

# Foundations of Artificial Intelligence

## A2. Introduction: What is Artificial Intelligence?

Malte Helmert

University of Basel

February 26, 2024

# Introduction: Overview

## Chapter overview: introduction

- A1. Organizational Matters
- A2. What is Artificial Intelligence?
- A3. AI Past and Present
- A4. Rational Agents
- A5. Environments and Problem Solving Methods

# What is AI?

# What is AI?

What do we mean by **artificial intelligence**?

↪ no generally accepted definition!

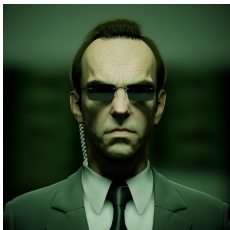
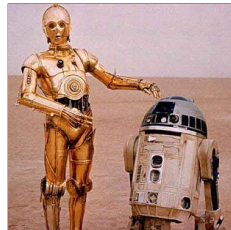
often pragmatic definitions:

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

**in this chapter:** some common attempts at defining AI

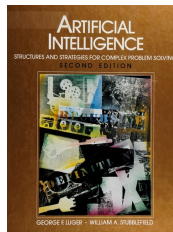
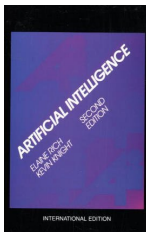
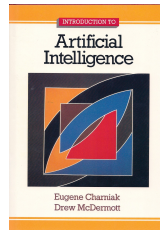
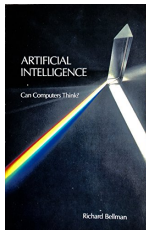
# What Do We Mean by Artificial Intelligence?

what **pop culture** tells us:



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

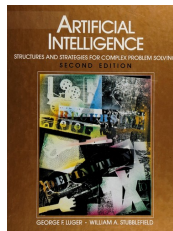
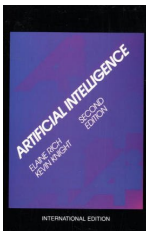
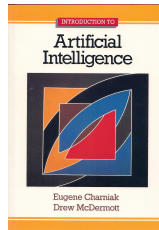
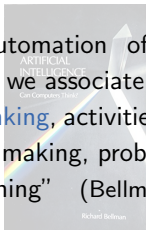
what **scientists** tell us:



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

what **scientists** tell us:

“[the automation of] activities that we associate with **human thinking**, activities such as decision-making, problem solving, learning” (Bellman, 1978)



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

what **scientists** tell us:

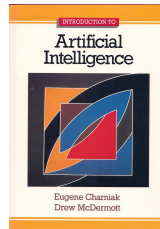
 <p>thinking like humans</p>	
	



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

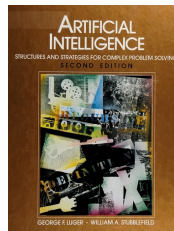
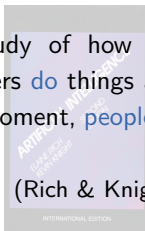
what **scientists** tell us:

thinking like humans



“the study of how to make computers **do** things at which, at the moment, **people** are better”

(Rich & Knight, 1991)



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

what **scientists** tell us:

 <p>thinking like humans</p>	
 <p>acting like humans</p>	

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

what **scientists** tell us:

 <p>thinking like humans</p>	 <p>“the study of <b>mental</b> faculties through the use of <b>computational models</b>”  (Charniak &amp; McDermott, 1985)</p>
 <p>acting like humans</p>	



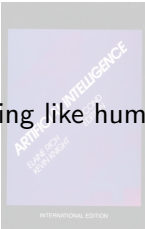
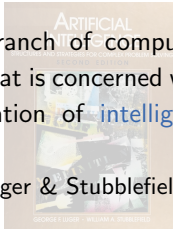
# What is AI: Humanly vs. Rationally; Thinking vs. Acting

what **scientists** tell us:

 <p>thinking like humans</p>	 <p>thinking rationally</p>
 <p>acting like humans</p>	 <p>acting rationally</p>

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

what **scientists** tell us:

 <p>thinking like humans</p>	 <p>thinking rationally</p>
 <p>acting like humans</p>	 <p>“the branch of computer science that is concerned with the automation of <b>intelligent behavior</b>” (Luger &amp; Stubblefield, 1993)</p>

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

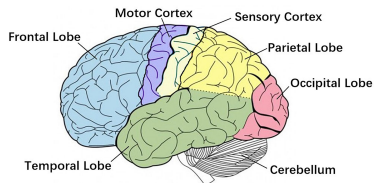
what **scientists** tell us:

 <p>thinking like humans</p>	 <p>thinking rationally</p>
 <p>acting like humans</p>	 <p>acting rationally</p>

# Thinking Like Humans

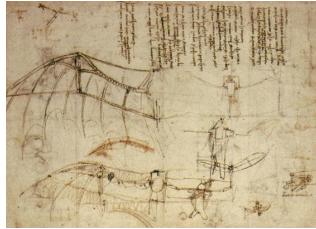
# Cognitive (Neuro-) Science

- requires knowledge of **how humans think**
- two ways to a scientific **theory of brain activity**:
  - **psychological**: observation of human behavior
  - **neurological**: observation of brain activity
- roughly corresponds to **cognitive science** and **cognitive neuroscience**
- today separate research areas from AI





# Machines that Think Like Humans



“brains are to intelligence as wings are to flight”



# What Do We Mean by Artificial Intelligence?

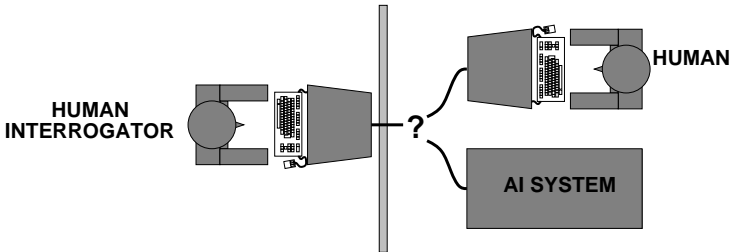
 <p>thinking like humans</p>	 <p>thinking rationally</p>
 <p>acting like humans</p>	 <p>acting rationally</p>

# Acting Like Humans

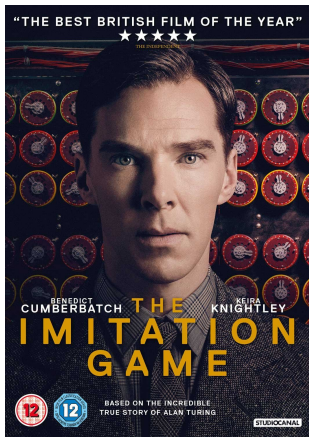
# The Turing Test

Alan Turing, Computing Machinery and Intelligence (1950):

- central question: Can machines think?
- hypothesis: yes, if they can act like humans
- operationalization: the imitation game



# Turing Test in Cinema



# Turing Test: Brief History

## ● Eliza

```
Welcome to
EEEEEE LL      IIII ZZZZZZ  AAAA
EE  LL      II      ZZ  AA  AA
EEEE  LL      II      ZZZ  AAAAAA
EE  LL      II      ZZ  AA  AA
EEEEEE LLLLLL IIII ZZZZZZ  AA  AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU:  Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:  They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:  Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:  He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:  It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:  █
```

- developed in 1966 by J. Weizenbaum
- uses combination of **pattern matching** and **scripted rules**
- most famous script mimics a **psychologist** ⇨ many questions
- fooled early users

# Turing Test: Brief History

- Eliza
- Loebner Prize



- annual competition between 1991–2019
- most human-like AI is awarded
- highly controversial

# Turing Test: Brief History

- Eliza
- Loebner Prize
- Eugene Goostman



- mimics a 13-year-old boy from Odessa, Ukraine with a guinea pig
- “not too old to know everything and not too young to know nothing”
- 33% of judges were convinced it was human in 2014  
    ↪ first system that passed the Turing test (?)



# Turing Test: Brief History

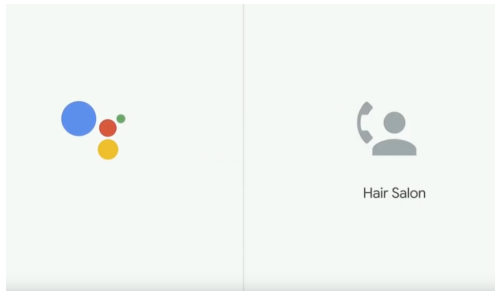
- Eliza
- Loebner Prize
- Eugene Goostman
- **Kuki** (formerly Mitsuku)



- **five times winner** of Loebner prize competitions (2015-2019)
- winner of "bot battle" versus Facebook's **Blenderbot**  
↪ <https://youtu.be/RBK5j0yXDT8>

# Turing Test: Brief History

- Eliza
- Loebner Prize
- Eugene Goostman
- Kuki (formerly Mitsuku)
- Google Duplex



- commercial product announced in 2018
- performs phone calls (making appointments) **fully autonomously**
- after criticism, it now starts conversation by **identifying as a robot**

# Turing Test: Brief History

- Eliza
- Loebner Prize
- Eugene Goostman
- Kuki (formerly Mitsuku)
- Google Duplex
- LaMDA & ChatGPT



- systems like LaMDA and ChatGPT would likely pass the Turing test
- example conversation: <https://www.nytimes.com/2023/02/16/technology/bing-chatbot-transcript.html>
- ChatGPT even [passed some exams](#) (but failed on others)

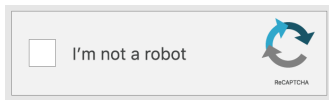
# Value of the Turing Test

- human actions **not always intelligent**
- **scientific value** of Turing test questionable:
  - Test for AI or for interrogator?
  - results not reproducible
  - strategies to succeed  $\neq$  intelligence:
    - **deceive** interrogator
    - **mimic** human behavior




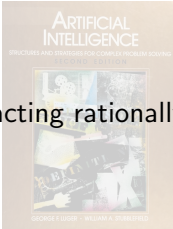
⇒ not important in AI “mainstream”



**practical** application: CAPTCHA  
 (“**C**ompletely **A**utomated **P**ublic Turing test to tell **C**omputers and **H**umans **A**part”)

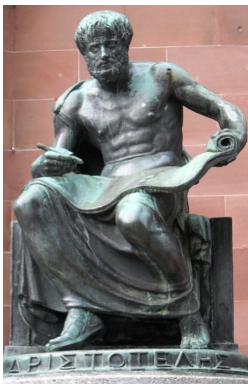


# What Do We Mean by Artificial Intelligence?

 <p>thinking like humans</p>	 <p>thinking rationally</p>
 <p>acting like humans</p>	 <p>acting rationally</p>

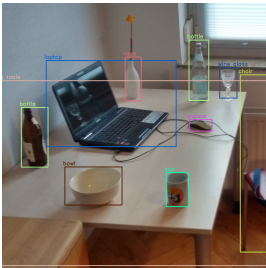
# Thinking Rationally

# Thinking Rationally: Laws of Thought

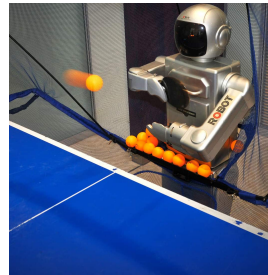
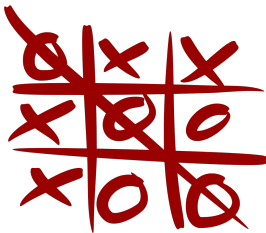


- **Aristotle:** What are correct arguments and modes of thought?
- **syllogisms:** structures for arguments that always yield correct conclusions given correct premises:
  - Socrates is a human.
  - All humans are mortal.
  - Therefore Socrates is mortal.
- direct connection to modern AI via mathematical **logic**

# Problems of the Logical Approach



not all intelligent behavior stems from **logical thinking** and **formal reasoning**





# What Do We Mean by Artificial Intelligence?

 <p>thinking like humans</p>	 <p>thinking rationally</p>
 <p>acting like humans</p>	 <p><b>acting rationally</b></p>

# 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 development of rational agents:

- more general than thinking rationally (logical inference only one way to obtain rational behavior)
- better suited for scientific method than approaches based on human thinking and acting

↪ most common view of AI scientists today

↪ what we use in this course

# 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 like humans: e.g., Turing test
- thinking like humans: cf. cognitive (neuro-)science
- thinking rationally: logic
- **acting rationally**: most common view today
  - $\rightsquigarrow$  amenable to scientific method
  - $\rightsquigarrow$  used in this course