Lecture 1
What is AI?

EECS 348

Intro to Artificial Intelligence

Doug Downey

Outline

1) What is AI: The Course
2) What is AI: The Field
3) Why to take the class (or not)
4) A Brief History of AI
5) “Predict the future” poll

What is AI: The Course

- Communication
  - Web site, e-mails to class
- Grading
  - 6 problem sets (55%)
    - Late assignments 10% off per day
  - Midterms (weeks 4 and 9) (30%)
  - Final: Othello Tournament (10%)
  - Collective lecture ratings (5%)

More on Problem Sets

- Programming (in pairs) and exercises (indv.)
  - PS 1: AI history and search
  - PS 2: Sudoku solver
  - PS 3: Logic and Agents
  - PS 4: Othello player I
  - PS 5: Machine Learning
  - PS 6: Othello player II

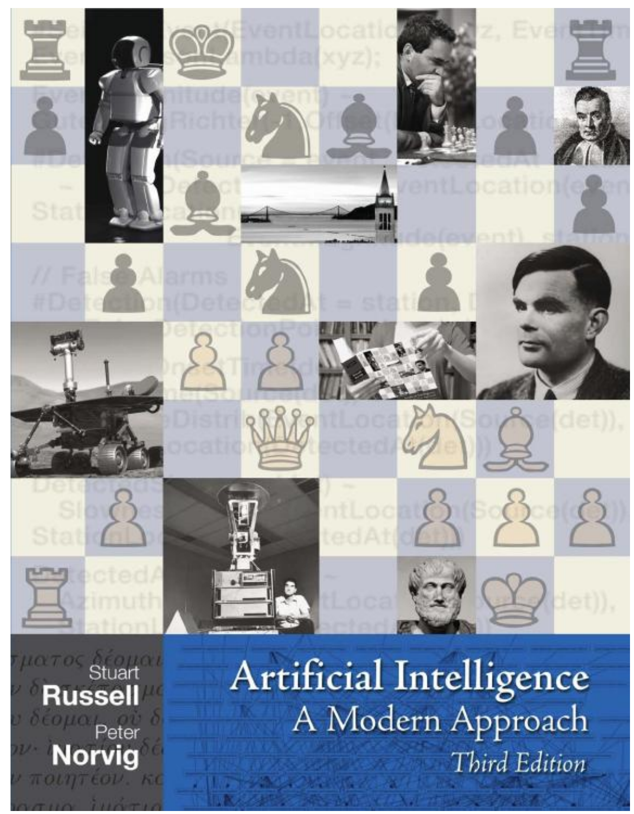
- Code:
  - Starter code in C++ or Python
  - Code in pairs – write reports individually

Topics

1. Introduction to AI, chapter 1.
2. Search, chapters 3, 4.
3. Constraint Satisfaction, Chapter 6.
4. Logic and agents, Ch 7-8.
5. Game playing, chapter 5.
6. Machine learning, chapters 18-20.
7. The Big Questions (final week) chapters 26, 27.

Textbook

Artificial Intelligence: A Modern Approach
Russell and Norvig



Goals of this Course

- To teach you the main ideas of AI
- To introduce you to a set of key techniques and algorithms from AI
- To introduce you to the applicability and limitations of these methods

Outline

1) What is AI: The Course
2) **What is AI: The Field**
3) Why to take the class (or not)
4) A Brief History of AI
5) “Predict the future” poll

- *What are the most fundamental scientific questions?*

What is Intelligence?

What is Artificial Intelligence?

- Dimensions:

  Human-like vs. rational

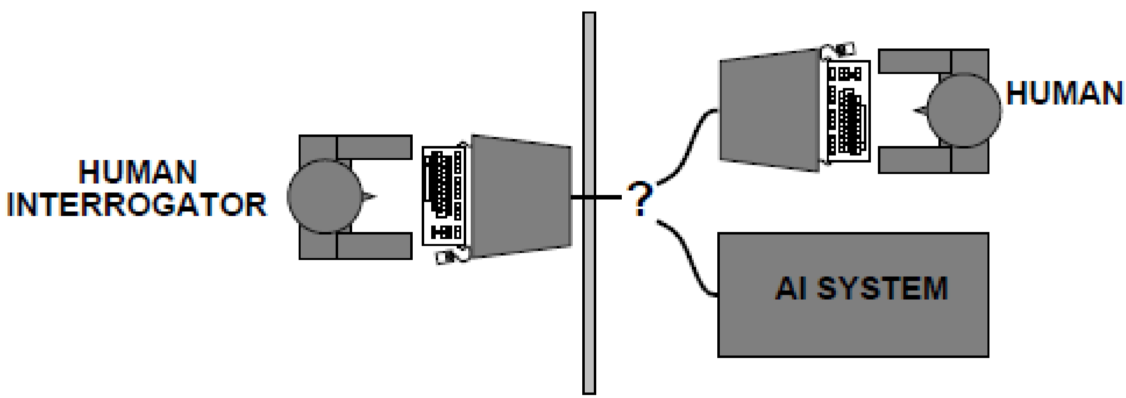
  Behavior vs. thought

Human-like behavior

- Dimensions:

  Human-like vs. rational　　　　**Behavior** vs. thought

- Turing test (1950):



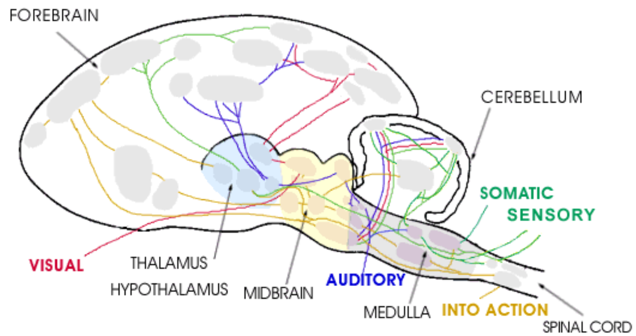
Human-like thought

- Dimensions:

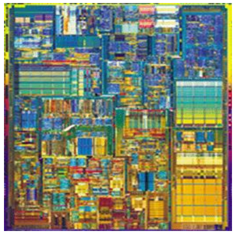
  Human-like vs. rational Behavior vs. **thought**

- Must choose level of abstraction
  - **Knowledge?**
  - **Neurons?**
- How to validate?
  - **Predict and test behavior from human subjects (top-down)**
  - **Measure neurological data (bottom-up)**
- **Cognitive Science** and **Cognitive Neuroscience**
  - Both fields distinct from AI today

Human vs. Computer Hardware

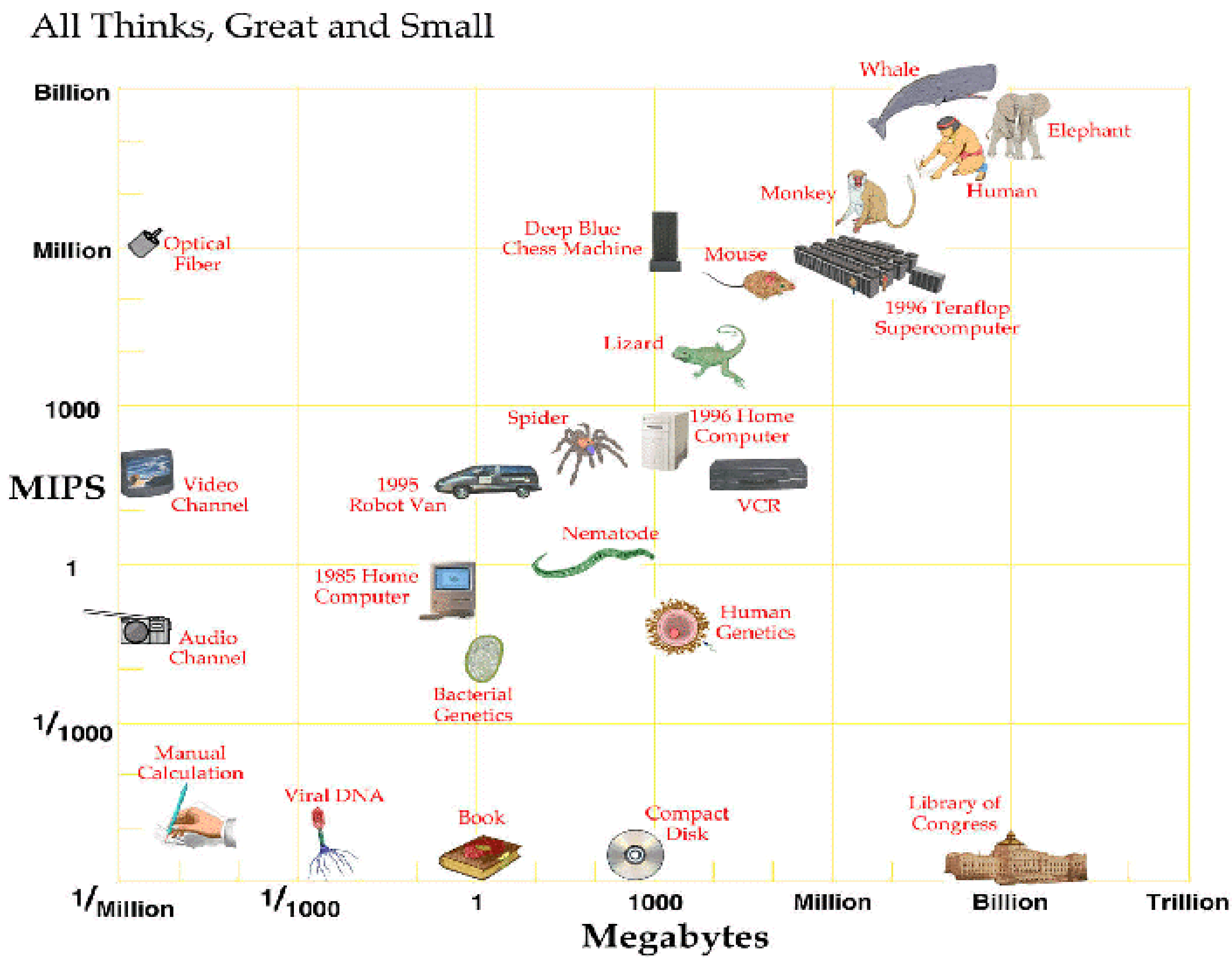


10^{11} neurons
10^{14} synapses
cycle time: 10^{-3} sec



10^{9} transistors
10^{11} bits of RAM
cycle time: 10^{-9} sec

Computer vs. Brain



- Conclusion
  - In near future we can have computers with as many processing elements as our brain, but:
    - fewer interconnections (wires or synapses)
    - much faster updates.

- Fundamentally different hardware may require fundamentally different algorithms
  - Very much an open question.
  - Neural net research.

Thinking Rationally

- Dimensions:

  Human-like vs. **rational** Behavior vs. **thought**

- *Prescriptive*: what would an *ideal* agent think?
  - vs. *descriptive* (what do people *actually* think)
- Harkens to ancient Greeks: logical notation and rules of derivation for thoughts
- Problems:
  - Lots of (rational) actions not due to thought at all
  - What thoughts *should* I think?

Acting Rationally

- Dimensions:

  Human-like vs. **rational** Behavior vs. thought

- Rational agents "do the right thing"
  - Take actions that are optimal for achieving *goals*

- Computational limits prohibit complete rationality
  - Thus, attempt to be "as rational as possible" given resource constraints
- Textbook focuses on "acting rationally" as the definition of AI

Outline

1) What is AI: The Course
2) What is AI: The Field
3) **Why to take the class (or not)**
4) A Brief History of AI
5) “Predict the future” poll

Why not to take the class

- It won’t be easy
- You have to like programming
- You’re best off if you already know:
  - A fair amount about algorithms and data structures
  - The basics of probability theory
  - The basics of first-order logic

Why to take the class

- Touches on a huge number of other fields
    - Mathematics, Philosophy, Neuroscience, Psychology, Cognitive Science, Economics, and of course Computer Science

- Get to play with fun algorithms

- Get to think about the future

- Material has potentially large impact

A question for our time

. . . Although trips around the moon and to neighboring planets may seem a long way off, the United States is probably in a better position at present to progress in this direction than any other nation. Since mastery of the elements is a reliable index of material progress, the nation which first makes significant achievements in space travel will be acknowledged as the world leader in both military and scientific techniques. To visualize the impact on the world one can imagine the consternation and admiration that would be felt here if the United States were to

Should the US Government embark on a 10-year, $1 trillion effort to bring about super-human-level AI?

http://www.fas.org/spp/eprint/origins/part05.htm

Summary of Last Time

- Course structure
  - 6 problem sets (3 programming, 3 written), 2 midterms, othello tournament, lecture ratings
- AI “Definition”

	Human-like	**Rational**
Thought	Cognitive science, cognitive neuroscience	Logic
Behavior	Turing Test	**This class (mostly)**

- 60+ people rated yesterday’s lecture ☺
  (you have until noon tomorrow to rate this one)

Outline

1) What is AI: The Course
2) What is AI: The Field
3) Why to take the class (or not)
4) **A Brief History of AI**
5) "Predict the future" poll

Going way back

- (4^{th} C BC+) Aristotle, George Boole, Gottlob Frege, Alfred Tarski
  – formalizing the laws of rational thought
- (16^{th} C+) Gerolamo Cardano, Pierre de Fermat, James Bernoulli, Thomas Bayes
  – formalizing probabilistic reasoning
- (1950+) Alan Turing, John von Neumann, Claude Shannon
  – thinking as computation
- (1956) John McCarthy, Marvin Minsky, Herbert Simon, Allen Newell
  – start of the field of AI

Classical AI

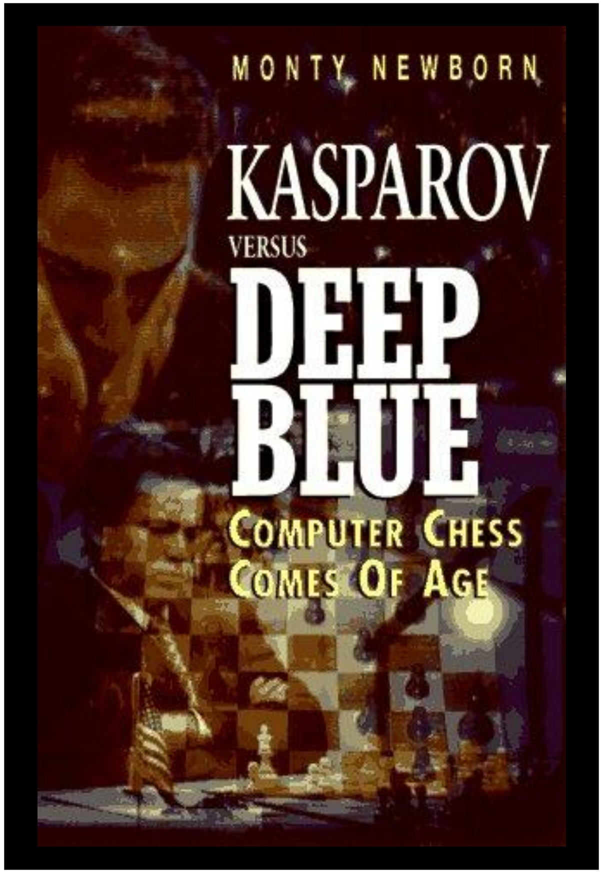
- The principles of intelligence are separate from any hardware / software / wetware implementation
- Look for these principles by studying how to perform tasks that require intelligence

Success Story: Expert Systems

- Gather knowledge from experts, codify it in software
- Example: Mycin (1980)
  - Expert level performance in diagnosis of blood infections

- Rose to prominence in early 80s. Today: 1,000's of systems
  - Everything from diagnosing cancer to configuring aircraft
  - Often outperform e.g. doctors in clinical trials

Success Story: Chess



I could feel – I could smell – a new kind of intelligence across the table
- Kasparov

Success story: IBM’s Watson

“I for one welcome our new robot overlords.”
-- Ken Jennings

Autonomous Systems

- In the 1990's there was a growing concern that work in classical AI ignored crucial scientific questions:
  - How do we integrate the components of intelligence (*e.g.* learning & planning)?
  - How do perception and action interact with reasoning?
  - How does the demand for real-time performance in a complex, changing environment affect the architecture of intelligence?

Provide a standard problem where a wide range of technologies can be integrated and examined

By 2050, develop a team of fully autonomous humanoid robots that can win against the human world champion team in soccer.

Key Hard Problem for AI

- Today's successful AI systems
  - operate in well-defined domains
  - employ narrow, specialized knowledge
    - IBM Watson an exception in *some* ways
- *Commonsense Knowledge*
  - needed to operate in messy, complex, open-ended worlds
    - Your kitchen vs. GM factory floor
  - understand unconstrained Natural Language

Role of Knowledge in Natural Language Understanding

- Speech Recognition
  - Massive investment, considerable progress
- Translation
  - Getting better. Classic mistake:

  The spirit is willing but the flesh is weak. (English)

  The vodka is good but the meat is rotten. (Russian)
- Understanding?

How to Get Commonsense?

- CYC Project (Doug Lenat, Cycorp)
  - Encoding 1,000,000 commonsense facts about the world by hand
  - Coverage not always adequate
- Shift from late 90s+ toward **learning from data**
  - E.g., mine common sense from **text** (IBM Watson, TextRunner do versions of this)
  - **Machine learning** from data enables many other applications as well (more on this later in course)

(Re-)Current Themes

- Combinatorial Explosion
- Micro-world successes don’t scale up.
- How to organize and accumulate large amounts of knowledge?
- How to translate from informal, ill-structured statements to formal reasoning (e.g., understand a story)?
- What are reasonable simplifying assumptions?

Outline

1) What is AI: The Course
2) What is AI: The Field
3) Why to take the class (or not)
4) A Brief History of AI
5) **"Predict the future" poll**

In the future…

- A computer will pass the Turing Test in:
  a) <20 years
  b) 20-50 years
  c) 50-100 years
  d) 100+ years
  e) Never

The future…

- 80% of households will have humanoid robots in…
    a) <20 years
    b) 20-50 years
    c) 50-100 years
    d) 100+ years
    e) Never

The future…

- The **most crucial** advance needed for progress in AI is:
  a) Better hardware (Faster CPUs/more RAM)
  b) Better software (algorithms)
  c) Better understanding of human intelligence/brains
  d) Better ways to harness human participation

The future…

- The US Government should spend **$1 trillion** to advance AI:
  a) Yes
  b) No