What is Artificial Intelligence?

What is ~~Artificial~~ Intelligence?

- We have a good intuitive idea of what “intelligence” is. (If I see it I recognise it)
- “Intelligence” is difficult to define rigorously:
  - Just clever behaviour is not enough.
  - Perhaps we also need some ingenuity/creativity.
  - Insight into problems, rather than mere mechanical application of rules.
- One important attempt to “solve” this is the Turing Test.

The Turing Test for Artificial Intelligence

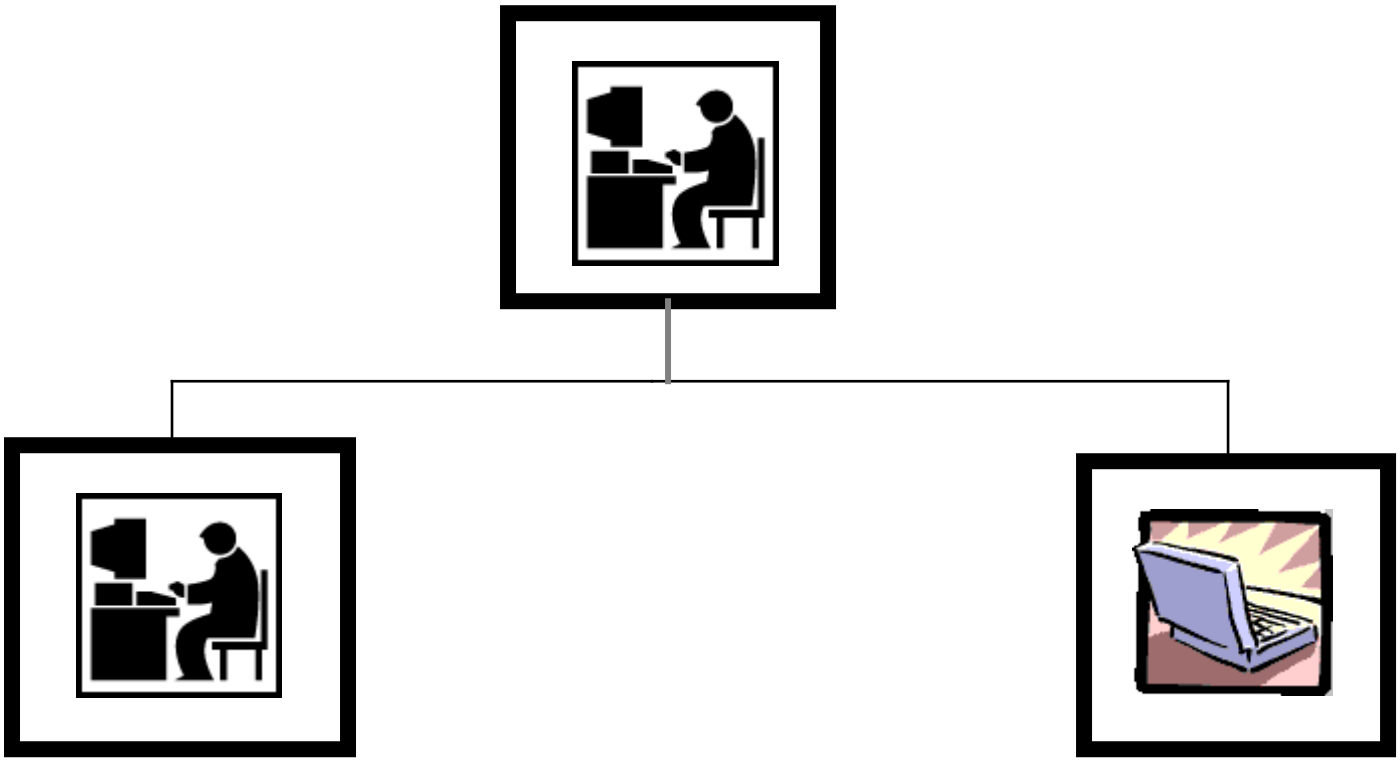
- Involves 3 participants, a human interrogator and two entities answering questions:
  - Another human
  - A computer

- The interrogator asks questions and gets answers via a keyboard and screen, without knowing if the answer was given by the other human or by the computer
  - Computer can give wrong answers to deceive interrogator

- The human interrogator has to figure out who is the other human and who is the computer, based on the answers.

The Turing Test for Artificial Intelligence

- Can a machine pass a behavioural test of intelligence?



How to program a machine to pass the test?
One possibility: the machine could learn…

Eliza – a psychoterapist program

- Written by Weizenbaum several decades ago
- Simulates a psychoterapist who encourages the patient to talk more about her/his problems
- The program uses a pattern matching technique
  - It scans the human user's sentence to detect some keyword(s)
  - It uses ready-made responses to detected keywords/patterns
  - E.g., keyword "everybody" is associated with a response such as "you have a specific person in mind, don't you"?
  - It is also capable of transforming the user's input sentence, using a list of transformations associated with patterns
  - As a last resource uses generic responses such as "tell me more about this"

Example of conversation between Eliza and a teenager (1)

Teenager: Men are all like.

Eliza: In what way?

Teenager: They're always bugging us about something.

Eliza: Can you think of a specific example?

Teenager: Well, my boyfriend made me come here.

Eliza: Your boyfriend made you come here.

Teenager: He says I am depressed much of the time.

Eliza: I am sorry to hear that you are depressed.

Teenager: It's true that I am unhappy.

Eliza: Do you think coming here will help you not to be unhappy?

Example of conversation between Eliza and a teenager (2)

Teenager: I need some help, that much seems certain.

Eliza: What would it mean to you if you got some help?

Teenager: Perhaps I could learn to get along with my mother.

Eliza: Tell me more about your family.

Teenager: My mother takes care of me.

Eliza: Who else in your family takes care of you?

Teenager: My father.

Eliza: Your father.

Teenager: You are like my father in some ways.

Eliza: What resemblance do you see?

. . . . .

Evaluating Eliza

- Eliza was used for several years, with many users
- Many users formed an “emotional bond” with Eliza (some people told their intimate secrets to the program), something that Weizenbaum found terribly wrong
- It was difficult to convince some people that Eliza was just a program, not a human psychiatrist.
- If you want to try it: https://www.masswerk.at/elizabot/
- A report on the Journal of Nervous and Mental Disease suggested that the program could be a therapeutic tool for hospitals and psychiatric centres with a shortage of therapists
- Weizenbaum wrote a book called “Computer Power and Human Reason”, criticizing AI and its goals, alerting to the danger of delegating human welfare decisions to computers

Has Eliza passed the Turing test?

- Despite these impressive results, Eliza has not passed the Turing test
  - Users were not trying to distinguish between a computer and a human, they simply assumed they were talking to a human
  - There were only two participants, rather than 3 (putting a human competing with the computer in the answers to the interrogator makes the test more challenging)
  - Eliza simulates only a psychiatric interview, not a general conversation. Weizenbaum chose this scenario because it is 'one of the few examples of . . . natural language communication in which one of the participating pair is free to assume the pose of knowing almost nothing of the real world'

Four Possible Objections to the Turing Test (1)

- The chimpanzee objection
  - Chimpanzees can think, but they would not pass the test
  - So, the test is too conservative and biased towards human intelligence
  - This objection shows a limitation, rather than flaw, of the test: Failing the test is inconclusive, but passing it shows intelligence

Four Possible Objections to the Turing Test (2)

- The sense organs objection
  - Test does not measure the computer's understanding of the words it is using in relation to real-world objects
    - So, passing the test is not proof of intelligence
  - Test should be strengthened by equipping the computer with artificial sense organs (e.g., cameras, arms, etc.)
  - However, many concepts are abstract and can be the subject of an interrogation to test intelligence without sense organs

Four Possible Objections to the Turing Test (3)

- The simulation objection
  - passing the test only shows that the computer is good at simulating thinking, not that it is thinking.
  - Assumes all simulations lack an essential feature of the thing being simulated – prejudiced answer to the test
  - This argument does not consider the possibility of a simulation having all essential features of the thing being simulated.
  - E.g., a simulated (artificial) voice is essentially a voice…
  - Raises the question: is it possible that a computer passes the test by using a simulation that does not have all the essential features of intelligence?

Four Possible Objections to the Turing Test (4)

- The black box objection
  - In the test, the computer is treated as a black box, the interrogator evaluates only the outward behaviour of the computer
    - Ignores how the computer program was designed and works
  - We evaluate the intelligence of people based on their outward behaviour, but we believe all people have similar brains (from a biochemical point of view)
  - Based on our analysis of the way that Eliza works, we can naturally conclude that Eliza does not “think”; so it seems important to consider the internal design of the program, not just its external behaviour

  Here is a *very hypothetical* way for a program that does not think to pass the Turing test:
  - Compute all meaningful English sentences having at most 100 words (a huge but finite number of sentences)
  - Associate with each of these sentences a ready-made meaningful response
  - The computer would pass the test, unless user questions have > 100 words

Four Possible Objections to the Turing Test (5)

- The black box objection (cont.)
  - The Turing test is based just on the output of the program
  - Should we use both an output criterion and a design criterion?
  - Possibilities for a design criterion
    - Program should do things it does in a way broadly similar to the way those things are done in the human brain
      - "strengthens" the test, and it is anthropocentric
    - Program should be modular – capable of being incorporated, as a building block, onto more complex programs
    - E.g., it should be possible to incorporate a conversation program that passed the test into the programs that control a robot, giving the robot an ability to talk about what it is doing

Summary

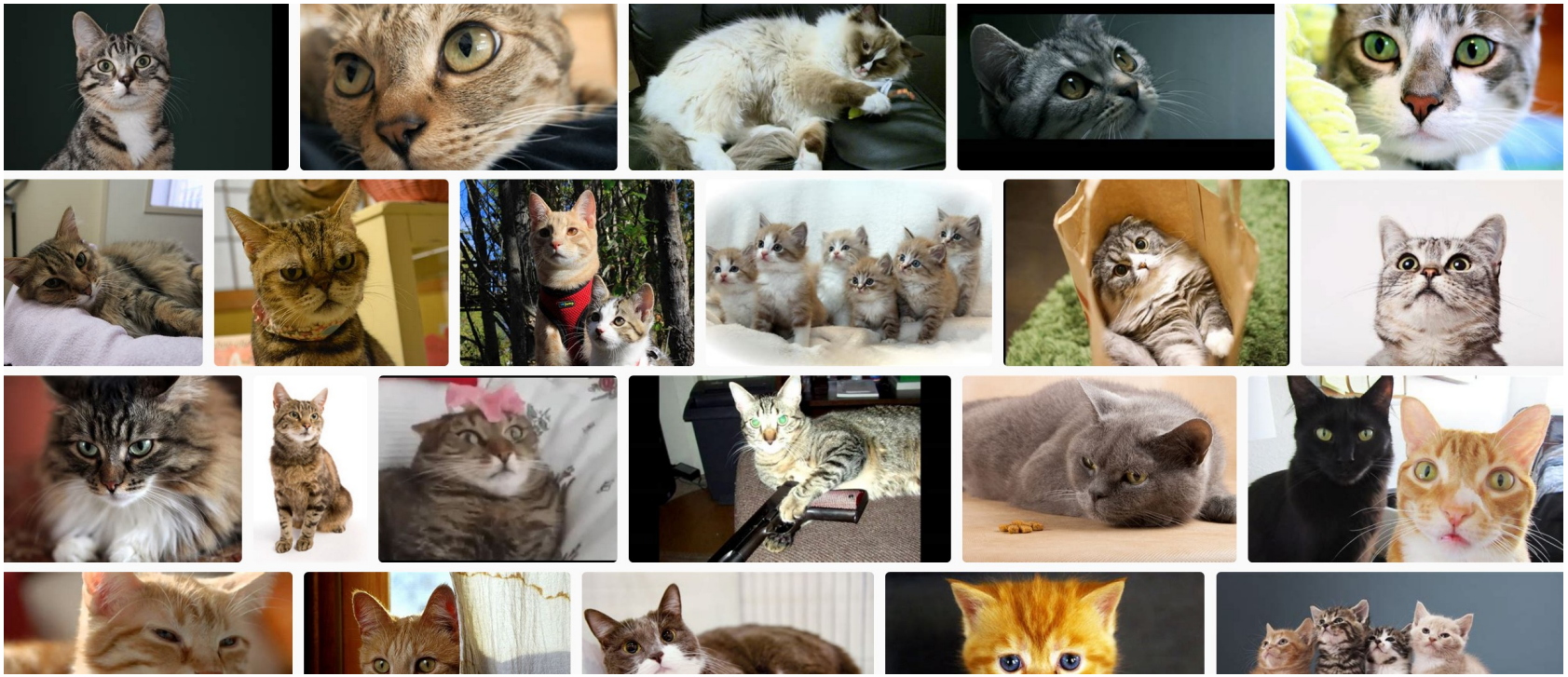
- The Turing Test is a behavioural test of computational intelligence
  - It considers only the output of a program, not its internal design
  - There are arguments to consider the internal design too…
- The usefulness of the test is controversial
  - Heavily criticized as “inadequate”, but no clearly better test for “computational intelligence” has been proposed yet
  - We don’t even have a good definition of human intelligence!

New Artificial Intelligence

- Nowadays there is less focus on modelling “intelligence” as such.
- Much of AI is about data-driven
  - **Prediction**
  - **analysis**
  - **Pattern extraction**
  - **(also) neuroscience/understanding how the brain works.**
- The philosophical idea of “intelligence” has for now receded to the background.

Statistical Machine learning

- Consider images of cats.
- There is something similar shared by all the cats (which makes it possible for us to recognise them).
- This similarity can be captured statistically.



Statistical Machine Learning

- In each image the cats look different
- The background is different (guns, sofas, etc...).
- There may be more than one cat.
- Indeed, most of the information in each image is not relevant for “cats”.

- The challenge is to extract from many images some “cat essence”, while neglecting all the accidental detail.
- This is sometimes called the **Information bottleneck**.

Statistical Machine Learning (Applications)

- There are many applications of statistical machine learning, including:
- Self-driving cars (extract relevant information from large amounts of environmental data).
- Unsupervised data-analysis (novelty detection, recommendation systems, feature detection)
- Image recognition

Statistical Machine learning models

- The general approach of statistical machine learning is as follows:
  - **Establish a general model that can be adjusted to many different problems.**
  - **These models usually have a large number of tuning parameters.**
  - **Learn the relevant relationship. This equates to tuning the parameters of the model.**
  - **Test the model on real data.**
  - **Apply the model once you are satisfied that it works well**

Reasons for the recent success of AI

- Availability of lots of data from:
  - **Ability to record and share data**
  - **Cheap storage**
  - **Measurement devices**
- Fast computers and cheap RAM
  - **Makes it possible to deal with the large amount of data.**
  - **AI models are often extremely large with billions (!) of adjustable parameters. Training is only feasible in large server farms.**
- Breakthrough algorithms
  - **Backpropagation**
  - **Contrastive divergence**
  - **...**

Impact of statAI

- Social impact:
  - **Better treatment, more efficient services, better entertainment**
  - **A higher ability to quantify individuals, monitor their behaviour and know about them.**
  - **Automated surveillance of individuals. New tools for the state to check on its citizens.**
  - **Quantifying everything and make automated decisions.**
  - **Expertise in AI is concentrating in large corporations. What does this mean for the power of the state/democracy?**
- Environmental Impact:
  - **Server farms require vast amounts of energy for computing and cooling.**
  - **AI enables more efficient production methods, which leads to more consumption, which is not sustainable.**

Seeing the larger picture

- AI has profound potential to automate tasks that at present require considerable human intervention.
- As such it could lead to huge productivity gains and increase economic growth.
- Productivity gain could be used in two ways:
  - **Go home earlier and watch TV.**
  - **Stay in the office and do more.**
- Is economic growth a good thing?

Reference

- J. Copeland. Artificial Intelligence: a philosophical introduction. Blackwell, 1993. Sections 2.3, 3.2–3.5
- Strand, Kovac, Volker: The Circular Economy in Europe, Critical Perspectives on Policies and Imaginaries; see ch 7; https://www.taylorfrancis.com/books/9780429061028