

# Introduction to Artificial Intelligence

## 1. Introduction to Artificial Intelligence

Prof. Jan Šnajder

Assoc. Prof. Marko Čupić

Prof. Bojana Dalbelo Bašić

University of Zagreb  
Faculty of Electrical Engineering and Computing

Academic Year 2022/2023

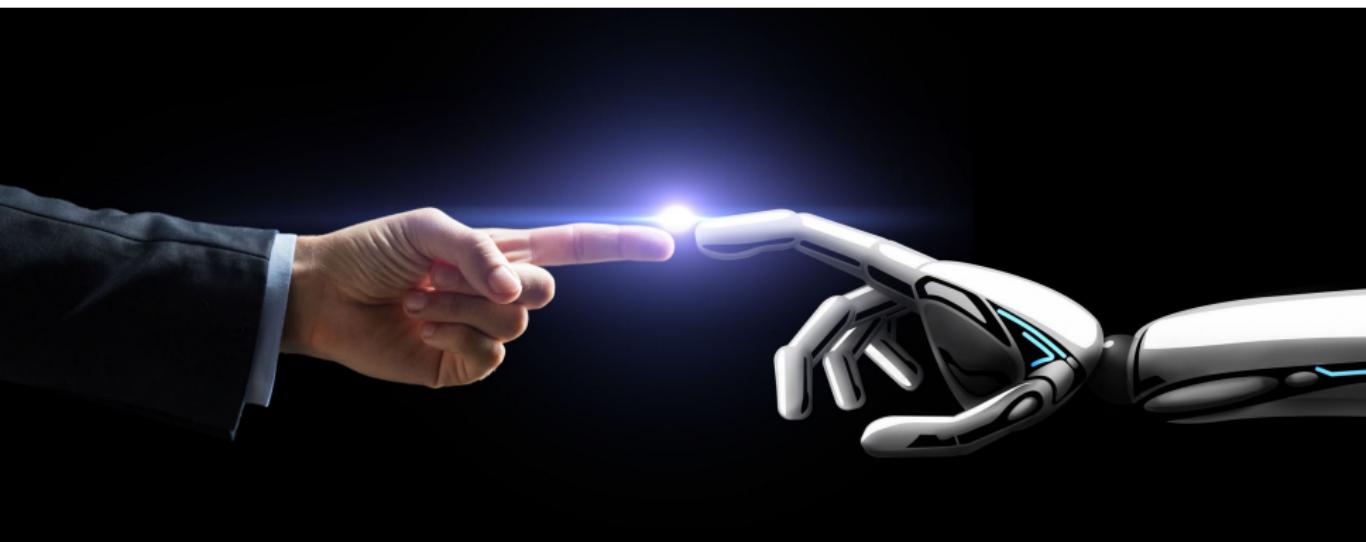


Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

v3.0

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”



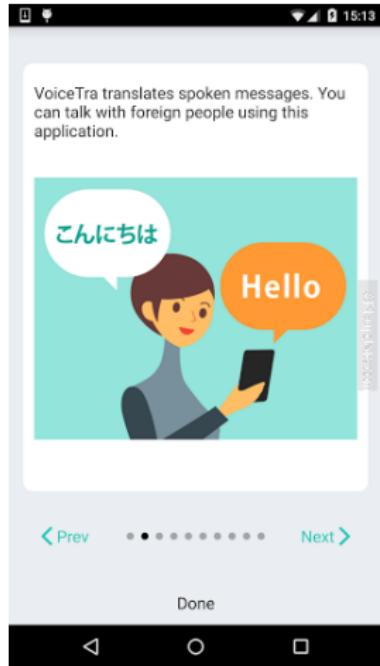
<https://www.youtube.com/watch?v=rVlhMGQgDkY>

# VoiceTra Real-Time Machine Translation

- Machine translation is the product of over 60 years of research
- Now entering its prime thanks to advances in machine learning and NLP
- 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

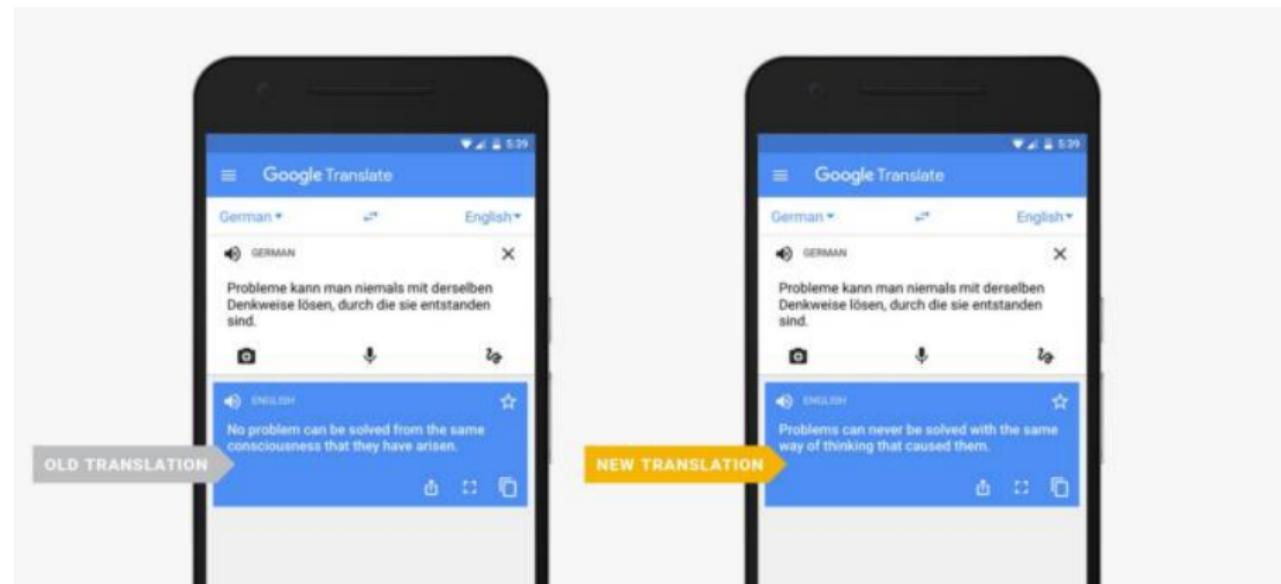
- Oct 2015: AlphaGo wins 5:0 against European champion Fan Hui
- Mar 2016: AlphaGo wins 4:1 against world champion Lee Sedol
- Two deep neural networks (DNN) trained to predict the next move and reduce the search space
- Reinforcement learning on top of DNN, to learn a playing strategy, trained by playing 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>

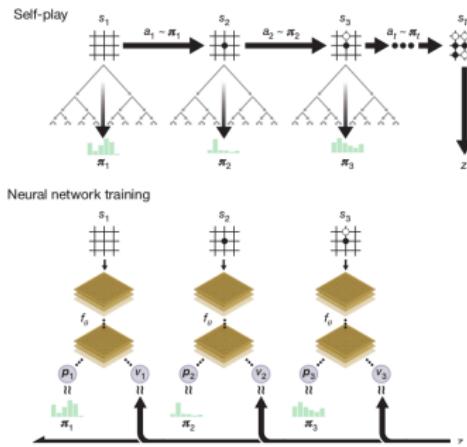
# Neural machine translation (NMT)

- 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](https://deepmind.com/documents/119/agz_unformatted_nature.pdf)

# ChatGPT



I'm teaching an intro to AI class next week. In my lecture, need to cover the latest breakthroughs in AI. What should I definitely cover? Give me three items.



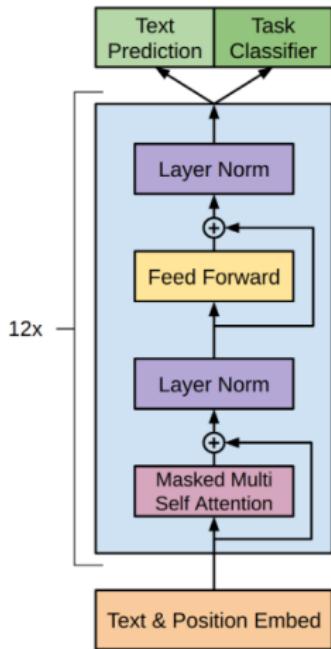
Sure, here are three latest breakthroughs in AI that you may want to cover in your lecture:



1. GPT-3: The Generative Pre-trained Transformer 3 (GPT-3) is a language model developed by OpenAI that has demonstrated impressive capabilities in natural language processing (NLP) tasks such as language translation, question answering, and text generation. It has 175 billion parameters, making it one of the largest and most powerful language models to date.
2. AlphaFold: AlphaFold is a deep learning model developed by Google's DeepMind that has revolutionized protein folding prediction. It was able to predict the 3D structure of proteins with remarkable accuracy, which has significant implications for drug discovery and understanding diseases.
3. Autonomous Driving: Autonomous driving has been a major focus of AI research in recent years. Breakthroughs in computer vision, deep learning, and sensor technology have enabled self-driving cars to navigate complex environments, making them safer and more reliable.

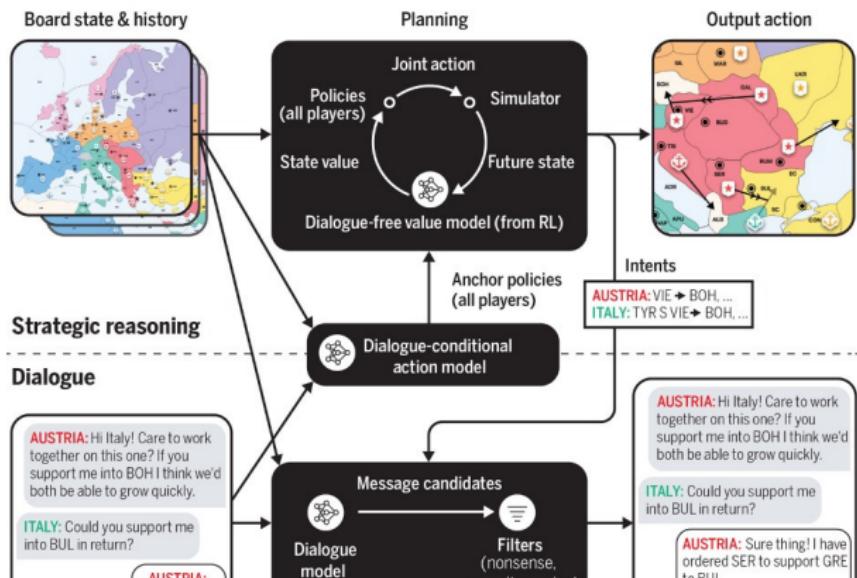
# ChatGPT

- AI chatbot launched in November 2022 by OpenAI
- Capable of chatting, answering questions, and creating text and code
- Based on a **large language model** (GPT) with a **transformer** trained on 375 billion of words ( $\sim$  4 million novels) and **reinforcement learning** with human feedback
- Demonstrates **emergent properties** (common-sense reasoning, zero/few-shot learning, chain-of-thought reasoning)
- Basis for conversational search (Bing)
- Potentially disruptive, raises philosophical and ethical questions



# Cicero

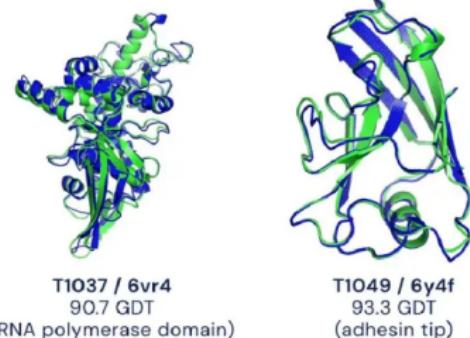
- AI system by Meta (November 2022) capable of human-level performance in the strategy game “Diplomacy”, which requires negotiation and tactical coordination between seven players
- Combines strategic reasoning and natural language processing (a neuro-symbolic approach to AI)



# AlphaFold

- A deep learning model developed by DeepMind for predicting 3D protein structures from its amino acid sequence
- Important for understanding the biological function of a protein and how to affect or modify it
- Two prediction modules, based on the transformer architecture
- November 2020: AlphaFold 2 wins the CASP14 competition (Critical Assessment of Structure Prediction), reducing the time for protein structure prediction from about 5 years down to seconds

II



- Experimental result
- Computational prediction

# What is this course about?



- ① An overview of **the fundamental AI methods and algorithms**
- ② **Advantages and shortcomings** of the different methods
- ③ The current **limitations and possibilities** of AI

# Today

- 1 Recent AI breakthroughs
- 2 Machines and us
- 3 Intelligence and artificial intelligence
- 4 A brief history of AI
- 5 Course overview

# Today

- 1 Recent AI breakthroughs
- 2 Machines and us
- 3 Intelligence and artificial intelligence
- 4 A brief history of AI
- 5 Course overview

# 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

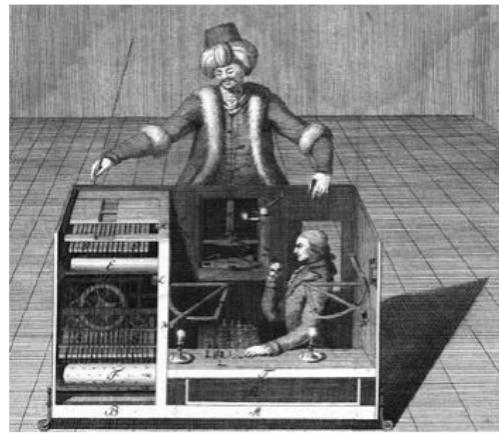
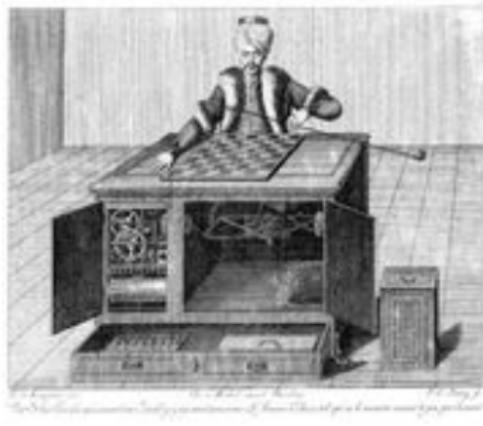
- Original story by Mary Shelley  
*“Frankenstein, or the modern Prometheus”*, published in 1818, describes an attempt by scientist Victor Frankenstein to create artificial life



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

The screenshot shows the homepage of the Amazon Mechanical Turk website. At the top, there's a navigation bar with links for 'Your Account', 'HITs', and 'Qualifications'. To the right, there are links for 'Already have an account?' and 'Sign in as a Worker | Requester'. Below the navigation, a banner states 'Mechanical Turk is a marketplace for work.' It explains that businesses and developers can access an on-demand, scalable workforce where workers select from thousands of tasks whenever it's convenient. A count of '389,697 HITs' available is shown with a link to 'View them now.'.

**Make Money**  
by working on HITs

HITs - Human Intelligence Tasks - are individual tasks that you work on. [Find HITs now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task → Work → Earn money

[Find HITs Now](#)

or [learn more about being a Worker](#)

**Get Results**  
from Mechanical Turk Workers

Ask workers to complete HITs - Human Intelligence Tasks - and get results using Mechanical Turk. [Register Now](#)

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results

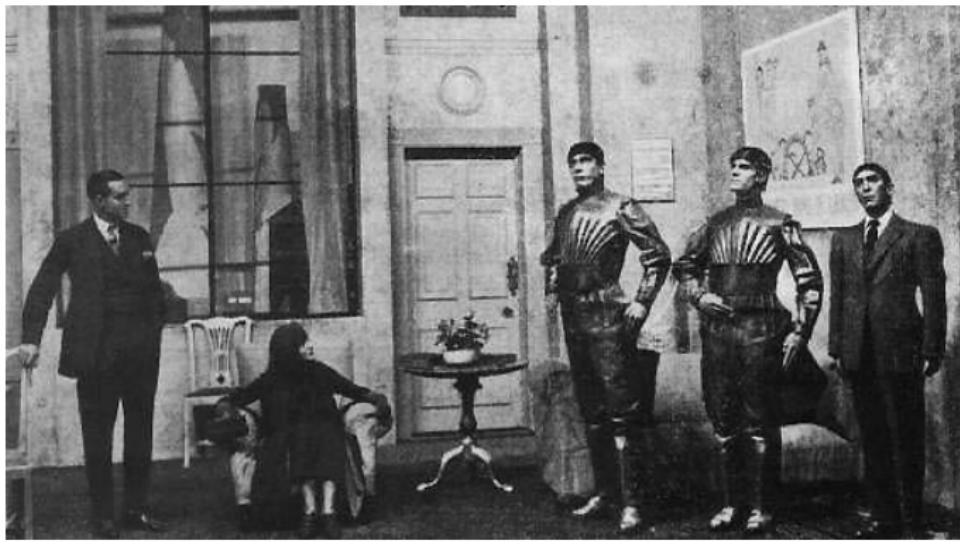
Fund your account → Load your tasks → Get results

[Get Started](#)

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

## Historical attempts: Robot

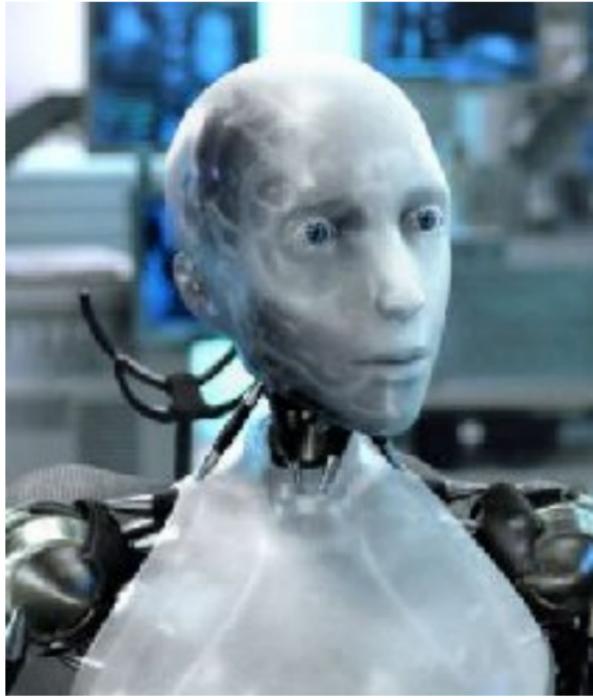
- In 1921 Czech writer Karel Čapek wrote the play *R. U. R. (Rossum's Universal Robots)*
- *Robot* (Czech *roboř*) – labour, forced labour



# Isaac Asimov: “I, robot”, 1942

Three Robot Laws:

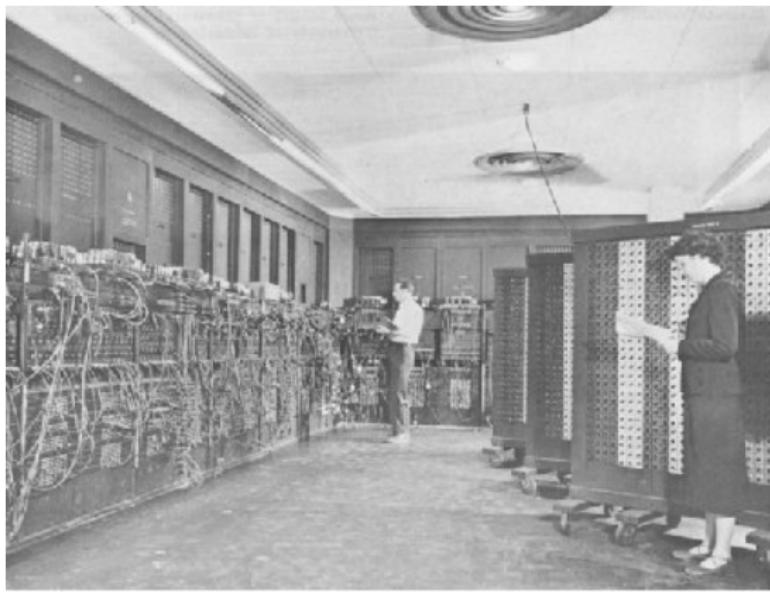
- ① A robot may not injure a human being or, through inaction, allow a human being to come to harm
- ② A robot must obey any orders given to it by human beings, except where such orders would conflict with the First Law
- ③ A robot must protect its own existence as long as such protection does not conflict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



# So, can machines think?

- We don't know, but we know we can build machines that **perform tasks that require intelligence**: control complex processes, solve complex problems, make decisions, reason, use natural language, recognize visual objects ...



*Rodney Brooks i robot Cog, MIT Media Lab*

# 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*?

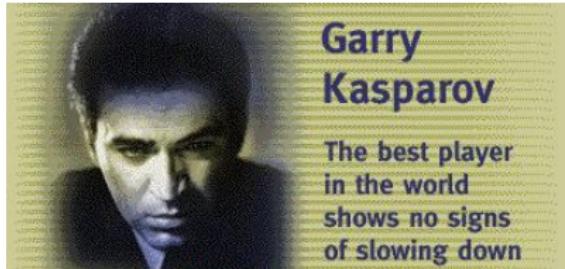


# Deep Blue vs. Garry Kasparov (2)



**Deep Blue**

This 1.4 ton  
8-year-old sure  
plays a mean  
game of chess



**Garry  
Kasparov**

The best player  
in the world  
shows no signs  
of slowing down

**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

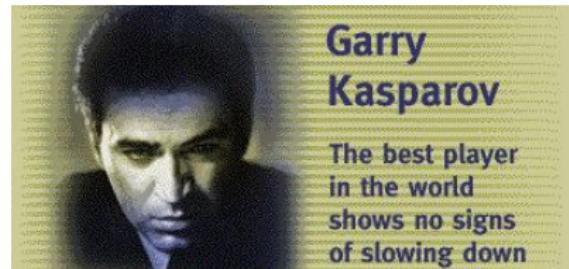
Has feelings and **brilliant intuition**, but can experience **fatigue and boredom** and loss of concentration

# Deep Blue vs. Garry Kasparov (3)



**Deep Blue**

This 1.4 ton  
8-year-old sure  
plays a mean  
game of chess



**Garry  
Kasparov**

The best player  
in the world  
shows no signs  
of slowing down

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

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

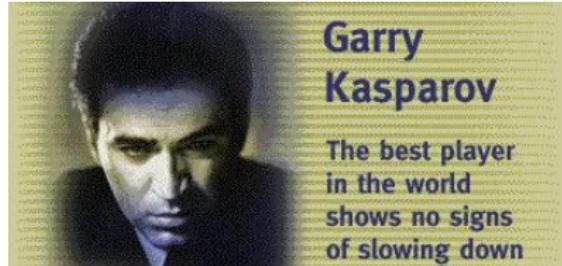
Garry Kasparov is **generally very intelligent**: he authored several books and speaks many languages

# Deep Blue vs. Garry Kasparov (4)



**Deep Blue**

This 1.4 ton  
8-year-old sure  
plays a mean  
game of chess



**Garry  
Kasparov**

The best player  
in the world  
shows no signs  
of slowing down

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

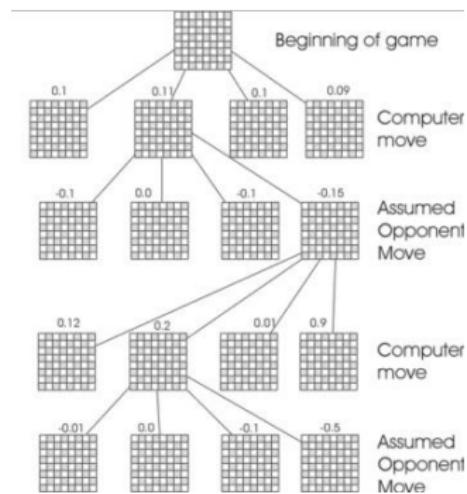
Garry Kasparov is capable of performing a **selective search** to determine the next move

# What's the problem?

- 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”)



# Combinatorial explosion is everywhere

- 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 **natural language sentences**

## 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 → Spanish → 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

The screenshot shows the Wolfram Alpha search interface. At the top, there is a red star logo followed by the text "WolframAlpha™ computational... knowledge engine". Below the logo is a search bar containing the query "How do you feel today?". To the right of the search bar are icons for saving, printing, and sharing, along with links for "Examples" and "Random". Below the search bar, there is a message: "Assuming 'How do you feel today' is a phrase | Use as a question about Alpha instead". The main content area is divided into two sections: "Input interpretation:" which contains the query "How are you?", and "Result:" which contains the response "I am doing well, thank you.". At the bottom left, it says "Computed by Wolfram Mathematica". On the bottom right, there is a "Download page" button.

How do you feel today?

Assuming "How do you feel today" is a phrase | Use as a question about Alpha instead

Input interpretation:

How are you?

Result:

I am doing well, thank you.

Computed by **Wolfram Mathematica**

Download page

# 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

# Today

- 1 Recent AI breakthroughs
- 2 Machines and us
- 3 Intelligence and artificial intelligence
- 4 A brief history of AI
- 5 Course overview

# 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**

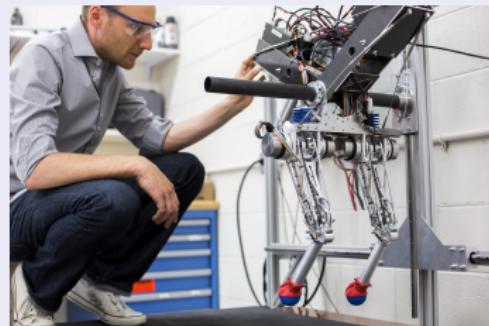
# A pragmatic (engineering-oriented) approach to intelligence

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

“Understanding by building”

*If a given kind of behavior (of a human, ant, elephant, robot, ...) is interesting, we ask:*

**How did it come to being?**



- Such an approach enables us to understand the fundamental principles of artificial intelligence

# What is *artificial* intelligence?

- A branch of computer science:  
**Technical Sciences → Computer Science → 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 Dartmouth 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
- Dartmouth 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”*

# Defining Artificial Intelligence (1)

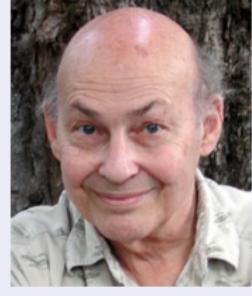
Patrick. H. Winston (MIT)

*"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)

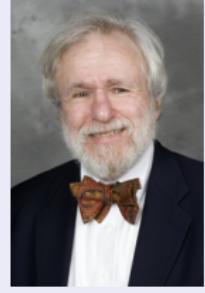
Elain Rich (University of Texas at Austin)

*"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."*



## Defining Artificial Intelligence (3)

D. W. Patterson (1990)

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

# How to test AI?

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."*



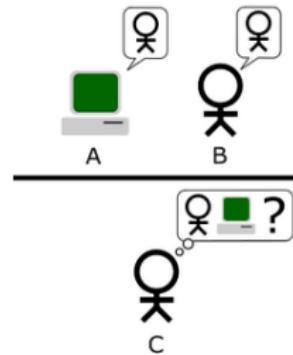
# Testing for intelligence

- Alan Turing ( “*Computing Machinery and Intelligence*”, *Mind* journal, 1950) 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

- Three players with different goals: A, B (who answer questions) and C (who asks questions). A is a machine, and B and C are humans
- **Goal of player C:**  
determine whether A or B is a machine
- **Goal of player B:**  
help player C in their task
- **Goal of player A:**  
trick player C into failing in their task
- Repeated several times and the success rate of player C is recorded
- Will player 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**



## Turing test – analysis

**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.

**Q:** Has any machine passed the TT?

- To date, no machine has (formally) passed the TT
- Early attempts: the controversial “Loebner prize”, held since 1990
- Recent LLMs (LaMDA, ChatGPT, etc.) come quite close

But, is it the TT valid?

# 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."*



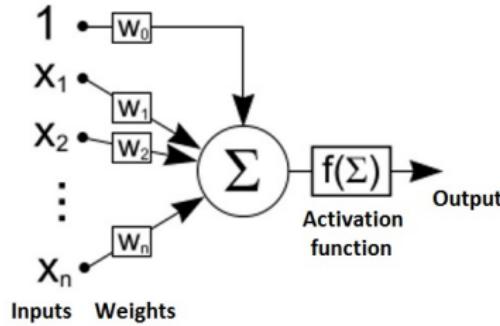
- TT is more significant for philosophy than for the development of AI

# Today

- 1 Recent AI breakthroughs
- 2 Machines and us
- 3 Intelligence and artificial intelligence
- 4 A brief history of AI
- 5 Course overview

## 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



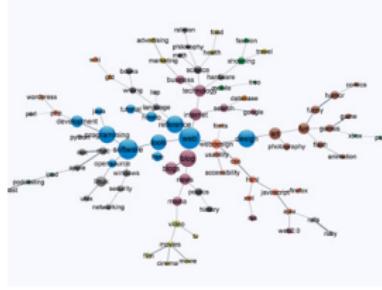
# 1952 – 1969: Early enthusiasm, high hopes (1)

- 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: Newell & 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



# 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"*  
→ *"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

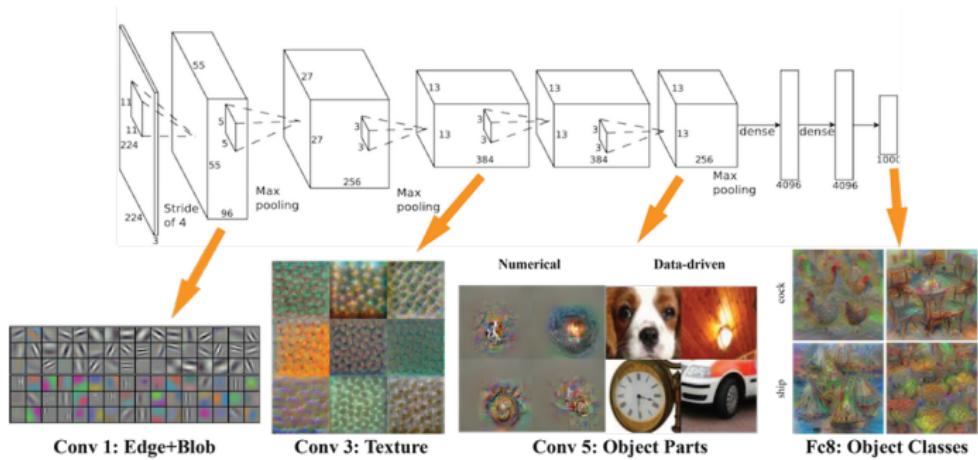
- DENDRAL, Fiegenbaum, Buchanan (Stanford) – a knowledge based system performs reasoning about molecular structures of organic compounds based on mass spectroscopy – 450 rules
- MYCIN, Shortliffe (Stanford), 550 rules, different from DENDRAL: no theoretical model as a foundation, introduces the “certainty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

## 1980 – 2010

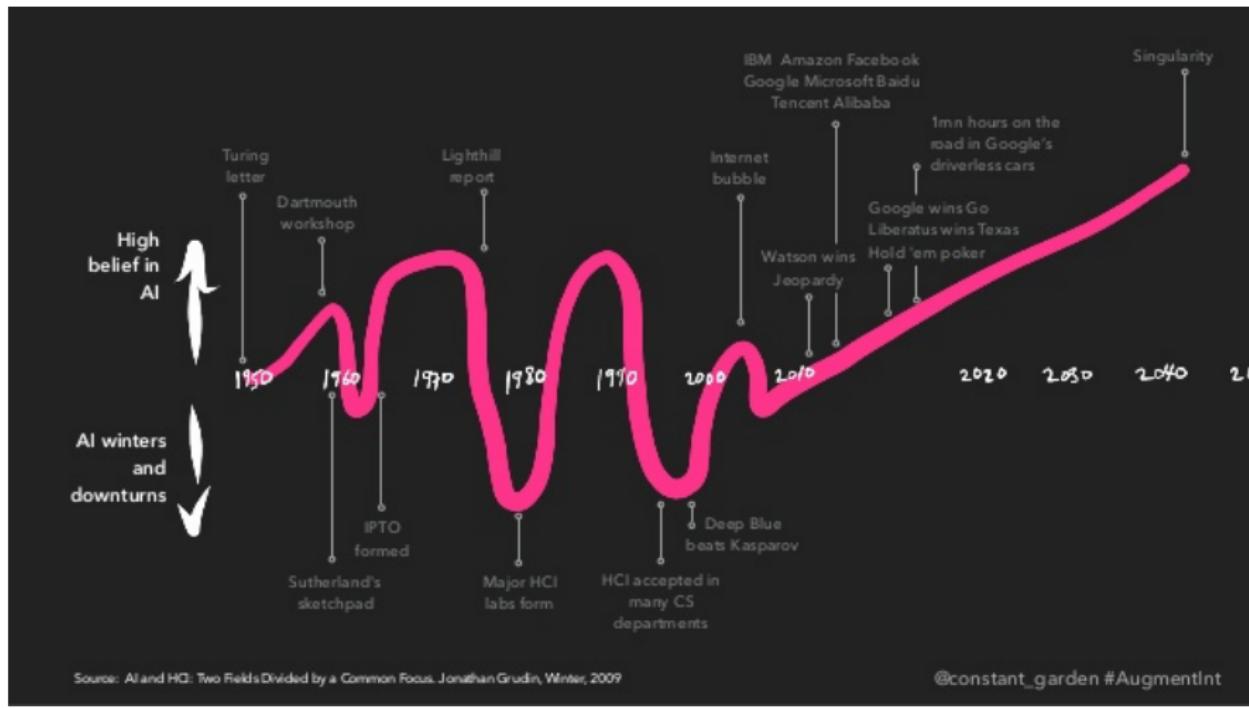
- 1980 – AI becomes an industry! (from several million dollars in 1980 up to a billion dollars in 1988)
- 1982 McDermott – DEC R1 expert system
- 1980 – Comeback of neural networks (Werbos – *backpropagation* algorithms)
- Intelligent agents (agent – perception of the environment through sensors and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



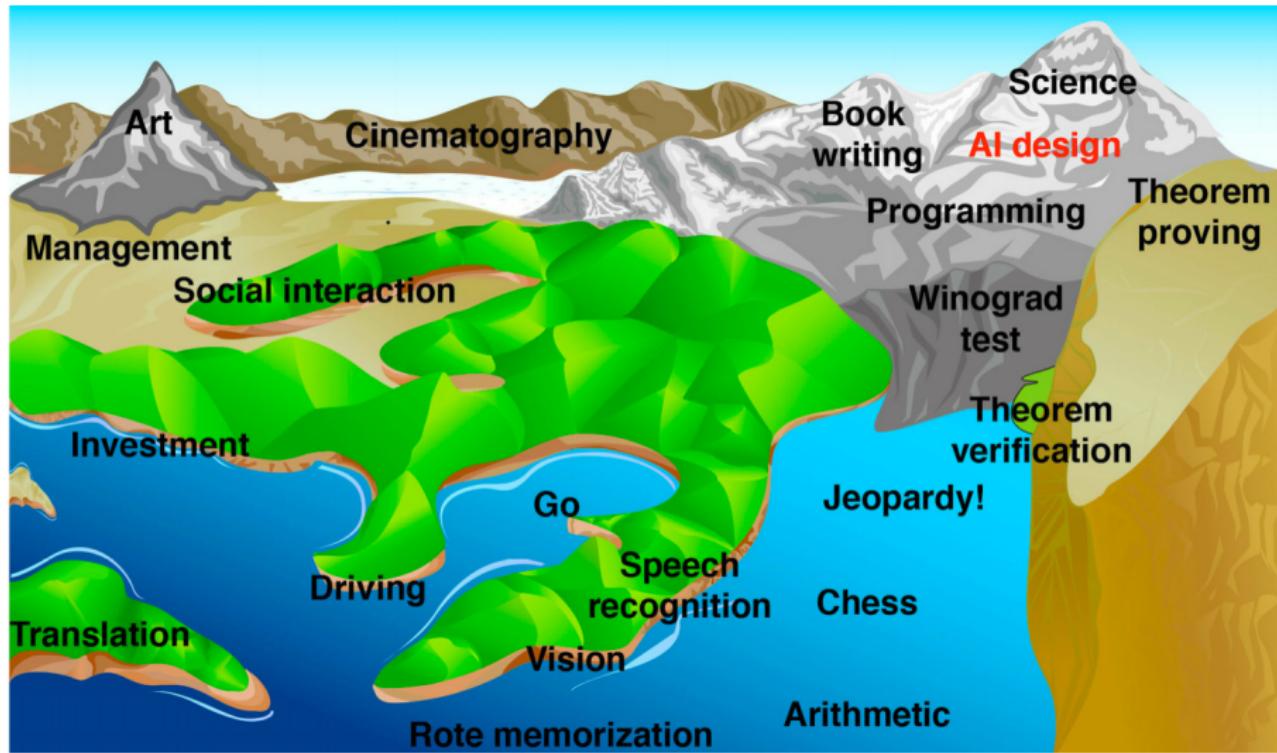
# Rough AI winters



Source: AI and HQ: Two Fields Divided by a Common Focus. Jonathan Grudin, Winter, 2009

@constant\_garden #AugmentInt

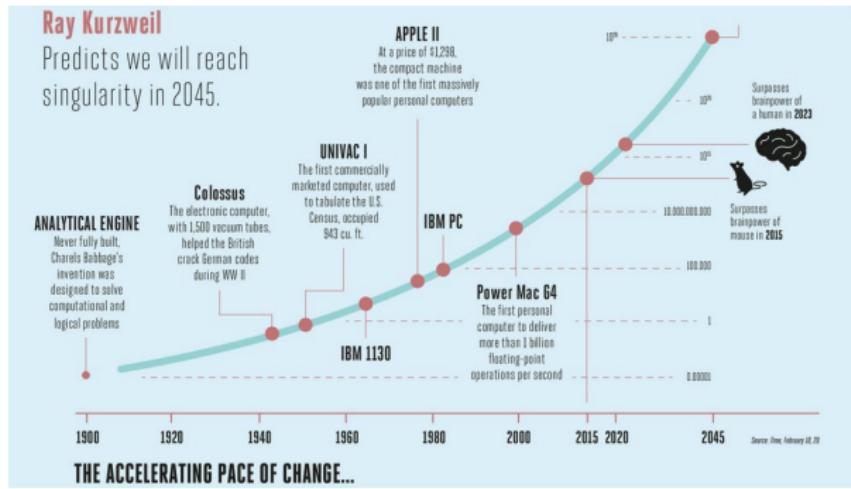
# What's coming tomorrow?



Hans Moravec: Landscape of human competence

# Main issues

- ① Short-term questions: safety, laws, weapons, jobs
- ② Mid-/long-term questions: AGI, superintelligence, singularity?
- ③ AI and consciousness (strong AI)?



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

# Today

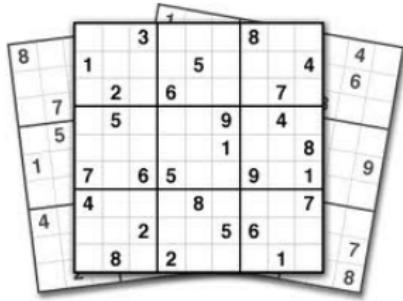
- 1 Recent AI breakthroughs
- 2 Machines and us
- 3 Intelligence and artificial intelligence
- 4 A brief history of AI
- 5 Course overview

# What is this course about?



- ① An overview of **the fundamental AI methods and algorithms**
- ② **Advantages and shortcomings** of the different methods
- ③ The current **limitations and possibilities** of AI

# State space search



# Heuristic search

initial state:

|   |   |   |
|---|---|---|
| 8 |   | 7 |
| 6 | 5 | 4 |
| 3 | 2 | 1 |

goal state:

|   |   |   |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 6 |
| 7 | 8 |   |

## What to use as heuristic function?

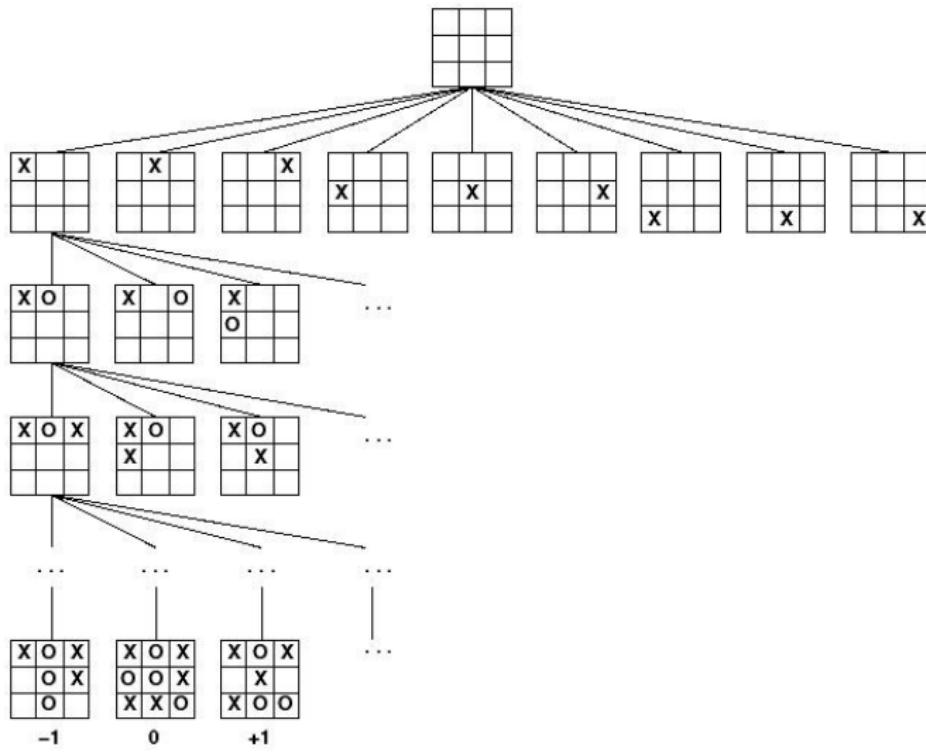
- Number of displaced squares:

$$h_1\left(\begin{smallmatrix} 8 & & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{smallmatrix}\right) = 7$$

- Sum of city-block (L1) distances between all squares and their final positions:

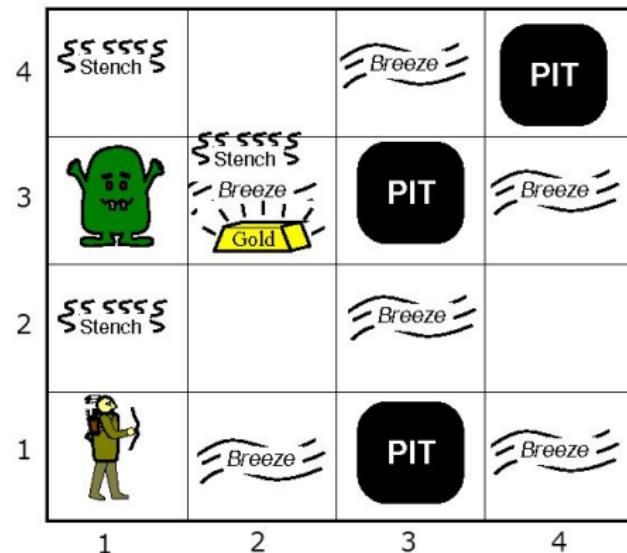
$$h_2\left(\begin{smallmatrix} 8 & & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{smallmatrix}\right) = 21$$

# Game playing



# Knowledge representation using formal Logic

## The Wumpus World



Percepts (facts):

$$\neg B_{1,1}$$

$$\neg B_{1,2}$$

$$B_{2,1}$$

$$\neg S_{1,1}$$

$$S_{1,2}$$

$$\neg S_{2,1}$$

$$\neg P_{1,1}$$

$$\neg W_{1,1}$$

$$\neg G_{1,1}$$

Knowledge (rules):

$$B_{2,1} \leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$$

$$S_{1,2} \leftrightarrow (W_{1,1} \vee W_{2,2} \vee W_{1,3})$$

⋮

# Automated reasoning

- A robot is delivering packages. The robot knows that all packages in room 27 are smaller than any of the packages in room 28. A and B are packages. Package A is in room 27 or 28, but the robot doesn't know where exactly. Package B is in room 27 and is not smaller than package A.
- Use refutation resolution to show how the robot can conclude that package A is in room 27
- Knowledge representation:



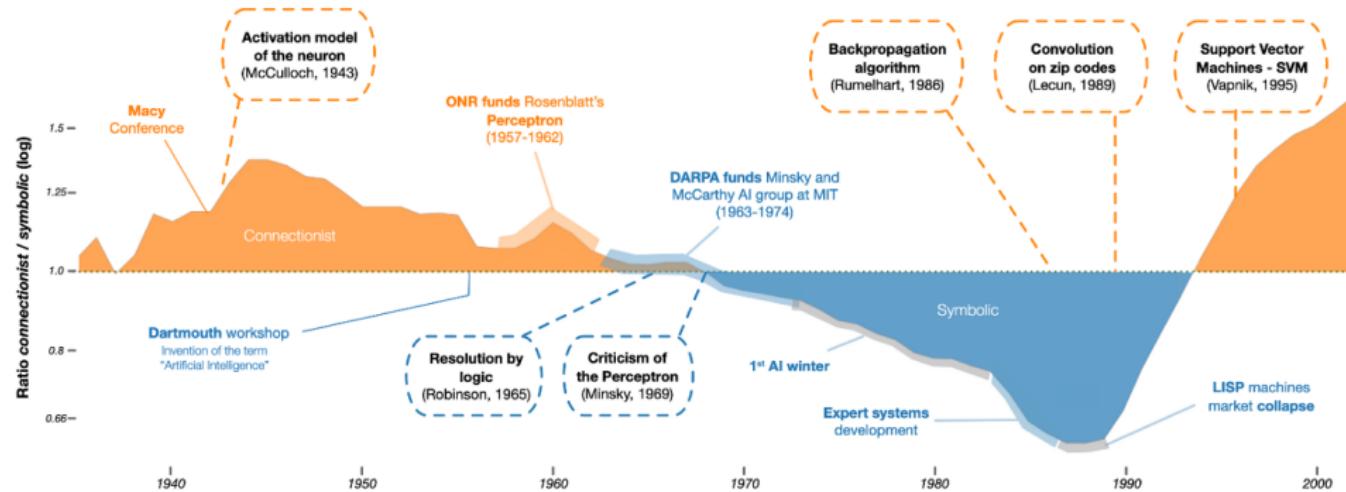
- (1)  $\forall x \forall y ((P(x) \wedge P(y) \wedge I(x, 27) \wedge I(y, 28)) \rightarrow S(x, y))$
  - (2)  $P(A) \wedge P(B)$
  - (3)  $I(A, 27) \vee I(A, 28)$
  - (4)  $I(B, 27) \wedge \neg S(B, A)$
- $\vdash I(A, 27)$

# Logic programming

The screenshot shows the SWISH Prolog IDE interface. The top menu bar includes File, Edit, Examples, and Help. The left pane displays a Prolog source code file named 'houses\_puzzle' containing simple test programs and knowledge bases. The right pane shows the execution trace for the query `?- trace, (jealous(X, Y)).`. The trace output is as follows:

```
trace, (jealous(X, Y)).  
Call: jealous(_3972, _3976)  
Call: loves(_3972, _4402)  
Exit: loves(vincent, mia)  
Call: loves(_3976, mia)  
Exit: loves(vincent, mia)  
Exit: jealous(vincent, vincent)  
X = Y, Y = vincent  
Redo: loves(_3976, mia)  
Exit: loves(marcellus, mia)  
Exit: jealous(vincent, marcellus)  
X = vincent,  
Y = marcellus  
Exit: loves(marcellus, mia)  
Call: loves(_3976, mia)  
Exit: loves(vincent, mia)  
Exit: jealous(marcellus, vincent)  
X = marcellus,  
Y = vincent  
Next 10 100 1,000 Stop  
?- trace, (jealous(X, Y)).
```

# Symbolism vs. connectionism



Cardon, D., Cointet, J. P., & Mazieres, A. (2018). **Neurons spike back: The Invention of Inductive Machines and the Artificial Intelligence Controversy.**

# Modeling uncertainty

$p$  = “Today is a sunny day”



P is  
true



P is  
false



?



?



?



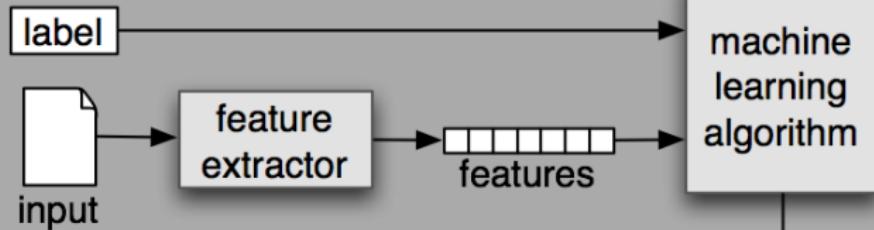
?



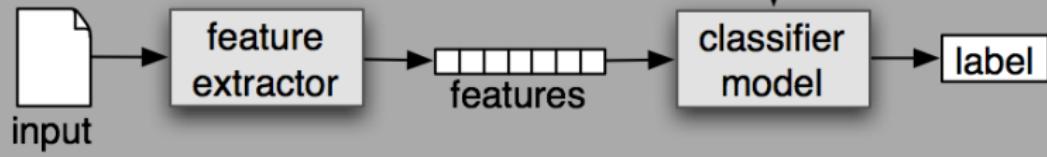
?

# Machine learning

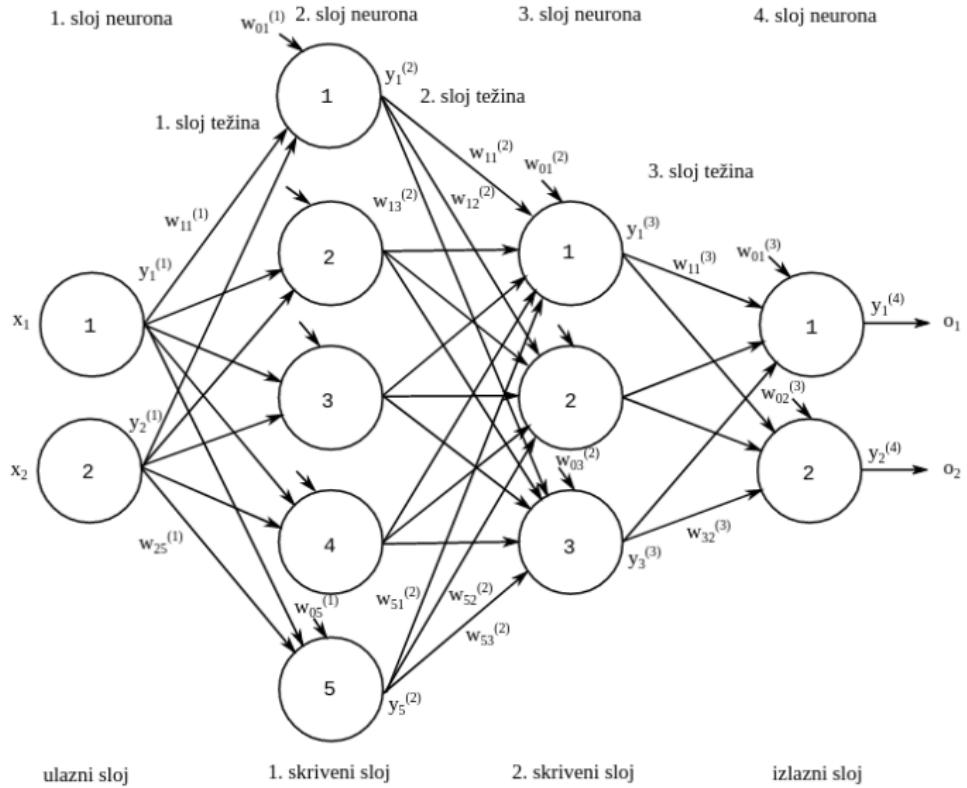
## (a) Training



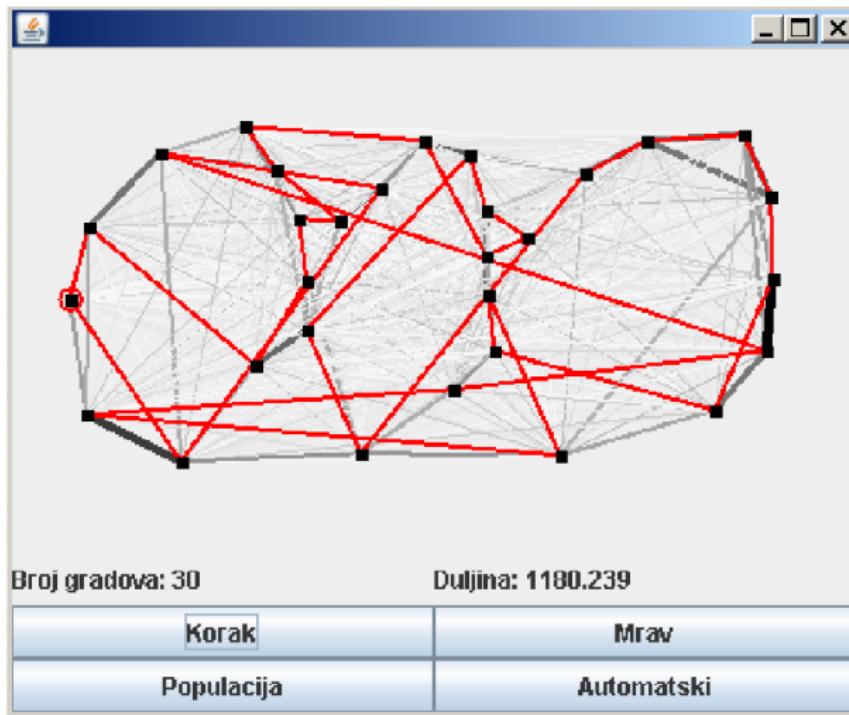
## (b) Prediction



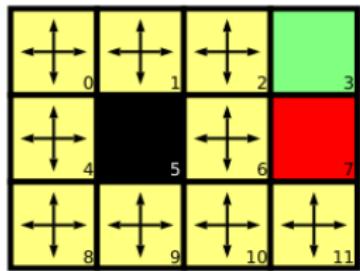
# Artificial neural networks



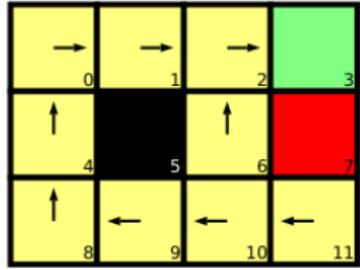
# Nature-inspired optimization algorithms



# Reinforcement learning



|        |        |        |        |
|--------|--------|--------|--------|
| -0.038 | 0.089  | 0.215  | 0.000  |
| 0      | 1      | 2      | 3      |
| -0.165 |        | -0.443 | 0.000  |
| 4      | 5      | 6      | 7      |
| -0.291 | -0.418 | -0.544 | -0.772 |
| 8      | 9      | 10     | 11     |



|       |       |       |       |
|-------|-------|-------|-------|
| 0.973 | 0.973 | 0.973 | 0.000 |
| 0     | 1     | 2     | 3     |
| 0.973 |       | 0.753 | 0.000 |
| 4     | 5     | 6     | 7     |
| 0.973 | 0.973 | 0.948 | 0.732 |
| 8     | 9     | 10    | 11    |

## 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 adopt a pragmatic “**understanding by building**” approach
- **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*