Artificial Intelligence (AI)

Mendlesham Computer Club

By Giles Godart-Brown

AI receives a lot of bad press. e.g deep fakes

But when they get it right

What is Artificial Intelligence and how does it work?

By Giles Godart-Brown

Can machines think like a human? – The Turing Test

- If a machine can engage in a conversation with a human without being detected as a machine, it has demonstrated human intelligence.

- Even the most advanced AI systems today do not consistently pass the Turing test, but for how long?

- This has a lot to do with the complexities and subtleties of the English language.

- Tasks that don't involve these subtleties are often better solved by AI than humans e.g.

  – Maths problems
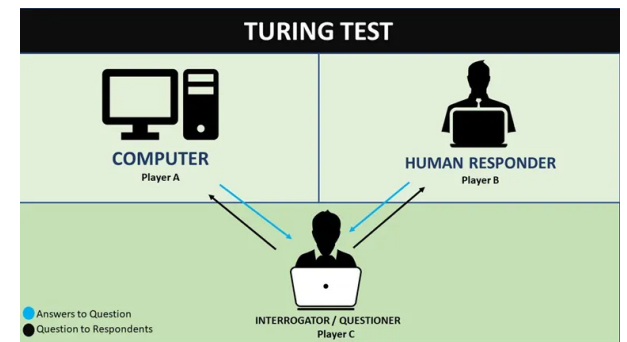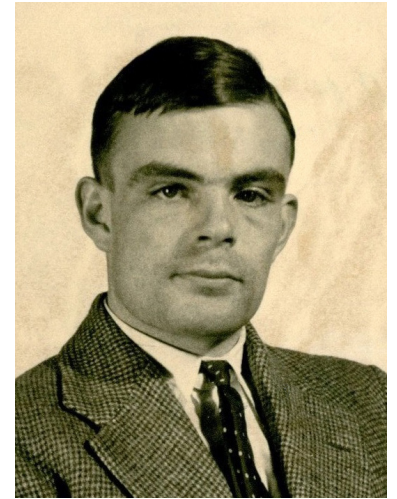
What technologies make this possible?

Mendlesham Computer Club

By Giles Godart-Brown

Enabling technologies

There are 4 major technologies used by AI

- Machine Learning

- Neural Networks

- Large Language Models

- Large data analysis

Machine Learning

- Machine learning has been around for a long time.
- Machines adapt how they work based on previous experience.
- A very simple example – Smart Heating.
  - Every day the system learns how quickly the room warmed up so on the following day it can know when to turn on the heating to get to the desired temperature at the right time. This saves fuel on warmer days.
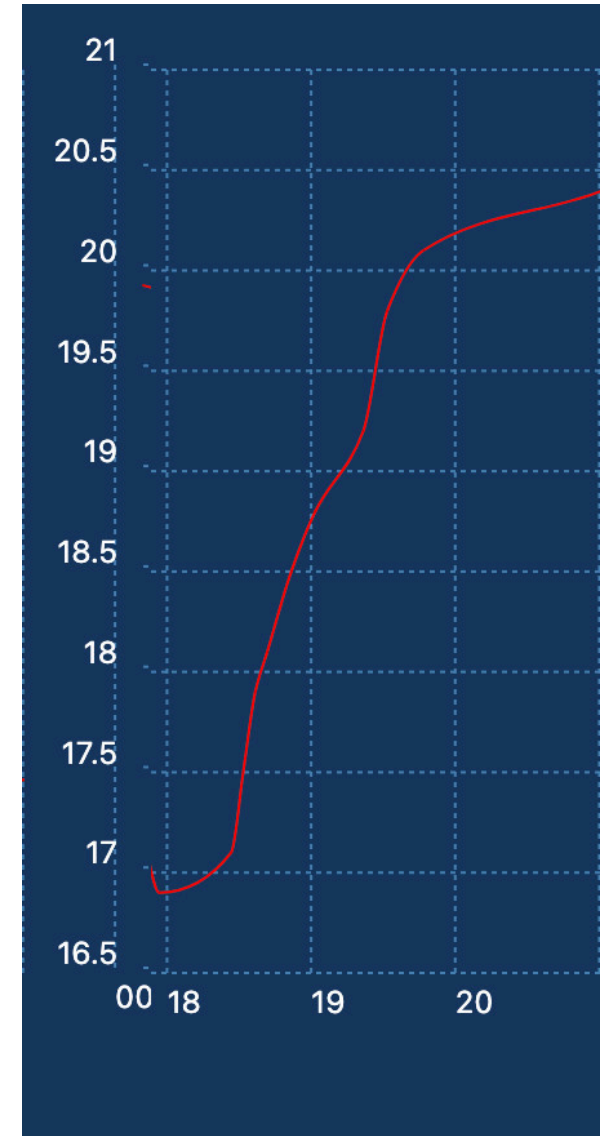
A much more complex example

From the Royal Institution Christmas Lectures 2023

How does it learn?

The more it is trained, the more accurate it gets until it reaches a threshold where it can make its own decisions.

- **Supervised learning**
  - The system is trained in a series of tasks and then a human supervisor tells it if it got them correct or not, the system remembers the results and can use this to make the right decisions in the future.
- **Un-supervised learning**
  - Like dog training - it does random actions until it gets it right, then it repeats the positive action.
  - For example, game playing where it tries many strategies and checks the score, eventually it works out the strategies that score the highest.

What is a Neural Network?

A network of interconnected nodes to process data and perform tasks.

Is it big?

Is it hairy?

Is it big and hairy?

OK, it's a mammoth
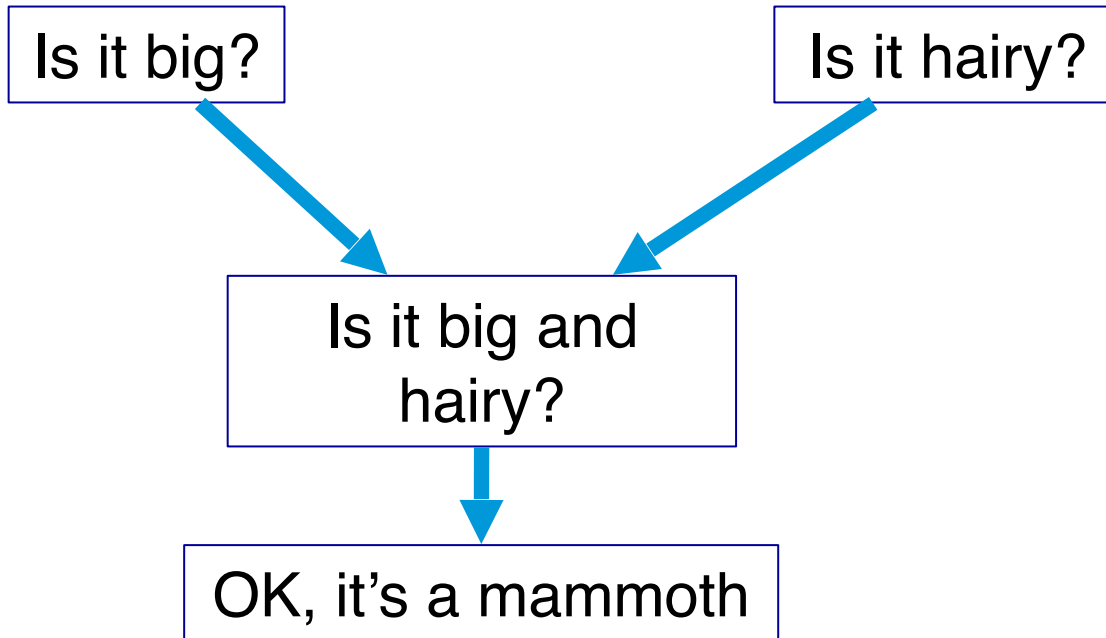
Large Language Models

A Large Language Model (LLM) is a type of Artificial Intelligence that can process, understand, and generate human language.

For example - Predictive text on mobile phones has now evolved from just predicting the current word from a learned dictionary to predicting the next word via Large Language Models.
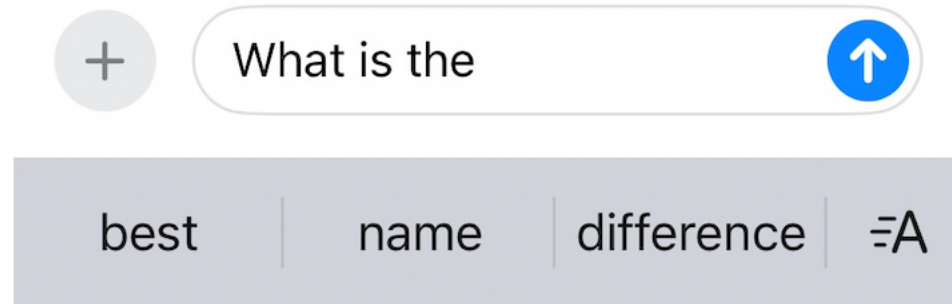
Large Language Models e.g. Siri, Alexa, ChatGPT

- Large Language Models work by recognising patterns in vast amounts of text.

- They use a sophisticated neural network architecture to predict the most likely next word or phrase, and then they generate contextually relevant responses based on their training.

- They can generate convincing but incorrect or biased information, as **they do not have real-world understanding or common sense**.

- They also cannot access real-time information or data beyond their training.

Large Language Models - text to speech

- They also learn how to translate text into speech.
- The accuracy of the predictions depends on the size of the data set it has used to learn, but new systems can clone your voice with less than 5 minutes of your real speech e.g. <u>Virtually Michael Parkinson</u>
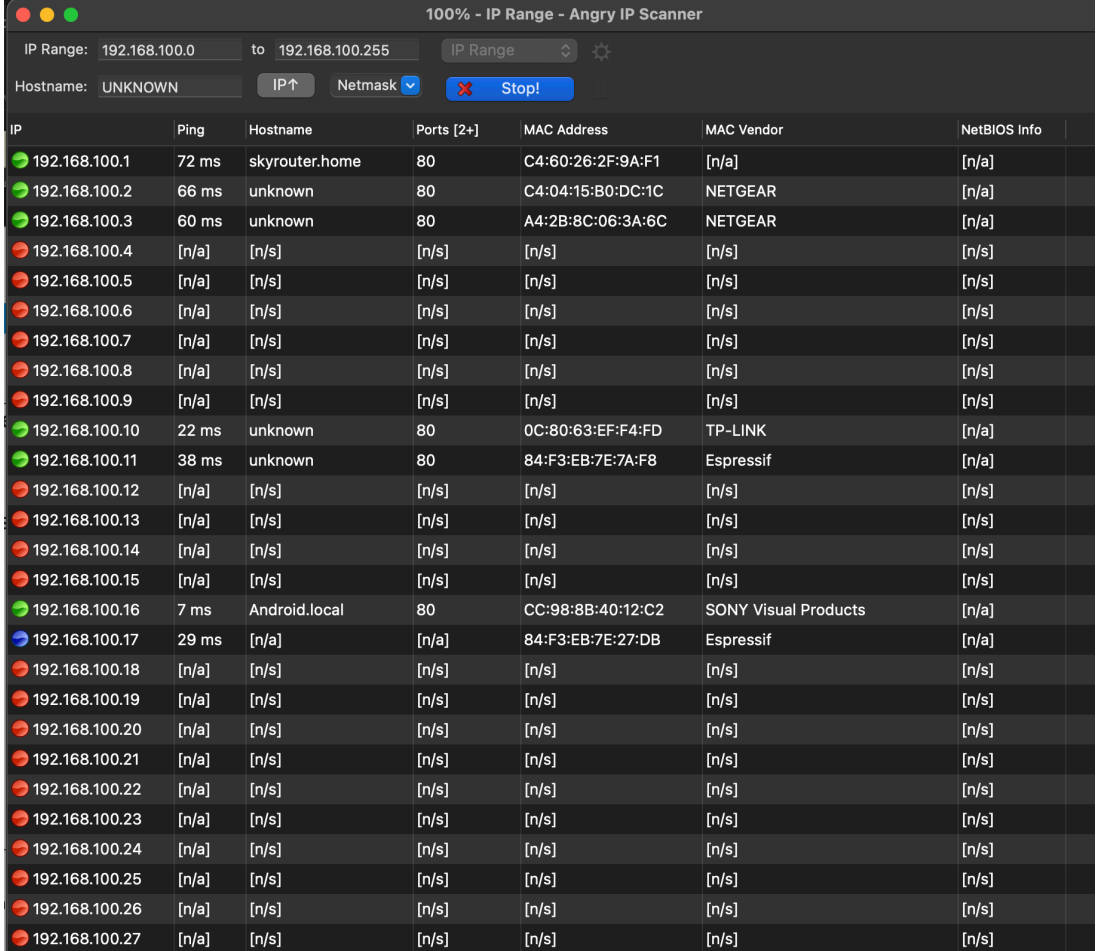
AI uses huge datasets to make its decisions

- Google search has a massive dataset of knowledge.
- Chat GPT uses a data set of about 45 TB, this is roughly equivalent to 450 Km of books on a shelf (280 miles or London to Sunderland).
- These datasets are getting bigger all the time
  - Whenever you upload a photo to social media along with your location, your name or your friends name, AI can discover this and use it for advertising to yourself, your friend or people in your demographic at that location.
  - Whenever you buy anything online, this information can be used by the vendor to show you related items.

How has this changed computer design?

By Giles Godart-Brown

Parallel Processing

- AI relies on analysing huge amounts of data

- This is only possible with Parallel Processing where a task is broken down into multiple smaller tasks each assigned to an individual processor, and these run in parallel
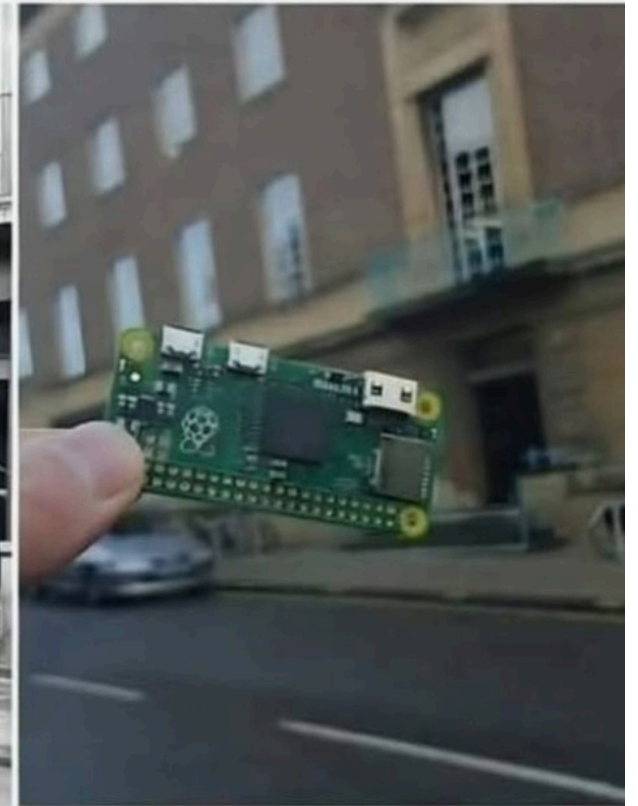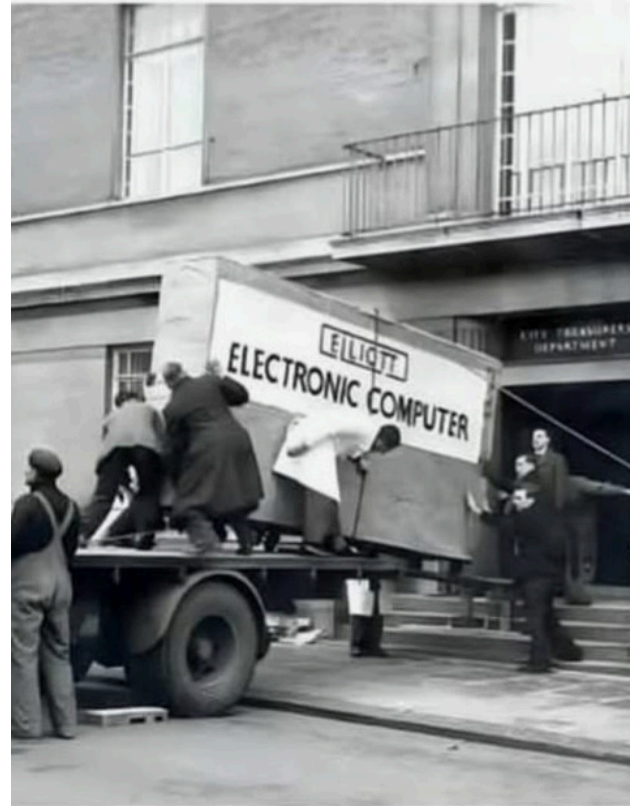
Effect on chip design

1957 **2024**

All modern computer chips have multiple processors

The Intelligent Processing Units

- Processors specifically designed to perform AI tasks and to handle huge amounts of data in parallel.

- Each Colossus MK2 GC200 IPU has 59.4 billion transistors, and 1,472 processing cores. Laptops typically have 4 cores.

- You could fit 188,416 processing cores in standard 19inch rack.

- The human brain has about 86 billion neurones

More demonstrations

Mendlesham Computer Club

By Giles Godart-Brown

AI in everyday life

- Maps and routing

- Weather forecasting

- Traffic light control

- Predictive text

- Smart speaker

- CCTV

- Translation apps

- Facial recognition

- Healthcare

- Autonomous vehicles

- Google search

- Research (ChatGPT) https://chatgpt.com/

- The arts https://youtu.be/p5vLTHXyTn8?t=2788

Questions

Mendlesham Computer Club

By Giles Godart-Brown