15-780: Welcome to Graduate AI

Aditi Raghunathan

Jan 13 2023

Outline

- What is this class about?
  - *What is AI?*
- (Some) history of AI
- Course overview
- Policies and logistics

Outline

- What is this class about?
  - *What is AI?*
- (Some) history of AI
- Course overview
- Policies and logistics

What is AI?

What is AI?

Artificial Intelligence (AI) is the field of computer science dedicated to creating systems that can perform tasks requiring human-like intelligence, such as learning, reasoning, perception, and decision-making

We'll go over the history of how what is *meant by AI* changed…

The philosophy of this class

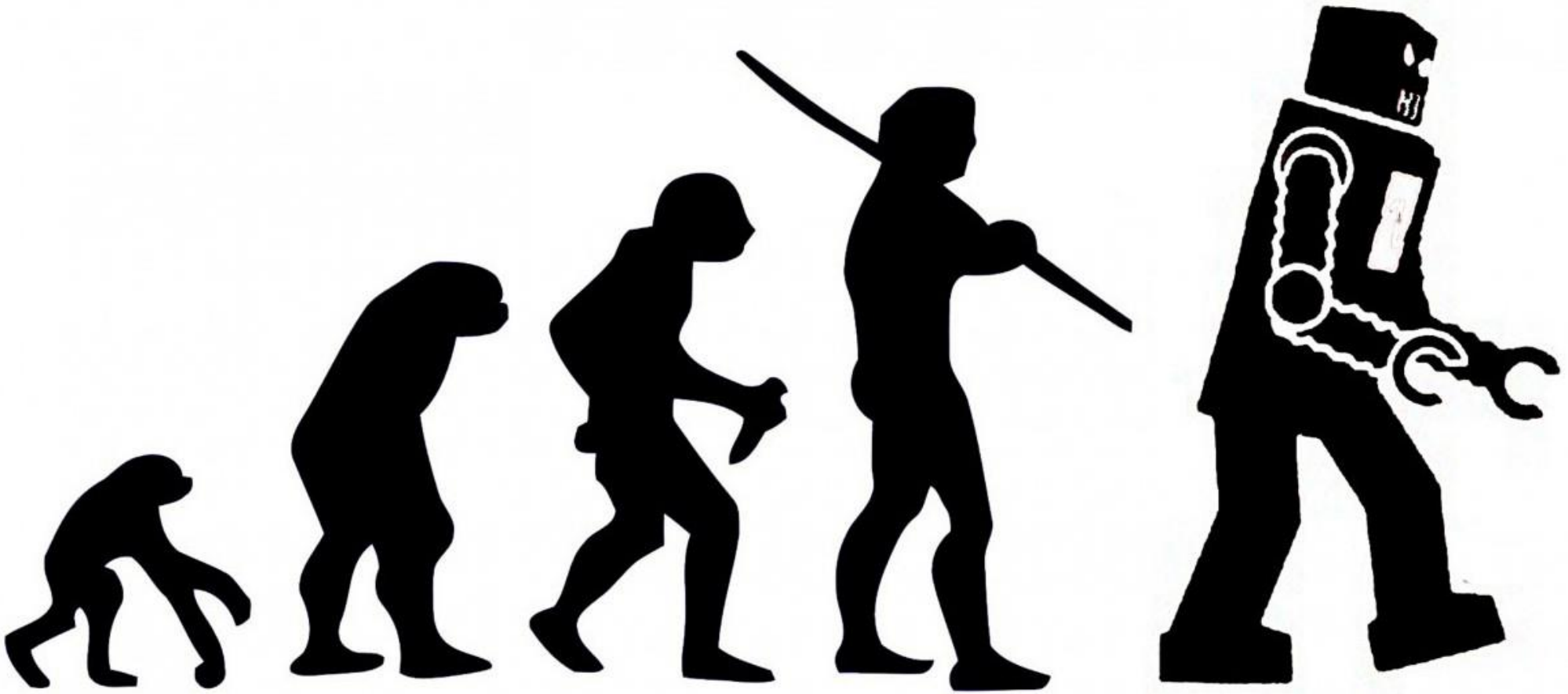
- Foundational understanding of **cutting-edge** AI techniques
  - When should we expect things to work?
  - What is the underlying principle?

- *Some* practical skills in designing and implementing AI systems

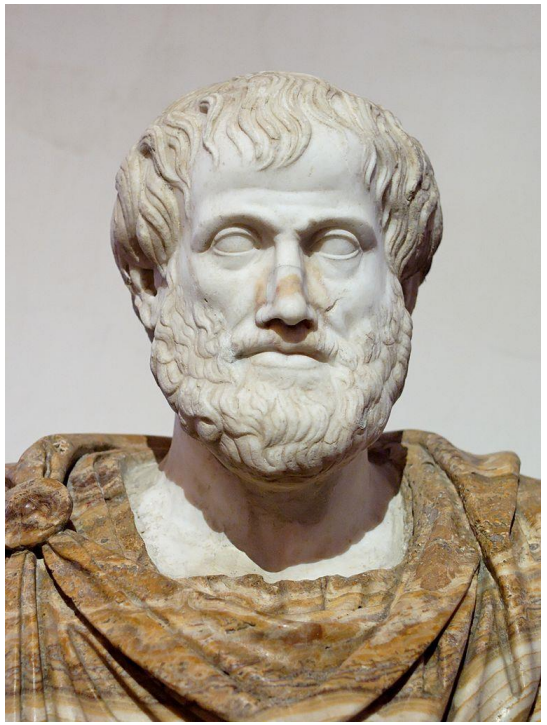
Outline

- What is this class about?
  - *What is AI?*
- (Some) history of AI
- Course overview
- Policies and logistics

A brief history of AI



Prehistory (400 B.C -)



Aristotle

Philosophy: mind/body dualism, materialism

Mathematics: logic, probability, decision theory, game theory

Cognitive psychology

Computer engineering

Optimistic era

1950 Turing test

- **Imitation Game**: A human judge interacts with both a machine and a human without knowing which is which; if the judge cannot reliably distinguish them, the machine is considered "intelligent."

- Took a fundamentally philosophical question and gave an empirical actionable test

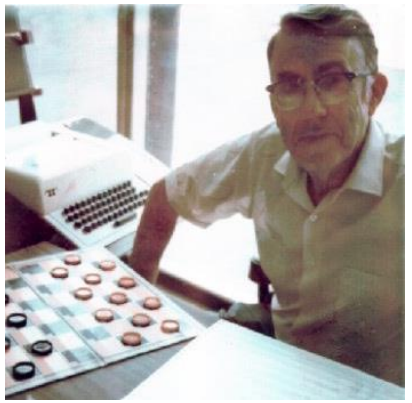
1956 Dartmouth workshop

- **Historical Event:** The Dartmouth Workshop, held in 1956, is widely considered the founding moment of artificial intelligence as a formal field of study

- **Participants:** Brought together pioneers like John McCarthy, Marvin Minsky, Allen Newell, and Herbert Simon, laying the groundwork for future AI research

- **Vision:** Researchers proposed that "every aspect of learning or intelligence" could, in principle, be so precisely described that a machine could simulate it

1952 Arthur Samuel's checkers



Arthur Samuel's program played at a strong amateur level

Combined symbolic reasoning and adaptive learning

Weights in the evaluation function were learned via self-play

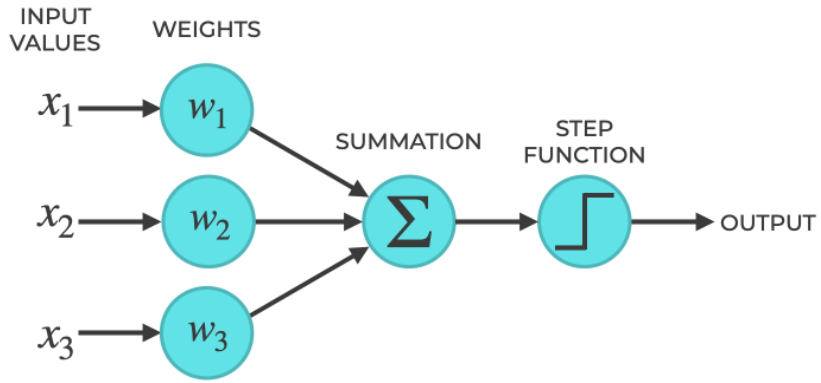
Popularized the usage of "machine learning" and the idea of improving from experience

(1955): Newell & Simon's Logic Theorist:

- Problem solving: prove theorems in Principia Mathematica using search + heuristics

- Later, developed General Problem Solver which promised to solve any problem (that can be encoded in logic)

1958 Rosenblatt's perceptron

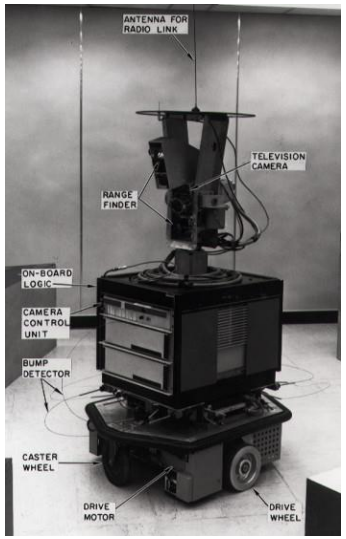


- Hardware implementation of a basic neural network
- Model inspired by the brain by 1943 McCullogh and Pitts

- **First Trainable Neural Network (1958)**: Frank Rosenblatt's perceptron introduced a learning algorithm to update weights based on errors, marking a significant milestone in machine learning

- **Linear Classifier**: The perceptron can classify linearly separable data

Early successes in AI (1950s – 60s)



[1958] McCarthy LISP, advice taker, time sharing

[1968-72] Shakey the robot

[1971-74] Blocksworld planning and reasoning domain

Early success in AI (1950s – 60s)

Overwhelming optimism

Machines will be capable of doing any work a man can do – Herbert Simon [1965]

Within a generation, I am convinced, few compartments of intellect will remain outside the machine's realm—the problems of creating "artificial intelligence" will be substantially solved – Marvin Misnky [1960s]

I visualize a time when we will be to robots what dogs are to humans. And I am rooting for the machines – Claude Shannon

AI winter (1970s– 80s)

First AI winter (Later 1970s)

AI did not live up to the promise

- 1966 ALPAC report cut off funding for machine translation
  - *we will not suddenly or at least quickly attain machine translation*

"Out of sight, out of mind" → (Russian) → "Invisible insanity"

"The spirit is willing, but the flesh is weak" → (Russian) → "The vodka is good but the meat is rotten"

First AI winter (Later 1970s)

- 1973 LightHill report
  - *In no part of the field have the discoveries made so far produced the major impact that was then promised*

- 1970s DARPA cut funding

What went wrong?

What went wrong?

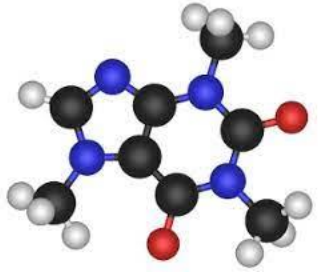
- Limited compute: search space grows exponentially
- Limited information about the complex world
  - Language translation example
- Limitations of perceptron: Minsky-Papert report criticized linear networks (no XOR)
  - The report's critique of perceptrons' limitations was interpreted by many as a dismissal of neural networks entirely, leading to reduced funding and interest in the field

- How to address this? The answer at the time was knowledge based systems or expert systems that encode prior knowledge

- Moved away from the optimism of generality…

Knowledge based systems (1970s-80s)

- [1971-74] Feigenbaum's DENRAL to infer molecular structure from mass spectrometry



- MYCIN: diagnose blood infections, recommend antibiotics

- 1981–Japan's "fifth generation" computer project, intelligence computers running Prolog

- [1982] XCON or R1 expert system to configure customer orders; deployed at DEC and saved $40 million a year

Second AI winter (late 1980s to early 1990s)

- Knowledge based systems also failed to deliver at the time
- Required considerable manual effort to develop and maintain
- Deterministic rules could not handle uncertainty
- [1987] DARPA cuts AI funding for expert systems
- [1991] Japan's fifth generation project fails to meet goals

SHRDLU (1971): Terry Winograd

- Executes commands and answers questions in blocks world
- Used symbolic AI techniques: grammar rules, semantic networks, and rule-based reasoning
- *"A number of people have suggested to me that large programs like the SHRDLU program for understanding natural language represent a kind of dead end in AI programming. Complex interactions between its components give the program much of its power, but at the same time they present a formidable obstacle to understanding and extending it"*
- *"The real challenges are not just technical, but social and ethical. We must design technology with a full awareness of its human impact."*

- Marked the transition from overly optimistic symbolic AI to exploring new paradigms

Splintering of AI

Splintering & changing of AI

- Many subfields and ideas: machine learning, computer vision, robotics, language processing, multiagent systems, ...
- Ideas from different fields
  - Bayes rule from probability
  - Cross-fertilization between search in AI and integer programming in operations research
  - Game theory from economics and mathematics
  - Stochastic gradient descent from statistics
  - Value iteration from control theory
  - Artificial neural networks from neuroscience
- AI becomes more mathematical
- Statistical rigor starts to be required in experimental results

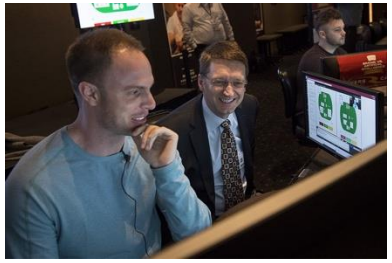
The AI renaissance: focus on applications

- [1997] DeepBlue defeats Gary Kasporov

- [2005, 2007] Stanford and CMU win DARPA grand challenges in autonomous driving

- [2011] IBM's Watson defeats human Jeopardy opponents

- [2017] CMU's Libratus defeats top players at two-player no-limit Texas Hold'em



A paradigm shift in “broad” AI

- [2012] AlexNet: first notable success with neural learning
  - Large scale data (ImageNet) + GPUs + training heuristics
  - Kickstarted the deep learning revolution
- Neural learning had its own optimism and winters
- Further reading: https://awards.acm.org/about/2018-turing

Very brief history of neural AI

- 1943 McClulloch and Pitts; 1958 Perceptron
- 1969: Perceptrons book killed neural nets research
- 1980s: Renewed interest under the banner of connectionism
- 1986: popularization of backpropagation for training multi-layer networks (Rumelhardt, Hinton, Williams)
- 1989: applied convolutional neural networks to recognizing handwritten digits for USPS (LeCun)
- 2006: unsupervised layerwise pre-training of deep networks (Hinton et al.)

A paradigm shift in “broad” AI

- [2016] AlphaGo expanded the horizons of deep learning
- Decision-making, creativity, and strategic planning
- Reinforcement learning + Monte-carlo tree search + deep learning integration

- Further reading: http://www.incompleteideas.net/IncIdeas/BitterLesson.html

[2022] ChatGPT

- **Emergent Conversational Abilities**: It demonstrated **unexpected fluency** and versatility in answering questions, writing, coding etc
- **Massive Public Impact**: a cultural phenomenon, with over **100 million users** within months of release, making conversational AI mainstream and influencing education, business, and creativity
- Intuitive interface for interacting with AI, democratizing AI's capabilities and sparking discussions about ethics, safety, and future directions in human-AI interaction
- **Foundation Models Era**: Exemplified the shift to **general-purpose AI systems**, trained on diverse tasks and adaptable to new one
  - *General methodology -> general models themselves*

Homework 1 (due 1/20)

- Come up with your own version of the Turing test for modern AI
  - Needs to be concrete and actionable!
- How do you use AI systems in your day-to-day?
- Do you trust AI systems in your day-to-day?

Outline

- What is this class about?
  - *What is AI?*
- (Some) history of AI
- Course overview
- Policies and logistics

Course topics

Prediction

Decision-making

- Machine learning
  - Supervised learning
  - Optimization
  - Neural network architectures
  - Unsupervised learning
- Large-scale foundation models
  - Large language models
  - Multimodal models

- Search
- Reinforcement learning
- Game theory
- Inference-time methods with LLMs

Focus on methods that scale with compute

Why should you take this class?

If you're intrigued by AI's vast domain,
Where machines learn, adapt, and train,
Come take this class, where concepts soar,
*In **15-780**, you'll learn and explore.*

From supervised paths to RL's scheme,
To neural nets and the dream of a beam,
Optimization, kernels, decisions vast,
You'll grasp foundations built to last.

With proofs that challenge, code to write,
You'll gain new vision, sharp and bright.
The math may daunt, the data twist,
But triumph lies in tackling the gist.

CMU's finest minds convene,
In this hall where knowledge gleans.
For theory and practice hand in hand,
Will shape the future, bold and grand.

So if you're eager to master AI,
To ask the hows, the whats, the why,
Step forth, enroll, and start the quest,
*In **15-780**, you'll learn from the best!*

Outline

- What is this class about?
  - *What is AI?*
- (Some) history of AI
- Course overview
- Policies and logistics

Class participation – 10%

- We will track in-person attendance
- We will also have periodic in-class short quizzes
- In case of health concerns, please email TAs **in advance**

Homework – 40%

- Both written and programming assignments
- Familiarity with linear algebra, probability and differential calculus, and python
- Homework policy:
  - Be reasonable
  - Collaborate with friends, generative AI but you are responsible for all the content you submit as part of your submission
  - Make sure you learn from the assignments!
  - Assignments would be graded mostly for completion

Midterm – 20%

- Closed laptop, internet, generative AI
- Open notes
- **Date: 2/26**

Class project – 30%

- Groups of 2-3
- Can span broad topics in AI
- Use it as an opportunity to work with cutting-edge systems
- More information to follow!

Student well-being

- CMU and courses like this one are stressful environments
- Don’t sacrifice quality of life for this course: still make time to sleep, eat well, exercise
- The goal of our homeworks and midterm is to facilitate learning
- Ask questions (in class, office hours)

Any questions?