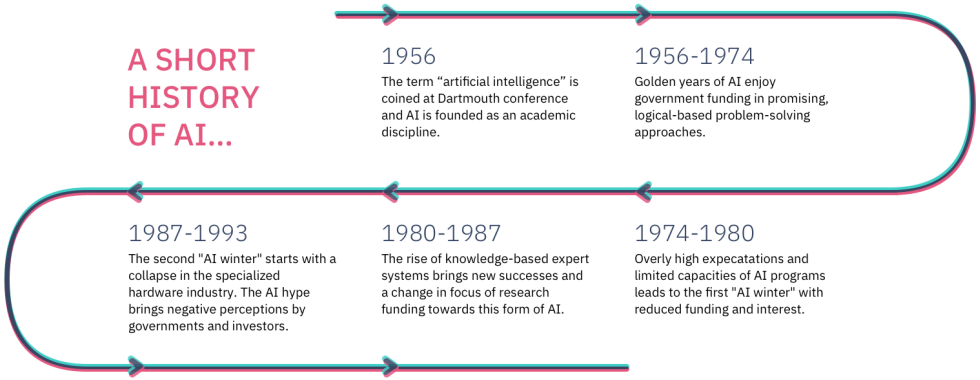
Artificial Intelligence

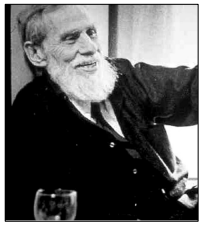
History of AI

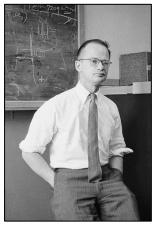
Ch1.3



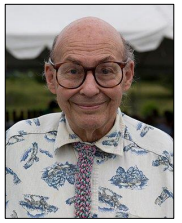
https://medium.com/ai-design-thinkers/patenting-ai-lets-start-with-a-history-lesson-af2cbc73a024

- The gestation of artificial intelligence (1943–1955):
  - The first work that is now generally recognized as AI was done by **Warren McCulloch and Walter Pitts** (1943). They proposed a **model of artificial neurons**.
  - Donald Hebb (1949) demonstrated a simple updating rule for modifying the connection strengths between neurons. His rule, now called **Hebbian learning**, remains an influential model to this day.
  - Two undergraduate students at Harvard, **Marvin Minsky and Dean Edmonds, built the first neural network computer** in 1950.

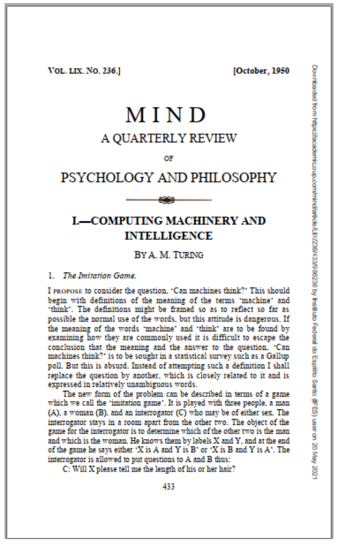

Warren McCulloch


Walter Pitts


Donald Hebb


Marvin Minsky

Minsky received the Turing Award in1969



VOL. LIX. NO. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

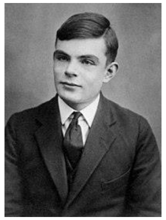
1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

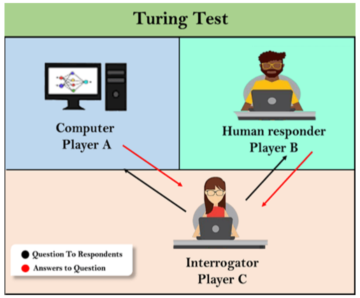
The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B thus:

C: Will X please tell me the length of his or her hair?

433

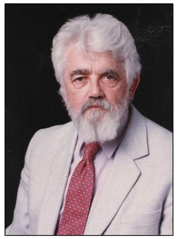


1950

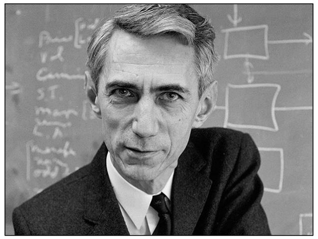


John McCarthy

Creator of the LISP Programming Language. McCarthy received the Turing Award in 1972.



Claude Shannon

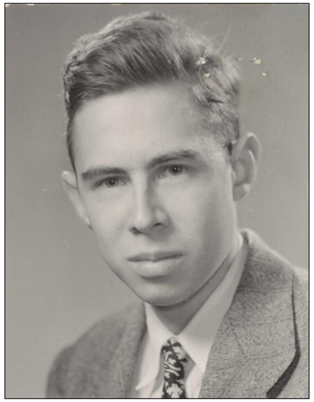


The "father" of information theory.

- The birth of artificial intelligence (1956)
  - John McCarthy convinces Minsky, Claude Shannon, and Nathaniel Rochester to help him bring together U.S. researchers interested in automata theory, neural nets, and the study of intelligence. They organized a two-month workshop at Dartmouth in the summer of 1956. **The term AI was used for the first time in the conference by McCarthy**.
  - The Dartmouth workshop did not lead to any new breakthroughs, but it did introduce all the major figures to each other. For the next 20 years, the field would be dominated by these people and their students and colleagues at MIT, CMU, Stanford, and IBM.

- Early enthusiasm great expectations (1952–1969)
  - The intellectual establishment preferred to believe that "a machine can never do X.". AI researchers naturally responded by demonstrating one X after another.
  - Hebb's learning methods were enhanced by Bernie Widrow (Widrow and Hoff, 1960; Widrow, 1962), who called his networks Adalines, and by **Frank Rosenblatt** (1962) with his **Perceptrons**.
  - **The perceptron convergence theorem** (Block et al., 1962) says that the learning algorithm can adjust the connection strengths of a perceptron to match any input data, provided such a match exists.

Frank Rosenblatt



A dose of reality (1966–1973)

- The fact that a program can find a solution in principle does not mean that the program contains any of the mechanisms needed to find it in practice.
  - **Even knowing that you can learn something, the question of how to learn remains.**
- Minsky and Papert's book Perceptrons (1969) proved that, although Perceptrons (a simple form of neural network) could be shown to learn anything they were capable of representing, they could represent very little.

Ironically, the new **back-propagation learning** algorithms for **multilayer perceptrons** that were to cause an enormous resurgence in neural-net research in the late 1980s were actually discovered first in 1969 (Bryson and Ho, 1969).

Perceptron Limitations

A single layer perceptron can only learn linearly separable problems.

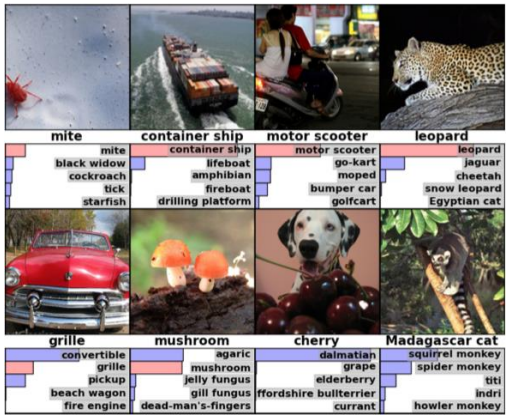
Boolean AND function is linearly separable, whereas Boolean XOR function (and the parity problem in general) is not.

- The return of neural networks (1986–present)
  - In the mid-1980s at least four different groups reinvented the back-propagation learning algorithm first found in 1969 by Bryson and Ho.
  - Since the late 2000s, the availability of data (from the web) and computation (particularly with the development of graphic processing units) allowed neural networks (rebranded as **deep learning**) to achieve unprecedent performance in a wide range of tasks becoming the most used machine learning model.

Received the Turing award in 2018 for conceptual and engineering breakthroughs that have made deep neural networks a critical component of computing.

Achievements of Deep Learning

AlexNet (a type of convolutional neural network) won the 2012 ImageNet competition with a top-5 error rate of 15.3%, compared to the second place top-5 error rate of 26.2%.



Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

Achievements of Deep Learning

- In October 2015, **AlphaGo**, an artificial bot for playing the ancient game GO, won the European Champion, Mr Fan Hui, with a score of 5-0.
- AlphaGo then competed against legendary Go player Mr Lee Sedol, the winner of 18 world titles, who is widely considered the greatest player of the past decade. AlphaGo's 4-1 victory in Seoul, South Korea, on March 2016 was watched by over 200 million people worldwide. This landmark achievement was a decade ahead of its time.
- AlphaGO is based on deep neural networks trained with data from previous human games and reinforcement learning.

https://www.theguardian.com/technology/2016/mar/15/alphago-what-does-google-advanced-software-go-next

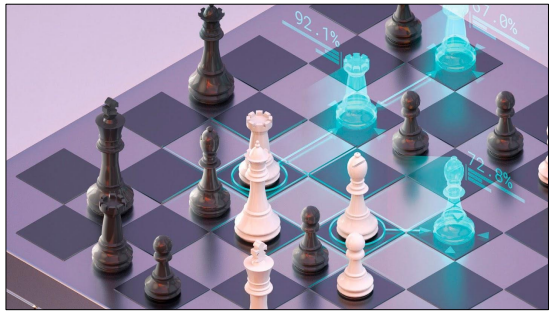
https://deepmind.com/research/case-studies/alphago-the-story-so-far

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S. (2016). Mastering the game of Go with deep neural networks and tree search. *nature*, *529*(7587), 484.

Achievements of Deep Learning

In late 2017, Google DeepMind introduced **AlphaZero**, a single system that **taught itself** (self-play; without human knowledge) from scratch how to master the games of chess, shogi, and Go, beating a world-champion program in each case.

1. Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., ... & Chen, Y. (2017). Mastering the game of go without human knowledge. *Nature*, *550*(7676), 354-359.
2. Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., ... & Lillicrap, T. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, *362*(6419), 1140-1144.
3. Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., ... & Lillicrap, T. (2019). Mastering atari, go, chess and shogi by planning with a learned model. *arXiv preprint arXiv:1911.08265*.



https://deepmind.com/research/case-studies/alphago-the-story-so-far

Achievements of Deep Learning

In 2019, OpenAI Five, wins back-to-back games versus Dota 2 world champions OG at Finals, becoming the first AI to beat the world champions in an esports game. OpenAI Five learned by playing over 10.000 years of games against itself.

https://openai.com/projects/five/

Achievements of Deep Learning

In 2020, Google DeepMind developed Agent57, the first deep reinforcement learning agent to obtain a score that is above the human baseline on all 57 Atari 2600 games. Agent57 combines an algorithm for efficient exploration with a meta-controller that adapts the exploration and long vs. short-term behaviour of the agent.

https://deepmind.com/blog/article/Agent57-Outperforming-the-human-Atari-benchmark

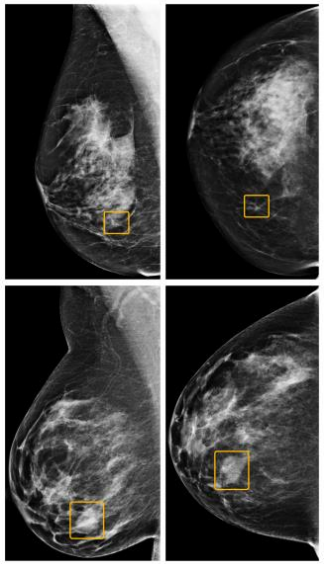
Achievements of Deep Learning

https://www.nvidia.com/pt-br/self-driving-cars/drive-platform/

In 2016, NVIDIA presented a neural network that was able of driving a car in diverse types of terrain and weather conditions.

Video: https://www.youtube.com/watch?v=qhUvQiKec2U

Paper: Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., ... & Zhang, X. (2016). End to end learning for self-driving cars. *arXiv preprint arXiv:1604.07316*.



Achievements of Deep Learning

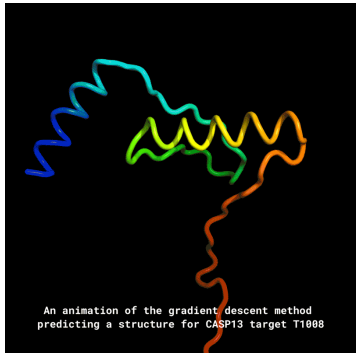
In 2020, Google presented an AI based on convolutional neural networks outperformed all of the human readers in the analysis of mammograms.

https://www.nature.com/articles/s41586-019-1799-6#Abs1

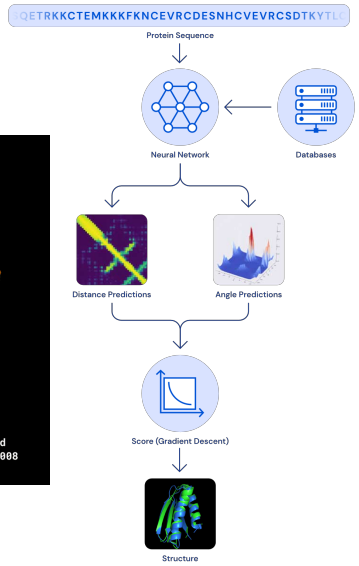
January 2020

Deep Mind's AlphaFold

"Bringing together experts from the fields of structural biology, physics, and machine learning to apply cutting-edge techniques to predict the 3D structure of a protein based solely on its genetic sequence. "

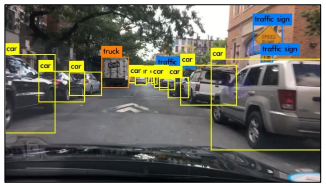


An animation of the gradient descent method predicting a structure for CASP13 target T1008



Achievements of Deep Learning

Deep learning have been used for in several robot perception tasks

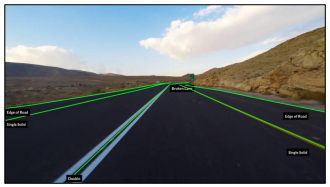


Object Detection

Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.

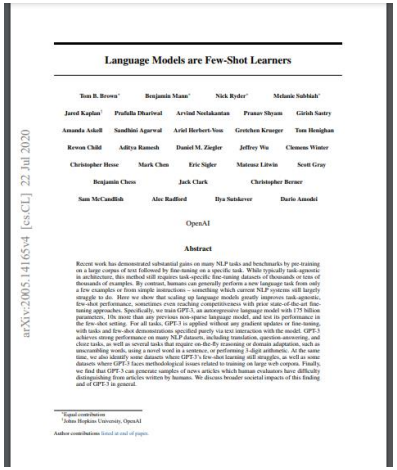
Semantic Segmentation

Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 801-818).



Lane Detection and Classification

Berriel, R. F., de Aguiar, E., De Souza, A. F., & Oliveira-Santos, T. (2017). Ego-lane analysis system (elas): Dataset and algorithms. *Image and Vision Computing*, *68*, 64-75.



Language Models are Few-Shot Learners

Tom B. Brown* Benjamin Mann* Nick Ryder* Melanie Subbiah*
Jared Kaplan† Prafulla Dhariwal Arvind Neelakantan Pranav Shyam Girish Sastry
Amanda Askell Sandhini Agarwal Ariel Herbert-Voss Gretchen Krueger Tom Henighan
Rewon Child Aditya Ramesh Daniel M. Ziegler Jeffrey Wu Clemens Winter
Christopher Hesse Mark Chen Eric Sigler Mateusz Litwin Scott Gray
Benjamin Chess Jack Clark Christopher Berner
Sam McCandlish Alec Radford Ilya Sutskever Dario Amodei
OpenAI

Abstract

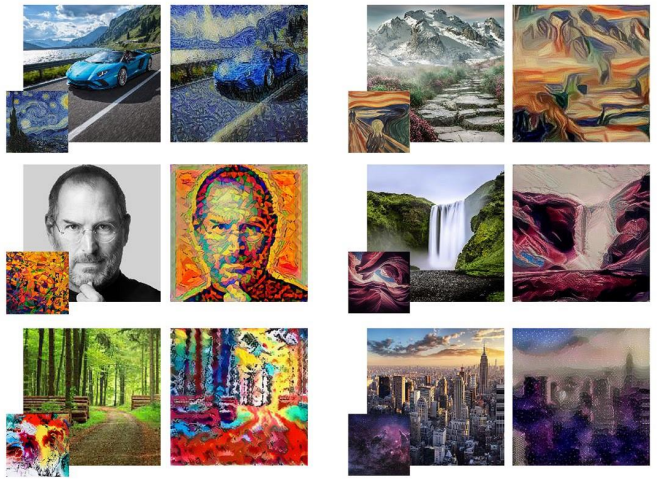
Recent work has demonstrated substantial gains on many NLP tasks and benchmarks by pre-training on a large corpus of text followed by fine-tuning on a specific task. While typically task-agnostic in architecture, this method still requires task-specific fine-tuning datasets of thousands or tens of thousands of examples. By contrast, humans can generally perform a new language task from only a few examples or from simple instructions – something which current NLP systems still largely struggle to do. Here we show that scaling up language models greatly improves task-agnostic, few-shot performance, sometimes even reaching competitiveness with prior state-of-the-art fine-tuning approaches. Specifically, we train GPT-3, an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model, and test its performance in the few-shot setting. For all tasks, GPT-3 is applied without any gradient updates or fine-tuning, with tasks and few-shot demonstrations specified purely via text interaction with the model. GPT-3 achieves strong performance on many NLP datasets, including translation, question-answering, and cloze tasks, as well as several tasks that require on-the-fly reasoning or domain adaptation, such as unscrambling words, using a novel word in a sentence, or performing 3-digit arithmetic. At the same time, we also identify some datasets where GPT-3's few-shot learning still struggles, as well as some datasets where GPT-3 faces methodological issues related to training on large web corpora. Finally, we find that GPT-3 can generate samples of news articles which human evaluators have difficulty distinguishing from articles written by humans. We discuss broader societal impacts of this finding and of GPT-3 in general.

arXiv:2005.14165v4 [cs.CL] 22 Jul 2020

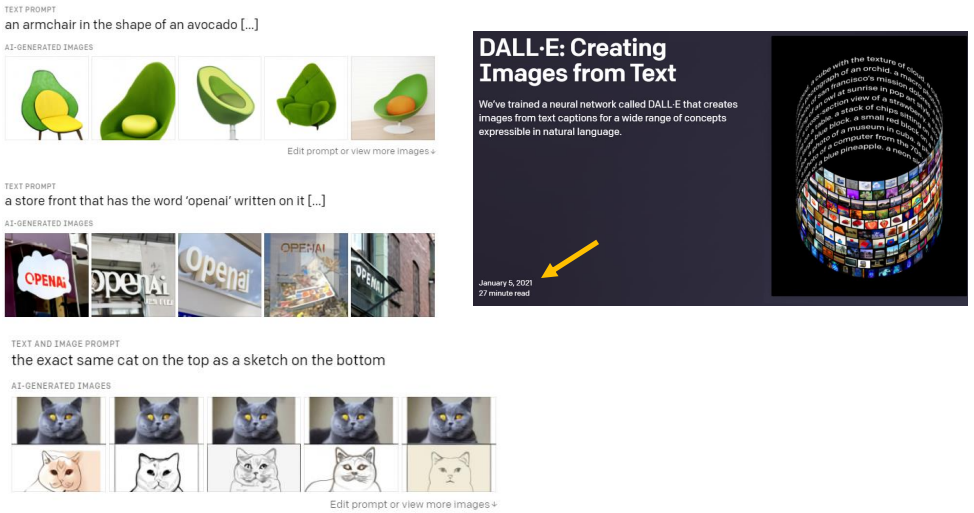
2020 – GPT-3

Title: United Methodists Agree to Historic Split
Subtitle: Those who oppose gay marriage will form their own denomination
Article: After two days of intense debate, the United Methodist Church has agreed to a historic split - one that is expected to end in the creation of a new denomination, one that will be "theologically and socially conservative," according to The Washington Post. The majority of delegates attending the church's annual General Conference in May voted to strengthen a ban on the ordination of LGBTQ clergy and to write new rules that will "discipline" clergy who officiate at same-sex weddings. But those who opposed these measures have a new plan: They say they will form a separate denomination by 2020, calling their church the Christian Methodist denomination.
The Post notes that the denomination, which claims 12.5 million members, was in the early 20th century the "largest Protestant denomination in the U.S.," but that it has been shrinking in recent decades. The new split will be the second in the church's history. The first occurred in 1968, when roughly 10 percent of the denomination left to form the Evangelical United Brethren Church. The Post notes that the proposed split "comes at a critical time for the church, which has been losing members for years," which has been "pushed toward the brink of a schism over the role of LGBTQ people in the church." Gay marriage is not the only issue that has divided the church. In 2016, the denomination was split over ordination of transgender clergy, with the North Pacific regional conference voting to ban them from serving as clergy, and the South Pacific regional conference voting to allow them.

Figure 3.14: The GPT-3 generated news article that humans had the greatest difficulty distinguishing from a human written article (accuracy: 12%).



Neural Style Transfer



Further Reading

- The History of Artificial Intelligence:
- https://courses.cs.washington.edu/courses/csep590/06au/projects/history-ai.pdf
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, *61*, 85-117.
- A Very Short History Of Artificial Intelligence (AI)
- https://www.forbes.com/sites/gilpress/2016/12/30/a-very-short-history-of-artificial-intelligence-ai/#32e669cb6fba
- Timeline of Machine Learning:
- https://en.wikipedia.org/wiki/Timeline_of_machine_learning
- A History of Machine Learning and Deep Learning:
- https://www.import.io/post/history-of-deep-learning/