CE213 Artificial Intelligence – Lecture 1

Module supervisor:

Prof. John Gan, Email: jqgan, Office: 4B.524

Homepage: https://www.essex.ac.uk/people/GANJO00207/john-qiang-gan

CE213 on Moodle: https://moodle.essex.ac.uk/course/view.php?id=3651&section=0

Module information, Assessment information, Teaching/reading materials, Other resources.

Acknowledgement:

Some of the lecture notes for this module were originally prepared by Dr. Paul Scott and updated by Prof. John Gan.

Information about this module

Lectures (Zoom webinars): Tuesdays, 2-4pm

Classes (Zoom meetings): Thursdays, 9-10am for CLAa01
3-4pm for CLAa02

Lab Exercises: at your own time, with detailed instructions and sample code provided on CE213 Moodle. Each lab exercise will be explained briefly in an appropriate class. *These are optional, but could be very useful.*

Assignment: one programming assignment with deadline in week 8

Progress Tests: through Moodle, one in week 6, another in week 11

Let's go to CE213 Moodle website and highlight some information provided there: https://moodle.essex.ac.uk/course/view.php?id=3651&section=0

Information about this module (2)

- CE213 is a challenging module, but we can work together to make it a success. I hope that you can do the following:

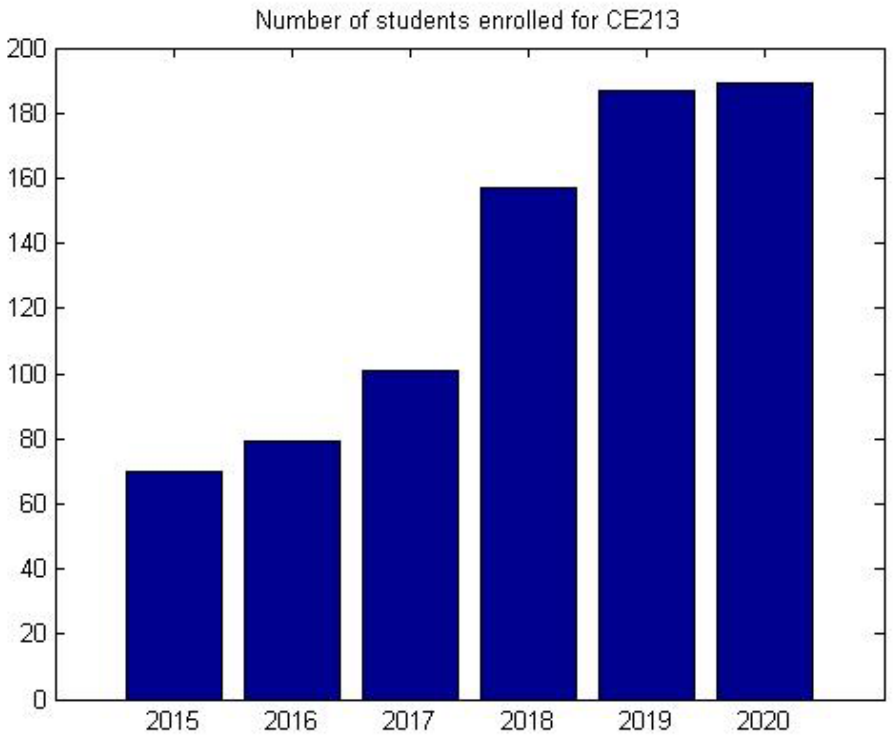
  1) Attend lectures and classes on time. Feel free to ask questions and take part in discussions.

  2) Skim relevant lecture notes and class problem sheet before a lecture or class.

  3) Have pen and paper at hand for problem solving exercises during classes.

  4) For lab exercises at your own time, at least try to run and understand the provided solutions.

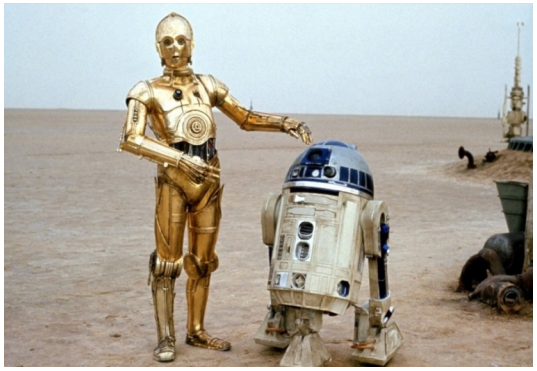
- Number of CE213 students in the past years:

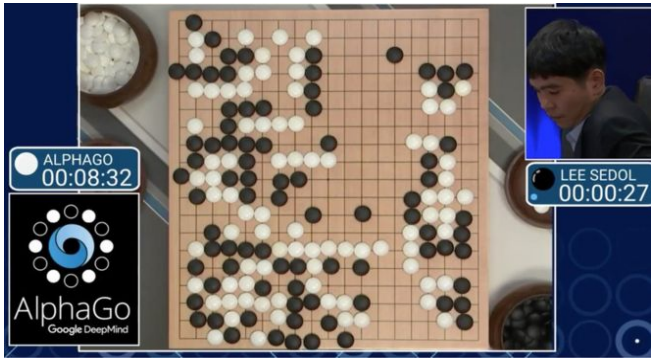


- In the recent years, average mark of this module is over 60%, and average SAMT score is over 4 out of 5.

Questions to be answered in this first lecture

1. What is AI?
2. Is AI possible?
3. How is AI possible?
4. What are the applications of AI?





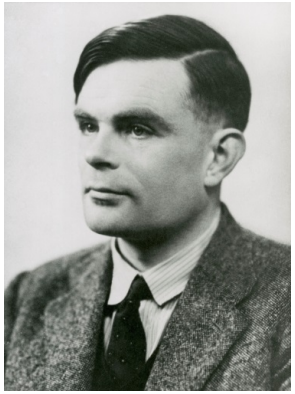
Do these pictures give you some ideas for answering the above questions?

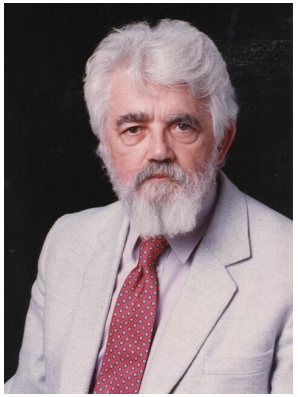
What Is AI? – A little bit history first

- 4th Century BC: Aristotle invented syllogistic logic, the 1st formal deductive reasoning system.

- 19th Century: Babbage's analytical Engine (1st general-purpose computer) and Lady Lovelace's first computer program (In her opinion, computers cannot really think).

- 20th Century: Alan Turing argued that machine intelligence is possible and proposed 'Turing Test' in 1950. John McCarthy coined the term 'Artificial intelligence' in 1955 to publicise a summer workshop.

Alan Turing 

John McCarthy 

What Is AI? – Some definitions

- There is still **no consensus** on the definition of AI. Here are two simple and famous definitions:

  (The science and engineering of making) intelligent machines. (McCarthy's definition)

  Machines that think. (Turing's definition)

- Strong AI: Machines that think.

- Weak AI: Machines that act as if they think.

- Specialised/Narrow AI: AI for one specific task.

- General AI or Artificial General Intelligence:

  Machines that can think, learn, act and feel like humans and are for multiple tasks. https://www.youtube.com/watch?v=x-QfL_BmZVE

Is AI Possible?

AI and The Philosophy of Mind

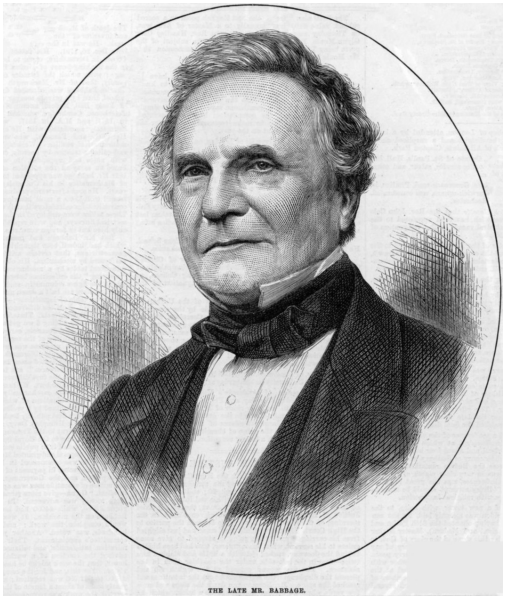
It seems that a ‘yes’ answer to this question is obvious now. However, is there a machine that is genuinely capable of thinking now?

For several hundred years people have been arguing about whether it was possible, in principle, to make a machine that can think. (Mainly argued by philosophers at the time when no computer was available)

The arguments took on a less abstract form about 180 years ago due to the design of the first general-purpose computer.

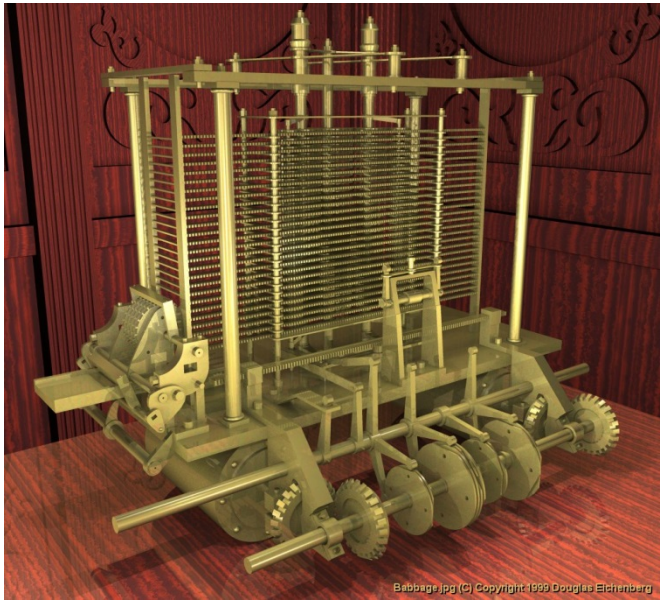
Babbage's Analytical Engine

The first general-purpose computer (it is mechanical!), which made the AI argument less abstract.



Charles Babbage

Lady Lovelace



Part of Babbage's Analytical Engine

The World's First Programmer

Lady Lovelace and Babbage collaborated on developing the Analytical Engine.

He designed the hardware.

She wrote the programs, was the world's first programmer.

Unfortunately the machine was never finished.

They ran out of money.

However Lady Lovelace did develop several key ideas that have been used by programmers ever since.

e.g., Subroutines

The Lady Lovelace's Objection

In her paper describing the Analytical Engine she said the following about the possibility that it could "think":

> "The Analytical Engine has no pretensions whatever to *originate* anything. It can do whatever we know how to order it to perform. It can *follow* analysis, but it has no power of *anticipating* any analytical relations or truths."

In other words:

> Computers cannot really think because they can only do what their programmers tell them to do.

Rebutting the Objection

In 1950, Alan Turing published a paper "Computing Machine and Intelligence", addressing a range of objections to the idea that computers could think. (http://www.csee.umbc.edu/courses/471/papers/turing.pdf)

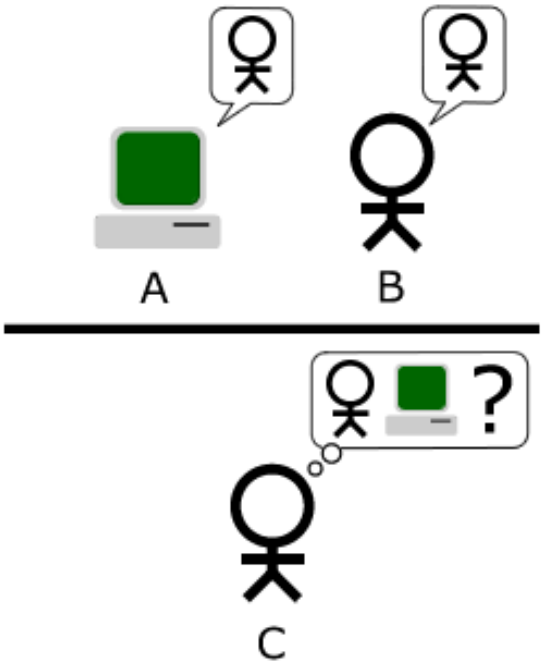
Including rebutting the Lady Lovelace's Objection.

Turing had two counterarguments:

- In practice, the results of programming often surprise us.
  (Lady Lovelace never got to run any of her programs!)
- The possibility that machines could be programmed to learn.

The Turing Test

In the scenario shown in the figure, A is a computer, B is a human, and C is an interrogator. C asks A and B questions, and determines on the basis of their replies which one is the computer. Communication is only by typewritten text.



Turing argues that if C cannot distinguish between A and B then we would have to concede that the computer is as capable of thought as a human being. This is essentially a **behavioural** argument - If an entity behaves as if it thinks then it thinks.

Should more interrogators be involved in Turing Test to make it more reliable? Even if so, is Turing Test valid for assessing AI? Is Turing Test itself a refutation of Lady Lovelace's Objection?

Searle's Chinese Room

(https://en.wikipedia.org/wiki/John_Searle)

The Turing Test is very famous, but also controversial. One of the criticism was given by John Searle. He proposed the so-called Searle's Chinese Room.

Consider the following two definitions of artificial intelligence:

*A machine that **thinks** like a person. (**strong AI**)*

*A machine that **acts as if it thinks** like a person. (**weak AI**)*

Searle has no quarrel with weak AI, but he profoundly disagrees with the view that Weak AI and Strong AI are equivalent, like what Turing argued for the Turing Test.

Furthermore, he does not believe Strong AI is possible.

Searle's Chinese Room (2)

The Chinese Room – A thought experiment

In order to demonstrate the difference between Weak AI and Strong AI, Searle developed an analogy known as the "Chinese Room".

> *Imagine an English speaking person locked in a room, who has no knowledge of Chinese or Russian.*
>
> *Also in the room is a (large) set of rules, written in English, for manipulating the symbols of Chinese text and producing Russian translations.*
>
> *Periodically this person is given texts written in Chinese.*
>
> *By mechanically following these rules, the person produces Russian translations of the Chinese text.*

Searle's Chinese Room (3)

Searle argues

The room+person system behaves as if it understands Chinese.

However, it is obvious that no understanding is involved.

Hence, behaviour does not imply the existence of a mind that understands.

In other words, weak AI does *not* imply strong AI.

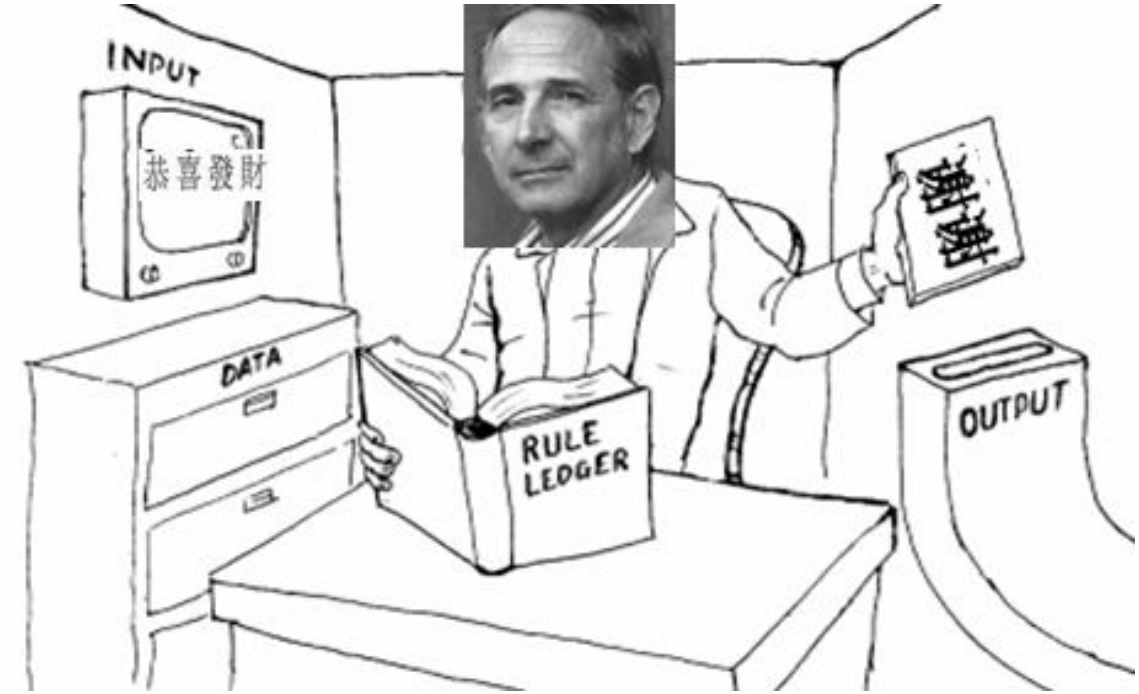
Thus, if we accept Searle's argument, then Turing test is not a valid criterion/method for assessing whether a machine can think.

What would be a better way to test whether (strong) AI is possible?

(Something like the experiment described in Ex Machina?)

(Robot college student test?)

Searle's Chinese Room (4)



Another version of the Searle's Chinese Room: Someone or a computer can communicate in Chinese without understanding Chinese in the 'Chinese Room', who could pass the Turing Test without understanding/thinking involved.

So, Is AI possible?

- There has been evidence of weak AI or specialised AI, e.g., AlphaGo, self-driving cars, Alexa, personalised search engine. We may say that weak AI is possible.

- However, no computer program has really passed the Turing Test (https://en.wikipedia.org/wiki/Loebner_Prize).

- And there are no well-recognised/accepted criteria for assessing strong AI or general AI.

- Anyway, if AI is defined as intelligent machines then AI is possible, but how is AI possible and what can AI do?

How Is AI Possible?

- **Three fundamental approaches:**

  Search as general problem solver – Unit 1

  The 'generate and evaluate' approach:

  state space representation of problems, search strategies (e.g., A*, minimax, MCTS), evaluation criteria

  Expert systems – Unit 2

  knowledge representation, rule interpretation, reasoning with uncertainty

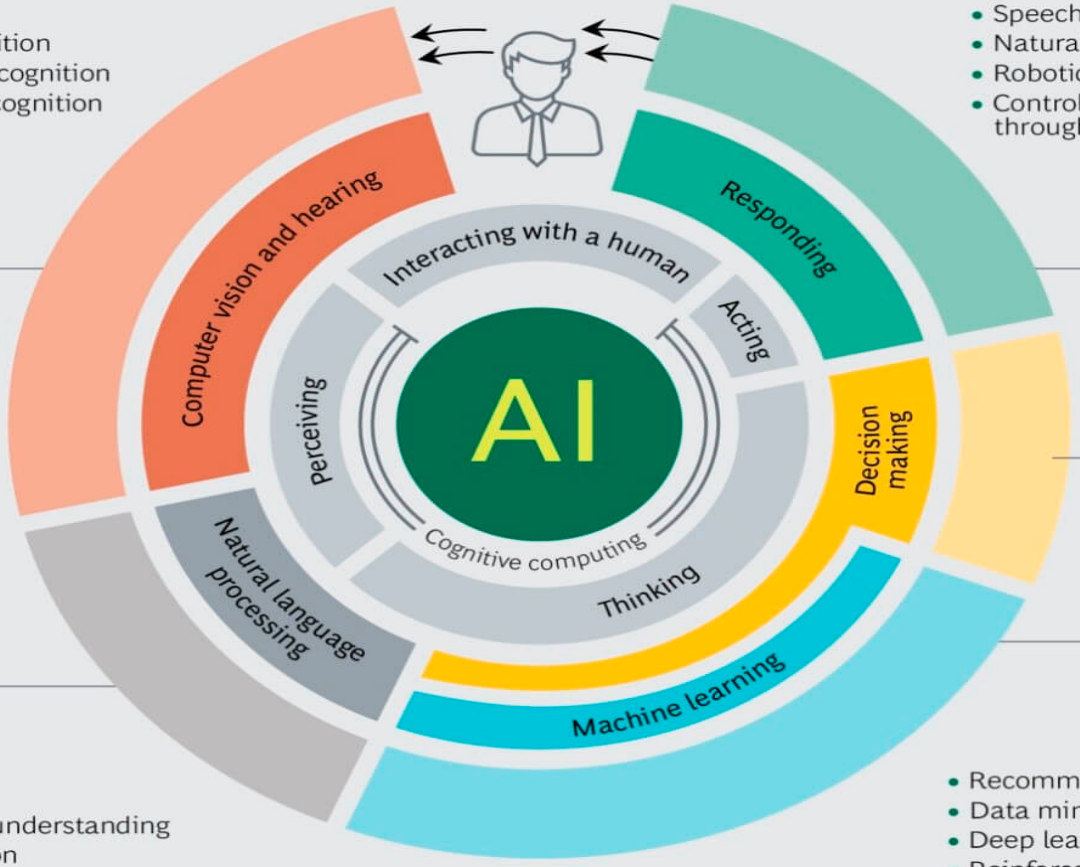
  Machine learning – Unit 3

  learning environment, models and algorithms

About 90% of an AI course would try to answer this question!

AI Applications

EXHIBIT 1 | AI and Robotics Technologies Come in Many Forms, Giving Rise to a Broad Variety of Applications



- Speech recognition
- Handwriting recognition
- Optical character recognition
- Image and video recognition
- Facial recognition

- Speech synthesis
- Natural language generation
- Robotic process automation
- Control of other systems through APIs

- Case-based reasoning
- Expert systems

- Natural language understanding
- Machine translation
- Sentiment analysis

- Recommender systems
- Data mining
- Deep learning
- Reinforcement learning
- Unsupervised learning
- Supervised learning

Source: BCG analysis.
Note: APIs = application programming interfaces.

Summary

What is AI?

Intelligent machines.

Machines that **think and learn** like people.

Machines that **act** rationally/intelligently.

Is AI possible?

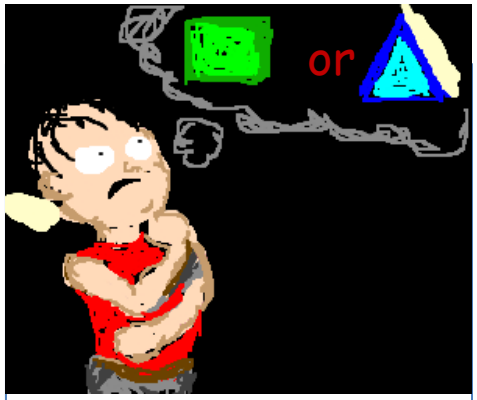
Lady Lovelace's Objection

The Turing Test

Searle's Chinese Room

How is AI possible?

What are the applications of AI?



Have you got the answers to these questions now?

What are you going to learn next?