AI :PHILOSOPHY & ETHICS
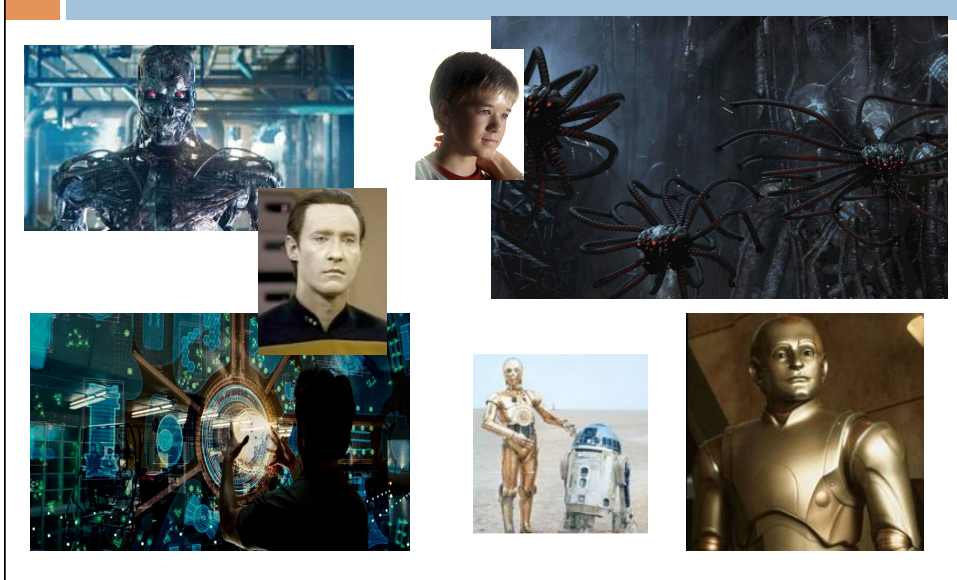
Today

- Reading
  - *Skim AIMA Chapter 26*

- Goals
  - Back to the beginning: what is AI?
  - Philosophy: Is it possible to create true AI?
  - Ethics: What are the ethical concerns in creating AI?

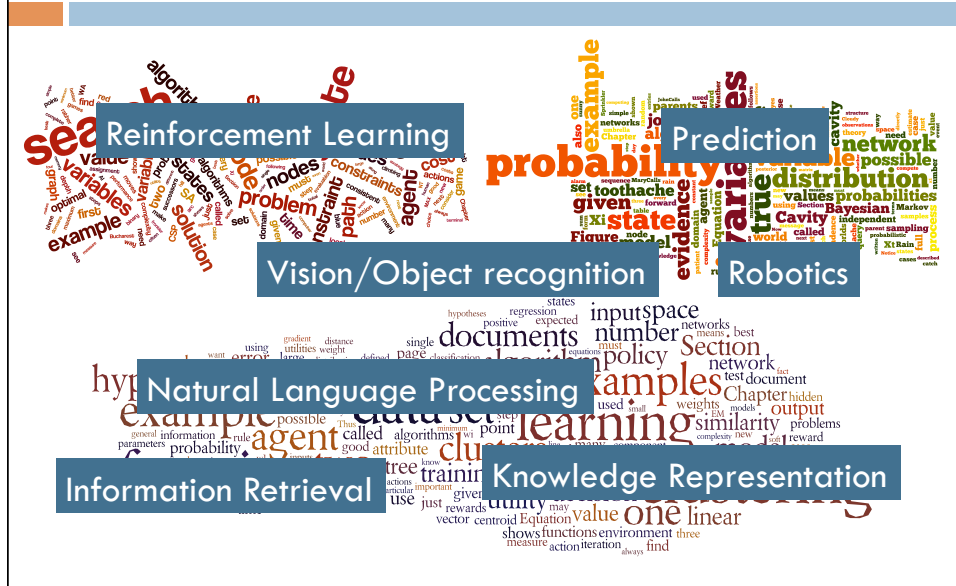# Artificial Intelligence: hopes and fears

Artificial Intelligence: reality

Artificial Intelligence: reality

Reinforcement Learning

Prediction

Vision/Object recognition

Robotics

Natural Language Processing

Information Retrieval

Knowledge Representation

Artificial Intelligence: reality
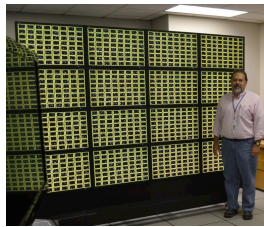
Robotics

Artificial Brains

A.L.I.C.E.

AI according to CS 151

"create a computer system that is capable of learning and reasoning like a human"

"making machines mimic what human beings are doing every day"

"overarching and general even when its sample size is 1"

"A computational model of the human brain"

What is "artificial intelligence"?

"solve problems that it has not been explicitly 'told' how to solve"

"an autonomous…problem-solving system that is designed…rather than evolved"

"respond to stimuli and situations that its creator did not envision it being in"

"ability of a machine to communicate and understand human language…combined with the ability to learn from 'experience'"

"a machine that independently makes choices based on some rational algorithm"

AI according to CS 151

Versatile, fast, inference algorithm

Handling new, novel, foreign experiences

Knowledge representation and common sense

Understanding the human brain first

What will be the biggest challenge?

Learning from mistakes (self-modifications)

Natural language understanding

Filtering out unimportant information

Earning people's trust and putting our confidence in new AI advances

AI according to CS 151

"unlikely"

"at least another 80-120 years to figure out the really deep complexities"

"I think it is very likely!"

"likely"

"not very likely"

"absolutely"

"50/50 chance"

How likely are you to see the creation of true artificial intelligence?

"we've already seen artificial intelligence"

"non-zero probability that we see a human conscious transplanted into a humanoid robot in the next 100 years"

"we absolutely have the potential to map out the human brain and create advanced AI well within my lifetime…the only thing I expect could delay its arrival is fear"

"beyond the capacity of the technology we will see in the next 50-100 years"

Weak AI vs. Strong AI

- Weak Artificial Intelligence
  - It is possible to create machines that act intelligent. Beyond that, who cares?
  - Simulated intelligence

- Strong Artificial Intelligence
  - Machines that do so are in fact intelligent
  - Real intelligence

Weak AI vs. Strong AI

"Surely, it may be said, however closely a computer might model the outward manifestations of human thought processes, the machine would not really be thinking, since that involves being able to have subjective, inner processes."
–Computers and Thought

"The question of whether *Machines Can Think*…is about as relevant as the question of whether *Submarines Can Swim*" -*Dijkstra*

"Not until a machine could write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain – that is, not only write it but know that it had written it" – G. Jefferson (1949)

Chinese Room (Searle 1980)

Ethics of Artificial Intelligence

- What are the ethical concerns surrounding artificial intelligence?

  - What are the concerns or issues highlighted in the story?

  - To what extent are these issues already present in today's society?

  - What, if anything, should be done to address or prevent these concerns?