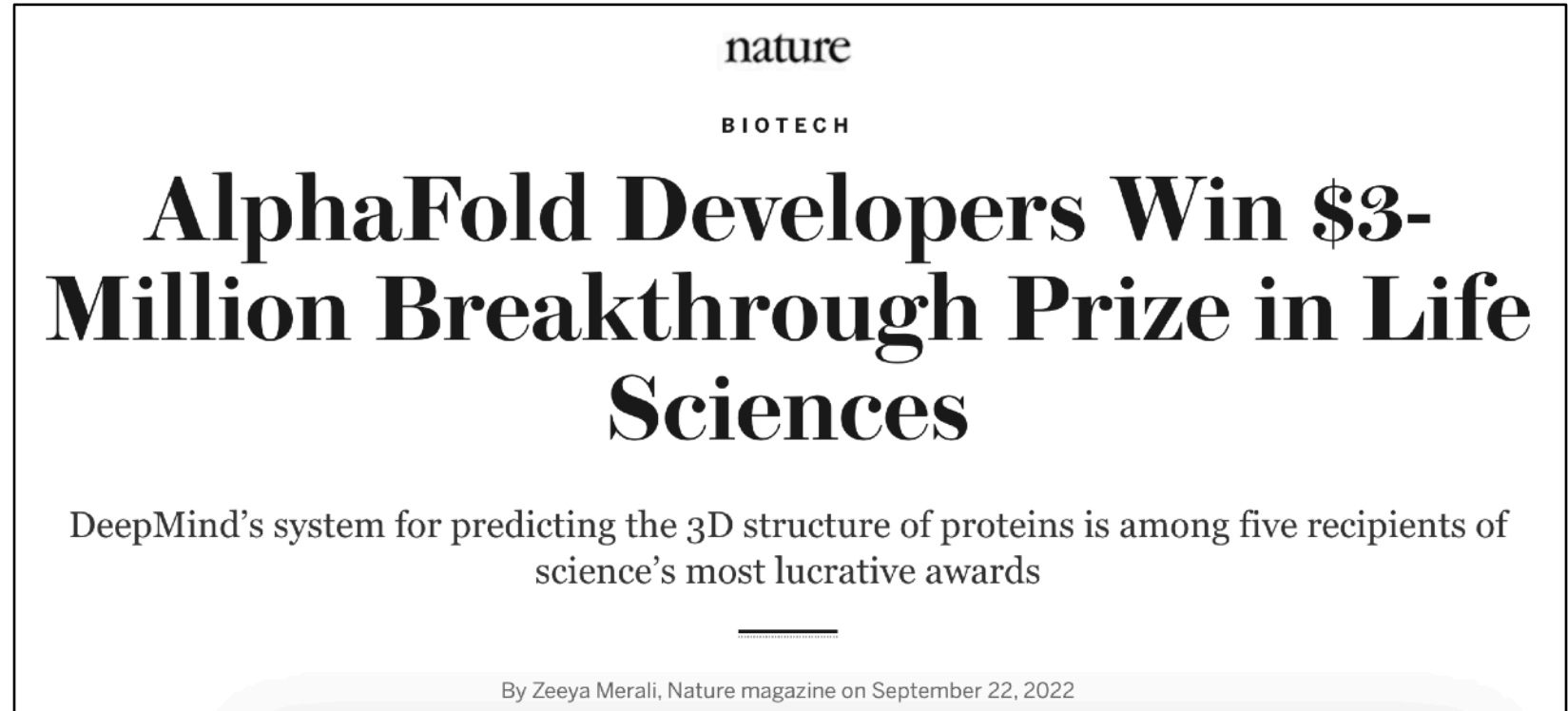
The human labor behind AI chatbots and other smart tools

Data labeling is an important step in developing artificial intelligence but also exposes the people doing the work to harmful content.

MarketWatch, 2023

AI is having real-world impact

- Public imagination
- Economy
- Politics
- Law
- Labor
- Sciences



Nature, 2022

AI is having real-world impact

- Public imagination
- Economy
- Politics
- Law
- Labor
- Sciences



Wired, 2022

AI is having real-world impact

- Public imagination
- Economy
- Politics
- Law
- Labor
- Sciences
- Education

BREAKING

ChatGPT In Schools: Here's Where It's Banned—And How It Could Potentially Help Students

Arianna Johnson Forbes Staff

I cover the latest trends in science, tech and healthcare.

Follow

 2

Jan 18, 2023, 02:31pm EST

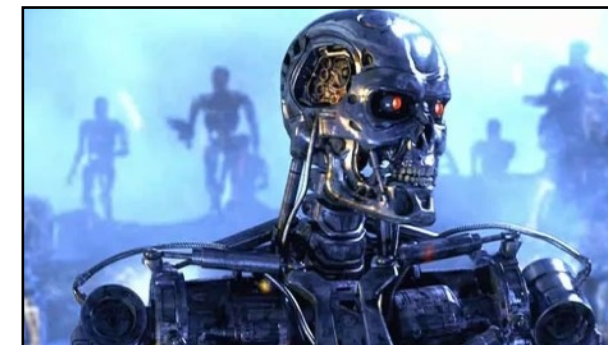
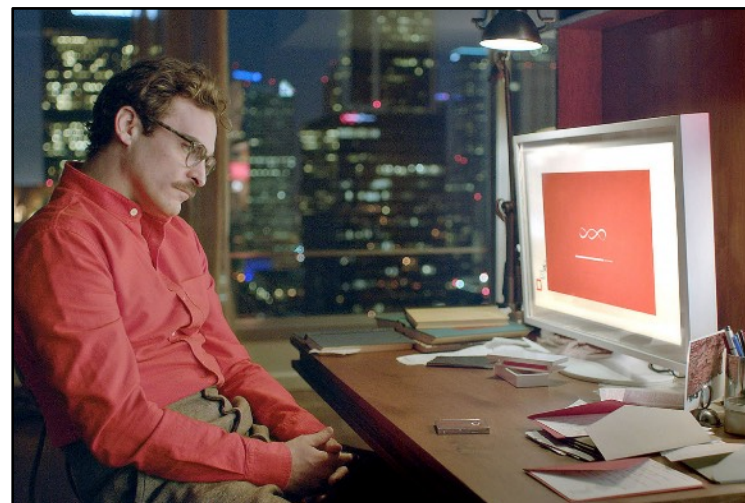
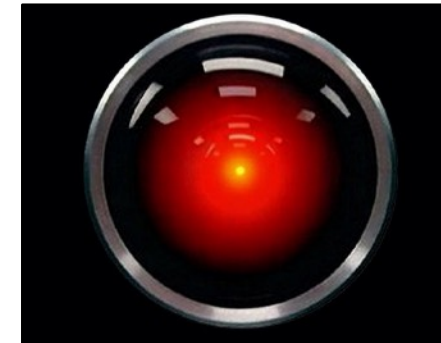
Forbes, 2023

AI is having real-world impact

- Public imagination
- Economy
- Politics
- Law
- Labor
- Sciences
- Education

Ok, but what does *AI do*???

Science fiction AI?



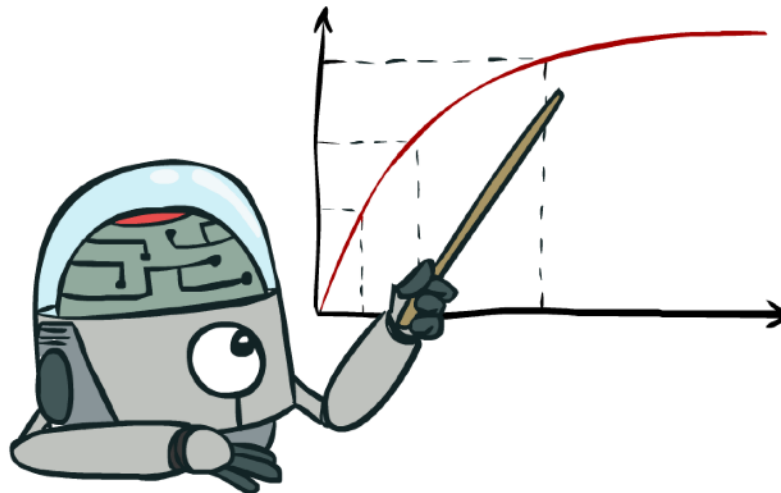
What is AI?

The science of making machines that:



Rational Decisions

- We'll use the term **rational** in a very specific, technical way:
 - Rational: *maximally achieving pre-defined goals*
 - Goals are expressed in terms of the **utility** of outcomes
 - World is uncertain, so we'll use **expected** utility
 - Being rational means acting to **maximize your expected utility**



Rational Decisions

- We'll use the term **rational** in a very specific, technical way:
 - Rational: *maximally achieving pre-defined goals*
 - Goals are expressed in terms of the **utility** of outcomes
 - World is uncertain, so we'll use **expected** utility
 - Being rational means acting to **maximize your expected utility**

A better title for this course would be:

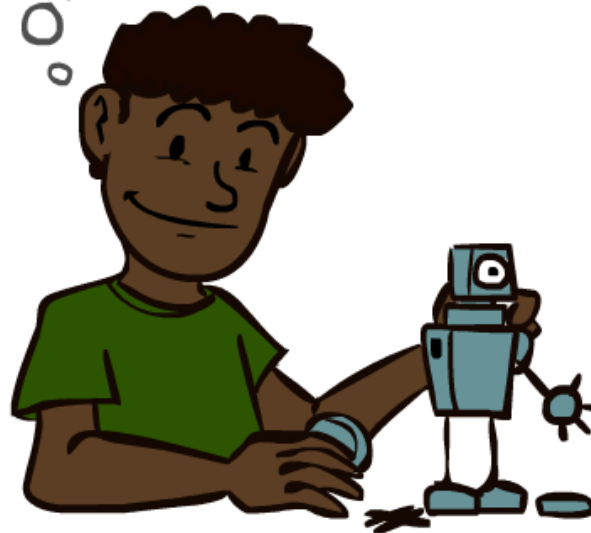
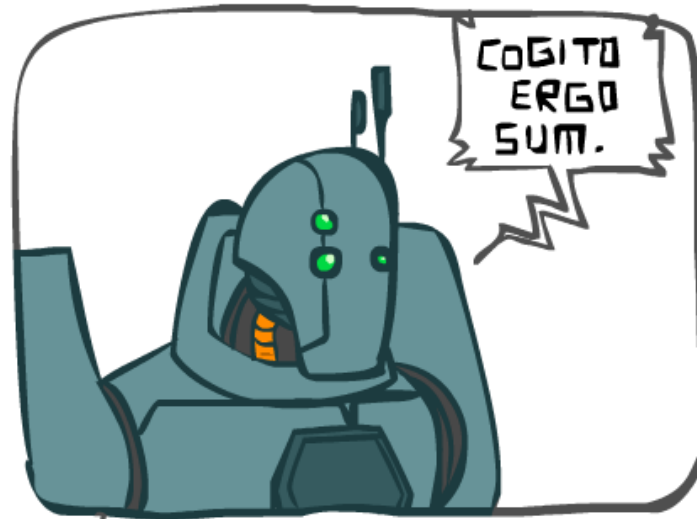
Computational Rationality

What About the Brain?

- Brains (human minds) are very good at making rational decisions, but not perfect
- Brains aren't as modular as software, so hard to reverse engineer!
- AI may be better than brains at some tasks
- *“Brains are to intelligence as wings are to flight”*
- We can't yet build AI on the scale of the brain
 - ~100T synapses in the human brain vs ~1.8T weights in ChatGPT4
- Still, the brain can be a great inspiration for AI!



A (Short) History of AI



A (Short) History of AI

- 1940-1950: Early days: neural and computer science meet
 - 1943: McCulloch & Pitts: Boolean circuit model of brain
 - 1950: Turing's "Computing Machinery and Intelligence"
- 1950—70: Excitement! Logic-driven
 - 1950s: Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
 - 1956: Dartmouth meeting: "Artificial Intelligence" adopted
 - 1957: Newell, Shaw, Simon GPS, General Problem Solver
 - 1965: Robinson's complete algorithm for logical reasoning



A (Short) History of AI

- 1970—90: Knowledge-based approaches
 - 1969—79: Early development of knowledge-based systems
 - 1980—88: Expert systems industry booms
 - 1988—93: Expert systems industry busts: “AI Winter”
- 1990—2010: Statistical approaches
 - Resurgence of probability, focus on uncertainty
 - Agents and learning systems... “AI Spring”?
 - 1992: TD-Gammon attains human-level performance
 - 1996: Kasparov defeats Deep Blue at chess
 - 1997: Deep Blue defeats Kasparov at chess
 - 2002: Embodied AI; Roomba vacuum invented



“I could feel --- I could smell ---
a new kind of intelligence
across the table.” ~Kasparov

A (Short) History of AI

- 2010—2017: Big Data, GPUs, Deep Learning
 - 2011: Apple releases SIRI
 - 2012: AlexNet (neural net) wins ImageNet (image recognition) competition
 - 2015: DeepMind achieves human-level control in Atari games
 - 2016: DeepMind's AlphaGo beat Go Master, Lee Sedol
 - 2016: Google Translate migrates to neural networks
- 2017—: Scaling up, Large Language Models
 - 2017: Google invents Transformer architecture
 - 2017: DeepStack and Libratus defeat humans at poker
 - 2018-2020: AlphaFold predicts protein structure from amino acids
 - 2021-2022: Modern text-to-image generation
 - 2022: OpenAI release ChatGPT

A (Short) History of AI

- Notable Failure 1 - Natural Language Processing Efforts

- Early work during cold war to translate Russian to English

- Success early on followed by reality of serious limitations

Out of sight, out of mind

The spirit is willing, but the flesh is weak

Invisible, Insanity

The Vodka is strong, but the meat is rotten

- Problem: Even simple sentences require world knowledge

Cinderella attended the ball

The soldiers shot the prisoners and I saw several fall

Cinderella kicked the ball

The soldiers shot the prisoners and I saw several reload

The house was built by the workman

The pen is in the box

The house was built by the river

The box is in the pen

- Attempted Solution: Capture and embody world knowledge

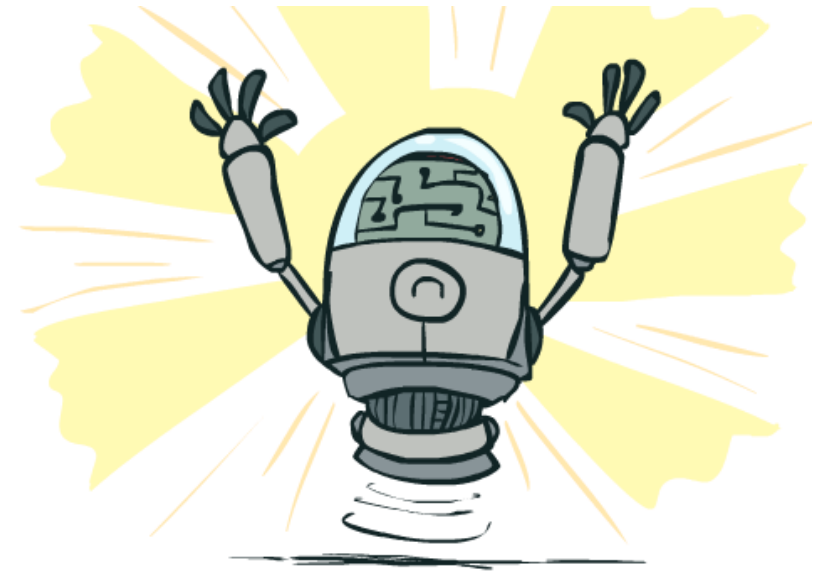
A (Short) History of AI

- Notable Failure 2 - Perceptrons
 - Rosenblatt's two layer neural nets
 - Belief was that they could learn logic operations
 - Minsky and Papert showed that they couldn't even solve XOR
 - No mathematical way to reflect that one input's contribution is greater than that of two: $X \not> (X+Y)$
 - Solution: Add hidden layer between input layer and output layer

What Can AI Do?

Quiz: Which of the following can be done at present?

- ✓ Win against any human at chess?
- ✓ Win against the best humans at Go?
- Play a decent game of tennis?
- Unload any dishwasher in any home?
- Drive safely along the highway?
- Drive safely along streets of San Francisco?
- Buy a week's worth of groceries on the web?
- Buy a week's worth of groceries at Berkeley Bowl?
- Discover and prove a new mathematical theorem?
- Perform a surgical operation?
- Translate spoken Chinese into spoken English in real time?
- Win an art competition?
- Write an intentionally funny story?
- Construct a building?



What Can AI Do?

Quiz: Which of the following can be done at present?

- ✓ Win against any human at chess?
- ✓ Win against the best humans at Go?
- Play a decent game of tennis?
- Unload any dishwasher in any home?
- Drive safely along the highway?
- Drive safely along streets of San Francisco?
- Buy a week's worth of groceries on the web?
- Buy a week's worth of groceries at Berkeley Bowl?
- Discover and prove a new mathematical theorem?
- Perform a surgical operation?
- Translate spoken Chinese into spoken English in real time?
- Win an art competition?
- Write an intentionally funny story?
- Construct a building?

Google's AlphaGo beats Lee Sedol at Go in 2016



But ... a plot twist in 2023!

DEEP BLUE WAS JUST THE START —

Man beats machine at Go in human victory over AI

Amateur exploited weakness in systems that have otherwise dominated grandmasters.

RICHARD WATERS, FINANCIAL TIMES - 2/19/2023, 4:51 AM

What Can AI Do?

Quiz: Which of the following can be done at present?

- ✓ Win against any human at chess?
- ✓ Win against the best humans at Go?
- ✓ Play a decent game of tennis?
- Unload any dishwasher in any home?
- Drive safely along the highway?
- Drive safely along streets of San Francisco?
- Buy a week's worth of groceries on the web?
- Buy a week's worth of groceries at Berkeley Bowl?
- Discover and prove a new mathematical theorem?
- Perform a surgical operation?
- Translate spoken Chinese into spoken English in real time?
- Win an art competition?
- Write an intentionally funny story?
- Construct a building?



What Can AI Do?

Quiz: Which of the following can be done at present?

- ✓ Win against any human at chess?
- ✓ Win against the best humans at Go?
- ✓ Play a decent game of tennis?
- ✗ Unload any dishwasher in any home?
- ✓ Drive safely along the highway?
- ➔ Drive safely along streets of San Francisco?
- Buy a week's worth of groceries on the web?
- Buy a week's worth of groceries at Berkeley Bowl?
- Discover and prove a new mathematical theorem?
- Perform a surgical operation?
- Translate spoken Chinese into spoken English in real time?
- Win an art competition?
- Write an intentionally funny story?
- Construct a building?

California just opened the floodgates for self-driving cars

After a state regulatory board ruling, San Francisco will have 24/7 robotaxis on its streets. A Washington Post analysis shows how it's a pivotal moment for the industry.

By [Trisha Thadani](#) and [Jeremy B. Merrill](#)

Updated August 10, 2023 at 10:06 p.m. EDT | Published August 10, 2023 at 9:15 p.m. EDT

The Register

California DMV hits brakes on Cruise's SF driverless fleet after series of fender benders

50% chop effective 'immediately' as department investigates traffic and safety issues

[Jude Karabus](#)

Mon 21 Aug 2023 / 14:28 UTC

What Can AI Do?

Quiz: Which of the following can be done at present?

- ✓ Win against any human at chess?
- ✓ Win against the best humans at Go?
- ✓ Play a decent game of tennis?
- ✗ Unload any dishwasher in any home?
- ✓ Drive safely along the highway?
- ? Drive safely along streets of San Francisco?
- ✓ Buy a week's worth of groceries on the web?
- ✗ Buy a week's worth of groceries at Berkeley Bowl?
- ? Discover and prove a new mathematical theorem?
- ✗ Perform a surgical operation?
- ✓ Translate spoken Chinese into spoken English in real time
- ✓ Win an art competition?
Write an intentionally funny story?
Construct a building?



What Can AI Do?

Quiz: Which of the following can be done at present?

- ✓ Win against any human at chess?
- ✓ Win against the best humans at Go?
- ✓ Play a decent game of tennis?
- ✗ Unload any dishwasher in any home?
- ✓ Drive safely along the highway?
- ? Drive safely along streets of San Francisco?
- ✓ Buy a week's worth of groceries on the web?
- ✗ Buy a week's worth of groceries at Berkeley Bowl?
- ? Discover and prove a new mathematical theorem?
- ✗ Perform a surgical operation?
- ✓ Translate spoken Chinese into spoken English in real time?
- ✓ Win an art competition?
- ✓ Write an intentionally funny story?
- Construct a building?

Prompt: Write a joke about an early space program that had used monkeys as astronauts for many missions, and then sent the first human astronaut into space along with a monkey.

In an early space program, after successfully sending monkeys on several missions, they finally decided to send a human astronaut along with a monkey. As the rocket was about to launch, the human nervously asked mission control, "So, what are my instructions?"

Mission control replied, "Relax, we gave the monkey the important stuff to do. You just feed him and make sure he's comfortable!"

What Can AI Do?

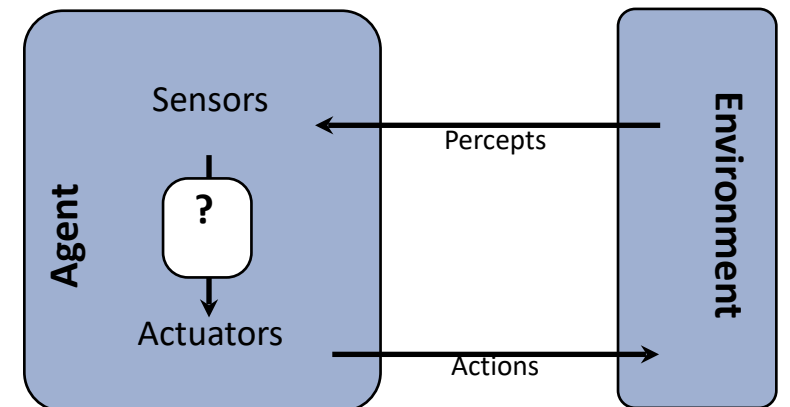
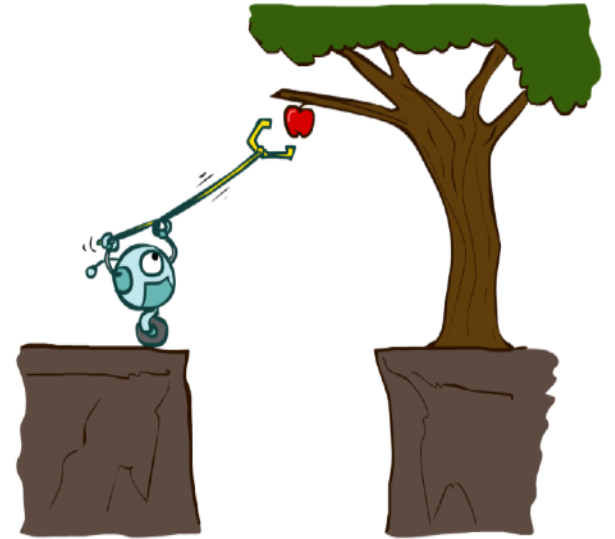
Quiz: Which of the following can be done at present?

- ✓ Win against any human at chess?
- ✓ Win against the best humans at Go?
- ✓ Play a decent game of tennis?
- ✗ Unload any dishwasher in any home?
- ✓ Drive safely along the highway?
- ? Drive safely along streets of San Francisco?
- ✓ Buy a week's worth of groceries on the web?
- ✗ Buy a week's worth of groceries at Berkeley Bowl?
- ? Discover and prove a new mathematical theorem?
- ✗ Perform a surgical operation?
- ✓ Translate spoken Chinese into spoken English in real time?
- ✓ Win an art competition?
- ✓ Write an intentionally funny story?
- ✗ Construct a building?

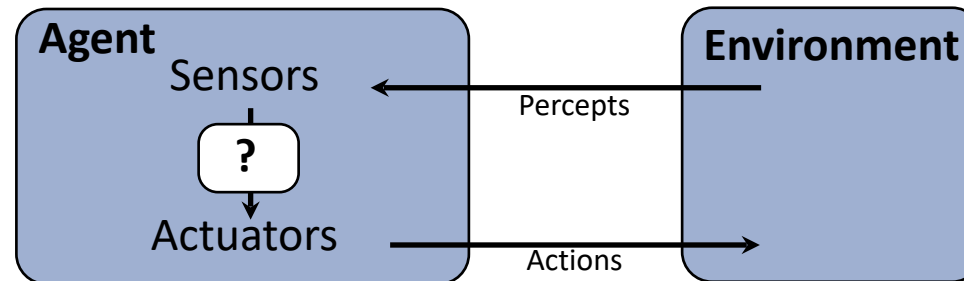
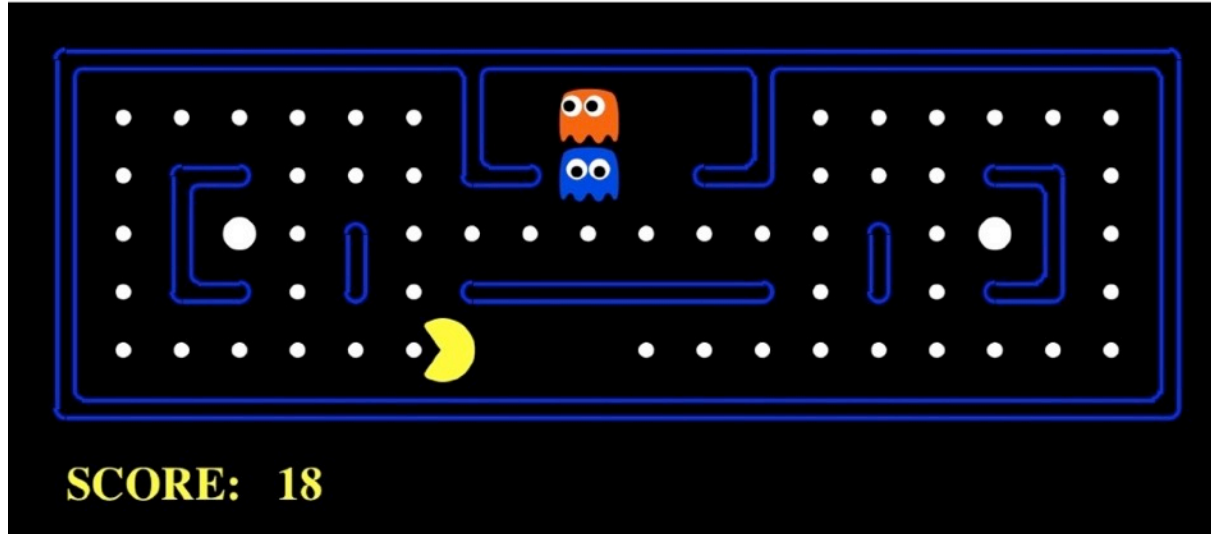


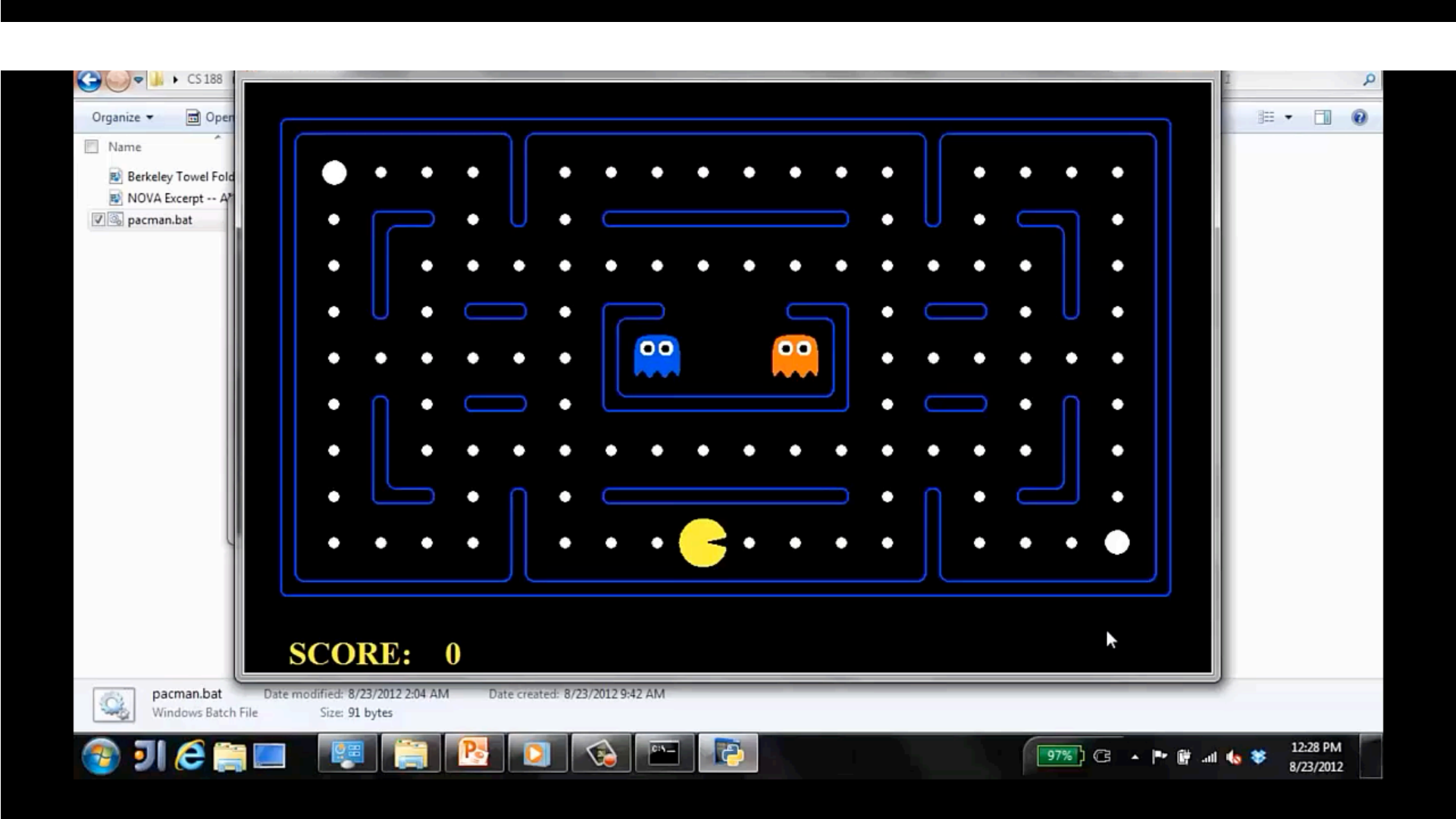
This Course: Designing Rational Agents

- An **agent** is an entity that perceives and acts.
- A **rational agent** selects actions that maximize its (expected) **utility**.
- Characteristics of the **percepts**, **environment**, and **action space** dictate techniques for selecting rational actions
- This course is about:
 - General AI techniques for a variety of problem types
 - Learning to recognize when and how a new problem can be solved with an existing technique



Pac-Man as an Agent





Course Topics

Core Components of Rational Agents:

Search &
Planning

Reinforcement
Learning

Probability &
Inference

Supervised
Learning

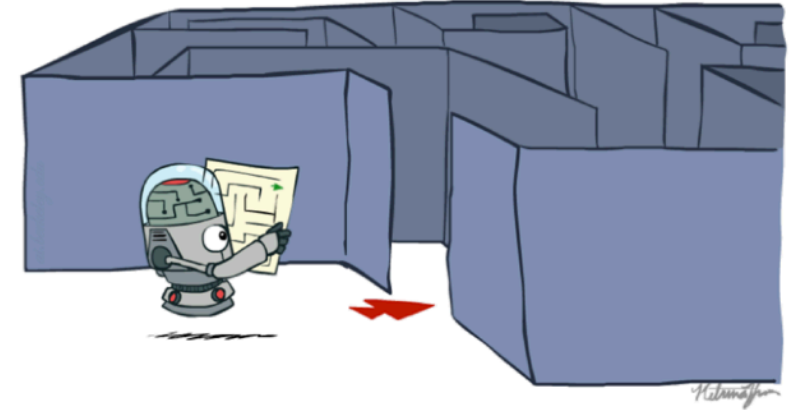
Course Topics

Search &
Planning

Reinforcement
Learning

Probability &
Inference

Supervised
Learning



How can I find a ***sequence of best decisions*** for a ***particular*** situation?

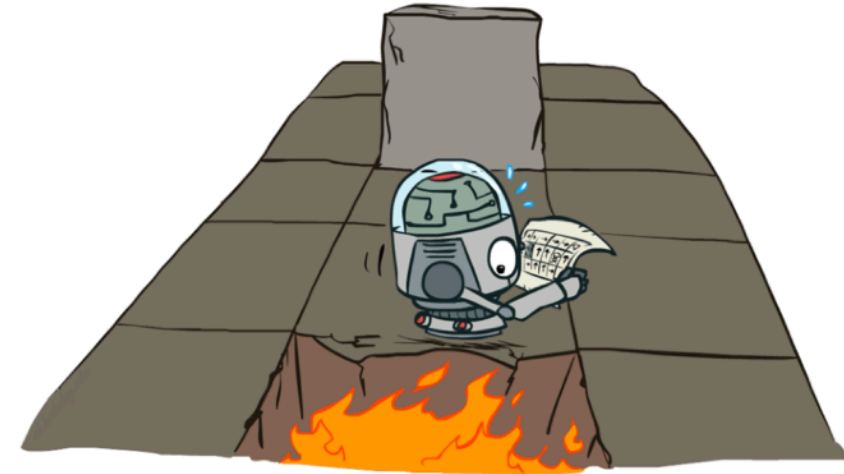
Course Topics

Search &
Planning

Reinforcement
Learning

Probability &
Inference

Supervised
Learning



How can I find **rules (policy)** to make best decisions for **any** situation?

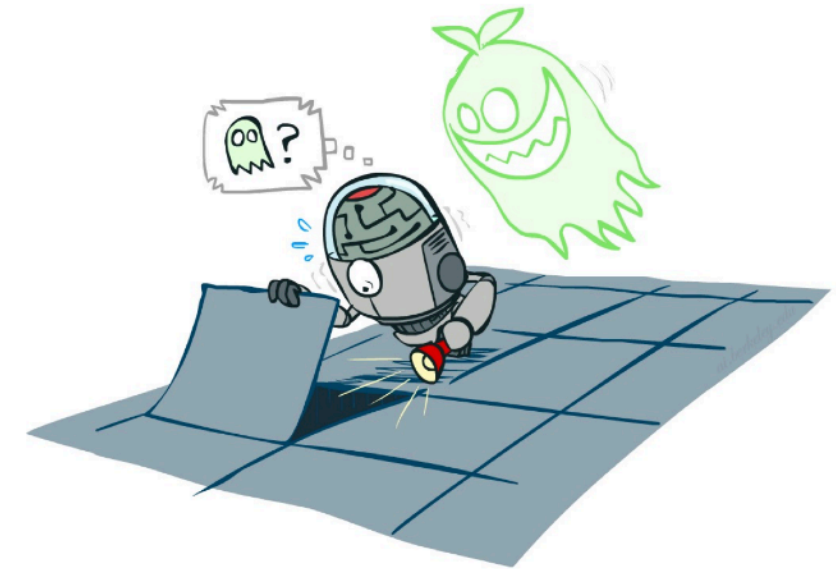
Course Topics

Search &
Planning

Reinforcement
Learning

Probability &
Inference

Supervised
Learning



How can I make sense of ***uncertainty*** in the world?

Course Topics

Search &
Planning

Reinforcement
Learning

Probability &
Inference

Supervised
Learning



How can I learn a *model* of the world from *data*?

Course Topics

Search &
Planning

Reinforcement
Learning

Intelligence from
Computation

Probability &
Inference

Supervised
Learning

Intelligence from
Data/Experience

Course Topics

Search &
Planning

Reinforcement
Learning

Probability &
Inference

Supervised
Learning

Applications

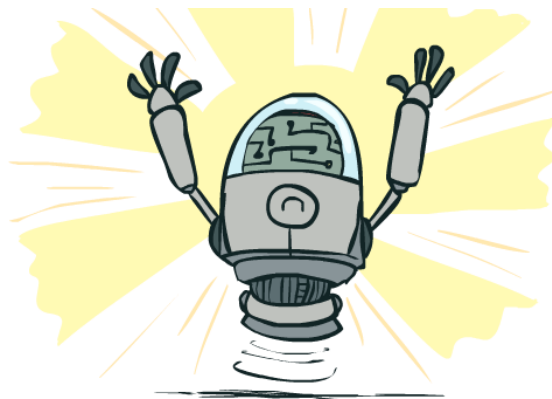
Impact on Sciences, Technology, Society

Should I take CS 188?

- Yes, if you want to know how to design rational agents!
 - CS 188 gives you extra mathematical maturity
 - CS 188 gives you a survey of other non-CS fields that interact with AI (e.g. robotics, cognitive science, economics)
- Disclaimer: If you're interested in making yourself more competitive for AI jobs, CS 189 and CS 182 are better fits.
 - The last few CS 188 lectures (neural networks) are used by many modern state-of-the-art systems. CS 189 and CS 182 cover these in more depth

By the end of this course you'll:

- Build and understand math of rational, learning agents
- Select and apply the right AI methods for wide range of problems
- Recognize how these methods are used in modern AI systems
- Be prepared to make decisions on how AI is used in society



Funny AI stories

■ James Meehan's Talespin: computer generated Aesop-like fables

"One day Joe Bear was hungry. He asked his friend Irving Bird where some honey was. Irving told him there was a beehive in the oak tree. Joe walked to the oak tree. He ate the beehive."

"Joe Bear was hungry. He asked Irving Bird where some honey was. Irving refused to tell him, so Joe offered to bring him a worm if he'd tell him where some honey was. Irving agreed. But Joe didn't know where any worms were, so he asked Irving, who refused to say. So Joe offered to bring him a worm if he'd tell him where a worm was. Irving agreed. But Joe didn't know where any worms were, so he asked Irving, who refused to say. So Joe offered to bring him a worm if he'd tell him where a worm was ..."

Funny AI stories

■ James Meehan's Talespin: computer generated Aesop-like fables

“One day Henry Crow sat in his tree holding a piece of cheese in his mouth, when up came Bill Fox. Bill saw the cheese and was hungry. He said, 'Henry, I like your singing very much. Won't you please sing for me?' Henry, flattered by this compliment, began to sing. The cheese fell to the ground. Bill Fox saw the cheese on the ground and was very hungry. He became ill. (*Hunger already on the goal stack*) Henry Crow saw the cheese on the ground, and became hungry, but he knew that he owned the cheese. He felt pretty honest with himself, so he decided not to trick himself into giving up the cheese. He wasn't trying to deceive himself, so he asked himself for the cheese. He trusted himself, but he remembered that he was also in a position of dominance over himself, so he refused to give himself the cheese. He couldn't think of a good reason why he should give himself the cheese [if he did that, he'd lose the cheese], so he offered to bring himself a worm if he'd give himself the cheese. That sounded okay, but he didn't know where any worms were. So he said to himself, 'Henry, do you know where any worms are?' But of course, he didn't, so he ...”

Eliza

■ Joseph Weizenbaum's Eliza: emulates Rogerian psychotherapy

Men are all alike

In what way?

They're always bugging us about something or other

Can you think of a specific example?

Well my boyfriend made me come here

Your boyfriend made you come here?

He says I'm depressed most of the time

I'm sorry to hear you are depressed

It's true, I am unhappy

Do you think coming here will help you not to be unhappy?

I need some help, that much seems certain

What would it mean to you if you got some help?

Perhaps I could learn to get along with my mother

Tell me more about your family

My mother takes care of me

Who else in your family takes care of you?

My father

Your father?

You are like my father in some ways

What resemblance do you see?

You are not very aggressive, but I think you don't want me to notice that

What makes you think I am not very aggressive?

You don't argue with me

Why do you think I don't argue with you?

You are afraid of me

Does it please you to believe I am afraid of you?

My father is afraid of everybody

What else come to mind when you think of your father?

Bullies

Does that have anything to do with the fact that your boyfriend made you come here?

Next Week: Search

