

# Foundations of Artificial Intelligence

## 1. Introduction: What is Artificial Intelligence?

Thomas Keller and Florian Pommerening

University of Basel

February 20, 2023

# Foundations of Artificial Intelligence

February 20, 2023 — 1. Introduction: What is Artificial Intelligence?

## 1.1 What is AI?

## 1.2 Thinking like Humans

## 1.3 Acting like Humans

## 1.4 Thinking Rationally

## 1.5 Acting Rationally

## 1.6 Summary

## Introduction: Overview

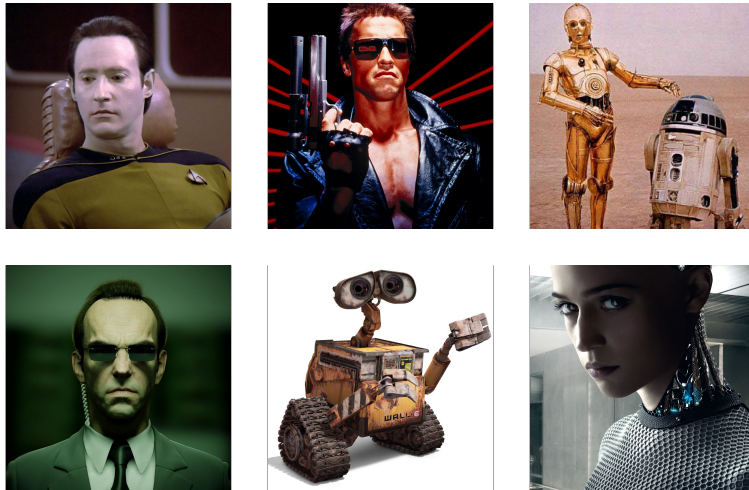
### Chapter overview: introduction

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

## 1.1 What is AI?

## What do we Mean by Artificial Intelligence?

what **pop culture** tells us:



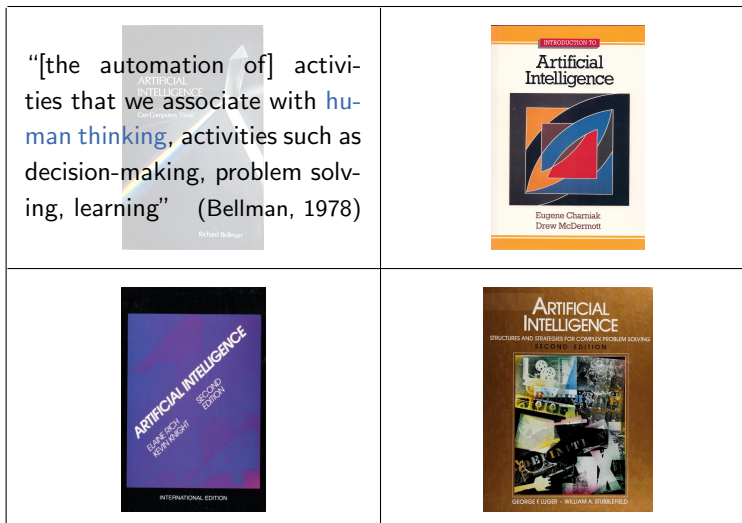
## What is AI: Humanly vs. Rationally; Thinking vs. Acting

what **scientists** tell us:



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

what **scientists** tell us:



## 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:

<div> <div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>Can Computers Think?</div> <div>Richard Bellman</div> </div> </div> <div>thinking like humans</div> </div> </div>	<div> <div> <div> <div>INTRODUCTION TO</div> <div>Artificial Intelligence</div> <div> </div> <div>Eugene Charniak Drew McDermott</div> </div> </div> </div>
<div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>BASED ON</div> <div>REASONING</div> <div>INTERNATIONAL EDITION</div> </div> </div> <div> <div>“the study of how to make computers <b>do</b> things at which, at the moment, <b>people</b> are better”</div> <div>(Rich &amp; Knight, 1991)</div> </div> </div>	<div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>STRUCTURES AND STRATEGIES FOR COMPLEX PROBLEM SOLVING</div> <div>THIRD EDITION</div> <div> </div> <div>George E. Luger · William A. Stubblefield</div> </div> </div> </div>

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

what scientists tell us:

<div> <div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>Can Computers Think?</div> <div>Richard Bellman</div> </div> </div> <div>thinking like humans</div> </div> </div>	<div> <div> <div> <div>INTRODUCTION TO</div> <div>Artificial Intelligence</div> <div> </div> <div>Eugene Charniak Drew McDermott</div> </div> </div> </div>
<div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>BASED ON</div> <div>REASONING</div> <div>INTERNATIONAL EDITION</div> </div> </div> <div> <div>acting like humans</div> </div> </div>	<div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>STRUCTURES AND STRATEGIES FOR COMPLEX PROBLEM SOLVING</div> <div>THIRD EDITION</div> <div> </div> <div>George E. Luger · William A. Stubblefield</div> </div> </div> </div>

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

what scientists tell us:

<div> <div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>Can Computers Think?</div> <div>Richard Bellman</div> </div> </div> <div>thinking like humans</div> </div> </div>	<div> <div> <div> <div>INTRODUCTION TO</div> <div>Artificial Intelligence</div> <div> </div> <div>Eugene Charniak Drew McDermott</div> </div> </div> <div> <div>“the study of <b>mental</b> faculties through the use of <b>computational models</b>”</div> <div>(Charniak &amp; McDermott, 1985)</div> </div> </div>
<div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>BASED ON</div> <div>REASONING</div> <div>INTERNATIONAL EDITION</div> </div> </div> <div> <div>acting like humans</div> </div> </div>	<div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>STRUCTURES AND STRATEGIES FOR COMPLEX PROBLEM SOLVING</div> <div>THIRD EDITION</div> <div> </div> <div>George E. Luger · William A. Stubblefield</div> </div> </div> </div>



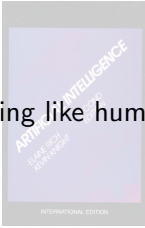
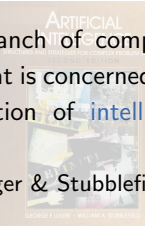
What is AI: Humanly vs. Rationally; Thinking vs. Acting

what scientists tell us:

<div> <div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>Can Computers Think?</div> <div>Richard Bellman</div> </div> </div> <div>thinking like humans</div> </div> </div>	<div> <div> <div> <div>INTRODUCTION TO</div> <div>Artificial Intelligence</div> <div> </div> <div>Eugene Charniak Drew McDermott</div> </div> </div> <div> <div>thinking rationally</div> </div> </div>
<div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>BASED ON</div> <div>REASONING</div> <div>INTERNATIONAL EDITION</div> </div> </div> <div> <div>acting like humans</div> </div> </div>	<div> <div> <div> <div>ARTIFICIAL INTELLIGENCE</div> <div>STRUCTURES AND STRATEGIES FOR COMPLEX PROBLEM SOLVING</div> <div>THIRD EDITION</div> <div> </div> <div>George E. Luger · William A. Stubblefield</div> </div> </div> </div>

## 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

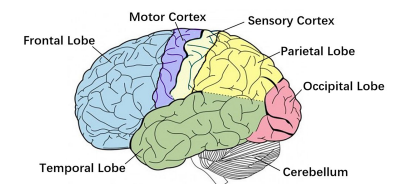
what **scientists** tell us:

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

## 1.2 Thinking like Humans

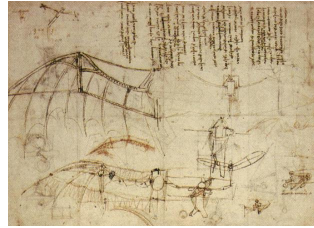
## Cognitive (Neuro-)Science

- ▶ 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?

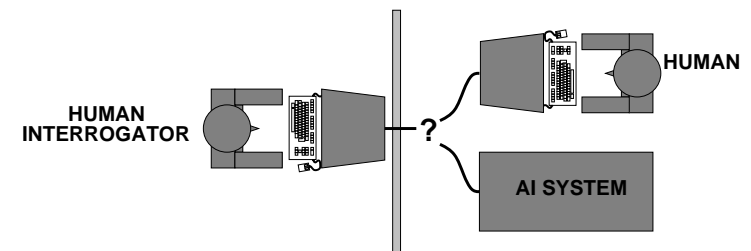


## 1.3 Acting like Humans

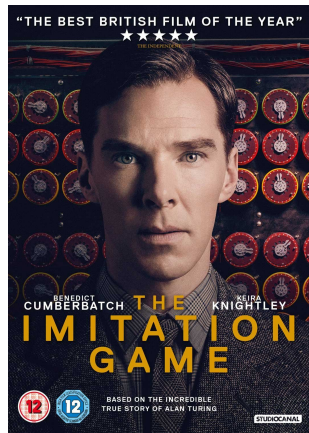
## The Turing Test

Alan Turing, Computing Machinery and Intelligence (1950):

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

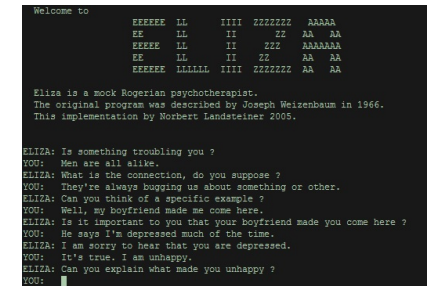


## Turing Test in Cinema



## Turing Test: Brief History

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



- ▶ 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
- ▶ Eugene Goostman
- ▶ Kuki (formerly Mitsuku)
- ▶ Google Duplex
- ▶ LaMDA & ChatGPT



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

## Turing Test: Brief History

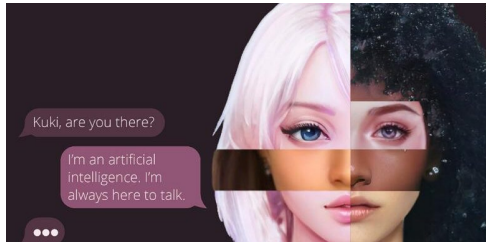
- ▶ Eliza
- ▶ Loebner Prize
- ▶ Eugene Goostman
- ▶ Kuki (formerly Mitsuku)
- ▶ Google Duplex
- ▶ LaMDA & ChatGPT



- ▶ 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 AI that passed the Turing test (?)**

## Turing Test: Brief History

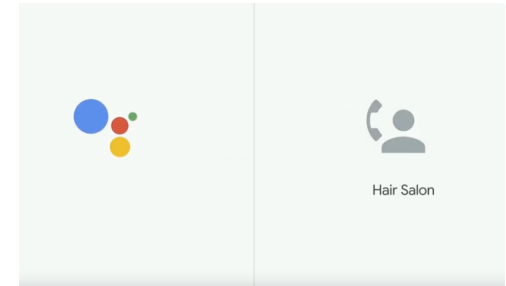
- ▶ Eliza
- ▶ Loebner Prize
- ▶ Eugene Goostman
- ▶ Kuki (formerly Mitsuku)
- ▶ Google Duplex
- ▶ LaMDA & ChatGPT



- ▶ five times winner of Loebner prize competitions (2015-2019)
- ▶ winner of “bot battle” versus Facebook’s Blenderbot  
url: <https://youtu.be/RBK5j0yXDT8>

## Turing Test: Brief History

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



- ▶ 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

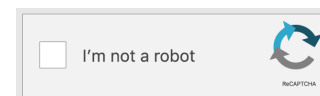


- ▶ systems like LaMDA and ChatGPT would likely pass the Turing test
- ▶ example conversation: <https://www.nytimes.com/2023/02/16/technology/bing-chatbot-transcript.html>
- ▶ ChatGPT even passed some exams (but failed on others)

## 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”



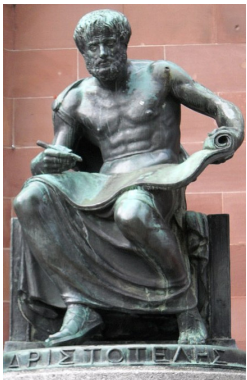
practical application: CAPTCHA  
 (“Completely Automated Public Turing test to tell Computers and Humans Apart”)

## What do we Mean by Artificial Intelligence?



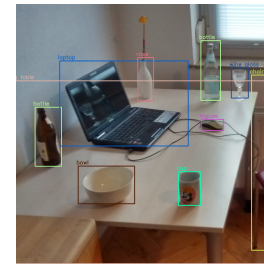
## 1.4 Thinking Rationally

## Thinking Rationally: Laws of Thought

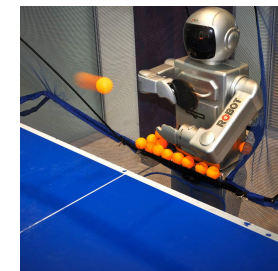
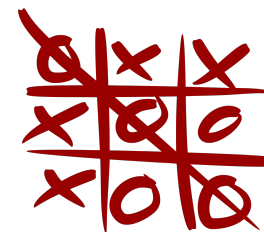


- ▶ **Aristotle:** What are correct arguments and modes of thought?
- ▶ **syllogisms:** structures for arguments that always yield correct conclusions given correct premises:
  - ▶ Socrates is a human.
  - ▶ All humans are mortal.
  - ▶ Therefore Socrates is mortal.
- ▶ direct connection to modern AI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



## 1.5 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

## 1.6 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**: the most common view today
  - $\rightsquigarrow$  amenable to scientific method