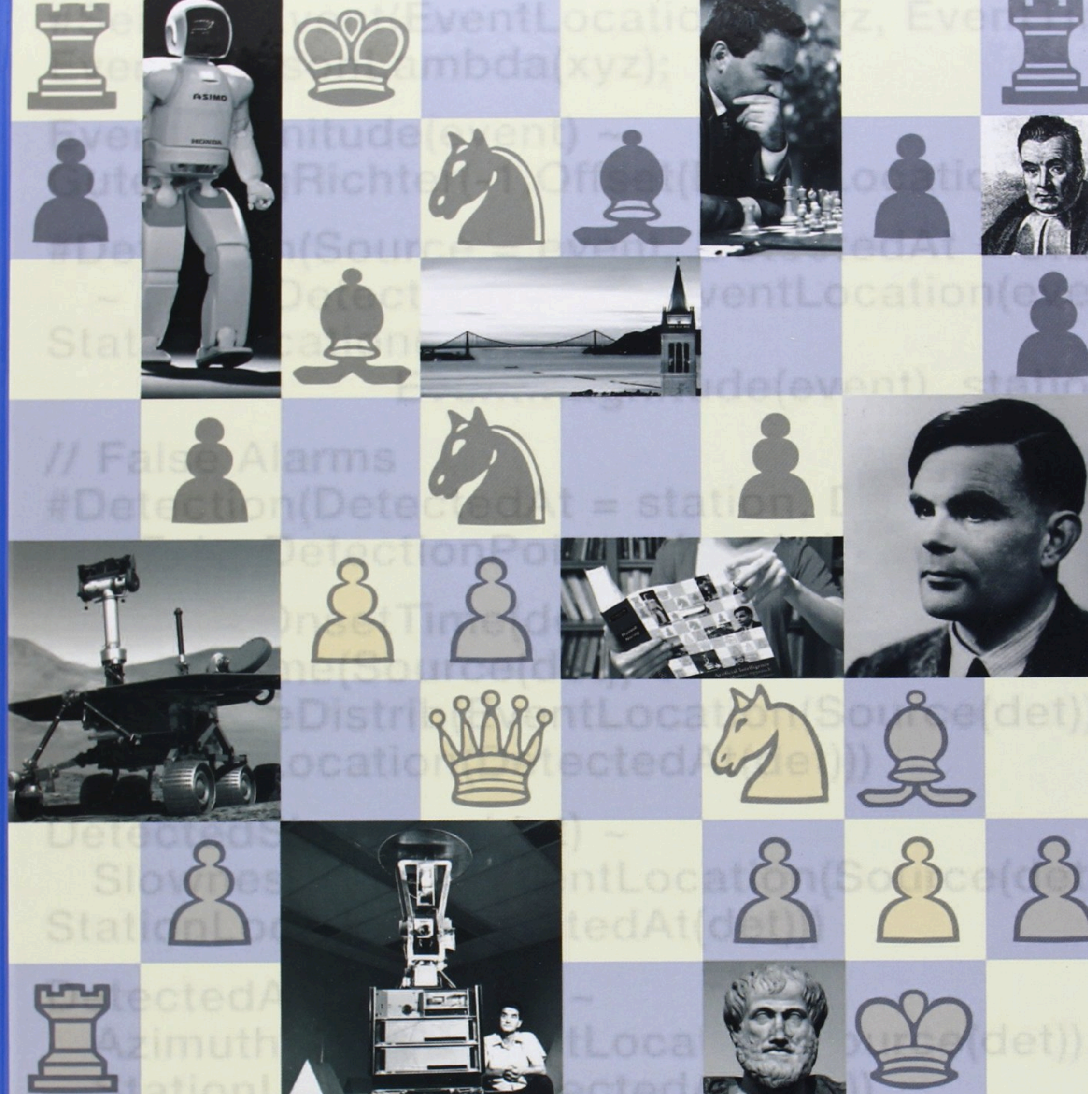
CS 4700: Foundations of Artificial Intelligence

Instructor:
Prof. Selman
selman@cs.cornell.edu

Introduction
(Reading R&N: Chapter 1)



Stuart
Russell
Peter
Norvig

Artificial Intelligence
A Modern Approach
Third Edition

Course Administration (separate slides) ✓

What is Artificial Intelligence?

Course Themes, Goals, and Syllabus

AI: Goals

Ambitious goals:

- **understand "intelligent" behavior**
- **build "intelligent" agents / artifacts**

autonomous systems

understand human cognition (learning, reasoning, planning, and decision making) as a computational process.

What is Intelligence?

Intelligence:

- capacity to learn and solve problems”
  (Webster dictionary)
- the ability to act rationally

Hmm… Not so easy to define.

What is AI?

Views of AI fall into four different perspectives --- two dimensions:

1) Thinking versus Acting
2) Human versus Rational (which is "easier"?)

	Human-like Intelligence	"Ideal" Intelligent/ Pure Rationality
Thought/ Reasoning ("modeling thought / brain)	2. Thinking humanly	3. Thinking Rationally
Behavior/ Actions "behaviorism" "mimics behavior"	1. Acting Humanly	4. Acting Rationally

Different AI Perspectives

2. Systems that ***think like humans***	3. Systems that ***think rationally (optimally)***
"The exciting new effort to make computers think ... *machines with minds*, in the full and literal sense" (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning ..." (Bellman, 1978)	"The study of mental faculties through the use of computational models" (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act" (Winston, 1992)
"The art of creating machines that perform functions that require intelligence when performed by people" (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better" (Rich and Knight, 1991)	"A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes" (Schalkoff, 1990) "The branch of computer science that is concerned with the automation of intelligent behavior" (Luger and Stubblefield, 1993)
1. Systems that *act like humans*	**4. Systems that *act rationally***

Note: A system may be able to *act like a human without thinking like a human!* Could easily "fool" us into thinking it was human!

1. Acting Humanly

	Human-like Intelligence	"Ideal" Intelligent/ Rationally
Thought/ Reasoning	2. Thinking humanly	3. Thinking Rationally
Behavior/ Actions	1. Acting Humanly →Turing Test	4. Acting Rationally

Universality of Computation

Mathematical Formulation of notion of Computation and Computability

23 June 2012
Turing Centenary

Abstract model of a **computer:**
rich enough to capture
any computational process.
Church-Turing Thesis (1936)

$M = \langle Q, \Gamma, b, \Sigma, \delta, q_0, F \rangle$ where

- Q is a finite, non-empty set of *states*
- Γ is a finite, non-empty set of the *tape alphabet/symbols*
- $b \in \Gamma$ is the *blank symbol* (the only symbol allowed to occu computation)
- $\Sigma \subseteq \Gamma \setminus \{b\}$ is the set of *input symbols*
- $q_0 \in Q$ is the *initial state*
- $F \subseteq Q$ is the set of *final* or *accepting states*.
- $\delta : Q \setminus F \times \Gamma \rightarrow Q \times \Gamma \times \{L, R\}$ is a partia left shift, R is right shift. (A relatively uncommon variant allows set.)

Hypotheses:

1) The brain performs some kind of computation.

2) Thinking is a computational process.

3) The brain is a computer.

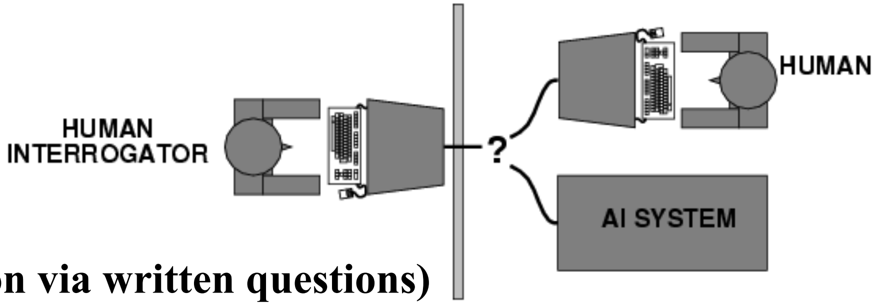
Acting humanly: Turing Test

Alan Turing

Turing (1950) "Computing machinery and intelligence"

"Can machines think?" "Can machines behave intelligently?"

– **Operational test for intelligent behavior: the Imitation Game**



(interaction via written questions)

AI system passes if interrogator cannot tell which one is the machine.

No computer vision or robotics or physical presence required!

Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes. **Achieved. (Siri! ☺)**

But, by scientific consensus, we are still several decades away from truly passing the Turing test (as the test was intended).

Trying to pass the Turing test: Some Famous Human Imitation "Games"

1960s ELIZA

- Joseph Weizenbaum
- Rogerian psychotherapist

1990s ALICE

Loebner prize

- win $100,000 if you pass the test

Still, passing Turing test is of somewhat questionable value.

Because, deception appears *required* and *allowed*!

Consider questions: Where were you born? How tall are you?

ELIZA:
impersonating a Rogerian psychotherapist

1960s ELIZA Joseph Weizenbaum

You: Well, I feel sad

Eliza: Do you often feel sad?

You: not very often.

Eliza: Please go on.

☺

Recent alternative

See: The New Yorker, August 16, 2013
Why Can't My Computer Understand Me?
Posted by *Gary Marcus*

http://www.newyorker.com/online/blogs/elements/2013/08/why-cant-my-computer-understand-me.html

Discusses alternative test by Hector Levesque:
http://www.cs.toronto.edu/~hector/Papers/ijcai-13-paper.pdf

The New York Times

News item --- Big Data vs. Semantics

Bits

Looking to the Future of Data Science

By STEVE LOHR

AUGUST 27, 2014

“Current knowledge bases are full of facts,” Mr. Etzioni observed, “but they are surprisingly knowledge poor.”

The “big” in big data tends to get all the attention, Mr. Etzioni said, but thorny problems often reside in a seemingly simple sentence or two. He showed the sentence: “The large ball crashed right through the table because it was made of Styrofoam.” He asked, What was made of Styrofoam? The large ball? Or the table? The table, humans will invariably answer. But the question is a conundrum for a software program, Mr. Etzioni explained, because the correct answer involves both grammar and background knowledge. And the latter is something humans acquire through experience of the world.

Link NYT

2. Thinking Humanly

	Human-like Intelligence	"Ideal" Intelligent/ Rationally
Thought/ Reasoning	2. Thinking humanly → **Cognitive Modeling**	Thinking Rationally
Behavior/ Actions	Acting Humanly →Turing Test	Acting Rationally

Thinking humanly: modeling cognitive processes

Requires scientific theories of internal activities of the brain.

1) **Cognitive Science (top-down) computer models + experimental techniques from psychology**

   → Predicting and testing behavior of human subjects

2) **Cognitive Neuroscience (bottom-up)**

   → Direct identification from neurological data

Distinct disciplines but especially 2) has become very active. Connection to AI: Neural Nets. (Large Google / MSR / Facebook AI Lab efforts.)

Neuroscience: The Hardware

The brain

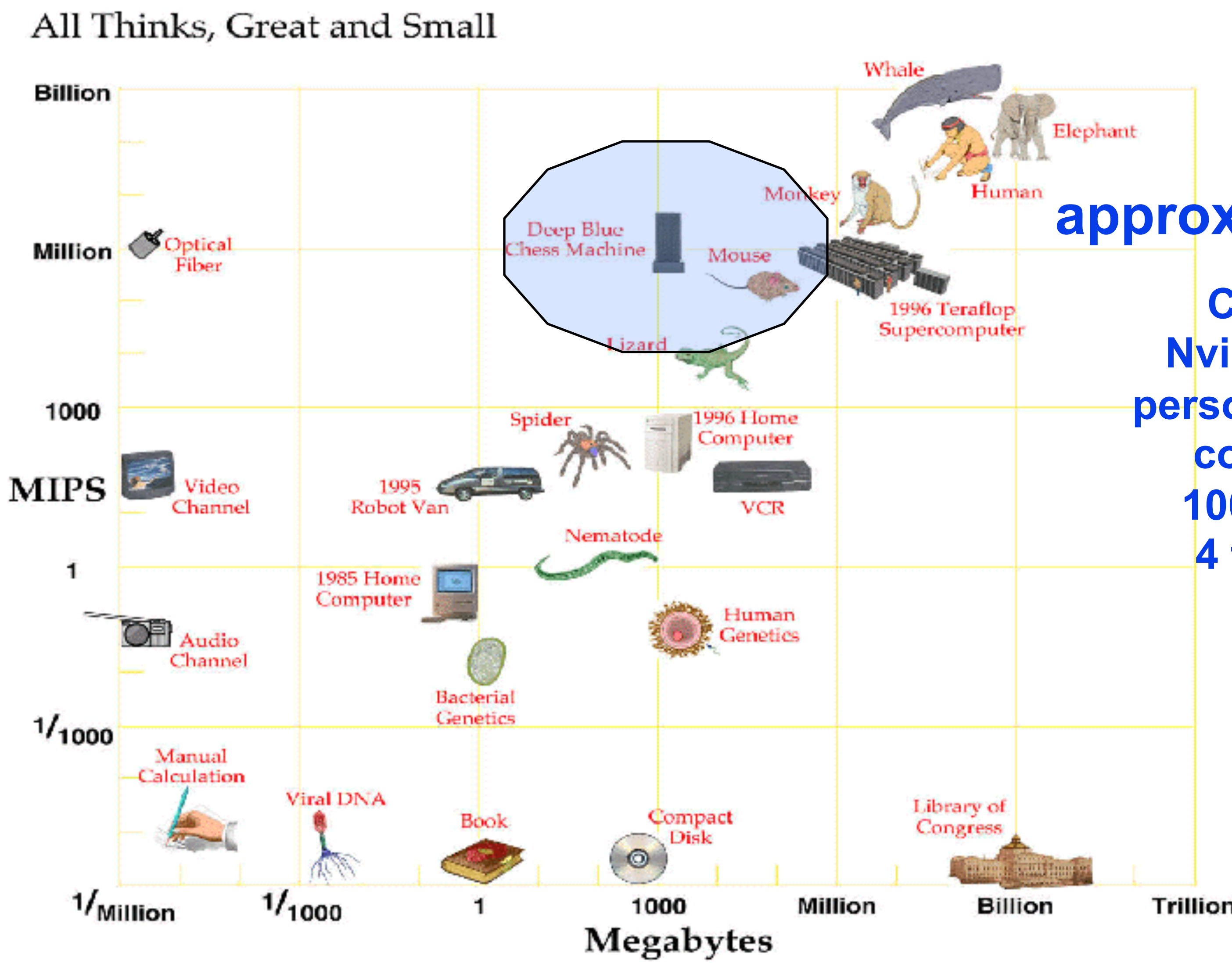
- a neuron, or nerve cell, is the basic information
- processing unit (10^11)
- many more synapses (10^14) connect the neurons
- cycle time: 10^(-3) seconds (1 millisecond)

How complex can we make computers?

- 10^9 or more transistors per CPU
- Ten of thousands of cores, 10^10 bits of RAM
- cycle times: order of 10^(-9) seconds

Numbers are getting close! Hardware will surpass human brain within next 20 yrs.

Computer vs. Brain



approx. 2025

Current:
Nvidia: tesla personal super-computer
1000 cores
4 teraflop

Aside: Whale vs. human brain

So,

- In near future, we can have computers with as many processing elements as our brain, but:

  far fewer interconnections (wires or synapses)

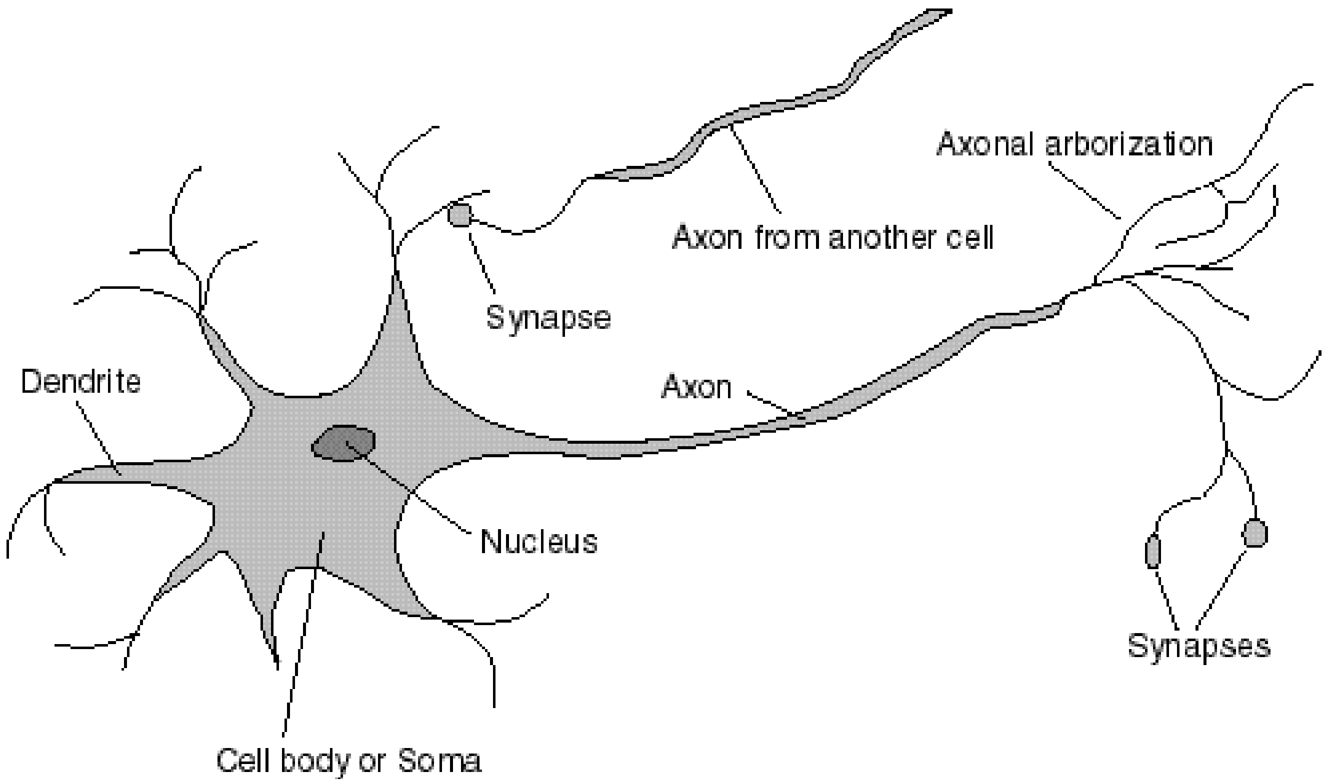
  then again, much faster updates.

Fundamentally different hardware may require fundamentally different algorithms!

- Still an open question.
- Neural net research.
- Can a digital computer simulate our brain?

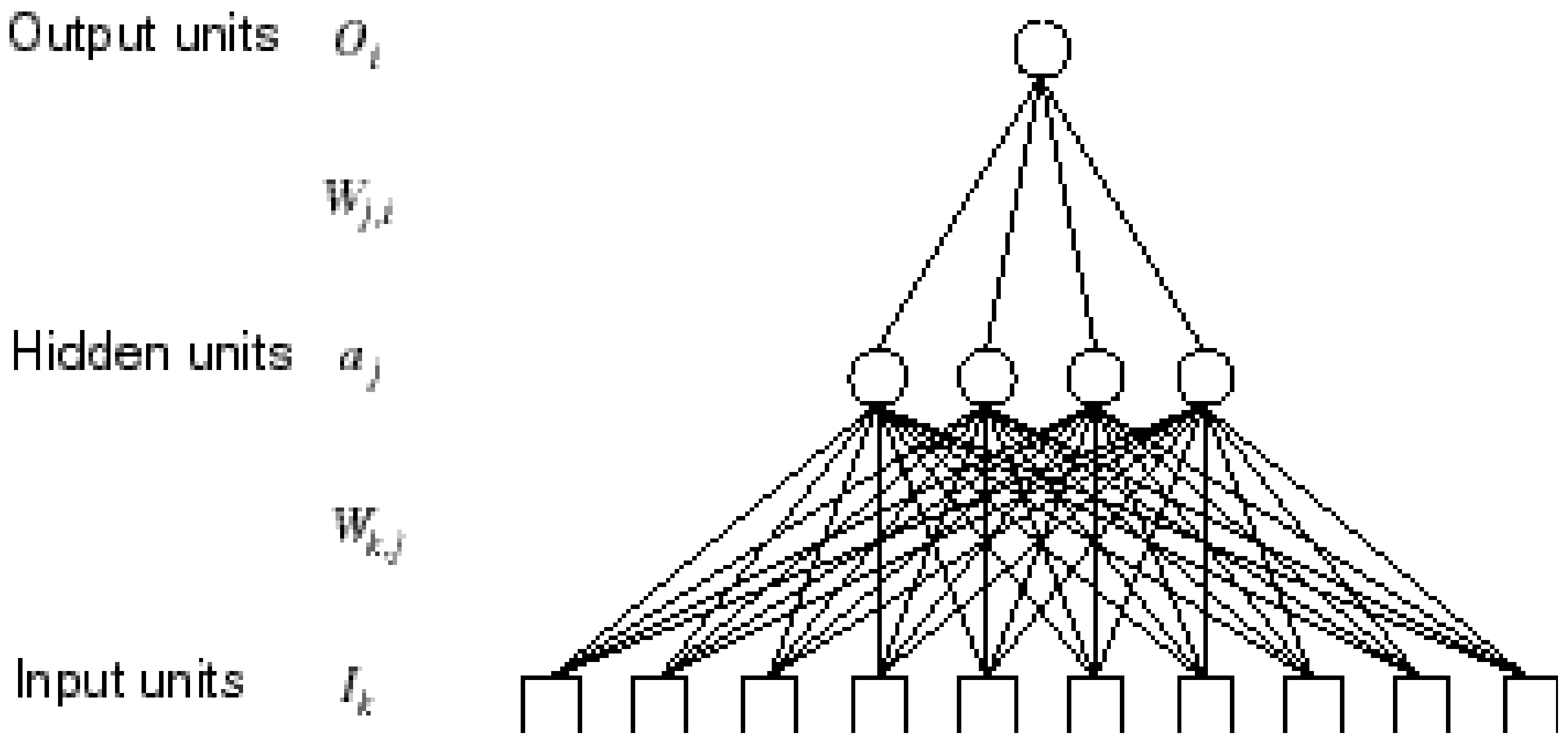
Likely: Church-Turing Thesis
(But, might we need quantum computing?)
(Penrose; consciousness; free will)

A Neuron


Axonal arborization
Axon from another cell
Synapse
Dendrite
Axon
Nucleus
Synapses
Cell body or Soma

An Artificial Neural Network (Perceptrons)

Output Unit



Input Units

An artificial neural network is an abstraction (well, really, a "drastic simplification") of a real neural network.

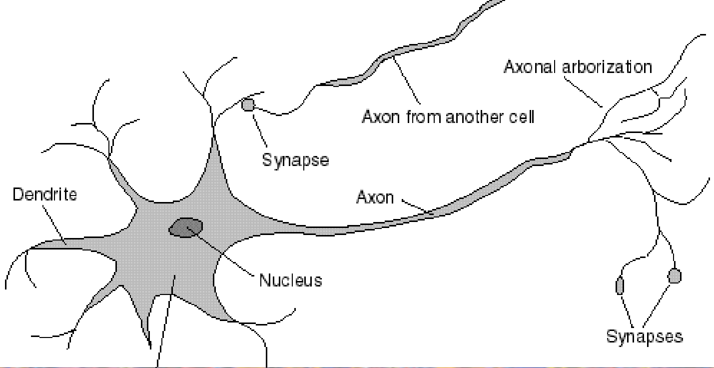
Start out with random connection weights on the links between units. Then train from input examples and environment, by changing network weights.

Recent breakthrough: Deep Learning

(automatic discovery of "deep" features by a large neural network.)

Deep learning is bringing perception (hearing & vision) within reach.

Neurons in the News



The Human Brain Project

European investment: 1B Euro (yeap, with a “b” ☺)

http://www.humanbrainproject.eu/introduction.html

“*… to simulate the actual working of the brain.* Ultimately, it will attempt to simulate the complete human brain.”

http://www.newscientist.com/article/dn23111-human-brain-model-and-graphene-win-sciences-x-factor.html

Bottom-line: Neural networks with machine learning techniques are providing new insights in to how to achieve AI. So, studying the brain seems to helps AI research.

Obviously?
Consider the following *gedankenexperiment*.

1) Consider a laptop running "something." You have no idea what the laptop is doing, although it is getting pretty warm… ☺

2) I give you voltage and current meter and microscope to study the chips and the wiring inside the laptop. Could you figure out what the laptop was doing?

3) E.g. is it running a quicksort or merge sort? *Could studying the running hardware ever reveal that?*

Seems difficult…
It's the challenge of neuroscience.

So, consider I/O behavior as an information processing task.

This is a general strategy driving much of current AI:
Discover underlying *computational process that mimics desired I/O behavior.*

E.g.
In: 3, -4, 5 , 9 , 6, 20   Out: -4, 3, 5, 6, 9, 20
In: 8, 5, -9, 7, 1, 4, 3   Out: -9, 1, 3, 4, 5, 7, 8

Now, consider hundreds of such examples.

A machine learning technique, called Inductive Logic Programming, can uncover a sorting algorithm that provides this kind of I/O behavior. So, it learns the underlying information processing task. (Also, Genetic Genetic programming.)

But, sorting numbers doesn't have much to do with general intelligence… However many related scenarios.

E.g., consider the area of activity recognition and planning.

Setting: **A robot observes a human performing a series of actions.**
Goal: ***Build a computational model of how to generate such action sequences for related tasks.***

Concrete example domain: **Cooking. Goal: Build household robot.**
Robot observe a set of actions (e.g., boiling water, rinsing, chopping, etc.). Robot can learn which actions are required for what type of meal.

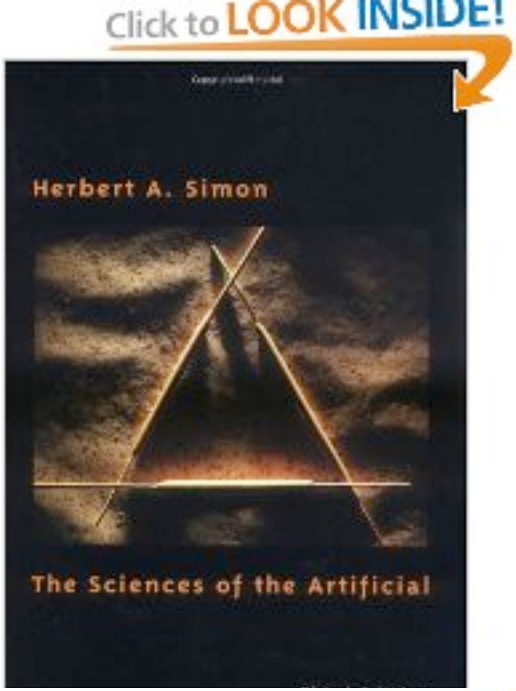
But, how do we get the right sequence of actions?
Certain orderings are dictated by domain, e.g. "fill pot with water, before boiling." Knowledge-based component (e.g. learn).

But how should robot decide on actions that can be ordered in different ways? Is there a *general* principle to do so?

Answer: Yes, minimize time for meal preparation.

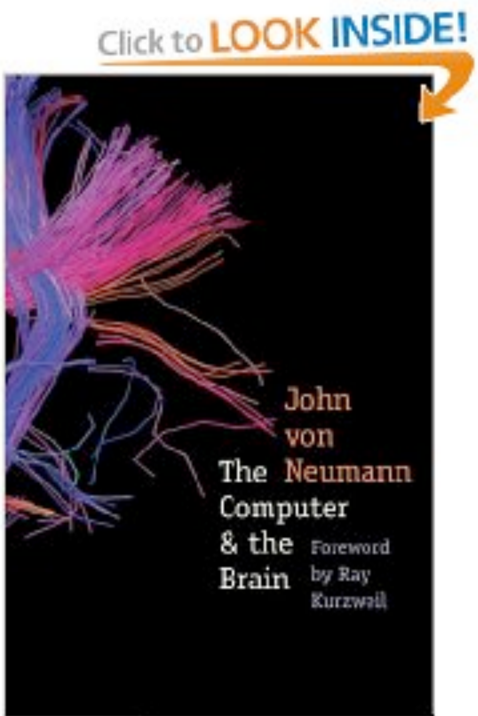
Planning and scheduling algorithms will do so. Works quite well *even though but we have no idea of how a human brain actually creates such sequences*. I.e., we viewed the task of generating the sequence of actions as an information processing task optimizing a certain objective or "utility" function (i.e., the overall duration). AI: We want to discover such principles!

General area: sequential decision making in uncertain environments. (Markov Decision Processes.)
Analogously: Game theory tells us how to make good decision in multi-agent settings. Gives powerful game playing agents (for chess, poker, video games, etc.).



Wonderful (little) book:
The Sciences of the Artificial
by Herb Simon

One of the founders of AI. Nobel Prize in economics. How to build decision making machines operating in complex environments. Theory of Information Processing Systems. First to move computers from "number crunchers" (fancy calculators) to "symbolic processing."



Another absolute classic:
The Computer and the Brain
by John von Neumann.

Renowned mathematician and the father of modern computing.

3. Thinking Rationally

	Human-like Intelligence	"Ideal" Intelligent/ Rationally
Thought/ Reasoning	Thinking humanly → Cognitive Modeling	**3. Thinking Rationally →formalizing "Laws of Thought"**
Behavior/ Actions	Acting Humanly →Turing Test	Acting Rationally

Thinking rationally: formalizing the "laws of thought"

Long and rich history!

Logic: Making the right inferences!

Remarkably effective in science, math, and engineering.

Several Greek schools developed various forms of *logic*: *notation* and *rules of derivation* for thoughts.

Aristotle: what are correct arguments/thought processes? (characterization of "right thinking").

Syllogisms
Aristotle

Socrates is a man

All men are mortal

Therefore, Socrates is mortal

Can we mechanize it? (strip interpretation)

Use: legal cases, diplomacy, ethics etc. (?)

More contemporary logicians (e.g. Boole, Frege, and Tarski).

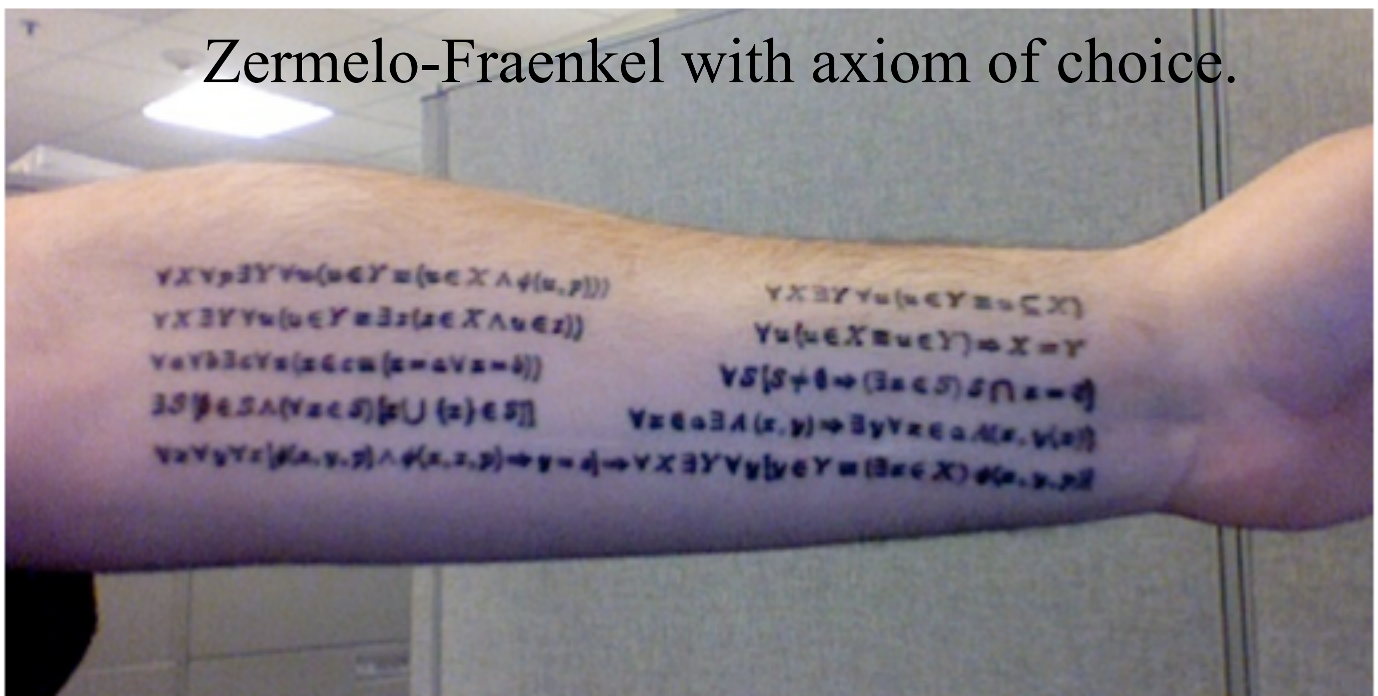
Ambition: **Developing the "language of thought."**

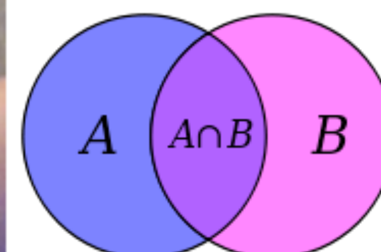
Direct line through mathematics and philosophy to modern AI.

Key notion:

***Inference derives new information from stored facts*.**

Axioms can be very compact. E.g. much of mathematics can be derived from the logical axioms of Set Theory.





Also, Godel's incompleteness.

Limitations:

- **Not all intelligent behavior is mediated by logical deliberation (much appears not…)**

- **(Logical) representation of knowledge underlying intelligence is quite non-trivial. Studied in the area of "knowledge representation." Also brings in probabilistic representations. E.g. Bayesian networks.**

- **What is the purpose of thinking?**
- **What thoughts should I have?**

4. Acting Rationally

	Human-like Intelligence	"Ideal" Intelligent/ Rationally
Thought/ Reasoning	Thinking humanly → Cognitive Modeling	Thinking Rationally →formalizing "Laws of Thought"
Behavior/ Actions	Acting Humanly →Turing Test	**Acting Rationally**

Rational agents

- **An agent is an entity that perceives and acts in the world (i.e. an "autonomous system" (e.g. self-driving cars) / physical robot or software robot (e.g. an electronic trading system))**

  This course is about designing rational agents

- **For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance**

- Caveat: computational limitations may make perfect rationality unachievable
  → design ***best*** program for given machine resources

Building Intelligent Machines

I Building exact models of human cognition

view from psychology, cognitive science, and neuroscience

II Developing methods to match or exceed human performance in certain domains, possibly by very different means

Main focus of current AI.

But, I) often provides inspiration for II). Also, Neural Nets blur the separation.

Key research areas in AI

Problem solving, planning, and search --- generic problem solving architecture based on ideas from cognitive science (game playing, robotics).

Knowledge Representation – to store and manipulate information (logical and probabilistic representations)

Automated reasoning / Inference – to use the stored information to answer questions and draw new conclusions

Machine Learning – intelligence from data; to adapt to new circumstances and to detect and extrapolate patterns

Natural Language Processing – to communicate with the machine

Computer Vision --- processing visual information

Robotics --- Autonomy, manipulation, full integration of AI capabilities