

Suvremena Filozofija (9)

Umjetna inteligencija

Predpočetci umjetne inteligencije (-1956)

- Idejno, barem od praškog Golema (16 st.):
 - rabin Judah Loew ben Bezalel
 - "Imao je znanja" o tome kako je Bog stvorio Adama, i to je koristio da bi stvorio Golema od blata koji je branio židove od antisemitskih napada
 - Golem je imao "vrata" na prsima, koja su se otvarala i stavile bi se upute

Što je AI?

- Je li AI stvaranje "novog" čovjeka?
 - Sasvim sigurno ne u bilo kojem smislenom značenju (zašto ne? **Superhuman...**), ali diskusija jest povijesno i filozofski zanimljiva
- Razmislite o "stvaranju umjetnog/novog čovjeka"
 - "prirodni"(apsolutno dozvoljen) i "neprirodni" način
 - Metaforički? (Nietzsche)
 - Uloga religije u osuđivanju (uloga čovjeka u stvaranju i "moralne" granice te uloge)
 - Je li Bog jedini koji smije "stvarati inteligenciju" (**katoličanstvo**) ili je Bog taj koji je čovjeku dao mogućnost (a samim time i dopuštenje (**evangelički protestantizam**), možda i obvezu (**kalvinizam**)) da stvori novu inteligenciju?
 - Diskusija usko povezana sa srednjovjekovnom diskusijom o predestinaciji

Što je AI?

- Uloga genetike i računarstva u omogućavanju
 - Zašto nam kloniranje nije toliko interesatno?
 - Klonirani čovjek je i dalje čovjek
 - Možda ima nedostataka, a sve augmentacije su stavljene na umjetan način
 - On nam je i dalje "blizak" i njegova inteligencija nam je srodna
 - Umjetna inteligencija, realizirana nad strojem nam djeluje fundamentalno drugačija
 - Koliko nas bi nas fascinirao razgovor s kloniranim čovjekom vs razgovor s ChatGPT
 - Od čovjeka, pa i kloniranog "očekujemo" ljudski razgovor, a od stroja ne

Human i superhuman

- Je li vam čudnije imati stroj koji:
 - Može razgovarati na zanimljivi način?
 - Može pobijediti bilo koga na svijetu u šahu?
- Human vs superhuman
 - Superhuman djeluje ono zbog čega želimo strojeve, od traktora do sustava za kontrolu satelita
 - Human je očekivano od ljudi, ali strojevima je strano
 - Očito je što bi bilo superhuman šah, ali što je "superhuman" razgovor? Ili prepoznavanje emocija?

AI

- U pozadini AI kao AI je jedna fundamentalna pretpostavka:
 - Descartes: Res Cogitans i Res Extensa
 - AI odbacuje ideju da je Res Cogitans supstancijalno vezana uz biološki supstrat
 - Pitanje je li moguće da se Res Cogitans na neki način realizira na neljudskom (i neživotinjskom) supstratu se malo pomalo u filozofiji počelo nazivati "The Mind-Body Problem"

AI prije AI-ja

- Pitts i McCulloch (1943): umjetne neuronske mreže
- Norbert Wiener (1948): kibernetika
- Yehoshua Bar-Hillel (1954): strojno prevođenje
- Svi ovi ljudi su po obrazovanju bili filozofi

Dartmouth (1956)

- Matematičar John McCarthy sazvaao
- "Dartmouth Summer Research Project on Artificial Intelligence"
- Namjerno nije pozvao Noberta Wienera jer kako je sam rekao "onda bi umjetna inteligencija bila Wienorovo područje"
- Marvin Minsky, Nathaniel Rochester (IBM, osigurao 10000 USD), Claude Shannon, Allan Newell i Herber Simon

Dartmouth (1956)

- "Pretpostavka ove konferencije jest da se svaki aspekt ljudske inteligencije se može toliko precizno definirati da se može pretvoriti u računalni program"
- Uočite da je po ovoj definiciji JEDNAKO važno:
 - Precizno definiranje (filozofija, psihologija, pa čak i sociologija)
 - Replikacija (računarstvo)

(LazyProgrammer, danas: "Bez eksperimentiranja umjetna inteligencija je tek filozofija")

Herbert Simon

- Sociolog, nobelovac
- Mentor mu je na postdocu iz filozofije znanosti bio R. Carnap
- Zajedno s Newellom, koristeći varijantu Fregeove logike napravili prvi AI program u povijesti (prezentiran u Dartmouthu), Logic Theorist koji je dokazao prvih 38 teorema iz Russellove Principie Mathematice
- "Allan and I have solved the famous mind-body problem"
- JSL nije htio objaviti rad jer se radilo o "dokazima već poznatih teorema", ali nisu skužili da je to bio znanstveni rad u povijesti gdje je koautor stroj

Sciences of the Artificial (Simon 1969)

- Vrlo interesantna i jako zanemarena knjiga iz filozofije znanosti
- Većina knjige je kibernetička, ali dio je čista filozofija znanosti
- Ukratko, (novovjekovne) znanosti do sada su analitičke
Razumijeti = rastaviti
- AI je prva sintetička znanost gdje
Razumijeti = sastaviti/reproducirati
- Reprodukcijska, kao i analiza ne moraju biti savršene (i nikad nisu), ali one definiraju kriterij znanstvenosti u AI i bit će glavni kriterij znanstvenosti u budućnosti

Claude Shannon

- Matematičar
- 1937. dokazao da se sve logičke formule (tzv. logički sklopi s fanout 1) mogu realizirati na sklopovlju telefonski
- 1948. "A Mathematical Theory of Communication"
 - Teorija informacija
 - Entropija, krosentropija

AI 1950-ih i 1960-ih

- Hladni rat, vojna tehnologija, vojni novac
 - Strojno prevođenje kao ključni moment
 - Računalna analiza špijunskih fotografija
 - Robotika i senzorska tehnologija
 - AI je jako važan i danas na senzorima, npr. Uključivanje ABS.
 - Najbanalnije, ako $\text{temp} < 4$ stupnja, uključi.
 - Ali, ima jako puno faktora i jako puno kombinacija, pa se danas za to jako uspješno koristiti AI
 - Puno nade za AI, ali posebno za primjene se i sada polaže na "strojno učenje", odnosno umjetne neuronske mreže

AI Winter (1969/1973-1986)

- Nakon 20 godina obilnog financiranja AI-ja, postavilo se pitanje što se time dobilo (1973, UK, Lighthill report)
 - Svi AI sustavi su bili mali sustavi koji se zbog eksponencijalne kompleksnosti nisu mogli proširiti na stvarne scenarije
 - Osim umjetnih neuronskih mreža
 - Imale problema, ali eksponencijalna kompleksnost nije bila jedan od njih
- 1969. Marvin Minsky i Seymour Papert, XOR problem

1970-e i rane 1980-e, ekspertni sustavi

- Razvoj robotike
- Otkrivanje alternativnih sustava ML, poput DT, ILP, SVM
- Razvoj ekspertnih sustava (prescribed, rule-based)
 - Respecijalizacija "AI startupova" Oracle i Symantec
- Christopher Lounget Higgins, kemičar, AI istraživač u Uni Edinburgh
 - "AI nam nije napravio sustave, ali smo jako puno toga naučili u definiranju intelinegije"
 - To je primarni uspijeh AI, i to AI čini jednom od šest vrlo uspješnih sastavnica "kognitivne znanosti"

1986

- Frank Rosenblatt 1960-ih napravio algoritam za automatizaciju učenja za Pittsove umjerne neuronske mreže, nazvao ga "perceptronskim algoritmom"
 - Dan danas se koristi jer **garantirano konvergira** na globalni minimum eror funkcije linearno separabilnih klasa
- 1969 XOR. Jedino rješenje osim feat engineering je otkrivanje novog algoritma učenja za višeslojne mreže
- 1986, Hinton, Rumelhart, Williams (Hinton glavni) otkrili kako napraviti backpropagation koristeći gradijentni spust
 - Problem za budućnost, nestajući gradijent

1997., 2006., 2010., 2017.

- 1997.:
 - LSTM (Schmidhuber)
 - Konvolucijske mreže (Le Cun)
- 2006.:
 - DBN, deep learning (Hinton)
- 2010.:
 - Autoenkoderi (Bengio)
 - Dropout (Hinton)
- 2017.:
 - Transformer ("Attention is all you need")

Turingova nagrada

- https://en.wikipedia.org/wiki/Turing_Award
- 2018 dobili Hinton, Le Cun, Bengio ali ne Schmidhuber!