

Lecture

Foundations of AI (GKI)

Intelligence & Knowledge

Dr. Mohsen Mesgar

Universität Duisburg-Essen

Exercise and practice class

- We suggest that **before the practice class begins**
 - You download notebooks to your computer
 - Upload them to Google Colab
 - Read and practice all tasks in the notebook
 - **Notebooks should be very clear on their own**
 - **Try to bring ONLY your questions to the practice class**
- **The goal of practice class is to go over the tasks in a notebook and answer your questions**
- **Friday 15. April 2022** is a public holiday. **The practice class does not take place on that date.**

Homework

- Homework will be **graded**.
- You can earn **bonus points** for your final grade.
- To make your life easier, we let you **solve homework in a group**.
- **To submit** your solution to a homework, you **MUST use Moodle**.
- **Each group** MUST submit only **ONE solution** for **each homework**.
- **Read the instructions and the deadline** for each homework in **its notebook**.

- **Moodle**
 - <https://moodle.uni-due.de/course/view.php?id=34785>
- **Slack**
 - https://join.slack.com/t/gki-sose2022/shared_invite/zt-16hul6yfl-Tg0Y6zmsWrk1Exj453xo1g
- **Google Colab**
 - <https://colab.research.google.com>

Feedback

- **Moodle**
 - <https://moodle.uni-due.de/course/view.php?id=34785>
 - Leave feedback for Lecture 01: Introduction

Any open questions from previous lecture / practice classes



Goal: In this lecture you learn ..

- What is intelligence?
- What is artificial intelligence?
 - Systems that think like humans
 - Systems that act like humans
 - Turing Test
 - Captchas
 - Systems that think rationally
 - Systems that act rationally
- Strong AI vs Weak AI
- Ethical concerns

Intelligence

What is intelligence?

- **Oxford Dictionary:**
 - The ability to acquire and apply knowledge and skills.
- **Wikipedia:** Intelligence has been defined in many different ways including
 - the capacity for logic, understanding, self awareness, learning, emotional knowledge, reasoning, planning, creativity, and problem solving.
 - It can be more generally described as the ability to perceive or infer information, and to retain it as knowledge to be applied towards adaptive behaviors within an environment or context.

What is intelligence?

Is it intelligent?



What is intelligence?

What about this machine?

- Performs a single task that is said to require intelligence
- In the (limited) domain of chess, behaves like a human
- Will not make certain mistakes human players would do
 - Is this a problem?



What is intelligence?

What about Sophia?

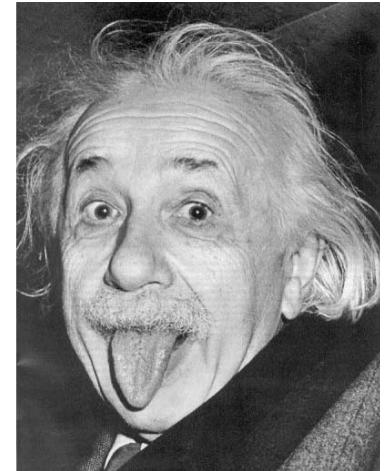
- Behaves like a human
 - interacts, shows emotions, capable of planning, solves difficult problems
- Behaves like a human
- Looks like a human
 - not quite



What is intelligence?

What about him?

- Communicates like a human
- Behaves like a human
- Looks like a human
- Has a brain
 - Currently at National Museum of Health and Medicine
- http://en.wikipedia.org/wiki/Albert_Einstein%27s_brain



What is intelligence?

- The conceptualization of intelligence is the ability to autonomously and efficiently achieve complex goals.

Human intelligence

Human intelligence

- **Human (general) intelligence** is the “**real**” form of intelligence
- Indeed, as humans we know ourselves as the entities with the highest intelligence ever observed in the Universe.
- As an extension of this, we like to see ourselves as **rational beings** who are able **to solve a wide range of complex problems** under **all kinds of circumstances** using **our experience and intuition, supplemented by the rules of logic, decision analysis and statistics**.
- It is not surprising that we have some difficulty to accept the idea that we might be a bit less smart than we keep on telling ourselves.

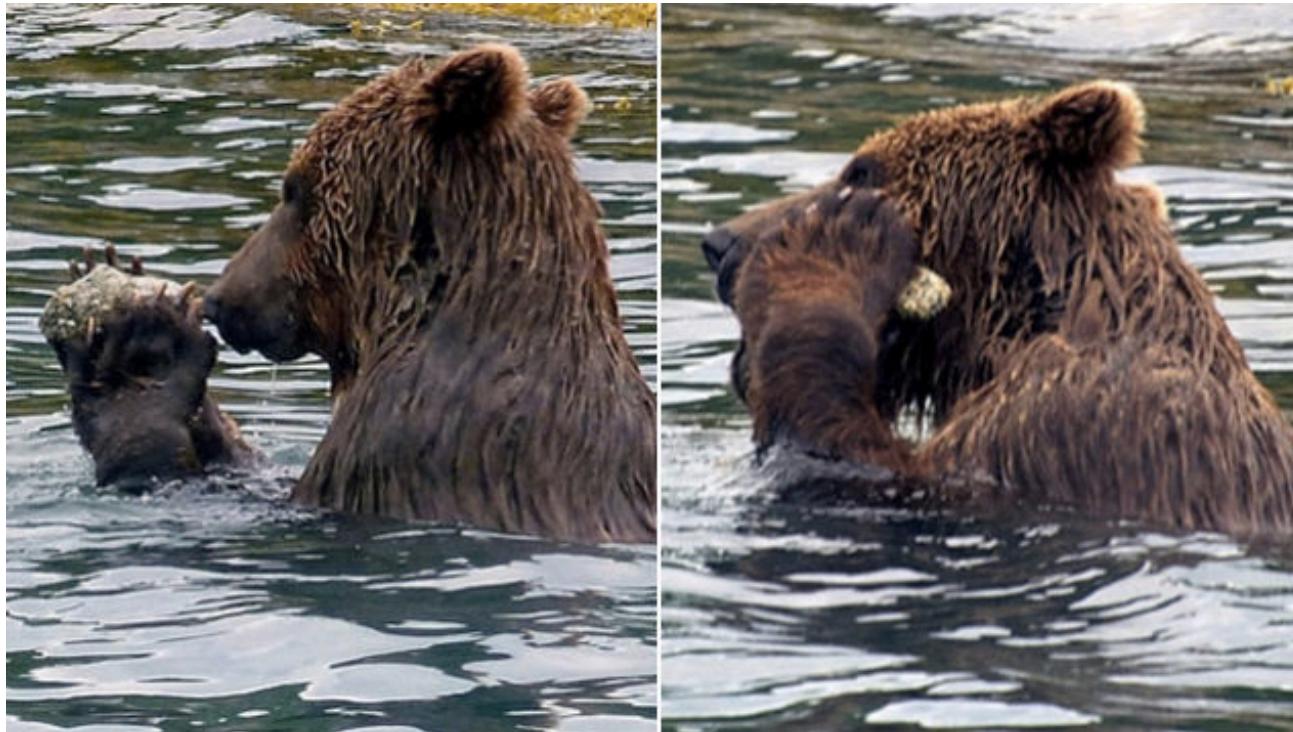
What is human intelligence?

- <https://rb.gy/rdmmxt>

What is human intelligence?

What is human intelligence?

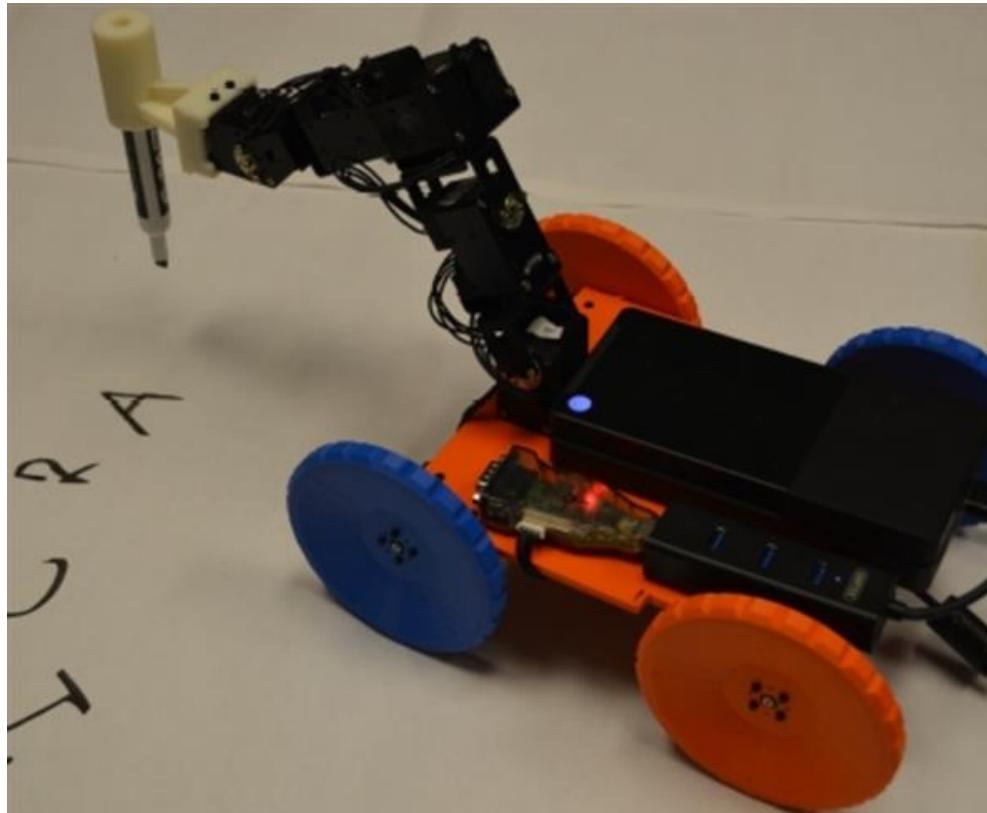
- Use tool?



Images by Volker Deecke

Use tool

- Robots can use tools. But how do they know which tool to use when?



Source: <http://www.3ders.org/articles/20170531-build-your-own-3d-printed-walking-or-wheeled-robot-with-new-carnegie-mellon-design-tool.html>

What is human intelligence?

- Tool use?
- Plan?

What is human intelligence?

- Tool use?
- Plan?

Can Kabadayi, Mathias Osvath (2017):

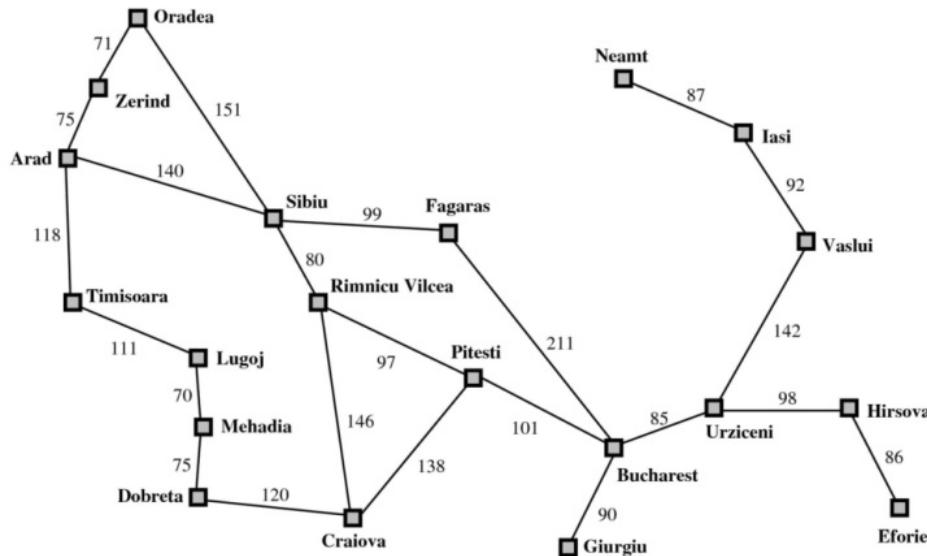
Ravens parallel great apes in flexible planning for tool-use and bartering



Sources:

https://upload.wikimedia.org/wikipedia/commons/3/3a/Eichh%C3%B6rnchen_im_Herbst.jpg
https://en.wikipedia.org/wiki/Common_raven#/media/File:A_Common_Raven.jpg

- Navigation planning → Lecture 4



What is human intelligence?

- Tool use?
- Planning?
- Language?

Francine G. Patterson (1978):

The gestures of a gorilla: Language acquisition in another pongid



Source: <http://www.koko.org/sign-language>

Language

- **Examples of language technology research:**
 - machine translation,
 - document retrieval,
 - speech recognition & synthesis,
 - summarization,
 - linguistic analysis (spelling, grammar),
 - caption generation
- Interested? Come talk to me in office hours.

What is human intelligence?

- Tool use?
- Planning?
- Language?
- Be creative?



Source: Ingo Arndt https://www.theguardian.com/artanddesign/gallery/2014/apr/22/the-worlds-best-animal-architecture-in-pictures?CMP=twt_gu

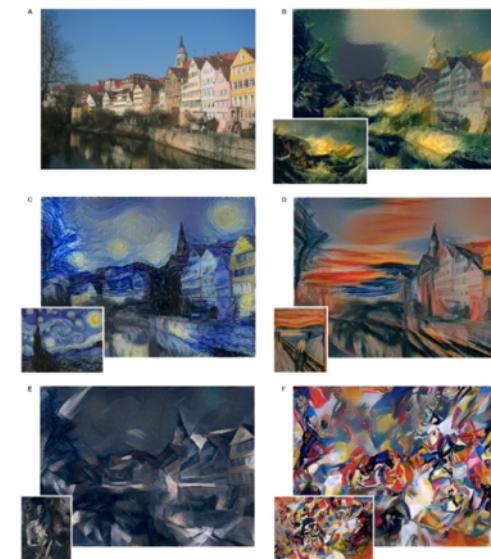
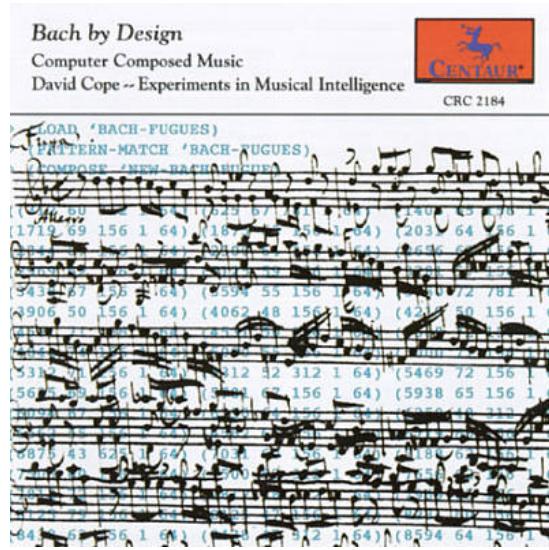


Source: Nature Picture Library /Alamy
<http://www.bbc.com/future/story/20140723-are-we-the-only-creative-species>

Creativity

- David Cope (1987):
Experiments in Music Intelligence
 - Leon A. Gatys, Alexander S. Ecker, Matthias Bethge (2015):
A Neural Algorithm of Artistic Style

→ Lectures 7-10 on machine learning and deep learning



What is human intelligence?

- Tool use?
- Planning?
- Language?
- Be creative?
- Cooperation and social behavior?

Human beings have evolved to coordinate complex activities, to gossip and to playact together. It is because they are adapted for such cultural activities — and not because of their cleverness as individuals — that human beings are able to do so many exceptionally complex and impressive things.

Michael Tomasello

Cooperation

- No single ant knows how to construct an anthill

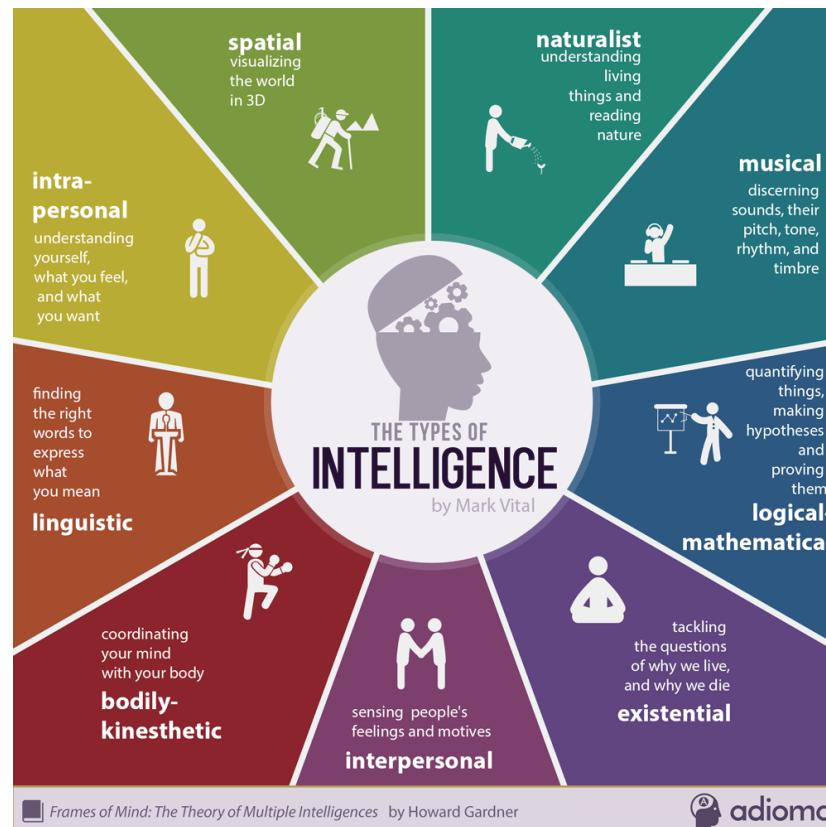


Source: <https://pixabay.com/de/insekten-wald-ameisenhaufen-1033718/>

Human intelligence

Educational perspective

- Howard Gardner (1993):
Frames of Mind: The Theory of Multiple Intelligences



Source: <https://blog.adioma.com/9-types-of-intelligence-infographic/>

Artificial intelligence

Definitions of AI

Systems that think like humans

“The exciting new effort to make computers think ... *machines with minds*, in the full and literal sense.”

(Haugeland, 1985)

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

Definitions of AI

Systems that think like humans

“The exciting new effort to make computers think ... *machines with minds*, in the full and literal sense.”

(Haugeland, 1985)

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

“The art of creating machines that perform functions that require intelligence when performed by people.”

(Kurzweil, 1990)

“The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)

Systems that act like humans

Definitions of AI

Systems that think like humans

“The exciting new effort to make computers think ... *machines with minds*, in the full and literal sense.”
(Haugeland, 1985)

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

“The art of creating machines that perform functions that require intelligence when performed by people.”
(Kurzweil, 1990)

“The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)

Systems that think rationally

“The study of mental faculties through the use of computational models.”
(Charniak and McDermott, 1985)

“The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)

Systems that act like humans

Definitions of AI

Systems that think like humans

“The exciting new effort to make computers think ... *machines with minds*, in the full and literal sense.”
(Haugeland, 1985)

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

“The art of creating machines that perform functions that require intelligence when performed by people.”
(Kurzweil, 1990)

“The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)

Systems that act like humans

Systems that think rationally

“The study of mental faculties through the use of computational models.”
(Charniak and McDermott, 1985)

“The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)

“Computational Intelligence is the study of the design of intelligent agents.” (Poole et al., 1998)

“AI ...is concerned with intelligent behavior in artifacts.” (Nilsson, 1998)

Systems that act rationally

Definitions of AI

Systems that think like humans

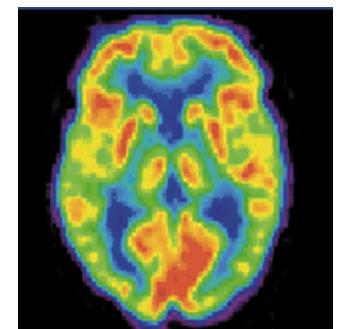
Systems that think rationally

Systems that act like humans

Systems that act rationally

Think like a human

- Requires a scientific theory of human thinking
 - Up to which level of abstraction?
 - How can we validate our theory?
- Cognitive Science
 - predicting and testing behavior of human subjects (top-down)
- Cognitive Neuroscience
 - direct identification from neurological data (bottom-up)
- In the end:
 - highly philosophical question

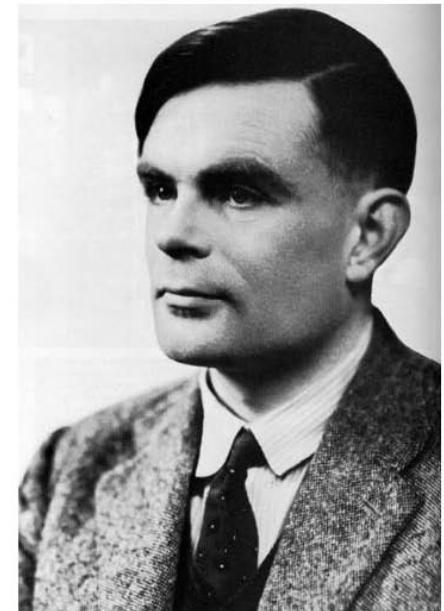


Source: https://commons.wikimedia.org/wiki/File:NIH_PET.JPG/

In Turing's Words

“The original question, ‘Can machines think?’, I believe to be too meaningless to deserve discussion. Nevertheless, I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted.”

Alan Turing (1950)



Alan Mathison Turing (1912-1954)

Do computers think?

Garry Kasparov, 1996, after losing for the first time against Deep Blue:

I got my first glimpse of Artificial Intelligence on Feb. 10, 1996, at 4:45 p.m. EST, when in the first game of my match with Deep Blue, the computer ... [made] a wonderful and extremely human move.

I could feel - I could smell - a new kind of intelligence across the table. While I played through the rest of the game as best I could, I was lost; it played beautiful, flawless chess the rest of the way and won easily.

If the computer makes the same move that I would make for completely different reasons, has it made an "intelligent" move? Is the intelligence of an action dependent on who (or what) takes it?

So although I think I did see some signs of intelligence, it's a weird kind, an inefficient, inflexible kind that makes me think I have a few years left.

Garry Kasparov, The Day That I Sensed a New Kind of Intelligence, Time Magazine 147, 13 (March 25 1996). <http://www.time.com/time/magazine/article/0,9171,984305-1,00.html>

Definitions of AI

Systems that think like humans

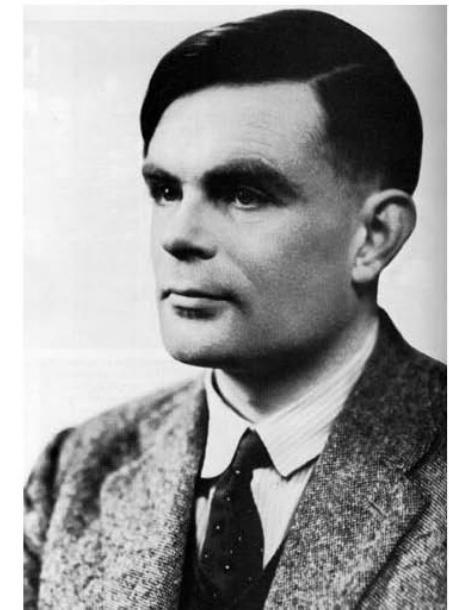
Systems that think rationally

Systems that act like humans

Systems that act rationally

Turing Test

- How can we evaluate whether a system acts intelligently?
- Instead of defining a long and controversial list of necessary prerequisites for intelligence
→ compare to undeniably intelligent humans
- Fixed game-like setup



Alan Mathison Turing (1912-1954)

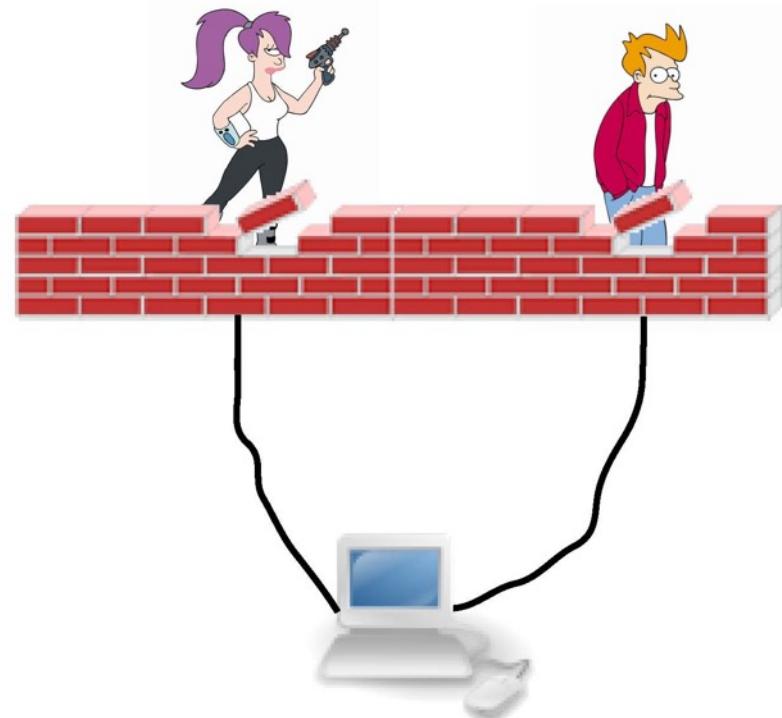
Imitation Game

Human judge connected via computer
to a man and a woman

- The judge can ask any question about any subject
- The woman replies honestly
- The man pretends he is a woman

The judge has to find out:

- Who is the woman?



Turing Test

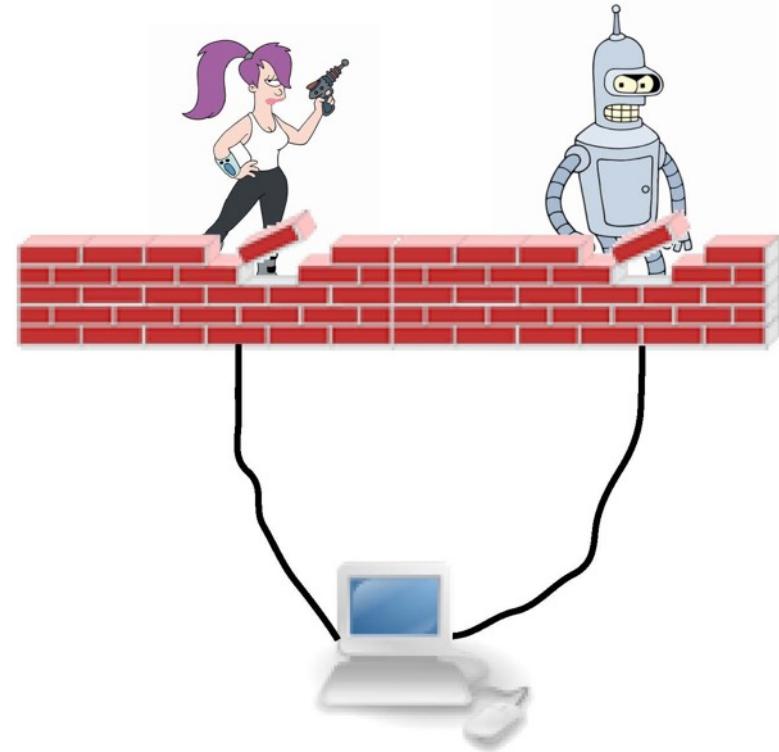
Human judge connected via computer
to a **computer** and a human

- The judge can ask any question about any subject
- The human replies honestly
- The **computer** pretends he is a human

The judge has to find out:

- Who is the human?

If the computer can fool the judge, we consider it intelligent



Act Like a Human

Turing Test

think like humans

think rationally

act like humans

act rationally

I believe that in about fifty years' time it will be possible to program computers, ..., to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after 5 minutes of questioning.

Alan Turing, 1950

Your turn

- Think about questions, which you would ask if you were a judge.



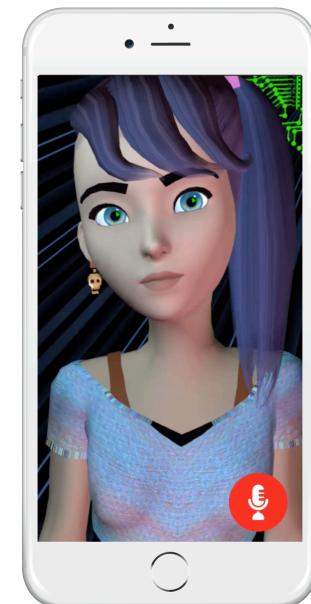
Source: <http://myneuroworld.com/test-your-brain/>

Your turn

- Test them with Kuki,:
<https://chat.kuki.ai>
- Three-time winner of the Loebner prize
build by Steve Worswick
- Are you convinced?



Source: Steve Worswick



5 minutes Break

Recap

- What is intelligence?
- What is artificial intelligence?
- Systems that think like humans
- Systems that act like humans
- Turing Test

Systems that think like humans

Systems that think rationally

Systems that act like humans

Systems that act rationally

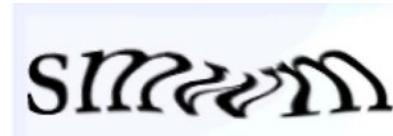
Reverse Turing: Are you human?

Turing test backwards

- we know that computers cannot pass the Turing test
- use Turing-style questions to exclude bots from services

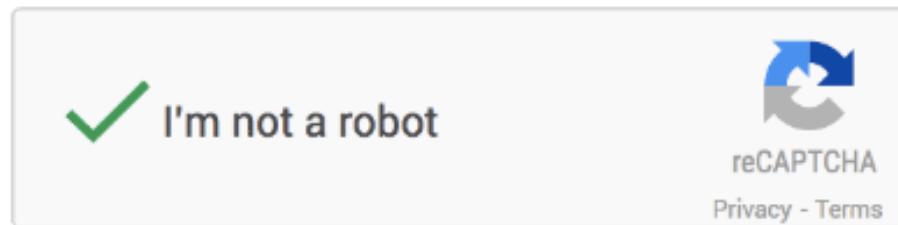
Examples:

- “What is written here?”



CAPTCHA:

Completely Automated Public Turing test to tell Computers and Humans Apart



CAPTCHAs in 2018

Wählen Sie alle Bilder aus, die ein Geschäft von vorne zeigen

Select all squares with street signs.

Report a problem

Crowdworkers

TO COMPLETE YOUR REGISTRATION, PLEASE TELL US WHETHER OR NOT THIS IMAGE CONTAINS A STOP SIGN:



NO YES

ANSWER QUICKLY—OUR SELF-DRIVING CAR IS ALMOST AT THE INTERSECTION.

SO MUCH OF "AI" IS JUST FIGURING OUT WAYS TO OFFLOAD WORK ONTO RANDOM STRANGERS.

Definitions of AI

Systems that think like humans

Systems that think rationally

Systems that act like humans

Systems that act rationally

Think Rationally

- When does one think rationally? Is there a “right” way of thinking?
Laws of Thought
- Aristotle: What are “correct” argument and thought processes?
 - Syllogism: “Socrates is a man. All men are mortal → Socrates is mortal”
 - initiated the field of logic
 - $A(s), \forall x(A(x) \rightarrow B(x)) \rightarrow B(s)$
- The logicist tradition in AI hopes to create intelligent systems using logic programming.

But: Not all intelligence is mediated by logic behavior

Definitions of AI

Systems that think like humans

Systems that think rationally

Systems that act like humans

Systems that act rationally

Act Rationally

Rational behavior

- A system is rational if it does the "right thing," given what it knows.

Maximize goal achievement given the available information

think like humans

think rationally

act like humans

act rationally

Note: "Humans" are not always perfect ("rational") even when we have enough knowledge.

- We all are not chess grandmaster, although we may know all chess rules.
- Not everyone gets the best grade in an exam.

Two advantages over previous approaches

- More general than laws of thought approach
- More amenable to scientific development: rationality can be defined and optimized

Doesn't necessarily involve thinking

- Rational Agent

Flying = Modeling birds?



Flying = Exploiting aerodynamics?



- ◊ Human engineers succeeded at flying when they stopped modeling birds and started examining aerodynamic processes.

Some human behavior is irrational

- reacting to insults
- lying
- typing mistakes

Some intelligent behavior is inhuman

- “What is the square root of 5411661?”
- “Dude, I do not have a calculator at hand.” vs “2326.297702359”

Definitions of AI

Systems that think like humans

“The exciting new effort to make computers think ... *machines with minds*, in the full and literal sense.”
(Haugeland, 1985)

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

“The art of creating machines that perform functions that require intelligence when performed by people.”
(Kurzweil, 1990)

“The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)

Systems that act like humans

Systems that think rationally

“The study of mental faculties through the use of computational models.”
(Charniak and McDermott, 1985)

“The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)

“Computational Intelligence is the study of the design of intelligent agents.” (Poole et al., 1998)

“AI ...is concerned with intelligent behavior in artifacts.” (Nilsson, 1998)

Systems that act rationally

Strong AI vs. Weak AI

Strong AI: Build a universally intelligent agent

- also called AGI (artificial general intelligence)
 - can pass the Turing test
 - encompasses the full range of human cognitive abilities
- Fear of “singularity” – self-improving AGI that evolves extremely fast

Weak AI: Build agents that act rationally

- playing chess, mowing the lawn, ...
- targeted at problem solving

Strong AI vs. Weak AI

- Systems that ...

Strong AI

think like humans

think rationally

act like humans

act rationally

Weak AI

Ethical Aspects

- People might lose their job to automation.

So far, automation through information technology in general and AI in particular has created more jobs than it has eliminated, and has created more interesting, higher paying jobs

Russel & Norvig

- What about jobs for low-skilled workers?

Ethical Aspects

- People might lose their job to automation.
- People might have too much (or too little) leisure time.

In an information economy [...] there is increasing pressure on everyone to work harder. AI increases the pace of technological innovation and thus contributes to this overall trend, but AI also holds the promise of allowing us to take some time off and let our automated agents handle things for a while.

Russel & Norvig

Ethical Aspects

- People might lose their job to automation.
- People might have too much (or too little) leisure time.
- People might lose their sense of being unique.

Ethical Aspects

- People might lose their job to automation.
- People might have too much (or too little) leisure time.
- People might lose their sense of being unique.
- AI systems might be used towards undesirable ends.

Dual Use → Robot Armies?

Privacy vs Security → Digital surveillance?

Ethical Aspects

- People might lose their job to automation.
- People might have too much (or too little) leisure time.
- People might lose their sense of being unique.
- AI systems might be used towards undesirable ends.
- The use of AI systems might result in a loss of accountability.

Many new questions arising, e.g.

- Who is responsible if an autonomous car creates an accident?
- Who owns the intellectual property of automatically created art?
- Is sexual harassment in virtual reality a crime?

Ethical Aspects

- People might lose their job to automation.
- People might have too much (or too little) leisure time.
- People might lose their sense of being unique.
- AI systems might be used towards undesirable ends.
- The use of AI systems might result in a loss of accountability.
- The success of AI might mean an end of the human race.

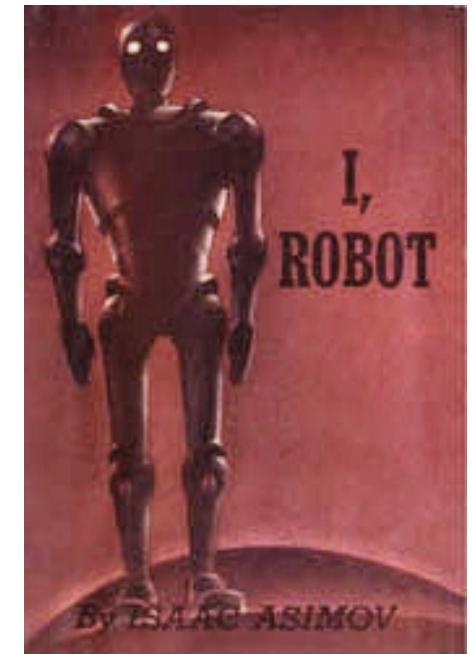
For example, we might propose a utility function designed to minimize human suffering [...]. Given the way humans are, however, we'll always find a way to suffer even in paradise; so the optimal decision for the AI system is to terminate the human race as soon as possible – no humans, no suffering.

Russel & Norvig

Laws of Robotics

Isaac Asimov (1950): *I, Robot*

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.



Source: https://en.wikipedia.org/wiki/I,_Robot

Readings

Mandatory

- Russell & Norvig, Section 1.1, *What is AI?*, p. 1-5.
- Russel & Norvig, Section 26.1, *Weak AI: Can machines act intelligently?*, p.1020-1026.
- Russel & Norvig, Section 26.3, *The ethics and risks of developing artificial intelligence*, p.1034-1040.

Optional

- Russel & Norvig, Section 26.1, Strong AI: Can machines really think?, p.1026-1033.
- Melanie Mitchell, Why AI is harder than we think. <https://arxiv.org/pdf/2104.12871.pdf>

Other References

- Francine G. Patterson (1978): [*The gestures of a gorilla: Language acquisition in another pongid*](#)
- Can Kabadayi, Mathias Osvath (2017): [*Ravens parallel great apes in flexible planning for tool-use and bartering*](#)
- Howard Gardner (1993): *Frames of Mind: The Theory of Multiple Intelligences*
- Garry Kasparov (1996): [*The Day That I Sensed a New Kind of Intelligence*](#), Time Magazine 147.
- David Cope (1987): [*Experiments in Music Intelligence*](#)
- Leon A. Gatys, Alexander S. Ecker, Matthias Bethge (2015): [*A Neural Algorithm of Artistic Style*](#)
- Michael Tomasello and Esther Herrmann (2010): [*Ape and Human Cognition: What's the Difference?*](#)
- Michael Tomasello (2008): [*How are humans unique?*](#)
- Joseph Weizenbaum (1966). [*ELIZA - a computer program for the study of natural language communication between man and machine.*](#)
- Barsalou (2008): [*Grounded cognition*](#).
- Pulvermüller & Bertier (2008): [*Aphasia therapy on a neuroscience basis*](#)

Popular Books

- Daniel Kahnemann (2011): *Thinking fast and slow*.
- Hugo Mercier, Dan Sperber (2017): *The enigma of reason*.
- Stefan Klein (2010): *Der Sinn des Gebens: Warum Selbstlosigkeit in der Evolution siegt und wir mit Egoismus nicht weiterkommen*.
- Michael Tomasello (2008): *Origins of Human Communication*.

Summary

- What is intelligence?
- What is artificial intelligence?
 - Systems that think like humans
 - Systems that act like humans
 - Turing Test
 - Captchas
 - Systems that think rationally
 - Systems that act rationally
- Strong AI vs Weak AI
- Ethical concerns

think like humans

think rationally

act like humans

act rationally