Introduction to AI

Hal Daumé III

Computer Science
University of Maryland

me@hal3.name

CS 421: Introduction to Artificial Intelligence

26 Jan 2012

Announcements

- **Very important stuff:**
  - HW0 due Tuesday!
  - P0 (Python tutorial) due next-next Tuesday!
  - Subscribe to Piazza *now!*
  - Handin is through our hand-rolled web service

- Waitlist....

- Questions?

Hal Daumé III (me@hal3.name) CS421: Intro to AI

Course Web Page

http://hal3.name/courses/2012S_AI/

Hal Daumé III (me@hal3.name)

Our "sister" course

- WashU St. Louis, taught by Kilian Weinberger:
  - Thanks to John Denero and Dan Klein for sharing all their work!

- Courses are ~90% identical

- Both culminate roughly simultaneously in a Pacman "capture the flag" competition
  - We will *try to* run cross-university Pacman servers

  (CTF demo)

Comments from previous offerings

➢ "There was far too much covered in this course for a single semester... the programming projects alone took too much time for one class."

➢ "Hal needs to learn how to write exams."

➢ "This was one of the best classes I've ever taken."

An experiment....

➢ http://u.hal3.name/ic.pl?q=xyz

Experiment: Item A

Experiment: Item B

Experiment: Item C

Experiment: Item D

Today

➢ What is AI?

➢ Brief history of AI

➢ What can AI do?

➢ What is this course?

Hal Daumé III (me@hal3.name)

Sci-Fi AI?


```
def getSolutionCosts (navigationCode):
        fuelStopCost = 15
        extraComputationCost = 8
  ⟶    thisAlgorithmBecomingSkynetCost = 999999999
        waterCrossingCost = 45
```

GENETIC ALGORITHMS TIP:
ALWAYS INCLUDE THIS IN YOUR FITNESS FUNCTION

Hal Daumé III (me@hal3.name)   CS421: Intro to AI
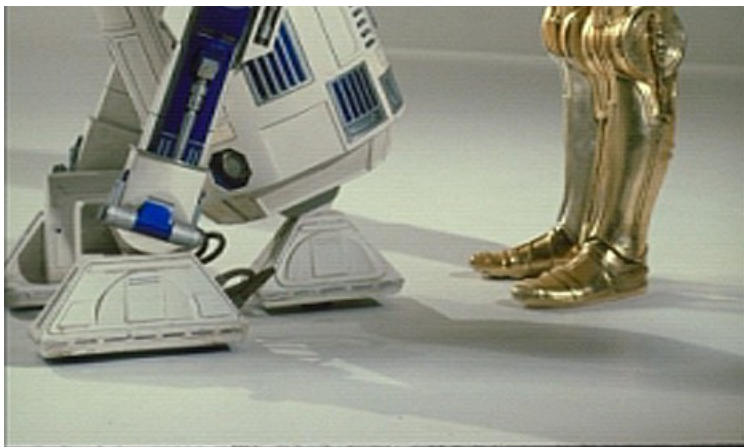
What is AI?

The science of making machines that:

Think like humans	Think rationally
Act like humans	Act rationally

Acting Like Humans?

➢ Turing (1950) "Computing machinery and intelligence"
  ➢ "Can machines think?" → "Can machines behave intelligently?"
  ➢ Operational test for intelligent behavior: the *Imitation Game*
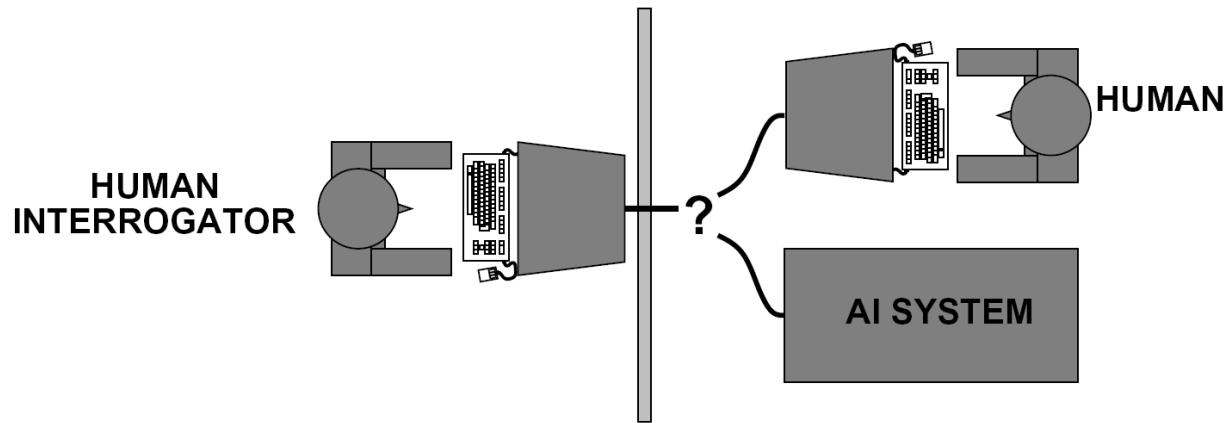
➢ Predicted ~2000, 30% chance of fooling lay person for 5 minutes
➢ Anticipated all major arguments against AI in following 50 years
➢ Suggested major components of AI: knowledge, reasoning, language understanding, learning
➢ Problem: Turing test is not reproducible or amenable to mathematical analysis

Thinking Like Humans?

➢ The cognitive science approach:
  ➢ 1960s ``cognitive revolution'': information-processing psychology replaced prevailing orthodoxy of behaviorism

➢ Scientific theories of internal activities of the brain
  ➢ What level of abstraction? "Knowledge" or "circuits"?
  ➢ **Cognitive science:** Predicting and testing behavior of human subjects (top-down)
  ➢ **Cognitive neuroscience:** Direct identification from neurological data (bottom-up)
  ➢ Both approaches now distinct from AI
  ➢ Both share with AI the following characteristic:
    The available theories do not explain (or engender) anything resembling human-level general intelligence

➢ Hence, all three fields share one principal direction!
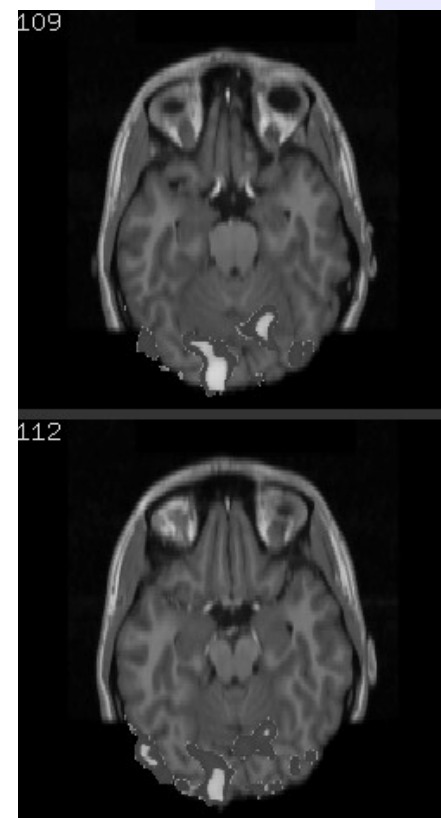
Images from Oxford fMRI center

Thinking Rationally?

➢ The "Laws of Thought" approach
  ➢ What does it mean to "think rationally"?
  ➢ Normative / prescriptive rather than descriptive

➢ Logicist tradition:
  ➢ Logic: notation and rules of derivation for thoughts
  ➢ Aristotle: what are correct arguments/thought processes?
  ➢ Direct line through mathematics, philosophy, to modern AI

➢ Problems:
  ➢ Not all intelligent behavior is mediated by logical deliberation
  ➢ What is the purpose of thinking? What thoughts should I (bother to) have?
  ➢ Logical systems tend to do the wrong thing in the presence of **uncertainty**
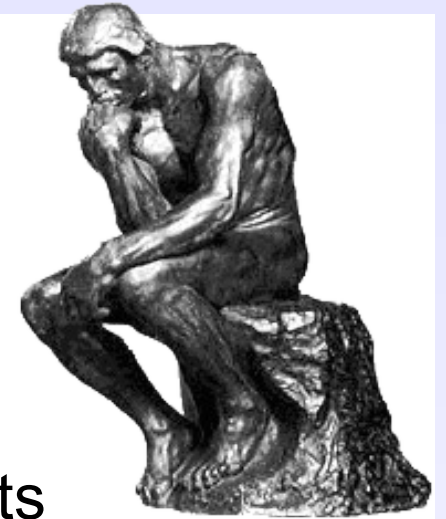
Acting Rationally

- Rational behavior: doing the "right thing"
  - The right thing: that which is expected to maximize goal achievement, given the available information
  - Doesn't necessarily involve thinking, e.g., blinking
  - Thinking can be in the service of rational action
  - Entirely dependent on goals!
  - Irrational ≠ insane, irrationality is sub-optimal action
  - Rational ≠ successful

- Our focus here: rational agents
  - Systems which make the best possible decisions given goals, evidence, and constraints
  - In the real world, usually lots of uncertainty
    - … and lots of complexity
  - Usually, we're just approximating rationality

- "Computational rationality" a better title for this course

Maximize Your Expected Utility

Hal Daumé III (me@hal3.name)

Rational Agents

➤ An **agent** is an entity that perceives and acts (more examples later)

➤ This course is about designing rational agents

➤ Abstractly, an agent is a function from percept histories to actions:

$$\mathcal{P}^* \rightarrow \mathcal{A}$$
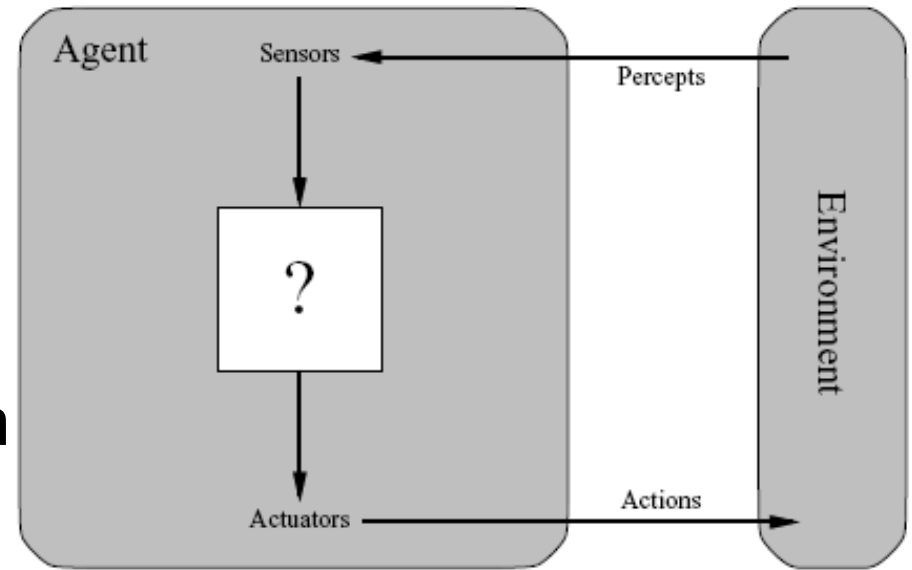
➤ For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance

➤ Computational limitations make perfect rationality unachievable

➤ So we want the best program for given machine resources

[demo: pacman]

AI Adjacent Fields

- Philosophy:
  - Logic, methods of reasoning
  - Mind as physical system
  - Foundations of learning, language, rationality
- Mathematics
  - Formal representation and proof
  - Algorithms, computation, (un)decidability, (in)tractability
  - Probability and statistics
- Psychology
  - Adaptation
  - Phenomena of perception and motor control
  - Experimental techniques (psychophysics, etc.)
- Economics: formal theory of rational decisions
- Linguistics: knowledge representation, grammar
- Neuroscience: physical substrate for mental activity
- Control theory:
  - homeostatic systems, stability
  - simple optimal agent designs

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—88: 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"

- 1988—: Statistical approaches
  - Resurgence of probability, focus on uncertainty
  - General increase in technical depth
  - Agents and learning systems… "AI Spring"?

What Can AI Do?

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

- ✔ Play a decent game of table tennis?
- ✔ Drive safely along a curving mountain road?
- ✘ Drive safely up to Baltimore?
- ✔ Buy a week's worth of groceries on the web?
- ✘ Buy a week's worth of groceries at Giant?
- ❓ Discover and prove a new mathematical theorem?
- ✘ Converse successfully with a person for an hour?
- ❓ Perform a complex surgical operation?
- ✘ Unload a dishwasher and put everything away?
- ✔ Translate spoken Chinese into English in real time?
- ✘ Write an intentionally funny story?

http://u.hal3.name/ic.pl?q=abc

Unintentionally Funny Stories

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. The End.

Henry Squirrel was thirsty. He walked over to the river bank where his good friend Bill Bird was sitting. Henry slipped and fell in the river. Gravity drowned. The End.

Once upon a time there was a dishonest fox and a vain crow. One day the crow was sitting in his tree, holding a piece of cheese in his mouth. He noticed that he was holding the piece of cheese. He became hungry, and swallowed the cheese. The fox walked over to the crow. The End.

[Shank, Tale-Spin System, 1984]

Natural Language

➤ Speech technologies
   ➤ Automatic speech recognition (ASR)
   ➤ Text-to-speech synthesis (TTS)
   ➤ Dialog systems

➤ Language processing technologies
   ➤ Machine translation:
   Aux dires de son président, la commission serait en
      mesure de le faire
   According to the president, the commission would be able
      to do so .

   Il faut du sang dans les veines et du cran .
   We must blood in the veines and the courage .

   ➤ Information extraction
   ➤ Information retrieval, question answering
   ➤ Text classification, spam filtering, etc…

Vision (Perception)

Images from Jitendra Malik
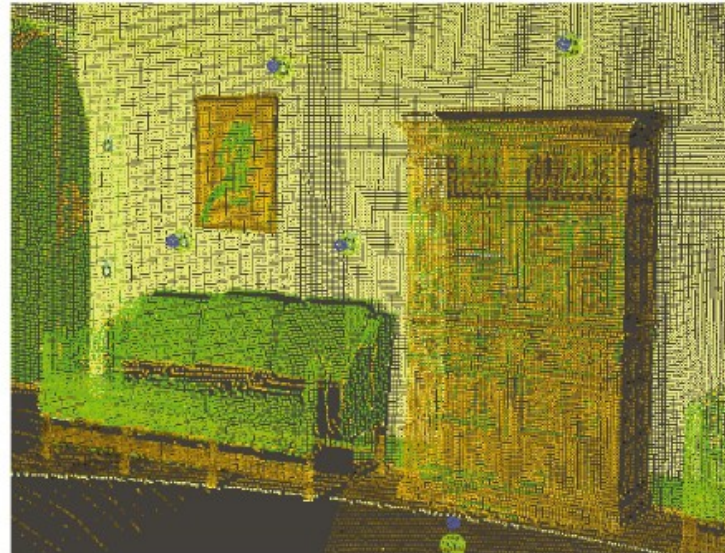
Robotics

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

- ➤ Technologies
  - ➤ Vehicles
  - ➤ Rescue
  - ➤ Soccer!
  - ➤ Lots of automation…

- ➤ In this class:
  - ➤ We ignore mechanical aspects
  - ➤ Methods for planning
  - ➤ Methods for control
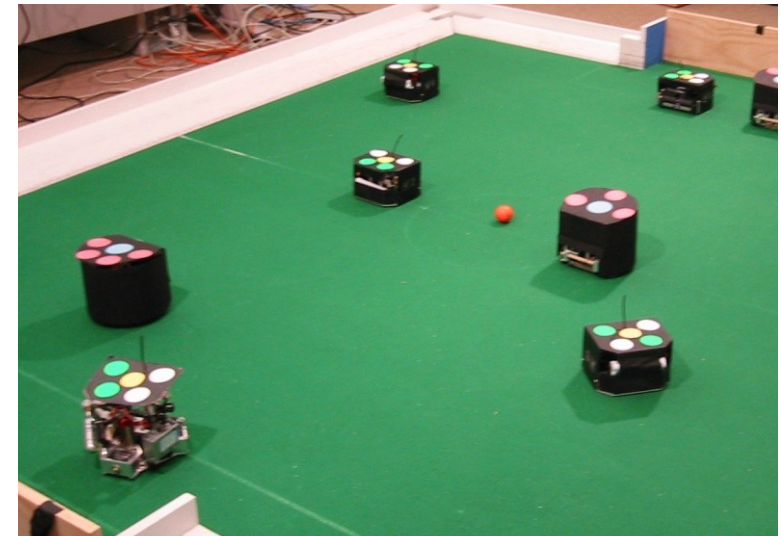
Images from stanfordracing.org, CMU RoboCup, Honda ASIMO sites

Logic

- Logical systems
  - Theorem provers
  - NASA fault diagnosis
  - Question answering

- Methods:
  - Deduction systems
  - Constraint satisfaction
  - Satisfiability solvers
    (huge advances here!)

Image from Bart Selman

Game Playing

- 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 big PC cluster

- Open question:
  - How does human cognition deal with the search space explosion of chess?
  - Or: how can humans compete with computers at all??

- 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

Decision Making

➢ Many applications of AI: decision making

   ➢ Scheduling, e.g. airline routing, military

   ➢ Route planning, e.g. mapquest

   ➢ Medical diagnosis, e.g. Pathfinder system

   ➢ Automated help desks

   ➢ Fraud detection

➢ … the list goes on.

Course Topics

➢ Part I: Optimal Decision Making
  ➢ Fast search
  ➢ Constraint satisfaction
  ➢ Adversarial and uncertain search
➢ Part II: Modeling Uncertainty
  ➢ Reinforcement learning
  ➢ Bayes' nets
  ➢ Decision theory
➢ Throughout: Applications
  ➢ Natural language
  ➢ Vision
  ➢ Robotics
  ➢ Games

Course Projects

➢ Pacman

➢ Robot control

Hal Daumé III (me@hal3.name)   CS421: Intro to AI

Some Hard Questions...

➢ Who is liable if a robot driver has an accident?

➢ Will machines surpass human intelligence?

➢ What will we do with superintelligent machines?

➢ Would such machines have conscious existence? Rights?

➢ Can human minds exist indefinitely within machines (in principle)?