Tuesday 23rd September – What is artificial intelligence?

Early influences

* Aristotle (c. 384 -322 BCE) formulated laws governing the rational mind.
* Ramon Llull (c. 1232 – 1315 ) devised a “machine” of paper wheels as a system with which to reason.
* Thomas Hobbes (1651) in his book *Leviathan*, talked of “artificial animals”.
* René Decartes (1596 - 1650) on **rationalism**, Francis Bacon (1561 – 1626) and John Locke (1632 – 1704) on **empiricism**, David Hume (1711 -1776) and Immanuel Kant (1785) on **induction**, and Ludwig Wittgenstein and Bertrand Russell on **logical positivism**, Jeremy Bentham (1832) and John Stuart Mill (1863)on **utilitarianism**.
* Formal logic of George Boole (1815 -1925) and Gottlob Frege (1848 -1925) led to first order logic which is a bedrock of earlier knowledge-based systems.
* Probability theory of Gerolamo Cardano (c. 1501 -1576) and Blaise Pascal (c. 1623-1662), and, later, the rules of Thomas Bayes (c. 1702 -1761) forms a crucial tool for AI.
* Some popular machine learning algorithms such as linear regression (we will learn more about this shortly) was first used for predicting planetary motions by Adrien- Marie Legendre (c. 1752 – 1833) and Karl Friedrich Gauss (1777 -1855), the term being coined by Francis Galton in 1886.

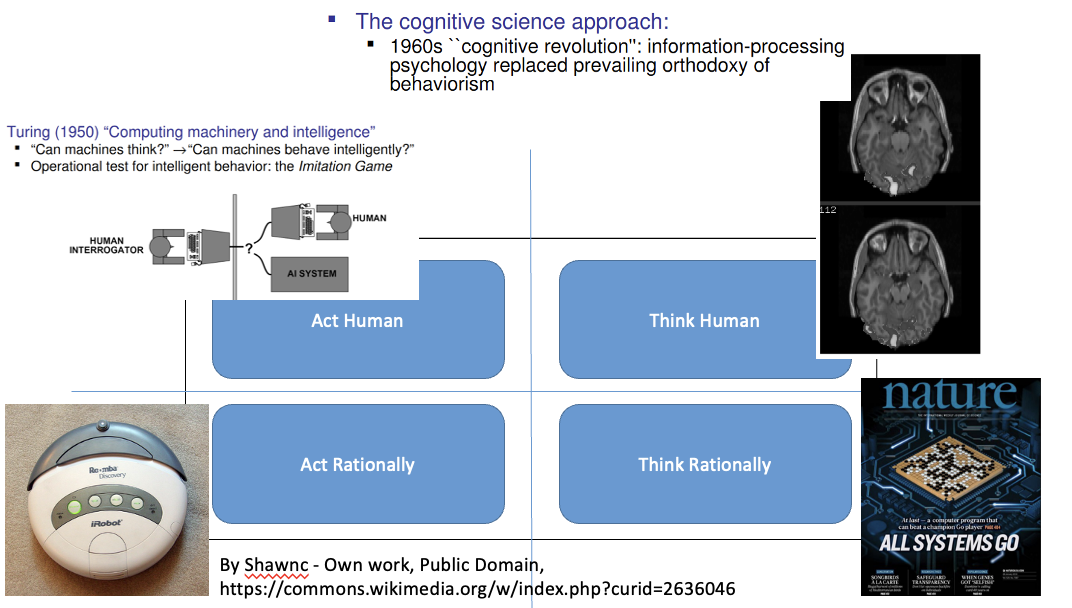
Defining the field

* “The field of artificial intelligence, or AI, is concerned with not just understanding but also building intelligent entities – machines that can compute how to act effectively and safely in a wide variety of novel situations.”
* Understand learning, reasoning, and perception while achieving specific objectives such as playing chess, proving mathematical theorems, writing poetry, driving a car, diagnosing disease.
* Tackle a variety of tasks such as describing what’s happening in a video, question answering, robots operating in a factory, real time translation, visual recognition of objects, optimization of investment and trading, and voice/speech recognition.

What is AI?

* “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 which, at the moment, **people are better**.” (Rich and Knight, 1991)
* “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)
* “AI is the ability of a machine to display human-like capabilities such as reasoning, learning, planning and creativity.” <https://www.europarl.europa.eu/news/en/headlines/society/20200827STO85804/what-is-artificial-intelligence-and-how-is-it-used>
* “Whilst there is no universal definition of AI, it can be considered as an evolving set of technologies that enable computers to simulate elements of human behaviour such as learning, reasoning and classification, by analysing data to model some aspect of the world and predict and anticipate possible future events.” <https://commonslibrary.parliament.uk/research-briefings/cbp-8152/>
* “The Artificial Intelligence Act (AIA) defines AI broadly as a suite of software development frameworks that encompass machine learning, expert and logic systems, and Bayesian or statistical approaches.” <https://datainnovation.org/2021/05/the-artificial-intelligence-act-a-quick-explainer/>

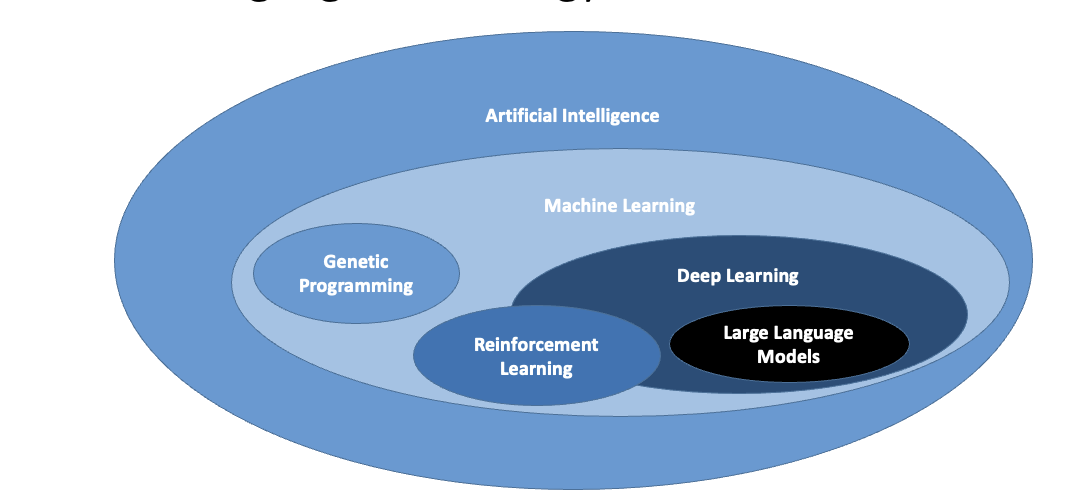
Taxonomy of AI



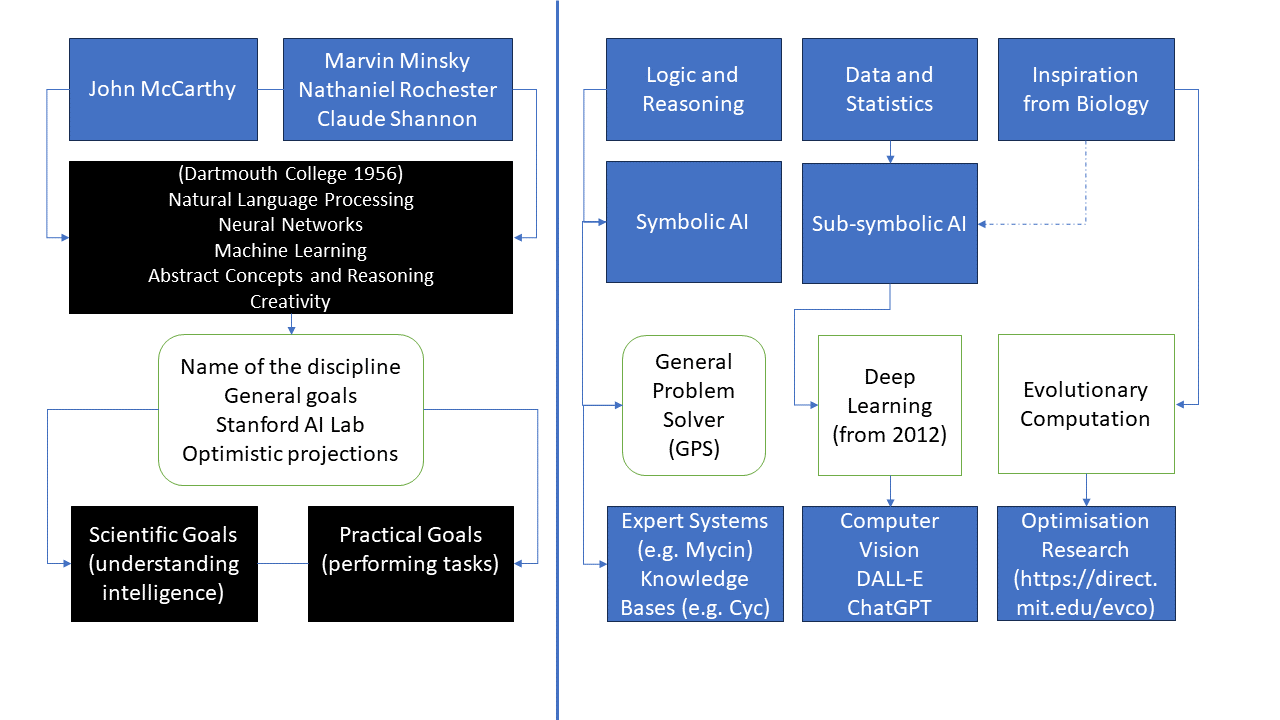
Measuring Intelligence

* Artificial **Narrow/Weak** Intelligence – AI trained to perform specific tasks.
* Artificial **General/Strong** Intelligence – trained to learn any task or at least wide variety of tasks and/or “where a machine would have an intelligence equal to humans; it would have a self-aware consciousness that has the ability to solve problems, learn, and plan for the future.”
* Artificial **Super Intelligence** – “would surpass the intelligence and ability of the human brain”

Disentangling Terminology

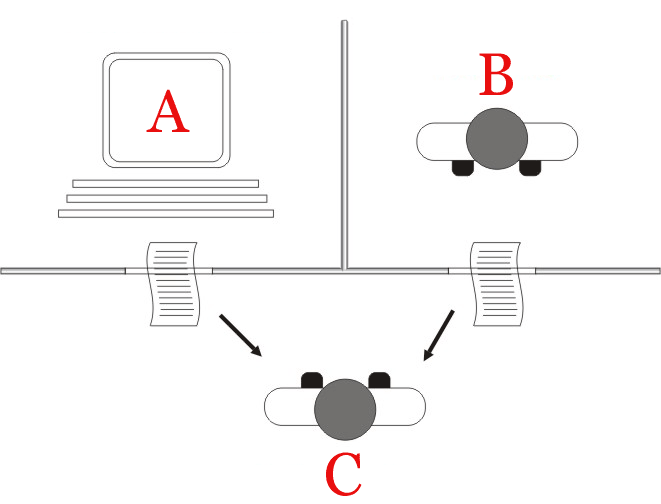


History in One Slide



AI Winters

Enter Turing and Asimov

 The "standard interpretation" of the Turing test, in which player C, the interrogator, is given the task of trying to determine which player – A or B – is a computer and which is a human. The interrogator is limited to using the responses to written questions to make the determination.

The fictional "Handbook of Robotics, 56th Edition, 2058 A.D.“ referenced in the story “Runaround” from I, Robot:

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

Early Artificial Neural Networks

* “McCulloch - Pitts neuron” - the first mathematical model of a neural network.
* Minsky also built, in 1951, the first randomly wired neural network learning machine, [SNARC](https://en.wikipedia.org/wiki/Stochastic_Neural_Analog_Reinforcement_Calculator).
* Rosenblatt was best known for the [Perceptron](https://en.wikipedia.org/wiki/Perceptron), an electronic device which was constructed in accordance with biological principles and showed an ability to learn.

Games and AI

* Deep Blue's victory is considered a milestone in the [history of artificial intelligence](https://en.wikipedia.org/wiki/History_of_artificial_intelligence) and has been the subject of several books and films.
* [AlphaGo Master](https://en.wikipedia.org/wiki/AlphaGo_Master) (white) v. [Tang Weixing](https://en.wikipedia.org/wiki/Tang_Weixing) (31 December 2016), AlphaGo won by resignation. White 36 was widely praised.
* Ken Jennings (left) faces off against supercomputer Watson (center) and his fellow Jeopardy champion, Brad Rutter (right).

AI Chat Agents

* **ELIZA** is an early [natural language processing](https://en.wikipedia.org/wiki/Natural_language_processing) [computer program](https://en.wikipedia.org/wiki/Computer_program) created from 1964 to 1967[[1]](https://en.wikipedia.org/wiki/ELIZA) at [MIT](https://en.wikipedia.org/wiki/MIT) by [Joseph Weizenbaum](https://en.wikipedia.org/wiki/Joseph_Weizenbaum).
* The **Loebner Prize** awarded prizes to the [computer programs](https://en.wikipedia.org/wiki/Computer_program) considered by the judges to be the most human-like. The prize is reported as defunct since 2020.
* A.L.I.C.E wins multiple times (2000 – 2004).
* **Tay** released by [Microsoft Corporation](https://en.wikipedia.org/wiki/Microsoft_Corporation) via [Twitter](https://en.wikipedia.org/wiki/Twitter) in 2016. Bot began to post inflammatory and offensive content. Shut down the service only 16 hours after its launch.

Increasing Interest in Responsible AI

* Lethal autonomous weapons (e.g. autonomous weaponized drones).
* Surveillance and persuasion (e.g. use of Cambridge Analytica for swaying votes).
* Biased decision making (e.g. unfair recruitment and selection).
* Impact on Employment (e.g. unemployment due to AI replacing humans).
* Safety-critical applications (e.g. nuclear; health; climate; epidemic).
* Cybersecurity (e.g. anything can be hacked).
* Impact on climate change and sustainability.
* Rise of EU AI Act, Scottish AI Register, AI auditing frameworks, taxonomies of risks and harms.

Summary

* Defining the field
* Taxonomy of AI
* Narrow, General, Super Intelligence
* Relationship between AI, Machine Learning, Deep Learning
* Brief history
* Evolution of themes