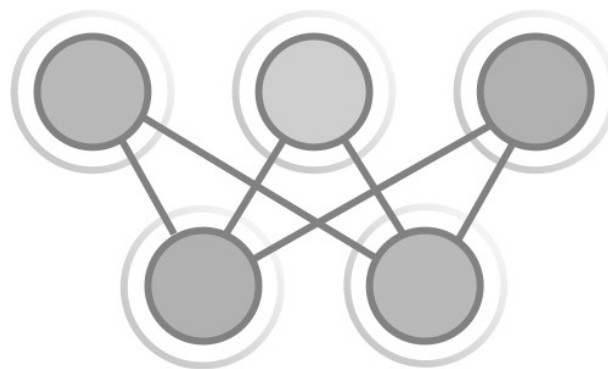


Prof.dr.sc. Bojana Dalbelo Bašić

Fakultet elektrotehnike i računarstva  
Zavod za elektroniku, mikroelektroniku, računalne i inteligentne sustave

[www.zemris.fer.hr/~bojana](http://www.zemris.fer.hr/~bojana)  
[bojana.dalbelo@fer.hr](mailto:bojana.dalbelo@fer.hr)

# Filozofski pogled na UI



v1.1

# Can machine think?

Kada su se rodila računala rodila se i želja da reproduciramo inteligenciju računalom.

**računala = elektronički mozak**

I zaista ... Računalom simuliramo rješavanja problema, donošenje odluka, zaključivanja, prirodni jezik, ...



# Can machine think?



Mogu li se strojevi ponašati inteligentno  
(kao ljudi) i imaju li onda umove?

# “Can machines think? ”

- Alan Turing u članaku *Mind* u časopisu “Computing Machinery and Intelligence” (1950) predložio: Pitanje “**Can machines think?** ” zamijeniti pitanjem može li stroj proći ponašajni test inteligencije - “*igra imaitcije*”.



## Can machines think? (Turing, 1950)

- Teološki/etički
  - “glava u pijesku”
- matematički
- Argument svijesti
- Lady Lavelace
  - neinventivnost strojeva

# Mogu li strojevi misliti?

- Pitanje: "*Mogu li strojevi misliti?*", moguće je interpretirati na dva temeljna načina koja su označila **dva temeljna smjera UI**:

## **JAKA vs. SLABA** **umjetna inteligencija**

- **SLABA UI**: "Mogu li se strojevi napraviti tako da izgleda da se ponašaju inteligentno?"
- **JAKA UI**: "Imaju li strojevi koji se ponašaju inteligentno svjesne umove?"

- **slaba UI** pretpostavlja da računala samo simuliraju razmišljanje te njihovo naizgled razumijevanje učinjenoga nije stvarno
- **Ponašajni** pristup inteligenciji
- npr. program za šah "Deep Blue"



# Jaka umjetna inteligencija

- Zagovara da neki oblik umjetne inteligencije može zaključivati i rješavati probleme, razviti sposobnost prosudbe, **postati samosvjesno** (ne treba nužno i predstavljati proces razmišljanja sličan ljudima )
- "Računalo nije samo alat u proučavanju uma; točnije, **pravilno programirano računalo jest um**, u smislu da se za računala kojima su dani odgovarajući programi može doslovno reći da razumiju i imaju ostala kognitivna stanja"  
(J. Searle:"Minds, Brains, and Programs", 1980)



# Slaba vs. Jaka UI

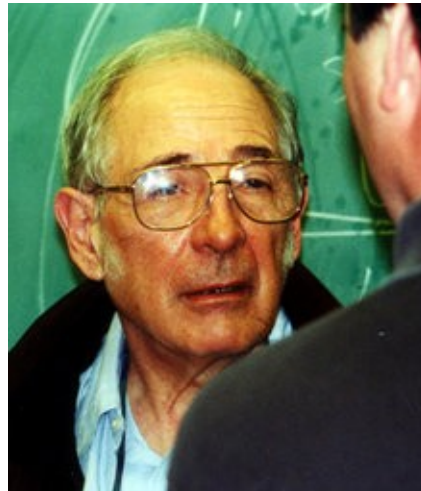
- "Razmišljanje o šahu kod Deep Bluea slično je *razmišljanju* jediničnih stanica o sintezi proteina; oboje nisu svjesni ničega te oboje slijede program koji je ugrađen u njih."
- S takvim razmišljanjem ne slaže se Drew McDermott, profesor računarske znanosti na Yaleu:

"Reći da "*Deep Blue*" zapravo ne misli o šahu, isto je kao i reći da zrakoplov ne leti zato što ne maše krilima."

Što vi mislite o tome?

# Misaoni eksperiment – argument protiv jake UI

- argument protiv “jake UI” i **funkcionalizma** –
  - Funkcionalizam: ideja da se stanja uma mogu ostvariti bez obzira na medij (biološki neuroni), bitni su samo procesi (funkcije)
- misaoni eksperiment koji je osmislio John Searle

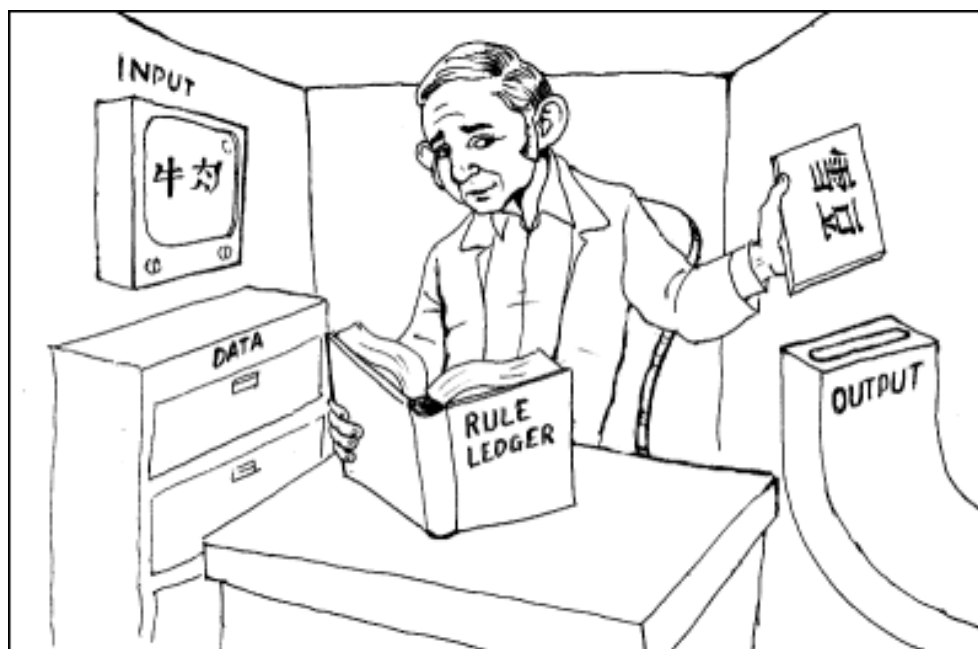


John Searle

# Kineska soba

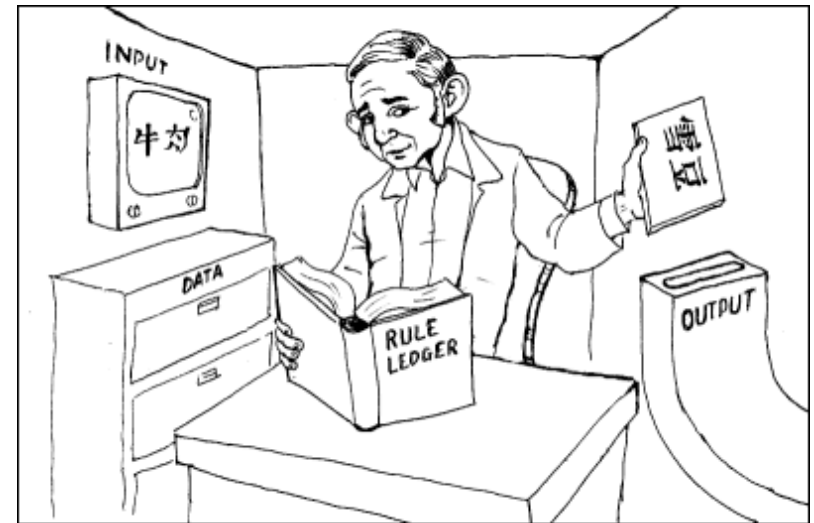
John Searle

- Ne može se suditi o inteligenciji sustava, dokaz je eksperiment nazvan *Kineska soba*



"Minds, Brains, and Programs," by John R. Searle, from *The Behavioral and Brain Sciences*, vol. 3. 1980.

- čovjek koji ne razumije kineski zatvoren je u sobi
- prosleđuju mu se kineski znakovi
- čovjek prevodi kineske znakove pomoću knjige pravila i šalje ih izvan sobe
- osoba unutar sobe je poput procesora, a zbirka pravila na engleskom program



- Osoba unutar sobe prošla bi Turingov test za razumijevanje kineskog (argument protiv TT)
- Naizgled inteligentno ponašanje sustava (računala i programa koji je omogućio takvo ponašanje), ALI računalo nema um, jer simboli koje ono obrađuje njemu **nemaju značenje!**
- **nema pravih mentalnih (intencionalnih) stanja!**

- Searlov argument **protiv** tvrdnje “*program je sve što je potrebno za inteligenciju i spoznajna stanja*”.
- Smatra da je za razumijevanje jezika potrebno:
  - znanje o objektima koji postoje u svijetu
  - **poznavanje riječi jezika**
  - znanje da bi se riječi mogle povezati s objektima (semantika)
  - **poznavanje gramatičkih pravila da bi se formirale smislene rečenice (sintaksa)**

- Searlov argument **protiv** tvrdnje “*program je sve što je potrebno za inteligenciju i spoznajna stanja*”.
- Smatra da je za razumijevanje jezika potrebno:
  - znanje o objektima koji postoje u svijetu
  - **poznavanje riječi jezika**
  - znanje da bi se riječi mogle povezati s objektima (semantika)
  - **poznavanje gramatičkih pravila da bi se formirale smislene rečenice (sintaksa)**
- 1. Računala mogu ostvariti 2. i 4. kriterij



- Premisa 1: Računala su u potpunosti definirana svojom formalnom, sintaktičkom strukturom.
- Premisa 2: Umovi imaju semantički sadržaj.
- Premisa 3: Sintaksa nije dovoljna za semantiku.
- Premisa 4: Mozgovi uzrokuju umove.

- Zaključak 1: Nijedan računalni program sam po sebi nije dovoljan za uzrokovanje uma. Programi nisu umovi.
- Zaključak 2: Išta drugo što uzrokuje umove moralo bi imati kauzalnu moć barem ekvivalentnu onoj od mozga.
- Zaključak 3: Sve što bi uzrokovalo umove, svaki umjetni mozak, morao bi imati mogućnost dupliciranja određenih kauzalnih moći mozga, što se ne bi moglo učiniti samo pokretanjem formalnog programa.
- Zaključak 4: Način na koji mozak funkcionira uzrokuje um i ne može biti samo djelovanje izvođenja računalnog programa.

# Odgovori na argument kineske sobe:

- *The Systems Reply*
  - čovjek ne razumije kineski, ali **sustav kao cjelina** koji razumije
- *The Robot Reply*
  - ne postoji razumijevanje, ali ono bi se moglo postići ako bi se “soba” prebacila u **robota koji interagira s okolinom**, tako da postoji **kauzalna veza** između simbola i njihove semantike
- *The Brain Simulator Reply*
  - pretpostavimo da postoji program koji do u detalje simulira svaki neuron mozga. Pojedinačni neuroni ništa ne razumiju, ali oni **očito uzorkuju um i raumijevanje**
  - *China brain, Brain prosthesis experiment*
- *Other Minds Reply*
  - **nije moguće dokazati mentalna stanja niti kod drugih ljudi** (pa tako ni razumijevanje), no ipak konvencionalno pretpostavljamo da ih imaju

# The Brain Prosthesis Experiment

- zamjenjuju se svi neuroni u ljudskom mozgu elektroničkim ekvivalentima, koji oponašaju neurone i sposobni su povezati se s njima
- zatim se radi obrnuti proces
- Proučava se vanjsko ponašanje i unutarnja stanja
- Hoće li svjesnost nestati ukoliko se zamijene svi neuroni?
- Moravec (robotičar, funkcionalist) vs. Searl (filozof)

## Filofija uma (philosophy of mind):

- Grana filozofije koja proučava prirodu uma, mentalnih događaja, mentalnih funkcija, mentalnih svojstava, svjesnosti te njihov odnos s fizičkim tijelom
- Problem um-tijelo, problem slobodne volje i problem ličnosti



Koji pristup prihvaćamo u ostvarivanju  
ciljeva UI?



# **Simbolički vs. Konektivistički pristup**

- **GOFAI** ("Good Old-Fashioned Artificial Intelligence"), tradicionalna, klasična, "top-down" pristup
- inteligentno ponašanje može se oblikovati pomoću sustava koji zaključuje logički na temelju skupa pravila i činjenica



# Konektivistički pristup

- “Subsymbolic” AI, “bottom-up” pristup, PDP

Intelligentno ponašanje postiže se povezivanjem jednostavnih procesnih elemenata koji rade paralelno (UNM).

Klasičan pristup dao je značajne rezultate no ipak nije mogao riješiti neke probleme...

Mozak ne funkcionira serijski kao računalo već **paralelno, može učiti**, stjecati znanje kroz iskustvo, neosjetljiv je na pogreške....

**Umjetne neuronske mreže** najznačajniji su predstavnik konektivističkog pristupa.

# Behavior based artificial intelligence and embodied cognitive science

- Nastala je krajem 80tih,
- Rodney Brooks (MIT, AI Lab)  
Pitanje “Can machine think?”  
zamijeniti “Kako nastaje inteligentno ponašanje?”
- Inteligencija je utjelovljena. Inteligencija je rezultat interakcije “organizma” sa sredinom.
- Inteligenciju možemo razumjeti gradnjom artefakata



I drugi pristupi...

I na kraju .....

# Područja umjetne inteligencije

## Klasifikacija prema **Association of Computing Machinery** (1998)

- Opće područje (kognitivno modeliranje, filozofske osnove)
- Ekspertni sustavi i primjene
- Dedukcija i dokazivanje teorema
- Formalizmi i metode prikaza znanja
- Strojno učenje (uvod)
- Razumijevanje i obrada prirodnih i umjetnih jezika
- Rješavanje problema, metode upravljanja i metode pretraživanja prostora stanja
- Robotika
- Računalski vid, raspoznavanje uzoraka i analiza scena
- Porazdijeljena umjetna inteligencija

# Što ćete naučiti na ovom kolegiju?

- Upoznat ćete sadašnja ograničenja i mogućnosti UI
- Pregled temeljnih metoda i tehnika UI
- Razumjet ćete prednosti i mane različitih metoda
- Moći ćete prepoznati probleme u kojima je pogodno primijeniti metode umjetne inteligencije
- Pregled različitih filozofskih pogleda na UI

# Jesmo li ostvarili cilj?

- Upoznat ćete sadašnja ograničenja i mogućnosti UI
- Pregled temeljnih metoda i tehnika UI
- Razumjet ćete prednosti i mane različitih metoda
- Moći ćete prepoznati probleme u kojima je pogodno primijeniti metode umjetne inteligencije
- Pregled različitih filozofskih pogleda na UI