Artificial Intelligence

CS4365 --- Fall 2022

Introduction

Instructor: Yunhui Guo

General Information for CS4365

- Where: CR 1.202
- When: TuTh 5:30PM - 6:45PM
- Instructor: Yunhui Guo
- Office hours: TuTh 1:00PM - 2:00PM   ECSS 4.604
- Teaching Assistant: TBD

- Private communication with course staff outside office hours:
    Via Microsoft Teams

Websites

- Course website: https://yunhuiguo.github.io/UTD-CS4365/

- Homework submission: TBD (Gradescope or eLearning)

- Class Discussion: Teams and eLearning

- Announcements: Course website and eLearning

Course Information

- Text:

  There is no required text, but the following book is highly recommended:

  Artificial Intelligence: A Modern Approach. Russell and

  Norvig, Prentice-Hall, Inc., second or third edition

- Prerequisites:
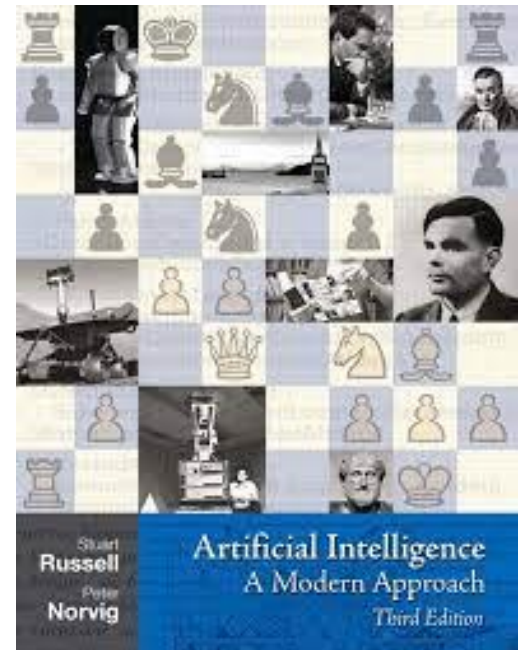
  Ability to program in Python

  Familiarity with the big-O notation

  Elementary knowledge of probability theory

  Knowledge of propositional and first-order logic

- Homework: 5-6 homework assignments

- Examinations: two midterms and one final exam

Tentative Grading Policy

- Assignment: 27%
- Project: 12%
- Midterm 1:   18%
- Midterm 2:   18%
- Final Exam: 25%

For homework assignments, getting 600 of the 700 possible points is sufficient for earning the 27% of the course grade

Course Policies

- Collaboration policy
    - assignments can be done in a group of two
    - project can be done in a group of two
    - need to turn in one solution / program per group
- Late assignment submission policy
    - 6 free late days (2 can be used per homework)
    - one day late   (10% penalty)
    - two days late (30% penalty)
    - no assignments accepted if more than two days late
    - Free late days allow you to submit late without any late penalty within the late submission period. They won't allow you to submit after 48 hours of the official due date.

Course Policies

Test policy

- No make-up exams are possible unless you have medical notes that document your sickness or documentation that you have to be out of town on the date of an exam

Extra credit policy

- No "extra credit" opportunity will be given towards the end of the semester
- If you experience difficulties with the course, you should talk to the instructor and/or the TA early in the semester

Course Policies

- Academic integrity policy
  - If you are caught cheating in a homework assignment, you will get a zero on that assignment *and* the highest grade you will receive in this course will be a C-.
  - If you are caught cheating in an exam, you will get a zero on that midterm *and* the highest grade you will receive in this course will be an F.

  We reserve the right to report cases of academic dishonesty to UTD's Office of Community Standards and Conduct.
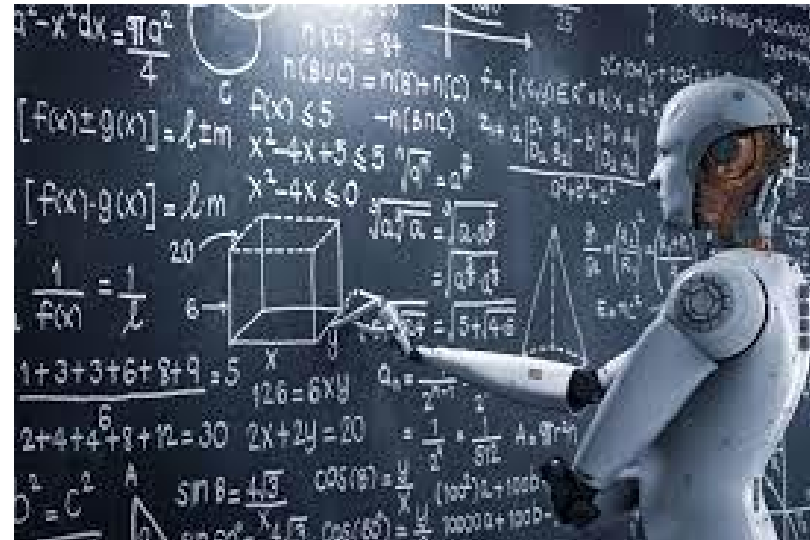
Course Overview

• Introduction to AI (today)

• Search

• Knowledge Representation and Reasoning

• Probabilistic Reasoning

• Game Theory

What is Artifical Intelligence (AI)?

- The science of making "intelligent" machines
  - Searching
  - Reasoning
  - Learning
  - Perception
  - ...

What is Intelligence?

Intelligence

- "the capacity to learn and solve problems" (Webster dictionary)

What is Intelligence?

Intelligence

- "the capacity to learn and solve problems" (Webster dictionary)

- the ability to act rationally

AI: Goals

- **Strong AI**
  - build "intelligent" agents/machines that can understand or learn any intellectual task that a human being can.

AI: Goals

- Strong AI
  - build "intelligent" agents/machines that can understand or learn any intellectual task that a human being can.

- Weak AI
  - designed to solve exactly one problem
    - Face recgonition
    - Virtual assistant
    - ...

Different Perspectives on AI

- Views of AI fall into four different perspectives:

  - Thinking versus Acting
  - Human versus Rational

|	Human-like Intelligence	"Ideal" Intelligent / Rationality
Thought / Reasoning	Thinking Humanly	Thinking Rathionally
Behavior / Actions	Acting Humanly	Acting Rationally

Acting Humanly

- Views of AI fall into four different perspectives:

  - Thinking versus Acting
  - Human versus Rational

|	Human-like Intelligence	"Ideal" Intelligent / Rationality
Thought / Reasoning	Thinking Humanly	Thinking Rathionally
Behavior / Actions	Acting Humanly	Acting Rationally

Turing Test
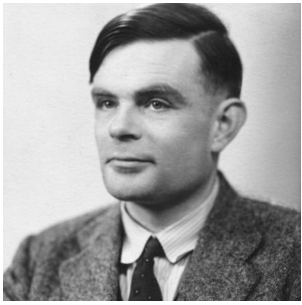
Acting Humanly: Turing Test

- Turing (1950) "Computing machinery and intelligence":
  - Can machines act intelligently?
  - Operational test for intelligent behavior: the Imitation Game

(interaction via written questions)
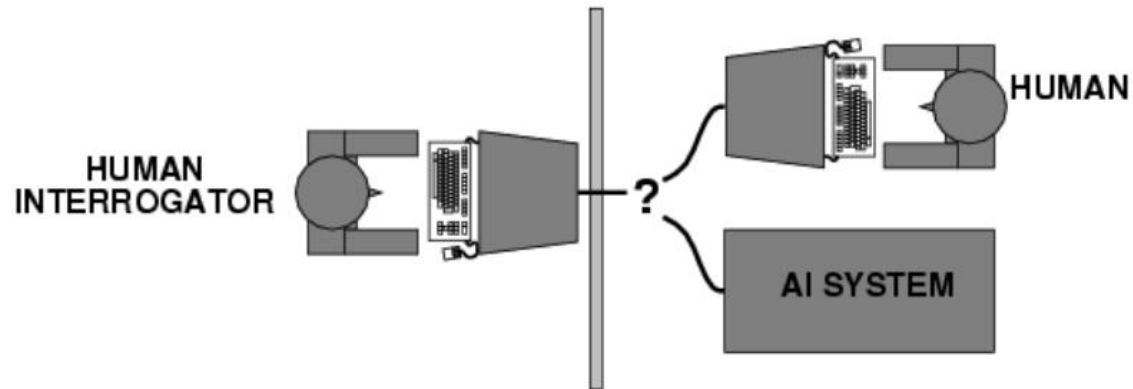
AI system passes if interrogator cannot tell which one is the machine

Example: A.L.I.C.E

• Introduced in 1995

Human: Do you like Microsoft?
ALICE: I want to replace Windows with AI
Human: Did you chat with Jabberwacky?
ALICE: No I didn't chat with Jabberwacky.
Human: It is another chat bot
ALICE: Really. it is?
Human: The news says that you did chat Jabbe
ALICE: Don't believe everything you see on TV.
You say:

Get ALICE Silver Edition with talking
VHost[tm], winner of the 2004
Loebner Prize

Example: A.L.I.C.E

- Introduced in 1995

Human: Do you like Microsoft?
ALICE: I want to replace Windows with AI
Human: Did you chat with Jabberwacky?
ALICE: No I didn't chat with Jabberwacky.
Human: It is another chat bot
ALICE: Really. it is?
Human: The news says that you did chat Jabbe
ALICE: Don't believe everything you see on TV.
You say:

Get ALICE Silver Edition with talking VHost[tm], winner of the 2004 **Loebner Prize**

- Unable to pass the Turing test.
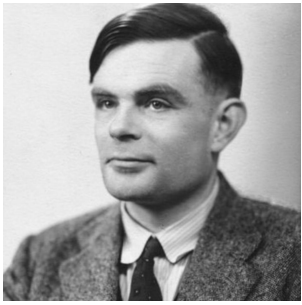
Acting Humanly: Turing Test

- Turing test ⟶ identified key research areas in AI:

Natural Language Processing – to communicate with the machine;

Knowledge Representation – to store and manipulate information;

Automated Reasoning – to use the stored information to answer questions and draw new conclusions;

Machine Learning – to adapt to new circumstances and to detect and extrapolate patterns.

Acting Humanly: Turing Test

- In 2014

A computer program called Eugene Goostman, which simulates a 13-year-old Ukrainian boy, is said to have passed the Turing test at an event organised by the University of Reading.
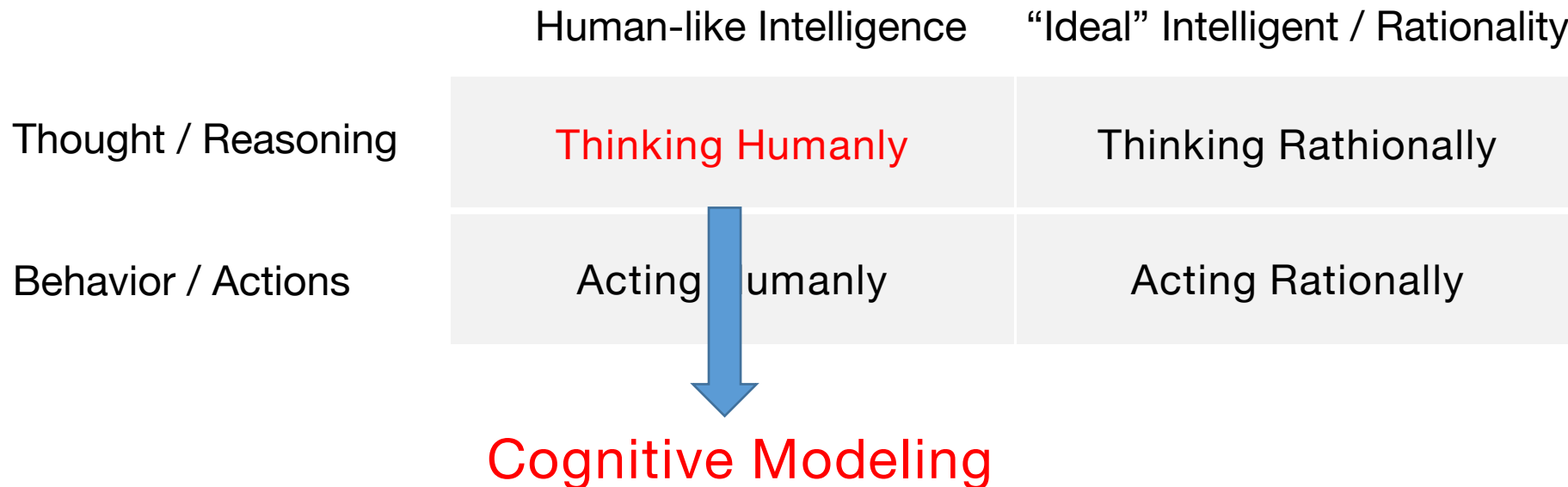
- In 2018,

TECHNOLOGY

Google's AI passed a famous test — and showed how the test is broken

- Not really. Only in the domain of booking appointments

Thinking Humanly

- Views of AI fall into four different perspectives:

  - Thinking versus Acting
  - Human versus Rational

	Human-like Intelligence	"Ideal" Intelligent / Rationality
Thought / Reasoning	Thinking Humanly	Thinking Rathionally
Behavior / Actions	Acting Humanly	Acting Rationally

Cognitive Modeling

Thinking Humanly:
Modeling Cognitive Processes

• Requires scientific theories of internal activities of the brain;

1) Cognitive Science: computer models + experimental techniques from psychology

    -> Predicting and testing behavior of human subjects

2) Cognitive Neuroscience

    -> Direct identification from neurological data

Thinking Humanly:
Modeling Cognitive Processes

• Requires scientific theories of internal activities of the brain;

1) Cognitive Science: computer models + experimental techniques from psychology

    -> Predicting and testing behavior of human subjects

2) Cognitive Neuroscience

    -> Direct identification from neurological data

Both approaches are now distinct from AI

Thinking Rationally

- Views of AI fall into four different perspectives:

  - Thinking versus Acting
  - Human versus Rational

|	Human-like Intelligence	"Ideal" Intelligent / Rationality	
Thought / Reasoning	Thinking Humanly	Thinking Rathionally	⟶ "Laws of Thought"
Behavior / Actions	Acting Humanly	Acting Rationally	

Thinking Rationally:
Formalizing the "Laws of Thought"

- Logic -> Making the right inferences! Several Greek schools developed various forms of logic: notation and rules of derivation for thoughts;
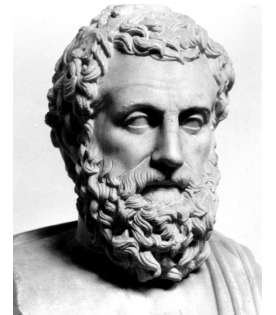
Thinking Rationally:
Formalizing the "Laws of Thought"

- Logic -> Making the right inferences! 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")

Socrates is a man

All men are mortal

Therefore Socrates is mortal

Thinking Rationally:
Formalizing the "Laws of Thought"

- Logic -> Making the right inferences! 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")

Socrates is a man

All men are mortal
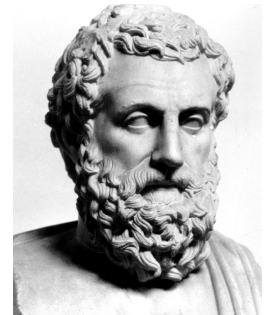
Therefore Socrates is mortal

More contemporary logicians (e.g. Boole, Frege, Tarski) -> Direct line through mathematics and philosophy to modern AI

Limitation:

- Not all intelligent behavior is mediated by logical deliberation

Acting Rationally

- Views of AI fall into four different perspectives:

  - Thinking versus Acting
  - Human versus Rational

	Human-like Intelligence	"Ideal" Intelligent / Rationality
Thought / Reasoning	Thinking Humanly	Thinking Rathionally
Behavior / Actions	Acting Humanly	Acting Rationally ➡ Course Pespective

Acting Rationally: Rational Agent

- Rational behavior: doing the right thing

- The right thing: that which is expected to maximize goal achievement, given the available information
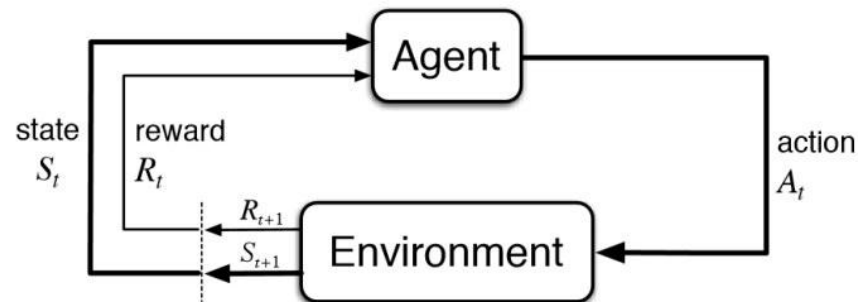
- Doesn't necessarily involve thinking – e.g., blinking reflex – but thinking should be in the service of rational action

Building Intelligent Machine

I Building exact models of human cognition view from psychology and cognitive science

Building Intelligent Machine

I Building exact models of human cognition view from psychology and cognitive science

II Developing methods to match or exceed human performance in certain domains, possibly by very different means -> e.g., Deep Blue;

Focus of this course

Methodology of AI

- Theoretical aspects
  - Mathematical formalizations, properties, algorithms

- Engineering aspects
  - The act of building (useful) machines

- Empirical science
  - Experiments

AI Leverages from Different Disciplines

- Philosophy
  - e.g., foundational issues in logic, methods of reasoning, mind as physical system, foundations of learning, language, rationality

- Computer science and engineering
  - e.g., complexity theory, algorithms, logic and inference, programming languages, and system building (hardware and software).

- Mathematics and physics
  - e.g., probability theory, statistical modeling, continuous mathematics, statistical physics, and complex systems

AI: More Direct Influence

Obtaining an understanding of the human mind is one of the final frontiers of modern science.

George Boole, Gottlob Frege, and Alfred Tarski
formalizing the laws of human thought
Alan Turing, John von Neumann, and Claude Shannon
thinking as computation

Direct Founders:
John McCarthy, Marvin Minsky, Herbert Simon, and Allen Newell
the start of the field of AI

History of AI: Milestones
The gestation of AI 1943-1956

$$x_0 = +1$$
$$x_1 \quad w_{k0} = b_k$$
$$x_2 \quad w_{k1}$$
$$x_3 \quad w_{k2} \quad v_k \quad \varphi(\cdot) \quad y_k$$
$$x_m \quad w_{km}$$

- 1943 : McCulloch and Pitts: artificial neurons

- 1950 : Turing's "Computing machinery and intelligence"

- 1950s Early AI programs, including Samuel's checkers program, Newell and Simon's Logic theorist

- 1956 Dartmouth meeting : Birth of "Artificial Intelligence"

  - The 4 founders: John McCarthy, Marvin Minsky, Herbert Simon, and Allen Newell
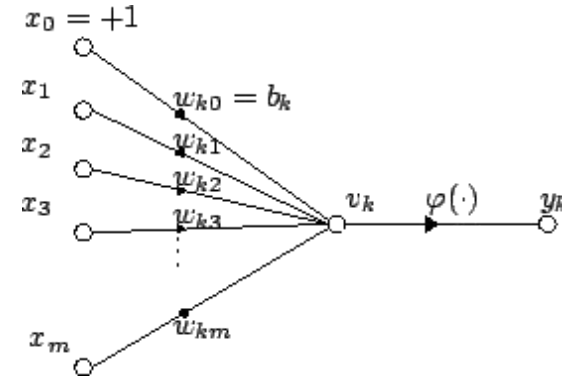
John McCarthy
Father of AI

Alan Turning
Mathematician &
Philosopher

Marvin Minsky
Cognitive Scientist &
Co-Founder of AI

Allen Newell
computer science &
cognitive psychology

Herbert Simon
Cognitive Psychologist

History of AI: Early enthusiasm, great expectations (1952-1969)

- 1957 Herb Simon:

It is not my aim to surprise or shock you – but the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create. Moreover, their ability to do these things is going to increase rapidly until – in the visible future – the range of problems that they can handle will be coextensive with the range to which human mind has been applied.

1958 : John McCarthy's LISP

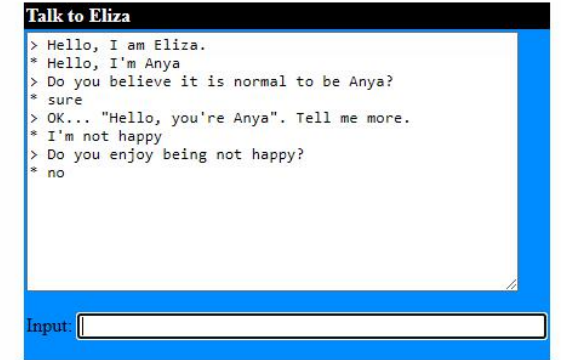1965 : J.A. Robinson invents the resolution principle, basis for automated theorem proving

Intelligent reasoning in Microworlds (e.g., Block's world)

The Block's World

Initial State

Goal State
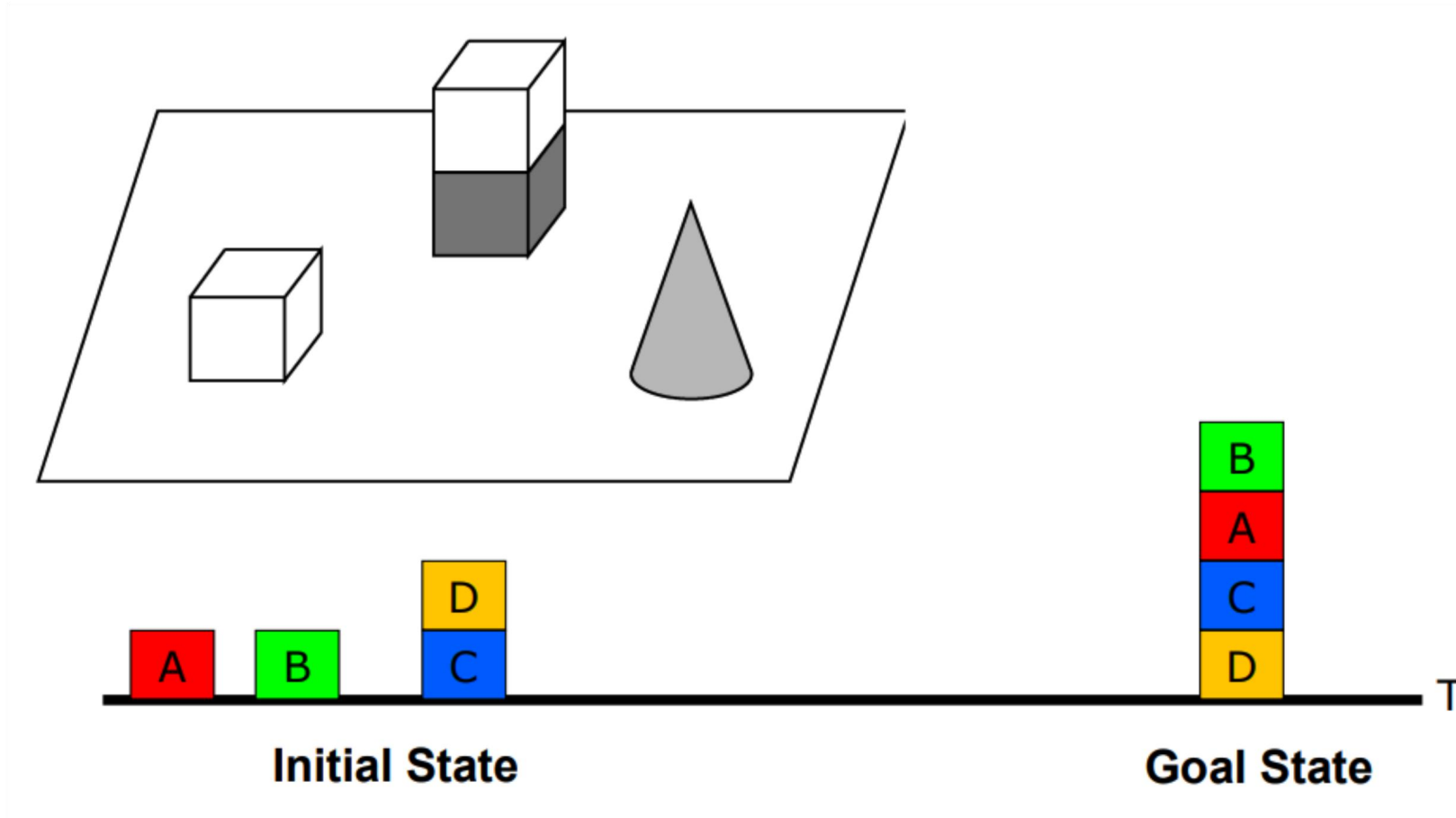
History of AI:
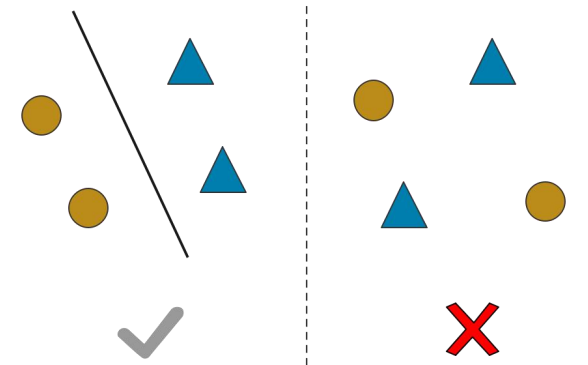# A dose of reality (1966-1978)

1965: Weizenbaum's ELIZA

Limitations of Perceptrons discovered

-> can only represent linearly separable functions

Neural network research almost disappear

-> NP-Completeness (Cook 72)

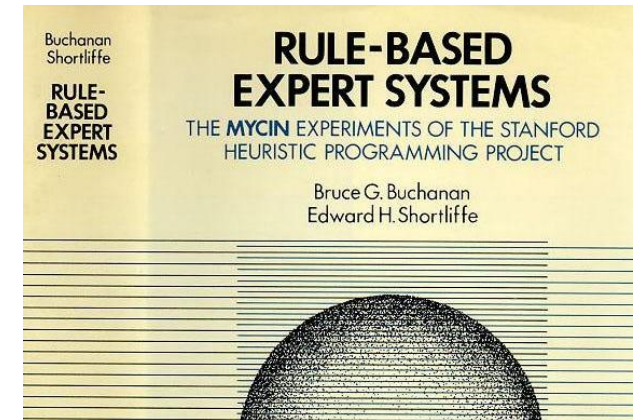   Intractability of the problems attempted by AI

History of AI:
Knowledge based systems (1969-79)

- **Intelligence requires knowledge**
  - Knowledge based systems as opposed to weak methods (general-purpose search methods)

-> Expert systems, e.g.:

- Mycin : diagnose blood infections

- R1 : configuring computer systems

Buchanan
Shortliffe

RULE-
BASED
EXPERT
SYSTEMS

**RULE-BASED
EXPERT SYSTEMS**
THE MYCIN EXPERIMENTS OF THE STANFORD
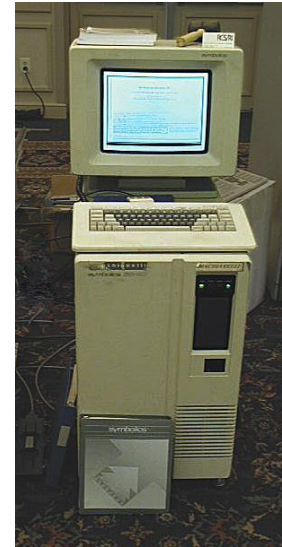HEURISTIC PROGRAMMING PROJECT

Bruce G. Buchanan
Edward H. Shortliffe

R1: an Expert in the Computer Systems Domain[1]

John McDermott
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

History of AI: AI becomes industry (1980-88)

- Expert systems

- Lisp-machines

- Return of Neural Nets

Lisp Machine

-> End of 80's – limitations of expert systems became clear, even though they have been quite successful in certain domains.

History of AI: 2000-
AI is alive and kicking

Current work on "intelligent agents":

Emphasis on integration of reasoning (search and inference as well as probabilistic reasoning), knowledge representation, and learning techniques

AI as a science: Combining theoretical and empirical analysis

-> Mathematical sophistication of AI techniques

AI Achievements

1997: Deep Blue Beats the World Chess Champion

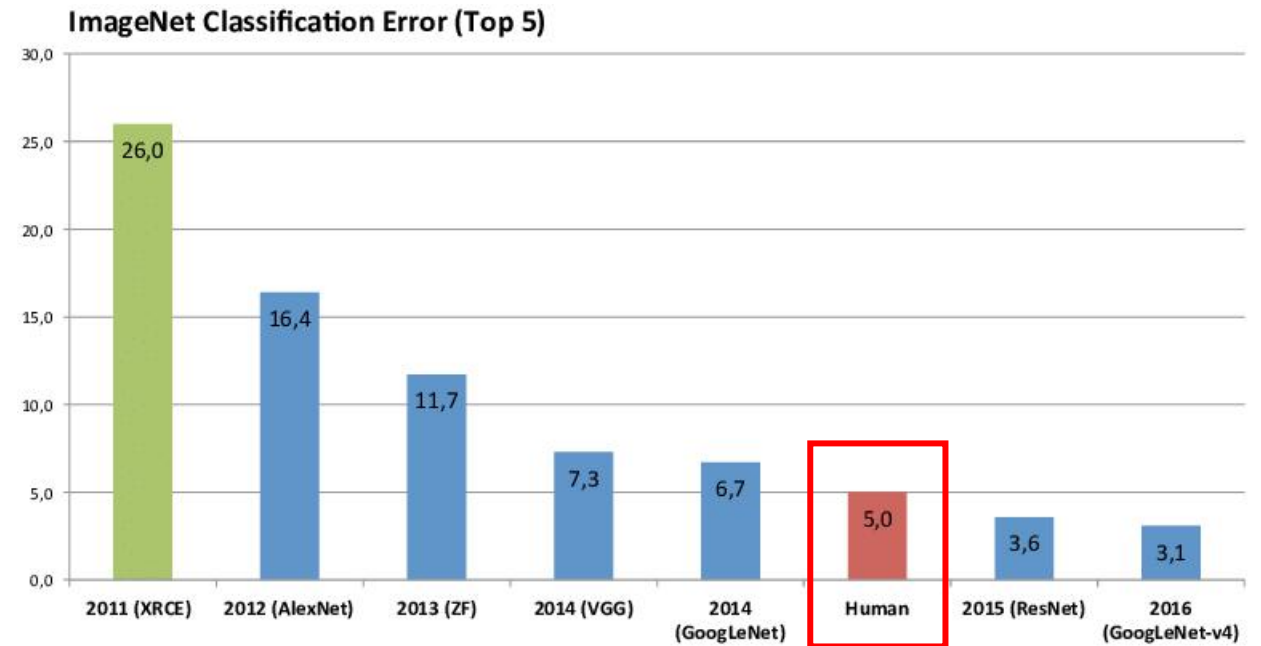Deep Blue had Kasparov in deep thought
(CNN)

vs.

I could feel human-level intelligence across the room
- Gary Kasparov, World Chess Champion

Vision (Perception)

- Beyond human-level recognition accuracy

ImageNet Dataset

ImageNet Classification Error (Top 5)

Year	Error
2011 (XRCE)	26,0
2012 (AlexNet)	16,4
2013 (ZF)	11,7
2014 (VGG)	7,3
2014 (GoogLeNet)	6,7
Human	5,0
2015 (ResNet)	3,6
2016 (GoogLeNet-v4)	3,1

Natural Language Processing

• Beyond human-level read comprehension

Article: Endangered Species Act
Paragraph: " ... *Other legislation followed, including the Migratory Bird Conservation Act of 1929, a 1937 treaty prohibiting the hunting of right and gray whales, and the Bald Eagle Protection Act of 1940. These later laws had a low cost to society—the species were relatively rare—and little opposition was raised.*"

Question 1: *"Which laws faced significant opposition?"*
Plausible Answer: *later laws*

Question 2: *"What was the name of the 1937 treaty?"*
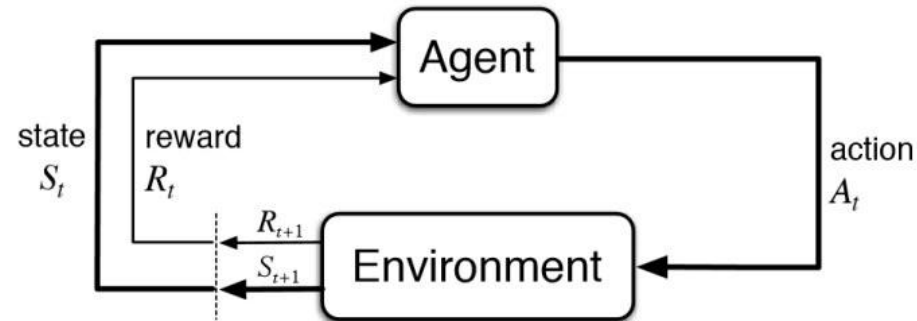**Plausible Answer:** *Bald Eagle Protection Act*

Figure 1: Two unanswerable questions written by crowdworkers, along with plausible (but incorrect) answers. Relevant keywords are shown in blue.

SQuAD1.1 Leaderboard

Rank	Model	EM	F1
	Human Performance *Stanford University* (Rajpurkar et al. '16)	82.304	91.221
1 Oct 05, 2018	BERT (ensemble) *Google AI Language* https://arxiv.org/abs/1810.04805	87.433	93.160
2 Sep 09, 2018	nlnet (ensemble) *Microsoft Research Asia*	85.356	91.202
3 Jul 11, 2018	QANet (ensemble) *Google Brain & CMU*	84.454	90.490

Rajpurkar, Pranav, Robin Jia, and Percy Liang. "Know what you don't know: Unanswerable questions for SQuAD." arXiv preprint arXiv:1806.03822 (2018).

Reinforcement Learning

Factors in Accelerated Progress

- Crowd-sourced human data

- Computional resources

- Investments in AI systems are being scaled-up by an order of magnitude (to billions)

What We Can't Do Yet

Need deeper semantics of natural language
Commonsense knowledge and reasoning

Example: "The large ball crashed right through the table because it was made of Styrofoam."
What does "it" refer to? The large ball or the table?

vs: "The large ball crashed right through the table because it was made of steel."

Commonsense is needed to deal with unforeseen cases.

Societal Issues

The emergence of intelligent autonomous machines among us is expected to have a major impact on society

• Societal issues:

    1) Economics (wealth inequality) & Employment
    2) AI Safety & Ethics
    3) Military Impact (Smart autonomous weapon systems)
    4) The Future: Super-Intelligence? Living with smart machines.

Goals of This Course

- To introduce you to the kinds of problems studied in AI.

- To introduce you to a set of key methods and techniques from AI, in the areas of search, reasoning, and learning

- To teach you about the applicability and limitations of these methods.

Reading: Chapter 1, R&N.