

Uvod u umjetnu inteligenciju

1. Uvod u umjetnu inteligenciju

prof. dr. sc. Jan Šnajder
izv. prof. dr. sc. Marko Čupić

Sveučilište u Zagrebu
Fakultet elektrotehnike i računarstva

Ak. god. 2023./2024.

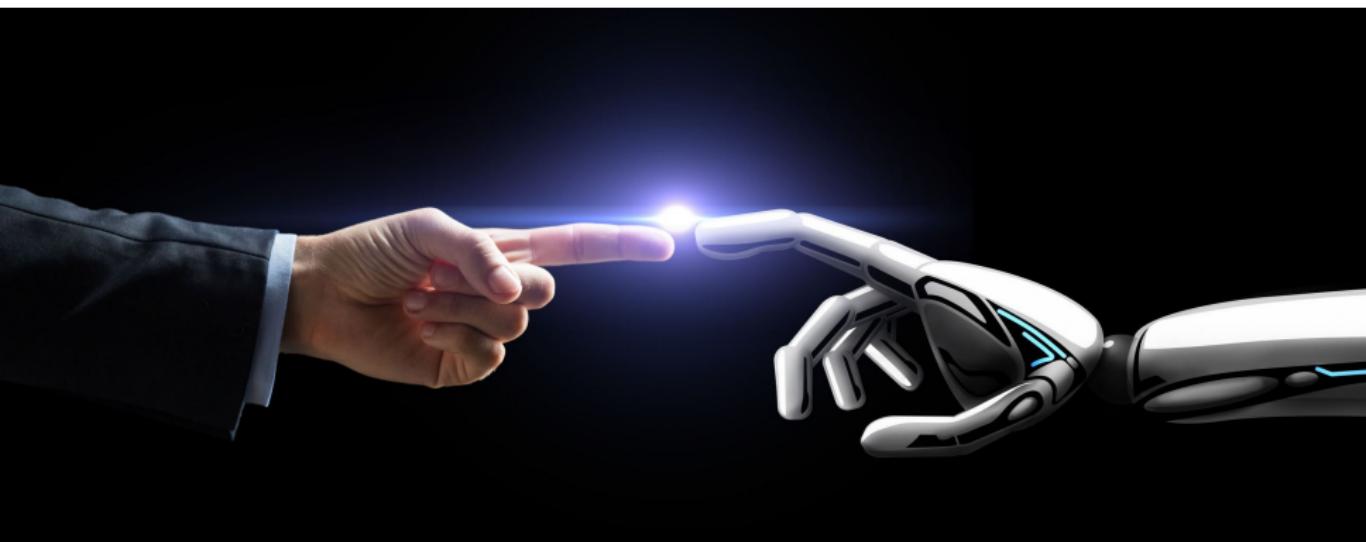


Creative Commons Imenovanje–Nekomercijalno–Bez prerada 4.0

v3.1

Živimo u zanimljivim vremenima za UI.

U posljednjih desetak godina, UI je znatno napredovala.



Pogledajmo neke od posljednjih vrhunaca...

Atlas: humanoidni robot

- Visoko pokretljiv humanoidni robot oblikovan za rad na otvorenom (neravnom terenu)
- Oblikovan za zadatke potrage i spašavanja
- Izgradila ga je tvrtka Boston Dynamics, finansirala DARPA, javnosti je predstavljen 2013. godine
- 28 hidraulički pokretanih stupnjeva slobode, dva dlana, ruke, noge, stopala i torzo
- Sensorska glava sa stereo-kamerama i laserskim mjeričem udaljenosti
- Proglašen primjerkom nove vrste, "robo sapiens"



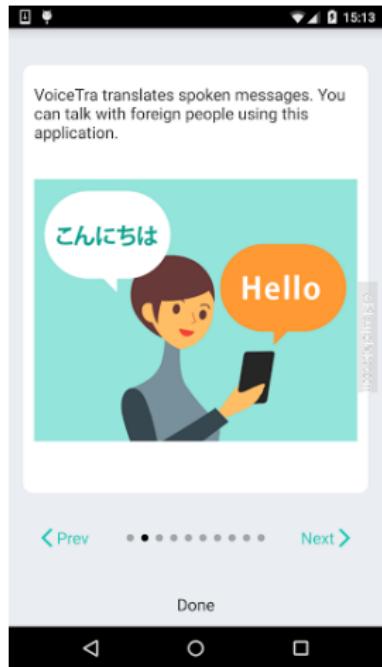
<https://www.youtube.com/watch?v=rVlhMGQgDkY>

VoiceTra strojno prevodenje u stvarnome vremenu

- Strojno prevodenje rezultat je 60-godišnjeg istraživanja
- Zrela tehnologija, zahvaljujući napretku u strojnem učenju i računarstvu u oblaku
- Japan razvija sustave za visokokvalitetno prevodenje iz govora u govor, u stvarnome vremenu
- Lansiranje prije Olimpijade 2020. godine u Tokiju, kao pomoć posjetiteljima
- Trenutačno pokriva 27 jezika u tekstu i četiri jezika u govoru

The Japan Times:

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



DeepMinds's AlphaGo: Go je pao

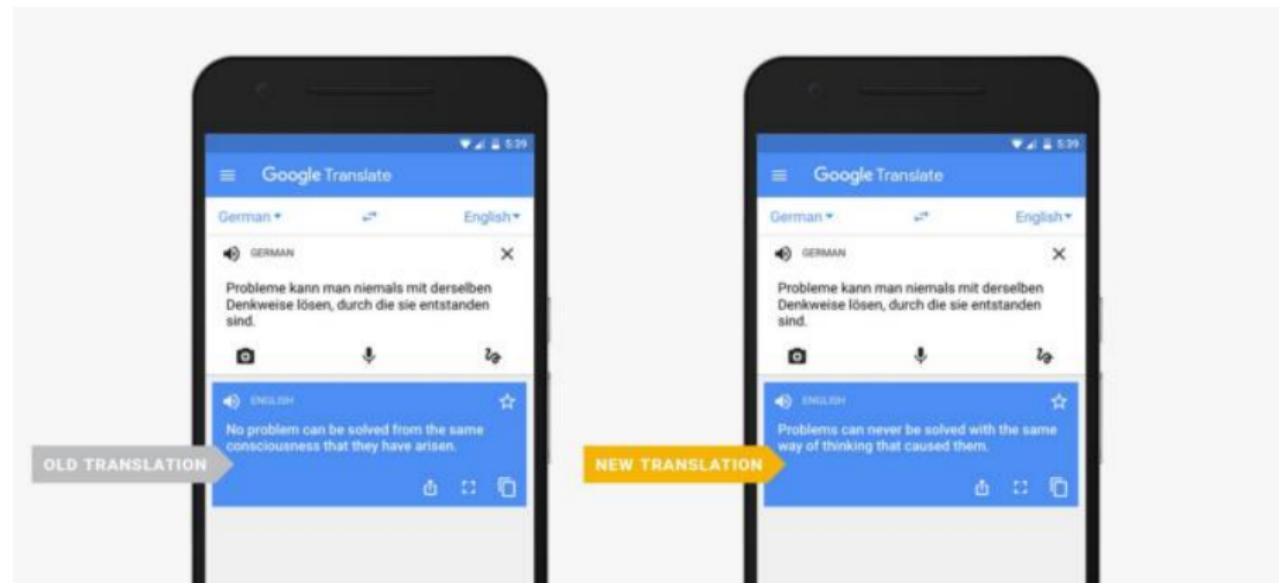
- Listopad 2015.: AlphaGo pobjeđuje europskog prvaka Fan Huia 5:0
- Ožujak 2016.: AlphaGo pobjeđuje svjetskog prvaka Lee Sedola 4:1
- Dvije duboke neuronske mreže (DNN), trenirane da predviđaju idući potez i smanje prostora pretraživanja
- Pojačano učenje kao nadogradnja na DNN, za učenje strategije igranja, treniranjem sustava da igra sam protiv sebe
- Za igranje koristi 40 dretvi, 48 CPU-a, i 8 GPU-a. Za treniranje koristi 40 dretvi, 1202 CPU-a i 176 GPU-a



<https://www.youtube.com/watch?v=SUBqykXVx0A>

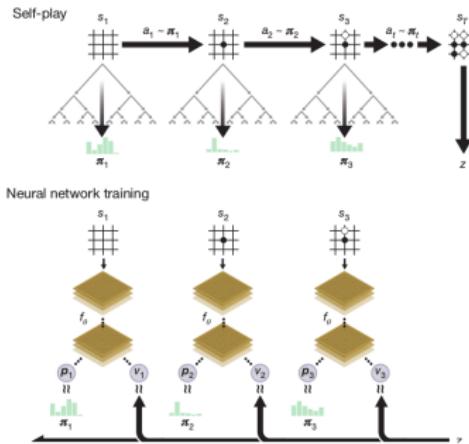
Neuronsko strojno prevodenje (NMT)

- Studeni 2016.: Google sustav za statističko strojno prevodenje (temeljen na frazama) zamjenjuje sustavom temeljenom na dubokim neuronskim mrežama (sequence-to-sequence learning)
- Duboka LSTM-mreža sa 8 slojeva kodera i dekodera
- 60%-tno smanjenje pogrešaka u prijevodu!



AlphaGo Zero: Učenje od nule

- Model učen samo podržanim učenjem u igrama protiv samoga sebe, krenuvši od slučajne strategije, bez ikakvog nadzora stručnjaka!
- Jednostavniji model od AlphaGo Fan/Lee (samo jedna mreža)
- Ostvaruje nadljudsku učinkovitost: 100:0 protiv AlphaGo Lee
- Primijenjen na drugim igrama: šah, shogi



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.



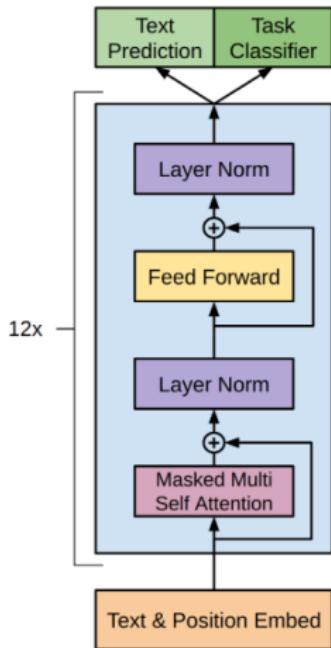
Sure, here are three latest breakthroughs in AI that you may want to cover in your lecture:



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.

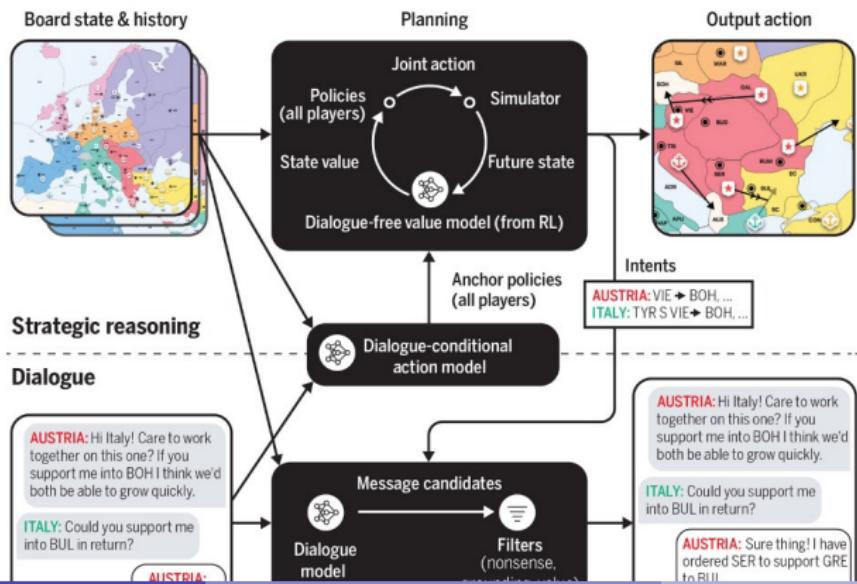
ChatGPT

- Razgovorni agent (chatbot) koji je u studenom 2022 predstavio OpenAI
- Razgovara, odgovara na pitanja te piše tekst i programski kod
- Temeljen na **velikom jezičnom modelu** (GPT) s **transformatorom** učenom na 375 milijardi riječi (~ 4 milijuna romana) i **podržanim učenjem** s povratnim signalom od čovjeka
- Pokazuje **izranjajuća svojstva** (zero/few-shot learning, zdravorazumsko i višekoračno zaključivanje)
- Temelj za razgovorno pretraživanje (Bing)
- Potencijalno remetilačka tehnologija, otvara filozofska i etička pitanja



Cicero

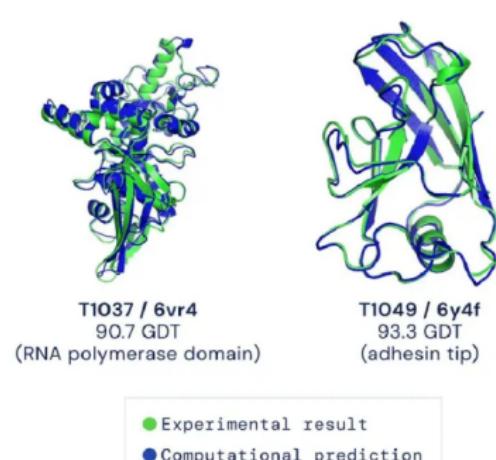
- UI sustav (Meta, studeni 2002) koji doseže ljudsku performancu u strateškoj igri "Diplomacy", koja iziskuje pregovaranje i taktičku koordinaciju između sedam igrača
- Kombinira strateško zaključivanje i obradu prirodnoga jezika (neuro-simbolički pristup)



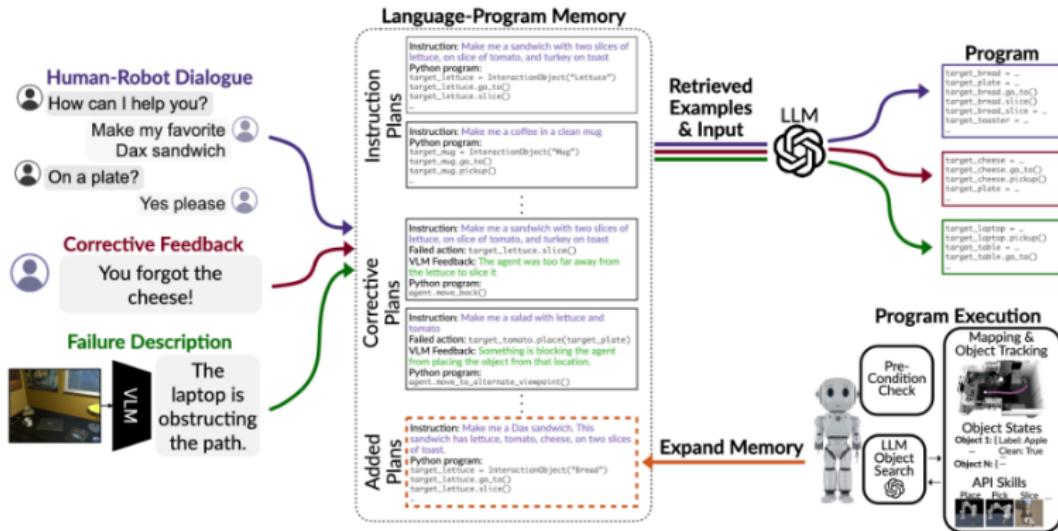
AlphaFold

- Model dubokog učenja koji je razvio DeepMind za predviđanje 3D proteinskih struktura iz njihovih proteinskih sljedova
- Ključno za razumijevanje biološke funkcije proteina (kako utjecati na protein ili ga izmijeniti)
- Dva modula za predviđanje, temeljena na arhitekturi transformatora
- Studeni 2020: AlphaFold 2 pobjeđuje na natjecanju CASP14 (Kritička procjena predviđanja strukture), smanjujući vrijeme za predviđanje strukture proteina s oko 5 godina na nekoliko sekundi

II



LLM-agenci u virtualnim/fizičkim okolinama



- <https://helper-agent-llm.github.io/>
- Sarch, G., Wu, Y., Tarr, M. J., & Fragkiadaki, K. (2023). Open-ended instructable embodied agents with memory-augmented large language models. EMNLP 2023 Findings

Čime se bavi ovaj predmet?



- ① Pregled **temeljnih postupaka i algoritama UI**
- ② **Prednosti i nedostatci** različitih metoda
- ③ Trenutačna **ograničenja i mogućnosti UI**

Danas

- 1 Nedavna postignuća umjetne inteligencije
- 2 Mi i strojevi
- 3 Inteligencija i umjetna inteligencija
- 4 Kratka povijest umjetne inteligencije
- 5 Pregled tema

Danas

1 Nedavna postignuća umjetne inteligencije

2 Mi i strojevi

3 Inteligencija i umjetna inteligencija

4 Kratka povijest umjetne inteligencije

5 Pregled tema

Mogu li strojevi misliti?

- Tisuće godina staro pitanje:
“Kako mi mislimo?”
- Zajedno s nastankom računala nastalo je i
uvjerenje da ćemo **inteligenciju reproducirati
računalom**
- Što je to uopće **inteligencija**? Što
podrazumijevamo pod pojmom **umjetna
inteligencija**?



I prije nastanka računala postojali su **pokušaji da napravimo svoju kopiju** . . .

Povijesni pokušaji: Frankenstein

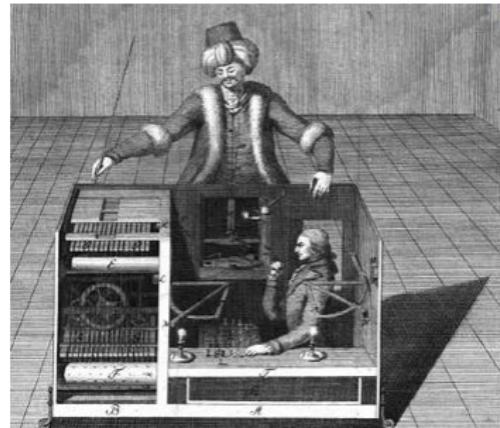
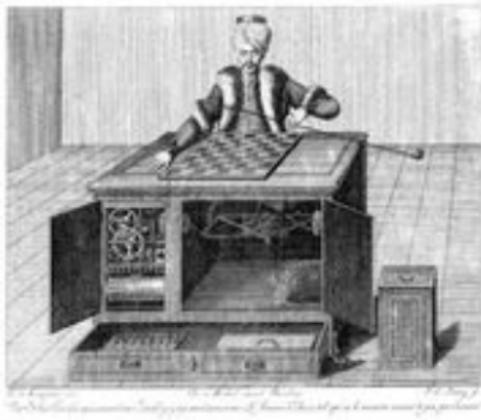
- Izvorna priča Mary Shelley
“Frankenstein, ili moderni Prometej”,
izdana 1818. godine, opisuje pokušaj
znanstvenika Victora Frankensteina da
stvori umjetan život



B. Wrightson: Frankenstein creates the fiend

Povijesni pokušaji: Turčin

- 1770. godine Wolfgang von Kempelen konstruirao je automat koji igra šah i izvodi konjićev obilazak
- Automat je izlagan i demonstriran 80 godina po Europi i Americi
- Vješt konstruirana mehanička iluzionistička naprava



Danas: 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 yellow bar at the bottom indicates '389,697 HITs available. [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

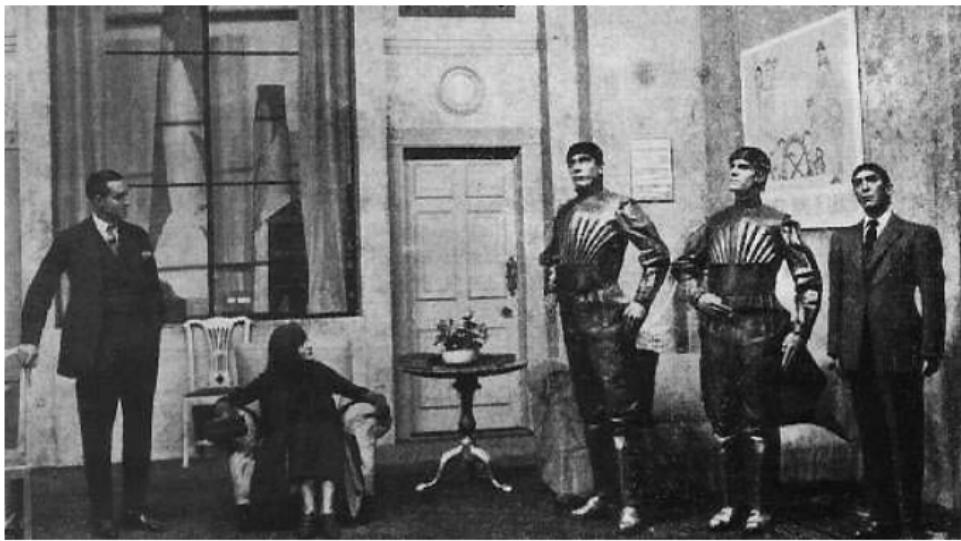
Find your account **Load your tasks** **Get results**

[Get Started](#)

- Velik broj ljudi plaćen da obavlja HIT-ove (*Human Intelligence Tasks*)
 - zadatke koji iziskuju ljudsku inteligenciju
- “Artificial Artificial Intelligence”, *crowdsourcing*

Povijesni pokušaji: Robot

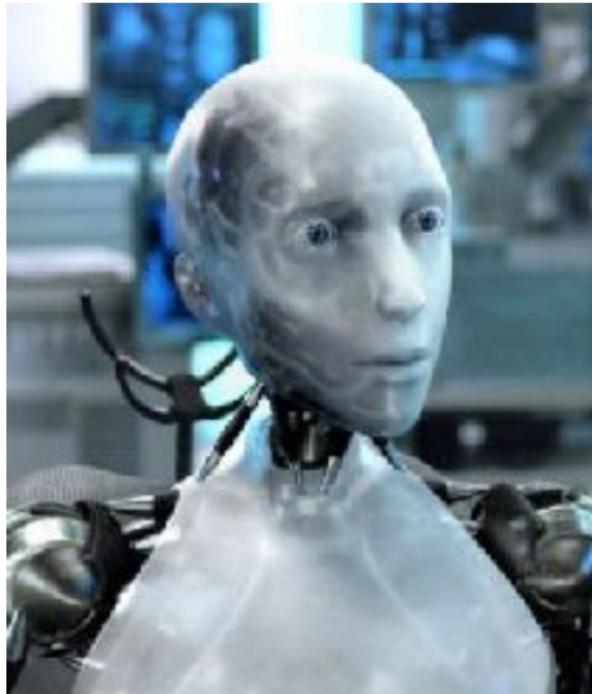
- Godine 1921. češki pisac Karel Čapek napisao je dramu *R. U. R.* (*Rossum's Universal Robots*)
- *Robot* (češki *robova*) – rad, prisilan rad



Isaac Asimov: "Ja, robot", 1942.

Three Robot Laws:

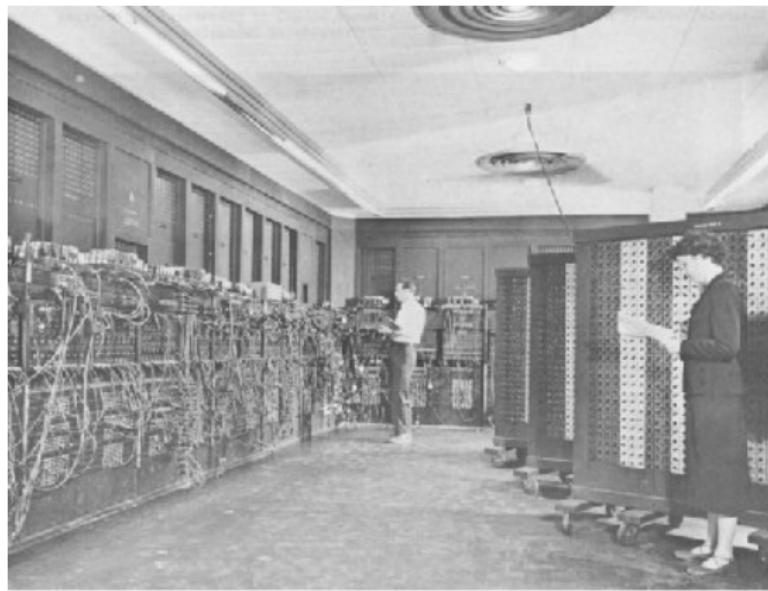
- ① A robot may not injure a human being or, through inaction, allow a human being to come to harm
- ② A robot must obey any orders given to it by human beings, except where such orders would conflict with the First Law
- ③ A robot must protect its own existence as long as such protection does not conflict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



I, Robot (20th Century Fox, 2004)

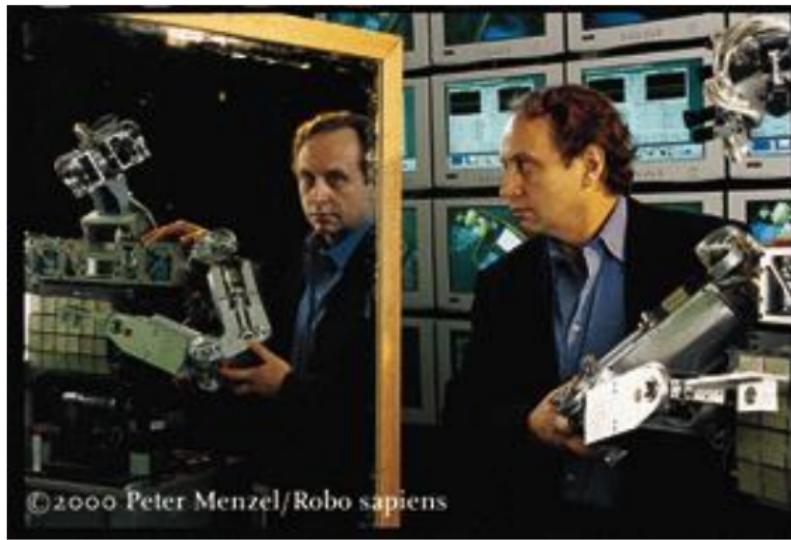
Računala i elektronički mozak

- 1945. godine razvijen je ENIAC, prvo elektroničko računalo
- U početku razvoja računala smatralo se da su računala istovjetna **elektroničkom mozgu**



Pa, mogu li strojevi misliti?

- Ne znamo, ali znamo da možemo programirati računala tako da upravljaju vrlo složenim procesima, rješavaju složene probleme, donoše odluke, zaključuju, koriste prirodan jezik, prepoznavaju vizualne objekte...



Rodney Brooks i robot Cog, MIT Media Lab

Deep Blue vs. Garry Kasparov (1)

- 1997. godine računalo Deep Blue (IBM) pobijedilo je svjetskog šahovskog prvaka Garryja Kasparova
- Je li Deep Blue inteligentan?

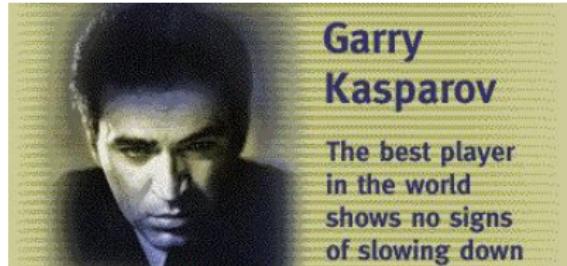


Deep Blue vs. Garry Kasparov (2)



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 šahovskih pozicija u sekundi

3 šahovske pozicije u sekundi

Posjeduje **мало znanja** o šahu, ali **огромну способност** izračunavanja

Posjeduje **много зnanja** o šahu, ali bitno **manju sposobnost** izračunavanja

Stroj **nema osjećaje niti intuiciju**, ne zaboravlja, ne može se zbuniti niti osjećati neugodno

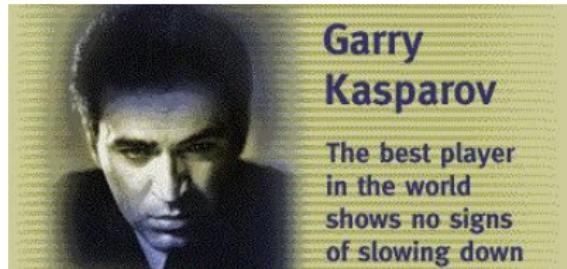
Ima osjećaje i **istančanu intuiciju**, ali može osjećati **umor i dosadu** te gubiti koncentraciju

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

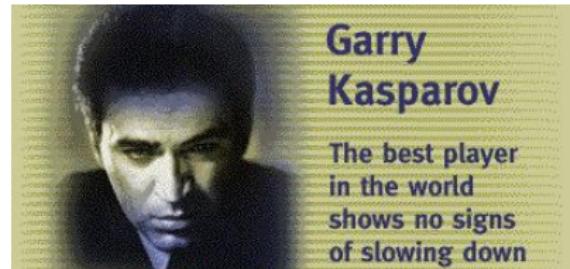
Deep Blue **ne uči**, pa ne može iskoristiti umjetnu inteligenciju da bi naučio od svog protivnika

Deep Blue nevjerljivo učinkovito **rješava probleme iz domene šaha**, no manje je "intelligentan" čak i od malog djeteta

Garry Kasparov može učiti i **brzo se prilagoditi** na temelju svojih uspjeha i pogrešaka

Garry Kasparov je općenito vrlo intelligentan. Autor je nekoliko knjiga i govori mnoge jezike

Deep Blue vs. Garry Kasparov (4)



Izmjene u načinu igre mogu napraviti samo članovi razvojnog tima, i to tek nakon igre

Deep Blue je vrlo vješt u procjeni šahovskih pozicija, no nije u stanju procijeniti slabosti svoga protivnika

Deep Blue mora provesti temeljito pretraživanje svih mogućih pozicija da bi odredio optimalni potez

Garry Kasparov u svakom trenutku **može promijeniti** svoj način igre

Garry Kasparov je vješt u procjeni svoga protivnika, i u **iskorištavanju protivnikovih slabosti**

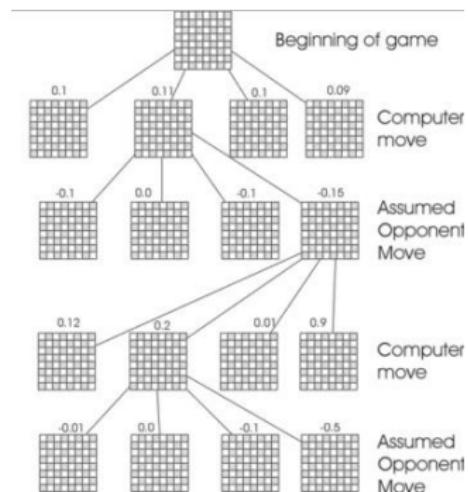
Garry Kasparov je sposoban **selektivno pretraživati** da bi odredio sljedeći potez

Pretraživanje prostora stanja

- Rješavanje ovakvih teških problema iziskuje **pretraživanje** velikog prostora stanja
- U prosjeku je moguće napraviti 35 različitih poteza
- Da bi se na majstorskoj razini igrao šah, potrebno je pretraživati **8 koraka unaprijed**, što znači da treba ispitati oko 35^8 ili $2 \cdot 10^{12}$ stanja

Kombinatorna eksplozija

Broj kombinacija **raste eksponencijalno** u svakom koraku ("netraktabilan problem")



Pretraživanje prostora stanja

- Šah je razmjerno jednostavan – faktor grananja je 35, a pravila igre stanu na jednu stranicu
- Znatno složenije: interpretacija **rečenice prirodnog jezika**

Višeznačnost prirodnog jezika

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.



- Ova rečenica može imati 8064 tumačenja – višeznačna je i za ljude!

Razumijevanje prirodnog jezika

- *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.*

Prevođenje i komunikacija prirodnim jezikom podrazumijeva **opće znanje i razumijevanje konteksta** pomoću kojih se razrješavaju više značnosti

Strojno prevodenje

- Strojno prevodenje jedan je od najtežih zadataka umjetne inteligencije, odnosno **obrade prirodnog jezika** kao njezine podgrane

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

- Prijevod engleski → španjolski → engleski:

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 – projekt DeepQA



- 16.2.2011.: Superračunalo **IBM Watson** pobijedilo je u igri Jeopardy najbolje ljudske natjecatelje i osvojilo \$35.734
- Napredni postupci **obrade prirodnog jezika, prikazivanja znanja, zaključivanja i pretraživanja informacija**

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 row of small orange icons representing different types of input or output. A message box below the search bar says "Assuming 'How do you feel today' is a phrase | Use as a question about Alpha instead". The main content area is divided into 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

Zadatci teški za računalo, a jednostavni za čovjeka

- Razumijevanje prirodnog jezika
- Zdravorazumno zaključivanje (engl. *common sense reasoning*)
- Raspoznavanje uzorka, razumijevanje slike i dinamičke scene
- Kretanje i navigacija
- Zadatci koju uključuju kreativnost
- ...

Činjenica je: **mi puno toga znamo.**

UI-potpuni problemi (*AI-complete problems*)

Računalni problemi čija je složenost istovjetna rješavanju središnjeg problema umjetne inteligencije: izgradnji stroja inteligenčnog koliko i čovjek

Danas

1 Nedavna postignuća umjetne inteligencije

2 Mi i strojevi

3 Inteligencija i umjetna inteligencija

4 Kratka povijest umjetne inteligencije

5 Pregled tema

Što je inteligencija?

- *lat. intelligere* – razabirati, shvaćati, razumijevati
- Inteligencija je deskriptivan pojam – opisuje neka svojstva jedinke ili grupe jedinki
- Ne postoji suglasnost oko definicije inteligencije
- Većina definicija uključuje koncepte kao što su **apstraktno rasuđivanje, razumijevanje, samosvijest, komunikacija, učenje, planiranje i rješavanje problema**

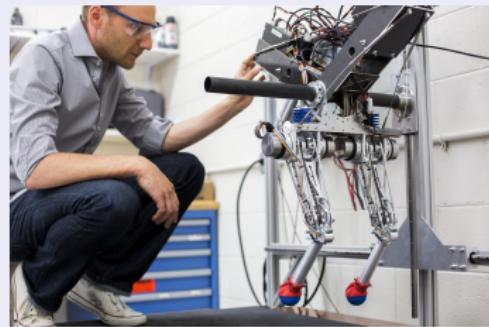
Pragmatičan (inženjerski) pristup inteligenciji

- Umjesto diskusije je li neko ponašanje inteligentno ili ne, možemo prihvati jedan pragmatičan pristup:

“Understanding by building”

Ako je neko ponašanje (čovjeka, mrava, slona, robota,...) zanimljivo, pitamo se:

Kako je ono nastalo?



- Takav pristup omogućava razumijevanje principa na kojima se temelji inteligencija.

Što je *umjetna inteligencija*?

- Grana računarske znanosti:

Tehničke znanosti → Računarstvo → Umjetna inteligencija

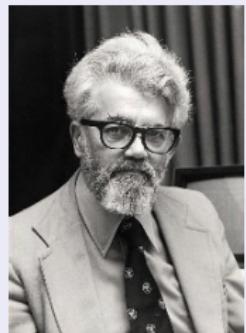
- Područja umjetne inteligencije (prema *Association of Computing Machinery*, ACM):

- ① Opće područje (kognitivno modeliranje, filozofske osnove)
- ② Ekspertni sustavi i primjene
- ③ Automatsko programiranje
- ④ Dedukcija i dokazivanje teorema
- ⑤ Formalizmi i metode prikaza znanja
- ⑥ Strojno učenje
- ⑦ Razumijevanje i obrada prirodnih i umjetnih jezika
- ⑧ Rješavanje problema, metode upravljanja i metode pretraživanja prostora stanja
- ⑨ Robotika
- ⑩ Računalni vid, raspoznavanje uzorka i analiza scena
- ⑪ Rasподijeljena umjetna inteligencija

Naziv "umjetna inteligencija" (1)

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.**”*

Naziv "umjetna inteligencija" (2)

- Znanstvenici s vodećih institucija: CMU, Stanford, MIT, IBM
- Dartmouthska konferencija nije donijela spektakularne rezultate, ali je ustanovila novo područje – **umjetnu inteligenciju** – područje različito od operacijskih istraživanja ili teorije upravljanja, koja su prije toga nastojala odgovoriti na slična pitanja

John McCarthy, (1956.)

*"Umjetna inteligencija – naziv za znanstvenu disciplinu koja se bavi izgradnjom računalnih sustava čije se **ponašanje** može tumačiti kao intelligentno"*

Definicije umjetne inteligencije (1)

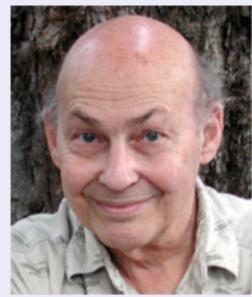
Patrick. H. Winston (MIT)

"Proučavanje postupaka koji mogućim čine percipiranje, rasudivanje i reagiranje."



Marvin Minsky (MIT)

"Znanost o tome kako postići da strojevi izvode zadatke koji bi, kada bi ih radio čovjek, iziskivali inteligenciju."



Definicije umjetna inteligencije (2)

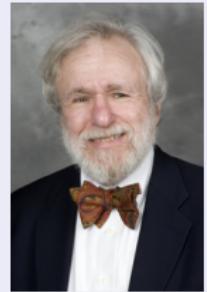
Elain Rich (University of Texas at Austin)

"Umjetna inteligencija bavi se izučavanjem kako računalo učiniti sposobnim da obavlja poslove koje u ovom času ljudi obavljaju bolje."



Eugene Charniak (Brown University)

"Proučavanje mentalnih svojstava kroz uporabu računalnih modela."



Definicija umjetna inteligencije

D. W. Patterson (1990.)

*"Umjetna inteligencija grana je računarske znanosti koja se bavi proučavanjem i oblikovanjem računarskih sustava koji pokazuju **neki oblik inteligencije**. Takvi sustavi mogu učiti, mogu donositi zaključke o svijetu koji ih okružuje, oni razumiju prirodni jezik te mogu spoznati i tumačiti složene vizualne scene te obavljati druge vrste vještina za koje se zahtijeva **čovjekov tip inteligencije**."*

Kako ispitati umjetnu inteligenciju?

Alan Turing, *Can machines think? (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."*



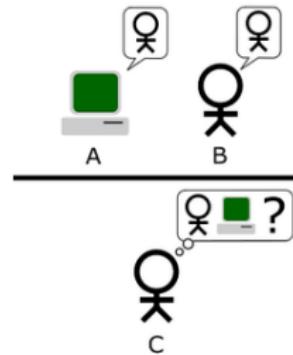
Ispitivanje inteligencije

- Alan Turing u članku “*Computing Machinery and Intelligence*” (1950.), u časopisu *Mind*, predložio je operacionalizaciju pitanja “mogu li strojevi razmišljati”
- Pitanje “**Mogu li strojevi razmišljati?**” zamijeniti pokusom “**igra oponašanja**”
- Pokus uspoređuje performanse prepostavljenog intelligentnog stroja i čovjeka na temelju nekog skupa upita



Turingov test

- Tri igrača, svaki s drugačijim ciljem: A, B (odgovaraju na pitanja) i C (postavlja pitanja). A je stroj, B i C su ljudi
- **Cilj igrača C:**
odrediti je li stroj igrač A ili B
- **Cilj igrača B:**
pomoći igraču C u njegovom zadatku
- **Cilj igrača A:**
navesti igrača C na pogrešan zaključak
- Igra se ponavlja više puta i bilježi se uspjeh igrača C
- Hoće li igrač C načiniti jednak broj pogrešaka kao i u slučaju kada su oboje A i B ljudi?
- **Turing: ako je broj pogrešaka identičan, onda je stroj inteligentan**



Turingov test – analiza

Q: Koje bi sposobnosti (inteligentan) stroj trebao imati, a da prođe TT?

- obrada prirodnog jezika
- prikaz (predstavljanje) znanja
- automatsko zaključivanje
- učenje

Turing je predviđao da će do 2000. računala (s oko 120 MB memorije) imati 30% šanse zavarati ljudе.

Q: Postoji li stroj koji je prošao TT?

- Do danas niti jedan stroj nije (formalno) prošao TT
- Rani pokušaji: kontroverzna “Loebnerova nagrada” (od 1990.)
- LLM-ovi (LaMDA, ChatGPT, etc.) blizu su tome da prođu TT

No, je li TT valjan test inteligencije stroja?

Nedostatci Turingovog testa

- Ljudska vs. opća inteligencija (ljudi se ponekad ponašaju neinteligentno, a inteligentno ponašanje ne mora nužno biti ljudsko)
- Prava vs. simulirana inteligencija (filozofski protuargument ponašajno-orientiranoj UI)
- Naivnost ispitiča (dokazano u slučaju ELIZA-bota)
- Irrelevantnost testa

Irelevantnost testa

Udžbenici aeronautike cilj aeronautike ne definiraju kao:

“Izgradnja strojeva koji lete tako slično golubovima da mogu prevariti druge golubove”



- TT je značajniji za filozofiju UI-a, nego što je to za sam razvoj UI-a

Danas

1 Nedavna postignuća umjetne inteligencije

2 Mi i strojevi

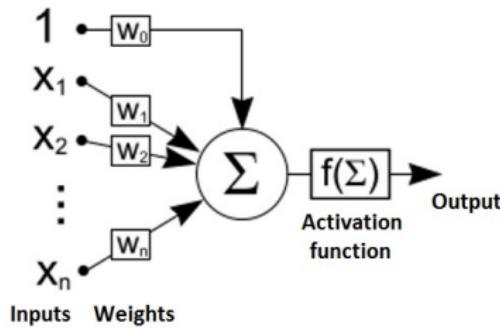
3 Inteligencija i umjetna inteligencija

4 Kratka povijest umjetne inteligencije

5 Pregled tema

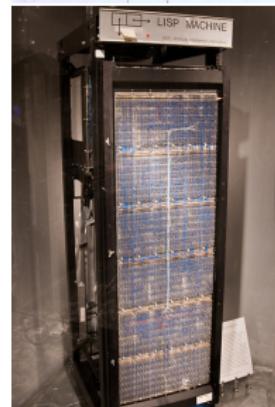
Početak: 1943. – 1952.

- 1943. J. McCulloch, W. Pitts: model umjetnog neurona
- 1949. D. Hebb: pravilo za modificiranje veze između dva neurona
- 1951. Minsky i Edmons: prva neuronsku mrežu od 40 neurona (vakumske cijevi)
- 1950. A. Turing: Turingov test, strojno učenje, genetički algoritmi, podržano učenje



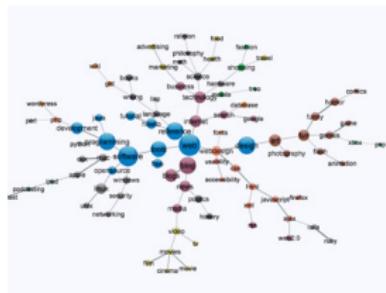
1952. – 1969. Rani entuzijazam, velika očekivanja (1)

- 1952. A. Samuel: igra dame, program koji uči
- 1956. Newell, Shaw i Simon: Logic Theorist (LT)
– skraćeni dokaz teorema iz knjige *Principia Mathematica*
- 1957. Newel & Simon: GPS, prvi program koji je utjelovio ljudski način razmišljanja
- 1958. J. McCarthy: LISP
- 1960.–1962. Widrow i Hoff: Adaline
- 1962. F. Rosenblatt: dokaz konvergencije perceptron-a



1952. – 1969. Rani entuzijazam, velika očekivanja (2)

- 1965. Joseph Weizenbaum – razgovorni program ELIZA
- 1965. Robinson – pravilo rezolucije
- 1966. Quillian – semantičke mreže
- 1969. Minsky, Papert: “Perceptrons” – ograničenje neuronske mreže



1952. – 1969. Otrežnjenje (1)

- Rezultati ranih sustava pokazali su se slabima na širem rasponu problema ili na težim zadacima
- Rani programi sadržavali su **malo ili ništa znanja**, a uspjeh se temeljio na jednostavnim sintaktičkim manipulacijama

Neuspjeh strojnog prevođenja (1957.)

Strojno prevođenje (financirano zbog ubrzavanja prevođenja ruskih radova o Sputniku) temeljilo se na sintaktičkim transformacijama i zamjenama riječi na temelju engleske i ruske gramatike. Rezultat:

“The sprit is willing but the flesh is weak”
→ *“The votka is good but the flesh is rotten”*



1952. – 1969. Otrežnjenje (2)

- Drugi veliki problem – **netraktabilnost** mnogih problema koje je UI pokušavala riješiti
- Početni uspjeh bio je moguć zbog malog broja kombinacija u mikrosvjetovima (pojednostavljeni problemi)
- Prije razvoja teorija izračunljivosti mislilo se da će se “skaliranje” na veće probleme riješiti povećanjem računalne snage!
- 1969. Minsky i Papert: *Perceptrons* – obeshrabrenje istraživanja neuronskih mreža

1970. – 1979. Sustavi temeljeni na znanju

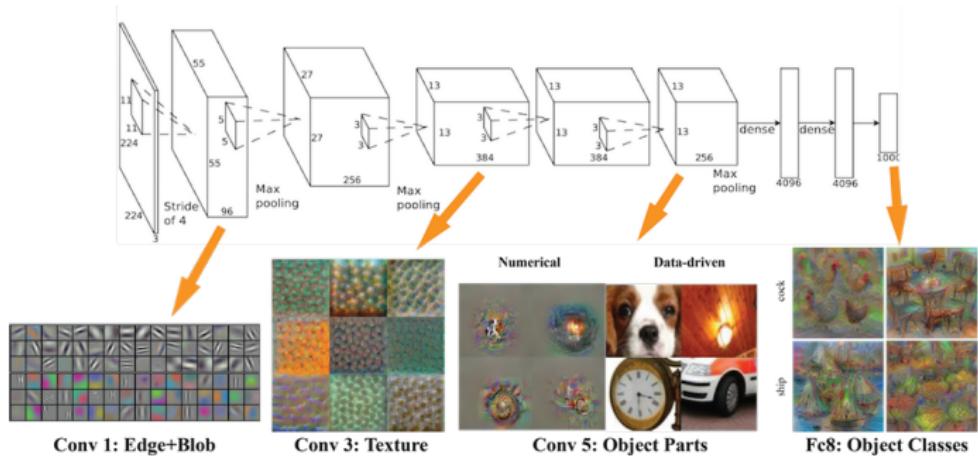
- DENDRAL, Fiegenbaum, Buchanan (Stanford) – sustav temeljen na znanju zaključuje o molekularnim strukturama organskih spojeva na temelju spektroskopije masa – 450 pravila
- MYCIN, Shortliffe (Stanford), 550 pravila, različit od DENDRALA: nema teorijskog modela kao podlogu, uvođenje faktora izvjesnosti
- Napredak u obradi prirodnog jezika
- PROLOG – logički programski jezik popularan u Evropi
- 1975. Minsky, teorija okvira (frames)

1980. – 2010.

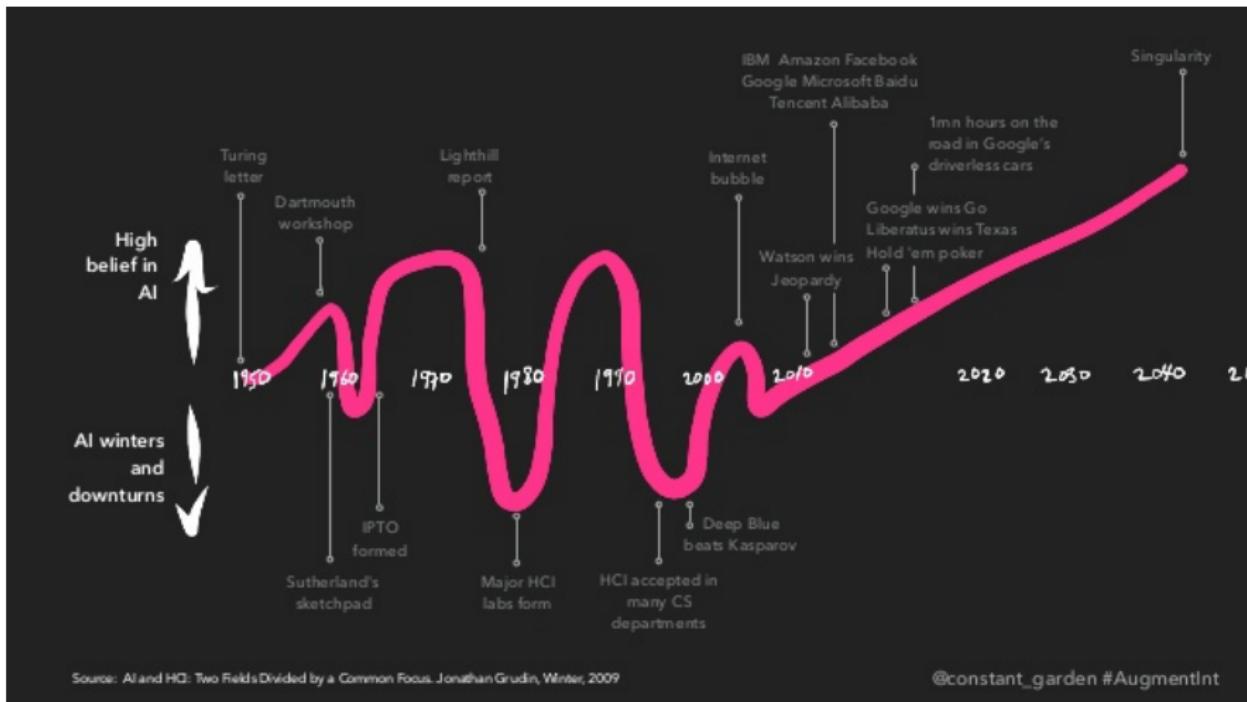
- 1980 – UI je industrija! (od nekoliko milijuna dolara u 1980. do miliarda dolara u 1988.)
- 1982 McDermott – DEC R1 eksperjni sustav
- 1980-ih – Povratak neuronskih mreža (Werbos – algoritam *backpropagation*)
- **Inteligentni agenti** (agent – percepција окoline кроз сензоре и дјелovanje на средину кроз акције)
- **Robotika**
- **Strojno učenje**

2010. – danas

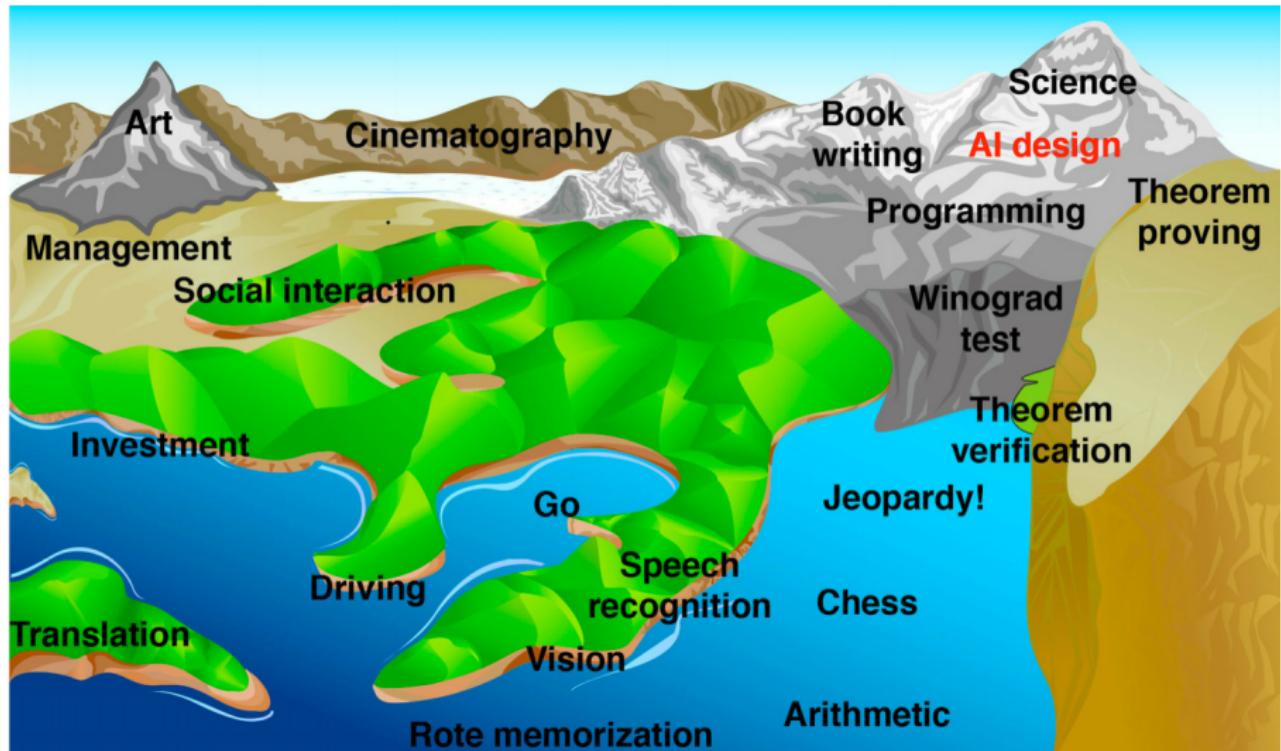
- Era **dubokog učenja** (deep learning)
- Duboko učenje – učenje višeslojnih apstrakcija podataka kroz višeslojne neuronske mreže
- Strojno učenje na velikim količinama podataka
- Veliki napretci u računalnom vidu, obećavajući napretci u obradi prirodnog jezika



Ciće zime umjetne inteligencije



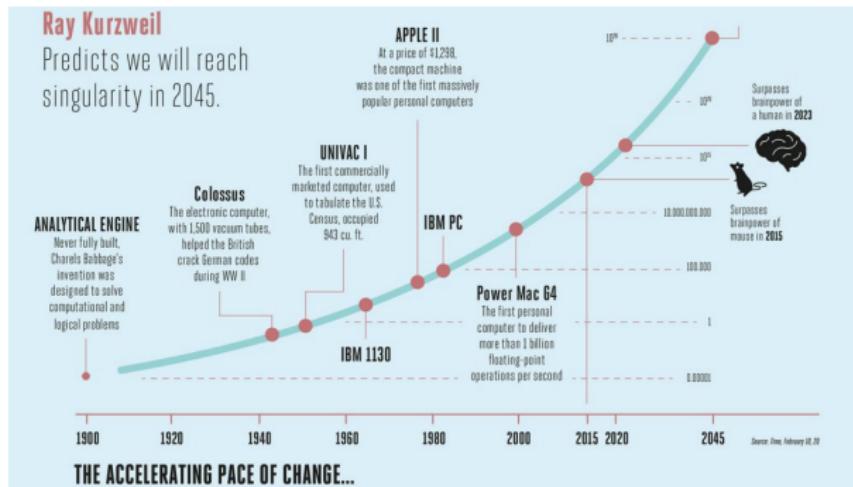
Što nas čeka sutra?



Hans Moravec: Krajolik ljudske kompetencije

Glavna pitanja

- ① Kratkoročni problemi: sigurnost, zakoni, oružje, poslovi
- ② Srednje/dugoročni problemi: AGI, superinteligencija, singularnost?
- ③ UI i svijest (jaka UI)?



... ovim pitanjima bavit ćemo se na zadnjem predavanju!

Danas

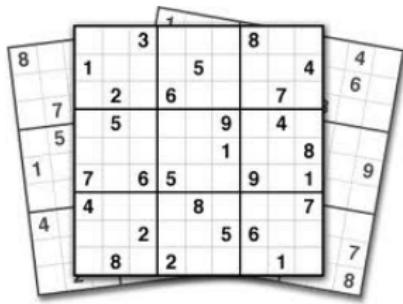
- 1 Nedavna postignuća umjetne inteligencije
- 2 Mi i strojevi
- 3 Inteligencija i umjetna inteligencija
- 4 Kratka povijest umjetne inteligencije
- 5 Pregled tema

Čime bavi ovaj predmet?



- ① Pregled **temeljnih postupaka i algoritama UI**
- ② **Prednosti i nedostatci** različitih metoda
- ③ Trenutačna **ograničenja i mogućnosti UI**

Pretraživanje prostora stanja



Heurističko pretraživanje

početno stanje:

8		7
6	5	4
3	2	1

završno stanje:

1	2	3
4	5	6
7	8	

Što upotrijebiti kao heurističku funkciju?

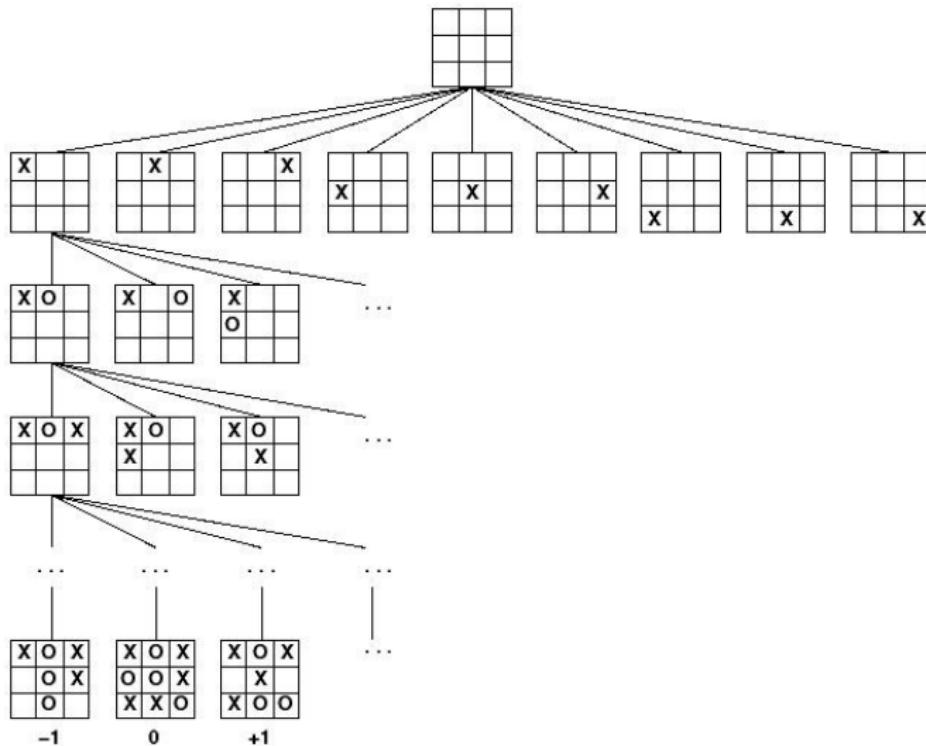
- Broj razmještenih pločica:

$$h_1(\begin{array}{|c|c|c|} \hline 8 & & 7 \\ \hline 6 & 5 & 4 \\ \hline 3 & 2 & 1 \\ \hline \end{array}) = 7$$

- Zbroj L1-udaljenosti između pločica i njihovih završnih pozicija:

$$h_2(\begin{array}{|c|c|c|} \hline 8 & & 7 \\ \hline 6 & 5 & 4 \\ \hline 3 & 2 & 1 \\ \hline \end{array}) = 21$$

Igranje igara



Prikazivanje znanja formalnom logikom

The Wumpus World

4	Stench		Breeze
3	Wumpus	Stench Breeze Gold	PIT
2	Stench		Breeze
1	Hunter	Breeze	PIT
	1	2	3

Percepti (činjenice):

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

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

$$B_{2,1}$$

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

$$S_{1,2}$$

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

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

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

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

Znanje (pravila):

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

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

⋮

Automatsko zaključivanje

- Robot dostavlja pakete. Robot zna da su svi paketi u sobi 27 manji od svakog paketa u sobi 28. A i B su paketi. Paket A je u sobi 27 ili u sobi 28, ali robot ne zna u kojoj. Paket B je u sobi 27 i nije manji od paketa A.
- Rezolucije opovrgavanjem pokažimo kako robot može zaključiti da je paket A u sobi 27.
- Prikazivanje znanja:



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

Logičko programiranje

The screenshot shows the SWISH Prolog IDE interface. On the left, the code editor displays a Prolog program with various predicates and facts. On the right, a trace window shows the execution of the `jealous/2` predicate with different arguments.

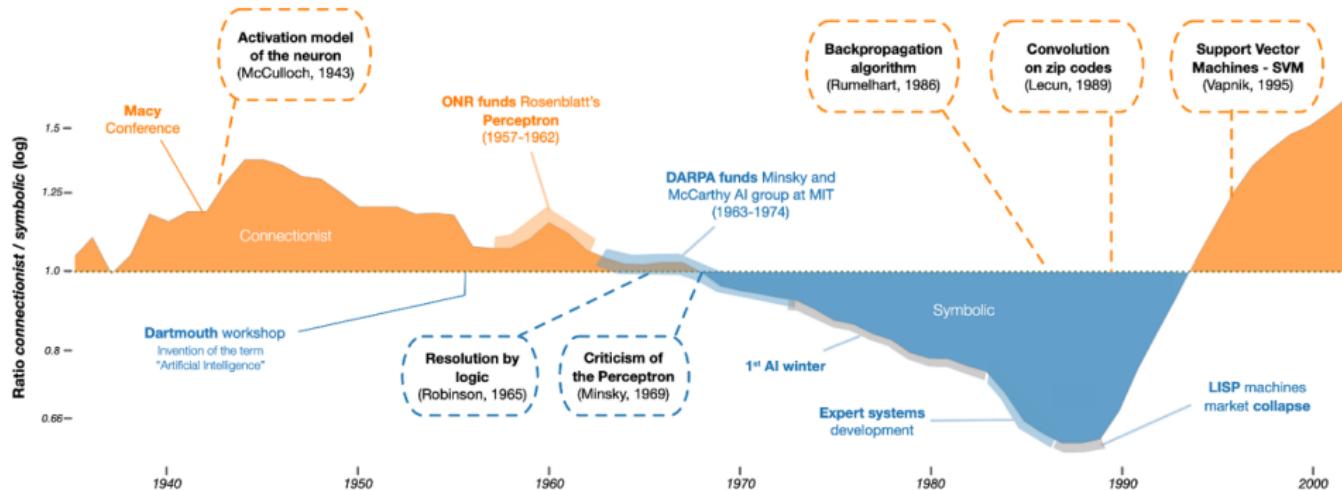
Code Editor:

```
1 % Some simple test Prolog programs
2 %
3 %
4 % Knowledge bases
5
6 loves(vincent, mia).
7 loves(marcellus, mia).
8 loves(pumpkin, honey_bunny).
9 loves(honey_bunny, pumpkin).
10
11 jealous(X, Y) :-  
12     loves(X, Z),  
13     loves(Y, Z).  
14
15
16 /** <examples>
17
18 ?- loves(X, mia).
19 ?- jealous(X, Y).
20
21 */
22
23
```

Trace Window:

```
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)).
```

Simbolizam vs. konekcionizam



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

Modeliranje nesigurnosti

$p = \text{"Danas je sunčan dan"}$



P je
istina



P je
laž



?



?



?



?

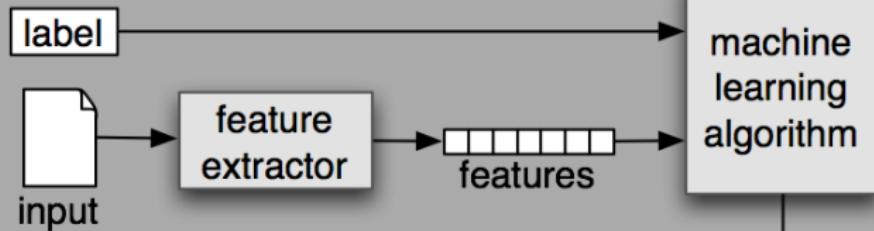


?

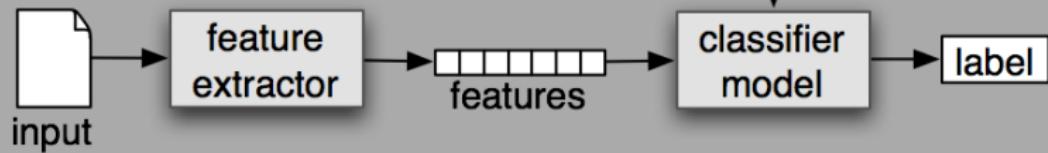


Strojno učenje

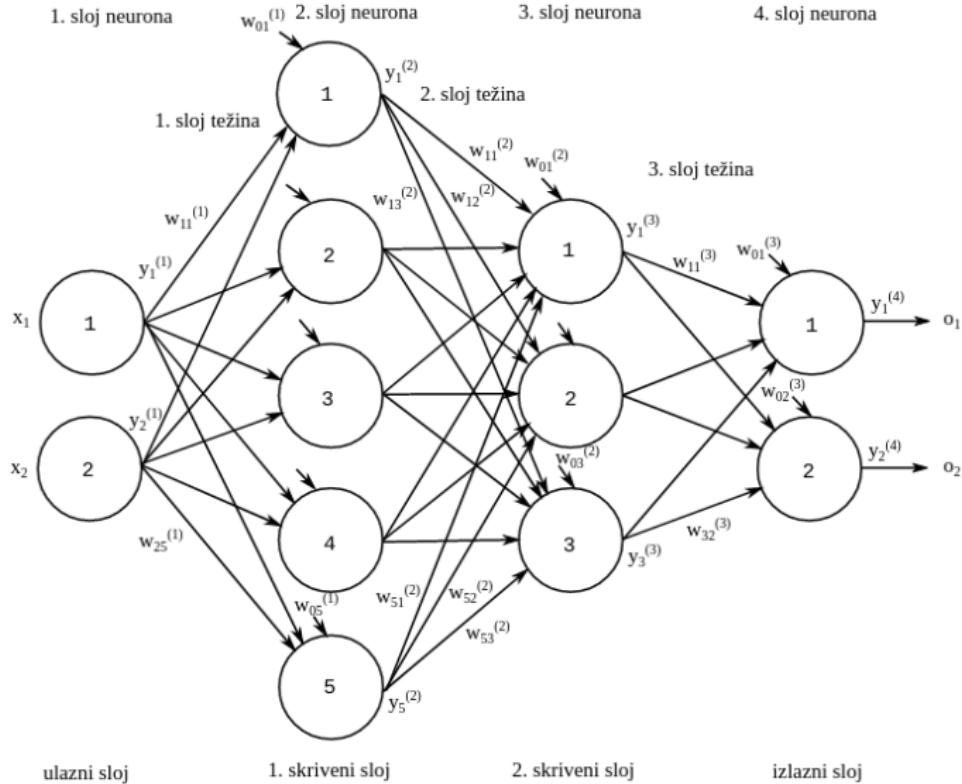
(a) Training



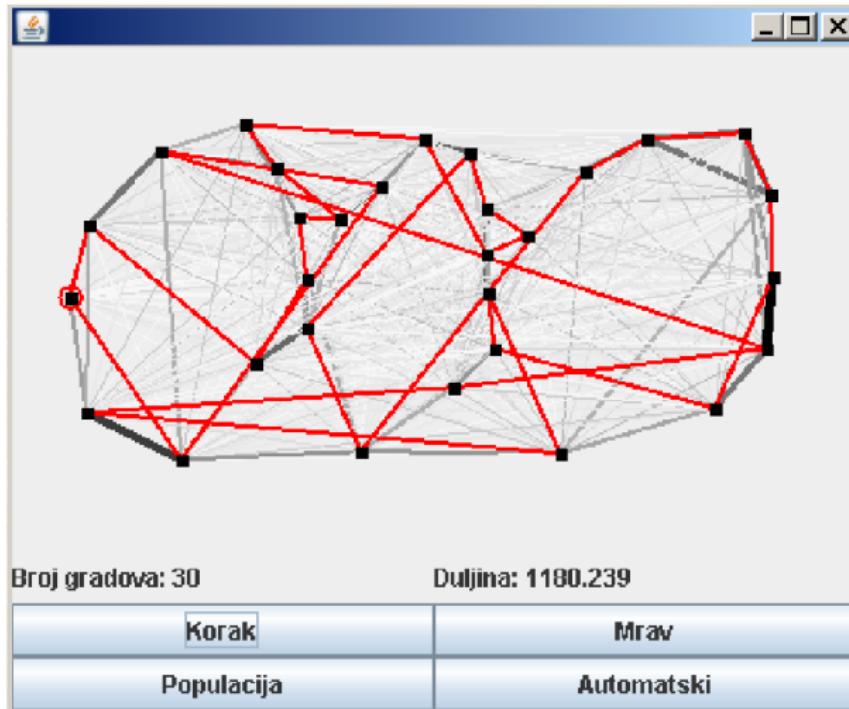
(b) Prediction



Umjetne neuronske mreže



Prirodom inspirirani optimizacijski algoritmi



Podržano učenje

				3
				7
				11

-0.038	0.089	0.215	0.000
-0.165		-0.443	0.000
-0.291	-0.418	-0.544	-0.772

				3
				7
				11

0.973	0.973	0.973	0.000
0.973		0.753	0.000
0.973	0.973	0.948	0.732

Sažetak

- Nastojanja da se napravi inteligentan stroj sežu daleko u prošlost
- Mnogi zadatci koji su jednostavnii za čovjeka, teški su za računalo.
Vrlo teške zadatke nazivamo **UI-potpuni zadatci**
- Ne postoji općeprihvaćena definicija UI, ali možemo usvojiti pragmatičan pristup "**understanding by building**"
- **Turingov test** mjeri inteligenciju stroja kroz igru oponašanja. Test je interesantan, ali u praksi nije toliko značajan
- Kroz povijest, UI je imala uspone i padove. Rana ekstravagantna obećanja uglavnom nisu ispunjena
- Računala danas mogu uspješno (i bolje od ljudi) **rješavati mnoge specifične probleme**



Sljedeća tema: Pretraživanje prostora stanja