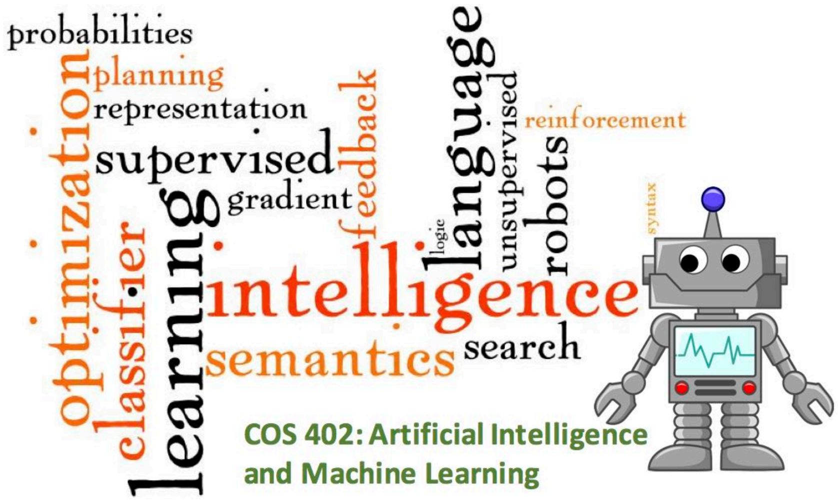


COS 402 – Machine Learning and Artificial Intelligence Fall 2016

Lecture 1: Intro

Sanjeev Arora

Elad Hazan

PRINCETON UNIVERSITY

Today's Agenda

- Defining intelligence and AI state-of-the-art, goals
- Course outline
- AI by introspection

Admin

- Course webpage (staff, policies, reading, etc.)
  https://www.cs.princeton.edu/courses/archive/fall16/cos402/

- Exercises – announced in class.
  Written part due one week later in class;
  programming part by CS dropbox.
  Learn Python!

- Minimal Blackboard usage.
  Register for Piazza forum.

Faculty Lead Preceptor:
Dr. Xiaoyan Li

TA:
Karan Singh

Admin

- Grading:
  - Final exam - 35%
  - Exercises (theory + programming) 50%
  - Midterm (lesson 13) – 15%
  - Bonus questions in class

- Movie, Oct 5th, 7:30 pm
  "Ex-Machina"
  Discussion panel w. Prof. Hasson @ PNI
  Garden theater, free/discounted tickets

Defining intelligence: by example

- Playing games:
  - Chess ("deep blue" beats world champion 1997), Checkers (game solved 2007), Go (world champion beaten 2016)
  - Atari games (Google deep mind)
- Driving a car (Tesla, Google, MobileEye)
- Recognize faces / songs / movies
- Prove mathematical theorems
- Play "jeopardy", translate (Google translate, IBM)

AI examples

- Crossword puzzle solving
  - “proverb” – better than most humans
  - Hard since requires much knowledge
  - Uses web
- Web search
  - Algorithmics/optimization (PageRank, Hubs and Authorities)
- House-cleaning robots, warfare droids
  - Robotics, control

Where are we today?

Examples:

Huge variety, successful.

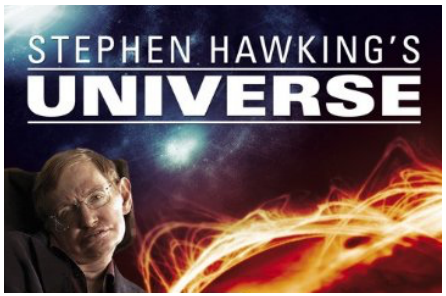
Limited domain, specific well-defined tasks

Definite progress (accelerated in recent years)

We cannot (yet):

1. converse on a human level
2. Translate as good as a human

The hype and the fear…



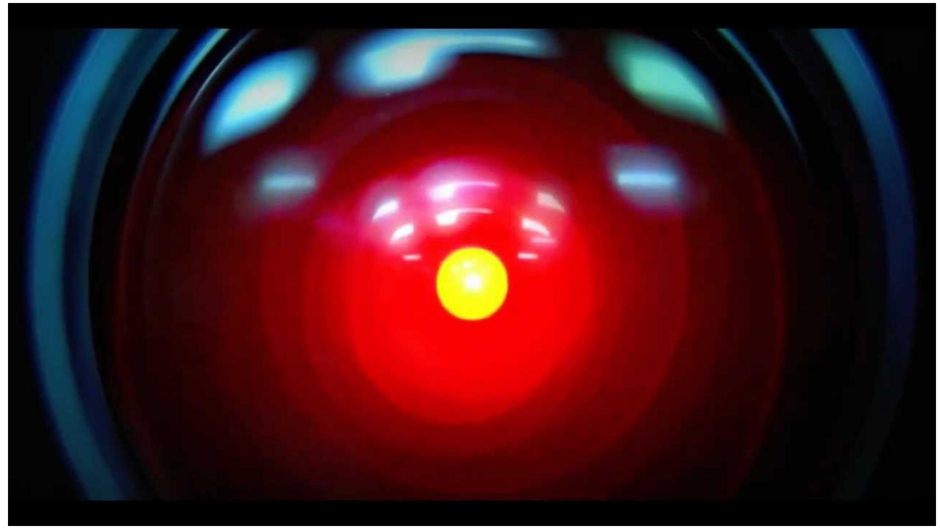
"The development of full artificial intelligence could spell the end of the human race."
"It would take off on its own, and re-design itself at an ever increasing rate. Humans, who are limited by slow biological evolution, couldn't compete, and would be superseded."

"In the movie "Terminator," they didn't create A.I. to—they didn't expect, you know some sort of "Terminator"-like outcome. It is sort of like the "Monty Python" thing: Nobody expects the Spanish inquisition. It's just—you know, but you have to be careful."

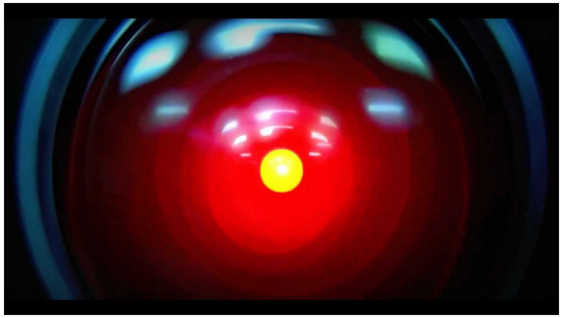
Where are we going?

" I am in the camp that is concerned about super intelligence. First the machines will do a lot of jobs for us and not be super intelligent. That should be positive if we manage it well. A few decades after that though the intelligence is strong enough to be a concern. I agree with Elon Musk and some others on this and don't understand why some people are not concerned."

“strong”
vs.
“weak” AI

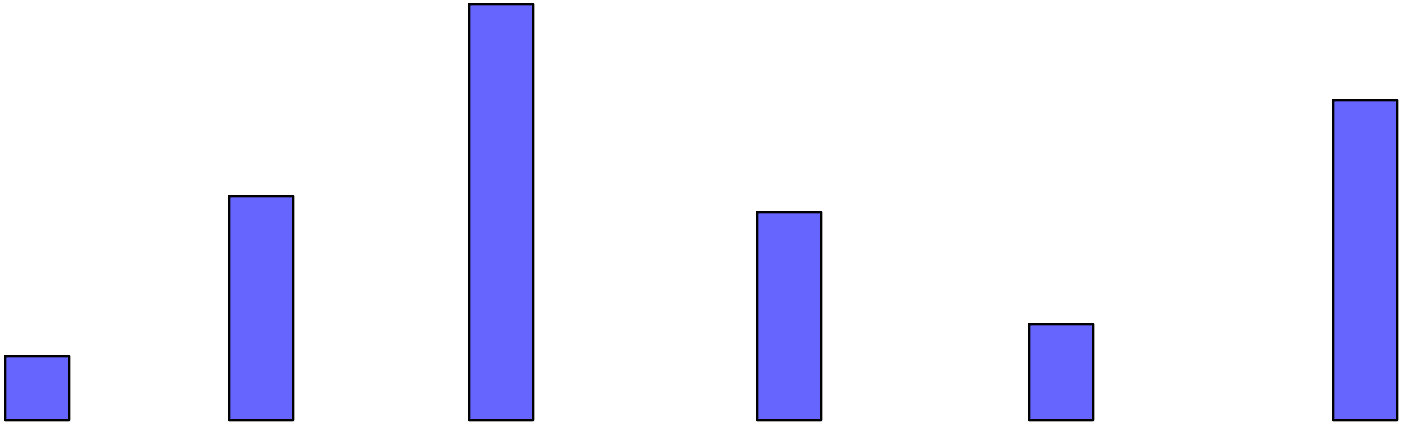


How long to get to



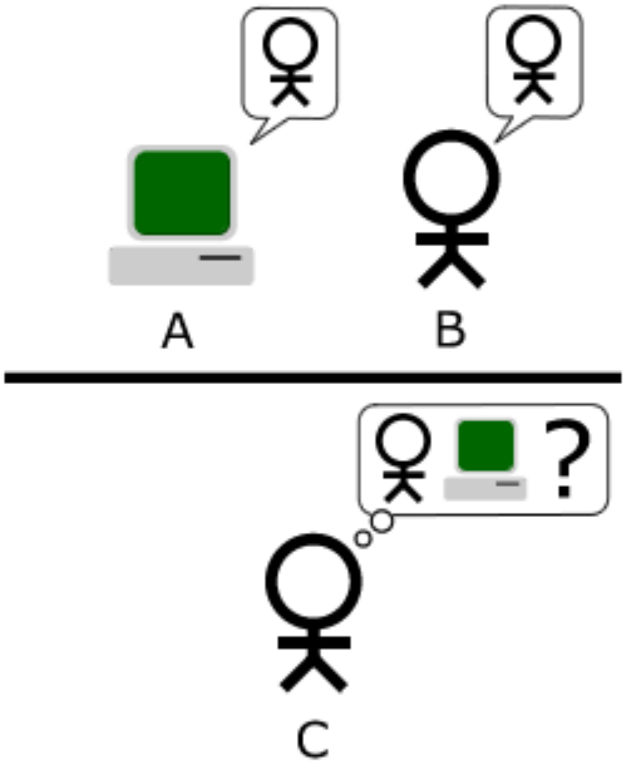
In years:

<10 25 50 100 1000+ ∞



Turing test: a functional, relativistic test of intelligence

test intelligence: if a person, interacting with the machine via text terminal, cannot distinguish it from a human [Turing 1950]



Various extensions (i.e. vision signal, robotic/mechanical component, etc.)

- No commitment to use brain methodology!
- Most computer scientists think this test will be passed by a machine within your lifetime, and possibly even within mine

This course: Basic principles of how to design machines/programs that act "intelligently."

- Will not try to debate what is or is not "intelligence." (Well, maybe a bit as and when needed)

- Suffices for machine/program to be useful (to us) and supplement our abilities.

"Is this story funny or sad?"
"Find images that contain an elephant at a picnic."
"Drive my car while I do my homework. Avoid obstacles, pedestrians, and other cars."

Key idea

Formalize "intelligence" as acting to maximize some mathematical "objective."
(= programmer's formalization of "correct" behavior)

- Find shortest path in a network
- Play tic-tac-toe optimally
- Prove or disprove a theorem

→ Permits algorithmic/optimization theory

Q:
(a) how to define such an objective
(b) how to optimize it
(c) how to show it leads to desired "intelligent" behavior

Example: Given an image, decide if it contains a chair

AI phase 1: 1950s -1990

Programmer defines machine's objective by introspection. ("How would I do it?")

Successes: Chess playing,
Decision Trees for medical diagnosis,
Simple robots, etc.

Defining objective by "introspection" alone proved tricky

Given an image, decide if it contains a chair?

AI phase 2: post-1990

- “Machine learning.” : Design more general purpose algorithms that allow machine to learn from experience (or by studying large data sets of examples).

  - Recently led to better-than-human performance on some tasks (e.g., labeling the content of images)

Largely Open: How to do learning from very few or even a single example.
(How does a human know this is a chair, though it looks unlike any other chair in history?)

Even more open: How can a machine create interesting and strong new chairs like a human designer?

Components of intelligence:

- Language processing
- Planning
- Reasoning
- Learning
- Perception

Course Topics:

1. AI by introspection (“Naive AI”)

2. "What does it mean for a machine to learn from examples?" A definition and its operational realization.

3. Efficient algorithms for learning, optimization.

4. Biological motivated machines: Neural nets for classification and image recognition.

5. Introduction to natural language processing, including using simple recurrent neural nets.

Topics:

- Logical representations, Probabilistic representations of knowledge (HMMs, Bayes nets,)

- Machines that take decisions and search for "rewards" in face of uncertainty (reinforcement learning). self-driving cars, robots etc.

AI by introspection

The intelligent crow

https://youtu.be/AVaITA7eBZE
https://youtu.be/AVaITA7eBZE

How can we formalize the planning and execution of such a sequence of actions?

A classic puzzle
(from 9th century France; variations arose independently in many cultures)

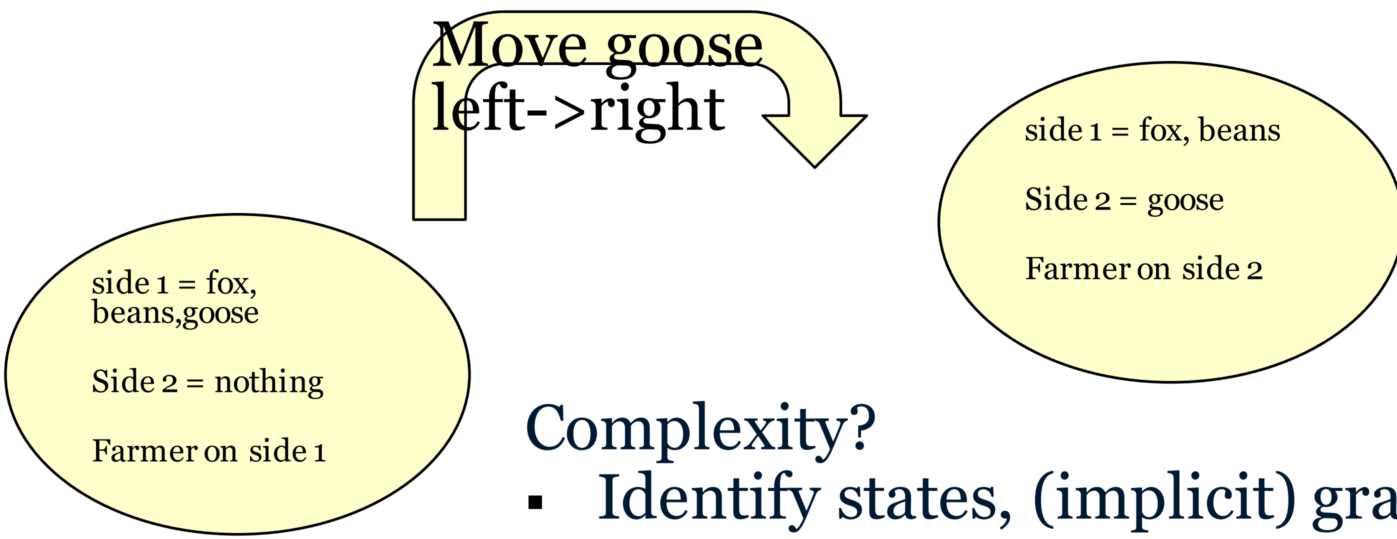
- A farmer must transport a fox, goat and box of beans from one side of a river to another using a boat which can only hold one item.
- The fox cannot be left alone with the goat, and the goat cannot be left alone with the beans.

How to reason out a plan?

problem solving

"Problem graph"

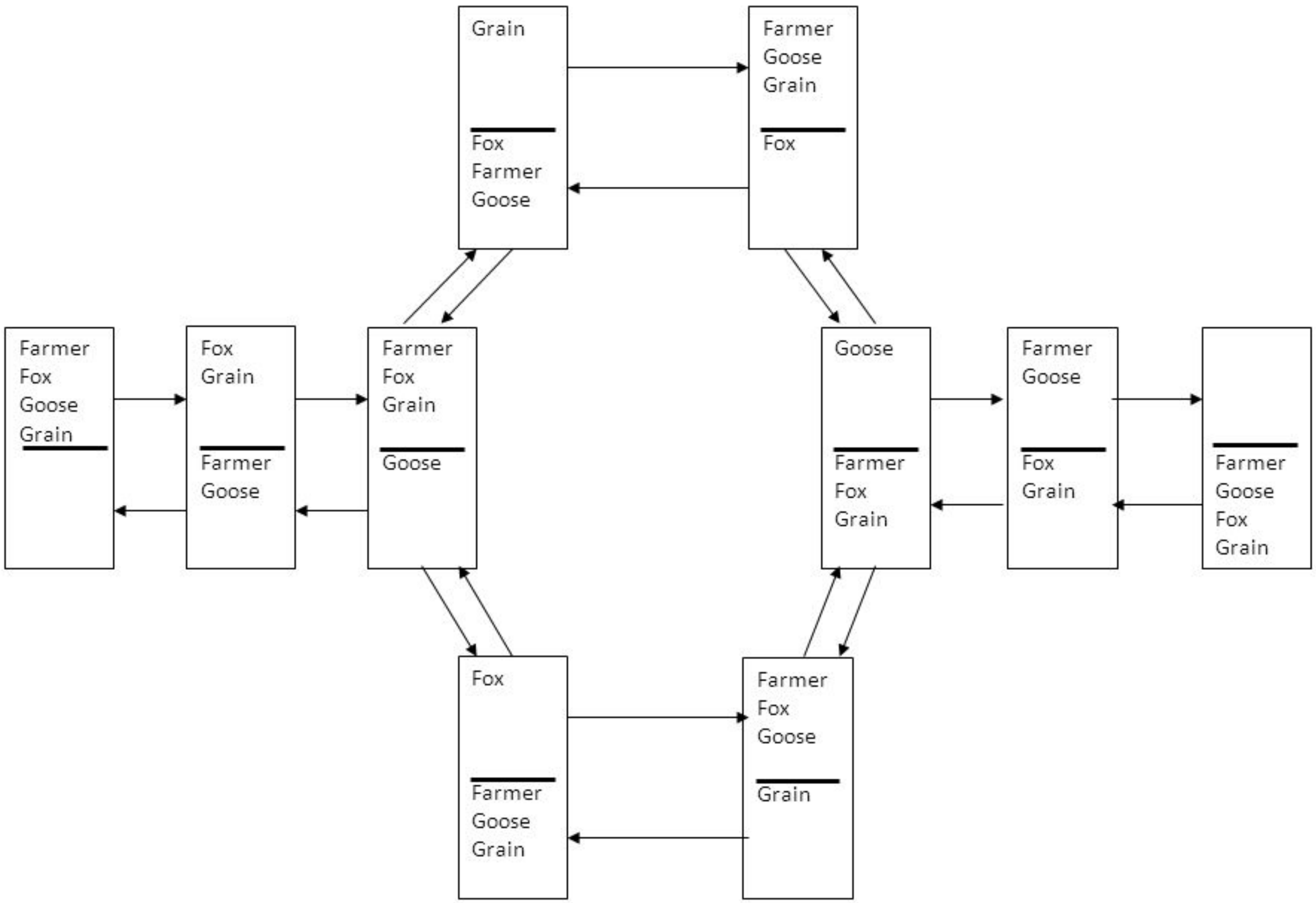
- Node = State = configuration
- Edge = action from state to state
- Solution = path from initial to final



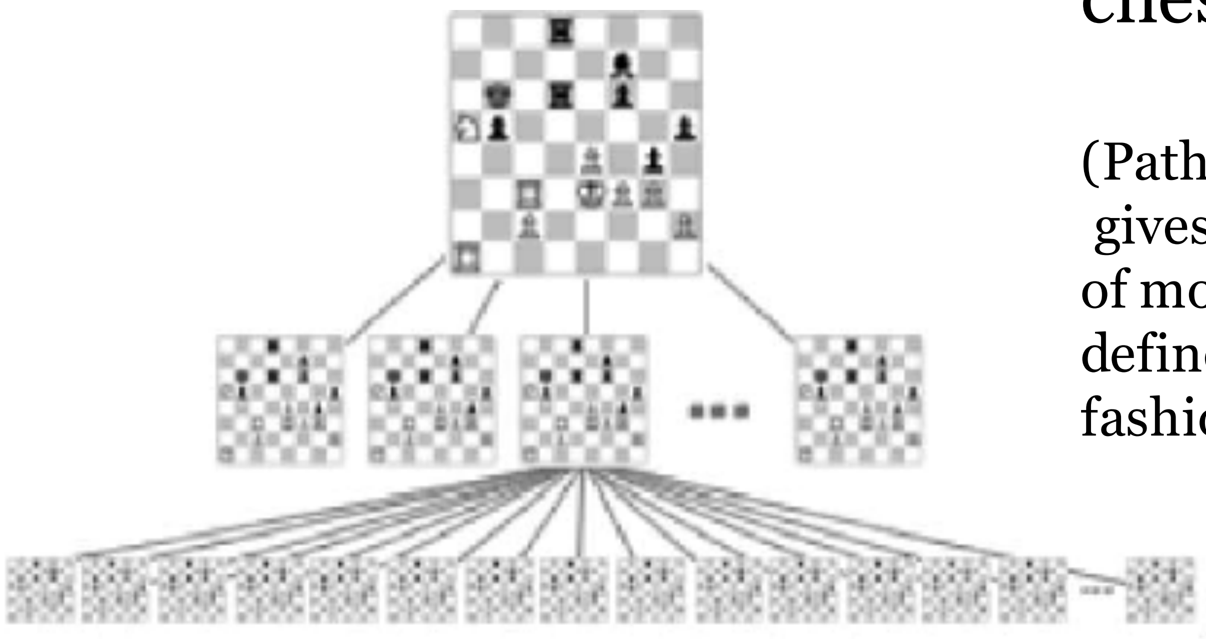
Complexity?

- Identify states, (implicit) graph
- Path finding (search)

The full graph of states and actions



Complicated versions of this are widely used for formal reasoning



(source: Stanford CS221)

Game tree for chess playing

(Path from root to leaf gives possible sequence of moves; “optimum” strategy defined in a bottom-up fashion)

Introspection ignores:

- How to model states?
- Obtaining an accurate representation of the world (perception, vision, scene understanding…)
- Feedback from environment
- Learning by deduction (communication with other humans)
- Next week: a more formal approach