CSE 473:
Artificial Intelligence

Hanna Hajishirzi

slides edited by Akari Asai, Yizhong Wang adapted from
Dan Klein, Pieter Abbeel ai.berkeley.edu
And Dan Weld, Luke Zettlemoyer
Website

- Website
  - tentative schedule
  - lecture slides
  - course policies, etc.
Course Staff (TAs)

TA: Akari Asai
Office hours: TBD
Location: TBD

TA: Yizhong Wang
Office hours: TBD
Location: TBD

TA: Skyler Hallinan
Office hours: TBD
Zoom link
Location: TBD

TA: Michael Lee
Office hours: TBD
Location: TBD

TA: Benlin Liu
Office hours: TBD
Location: TBD

Ronan Gupta
Kevin Farhat

TA: Wisdom Ikezogwo
Office hours: TBD
Location: TBD

TA: Khushi Khandelwal
Office hours: TBD
Location: TBD

TA: William Howard-Snyder
Office hours: TBD
Location: TBD

TA: Stefan Todoran
Office hours: TBD
Location: TBD
Course Staff

• Office hours
  • Schedule on the website
  • **TAs**: concepts, projects, homework
  • **Hanna**: concepts, high level guidance
Logistics

- **Canvas**: grades, submitting programming assignments:

- **GradeScope** for written assignments

- **Ed**: Discussion board: ask and answer questions; announcements
Course Format

- Programming Assignments
  - 4 projects
  - Python
  - Autograded
  - Give you hands-on experience with the algorithms
  - I expect you to get 100% on projects

- Written homeworks
  - 4 written homeworks
  - Gives you a more conceptual understanding of the material
Course format

- Quizzes
  - 2 quizzes
  - Mainly short-answer questions; evaluates your understanding of the class, homeworks, and programming assignments
Prerequisites

- Data Structure or Equivalent: CSE 332

- Math:
  - Basic exposure to probability and data structures

- Programming – Familiar with Python
  - There is a 0th project (P0)
Textbook

- Not required, but for students who want to read more we recommend

- Warning: Not a course textbook, so our presentation does not necessarily follow the presentation in the book.
Course Policies

- Grade:
  - Your grade will be:
    - 45% programming assignments,
    - 45% written homeworks,
    - 10% (2 quizzes)

- Assignments should be done individually unless otherwise specified.

- Late Policy: Six penalty-free late day for the whole quarter; maximum 4 days per assignment.
Today

- What is artificial intelligence (AI)?
- What can AI do?
- What is this course?
Science Fiction AI
Artificial Intelligence in Action
What is AI?

Explain Artificial Intelligence using precise definitions for third-year computer science undergraduate students.

Certainly! Artificial Intelligence (AI) refers to the development of computer systems that can perform tasks that typically require human intelligence. These tasks include things like visual perception, speech recognition, problem-solving, and decision-making.

Machine Learning is a subset of AI that focuses on training computers to learn from data and improve their performance on a specific task over time, without being explicitly programmed for that task. This is done through the use of algorithms that allow computers to find patterns and make predictions or decisions based on data.

Deep Learning is a specific type of Machine Learning that uses artificial neural networks with many layers (hence "deep") to analyze and process data. It has been particularly successful in tasks like image and speech recognition.
What is AI?

Explain Artificial Intelligence in simple words so that my six year old kid can understand.

Sure! Imagine if we could teach a computer to think and learn, kind of like how you learn new things in school. That's what we call Artificial Intelligence, or AI for short.

Just like you, when you see a picture of a cat, you know it's a cat because you've seen lots of cats before. With AI, we can teach a computer to recognize cats, dogs, and many other things too!

AI can also help us do things really fast, like finding the best way to get to a friend's house or even telling us what the weather will be like.

It's like having a really smart friend inside the computer that can help us with all sorts of things!
What is AI?

The science of making machines that:

- Think like people
- Act like people
- Think rationally
- Act rationally
Rational Decisions

We’ll use the term *rational* in a very specific, technical way:

- Rational: maximally achieving pre-defined goals
- Rationality only concerns what decisions are made (not the thought process behind them)
- Goals are expressed in terms of the utility of outcomes
- Being rational means *maximizing your expected utility*

A better title for this course would be:

*Computational Rationality*
Maximize Your Expected Utility
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!
- “Brains are to intelligence as wings are to flight”
- Lessons learned from the brain: memory and simulation are key to decision making
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
Topics in This Course

- **Part I: Intelligence from Computation**
  - Fast search
  - Adversarial and uncertain search

- **Part II: Reasoning under Uncertainty**
  - Decision theory: Reinforcement Learning, Markov Decision Processes
  - Graphical Models - Bayes Nets; HMMs
  - Machine Learning

- **Throughout: Applications**
  - Natural language, vision, robotics, games, ...
This course w.r.t. Current AI Research

- Rational Agents [decisions]
  - Robot [physically embodied]
- Machine Learning [learning decisions; sometimes independent]
- Human-AI Interaction
- NLP
- Computer Vision
This course w.r.t. Current AI Research

Raw Data

Modeling

- Probabilistic (Graphical Models)
- Neural models
- Feature space & Linear functions

Inference

Decision Making & Planning

Machine Learning
This course vs. others

- Stat methods
- NLP
- ML
- Robotics
- Vision
- Advanced RL
- Deep Learning
Today

- Course overview
- What is artificial intelligence (AI)?
- What can AI do?
- What is this course?
A (Short) History of AI
A (Short) History of AI

1940-1950: Early days
- 1943: McCulloch & Pitts: Boolean circuit model of brain
- 1950: Turing’s “Computing Machinery and Intelligence”

1950—70: Excitement: Look, Ma, no hands!
- 1950s: Early AI programs, including Samuel’s checkers program, Newell & Simon’s Logic Theorist, Gelernter’s Geometry Engine
- 1956: Dartmouth meeting: “Artificial Intelligence” adopted
- 1965: Robinson’s complete algorithm for logical reasoning

1970—90: Knowledge-based approaches
- 1969—79: Early development of knowledge-based systems
- 1980—88: Expert systems industry booms

1990—2012: Statistical approaches
- Resurgence of probability, focus on uncertainty
- General increase in technical depth
- Agents and learning systems... “AI Spring”?

- Big Data, big compute, neural networks
- Some re-unification of subfields

2022-present: More excitement; Wow!!!
- AI is being used by public.
What Can AI Do? [Updated 2023]

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

- Play a decent game of Jeopardy?
- Win against any human at chess?
- Win against the best humans at Go?
- Play a decent game of tennis?
- Grab a particular cup and put it on a shelf?
- Unload any dishwasher in any home?
- Drive safely along the highway?
- Drive safely along University Avenue?
- Buy a week's worth of groceries on the web?
- Buy a week's worth of groceries at QFC?
- Discover and prove a new mathematical theorem?
- Perform a surgical operation?
- Unload a known dishwasher in collaboration with a person?
- Translate spoken Chinese into spoken English in real time?
- Write an intentionally funny story?
Computer Vision

Scene Segmentation

Object Recognition
Object Recognition

YOLO v8 (2023)
An illustration of an avocado sitting in a therapist’s chair, saying ‘I just feel so empty inside’ with a pit-sized hole in its center. The therapist, a spoon, scribbles notes.

Tiny potato kings wearing majestic crowns, sitting on thrones, overseeing their vast potato kingdom filled with potato subjects and potato castles.
AI artwork?

Natural Language

- Speech technologies (e.g. Siri)
  - Automatic speech recognition (ASR)
  - Text-to-speech synthesis (TTS)
  - Dialog systems

- Language processing technologies
  - Question answering
  - Machine translation

"Il est impossible aux journalistes de rentrer dans les régions tibétaines"
Bruno Philip, correspondant du "Monde" en Chine, estime que les journalistes de l'AFP qui ont été expulsés de la province tibétaine de Qinghai "n'étaient pas dans l'illegalité".


"It is impossible for journalists to enter Tibetan areas"
Philip Bruno, correspondent for "World" in China, said that journalists of the AFP who have been deported from the Tibetan province of Qinghai "were not illegal."

Facts: The Dalai Lama denounces the "exile" imposed since he fled Tibet in 1959.

Video: Anniversary of the Tibetan rebellion: China on guard.

- Web search
- Text classification, spam filtering, etc...
Jeopardy - 2011

http://www.youtube.com/watch?v=WFR3lOm_xhE
Start at 20 sec
Introducing ChatGPT

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.

Try ChatGPT  Read about ChatGPT Plus
AI-powered Search

What are the best restaurants in Seattle?

Here are some of the best restaurants in Seattle, according to Google Maps:

- **Canlis**: Landmark fine-dining destination (since 1950s) offering Pacific NW fare in a midcentury-modern home. It has a 4.6 star rating on Google Maps.

- **RockCreek Seafood & Spirits**: Bi-level eatery serving eco-friendly seafood from worldwide sources in airy, cabinlike quarters. It has a 4.6 star rating on Google Maps.

- **The Pink Door**: Italian restaurant with burlesque shows & trapeze entertainment plus terrace with Elliott Bay views. It has a 4.6 star rating on Google Maps.

- **Art of the Table**: Intimate eatery offering innovative locavore fare, with each course introduced by the chef-owner. It has a 4.5 star rating on Google Maps.

- **Shaker + Spear**: At the Kimpton Palladian Hotel, this chic eatery offers Pacific NW cuisine amid pale-wood decor. It has a 4.5 star rating on Google Maps.

These are just a few of the many great restaurants in Seattle. With so many options to choose from, you're sure to find the perfect place to enjoy a delicious meal.
Game Agents

- **Classic Moment: May, '97: Deep Blue vs. Kasparov**
  - First match won against world champion
  - “Intelligent creative” play
  - 200 million board positions per second
  - Humans understood 99.9 of Deep Blue's moves
  - Can do about the same now with a PC cluster

- **1996: Kasparov Beats Deep Blue**
  “I could feel --- I could smell --- a new kind of intelligence across the table.”

- **1997: Deep Blue Beats Kasparov**
  “Deep Blue hasn't proven anything.”

Text from Bart Selman, image from IBM’s Deep Blue pages
Game Agents

- Reinforcement learning

- Pong
- Enduro
- Beamrider
- Q*bert
2016

AlphaGo deep RL defeats Lee Sedol (4-1)
OpenAI wins OG world champions team of 5 pros
AI controlled 5 bots using different layers of same network
  • Trained with RL and self play
  • Equivalent to 45000 years (over 10 months)
Board Game - 2022

- Meta developed CICERO, which achieved human-level performance in the popular strategy game Diplomacy.
Robotics

- Robotics
  - Part mech. eng.
  - Part AI
  - Reality much harder than simulations!

- Technologies
  - Vehicles
  - Rescue
  - Help in the home
  - Lots of automation…

- In this class:
  - We ignore mechanical aspects
  - Methods for planning
  - Methods for control

Images from UC Berkeley, Boston Dynamics, RoboCup, Google
Simulated Agents

Iteration 0

[Schulman, Moritz, Levine, Jordan, Abbeel, ICLR 2016]
Robots
Robocup

https://www.youtube.com/watch?v=_PC-V5GJP6Q
Self-Driving Car 2023

L2: partial driving automation

L4: high driving automation

Robotaxis score a huge victory in California with approval to operate 24/7

/ After a six hour hearing, the California Public Utilities Commission voted to open the floodgate to autonomous vehicles. Public comment was sharply divided over the expansion of self-driving cars in San Francisco.

Tools for Predictions & Decisions
Decision Making

- Applied AI in many kinds of automation:
  - Scheduling, airline routing
  - Route planning
  - Medical diagnosis
  - Web search
  - Spam classification
  - Automated help desks
  - Smarter devices, like cameras
  - Fraud detection
  - Product recommendation
  - … Lots more!
AI is Having Real-World Impact

- Economy
- Politics
- Law
- Sciences
- Educations
Today

- Course overview
- What is artificial intelligence (AI)?
- What can AI do?
- What is this course?
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
Maximize Your Expected Utility
Pac-Man as an Agent

Pac-Man is a registered trademark of Namco-Bandai Games, used here for educational purposes.
Assignments: Pac-man

Originally developed at UC Berkeley:
http://www-inst.eecs.berkeley.edu/~cs188/pacman/pacman.html
Important This Week

- **Important this week:**
  - **Check out** canvas--- our main resource for assignments and grades
  - **Check out** website – for schedule and slides
  - **Check out** Ed – for discussions; we have added everyone to Ed
  - **Check out** Gradescope -- for written assignments
  - **P0: Python tutorial** is out