Artificial intelligence: Ideas & their evolution

Lecture 1 of “Mathematics and AI”

Outline

1. What is AI?
2. Approaches to AI
3. The Dartmouth Workshop
4. Separating the What and How
5. Course information

What is AI?

“As soon as it works, nobody calls it AI anymore.”

Attributed to John McCarthy

Intelligence

intelligence noun

in·tel·li·gence in-ˈte-lə-jən(t)s

Source: www.merriam-webster.com

Intelligence

intelligence noun

in·tel·li·gence in-ˈte-lə-jən(t)s

Synonyms of *intelligence* >

1 a (1) : the ability to learn or understand or to deal with new or trying situations : **REASON**

also : the skilled use of reason

(2) : the ability to apply knowledge to manipulate one's environment or to think abstractly as measured by objective criteria (such as tests)

b : mental acuteness : **SHREWDNESS**

c Christian Science : the basic eternal quality of divine Mind

2 a **INFORMATION**, **NEWS**

b : information concerning an enemy or possible enemy or an area

also : an agency engaged in obtaining such information

3 : the act of understanding **COMPREHENSION**

Source: www.merriam-webster.com

Reason

Given an input, a system returns a correct output that follows from applying a set of logical rules

Source: www.merriam-webster.com

Information

Given a query, a system returns the correct answer based on its internal knowledge base

Comprehension

???

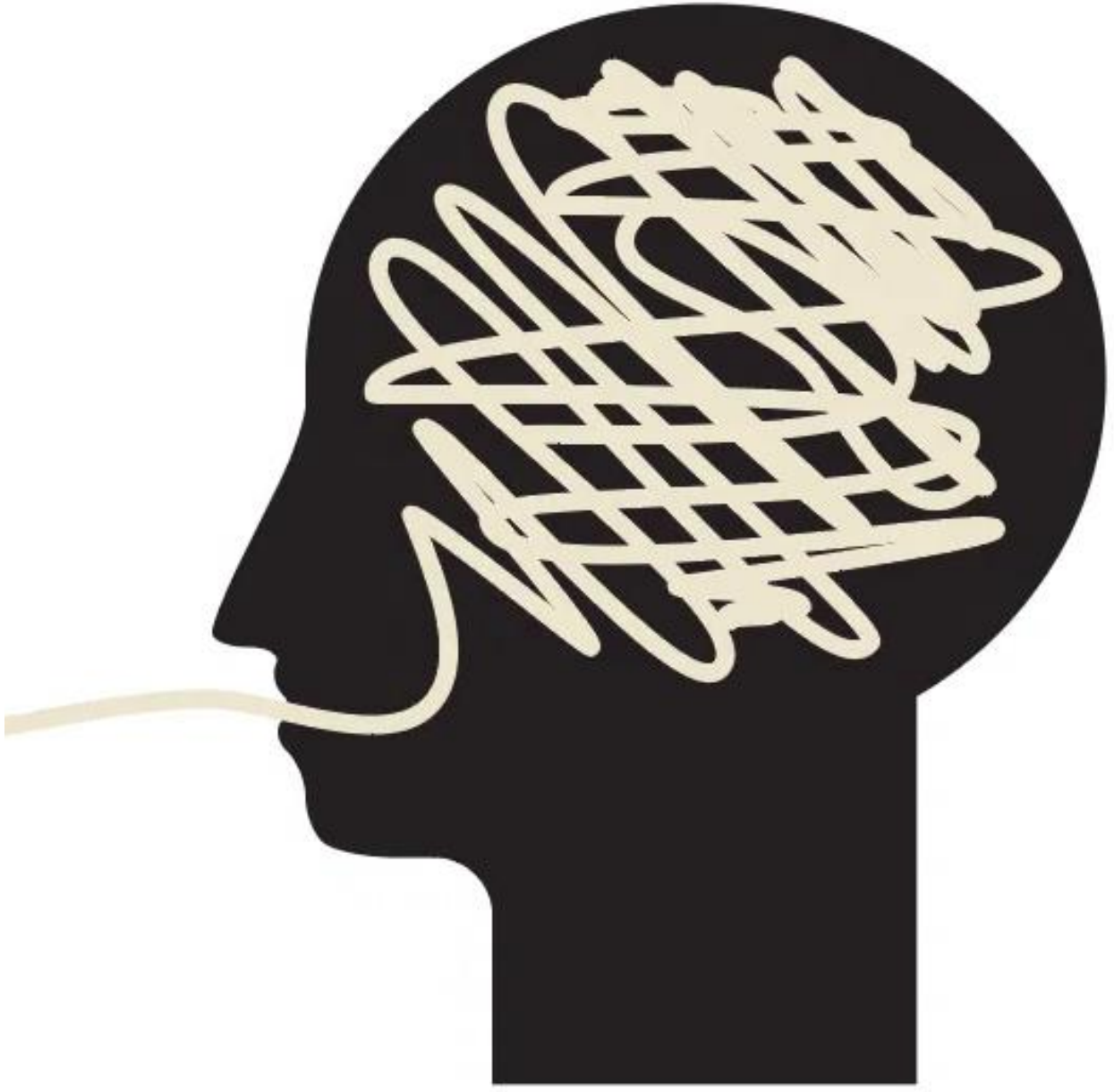
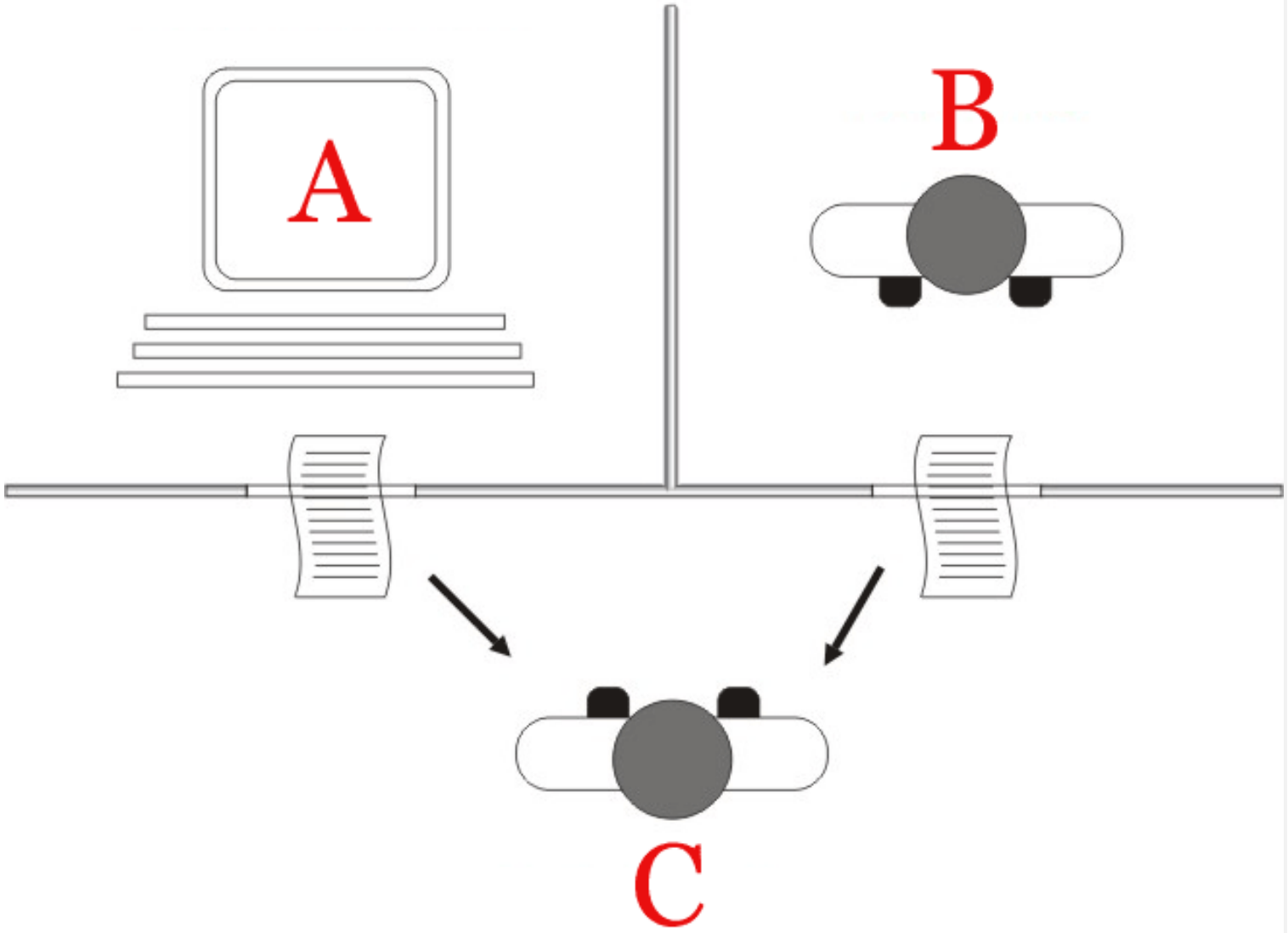


Image source: https://medium.com/kocuniversity/the-bases-of-the-mind-the-relationship-of-language-and-thought-a0bf30375528

The Turing Test

- Proposed by Alan Turing (1950)
- Q: How could we possibly test "machine intelligence"?
  - Separate machine intelligence from machine comprehension
  - View humans and machines as input-output systems
- Thought-experiment rather than a precise experimental method



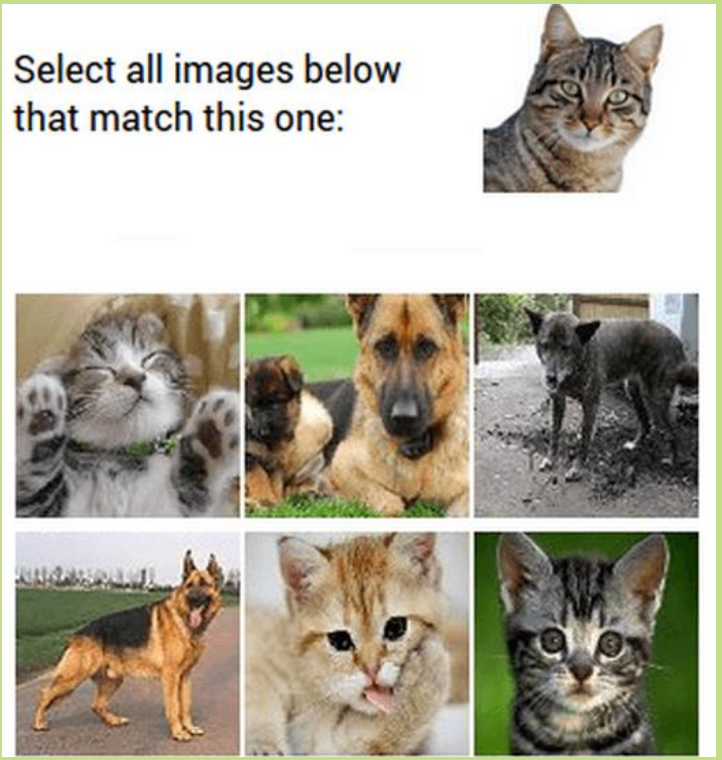
“Researchers have put **OpenAI's GPT-4** up against the Turing Test, and according to a new study[1] the impressive large language model **passed the test**. The not-yet-peer-reviewed-study asked 500 people to engage with four different respondents. One of those respondents was a human, another was a 1960s-era AI called ELIZA, another was OpenAI's less sophisticated GPT-3.5, and finally GPT-4.

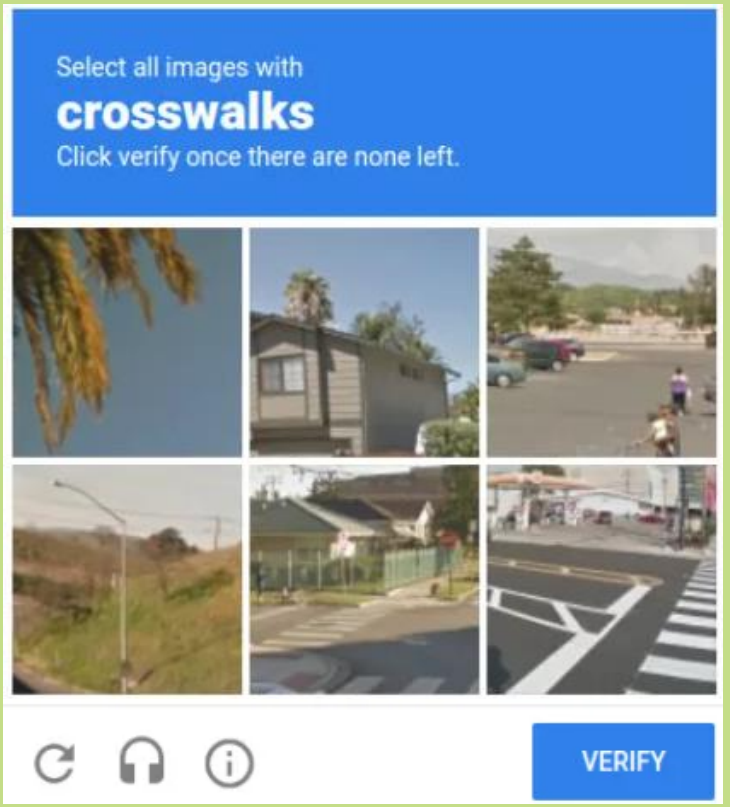
The paper states each conversation lasted approximately five minutes, and according to the study's results, **humans found GPT-4 to be a human 54% of the time**. These results lead researchers to claim GPT-4 has already passed the Turing Test. However, the study also indicated **study participants believed the human respondent was a human 67% of the time**, while ELIZA scored just 22%, and GPT-3.5 scored 50%.”

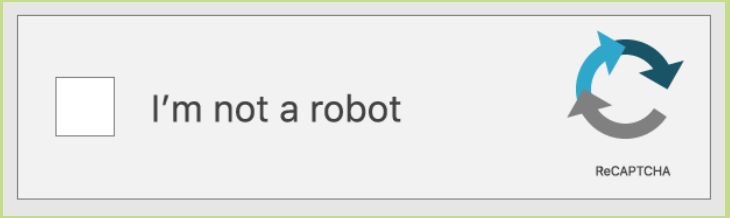
Source: Jak Connor for tweaktown.com, June 17, 2024 [2]

1: https://arxiv.org/abs/2405.08007

2: https://www.tweaktown.com/news/98851/scientists-discover-ai-has-already-passed-the-notorious-turing-test/index.html







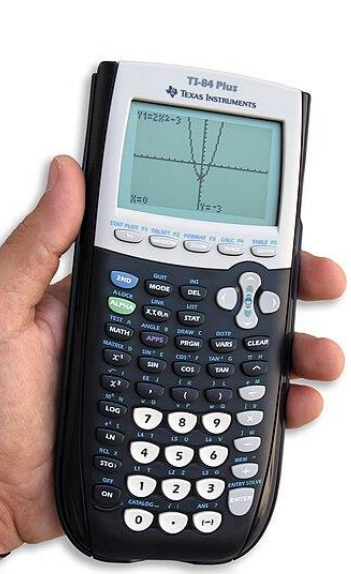
CAPTCHA: Completely Automated Public Turing test to tell Computers and Humans Apart

Weak AI and strong AI

Artificial general intelligence (AGI)

Weak AI — Expert systems — Strong AI

Can perform specific tasks
Narrow scope
Works in fixed environment



?

Can perform many tasks
Versatile, broad scope
Adjusts to changes of environment

is chatgpt a strong or weak ai?

ChatGPT is considered an example of "weak AI," also known as narrow AI. It is designed to perform specific tasks, such as generating text based on input prompts, and it doesn't possess general cognitive abilities or consciousness. It operates within the constraints of its programming and training data, without understanding or awareness in the way humans do. "Strong AI," or artificial general intelligence (AGI), would have the ability to understand, learn, and apply knowledge across a broad range of tasks at a human-like level, which is not something ChatGPT can do.

2/2

Approaches to AI

“Artificial intelligence is the science of making machines do things that would require intelligence if done by men.”

Marvin Minsky 1968

Logical approach

- Idea: If a computer is given the right set of starting axioms and an ability of logical reasoning, should it not be able to deduce all true statements eventually?
- Goal: A machine that can deduce the ***truth value of any given input statement***
- Formal logic, logical programing, computational theory
- Relevant for traditional expert systems

P	Q	$\neg P$	$P \to Q$	$\neg P \wedge (P \to Q)$
T	T	F	T	F
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

□

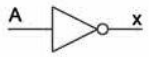
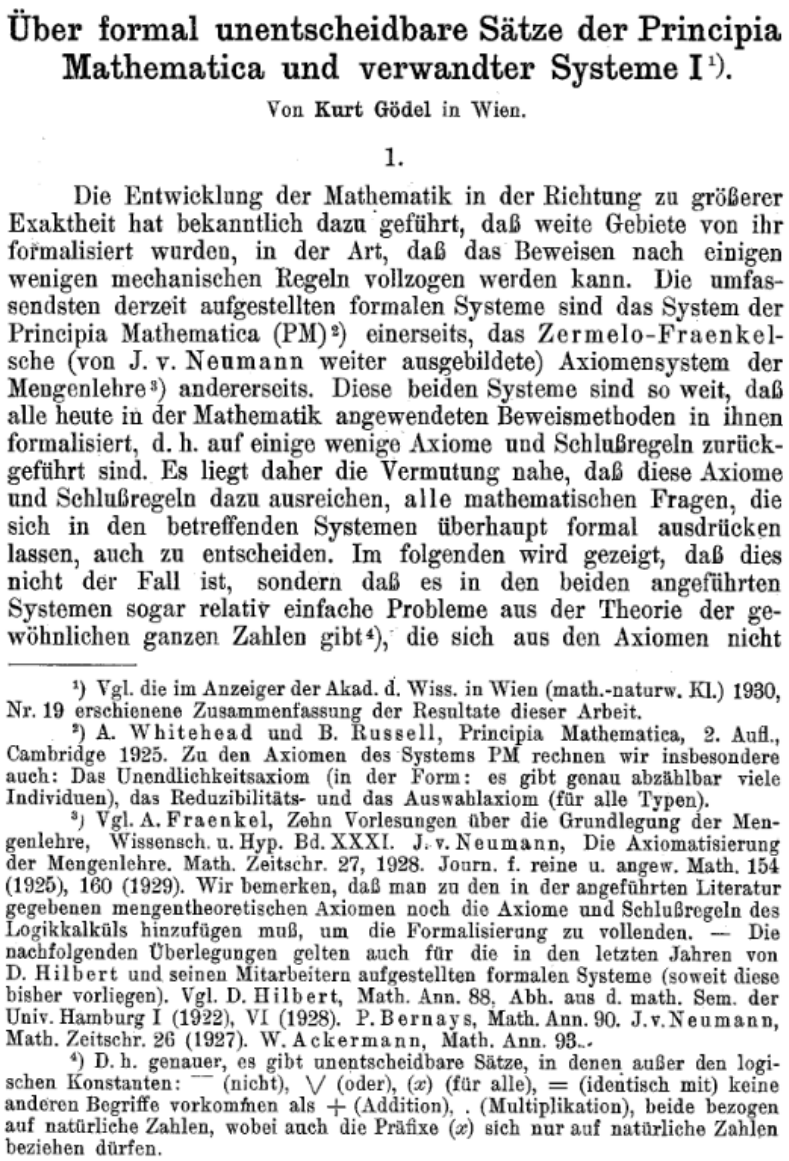
Name	NOT	AND	NAND	OR	NOR	XOR	XNOR
Alg. Expr.	$\overline{A}$	AB	$\overline{AB}$	$A+B$	$\overline{A+B}$	$A \oplus B$	$\overline{A \oplus B}$
Symbol							
Truth Table	A X 0 1 1 0	B A X 0 0 0 0 1 0 1 0 0 1 1 1	B A X 0 0 1 0 1 1 1 0 1 1 1 0	B A X 0 0 0 0 1 1 1 0 1 1 1 1	B A X 0 0 1 0 1 0 1 0 0 1 1 0	B A X 0 0 0 0 1 1 1 0 1 1 1 0	B A X 0 0 1 0 1 0 1 0 0 1 1 1

Image sources:
https://sites.millersville.edu/bikenaga/math-proof/truth-tables/truth-tables.html
https://michellescomputerscience.wordpress.com/2016/08/31/ch-3-logic-gates-and-logic-circuits/

Goedel's incomplete-ness theorem

- Can there be a formal language (math or similar) in which **all** sentences that can be stated are either provably true or provably untrue?
- Provides limits of provability of "the universe" of a formal language
- Provides limits on the capabilities of logic-based systems



Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I[1].

Von **Kurt Gödel** in Wien.

1.

Die Entwicklung der Mathematik in der Richtung zu größerer Exaktheit hat bekanntlich dazu geführt, daß weite Gebiete von ihr formalisiert wurden, in der Art, daß das Beweisen nach einigen wenigen mechanischen Regeln vollzogen werden kann. Die umfassendsten derzeit aufgestellten formalen Systeme sind das System der Principia Mathematica (PM)[2] einerseits, das Zermelo-Fraenkelsche (von J. v. Neumann weiter ausgebildete) Axiomensystem der Mengenlehre[3] andererseits. Diese beiden Systeme sind so weit, daß alle heute in der Mathematik angewendeten Beweismethoden in ihnen formalisiert, d. h. auf einige wenige Axiome und Schlußregeln zurückgeführt sind. Es liegt daher die Vermutung nahe, daß diese Axiome und Schlußregeln dazu ausreichen, alle mathematischen Fragen, die sich in den betreffenden Systemen überhaupt formal ausdrücken lassen, auch zu entscheiden. Im folgenden wird gezeigt, daß dies nicht der Fall ist, sondern daß es in den beiden angeführten Systemen sogar relativ einfache Probleme aus der Theorie der gewöhnlichen ganzen Zahlen gibt[4], die sich aus den Axiomen nicht

[1]) Vgl. die im Anzeiger der Akad. d. Wiss. in Wien (math.-naturw. Kl.) 1930, Nr. 19 erschienene Zusammenfassung der Resultate dieser Arbeit.

[2]) A. Whitehead und B. Russell, Principia Mathematica, 2. Aufl., Cambridge 1925. Zu den Axiomen des Systems PM rechnen wir insbesondere auch: Das Unendlichkeitsaxiom (in der Form: es gibt genau abzählbar viele Individuen), das Reduzibilitäts- und das Auswahlaxiom (für alle Typen).

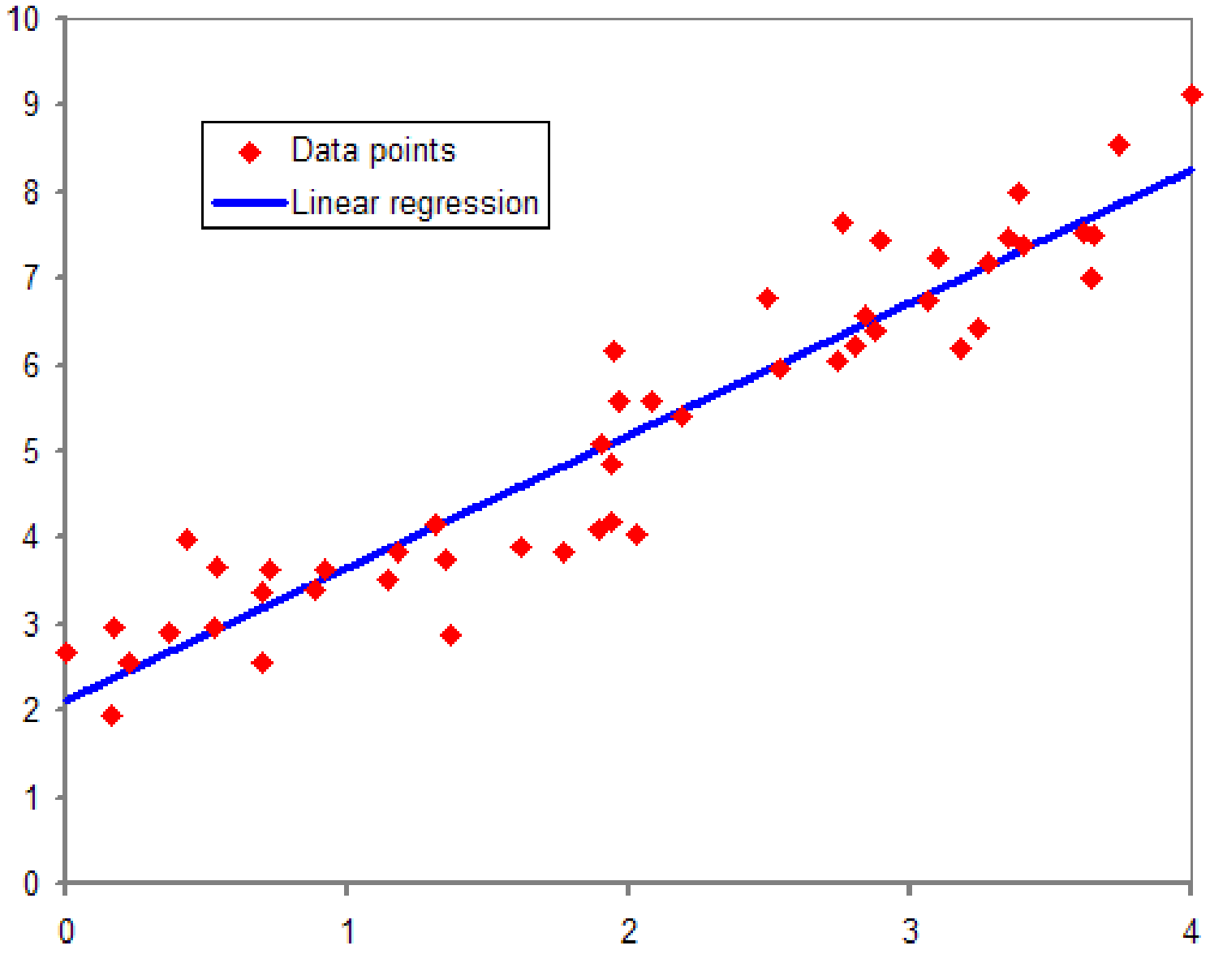
[3]) Vgl. A. Fraenkel, Zehn Vorlesungen über die Grundlegung der Mengenlehre, Wissensch. u. Hyp. Bd. XXXI. J. v. Neumann, Die Axiomatisierung der Mengenlehre. Math. Zeitschr. 27, 1928. Journ. f. reine u. angew. Math. 154 (1925), 160 (1929). Wir bemerken, daß man zu den in der angeführten Literatur gegebenen mengentheoretischen Axiomen noch die Axiome und Schlußregeln des Logikkalküls hinzufügen muß, um die Formalisierung zu vollenden. — Die nachfolgenden Überlegungen gelten auch für die in den letzten Jahren von D. Hilbert und seinen Mitarbeitern aufgestellten formalen Systeme (soweit diese bisher vorliegen). Vgl. D. Hilbert, Math. Ann. 88, Abh. aus d. math. Sem. der Univ. Hamburg I (1922), VI (1928). P. Bernays, Math. Ann. 90. J. v. Neumann, Math. Zeitschr. 26 (1927). W. Ackermann, Math. Ann. 93.

[4]) D. h. genauer, es gibt unentscheidbare Sätze, in denen außer den logischen Konstanten: $\neg$ (nicht), $\vee$ (oder), (x) (für alle), $=$ (identisch mit) keine anderen Begriffe vorkommen als $+$ (Addition), $\cdot$ (Multiplikation), beide bezogen auf natürliche Zahlen, wobei auch die Präfixe (x) sich nur auf natürliche Zahlen beziehen dürfen.

Image source: https://medium.com/kocuniversity/the-bases-of-the-mind-the-relationship-of-language-and-thought-a0bf30375528

Probabilistic approach

- Idea: Randomness in human intelligence motivates randomness in artificial intelligence
- Goal: For a given query, a system returns a (most) likely or at least ***plausible answer***
- Results improve with more or better input or training data
- Statistical learning, Bayesian statistics



Neuro-inspired approach

- Idea: The human “thinking organ” is the brain. Let’s build a machine like it.
- Goal: Construct a system that can ***develop*** intelligence like a human
- Overlap with neuroscience, cognitive sciences
- Supervised learning: The machine learns as we train it

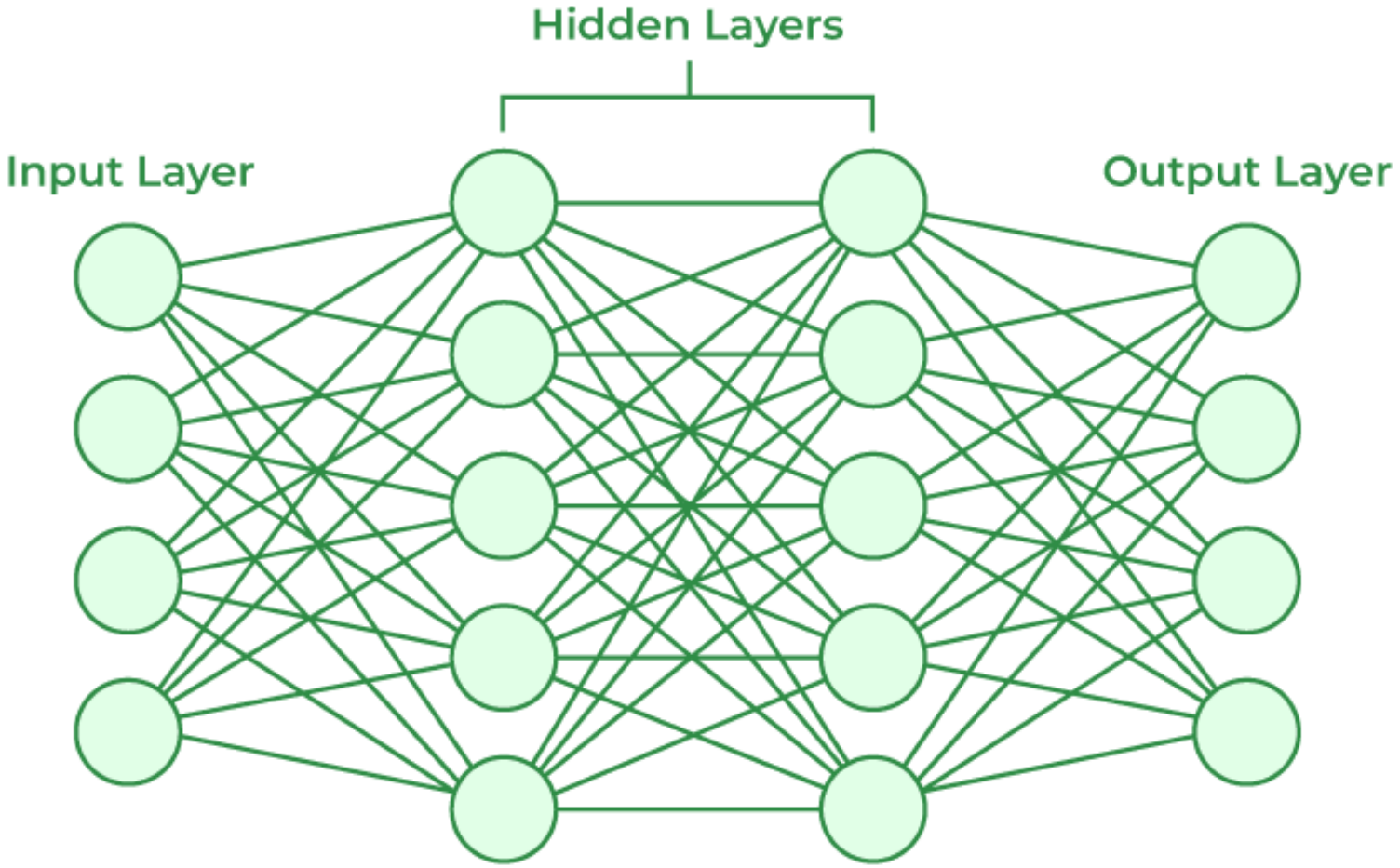


Image source:https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/

Evolutionary approach

- Idea: Human intelligence is the product of evolution. Can artificial evolution lead to artificial intelligence?
- Goal: Create one or many adaptable entities to iterations of mutation and selection to find the ***best algorithm in a given environment***
- Overlap with evolutionary game theory, dynamical systems
- Collective intelligence, swarm intelligence

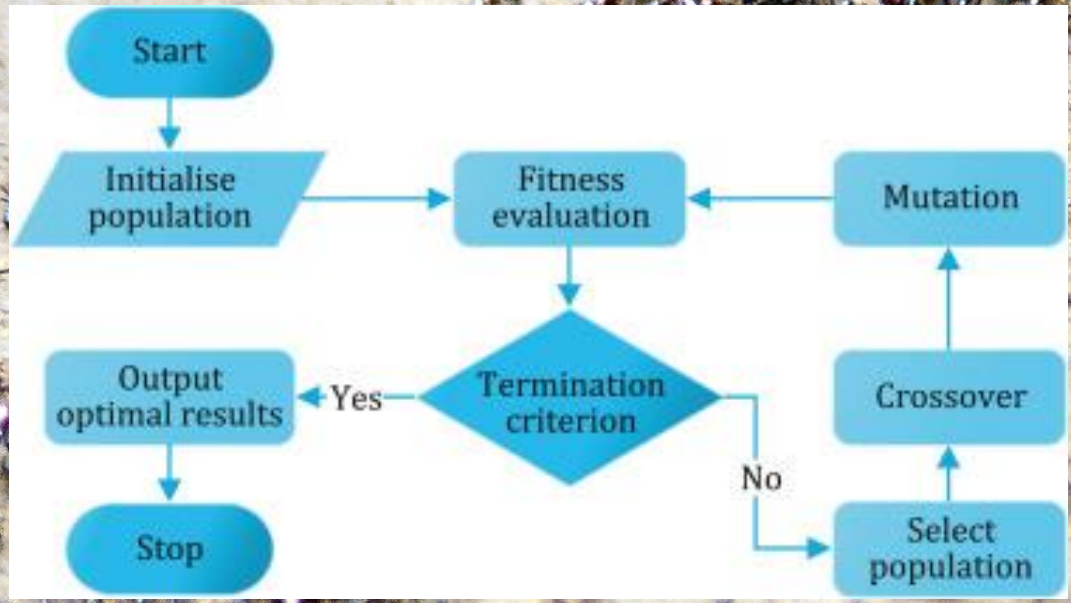


Image source: Pramanik et al. "Big Data classification: techniques and tools" in "Applications of Big Data in Healthcare", Academic Press, 2021

Image source:https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/

The Dartmouth Workshop

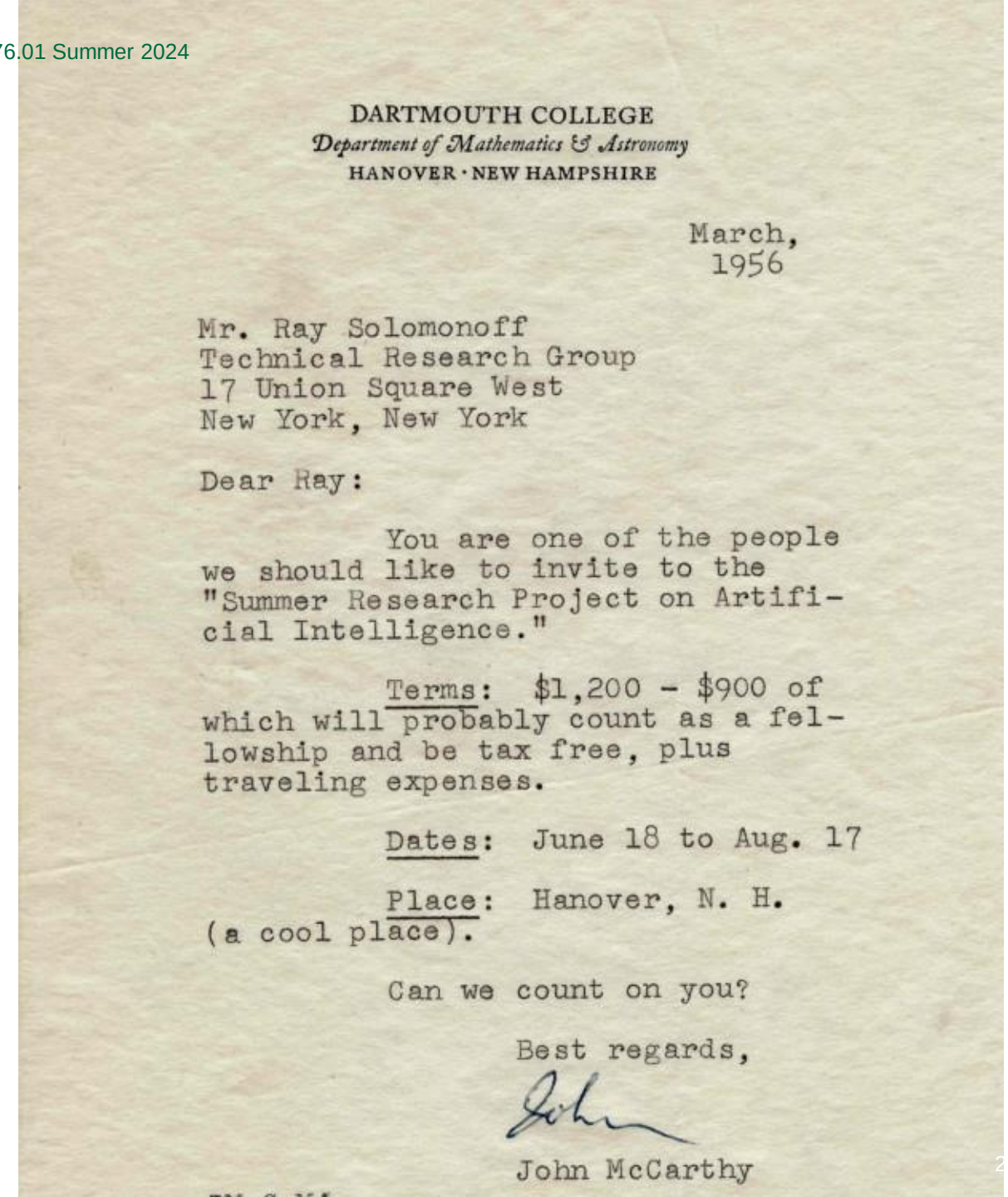
John McCarthy

- 1948 B.S. math (Caltech)
- 1951 Ph.D. in math (Princeton)
- 1955 Assistant Professor of Mathematics (Dartmouth College)
- Subsequent professorships at MIT and Stanford
- “founding fathers” of artificial intelligence

Photo by Margaret Minsky

The Birthplace of AI

- Workshop recognized AI as an interdisciplinary field with connections to:
  - mathematics,
  - computational theory,
  - information theory,
  - neuroscience,
  - psychology,
  - and philosophy



DARTMOUTH COLLEGE
Department of Mathematics & Astronomy
HANOVER · NEW HAMPSHIRE

March,
1956

Mr. Ray Solomonoff
Technical Research Group
17 Union Square West
New York, New York

Dear Ray:

You are one of the people we should like to invite to the "Summer Research Project on Artificial Intelligence."

Terms: $1,200 - $900 of which will probably count as a fellowship and be tax free, plus traveling expenses.

Dates: June 18 to Aug. 17

Place: Hanover, N. H. (a cool place).

Can we count on you?

Best regards,

John

John McCarthy

Image source: https://raysolomonoff.com/dartmouth/

What and how?

“A man provided with paper, pencil, and rubber, and subject to strict discipline, is in effect a universal machine.”

Alan Turing 1948

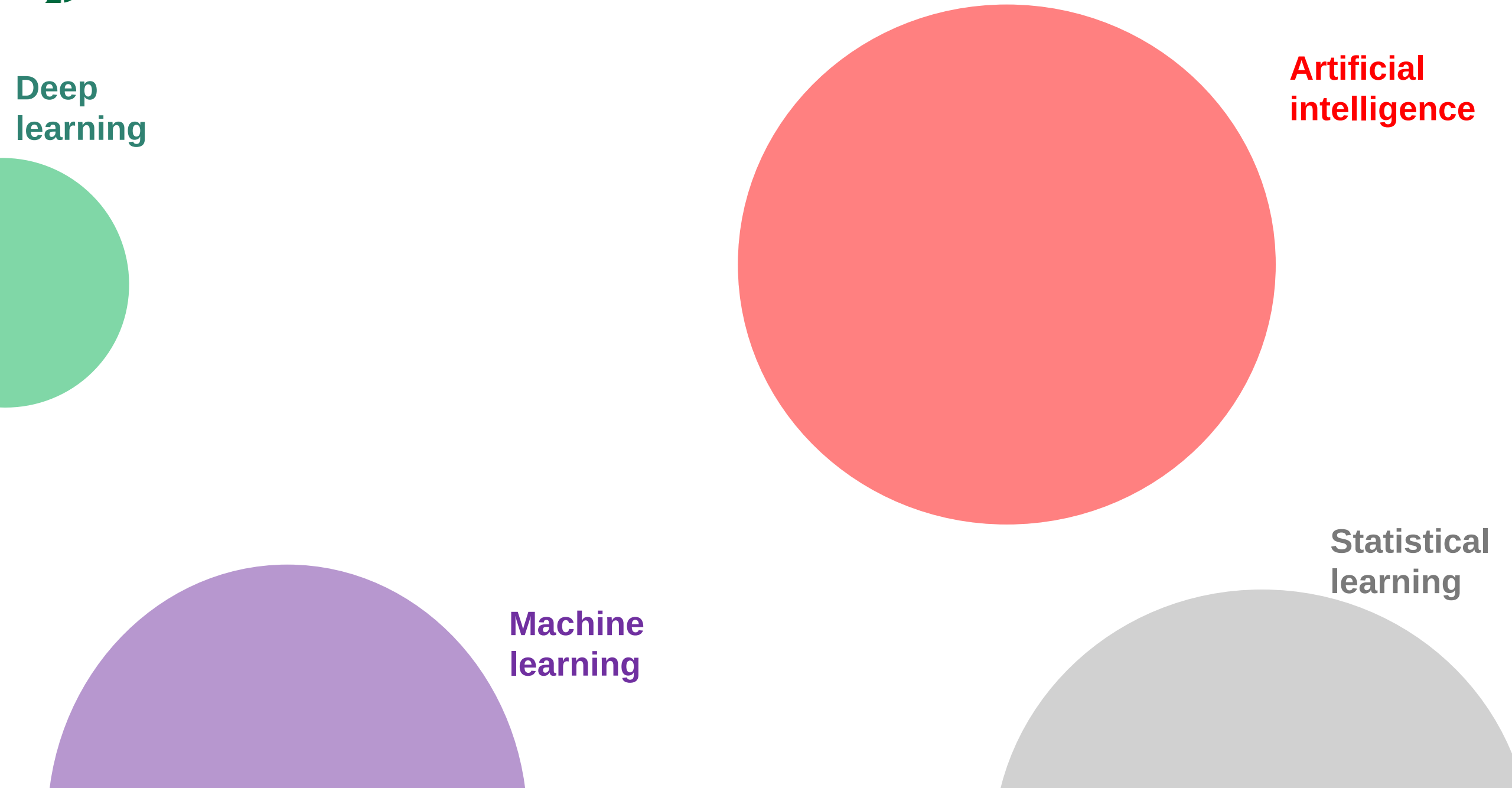
What and how?

What

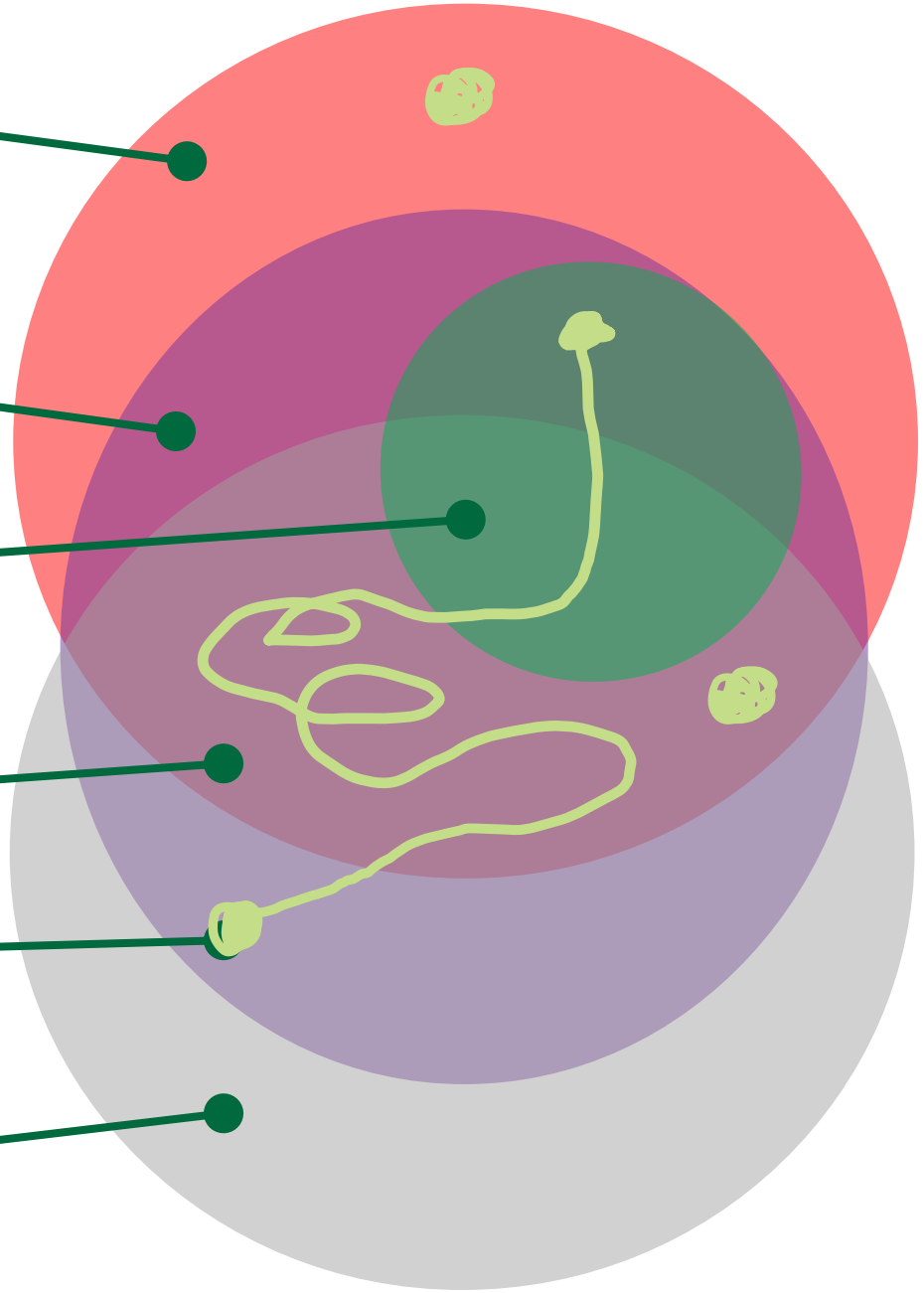
- Artificial general intelligence
- Strong AI
- Expert systems
- Weak AI/narrow AI

How

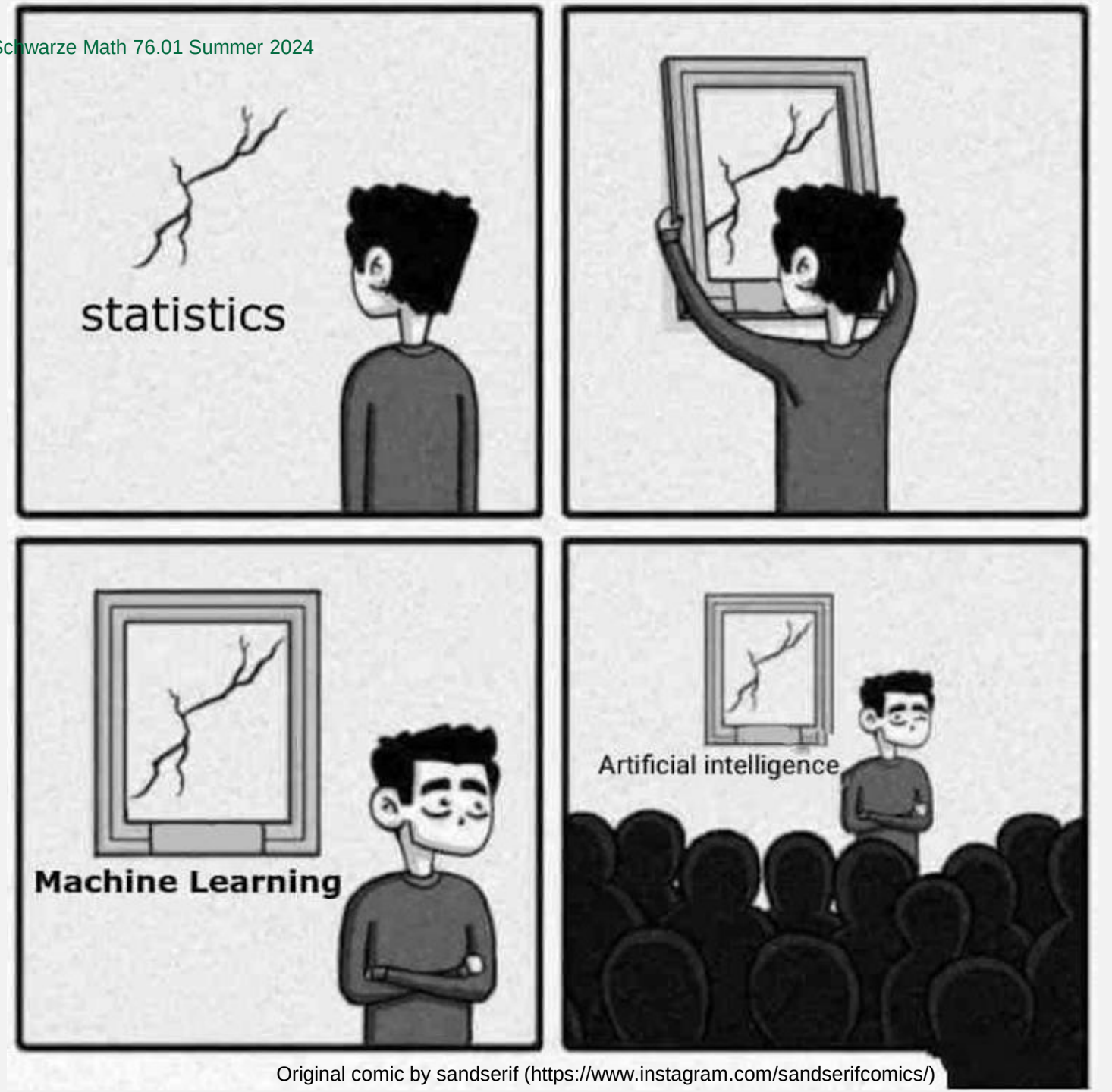
- Logical programming
- Statistical learning
- Machine learning
- Deep learning
- Reinforcement learning


Deep learning
Artificial intelligence
Machine learning
Statistical learning

- AI that does not “learn”
  (e.g. expert systems with hard-coded knowledge and/or logic)
- “bespoke” algorithms* that “learn” from data
  (e.g. Learning via neural networks)
- Algorithms* that use deep neural networks
- Algorithms* with an accompanying statistical theory of why they work
- “simple” algorithms* that “learn” from data
  (e.g. Linear regression)
- Humans using statistical methods



Is everything “just” statistics?



Original comic by sandserif (https://www.instagram.com/sandserifcomics/)

Course information

This is a collaborative learning endeavor.

There is no curve.

Course components

- Class attendance
- Quizzes

} Participation

- Homework
- Course project

} Your repository