

# Lecture 1

## What is AI?

EECS 348

Intro to Artificial Intelligence

Doug Downey

With material adapted from Oren Etzioni (UW) and Stuart Russell (UC Berkeley)

# 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
  - In class
  - Send feedback early and often
- Grading
  - 6 problem sets
  - Late assignments not graded
    - Can drop lowest grade
- Midterm (week 6 or 7), no final

# More on Problem Sets

- Programming and exercises
  - PS 1: AI history and search
  - PS 2 (pairs): Othello player I
  - PS 3 (pairs): Othello player II
    - Tournament!
  - PS 4: TBD
  - PS 5: Sudoku solver
  - PS 6: TBD
- Code:
  - Language of your choice
  - Write your own code & explain how it works

# Topics

1. Introduction to AI, chapter 1.
2. Search, chapters 3, 4.
3. Game playing, chapter 5.
4. Constraint Satisfaction, Chapter 6.
5. Logic and agents, Ch 7-8.
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) “Predict the future” poll
- 5) A Brief History of AI

# What is AI

- *What are the most fundamental scientific questions?*

# What is Intelligence?

# What is Artificial Intelligence?

- Dimensions:

Human-like vs. rational

Behavior vs. thought

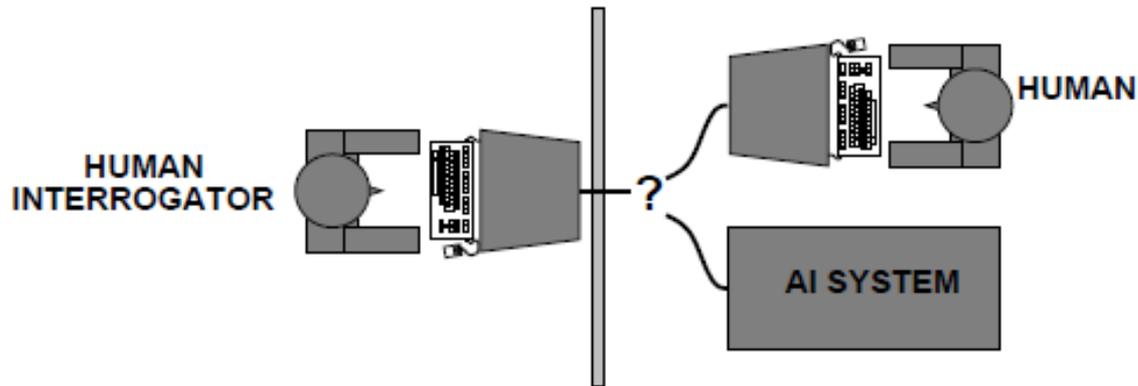
# Human-like behavior

- Dimensions:

**Human-like** vs. rational

**Behavior** vs. thought

- Turing test (1950):



# Human-like thought

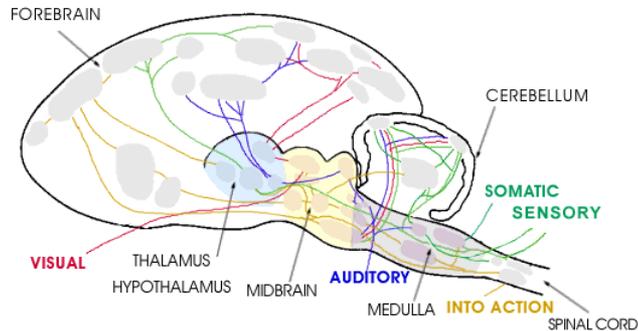
- Dimensions:

**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
  - Like AI, can't explain or produce human-level intelligence yet

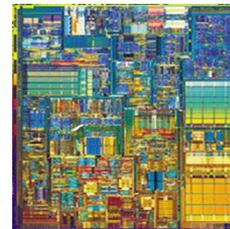
# Human vs. Computer Hardware



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

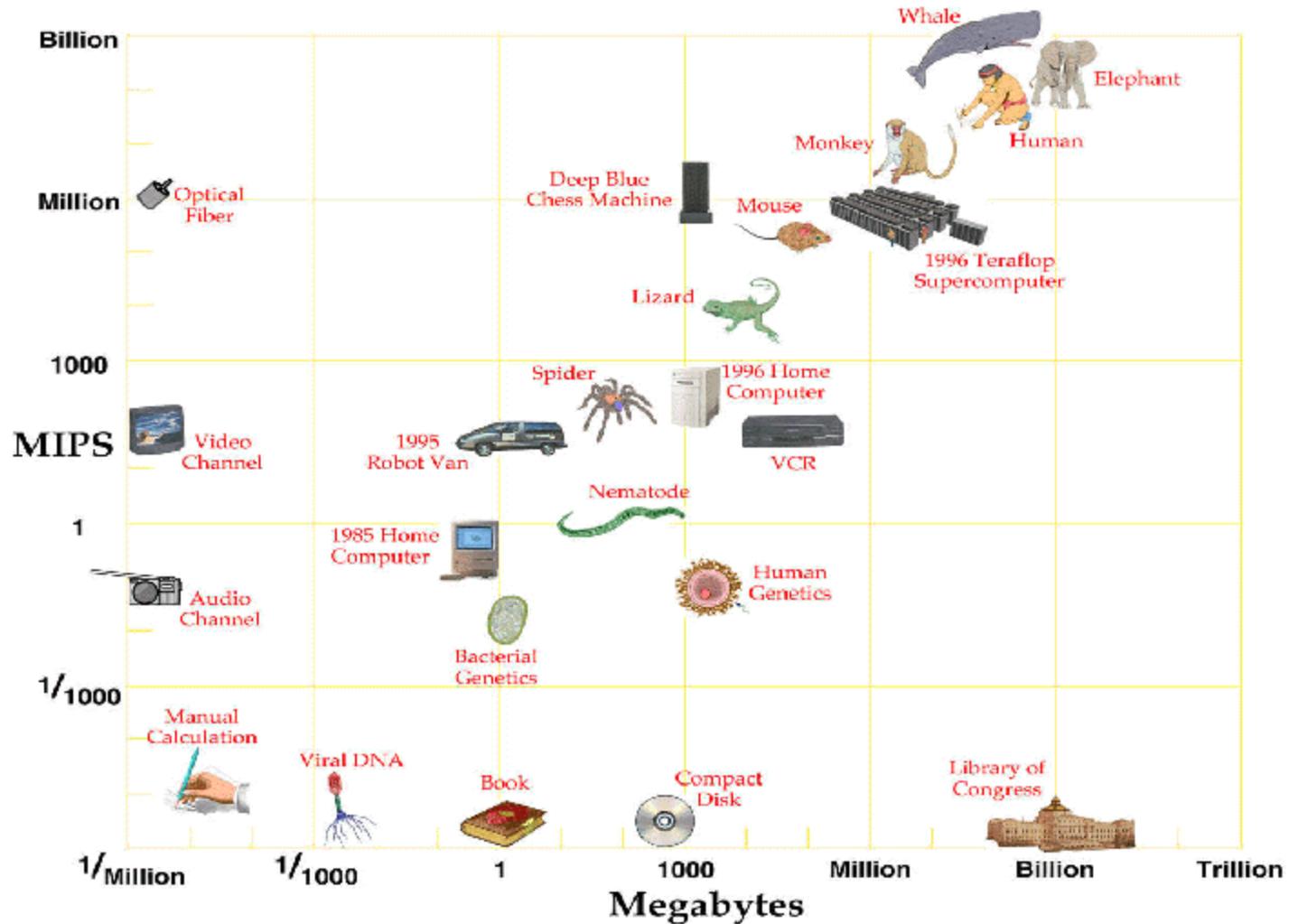
—

$10^7$  transistors  
 $10^{10}$  bits of RAM  
cycle time:  $10^{-9}$  sec



# Computer vs. Brain

All Things, Great and Small



- 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**
- *Perscriptive*: 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

- My first time teaching it
- You have to like/tolerate 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, Cognitive Science, and of course Computer Science
- Get to play with fun algorithms
- Get to think about the future
- Material has potentially large impact

# 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<sup>th</sup> C BC+) Aristotle, George Boole, Gottlob Frege, Alfred Tarski
  - formalizing the laws of human thought
- (16<sup>th</sup> 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



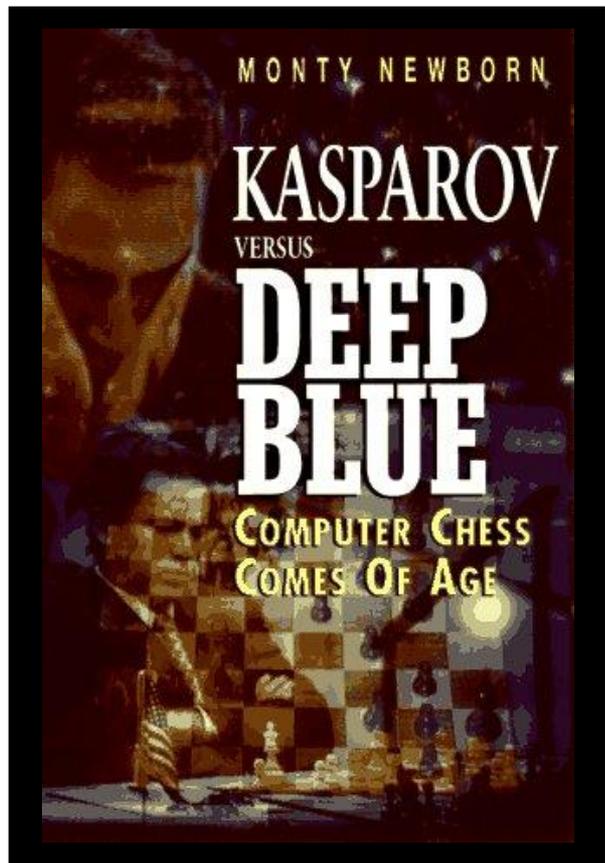
# Classical AI

- 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*

- Examines 5 billion positions / second
- Intelligent behavior **emerges** from brute-force search

# 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
- *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
  - “word spotting” feasible today
  - continuous speech – rapid progress
  - turns out that “low level” signal not as ambiguous as we once thought
- Translation / Understanding
  - very limited progress

*The spirit is willing but the flesh is weak. (English)*  
*The vodka is good but the meat is rotten. (Russian)*

# Syntactic, Semantic, Analogical Knowledge

- Time flies like an arrow.
- Fruit flies like a banana.
- Fruit flies like a rock.

# How to Get Commonsense?

- **CYC Project** (Doug Lenat, Cycorp)
  - Encoding 1,000,000 commonsense facts about the world by hand
  - Coverage still too spotty for use!
- Alternatives?
- Open Mind
- KnowItAll

# Recurrent Themes

- **Neural nets vs Logic**
  - McCulloch & Pitts 1943
  - Died out in 1960's, revived in 1980's
    - Neural nets vastly simplified model of real neurons, but still useful & practical – massive parallelism
    - particular family of learning and representation techniques
- **Logic vs Probability**
  - In 1950's logic seemed more computationally & expressively attractive (McCarthy, Newell)
    - attempts to extend logic “just a little” to deal with the fact that the world is uncertain!
  - 1988 – Judea Pearl's work on Bayes nets
    - provided efficient computational framework
  - Today – no longer rivals
    - hot topic: combining probability & first-order logic

# Recurrent Themes, cont.

- **Weak vs Strong Methods**
  - Weak – general search methods
    - A\* search, constraint propagation, ...
  - Rise of “knowledge intensive” approach
    - expert systems
    - more knowledge, less computation
  - Today: resurgence of weak methods
    - desktop supercomputers
    - in highly competitive domains (Chess) **exceptions** to the general rules are most important!
  - **How to combine weak and strong methods seamlessly?**

# (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?

# What we can do, what we can't

- Write a funny joke?
- Play a decent game of ping-pong?
- Converse coherently with a human for an hour?
- Drive across a windy mountain road?
- Drive safely down Sheridan Road?

# 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**