

Filozofija kognitivnih znanosti

izr. prof. dr. Olga Markič

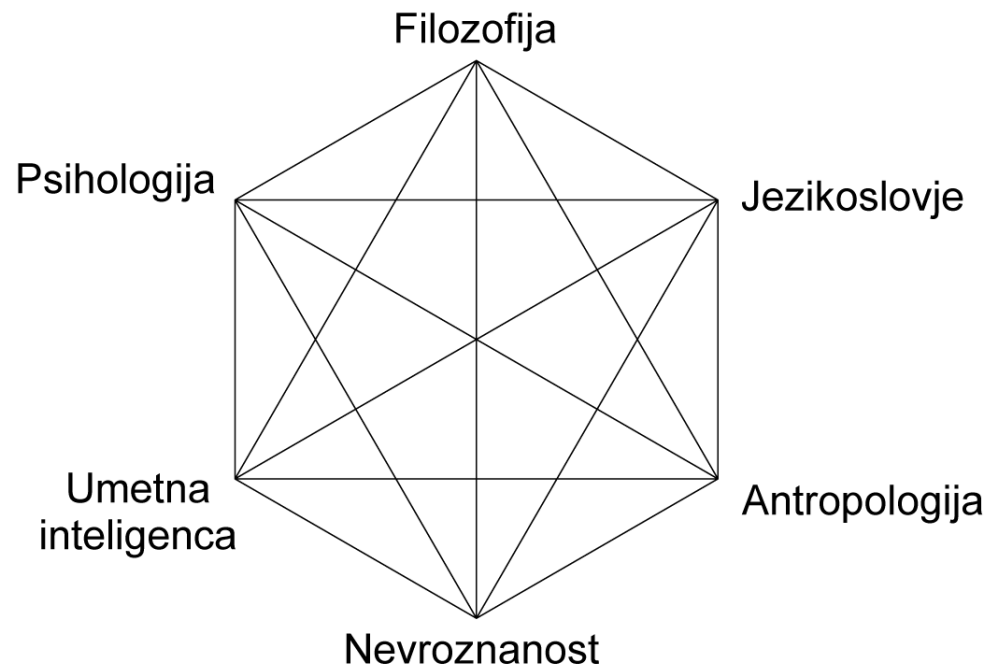
olga.markic@guest.arnes.si

- Interdisciplinarna struktura kognitivne znanosti
- Paradigme
- Vprašanja filozofije duha in filozofije znanosti
- Etična vprašanja

Raziskovalno področje

- Kognitivni procesi v ožjem smislu
 - Zaznavanje, pomnjenje, učenje, reševanje problemov
- Čustva
- Volja, hotenje

Kognitivna znanost



KOGNITIVNA ZNANOST



- splošna opredelitev kognitivne znanosti:
Kognitivna znanost je znanosti o duševnih procesih. (Bermudez, 2010)
- Definicija KZ iz začetnega obdobja:
»Kognitivna znanost je multidisciplinarni znanstveni študij kognicije in njene vloge v inteligentnih (razumnih) dejavnostih. Raziskuje kaj kognicija je, kaj počne in kako deluje«
(Bechtel, Abrahamsenova in Graham, 1998: 3)

Kratek zgodovinski oris

- Raziskovanja v okviru filozofije:
od starogrških filozofov do sodobne
filozofije duha (Platon, Aristotel,
Descartes, Spinoza, Hume, Locke, Kant,
Brentano, Husserl, Merleau-Ponty, Fodor,
Dennett, Clark, ...)

Raziskovanja v okviru psihologije kot samostojne znanstvene discipline

- Introspekcijem

temeljna znanstveno metoda: introspekcija

- temeljna hipoteza: psihologija je raziskovanje “fenomenologije” človekove duševnosti
- popoln opis zornega sveta duševnosti, kot se prikazuje subjektu (“čutni atomi”)

Najbolj znana laboratorija:

Cornell šola v Ithaci (ZDA) - Titchener

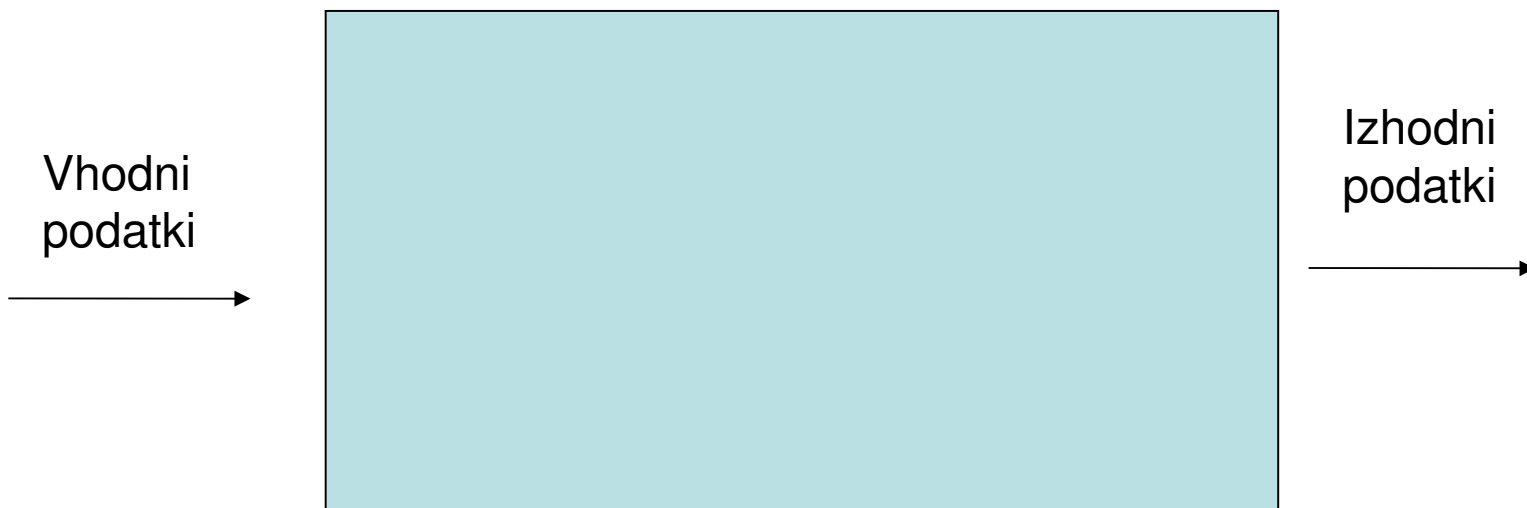
Würzburška šola v Leipzigu (Nemčija) - Külpe.

- Problemi:
 - Notranji vzroki:
različni rezultati laboratorijev,
ni bilo splošno sprejete metode o falsificiranju
rezultatov.
 - Zunanji vzroki:
sprememba v raziskovalni klimi v Evropi in ZDA.
pozitivistična naravnost tako v znanosti kot v
humanistiki, izogibanje vsemu “duševnemu,
mentalnemu”

- Behaviorizem

- sam pojav, ki ga psihologija preučuje, ni nič drugega kot opazovano vedenje; zavest in introspekcija sta izrinjeni iz znanstvenega preučevanja
- "Behaviorizem trdi, da zavest ni jasen in uporaben pojem. Behaviorist, ki se je izšolal kot eksperimentator meni, da izvira prepričanje v obstoj zavesti iz starodavnih časov praznoverja in magije. (Watson, *Behaviorism*, cit. po Güzeldere, 1995, str. 39)

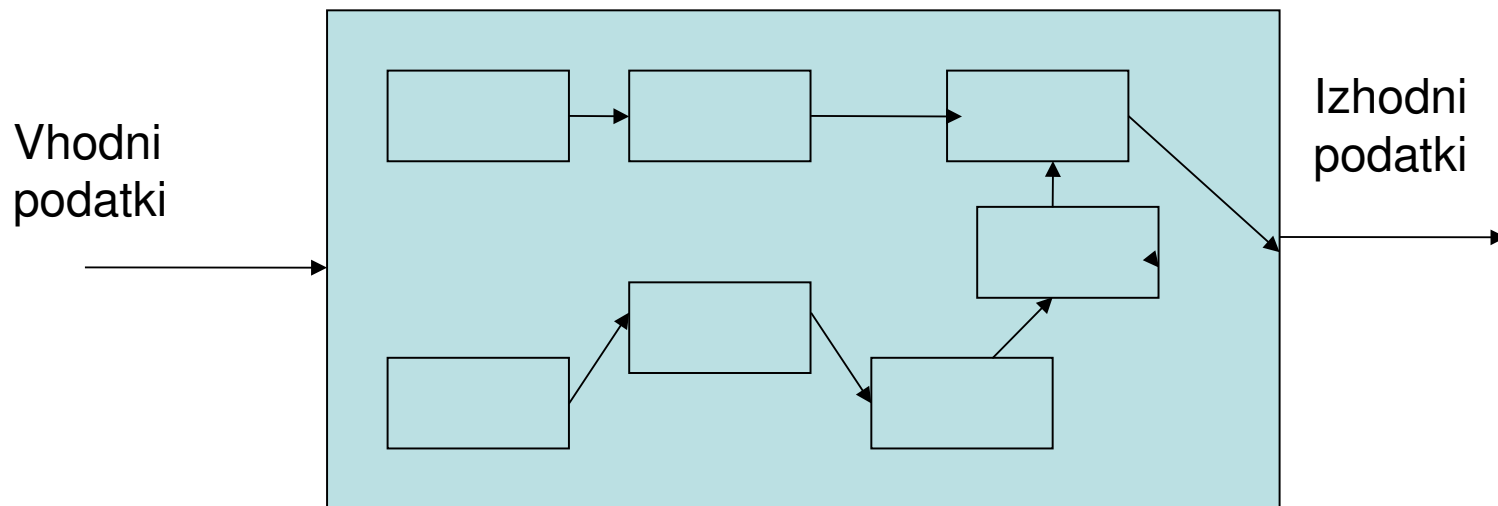
Črne škatla



Kognitivna psihologija

- nastanek nove raziskovalne paradigme
- Psihologi začeli raziskovati spoznavne (kognitivne) procese - odločilni za razumevanje človekovega vedenja

Kognitivna škatla/ škatla prepričanj, želja in namer



Kognitivna znanost

- Hixon Symposium na California Institute of Technology (1948)
 - J. von Neumann, W. Mc Culloch, K. Lashley
- 1956 – rojstno leto - “kognitivna revolucija” (Gardner, 1987)
 - Psihologija: Bruner, Goodnow, Austin: *Study of Thinking* (1956) in Miller *The Magical Number Seven* (1956)
 - Jezikoslovje: Chomsky: *Three Models of Language* (1956)
 - Računalništvo: Newell in Simon: *Logical Theory Machine* (1956)
- Institucionalno v 70.tih letih 20. st.

- Warren Mc Culloch in Walter Pitts (1943) - preučevala, kako možgani obdelujejo informacije in postavila vzporednico med živčnim sistemom in logičnimi napravami
- Temeljna hipoteza kognitivne znanosti:

kognitivni procesi so informacijski procesi

- izhodišče za dva pristopa, ki sta kasneje zaznamovala kognitivno znanost
 - eden vodi k nevronske mrežam, sestavljenim iz velikega števila med seboj povezanih nevronov
 - drugi pa k simbolnim modelom klasične kognitivne znanosti

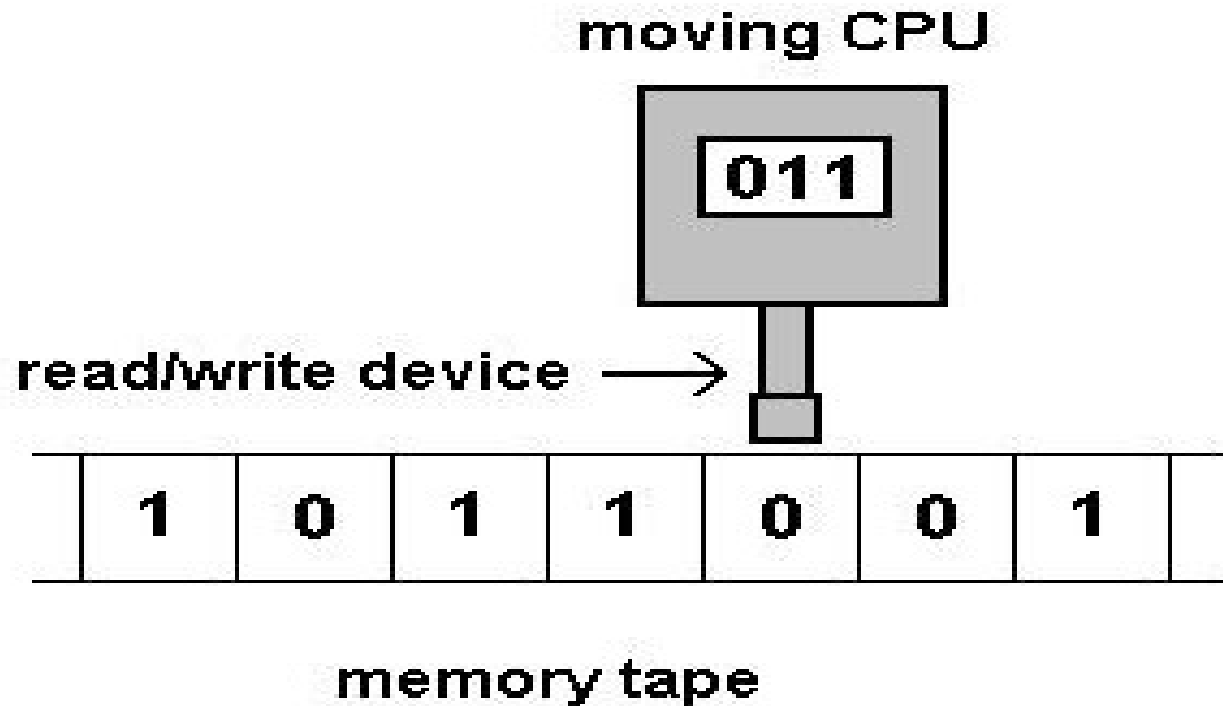
Pristopi in paradigme v kognitivni znanosti

- Model / simulacija
- Pristopi in paradigme
 - Klasični simbolni modeli
 - Fizični simbolni sistemi
 - Konekcionistični modeli
 - Nevronske mreže
 - Teorija dinamičnih sistemov
 - Utelešena kognicija, umetno življenje, roboti

- kognitivizem:
 - Tako kot behaviorizem
 - sprejema cilj psihologije razlaga in napoved vedenja
 - za to je ustrezna raven funkcionalna raven, ne pa nevrofiziologija ali kakšna druga naravoslovna znanost (npr. fizika, biologija).
 - Razlikuje se od behaviorizma
 - odprl je "črno škatlo" in vanjo postavil duševna stanja kot člene v vzročni verigi razlage in napovedi vedenja

- Kognitivna revolucija, nov pristop v proučevanju duha, ne bi bila možna brez teoretskih prispevkov:
 - dosežki matematike in logike
 - razvoj informacijske teorije, kibernetika
 - modeliranje nevronov
 - raziskovanje spoznavnih nezmožnosti, ki so posledice poškodb možganov.

Turingov stroj



- neskončno dolg trak s polji, na katera so vpisani znaki iz naprej določene abecede
- glava, ki bere, kar je na traku zapisano
- Glava stroja je v določenem notranjem stanju iz vnaprej določene končne množice stanj. Stanje glave se v vsakem koraku, ko prebere polje na traku lahko spremeni, dokler se ne znajde v stanju “stoj”, kjer se stroj ustavi in pusti zapisan izhodni podatek na traku.

Delovanje glave je popolnoma določeno z dvema dejavnikoma:

1. z znakom, ki ga prebere na traku;
2. z notranjim stanjem, v katerem se trenutno nahaja.

Ta dva dejavnika določata vsako od treh posledic:

1. kateri znak bo glava napisala na trenutno polje in s tem zamenjala
prejšnji znak;
2. na katero polje se bo premaknila - ali bo ostala na istem mestu, ali se bo premaknila za eno polje v levo, ali pa v desno;
3. v katerem notranjem stanju bo glava naslednji korak ali pa se bo morda ustavila.

Preprost primer Turingovega stroja

EN+1

številki, zapisani v eniškem sistemu, prišteje ena
 $00 \rightarrow 00D$, $01 \rightarrow 11D$, $10 \rightarrow 11STOP$, $11 \rightarrow 11D$.

Prva številka nam pove, v katerem notranjem stanju je glava, številka v mastnem tisku pove, kaj je zapisano na traku, D in L označujeta smer, v katero se premakne glava, znak STOP pa pomeni, da se računanje ustavi.

Če navodila uporabimo na traku0000001111000000...., ki predstavlja število 4, se bo stroj ustavil po upoštevanju navodila $10 \rightarrow 11STOP$. Na traku bo napisano ...00000011111000000..., kar predstavlja število 5.

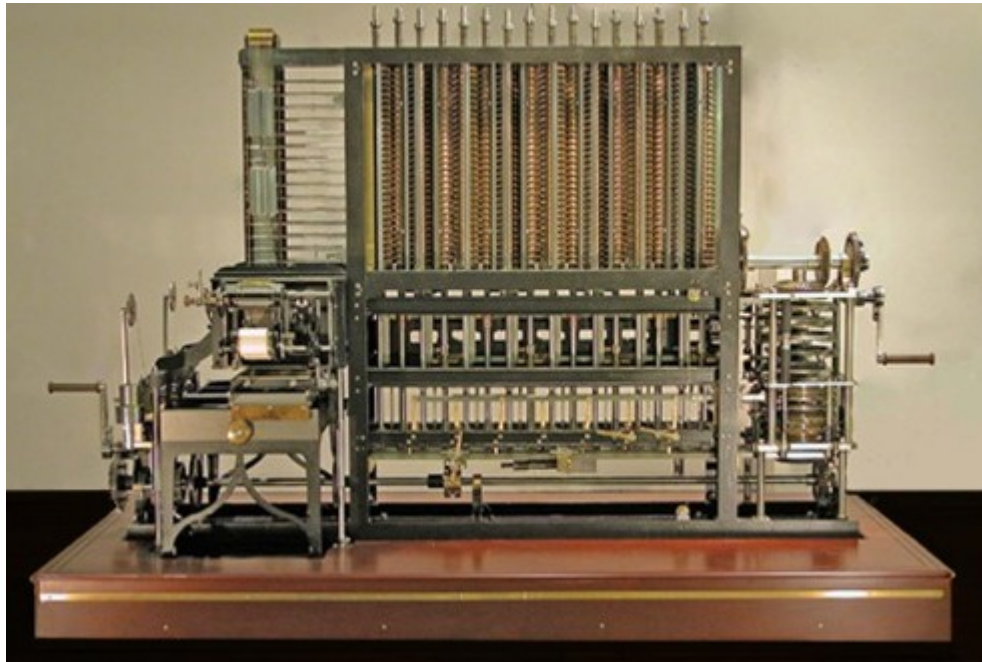
- Turing je pokazal, da lahko s takim preprostim strojem izvedemo vsako nalogo, za katero lahko jasno navedemo korake, ki so potrebni za izpolnitev naloge. Takemu učinkovitemu postopku pravimo algoritem.
- Do ekvivalentnega odkritja je prišel tudi ameriški logik Church: Church - Turingova teza
 - pojem Turingovega stroja dejansko definira to, kar matematično pojmuje kot algoritmična (rekurzivna, mehanična) procedura.

- Turingov stroj lahko programiramo tako, da bo oponašal druge Turingove stroje.
univerzalni Turingov stroj.
- Gre za **digitalno** obliko reprezentacije (npr. digitalna ura), za razliko od **analogne** oblike, kjer so mogoča tudi vmesna stanja (npr. analogna ura).

- Turingov stroj je torej naprava, ki dobi določene vhodne podatke (input), te nato obdeluje v skladu z navodili in nato odda izhodne podatke (output). Takemu načinu procesiranja običajno pravijo manipuliranje s simboli (reprezentacijami) v skladu s pravili, računalniku, univerzalnemu Turingovemu stroju pa manipulator simbolov.
- Vhodni in izhodni podatki, tj. znaki na traku, so simboli (reprezentacije) - predstavljajo nekaj oziroma stojijo za nekaj. Ko je Turingov stroj EN+1 računal funkcijo "prištej 1", je dobil "1111" kot vhodni podatek in "11111" kot izhodni podatek, torej reprezentacijo za število 4 oziroma 5.
- dejstvo, da pri računanju funkcije potrebujemo reprezentacije, ima pomembno vlogo pri vzpostavitvi analogije med miselnimi procesi in računanjem, ki je v temelju kognitivne znanosti.

Elektronski digitalni računalniki

- Angleški matematik Charles Babbage (1791 - 1871) je izdelal natančne načrte za mehanski računalnik - Analitični stroj (Analytical Engine).



- Babbage je pri načrtovanju upošteval dve ideji, ki sta temeljni za računalništvo:
 1. vse operacije lahko programiramo;
 2. programi lahko vsebujejo pogojne razvejitve.

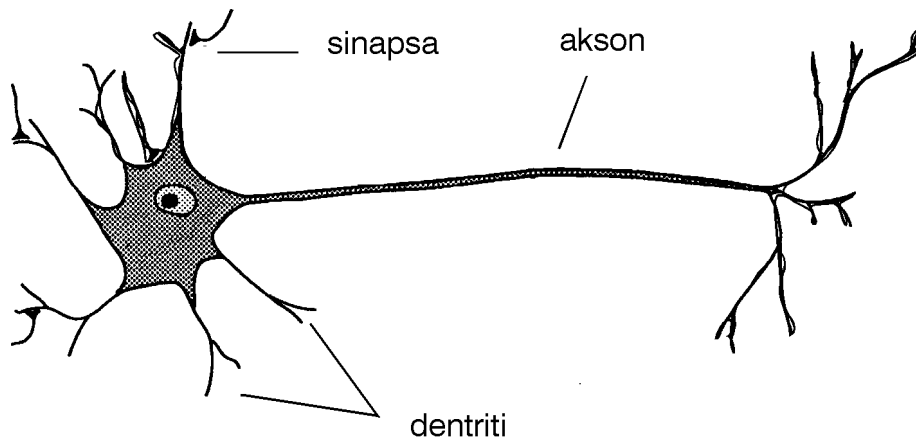
Navodila, kako naj stroj deluje, so bila zapisana na luknjastih karticah, povezanih v trakove.

Babbage je tako lahko spreminjal splošno namenski sistem v specifičen avtomatski sistem in tako vpeljal idejo programiranja, eno najmočnejših idej v zgodovini tehnologije (Haugland, 1985).

- Razvoj elektronskih digitalnih računalnikov
- Von Neumannove arhitekture – pomnilnik, osrednja procesna enota, nadzorna enota
- S programskimi jeziki lahko simuliramo virtualne stroje, ki imajo drugačno arhitekturo kot je Von Neumannova arhitektura na nivoju strojne opreme in so primernejši za modeliranje izbranih kognitivnih funkcij.

Modeliranje nevronov

- Ramon y Cajal (1906)



- McCullocha in Pitts: “A logical calculus of the ideas immanent in nervous activity” (1943)
- operacije nevronov in njihove povezave z ostalimi nevroni lahko modeliramo s pojmi logike
- Nevrone si lahko zamislimo kot stavke. Stanja posameznih nevronov, ki so lahko ali vzburjeni ali ne (lastnost vse ali nič, 1 ali 0), pa primerjamo z resničnostno vrednostjo stavka. Ustrezno obtežene povezave med nevroni predstavljajo logične veznike.

Shematski prikaz nevrona

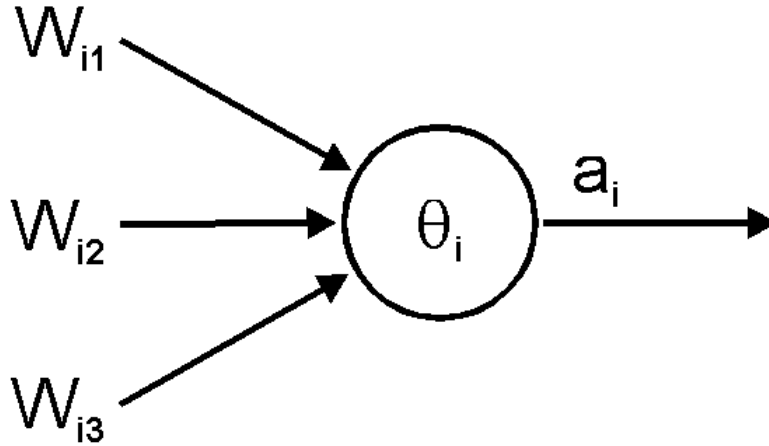


Diagram of a McCulloch-Pitts neuron

$$a_i(t+1) = F(\sum w_{ji} o_j(t) - \theta_i)$$

$$F(x) = 1, \text{ \u010d}e x \geq 0$$

$$F(x) = 0, \text{ \u010d}e x < 0$$

McCulloch in Pitts sta pokazala, da je s sinhrono povezavo takih preprostih nevronov (nevroni se prižigajo istočasno) in izbiro ustreznih uteži v načelu možno izračunati vse, kar lahko izčrpno in nedvoumno opišemo (vse, kar lahko izračuna digitalni računalnik).

Abstrakten model živčnega sistema tako gleda v dve smeri:

1. k računalniku, ki lahko realizira vse tiste procese, ki jih lahko nedvoumno opišemo z logiko;
2. k nevronske mreži oziroma konekcionističnim modelom, ki poudarjajo povezanost in dinamiko.

Model/simulacija

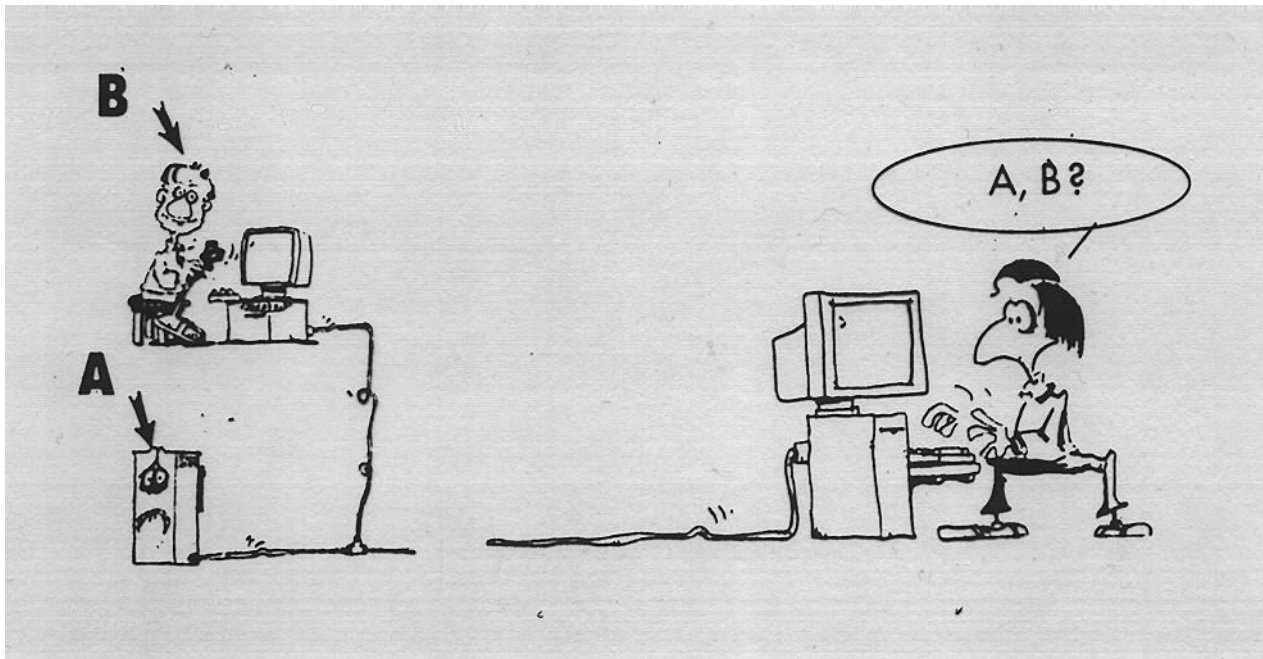
Model – temelji na analogiji, ki je bistvena za raziskovano področje

- metafore iz mehanike - La Mettrie "Človek stroj"
- pretok energije, elektromagnetizem – Freud
- metafora telefonske centrale za centralni živčni sistem – Sherrington
- računalnik - ena od vrste metafor v zgodovini filozofije in psihologije

- metafore izražajo nove teoretične trditve in pomagajo pri konstituciji znanstvene teorije - *teorijsko konstitutivne metafore*
- imajo *induktivni odprti konec* - usmerjenost v nadaljnje raziskovanje, ki bo omogočilo novo razumevanje teoretsko relevantnih vidikov podobnosti ali analogije

Ali lahko stroji mislijo?

- Turingov test kot kriterij za mišljenje



Vsi ljudje so enaki.

NA KAKŠEN NAČIN

Vedno težijo na tak ali drugačen način.

ALI SE SPOMNIŠ DOLOČENEGA PRIMERA

No ja, moj fant me je ravnokar poslal sem.

TVOJ FANT TE JE POSLAL SEM

Pravi, da sem preveč depresivna.

OBŽALUJEM, DA SI DEPRESIVNA

Res je, da sem nesrečna.

ALI MISLIŠ, DA TI BO TO, DA SI PRIŠLA SEM, KAJ POMAGALO IN BOŠ
MANJ NESREČNA

Prav gotovo potrebujem pomoč.

KAJ BI TI POMENILO, ČE BI DOBILA POMOČ

Morda bi se naučila shajati z mojo mamo.

POVEJ MI VEČ O SVOJI DRUŽINI

Mama skrbi zame.

KDO ŠE SKRBI ZATE

Moj oče.

TVOJ OČE

Na nek način ste podobni mojemu očetu.

(Primer terapevtskega razgovora med programom Eliza in najstnico je
povzet po Copeland (1993)).

- Ali je Turingov test primeren
 - 1. Ali test ne zahteva preveč?
 - 2. Ali test ne zahteva premalo?