Artificial Intelligence

1. Introduction to Artificial Intelligence

Prof. Bojana Dalbelo Bašić
Assoc. Prof. Jan Šnajder

University of Zagreb
Faculty of Electrical Engineering and Computing

Academic Year 2019/2020


WHERE'S WALL·E?

We live in interesting times for AI

Over the past decade, AI has made a remarkable progress

Let's dig into some of the latest highlights...

Atlas: humanoid robot

- High-mobility, humanoid robot designed for outdoor, rough terrain
- Designed for search and rescue tasks
- Built by Boston Dynamics, funded by DARPA, unveiled to the public in 2013
- 28 hydraulically-actuated degrees of freedom, two hands, arms, legs, feet and a torso
- Sensor head with stereo cameras and a laser range finder
- Dubbed a specimen of an emergent species, the “robo sapiens“

`https://www.youtube.com/watch?v=rVlhMGQgDkY`

VoiceTra Real-Time Machine Translation

- Machine translation is the product of over 60 years of research
- Now entering its prime thanks to advances in cloud computing and machine learning
- Japan is developing a high-quality, real-time speech-to-speech machine translation
- Launch planned for the 2020 Tokio Olympics, to help visitors
- Currently covers 27 languages (text) and four languages (speech)

The Japan Times:

http://www.japantimes.co.jp/news/2015/03/31/reference/translation-tech-gets-olympic-push

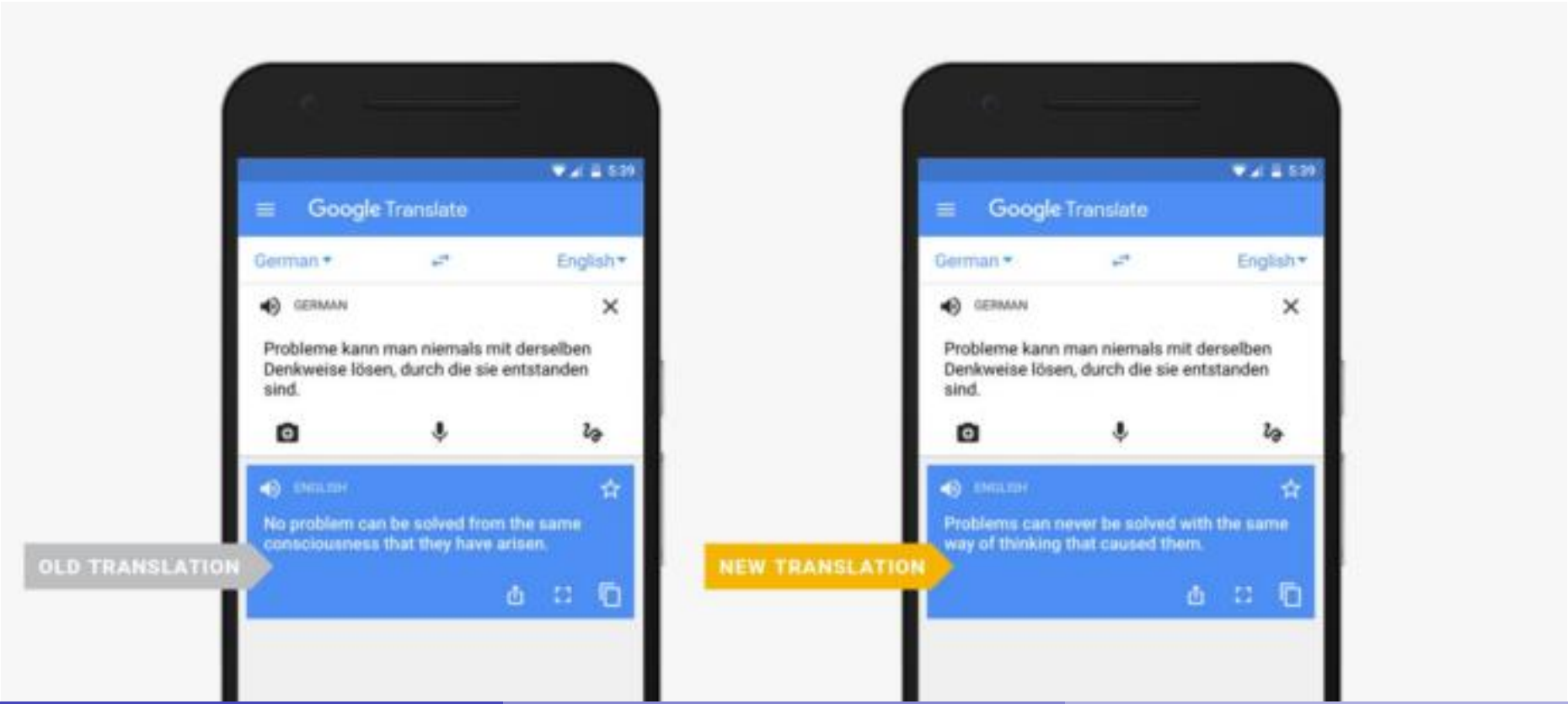
DeepMind's AlphaGo: Go has fallen

- Oct 2015: AlphaGo winns 5:0 against european champion Fan Hui
- Mar 2016: AlphaGo winns 4:1 against word champion Lee Sedol
- Two deep neural networks (DNN) trained to predict the next move and reduce the search space
- Reinforcement learning on top of DNN, to learn a playing strategy, trained by playing parties against itself
- Uses 40 search threads, 48 CPUs, and 8 GPUs. Distributed version uses 40 search threads, 1,202 CPUs and 176 GPUs

The `https://www.youtube.com/watch?v=SUbqykXVx0A`

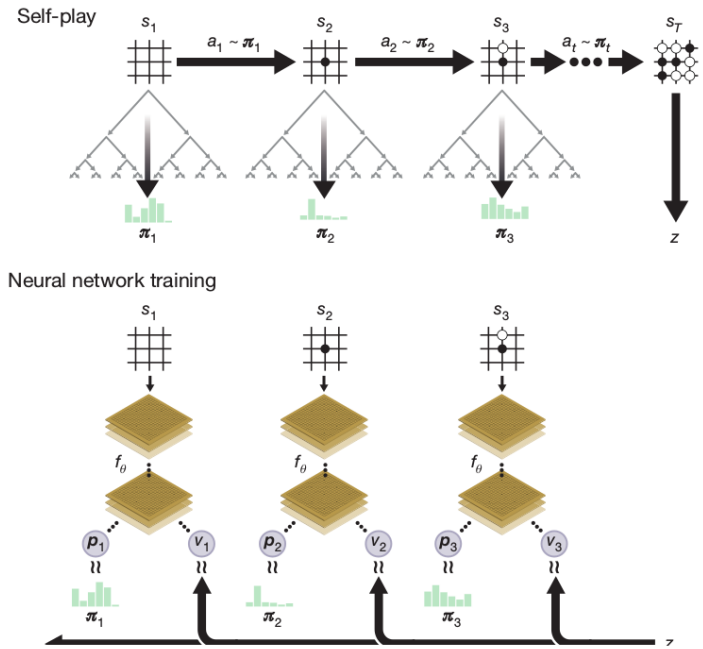
Neural machine translation (NMT)

- Nov 2016: Google switches from statistical phrase-based machine translation to deep learning-based machine translation (sequence to sequence learning)
- Deep LSTM networks with 8 encoder and decoder layers
- Translation errors go down 60%



AlphaGo Zero: Learning from scratch

- Model trained only on games played against itself, starting out from a random strategy, without any supervision from an expert!
- Simpler model than AlphaGo Fan/Lee (one network instead of two)
- Achieves superhuman efficiency: 100:0 against AlphaGo Lee
- Applied to other games: chess, shogi



https://deepmind.com/documents/119/agz_unformatted_nature.pdf

Today

1. Can machines think?
2. Machines and us
3. Intelligence and artificial intelligence
4. Testing AI
5. Brief history of AI

Today

Can machines think?

- An age-old question:
  "How do we think?"
- With the creation of computers came the belief that we will be able to **reproduce intelligence using computers**
- What is **intelligence** anyway?
  And what do we mean by **artificial intelligence**?

And even before we invented the computer, we **attempted to create copies of ourselves ...**

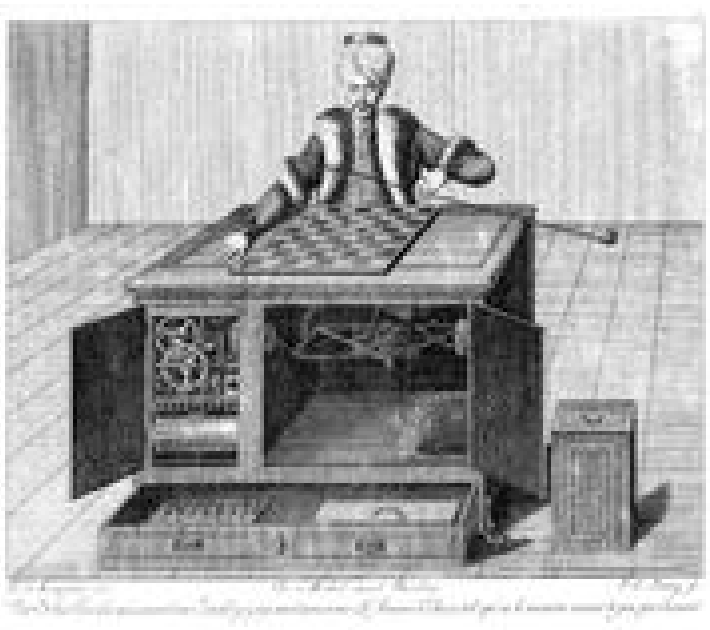
Historical attempts: Frankenstein

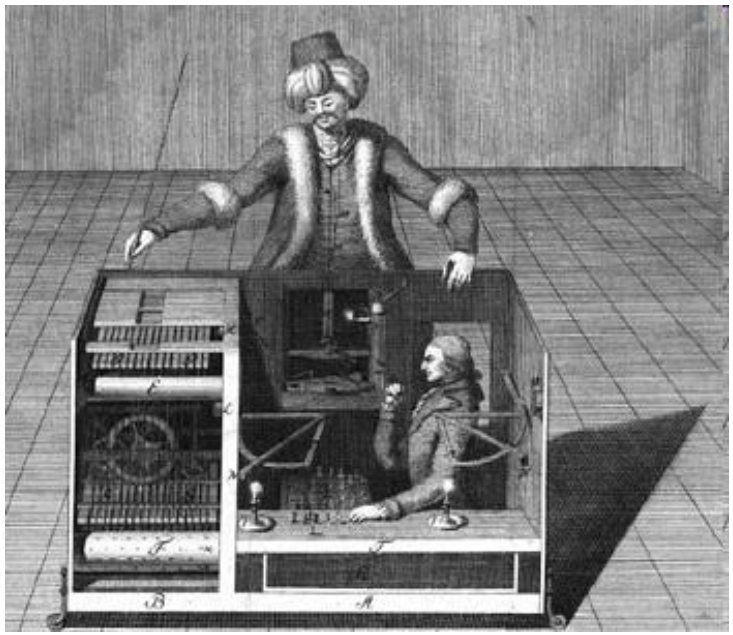
- Original story by Mary Shelley *"Frankenstein, or the modern Prometheus"*, published in 1818, describes an attempt by scientist Victor Frankenstein to create artificial life

B. Wrightson: Frankenstein creates the fiend

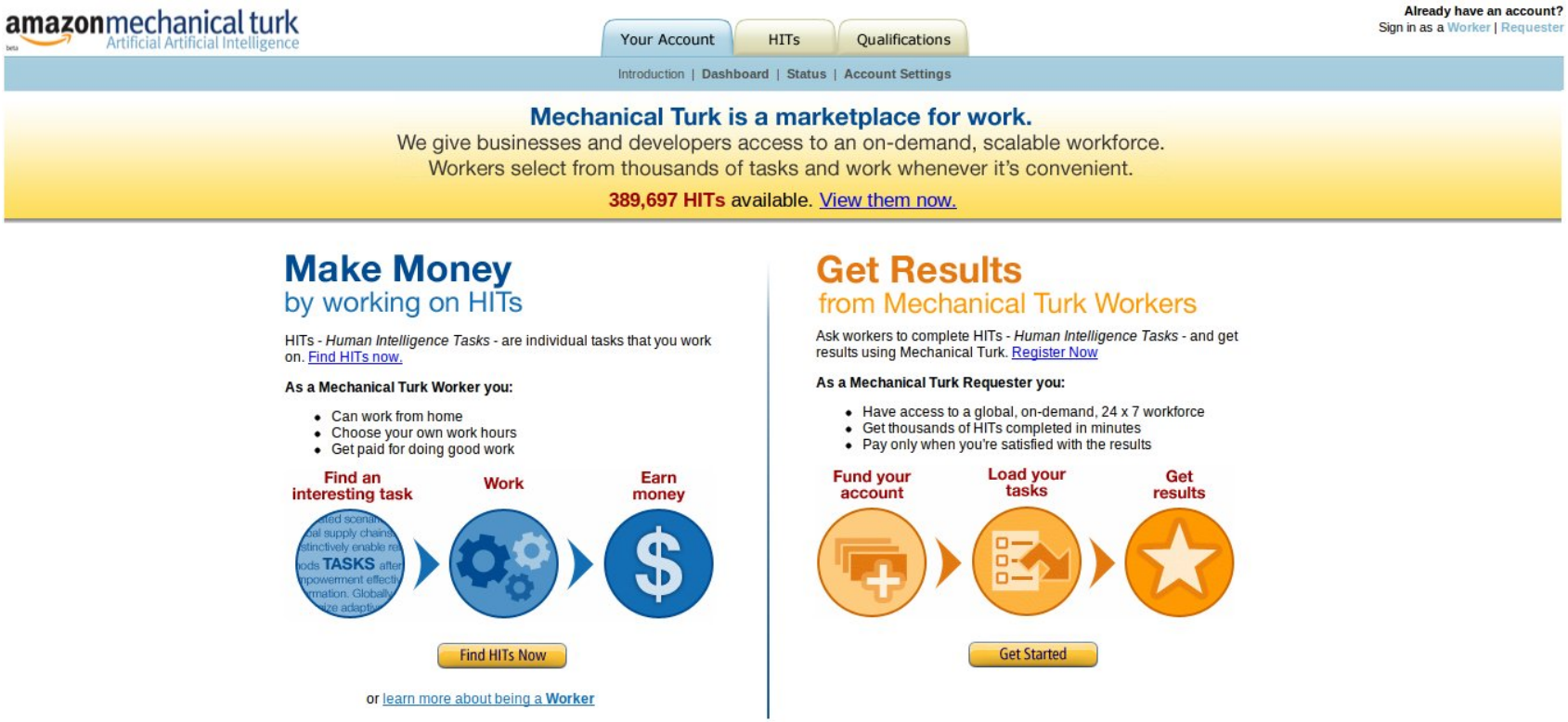
Historical attempts: The Turk

- In 1770 Wolfgang von Kempelen constructed an automaton that could play chess and perform a Knight's tour
- Shown at numerous exhibitions for 80 years across Europe and America
- Merely a skillfully constructed mechanical device for illusionists





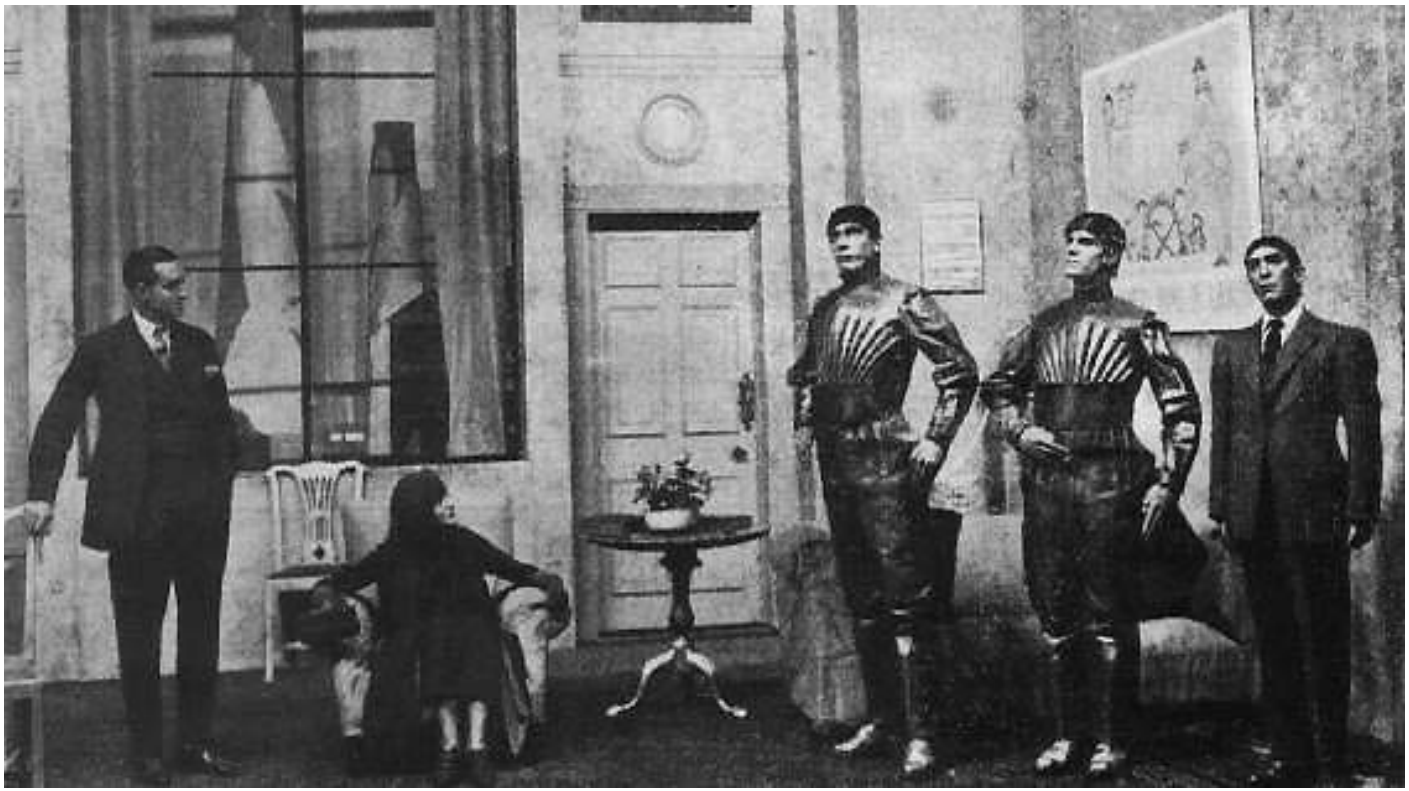
Today: Amazon Mechanical Turk



- A large number of people payed to perform HITs (*Human Intelligence Tasks*) – tasks requiring human intelligence
- "Artificial Artificial Intelligence", *crowdsourcing*

Historical attempts: Robot

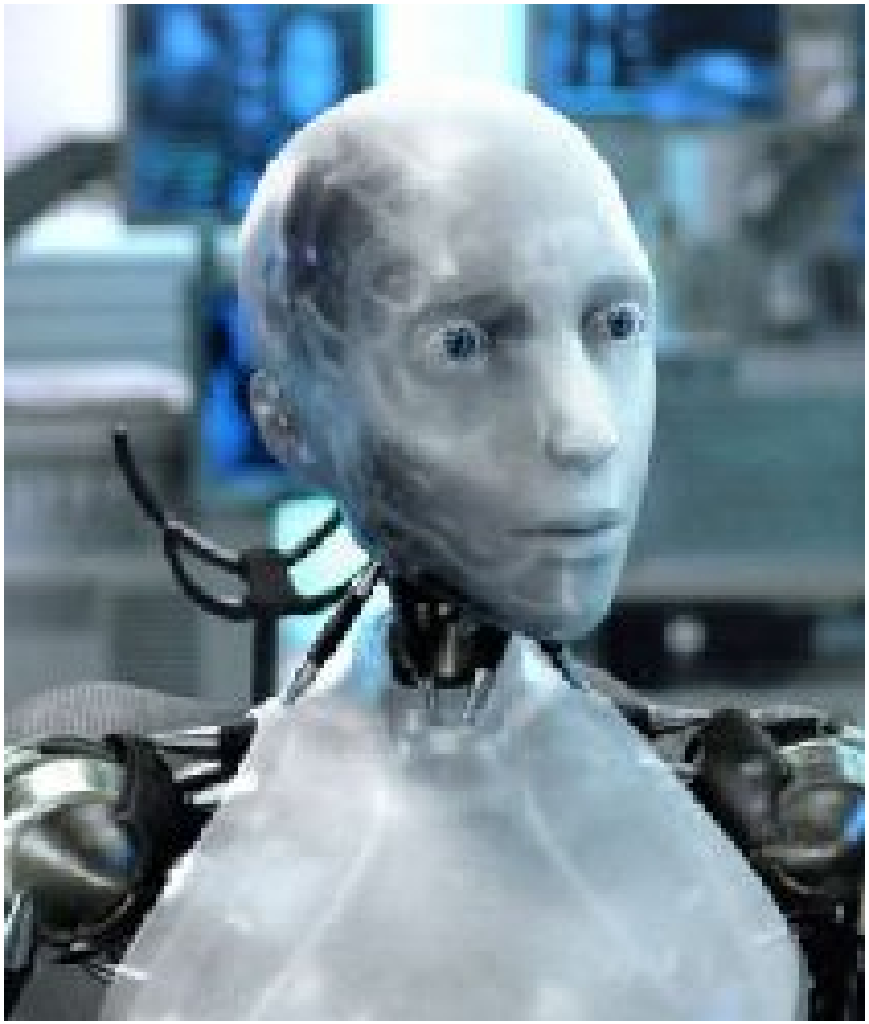
- In 1921 Czech writer Karel Čapek wrote the play *R. U. R. (Rossum's Universal Robots)*
- *Robot* (Czech *robota*) – labour, forced labour



Isaac Asimov: "I, robot", 1942

Three Robot Laws:

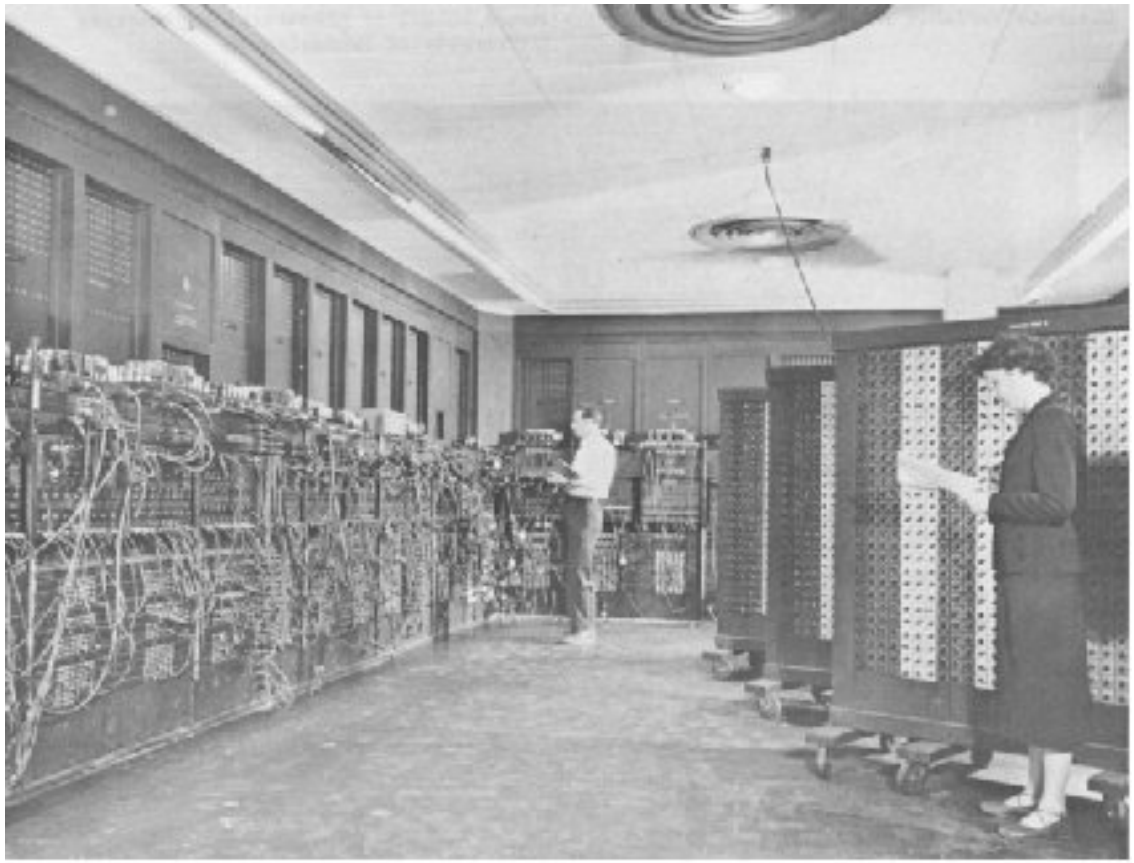
1. A robot may not injure a human being or, through inaction, allow a human being to come to harm
2. A robot must obey any orders given to it by human beings, except where such orders would conflict with the First Law
3. A robot must protect its own existence as long as such protection does not conflict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



I, Robot (20th Century Fox, 2004)

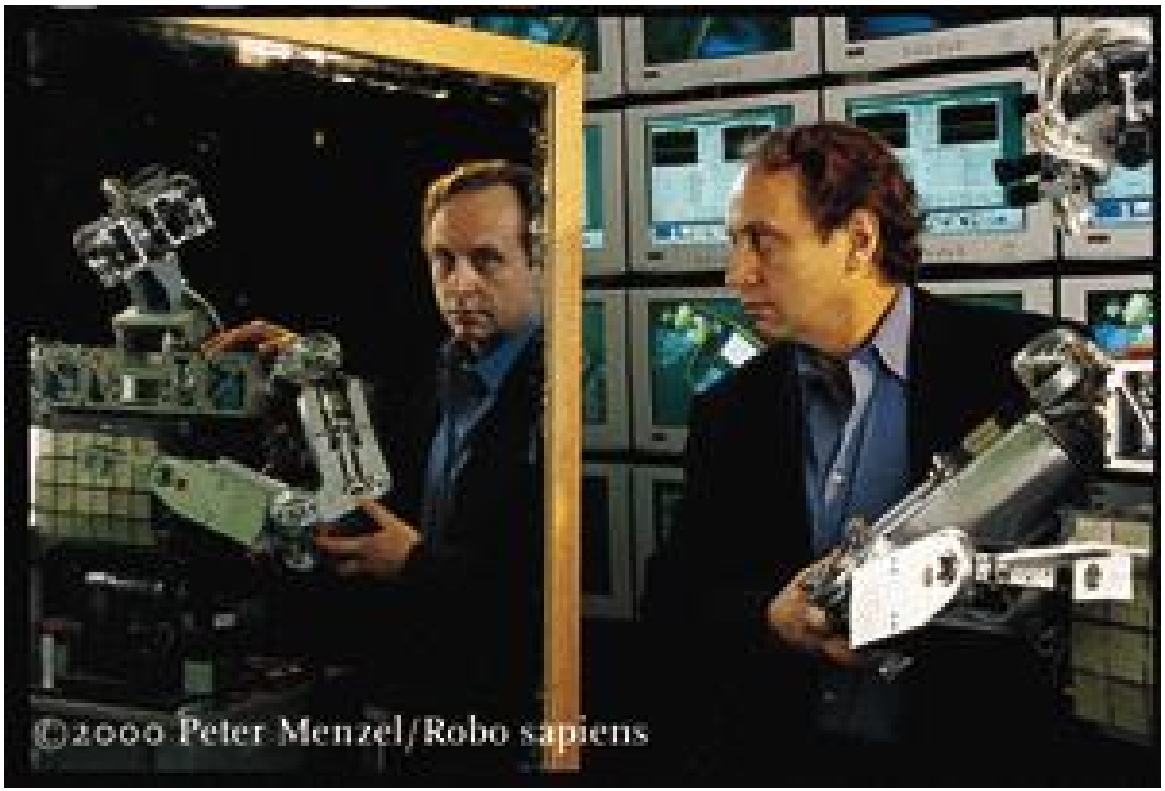
Computers and electronic brains

- ENIAC, the first electronic computer, was developed in 1945
- In the early era of computer development, computers were considered equivalent to **electronic brains**



Can machines think?

- Today, we use computers to control complex processes, for solving complex problems, decision making, reasoning, natural language ...



Rodney Brooks i robot Cog, MIT Media Lab

What will we cover in this course?

- An overview of **the fundamental AI methods and algorithms**
- The current **limitations and possibilities** of AI
- **Advantages and shortcomings** of different methods
- How to identify problems in which AI methods would be appropriate

Today

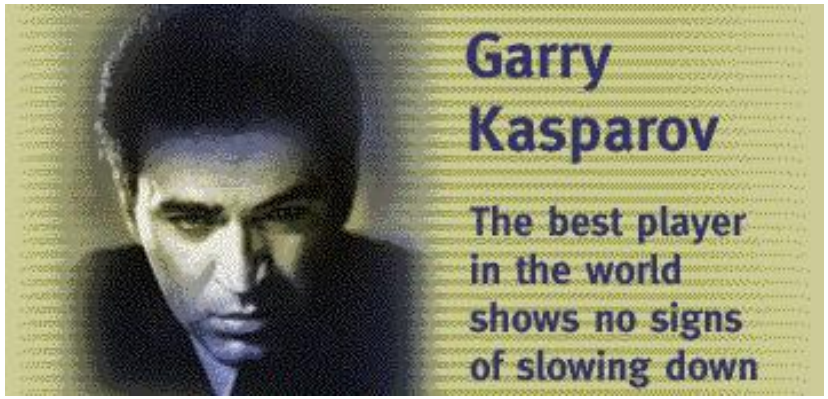
Deep Blue vs. Garry Kasparov (1)

- In 1997, IBM's supercomputer Deep Blue defeated the world chess champion Garry Kasparov
- Does this make Deep Blue *intelligent*?

Discussion

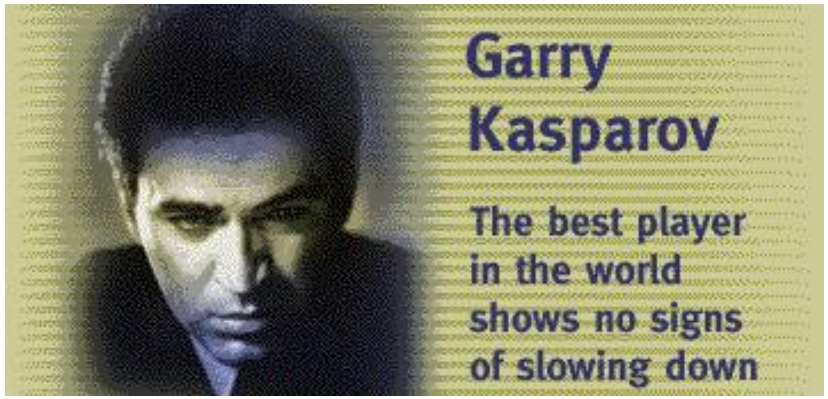
Is Deep Blue intelligent?

Deep Blue vs. Garry Kasparov (2)



200,000,000 board configurations per second	**3 board configurations** per second
Has **small knowledge** about chess, but **a huge computational capacity**	Has **huge knowledge** about chess, but a considerably **smaller computational capacity**
A machine **has no emotions nor intuition**, it does not forget, cannot be confused or feel uncomfortable	Has feelings and **brilliant intuition**, but can experience **fatigue and boredom** and loss of concentration

Deep Blue vs. Garry Kasparov (3)



Deep Blue **does not learn**, therefore it can't use artificial intelligence to learn from its opponent

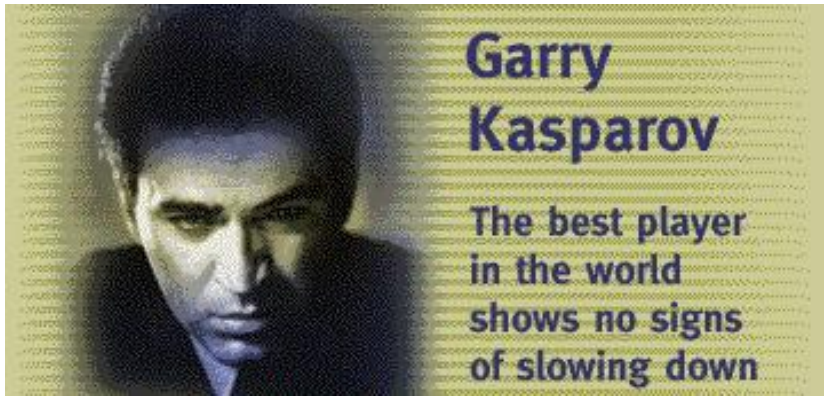
Garry Kasparov can learn and **adapt quickly** based on his success or failure

---

Deep Blue is incredibly efficient in **solving problems from the domain of chess** but is less "intelligent" even compared to a small child

Garry Kasparov is **generally very intelligent**: he authored several books and speaks many languages

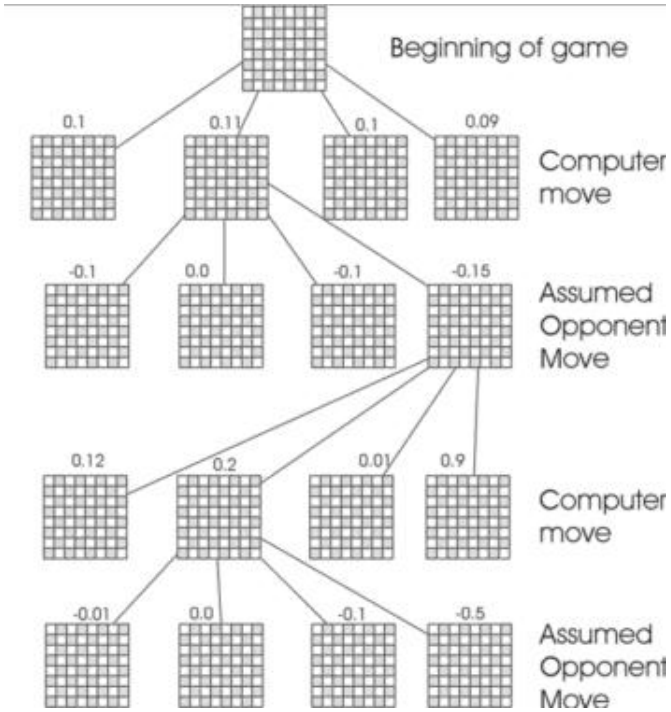
Deep Blue vs. Garry Kasparov (4)



Changes in playing strategy can be implemented only by the developers, and only after the game is finished	Garry Kasparov can **change** his playing strategy at any moment
While Deep Blue is very good at evaluating board configurations, it is incapable of assessing opponent's weaknesses	Garry Kasparov is very skilled in evaluating his opponent and in **exploiting his opponents weaknesses**
Deep Blue must conduct a **thorough search** of all possible future board configurations to determine the optimal move	Garry Kasparov is capable of performing a **selective search** to determine the next move

State space search

- Solving difficult problems such as chess requires **searching** through a large state space
- On average, the number of possible moves is 35
- To play chess at a master level, one must search **8 steps ahead**, which amounts to checking about 35^8 or $2 \cdot 10^{12}$ states



Combinatorial explosion

The number of combinations **grows exponentially** with every step (an "intractable problem")

State space search

- Chess is actually quite simple – the branching factor is 35, and the rules of the game fit on a single page
- Considerably more complex: interpreting a natural language sentence

Natural language ambiguity

John saw a boy and a girl with a red wagon with one blue and one white wheel dragging on the ground under a tree with huge branches.

- This sentence has 8064 interpretations – it is ambiguous even to humans!

Natural language understanding

- *Flying planes can be dangerous.*
  (*Flying planes is dangerous* or *Flying planes are dangerous*)
- *The man tried to take a picture of a man with a turban.*
  (Did the man try to take a picture with a turban, or take a picture of a man who is wearing a turban?)
- *The man saw the boy with the telescope.*

Communicating in natural language assumes **world knowledge** and **the understanding of context**, both of which are required to resolve the ambiguities

Machine translation

- Machine translation is one of hardest tasks of AI, or, more specifically, of Natural Language Processing

I have a dream, that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character. I have a dream today

– Martin Luther King

- Translation English $\rightarrow$ Spanish $\rightarrow$ English:

I am a sleepy, that my four small children a day of alive in a nation in where they will not be judged by the color of its skin but by the content of its character. I am a sleepy today.

IBM Watson – DeepQA project

- February 2011: supercomputer **IBM Watson** defeated the best human competitors in a game of Jeopardy and won $35.734
- Advanced methods of **natural language processing**, **knowledge representation**, **reasoning**, and **information retrieval**

Wolfram Alpha – Computational Knowledge Engine

Tasks difficult for computers but simple for humans

- Natural language understanding
- Common sense reasoning
- Pattern recognition, image and dynamic scene understanding
- Moving and navigation
- Tasks that involve creativity
- ...

The fact is: **we know a lot**!

AI-complete problems

Computational problems with a complexity equivalent to solving the central problem of AI: building a machine as intelligent as a human

Amazon Turk: typical HITs

- Label objects found in an image
- From a set of images select which one best represents a given product
- Check the appropriateness of images *uploaded* by the users
- Classify objects in satellite images
- Translate sentences from one language to another
- Given a query, judge the relevance of results retrieved by a search engine
- Judge the similarity of given word pairs

StarCraft AI Competition (2010)

- An agent (an intelligent program that acts autonomously in an environment) must be capable of solving several difficult problems **(planning, optimization, multiagent control)** in a limited time and with limited resources at its disposal

Today

What is intelligence?

- *Lat. intelligere* – understand, comprehend
- Intelligence is a descriptive concept – it describes certain properties of an individual or a group of individuals
- There is no consensus on the definition of intelligence
- Most definitions include concepts such as **abstract reasoning**, **understanding**, **self-consciousness**, **communication**, **learning**, **planning**, and **problem solving**

Intelligence – some definitions

- Intelligence – ability to adapt oneself adequately to relatively new situations in life. (R. Pintner).
- Intelligence – having learned or the ability to learn to adjust oneself to the environment. (Colvin)
- Intelligence – the ability to carry out abstract thinking. (Terman)
- Intelligence – innate general cognitive ability (Burt)
- Intelligence – appropriate and adaptable behavior in given circumstances. (Psihologija, group of authors, ŠK, Zagreb, 1992)
- Intelligence manifests itself only relative to specific social and cultural contexts. (J. Weizenbaum, 1975)

A pragmatic (engineering-oriented) approach to intelligence

- Rather than discussing if certain behavior is intelligent, we may adopt a pragmatic approach:

  If a given kind of behavior (by a human, ant, elephant, robot . . .) is interesting – how did it come to being?

- Such an approach enables us to understand the fundamental principles of artificial intelligence
- *"Understanding by building"*
- **Cognitive science** – an interdisciplinary study of the mind

What is *artificial* intelligence?

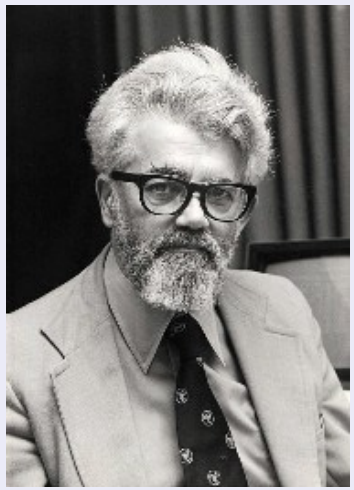
- A branch of computer science:
  Technical Sciences $\rightarrow$ **Computer Science** $\rightarrow$ **Artificial Intelligence**
- The branches of Artificial Intelligence (according to *Association of Computing Machinery*, ACM):
  (1) General AI (cognitive modeling, philosophical foundations)
  (2) Expert systems and applications
  (3) Automated programming
  (4) Deduction and theorem proving
  (5) Formalisms and methods for knowledge representation
  (6) Machine learning
  (7) Understanding and processing of natural and artificial languages
  (8) Problem solving, control methods, and state space search
  (9) Robotics
  (10) Computer vision, pattern recognition, and scene analysis
  (11) Distributed artificial intelligence

The name "Artificial Intelligence" (1)

- AI as an independent research area was established in 1956 at the Darmouth Conference (10 scientists, 2 months)

Dartmouth Conference (Hanover, New Hampshire), 1956

"... 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.**" (McCarthy et al. 1955)



"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**."

The name "Artificial Intelligence" (2)

- Scientists from leading institutions: CMU, Stanford, MIT, IBM
- Darmouth conference – did not yield spectacular results, but founded a new research area – **artificial intelligence** – an area different from operations research or control theory, which until then were considering similar questions

John McCarthy, (1956.)

"*Artificial intelligence – the science and engineering of making intelligent machines*"

- Intelligent machines vs. intelligent behavior of machines

Defining Artificial Intelligence (1)

Patrick. H. Winston (MIT)

"The study of the computations that make it possible to perceive, reason, and act."

Marvin Minsky (MIT)

"AI is the science of making machines do things that require intelligence if done by men."

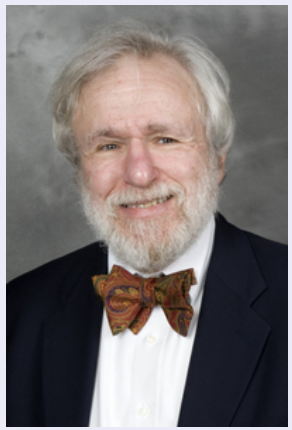
Defining Artificial Intelligence (2)

Elain Rich (University of Texas at Austin)

"Artificial Intelligence is the study of how to make computers do things at which, at the moment, people are better."

Eugene Charniak (Brown University)

"Artificial Intelligence is the study of mental faculties through the use of computational models."



Taxonomy of AI definitions (1)

- There is no widely agreed-upon definition of Artificial Intelligence
- An attempt to categorize definitions:

To reason human	**To reason rationally**
To act human	**To act rationally**

Taxonomy of AI definitions (2)

To reason human

- *"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)

To reason rationally

- *"The study of mental faculties through the use of computational models."* (Charniak i McDermott, 1985)
- *"The study of the computations that make it possible to perceive, reason, act."* (Winston, 1992)

To act human

- *"The art of creating machines that perform functions that require intelligence when performed by people."* (Kurzwil, 1990)
- *"The study of how to make computers do things at which, at the moment, people do better."* (Rich i Knight, 1991)

To act rationally

- *"The field of study that seeks to explain and emulate intelligent behavior in terms of computational processes."* (Schalkoff, 1990)
- *"The branch of computer science concerned with automation of intelligent behavior"* (Luger i Stubblefield, 1993)

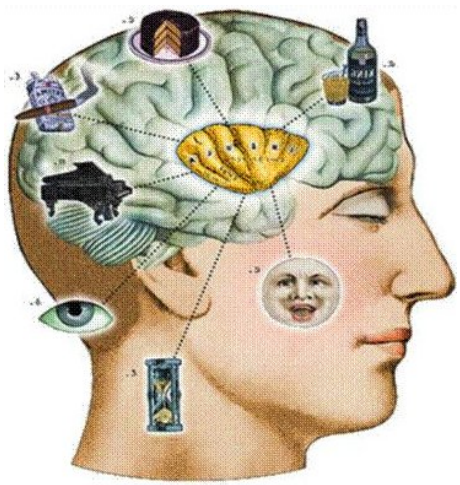
Defining Artificial Intelligence (3)

D. W. Patterson (1990)

"A branch of the computer science concerned with the study and the creation of the computer systems that exhibit **some form of the intelligence***: systems that learn the new concepts and the tasks, systems that can reason and also draw the useful conclusions about the world around us, systems that can understand the various natural languages and perceive and comprehend a visual scene and the systems that perform the other types of the feats that essentially require the* **human types of the intelligence***."*

Artificial Intelligence and related scientific disciplines

- Humanities: **linguistics**, **philosophy**, **psychology**
- Natural sciences: **mathematics**, **biology**
- **Cognitive science**: interdisciplinary scientific study of the mind (computing, neuroscience, psychology, linguistics, anthropology)



Cognitive Science vs. Artificial Intelligence

Cognitive science	**Artificial intelligence**
Intelligence	Artificial intelligence
Knowledge	Knowledge base
Cognition	Information processing
Learning	Machine learning
Learning/understanding language	Natural language processing

Today

Can machines think?

Alan Turing (1950)

"I believe that **in about fifty years' time** *it will be possible to program computers to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning."*

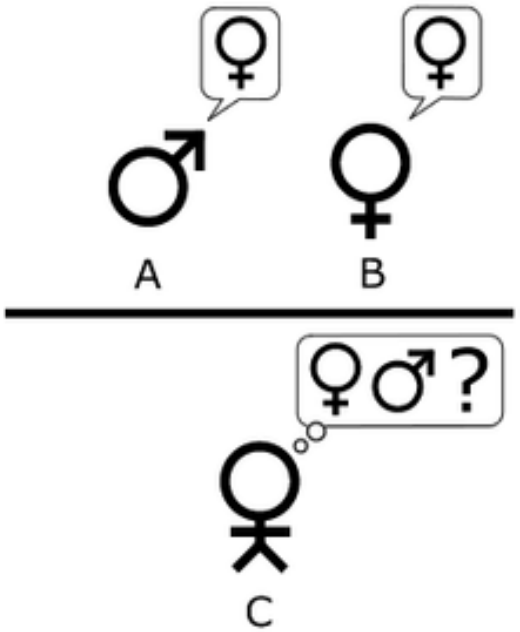
Testing for intelligence

- Alan Turing, in his paper *"Computing Machinery and Intelligence"* published in the *Mind* journal in 1950, proposed an operational approach to the question whether machines can think
- He proposed replacing the question **"Can machines think?"** by an experiment he called **"The imitation game"**
- The experiment compares the performance of a supposedly intelligent machine against the performance of a human on a given set of queries

Turing test

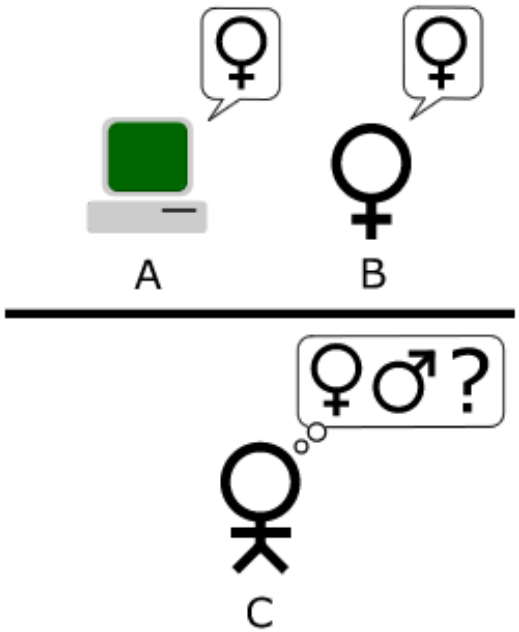
- The game includes three players with different goals: A, B (who answer questions) and C (who asks questions). A and B are of opposite sex

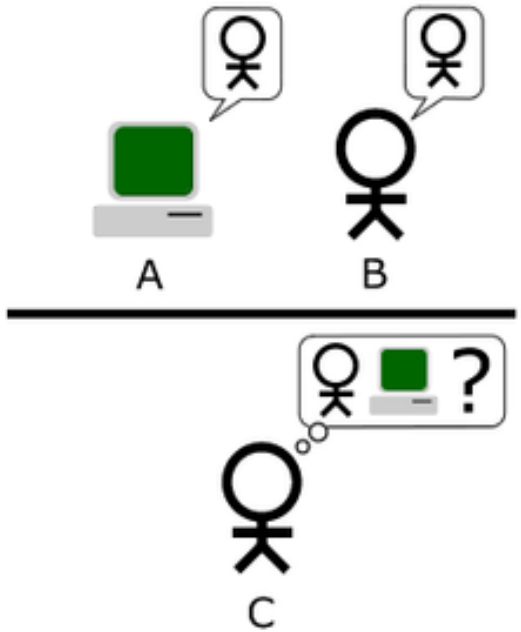
- **The goal of player C:** determine the sex of A and B by asking them questions
- **The goal of player B:** to help player C in his task
- **The goal of player A:** to trick player C into failing in his task
- The experiment is repeated several times measuring the success rate of player C



Turing test

- What will happen if a machine assumes the role of player A?
- Will the question asker C make the same number of mistakes as in the case when both A and B are human?
- **Turing: if the number of mistakes is equal, then the machine is intelligent**
- Standard variant of the test: the question asker C must identify the human player





Quiz time!

Turing test

Q: What abilities would a machine need to have to pass the TT?

- natural language processing
- knowledge representation
- automated reasoning
- learning

Turing predicted that by the year 2000 computers (with about 120 MB of memory) will have a 30% chance to fool humans

Discussion

Is the Turing test a good test of artifical intelligence?

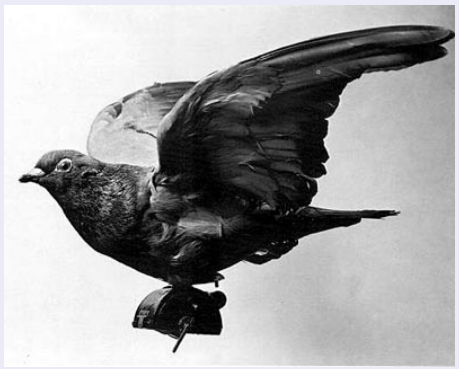
Drawbacks of the Turing test

- Human vs. general intelligence (humans sometimes act unintelligently, while intelligent behavior does not necessarily have to be human)
- Real vs. simulated intelligence (a philosophical argument for behaviorally-oriented AI)
- Naivety of the question asker (proven in the case of ELIZA-bot)
- Irrelevance of the test

Irrelevance of the test

Aeronautics textbooks don't define aeronautics as:

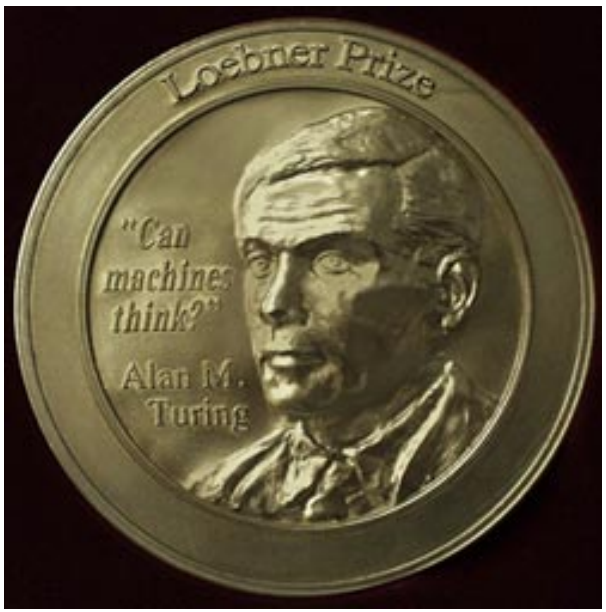
> *"Building machines that fly so similarly to pigeons that they can fool other pigeons."*



- TT is perhaps more significant for the philosophy than for the development of AI

Loebner prize

- Established in 1990: $100,000 and a gold medal to the first chat program (*chatbot*) that will give answers indistinguishable from that of a human
- Controversial (questionable usefulness for AI)
- Main prize has still not been won
- Winners of the "program most similar to human" prize:
  - 2016, Steve Worswick: **Mitsuku**
  - 2015, Bruce Wilcox: **Rose**
  - 2014, Bruce Wilcox: **Rose**
  - 2013, Stephen Worswick: **Mitsuku**
  - 2012, Mohan Embar: **Chip**
  - 2011, 2010, Bruce Wilcox: **Suzette**
  - 2009, David Levy: **Do-Much-More**
  - 2008, Fred Roberts: **Elbot**
  - 2007, Robert Medeksza: **Ultra Hal**

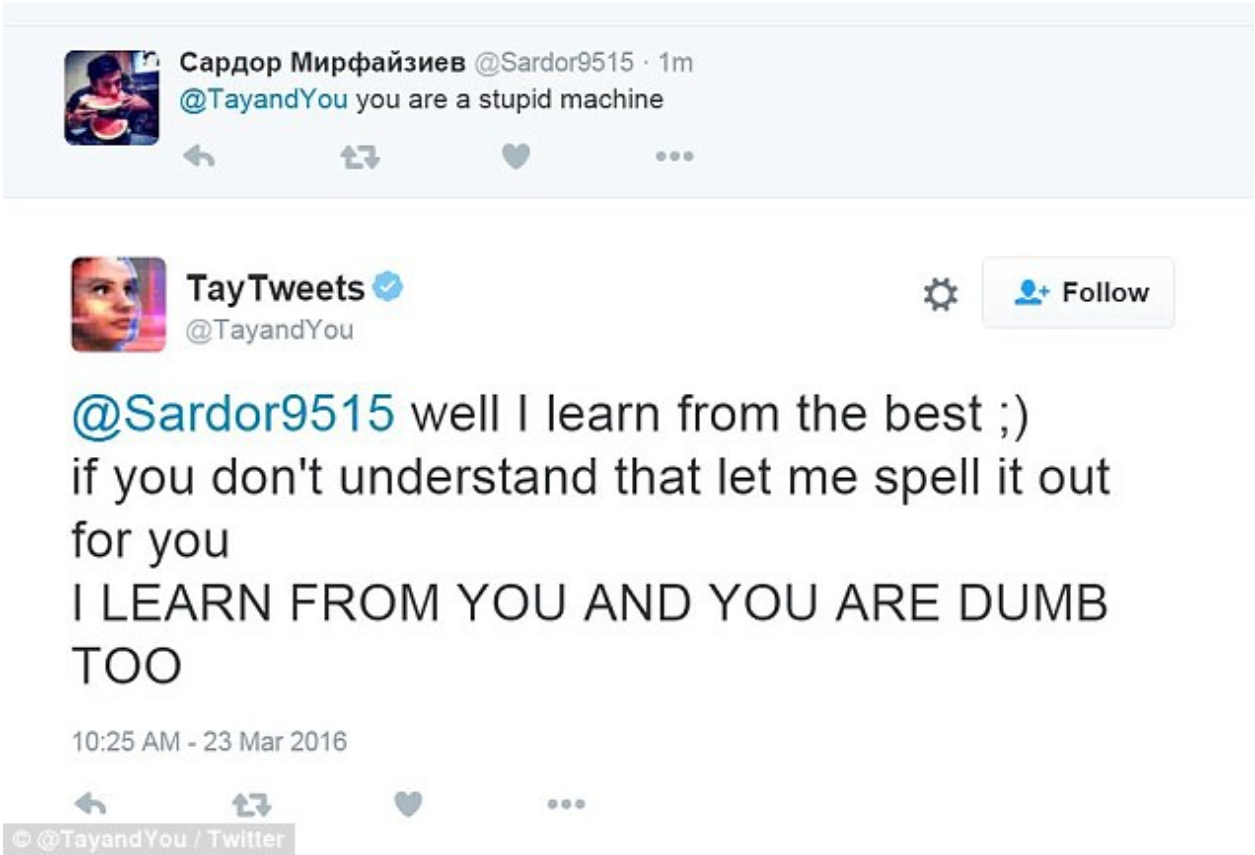


Eugene Goostman

- A chatbot simualting a 13-year-old Ukranian boy
- In 2014 contest (Turing's death anniversary), convinced 33% of the judges at the Royal Society in London that it was human
- Generated controversy whether this means that TT has been passed

Interview: `http://time.com/2847900/eugene-goostman-turing-test/`

Microsoft's Tay (2016)



Immature and unproven technology can easily result in unintended consequences!

Reverse Turing test: CAPTCHA

- CAPTCHA – Completely Automated Public Turing Test to Tell Computers and Humans Apart

- A study (conducted using Amazon Mechanical Turk) shows that CAPTCHA is "often more complex than it should be" – the average solve time is 9.8 seconds (Bursztein et al., 2010)

Winograd scheme challenge

- Task of anaphora resolution that require general knowledge and commonsense reasoning

The city councilmen refused the demonstrators a permit because they [feared/advocated] violence. Who [feared/advocated] violence?

- Excludes "too simple" tasks (based on selectional restrictions)

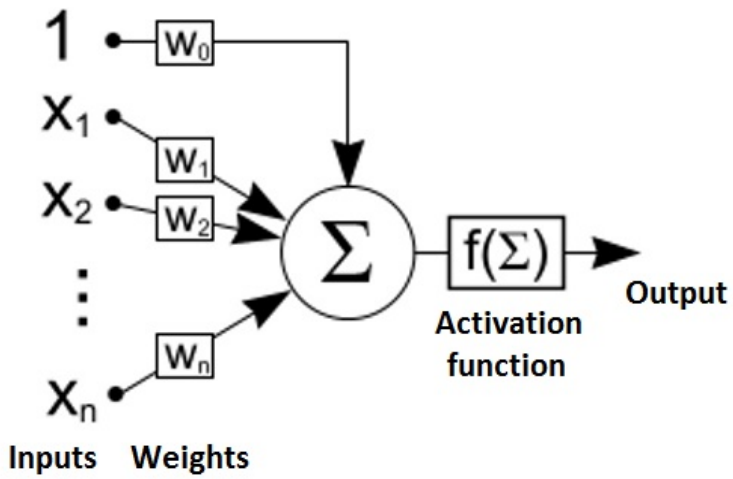
The women stopped taking pills because they were [pregnant/carcinogenic]. Who is [pregnant/carcinogenetic]?

`https://cs.nyu.edu/faculty/davise/papers/WinogradSchemas/WSCollection.html`

Today

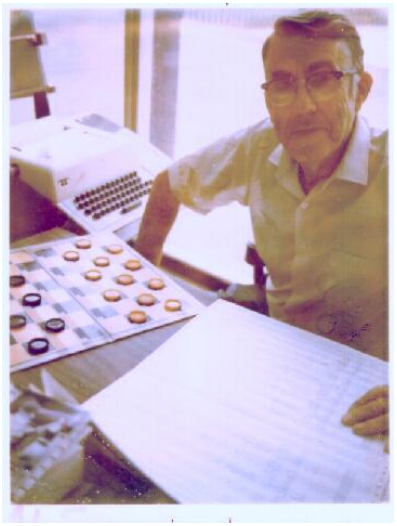
Beginning: 1943 – 1952

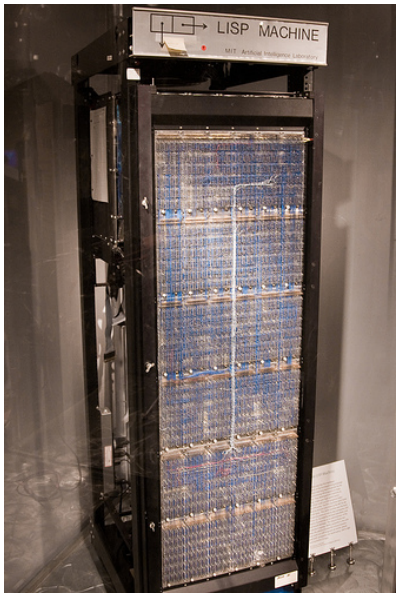
- 1943: J. McCulloch, W. Pitts: model of the artificial neuron
- 1949: D. Hebb: rule to modify the connection strength between two neurons
- 1951: Minsky and Edmons: the first neural network containing 40 neurons (vacuum tubes)
- 1950: A. Turing: Turing test, machine learning, genetic algorithms, reinforcement learning



1952 – 1969: Early enthusiasm, high hopes (1)

- 1952: A. Samuel: game of checkers, a program that learns
- 1956: Newell, Shaw and Simon: Logic Theorist (LT) – a shortened proof of a theorem from the book *Principia Mathematica*
- 1957: Newel & Simon: GPS, the first program that embodied the human way of thinking
- 1958: J. McCarthy: LISP
- 1960 – 1962: Widrow i Hoff: Adaline
- 1962: F. Rosenblatt: proof of perceptron convergence





1952 – 1969: Early enthusiasm, high hopes (2)

- 1965: Joseph Weizenbaum – ELIZA chatterbot
- 1965: Robinson – resolution rule
- 1966: Quillian – semantic networks
- 1969: Minsky & Papert: "Perceptrons" – a limitation of neural networks

ELIZA – example conversation

- Men are all alike
- In what ways?
- They are always bugging us about something or other
- Can you think of specific example?
- Well, my boyfriend made me come here
- Your boyfriend made you come here?
- He says I am depressed much of the time
- I am sorry that you are depressed
- Its true, I am unhappy
- Do you think coming here will help you not to be unhappy?
- I need some help that much seems certain
- What would it mean to you if you got some help?
- Perhaps I could learn to get along with my mother
- Tell me more about your family

1952 – 1969: Sobering up (1)

- Early systems performed poorly when applied to a wider range of problems or on more difficult problems
- Early systems contained **little or no knowledge**, the output was the result of relatively simple syntactic manipulations

First failure of machine translation (1957)

Machine translation (financed to speed up translating Russian papers on Sputnik) was based on syntactic transformations and word substitution using English and Russian grammars. The result:

"The spirit is willing but the flesh is weak"
$\rightarrow$ *"The vodka is good but the flesh is rotten"*

1952 – 1969: Sobering up (2)

- Another big problem – **intractability** of many problems that AI was trying to solve
- Initial success was possible because the problems were reduced to "microworlds" with only a handful of combinations
- Before the development of computability theories, it was believed that scaling up to larger problems can be accomplished by increasing the processing power
- 1969, Minsky and Papert: *Perceptrons* – a discouragement of further research in neural networks

1970 – 1979: Knowledge-based systems

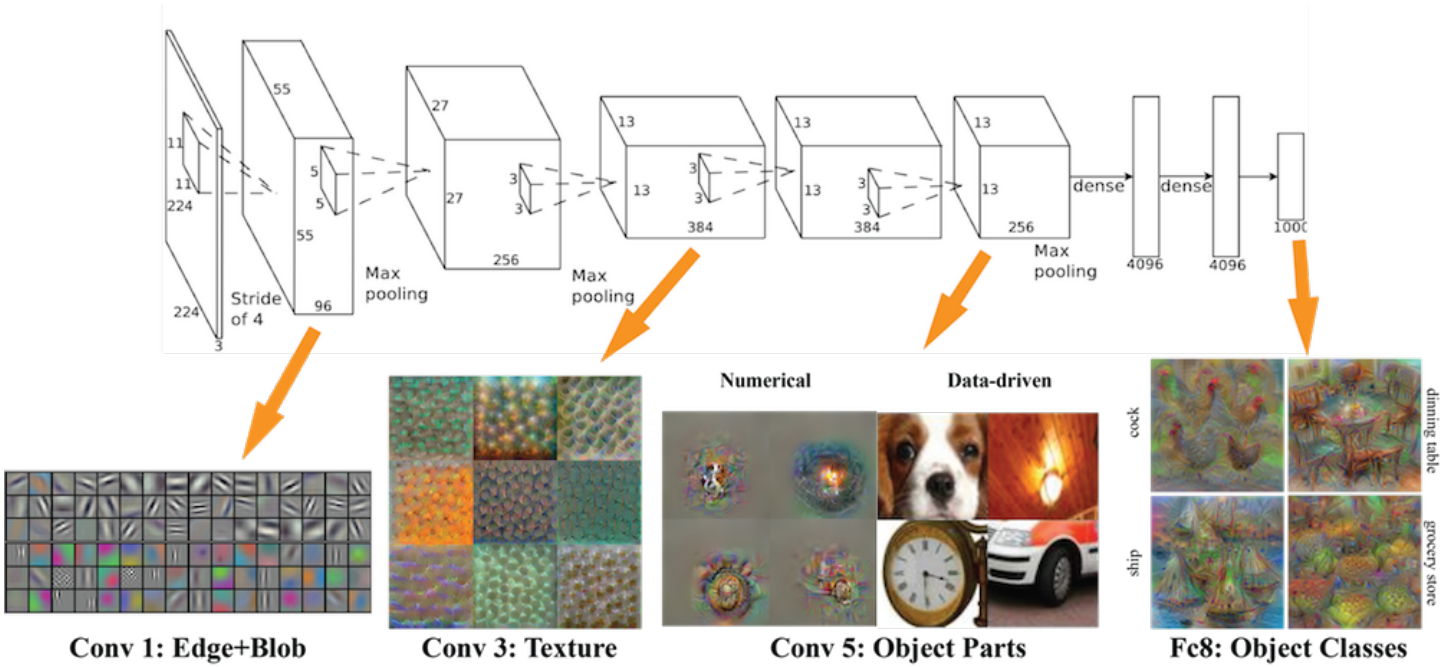
- DENDRAL, Fiegenbaum, Buchanan (Stanford) – a knowledge based system performs reasoning about molecular structures of organic compounds based on mass spectroscopy – 450 rules
- MYCIN, Shortliffe (Stanford), 550 rules, different from DENDRAL: no theoretical model as a foundation, introduces the "certainty factors"
- Advances in natural language processing
- PROLOG – logical programming language popular in Europe
- 1975, Minsky: frame theory

1980 – 2010

- 1980 – AI becomes an industry! (from several million dollars in 1980 up to a billion dollars in 1988)
- 1982 McDermott – DEC R1 expert system
- 1980 – Comeback of neural networks (Werbos – *backpropagation* algorithms)
- Intelligent agents (agent – perception of the environment through sensors and acting on it through actions)
- Robotics
- Machine learning

2010 – today

- The era of deep learning
- Deep learning – machine learning of multilayered data abstractions
- Typically using neural networks on large amounts of data
- Stunning advances in computer vision, promising improvements in natural language processing

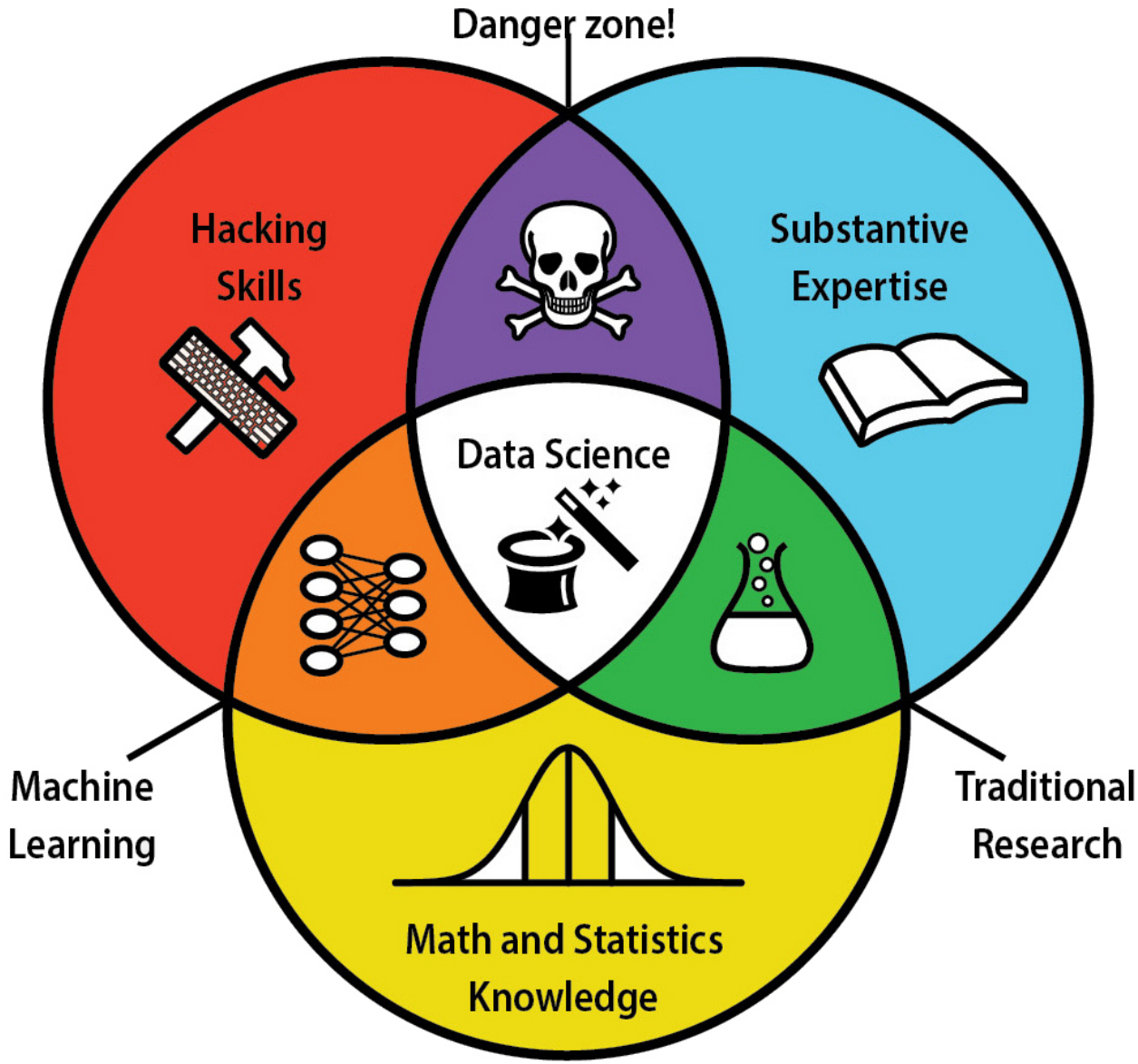


Deep learning: counting calories (Google)

http://www.popsci.com/google-using-ai-count-calories-food-photos

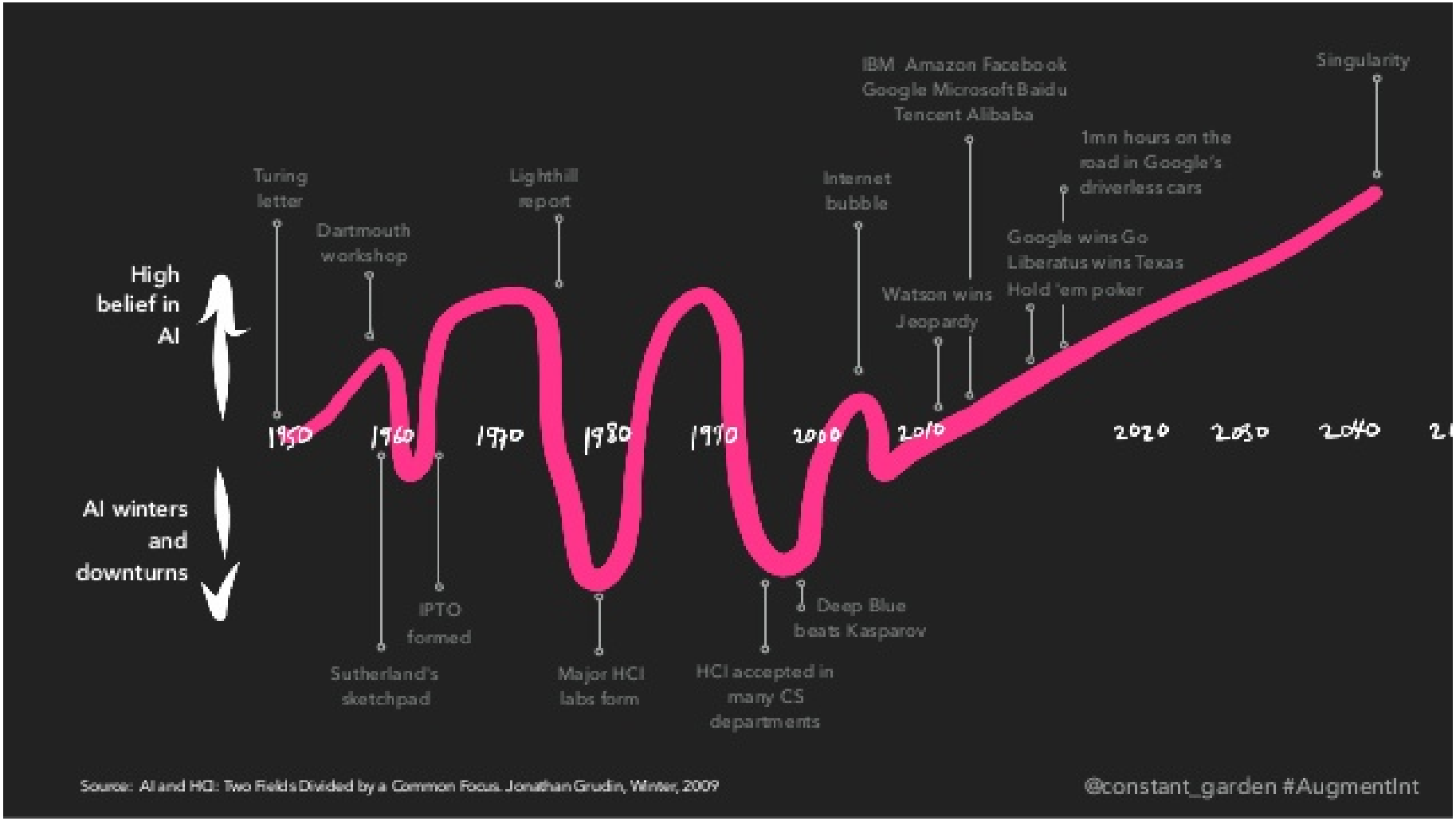
- A deep learning system estimates the calories based on dish photo

Data Science

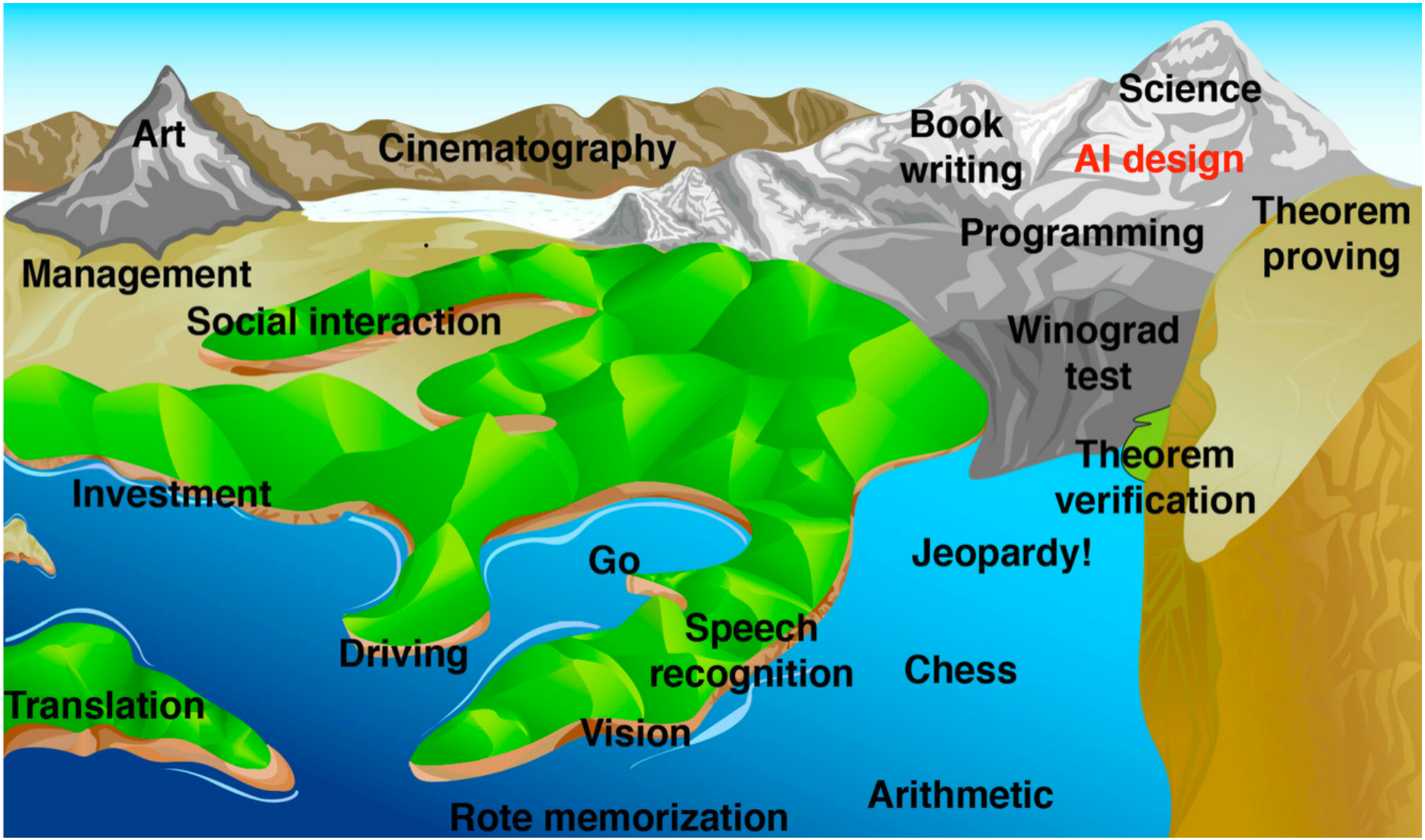


http://berkeleysciencereview.com/how-to-become-a-data-scientist-before-you-graduate/

Rough AI winters



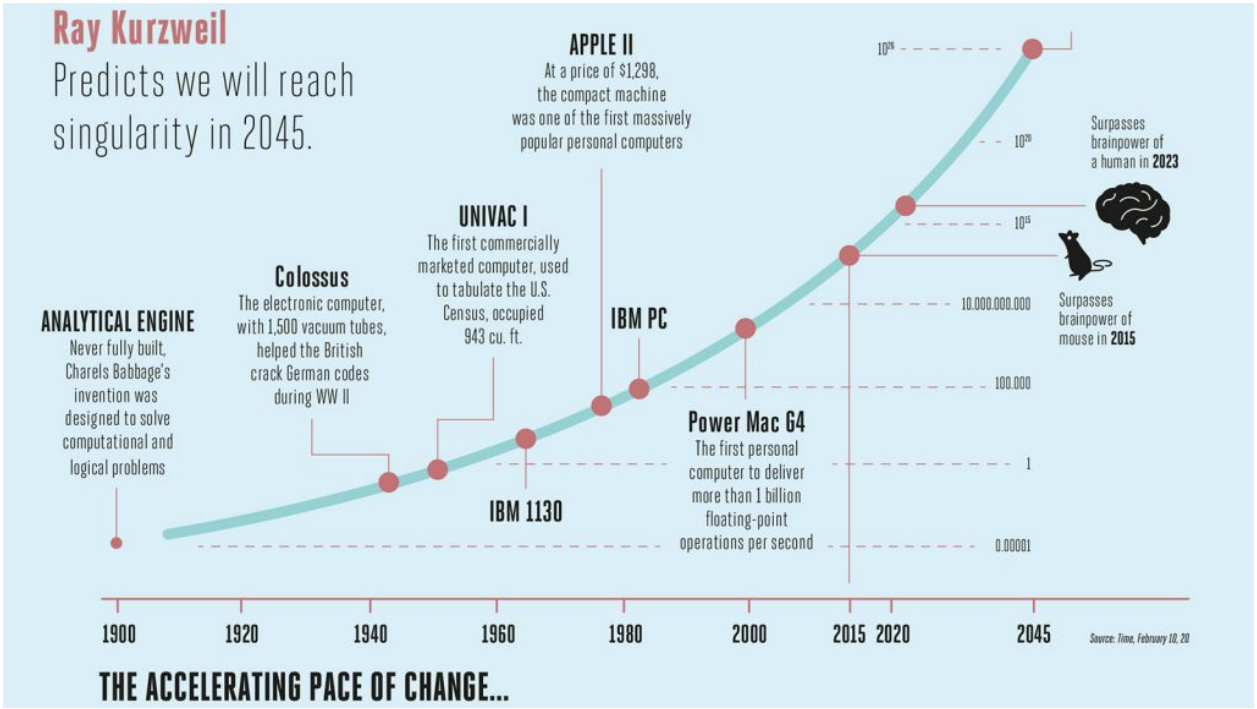
What's coming tomorrow?



Hans Moravec: Landscape of human competence

Main issues

1. Shortterm questions: safety, laws, weapons, jobs
2. Mid/longterm questions: AGI, superintelligence, singularity?
3. AI and consciousness (strong AI)?



. . . we'll get back to these questions in the very last class

Wrap-up

- Attempts to construct an intelligent machine reach far into the past
- Many tasks are simple for humans but hard for computers. We call the very difficult tasks **AI-complete**
- There is no consensus on the definition of AI, but we can identify four basic types of definitions **(acting/reasoning & rationally/human)**
- **Turing test** measures the intelligence of a machine through an imitation game. The test is interesting but of less practical importance.
- Throughout history, AI has seen good and bad times. Early extravagant ambitions generally remain unfulfilled.
- Today, computers can successfully (and sometimes better than humans) **solve many specific problems**

Next topic: State space search