Foundations of Artificial Intelligence

1. Introduction: What is Artificial Intelligence?

Malte Helmert and Thomas Keller

University of Basel

February 17, 2020

Introduction: Overview

Chapter overview: introduction

1. What is Artificial Intelligence?
2. AI Past and Present
3. Rational Agents
4. Environments and Problem Solving Methods

What is AI?

What is AI?

What do we mean by artificial intelligence?

$\rightsquigarrow$ no generally accepted definition!

often pragmatic definitions:

- "AI is what AI researchers do."
- "AI is the solution of hard problems."

in the following: some common attempts at defining AI

What is AI: Humanly vs. Rationally; Thinking vs. Acting

"[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 & McDermott, 1985)
"the study of how to make computers do things at which, at the moment, people are better" (Rich & Knight, 1991)	"the branch of computer science that is concerned with the automation of intelligent behavior" (Luger & Stubblefield, 1993)

four typical categories:

thinking humanly	thinking rationally
acting humanly	acting rationally

What is AI: Humanly vs. Rationally; Thinking vs. Acting

"[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 & McDermott, 1985)
"the study of how to make computers do things at which, at the moment, people are better" (Rich & Knight, 1991)	"the branch of computer science that is concerned with the automation of intelligent behavior" (Luger & Stubblefield, 1993)

four typical categories:

thinking humanly	thinking rationally
acting humanly	acting rationally

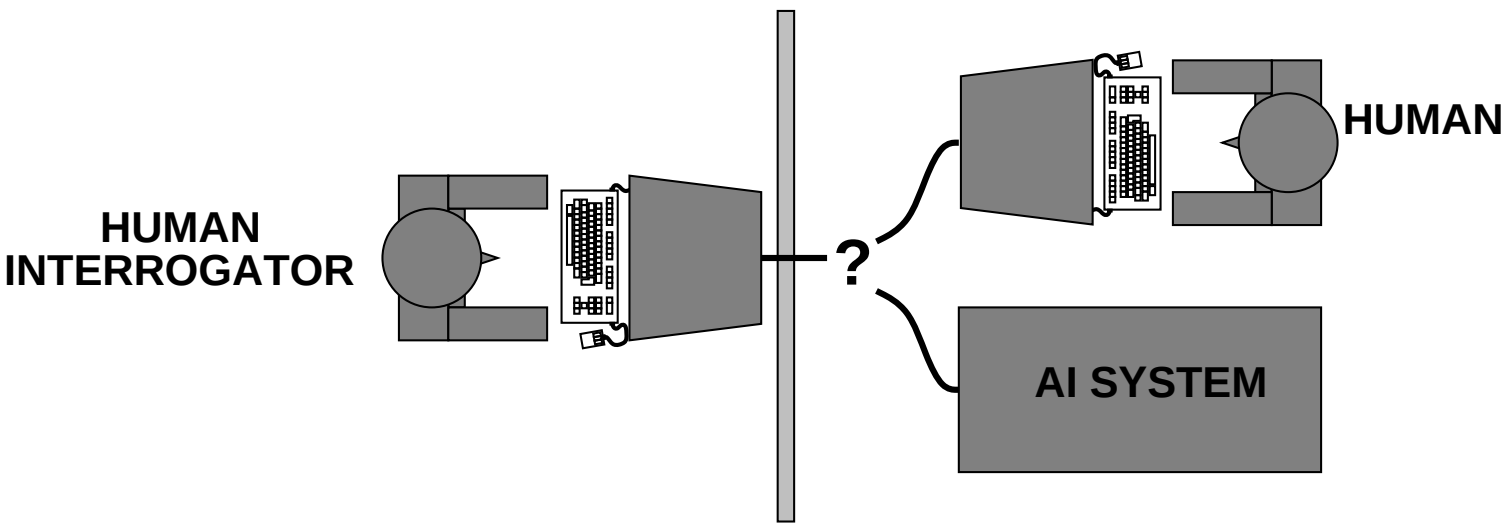
$\rightsquigarrow$ here (and most widespread these days): acting rationally

Acting Humanly

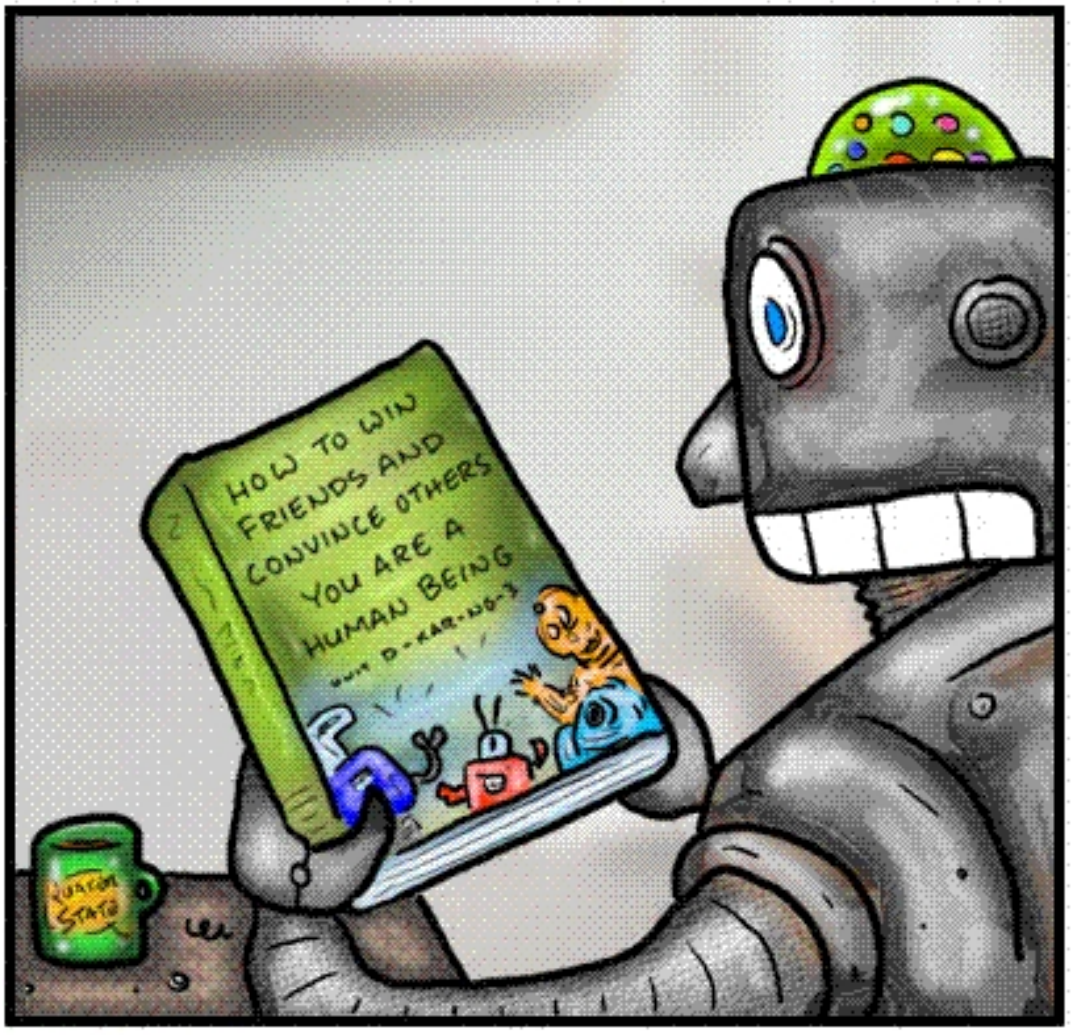
Acting Humanly: the Turing Test

Alan Turing, Computing Machinery and Intelligence (1950):

- from "can machines think?"
  to "can machines act intelligently?"
- operationalization: the imitation game



Cartoon



Unit Bob crams for his Turing Test.

How Useful is the Turing Test?

Turing Test:

- scientific usefulness is questionable
- not important in AI "mainstream"
- but: annual competitions (Loebner Prize): `https://en.wikipedia.org/wiki/Loebner_Prize`
- practical application: CAPTCHA ("Completely Automated Public Turing test to tell Computers and Humans Apart")

Turing and the Turing Test in Cinema

https://www.imdb.com/title/tt2084970/

More Turing Test in Cinema

https://www.imdb.com/title/tt0470752/

Turing's "Computing Machinery and Intelligence"

Turing's Computing Machinery and Intelligence:

- already discussed all important arguments of the 20th century against possibility of AI
- suggested core aspects of AI: knowledge representation, reasoning, language understanding, learning
- prediction: in the year 2000, a machine will be able to fool a layperson for 5 minutes with 30% probability
- in the news: `https://www.engadget.com/2014/06/08/supercomputer-passes-turing-test/`

Thinking Humanly

Thinking Humanly: Cognitive Science

- cognitive revolution of the 1960s: information processing supplants dominant behaviorism in psychology
- Which cognitive abilities are necessary for intelligent behavior?
- requires scientific theory of brain activity
  - ⇝ which level of abstraction? "knowledge" or "circuits"?
- How to test? Requires
  - prediction/test of human behavior (top-down) or
  - identification from neurological data (bottom-up)
- roughly corresponds to cognitive science and cognitive neuroscience
  - today separate research areas from AI

Thinking Rationally

Thinking Rationally: Laws of Thought

- normative (prescriptive) rather than descriptive
- Aristotle: What are correct arguments/modes of thought?
- syllogisms: structures for arguments that always yield correct conclusions given correct premises:
  - Socrates is a human.
  - All humans are mortal.
  - Therefore Socrates is mortal.
- several Greek schools of thought developed various forms of logic:
  - notations (syntax) and
  - derivation rules (calculi) for "correct" thinking
- direct connection to modern AI via mathematical logic (early 20th century)

Problems of the Logical Approach

problems:

- not all intelligent behavior stems from logical thinking
- Which conclusions are relevant?
- How to deal with uncertainty?
- How to deal with contradictions?

Acting Rationally

Acting Rationally

acting rationally: "doing the right thing"

- the right thing: maximize utility
  given available information
- does not necessarily require "thought" (e.g., reflexes)

advantages of AI as design of rational agents:

- more general than "laws of thought":
  - ⇝ logical inference only one mechanism
    for obtaining rational behavior
- better suited for scientific method
  than approaches based on acting/thinking humanly

Summary

Summary

What is AI? $\rightsquigarrow$ many possible definitions

- guided by humans vs. by utility (rationality)
- based on externally observable actions or inner thoughts?

$\rightsquigarrow$ four combinations:

- acting humanly: e.g., Turing test
- thinking humanly: cf. cognitive science
- thinking rationally: logic
- acting rationally: the most common view today
  $\rightsquigarrow$ amenable to scientific method