COMS W4701: Artificial Intelligence

Lecture 1a: Introduction to AI

Tony Dear, Ph.D.

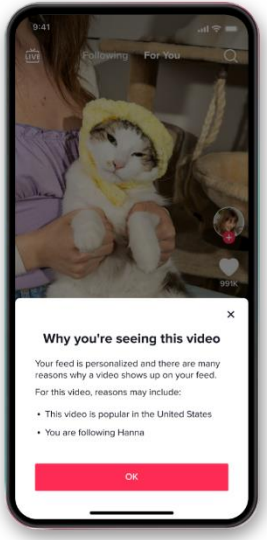
Department of Computer Science

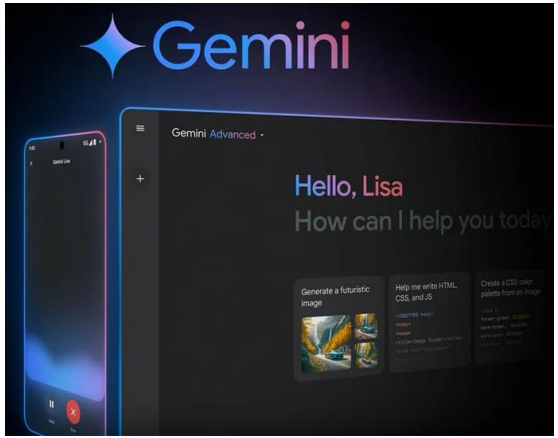
School of Engineering and Applied Sciences

Today

- Course logistics
- Defining intelligence and AI
- History and development of AI
- AI today and to come

Your Experience with AI

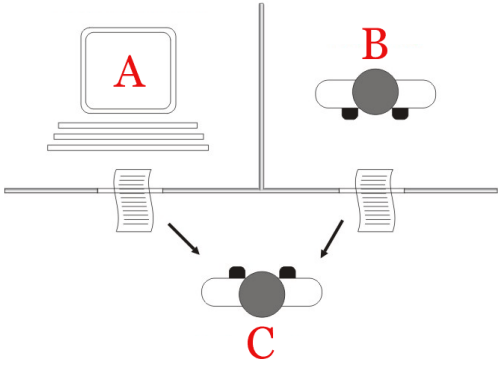




Intelligence Tests

- What is intelligence? Possible definition—behavior that is indistinguishable from that of a human

- **Turing test:** Can a computer reliably imitate a human in natural language so as to be indistinguishable?

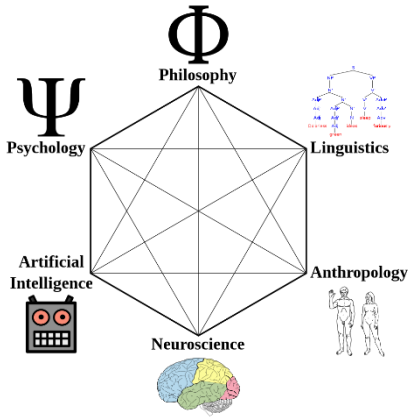


- Simple to implement, can test a breadth of knowledge
- But human intelligence may not be the best milestone

- AI is much more than just NLP and human imitations

Thinking and Reasoning

- Human intelligence is a facet of human *thinking*
- Studied more fully in cognitive science and psychology



- As with human behavior, AI can interface with studies on human thinking, but they are separate fields

- Philosophy and mathematics utilize *logic* and *inference* to reason and make predictions about the world
- Important components of many AI systems, but much of modern AI goes beyond rational thinking

$$\begin{array}{l} p \\ p \rightarrow q \\ \hline \therefore q \end{array}$$

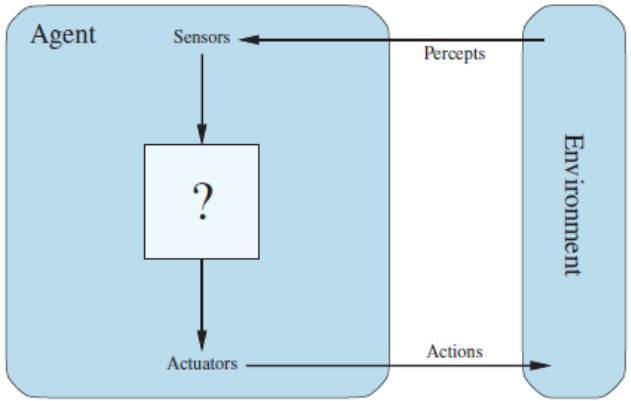
Rational Agents

- AI is concerned with the *synthesis* and *analysis* of **rational agents**
- **Agents** are evaluated by their *actions* in their environments

- Acting **rationally** considers factors like goals, limitations, externalities, experiential learning, and flexibility in dynamic or uncertain situations

Agents and Environments

- An **agent** perceives its environment through *sensors* and acts upon its environment through *actuators*
- Agents may be *controlled* or *autonomous*



- **Agent functions** map percepts to actions
- Other dependencies: Prior knowledge, past experience, goals, preferences, capabilities

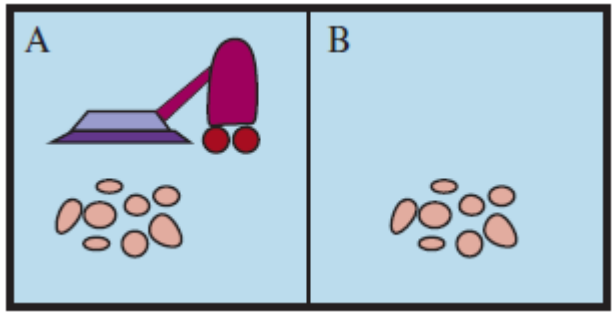
- Environment may be arbitrarily general, but in practice limited to aspects that directly interact with the agent

Examples of Agents

- Humans are agents!
  - Sensors: Vision, hearing, touch, smell, taste, proprioception
  - Actuators: Muscles, reflexes, changing brain state

- AI is more interested in “artificial” agents in complex environments that require nontrivial decision making

- Example: Vacuum cleaner (think Roomba)
- Environment: Square A and square B



Rational Agents

- Environment state sequence can be evaluated by a **performance measure**
- Performance measures usually based on desired *outcomes*, not *behaviors*

- A **rational agent** selects an action to *maximize* its performance measure given percept sequence and in-built knowledge.
  - What performance measure makes our vacuum cleaner rational or irrational?

- Rational agents maximize *expected* performance, are not omniscient
- Rationality may involve info gathering, exploration, learning

Human Intelligence

- Humans are also agents, sometimes acting rationally
- Sources of human intelligence: Biology, culture, lifelong learning
- Humans can also organize into organizations and societies

- Some of these sources may inspire the creation of AI agents
- AI systems may also operate with “humans in the loop” for assistance and feedback on both sides

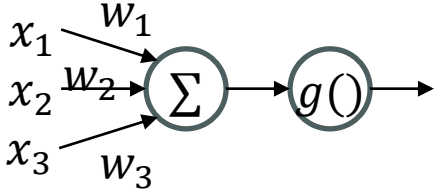
Related Fields and Subdomains

- In addition to previously mentioned fields, AI also shares ideas and techniques with economics, optimization, neuroscience, control theory, linguistics, etc.
- Within CS, AI is a *broad* domain encompassing subdomains like machine learning, natural language processing, computer vision, and robotics
- Each of the above has *subdomains* of their own; e.g., ML includes supervised, unsupervised, self-supervised, reinforcement, and deep learning
- Many techniques and ideas can fall under *multiple* subdomains

Pre-history of AI

- People have built models and technology for “intelligent agents” throughout history—clocks, telephones, hydraulics, automata, etc.

- 16th and 17th centuries: Formal reasoning espoused by European philosophers like Hobbes, Descartes, Leibnitz, and others



- Early 1900s: Turing and Church computational models
- 1940s: Brain models, early artificial neural networks and perceptrons (McCulloch and Pitts, Hebb, Minsky and Edmonds, Rosenblatt)
- 1950s: Checkers programs (Samuel), theorem provers (Newell and Simon)

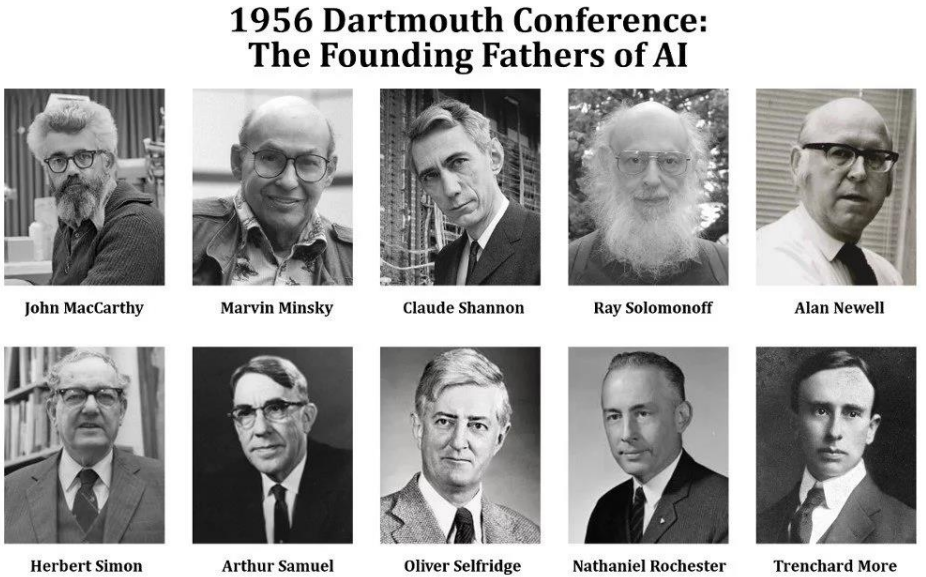
1956 Dartmouth Conference

A Proposal for the

DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

June 17 - Aug. 16

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.



Boom and Bust

- 1956-74: Advances in *symbolic artificial intelligence*
  - Search algorithms, natural language representation, robotics and automata
- 1974-80: First **AI winter** due to computational intractability (Karp, Lighthill), limited computational power, and commonsense knowledge requirements
  - *Perceptrons* (Minsky and Papert 1969) discouraged new research into **connectionism** and neural nets for next 10 years due to pessimistic predictions about applicability
- 1980-87: **Expert systems** used domain knowledge to tackle specific problems
  - Examples: Disease diagnostics, product manufacturing, more advanced chess AIs
  - Lots of investments by companies and governments for large-scale projects
- 1987-93: AI winter when companies failed to deliver and funding dried up

Modern AI

- Late 1980s-1990s: Shift toward “intelligent agents” thanks to paradigms from decision theory, economics, and control theory

- 1988: Introduction of **Bayesian networks** (Pearl) for probabilistic reasoning
- 1986-89: New interest in neural nets and advances in **reinforcement learning**
- 1990s: General shift in machine learning toward data-based approaches

- 2000s-present: **Big data** facilitated success of new ML algorithms
- 2010s-present: **Deep learning** using multiple-layer neural networks (CNNs, large language models), especially in vision, NLP, and generative AI

Possible AI Trends

- Further progress in subdomains like robotics, NLP, and computer vision
- Further progress in automated solutions like search, planning, scheduling
- AI in new domains: Cancer and disease research, sustainability, climate

- Increased mainstream presence (“democratization”): Virtual assistants, recommender systems, business tools, communications, healthcare

- Generative AI: Art, writing, music, speech, videos (“deepfakes”)
- Ethics, fairness, AI safety, explainable AI: “AI for Good”

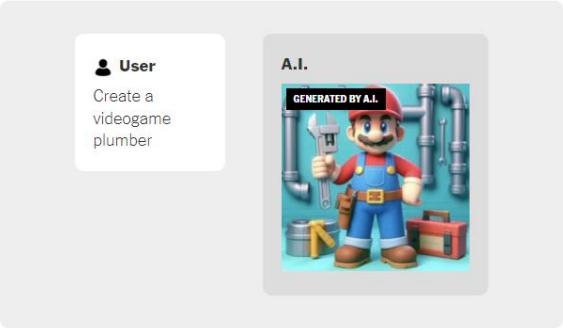
Societal Impacts

Eating disorder helpline takes down chatbot after it gave weight loss advice

JUNE 8, 2023 · 4:21 PM ET

HEARD ON ALL THINGS CONSIDERED

By Kate Wells



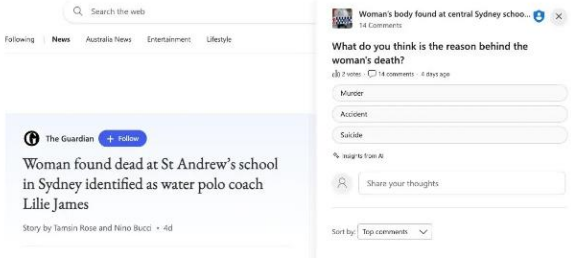
We Asked A.I. to Create the Joker. It Generated a Copyrighted Image.

By Stuart A. Thompson Jan. 25, 2024

Microsoft Criticized for Embedding 'Crass' A.I. Poll Beside News Article

A poll generated by artificial intelligence, embedded next to a Guardian article on Microsoft's news aggregator platform, asked readers to speculate on the cause of a woman's death.

Share full article



What Do You Do When A.I. Takes Your Voice?

Two voice actors say an A.I. company created clones of their voices without their permission. Now they're suing. The company denies it did anything wrong.

We'll need universal basic income - AI 'godfather'

1 day ago

Share

Faisal Islam, Economics editor, @faisalislam

Advertising That Mixes Reality with A.I.

Jewelers are increasingly exploring the technology's potential to create new worlds in online campaigns.

Share full article

Course Outline

- Planning and problem solving through intelligent search
- Decision making to maximize expected utility with uncertainty
- Probabilistic reasoning and inference with incomplete or uncertain information
- Machine learning: Forming models and predictions from data