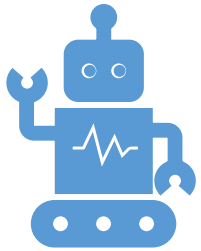
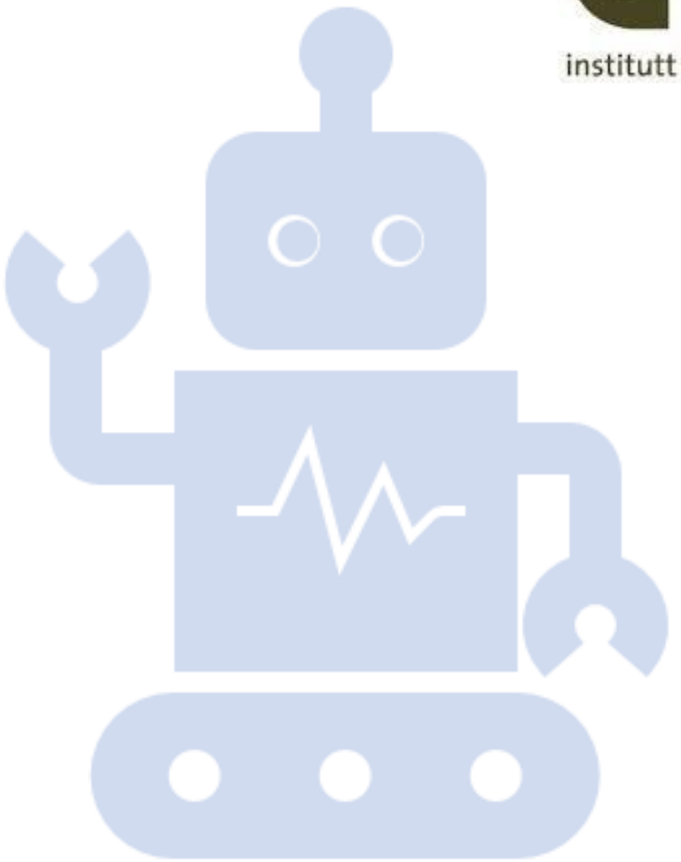
UiO : University of Oslo

institutt for informatikk



IN3050/IN4050 - Introduction to Artificial Intelligence and Machine Learning



Lecture 12

The History and Philosophy of Artificial Intelligence

Slides by Jan Tore Lønning – Presented by Yves Scherrer



WHAT IS A.I.?

Source: Wikipedia

What is AI?

History

- What are AI researchers doing?
- And what have they done?

Philosophy

- What is intelligence?
- Relationship between
  - Artificial intelligence
  - Natural (human) intelligence

Program

1. The birth of AI
   - (1956-1970)
2. The Turing test and a little more philosophy related to AI
3. Approaches to AI
   - (→ 1990)
4. More recent trends
   - (1990 →)

The two first videos were recorded in 2020
Some cross-references may be inaccurate

14.1 The birth of AI

IN3050/IN4050 Introduction to Artificial Intelligence and Machine Learning

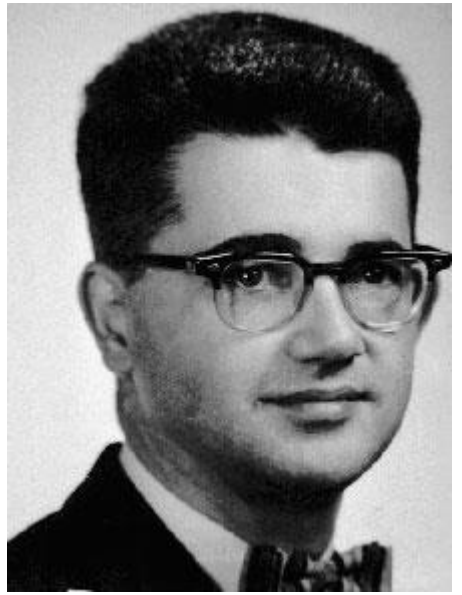
The birth of (the term) Artificial Intelligence

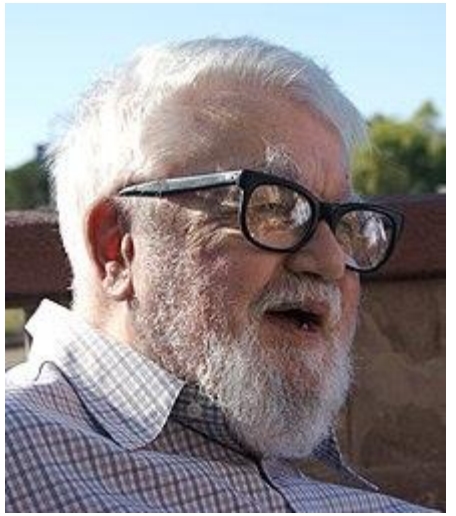
- The Dartmouth Summer Research Projects, Summer of 1956
- Arranged by
  John McCarthy, Marvin Minsky, Nathaniel Rochester, Claude Shannon
- Other participants:
  Herbert Simon, Allen Newell, Arthur Samuel, John Nash

- The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.
- An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.
- We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer. (2-month, 10 man)

John McCarthy, 1927-2011

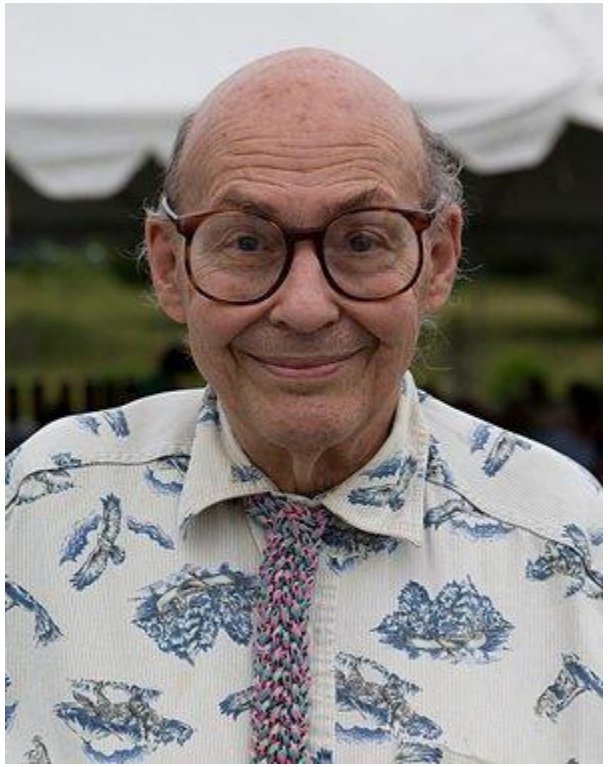
- Invented around 1958
  - LISP (programming language)
  - Garbage collection
  - Time sharing
- MIT 1956-1962
- Stanford 1962 →
  - Established Stanford AI Lab, 1963
- Turing Award, 1971 for his AI work





Marvin Minsky, 1927-2016

- MIT 1958 →
- Founded MIT's AI Lab together with John McCarthy, 1958
- Inventions:
  - Hardware: Head-mounted graphical display, etc.
  - w/ Papert: LOGO programming language
- Perceptrons:
  - PhD thesis on perceptrons, 1954
  - *Perceptrons* (w/ Seymour Papert), 1969
- Logically oriented work
  - e.g. on Turing machines
- *The Society of Minds*
- Turing Award, 1969



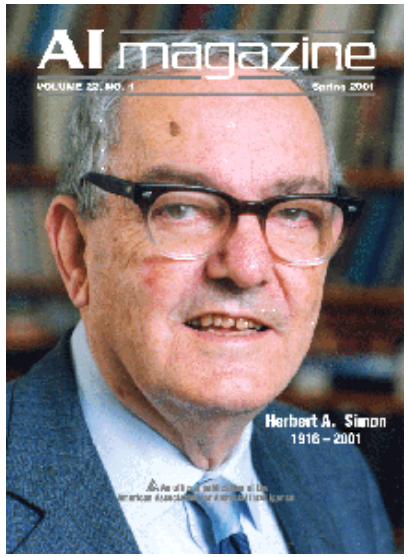
Marvin Minsky (1927-2016)
AI pioneer, MIT AI Lab

source

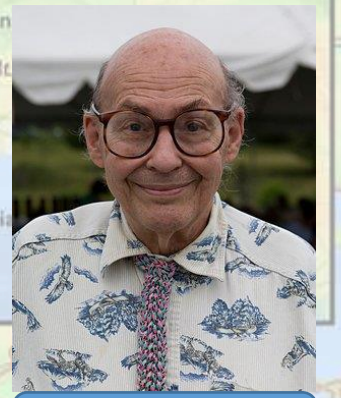
Allen Newell (1927-1992)
Herb Simon (1916-2001)



- *Logic Theorist*, 1956
  - Working program, demonstrated at Dartmouth
  - Proved logical theorems from *Principia Mathematica*
- *General Problem Solver*, 1957
  - A program for solving tasks in general
- Physical Symbols System Hypothesis, 1976 (1963?)
  - Theory about AI
- Turing Award, 1976
- Simon, Nobel prize in economy, 1978



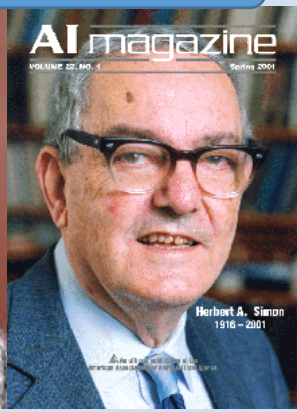
The strongholds of AI in USA



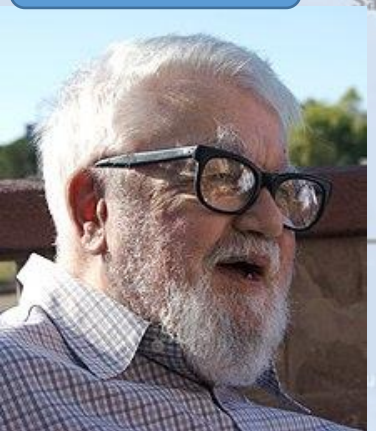
MIT

CMU





Stanford

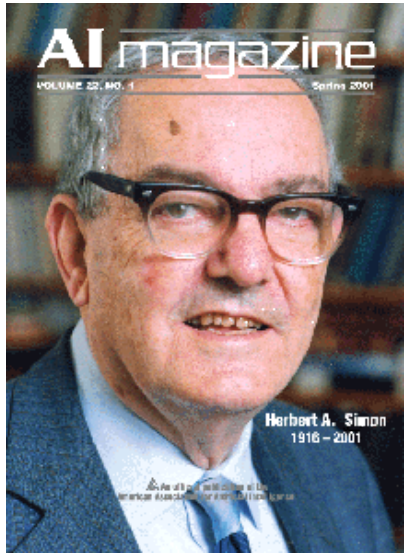


- They remained the strongholds for AI for 50 years
- What are todays' AI strongholds?
  - Facebook, Google, Amazon, Baidu, …

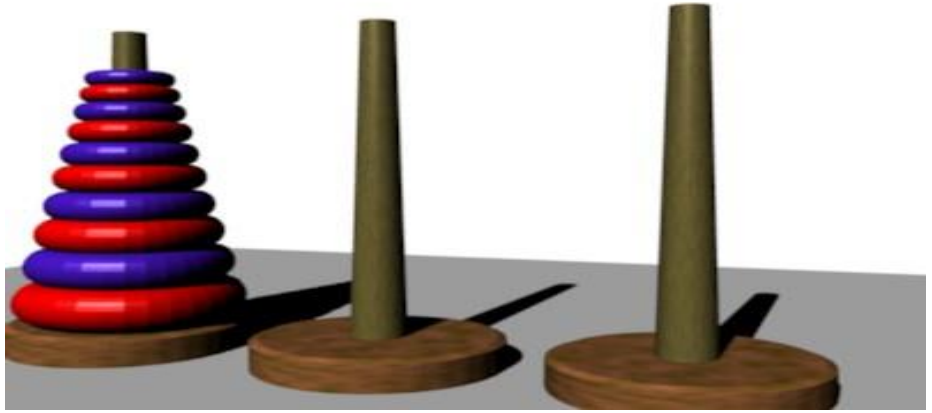
General Problem Solver (1957, GPS)



- Goals-mean analysis:
  1. Current situation
  2. Goal to achieve
  3. A set of available operations
- Task:
  - Put together a sequence of operation leading from current situation to goal.
  - Some conditions must be fulfilled.
  - Earlier operations must establish these conditions.
  - This establishes sub-goals.



GPS example: the towers of Hanoi

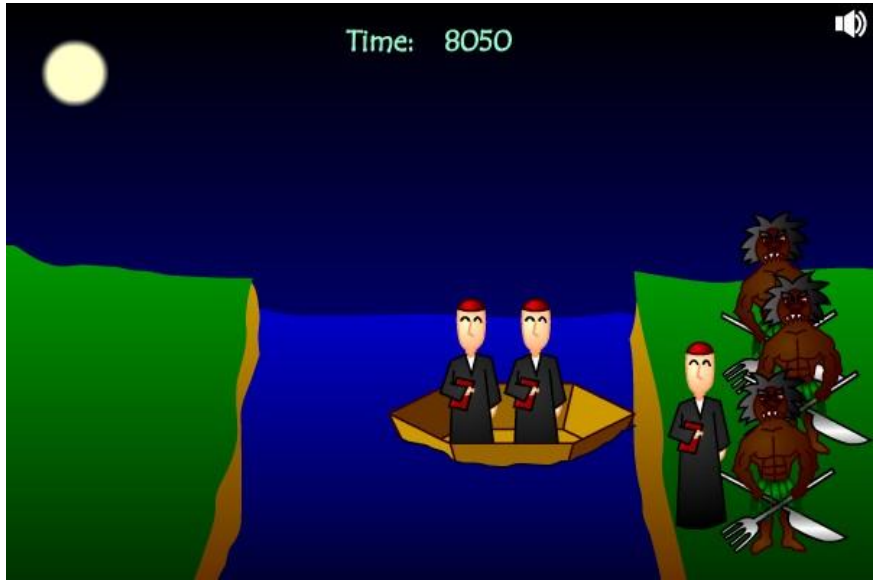


- Goal:
  - Move the stack from A to C
- Rules:
  - Move one ring at a time
  - A ring cannot be placed on top of a smaller one

- Then ring 8 must be at the bottom of C
- Preconditions:
  - C must be free
  - Nothing on top of 8
  - Hence: 1-7 must be on B
- New sub-goal:
  - 1-7 on B
- etc.

GPS - evaluation

- Other tasks solved by GPS:
  - Logical proofs
  - Missionaries and cannibals
- Two types of rules:
  - General rules
  - Task specific rules
- Compared to humans:
  - Modelled on human problem solving
  - Results evaluated against human performance



- Solves some problems
- The general rules can reduce the search space compared to the domain specific rules
- Restricted applicability:
  - Sometimes stuck in local optima
  - Combinatorial explosion, cf. chess
- Project closed

Other voices

Symbolic AI/Rule-based

- Properties of GPS:
  - Symbolic
  - Rule-based
  - Based on/related to logic and proofs
  - Search
- Typical for the approaches of the founding fathers

Other early pioneers

- Arthur Samuel (1901-1990)
  - (attended the Dartmouth workshop)
  - checkers playing program
  - coined the term "machine learning", 1959
- William Grey Walter (1910-1977)
  - turtles

Samuel's checkers playing program





- Based on search
  - (1952) Started with giving rewards to positions based on recorded earlier games:
    - based on Christopher Strachey' s 1951 program
      - the first AI-program according to Jack Copeland
  - (1955) Let the program play against itself and humans and improved the reward function
  - The term Machine learning, 1959
- The program beat a local champion, 1959,
  - but was beaten by stronger programs in the 1970s

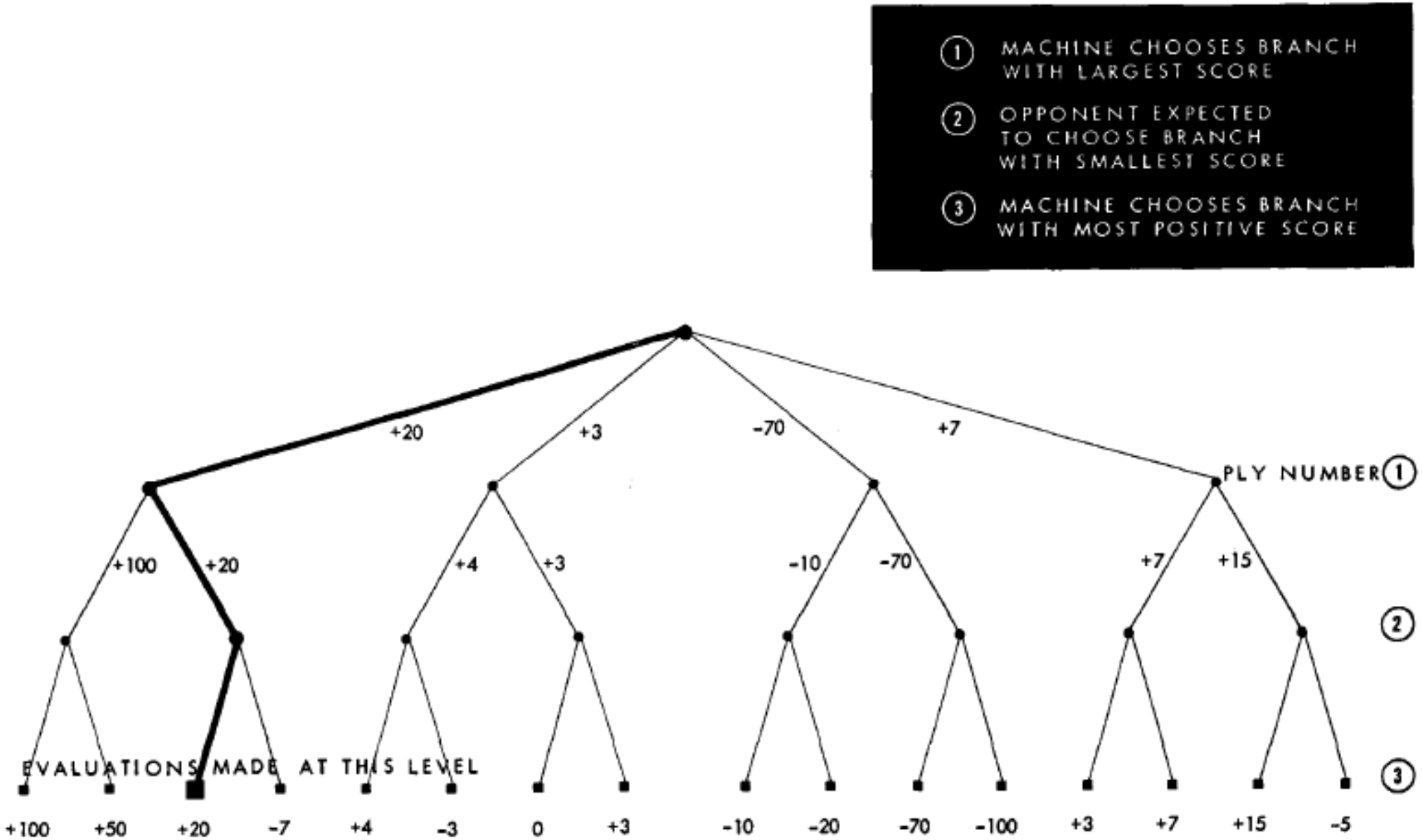


Figure 2 **Simplified diagram showing how the evaluations are backed-up through the "tree" of possible moves to arrive at the best next move. The evaluation process starts at ③.**

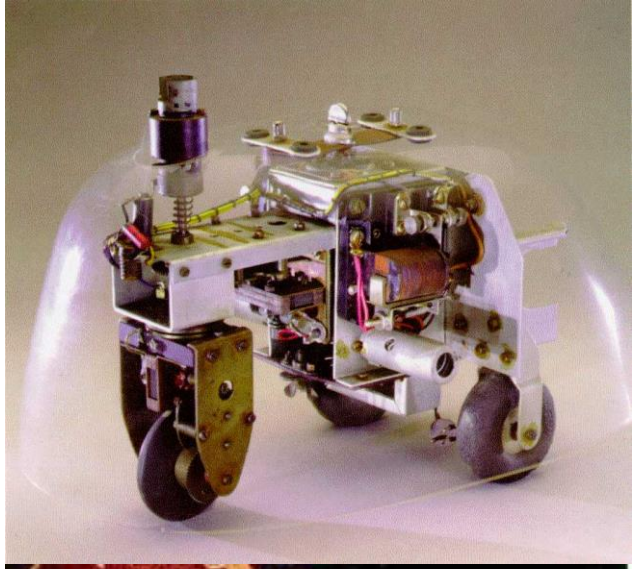
Grey Walter (1910-1977)

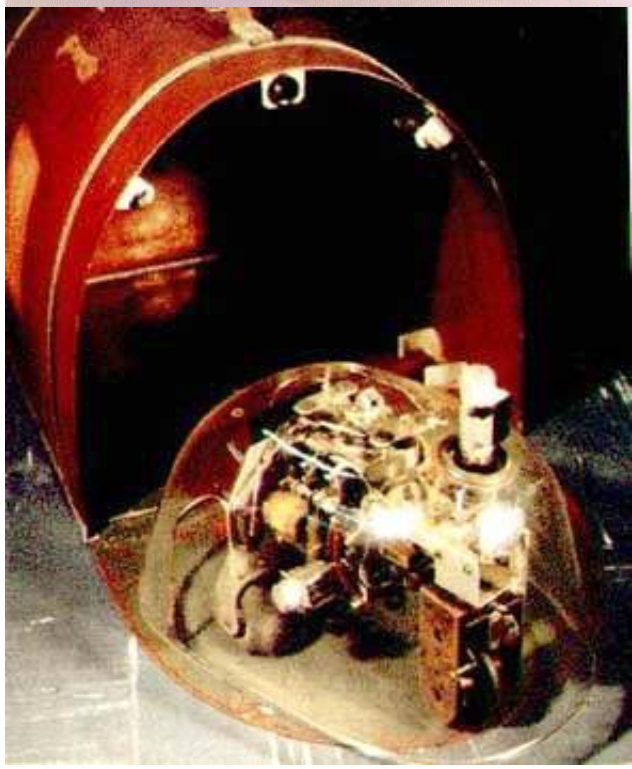
- Worked mainly in GB, also in Soviet Union and USA
- Physiologist:
  - Early use of EEG, several discoveries
- Turtles, 1951:
  - Simple robots
  - Demo



Turtles, 1951

- Three wheels, two engines
- Two sensors:
  - Touch – avoid collision
  - Light – attracted by light, but not too sharp
- When battery got week, more sensible to light
  - A strong light in its home
  - Returned to home
- Goal-oriented behavior?
- Properties: brain-inspired, simple, analogue
- Compare to modern lawn movers and vacuum cleaners





Artificial Intelligence from 1956 →1970 (and beyond)

Methods

- "An anarchy of methods"
  - according to Melanie Mitchell
- Mostly:
  - Symbolic
  - Rule-based
  - Combined with search
  - Logic
- But also, e.g.
  - Perceptron

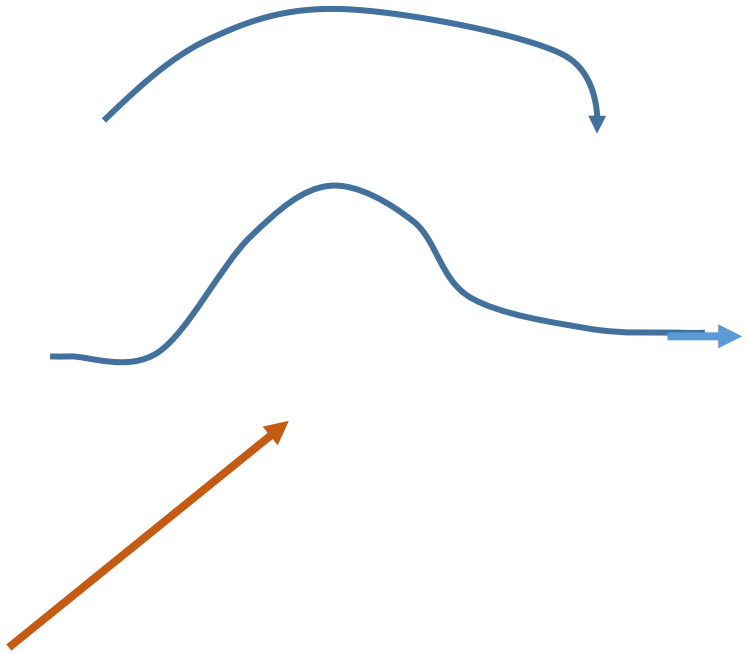
Tasks

- Problem solving
  - Search, Game playing
- Knowledge and Reasoning
  - Logic, Theorem proving, Knowledge representation
- Planning
- Learning
- Natural language understanding
- Perception
- Motion and manipulation

History of neural networks

Three main epochs:

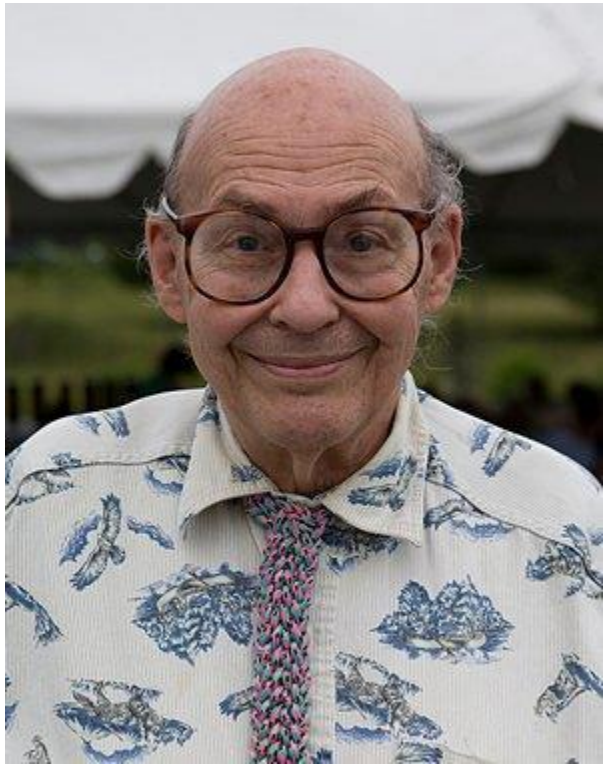
1. The beginning (→ 1969)
2. Backpropagation (1986-)
3. Deep learning (2011→)



- Marsland, originally 2009, lacks (3)

Minsky & Papert, *The perceptron* (1969):

- Showed:
  - Networks without hidden layers can only solve linearly separable problems
  - Many simple problems, like logical XOR, are not linearly separable
- Speculated
  - Networks with hidden layers are probably impossible to train
- Effect:
  - Halted the development of neural networks
- Why such an effect?
  - Minsky's position
  - A growing skepticism towards AI (funding)

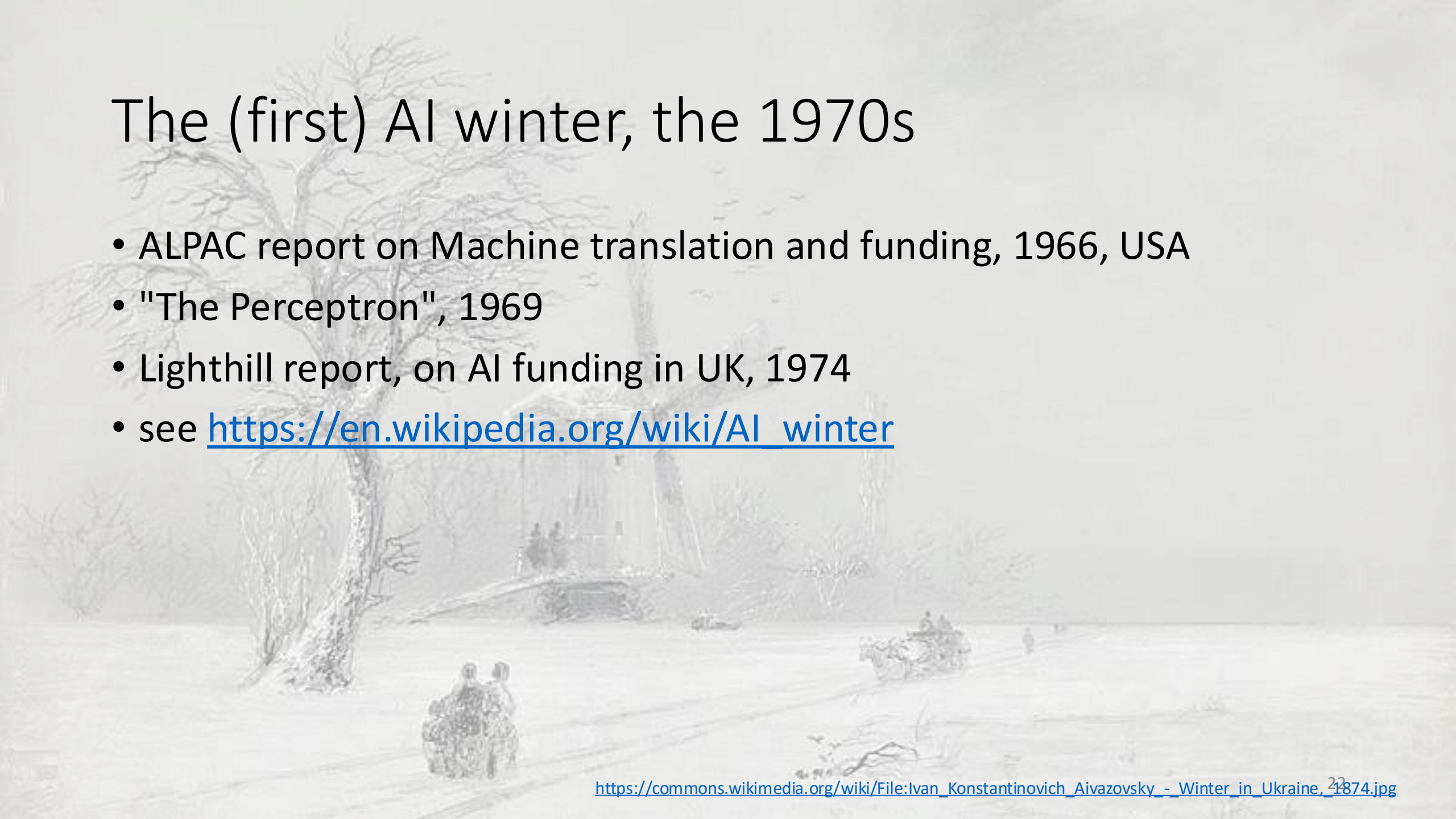


Marvin Minsky (1927-2016)
AI pioneer, MIT AI Lab

source

The (first) AI winter, the 1970s

- ALPAC report on Machine translation and funding, 1966, USA
- "The Perceptron", 1969
- Lighthill report, on AI funding in UK, 1974
- see https://en.wikipedia.org/wiki/AI_winter

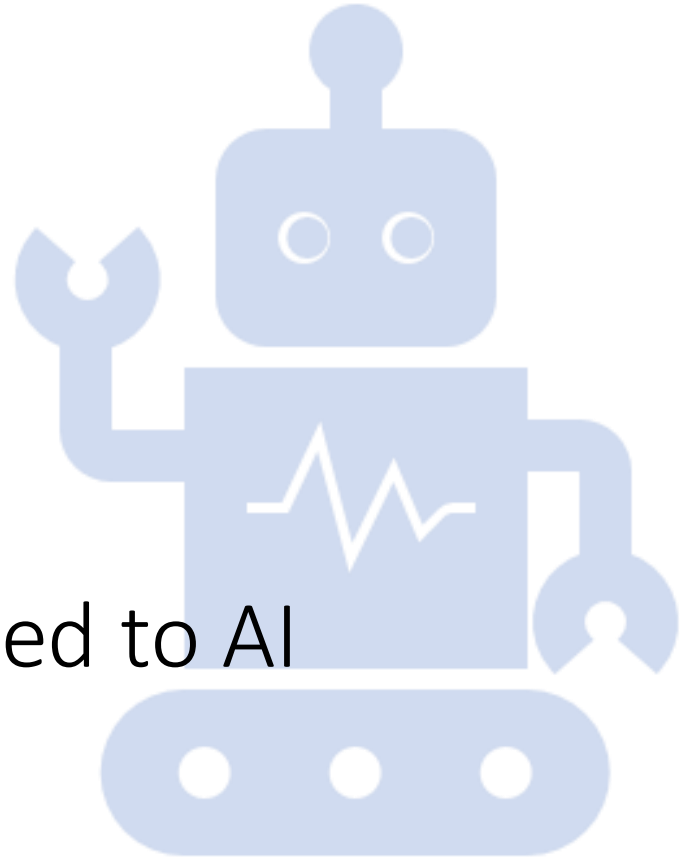


https://commons.wikimedia.org/wiki/File:Ivan_Konstantinovich_Aivazovsky_-_Winter_in_Ukraine,_1874.jpg

Overselling

https://en.wikipedia.org/wiki/History_of_artificial_intelligence

- 1958, H. A. Simon and Allen Newell: "within ten years a digital computer will be the world's chess champion" and "within ten years a digital computer will discover and prove an important new mathematical theorem."[69]
- 1965, H. A. Simon: "machines will be capable, within twenty years, of doing any work a man can do."[70]
- 1967, Marvin Minsky: "Within a generation ... the problem of creating 'artificial intelligence' will substantially be solved."[71]
- 1970, Marvin Minsky (in *Life* Magazine): "In from three to eight years we will have a machine with the general intelligence of an average human being."[72]



14.2 The Turing Test

- and a little more philosophy related to AI

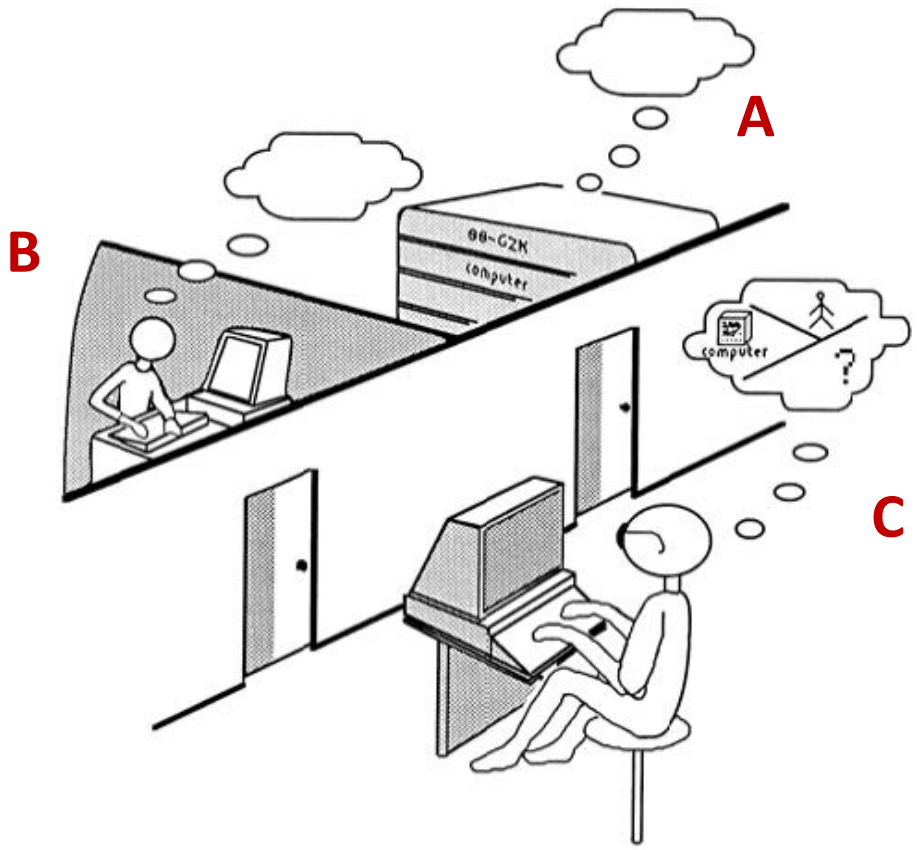
IN3050/IN4050 Introduction to Artificial Intelligence and Machine Learning

Alan Turing (1912-1954)

- 1936: The Turing machine
  - the theoretical foundation of the computer
- 1939-1945: Codebreaking
  - cf. "The Imitation Game"
- 1945→: Developed computers
- 1950: The Turing test
- 1952-1954: Mathematical biology
- The Turing Award named in honor of Turing



The Turing test



A. Try to fool the interrogator to think it is a human
B. Try to help the interrogator to see that he/she is a human
C. The interrogator should guess who is human and who is machine

The Turing test



Turing's presentation

- *Can machines think?*
- *…I shall replace the question by another, which is closely related to it…*
  - Game 1: A:man, B:woman, C:who is what?
  - Game 2: A:machine, B:human, C: who is what?
- *Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman?*

Turing's view

- In 50 years (2000):
  - The machine's memory 10^9 (=1 Gigabyte)
  - Less than 70% chance of correct identification
- But he says this is a guess
  - (other guesses in other articles/interviews)

Evaluating the Turing test

1. Is it adequate?
   - Will we say that a machine that passes the test can think?
2. Can a computer pass the test (in the future)?
3. Is it a goal that a machine passes the test?

- The test has been much discussed
- Turing anticipated 9 objections in his original paper which he tried to rebut

(4.) The argument from Consciousness

Jefferson according to Turing

"Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain - that is, not only write it but know that it had written it. No mechanism could feel […]pleasure at its successes, grief when its valves fuse, […] be angry or depressed when it cannot get what it wants."

Turing's answer

- *According to the most extreme form of this view the only way by which one could be sure that a machine can think is to be the machine…*
- To get convinced e.g., that somebody else compose because of felt emotions, we would interrogate them, as in the imitation game

Is the Turing test adequate?

Too strong

- Animals are intelligent, but they don't pass the test.
- The machine does not only have to be intelligent, it must also mimic a human

Answers:

- The test only shows a sufficient condition not a necessary one

Inadequate

- Simulation is not the real thing, e.g., a man could simulate being a woman
- We ascribe consciousness to other humans because we know they are made in the same ways we are.

Chat bots

- Eliza (Weizenbaum, 1966)
- Planned architecture:
  - An overarching program for dialogue management
  - Domain specific modules for various applications
  - Doctor, the first example of an application
  - A psychotherapist (little domain knowledge necessary)
- Eliza/Doctor, principles:
  - looking for key words in the questions
  - transform input to answers, e.g. *Why do you + ….*
    - exchange *me* with *you*
  - vary the answers



Joseph Weizenbaum (1923-2008)

- Scared by the reactions to ELIZA

1. Some psychiatrists believed DOCTOR (ELIZA) could be used in therapy
2. Users developed a personal relationship to ELIZA
3. Some believed this was a model for successful NLP



https://en.wikipedia.org/wiki/Joseph_Weizenbaum

- Weizenbaum became skeptical towards AI
- *Computer Power and Human Reason*, 1976
  - Great belief in humans, little belief in machines
  - c.f. similar views by Stephen Hawking (1942-2018) a.o. around 2015

Has the Turing test been passed?

- Now and then, there are stories in the news that the Turing test has been passed.
- It is normally a combination of
  - A program that imitates a mad, nasty or uncooperative person
  - Judges that don't ask proper questions
- Observe that according to the rules, the human should be cooperative.

- There have been various attempts of amplifying the rules of the test with respect to interrogators, etc.

Questions to ask



https://en.wikipedia.org/wiki/Terry_Winograd

- Try to ask, *Where is New York times published?*
- Or try the Winograd schemas
  - (Terry Winograd, 1946 →, AI/NLP pioneer turned skeptical)

The city councilmen refused the demonstrators a permit because they feared violence.

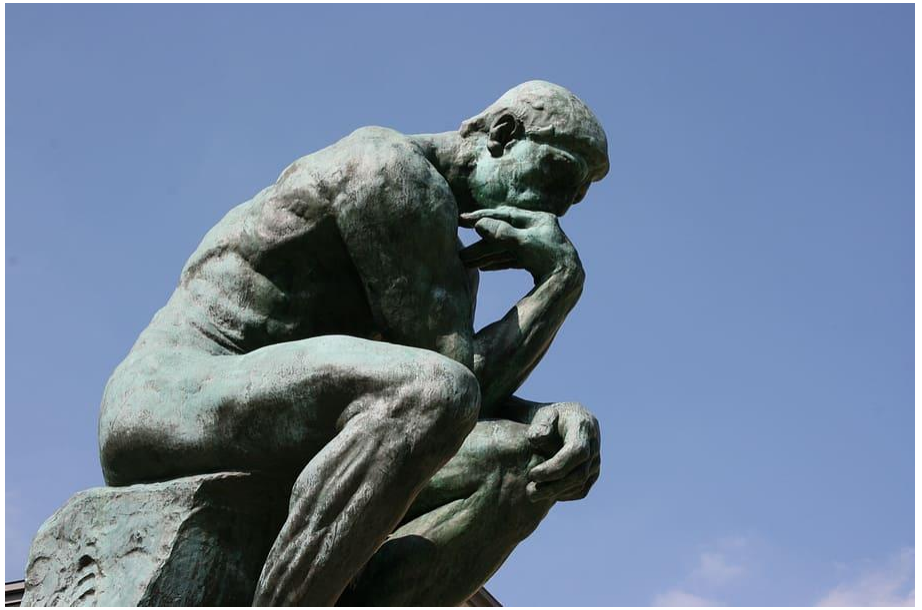
Q: *Who feared violence?*

The city councilmen refused the demonstrators a permit because they advocated violence.

Q: *Who advocated violence?*

Philosophy of AI

A. Can a machine pass the Turing test?
B. Can a machine act intelligently?
C. Can it solve any problem that a person would solve by thinking?
D. Can a machine have a mind and consciousness?
E. Are human intelligence and machine intelligence the same?



c.f. https://en.wikipedia.org/wiki/Philosophy_of_artificial_intelligence

Intelligent machines?

B. Can a machine act intelligently?

- If "A machine is considered intelligent if it can perform tasks which are considered intelligent when carried out by a human being.“ (IN3050, lect. 1)
- then yes

C. Can it solve *any* problem that a person would solve by thinking?

- According to the invitation to the Dartmouth conference, yes.
- We aren't there yet!

E. Human intelligence = Machine intelligence?

- Physical symbol system hypothesis:
- *"A physical symbol system has the necessary and sufficient means of general intelligent action."* (Newell and Simon, 1976)

A physical symbol system e.g.,

- Computer
- Chess
- Writing on a whiteboard

- Implies that machines are intelligent (jfr. Dartmouth)
- Interpreted as only symbol systems are intelligent, i.e., human intelligence is equivalent to executing computer programs.
- The computational theory of mind

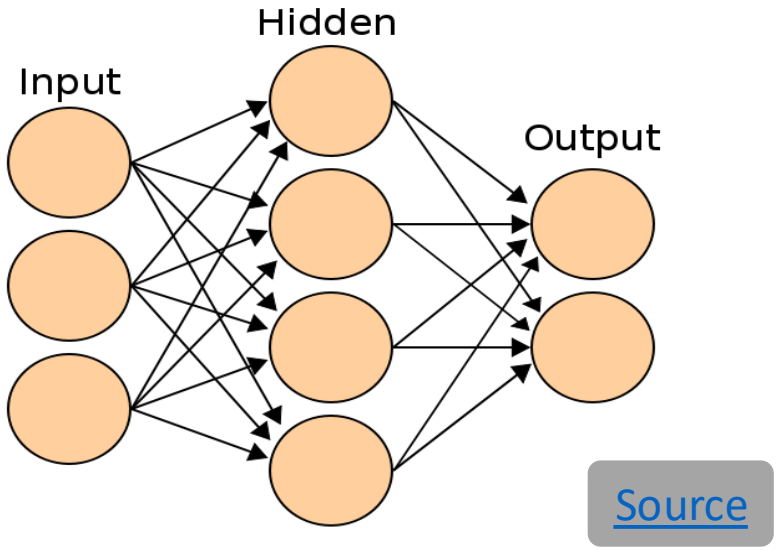
D. Can a machine have consciousness?

- A large philosophical discussion related to this in the 1980-1990s
- To a large degree arguments against the Turing test and the PSSH:
  - cf. The consciousness argument above
  - Chinese room argument (Searle).
- Argued that strong AI means, not only to do the same as a human, but also do it the same way, which includes having a mind and consciousness that machines don't have.



Connectionism

- After the publication of the backpropagation paper in 1986,
- some psychologist and philosophers argued for modelling the human mind in terms of neural networks
- Connectionism
- Sub-symbolic computing



- Argues that this avoids some of the criticism directed towards the computational theory of mind

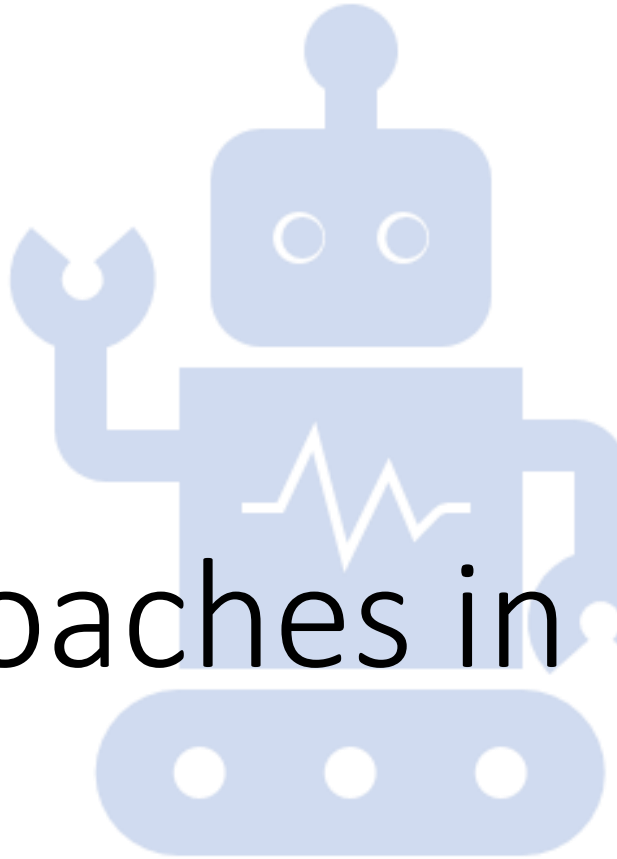
Main approaches to AI

Symbolic, Rule-based

- Logic, deduction
- Explicit coding of knowledge as formulas or rules
- Dominated AI-books until the end of the last century
- Compatible philosophy:
  - Computational theory of mind (PSSH)

Machine learning, Neural nets

- Induction rather than deduction
- Adapt to the environment
- Main-focus in this course
- Dominates AI today
- Compatible philosophy:
  - Connectionism



14.3 Traditional approaches in AI

IN3050/IN4050 Introduction to Artificial Intelligence and Machine Learning

Artificial Intelligence from 1956 →1970 (and beyond)

Methods

- "An anarchy of methods"
  - according to Melanie Mitchell
- Mostly:
  - Symbolic
  - Rule-based
  - Combined with search
  - Logic
- But also, e.g.
  - Perceptron

Tasks

- Problem solving
  - Search, Game playing
- Knowledge and Reasoning
  - Logic, Theorem proving, Knowledge representation
- Planning
- Learning
- Natural language understanding
- Perception
- Motion and manipulation

Logic

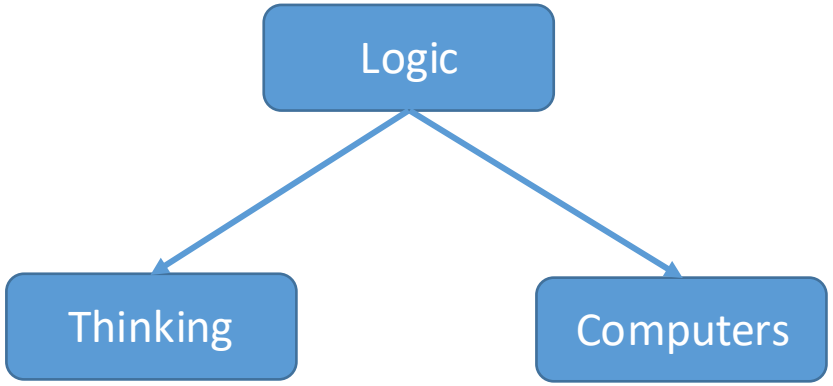
- Aristotle 384- 322 BC
  - *All humans are mortal*
  - *Socrates is a human*
  - *Socrates is mortal*
- Modern symbols:
  - $\forall x(human(x) \rightarrow mortal(x))$
- Rationality, intelligence
- Correct reasoning



image: Wikipedia

Logic and computation

- The computer is based on logic:
  - A computation can be considered a logical proof
  - Turing machines
  - C.f. also McCulloch & Pitts' neural model of logic
- Hence logic seems to be the perfect link between:
  - Human intelligence
    - (beyond numerical calculations)
  - Computers



- E.g., LISP
  - Based on logic
    - (expressed in lambda calculus)
  - Symbolic computing
    - In contrast to numeric computing

Challenges for the logical approach 1

A framework without content

- *Kim bought a rose*
- *Kim bought a flower*

- Not a valid inference
- One needs an additional axiom
  - $\forall x(Rose(x) \rightarrow Flower(x))$

Knowledge representation

- An explosion of facts that must be represented:
  - AI-applications focus on limited domains
- Knowledge representation a subfield of AI
  - Other approaches than logic:
    - Semantic nets
    - Ontologies
- While logic is neat, the represented knowledge can be ad hoc

Challenges for the logical approach 2

Logic provides proofs

- To see that a conclusion follows from premises, logic provides proofs and proof-procedures

Search problem

- To find a proof, however, is a challenge
- An enormous search space

Search

- Initially AI was to a large degree focusing on search
  - (discrete structures)
- Common part of various problems:
  - Logical proofs
  - Newell and Simon's GPS
  - Game playing, like Checkers
  - *Travelling salesman*

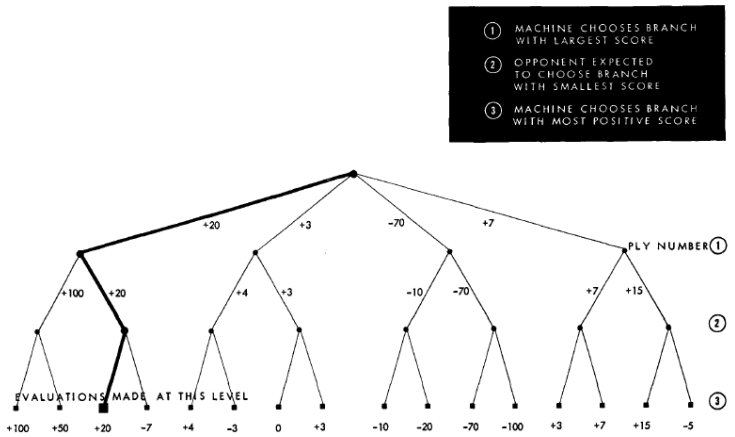
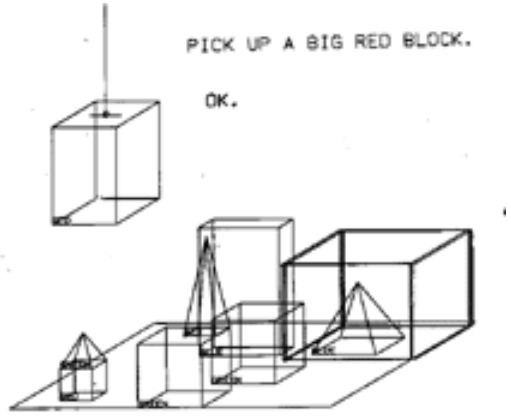


Figure 2 **Simplified diagram showing how the evaluations are backed-up through the "tree" of possible moves to arrive at the best next move. The evaluation process starts at ③.**

- First principles:
  - Ways to search a large search space efficiently
- The search space often still too large:
  - Task specific heuristics:
    - E.g., Chess

General symbolic approaches

- Symbolic approaches could use other representations than logic:
  - cf., GPS, checkers
- One approach bases the system on observations of human behavior
  - (in contrast to logic)
  - E.g., Newell and Simon (eventually)
- Or whatever works
  - Minsky, became anti-logic, ''scruffy''
    - Many projects that were impressive in the small, e.g., ''Blocks world''



- Ad-hoc rules and blocks worlds:
  - Impressive in the small
  - Problems with scalability

The 1980s a new spring

- Expert systems
- Logic and the 5th generation program
- Revival of Neural Nets 1986

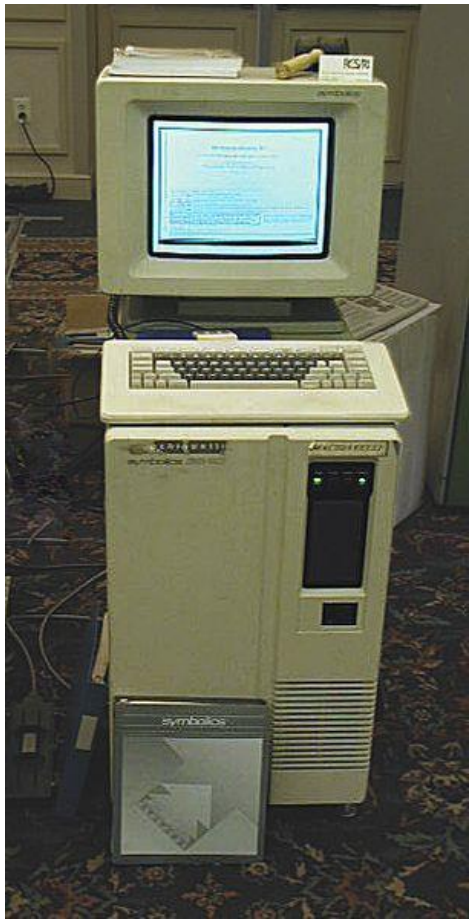
Expert systems

- The system tries to reproduce human expertise
- In its simplest forms:
  - A set of if-then-else clauses:
    - *If red dots, check for fever*
- The system is built by interviewing human experts
- A system made by interviewing medical experts, was reported to perform between a GP and an expert
- Expert system grew into a commercial success and were adopted by many companies



Side remark: Lisp machines, the 1980s

- The decades when PCs and personal workstation became common
- Dedicated machines for AI:
  - LISP machines
  - Symbolics
  - XEROX
  - Texas instruments
  - A.o.
- LISP as operating system



https://en.wikipedia.org/wiki/File:Symbolics3640_Modified.JPG

Logic

- Larger interest for logical (‘’neat’’) approaches
- Partly because of the limitations of the ad hoc approaches
- Partly because of the developments
  - The resolution procedure (Robinson 1963)
  - The programming language PROLOG based on the procedure:
    - Computation as proofs

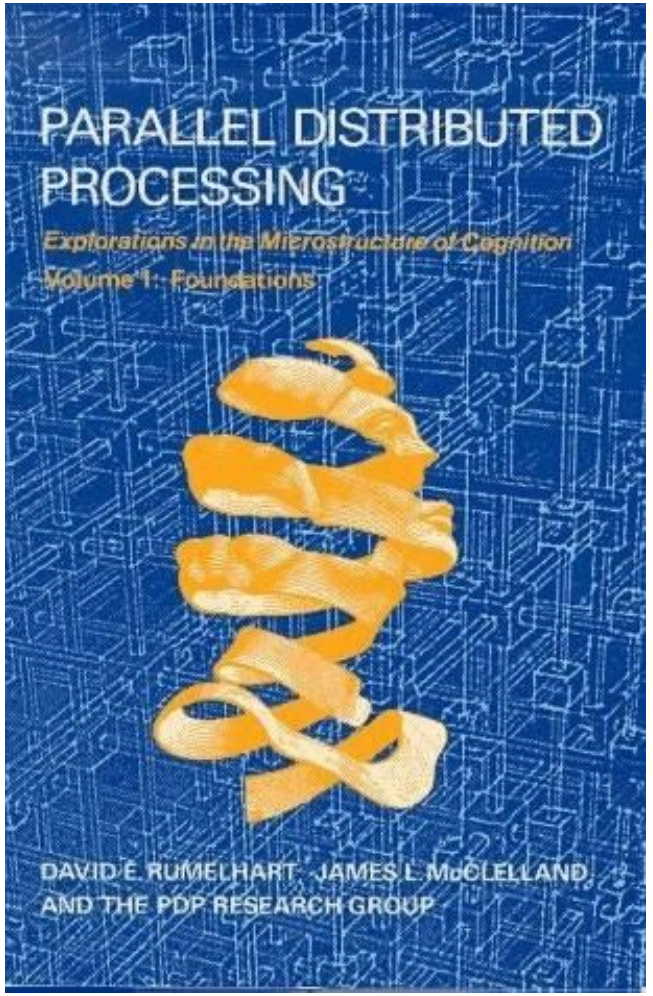
Fifth Generation Program (1982-1992)

- Large Japanese governmental funded research program
- Goals/approaches:
  - Hardware: parallel
  - Software: logic-based
- Lead to AI-funding also in the rest of the world
- Failure?
  - In particular, the AI promises
- Ahead of its time?



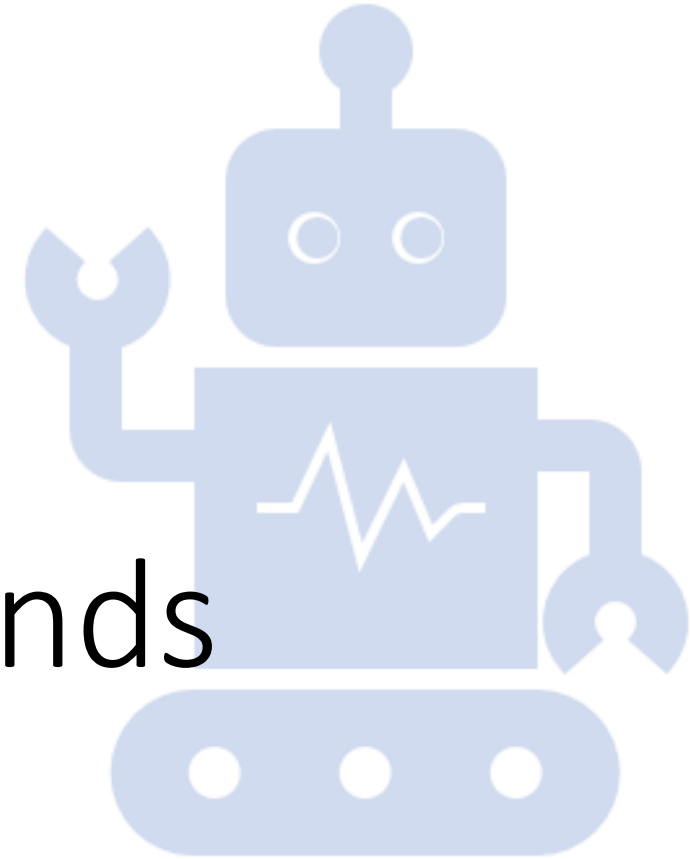
NN.2: Backpropagation (1986-)

- 1986, Rumelhart, Hinton, Williams (re)invented backpropagation
- An immediate enormous interest by researchers
- But the practical results weren't impressing, and the interest diminished



Around 1990

- The commercial interest in AI diminished
  - Expert systems were partly considered mainstream application of computers
  - The market for AI-workstations collapsed
- Research funding shrunk
- The 1990s (1987-1993) sometimes called the second AI winter



14.4 More recent trends

IN3050/IN4050 Introduction to Artificial Intelligence and Machine Learning

Trends since 1990

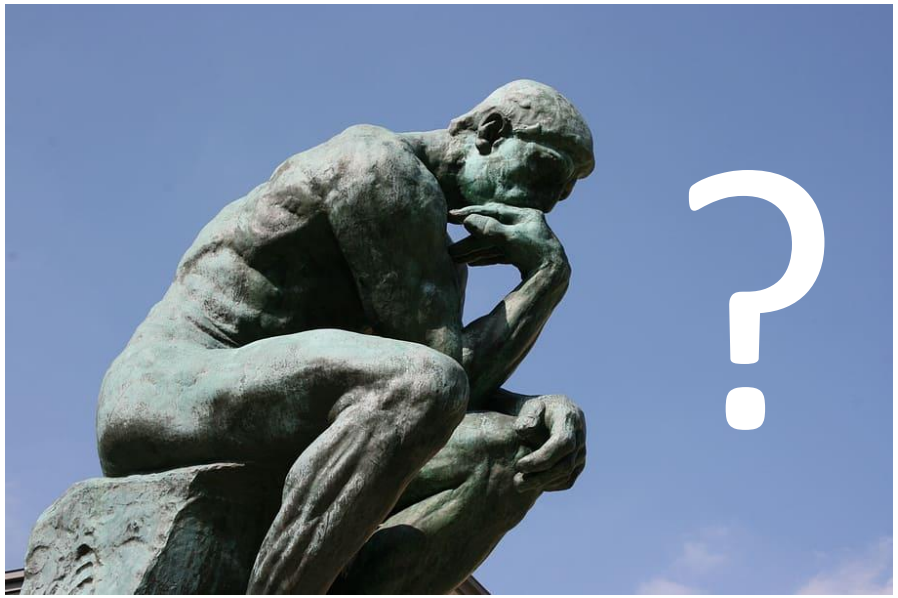
Nouvelle AI/Behavior-based robotics

AI becomes an empirical science

Deep Learning

What is Artificial Intelligence?

"A machine is considered intelligent if it can perform tasks which are considered intelligent when carried out by a human being." (Definition?)



Hence AI has focused on what is typical for only humans:

- Language
- Mathematics

Intelligence?

Higher-level intelligence

- Playing chess
- First year university mathematics
- Machines are good at this

Lower-level intelligence (?)

- Face recognition
- Moving around
- Animals are good at this
- Machines are not

Difficult things are easy – Easy things are difficult

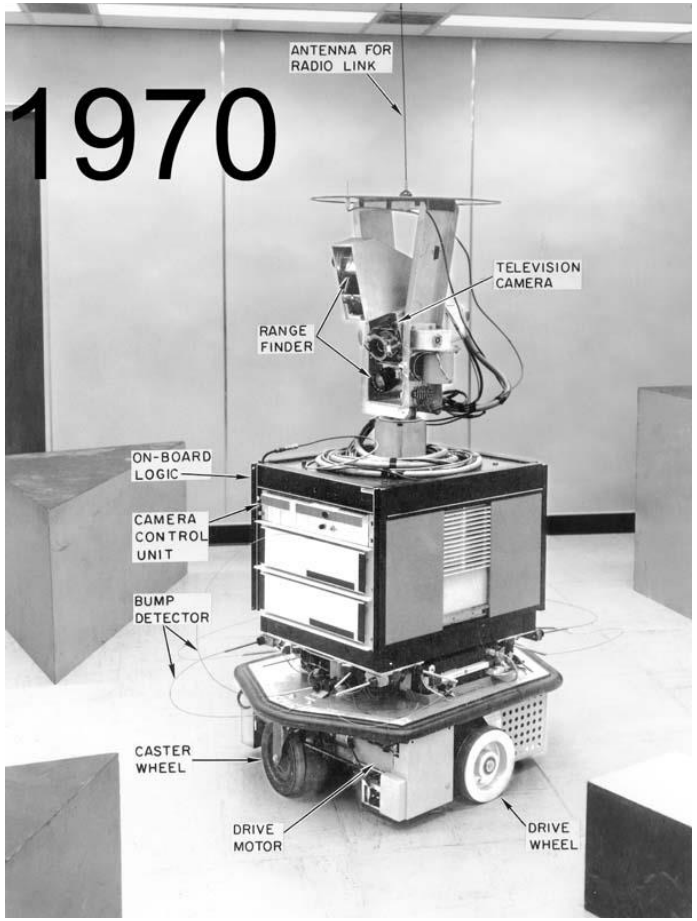
Rodney Brooks

- Robot researcher
- MIT: 1984-2007
  - Director for AI Lab, MIT
- *Intelligence without*
  - *representation*, 1987/1991
  - *reason*, 1991
- *Elephants don't play chess*, 1990



Critical towards traditional AI

- Studied isolated problems: chess, language, etc.
  - How to put them together?
- Split the AI-part off from the rest
  - No such separation in the real world
- Controlled, restricted environments
  - Instead of real word
- The *Sense-model-plan-act*-framework
  - E.g., Shakey, ca 1970 (Stanford)
  - Too much emphasis on inner representations
  - Preplanned tasks
- In general, AI is too much influenced by the von Neuman architecture: a CPU



Brooks' alternative proposal: Behavior-based Robotics



- Inspired by biology and the evolution,
  - e.g., vertebrates 550 mill. years
  - mammals for 250 mill. years.
  - Humans 1.5. mill years
- Walter's turtles

This suggests that problem solving behavior, language, expert knowledge and application, and reason, are all pretty simple once the essence of being and reacting are available. (Brooks)

- Animal-inspired robots acting in the real world
- Decomposed by activity:
  - One system for avoiding collision
  - Another system for goal-directedness
- No central representation
  - "*The world is its own best model*"

Commercialization

- Brooks with colleagues commercialized the technology (iRobot)
  - Vacuum cleaners
  - Military robots
- These ideas are also essential for the development of e.g., self-driving cars







Trends since 1990

Nouvelle AI/Behavior-based robotics

AI becomes an empirical science

Deep Learning

Example: Natural Language Processing

Laboratory - traditional

- Write neat rules which can handle a limited fragment of, say, English, very well
  - (High precision)
- Low applicability (recall)
- Example
  - *All humans are mortal*
  - *Socrates is a human*
  - *Socrates is mortal*

Out in the world

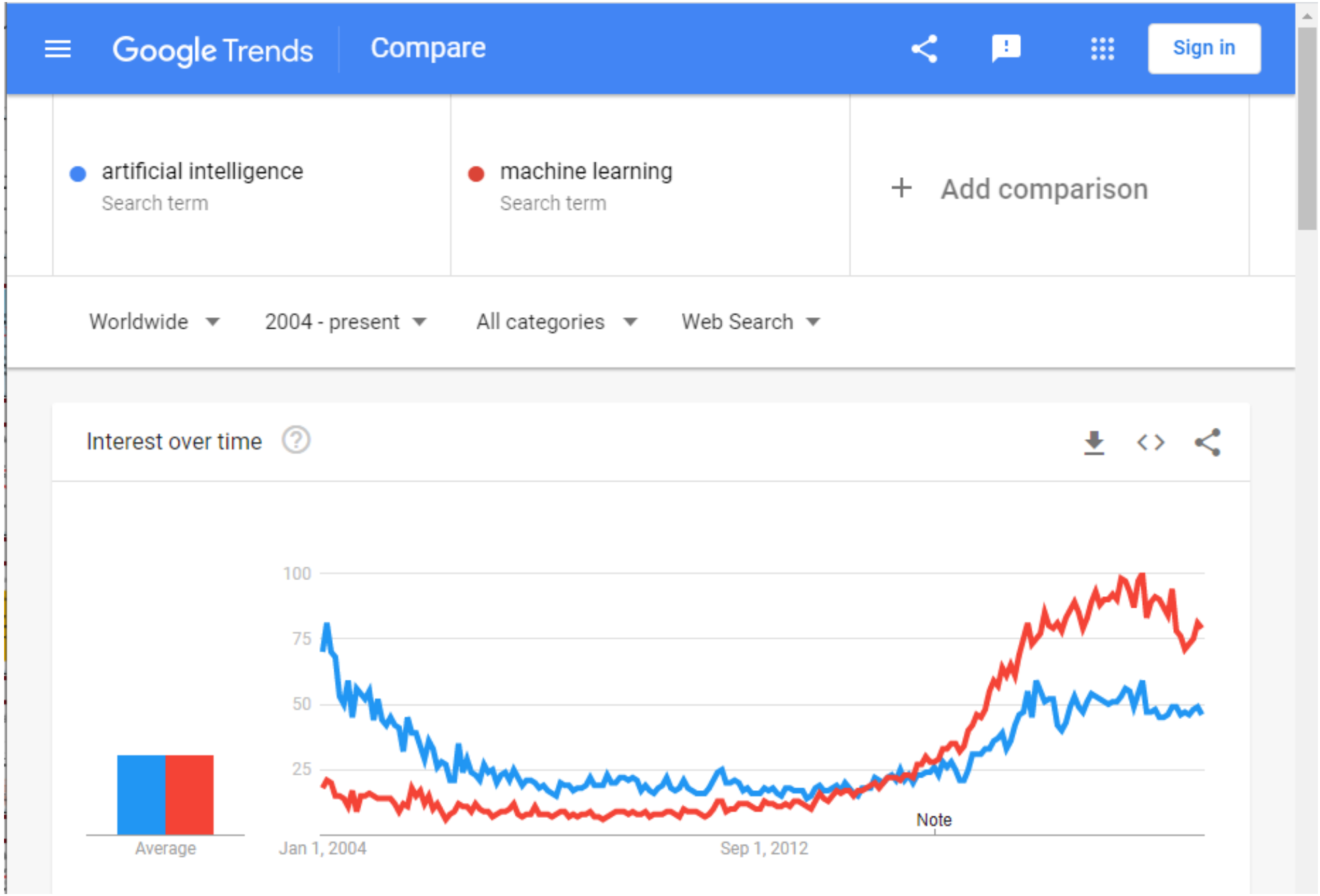
- Consider texts in the real world:
  - What can you do to them?
- Inspired by speech recognition
- High applicability, but lower precision
- Example
  - *Kim bought a rose*
  - *Kim bought a flower*

Development

- Real-world data
- A bottom-up approach compared to a top-down approach which were common in AI/NLP
  - Induction rather than deduction
- Probabilities:
  - What is the most probably translation of this sentence?
  - Which led to: Numerical methods
  - No longer only symbolic computing

Development ctd.

- Machine learning
- Rigid evaluation
  - Took over methods from empirical sciences, experimental method
  - Shared tasks
- Large amounts of available data
  - Data science
- Stopped to call it AI!
  - (*If it works, it is no longer AI*)
  - The AI-effect


Google Trends
Compare
Sign in
artificial intelligence
Search term
machine learning
Search term
Add comparison
Worldwide
2004 - present
All categories
Web Search
Interest over time
100
75
50
25
Note
Average
Jan 1, 2004
Sep 1, 2012

Trends since 1990

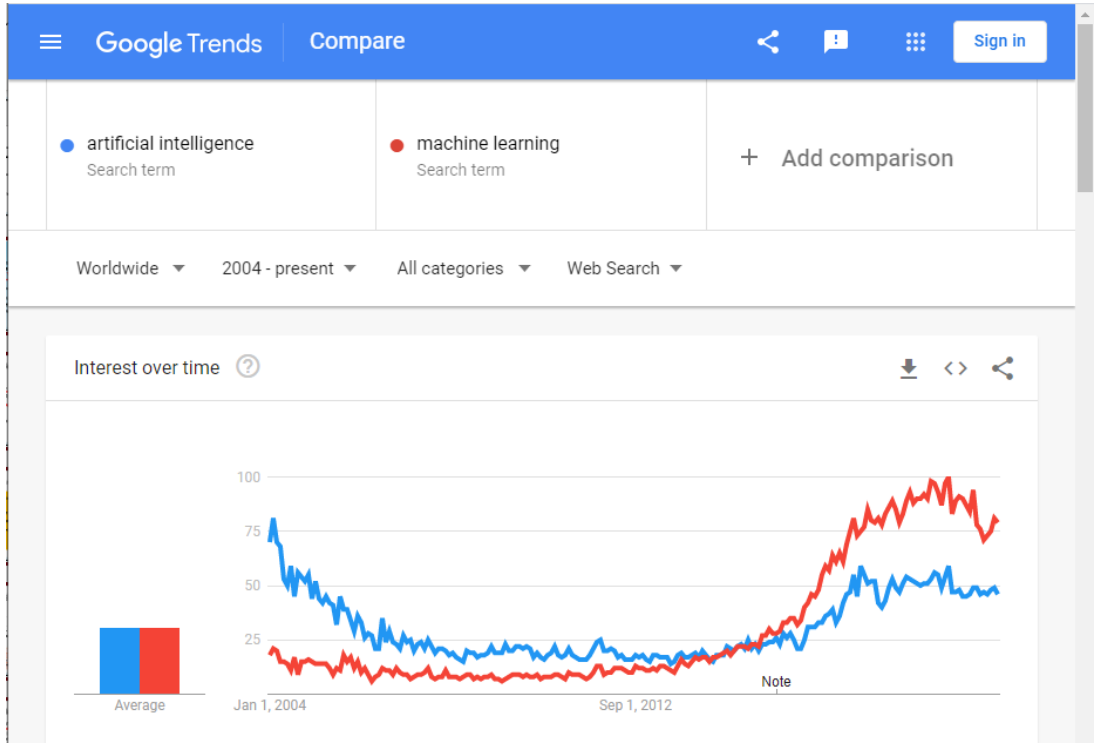
Nouvelle AI/Behavior-based robotics

AI becomes an empirical science

Deep Learning

Deep learning - Neural nets

- The third large change to AI since 1990 is the Deep learning revolution since 2012
- The revival of neural networks
- This made the term AI popular again
- And then even more popular with the advent of generative language models (ChatGPT, etc)



An additional observation

- The philosophical discussions:
  - Can machine think?
  - Are machines intelligent
- are less active today than in the last century

- But the ethical discussions concerning AI. E.g.
  - bias and fairness
  - Intrusion into our lives
- have been become more relevant now that AI systems are everywhere.

Main approaches to AI

Symbolic, Rule-based

- Logic, deduction
- Explicit coding of knowledge as formulas or rules
- Dominated AI-books until the end of the last century
- Compatible philosophy:
  - Computational theory of mind (PSSH)

Machine learning, Neural nets

- Induction rather than deduction
- Adapt to the environment
- Main-focus in this course
- Dominates AI today
- Compatible philosophy:
  - Connectionism

Do we still need symbolic, rule-based AI?

- Neural nets are good at many tasks, but not all
- Neural nets are often black boxes, they give prediction but no explanations
- There is a demand for explainable AI
- And don't forget, neural nets did not get much attention for 15 years.
- The rule-based symbolic AI may strike again.

