

Nevrološka verodostojnost in kognitivne teorije

V tem prispevku bom skušala odgovoriti na vprašanje, zakaj kognitivna teorija ne more biti neodvisna od nevroznanosti. Mnogi zagovorniki klasične kognitivne znanosti¹ menijo, da je za konstrukcijo kognitivne teorije dovolj, da znanstvenik ostane na kognitivni ravni, t.j. na ravni reprezentacij. V nasprotju s takim stališčem popolne avtonomije bom z argumentom iz izbire med teorijami pokazala, zakaj je za konstrukcijo teorije, ki resnično opisuje kognitivne procese pri človeku, potrebno poznati delovanje možganov. Kriterij nevrološke verodostojnosti (angl. plausibility) in spoznanje, da je boljša tista kognitivna metodologija, ki ni metodološko avtonomna glede na nevroznanost, bom nato uporabila pri primerjavi dveh pristopov h kognitivnemu modeliranju.

¹ npr. Fodor & Pylyshyn (1988).

Argument iz izbire teorij

Argument ima naslednjo strukturo:

1. Iščemo teorijo, ki resnično opisuje človekove kognitivne procese.
2. Obstaja lahko več kognitivnih teorij, ki so enako uspešne pri napovedovanju vedenja.
3. Kognitivne teorije se razlikujejo glede na to, kakšni so predpostavljeni procesi, ki izbrano kognitivno funkcijo realizirajo, in kako so ti procesi fizično realizirani.

² Npr. obstaja več teorij kratkoročnega spomina: modalni model Atkinsona in Shiffrina, Baddeleyjev sistem delujočega spomina, Andersonov ACT* sistem (Cohen, Kiss & LeVoi, 1993).

4. Pri odločanju, katera teorija je boljša, moramo upoštevati fizikalno strukturo in mehanizme, ki realizirajo predpostavljene procese. Boljša je tista teorija, ki nam zagotavlja fizikalni mehanizem (t.j. dejanske vzročne procese), katerega funkcionalni opis je funkcija, ki jo sistem izvaja.
5. V primeru človeka fizikalno strukturo, ki realizira predpostavljene procese, predstavljajo možgani.
6. Boljša je tista teorija, ki pri realizaciji predpostavljenih procesov bolj upošteva omejitve iz delovanja možganov.

Cilj znanstvenega raziskovanja je konstrukcija take teorije, ki bo resnično opisovala procese. Tako stališče zavzemajo znanstvenice in znanstveniki, ki sprejemajo objektivno resnico kot cilj, ki ga s svojim raziskovanjem želijo doseči. Empirično ustreznost hipotez in teorij običajno preverjajo eksperimentalno. Vendar se lahko zgodi, da z eksperimenti, ki so dostopni opazovanju, ne morejo določiti, katera od dveh teorij (T_1 in T_2) je boljša, saj obe "ohranjata pojave". V takih primerih pravimo, da je teorija poddoločena z evidenco, ki je trenutno dostopna opazovanju znanstvenika. Do take situacije je prišlo v astronomiji, ko sta oba modela, narejena eden v skladu s Kopernikovo heliocentrično in drugi s Ptolomejevo geocentrično teorijo, kazala in napovedovala retrogradno gibanje planetov Marsa, Jupitra in Saturna, ko so v opoziciji s Soncem.

Tudi kognitivni znanstvenik, ki skuša poiskati teorijo, ki bo resnično opisovala kognitivne procese, se lahko znajde pred podobno težavo. Na primer, dve različni kognitivni teoriji T_1 in T_2 sta lahko povsem enaki pri napovedi vedenja, hkrati pa ni na voljo nobenega eksperimenta (na ravni vedenja), ki bi odločil v prid eni ali drugi teoriji. Ker je resnična lahko največ ena od teorij, je treba poiskati kriterij, ki bo take "vedenjsko enakovredne" teorije medsebojno razločeval in na osnovi katerega bo mogoče ugotoviti, katera teorija je boljša, to je, bliže resnici.

Ko konstruiramo kognitivno teorijo, moramo določiti funkcionalno organizacijo sistema. To pomeni, da moramo določiti osnovno funkcionalno arhitekturo, ki pove, s kakšnimi procesi so kognitivne funkcije realizirane. Ti procesi so nato realizirani na fizikalni ravni. Kognitivne teorije se v širšem smislu razlikujejo glede na naravo procesov, ki realizirajo kognitivne funkcije (npr. pri klasični kognitivni znanosti so ti procesi manipulacija s simboli). Kadar imamo v mislih tako razlikovanje, običajno govorimo o različnih pristopih (npr. klasični in konekcionistični pristop). Teorije znotraj različnih pristopov se nato razlikujejo glede na to, kakšni so konkretni postopki, s katerimi je funkcija ob predpostavljenih procesih realizirana, in kako so te funkcije realizirane na fizikalni ravni.²

V primerih, ko je z vedenjskim eksperimentom nemogoče ugotoviti razliko med tisto kognitivno teorijo, ki vedenje resnično pojasnjuje, in tisto, ki ga zgolj oponaša, mora znanstvenik pri iskanju odločitve upoštevati še druge kriterije. Iz filozofije znanosti je znano, da si lahko pri reševanju problema empirične poddoločenosti teorij pomagamo tako, da kot boljšo rešitev izberemo tisto, ki je ontološko preprostejša, bolj elegantna, ali ima večjo združevalno moč. Pri prvem kriteriju gre za znano načelo "Occamove britve", ki pravi, naj brez potrebe ne pomnožujemo ontoloških entitet. Za razlago bi se to načelo gospodarnosti glasilo: če zadostuje preprosta in elegantna razlaga, potem ni potrebe, da bi iskali zapleteno. Vrnimo se na primer h Kopernikovi in Ptolemejevi teoriji. Po Kopernikovi teoriji pride do omenjenega pojava zaradi premikanja zemlje, ki prehiti planete v njihovih orbitah okoli sonca. Teorija pojasni, da je retrogradno gibanje zgolj navidezno in se nujno zgodi, kadar so planete v opoziciji s Soncem. Pri Ptolemejevi geocentrični teoriji prihaja do retrogradnega gibanja zaradi specifične dolžine epiciklov in deferentov³ ter njihovih relativnih hitrosti vrtenja. Da bi model ustrezno prikazal in napovedal pojav, je potrebno določiti več parametrov, kar se zdi bolj ad hoc rešitev, ne pa elegantna teoretska pojasnitev. Poleg tega je Kopernikova teorija nudila enotno pojasnitev za večje število pojavov kot njena tekunica in tako boljše zadostila tudi zahtevi po večji splošnosti.

Pri konstrukciji teorije gre v mnogih primerih za postuliranje teoretskih entitet ali lastnosti na mikroravni, ki nato pojasnjujejo opazljive posplošitve. Na ta način se doseže enotna obravnava in globlja pojasnitev pojavov. Eden od zgledov, ki jih za mikro-teoretsko pojasnitev najdemo v zgodovini znanosti, je pojasnitev valence kemijskih elementov s pomočjo kvantne mehanike. Če ne bi uspeli najti mehanizmov na ravni teoretske mikrostrukture, ki pojasnjujejo celotno dogajanje, bi morali verjeti, da je dejstvo, da teorija deluje, nekakšen čudež. Podobno si tudi pri reševanju problema vedenjske poddoločenosti kognitivne teorije pomagamo tako, da poiščemo mikrostrukturo, to je mehanizme na nižji ravni. Pri kognitivnih teorijah predstavlja tako mikrostrukturo fizična realizacija predpostavljenih procesov. Tista teorija, ki zagotavlja fizikalni mehanizem (t.j. dejanske vzročne procese), katerega funkcionalni opis je funkcija, ki jo sistem izvaja, nudi "globljo" pojasnitev delovanja sistema. Zato je taka teorija boljša.

Kognitivni znanstvenik, ki je pred dilemo pri izbiri teorije zaradi poddoločenosti z vedenjsko evidenco, dejansko ni v tako težkem položaju kot znanstvenik, ki se sooča s splošnim problemom empirične poddoločenosti. Že dolgo je znano, da so pri človeku možgani tisti, ki predstavljajo fizikalno strukturo, ki realizira predpostavljene procese. Iz tega dejstva in iz prejšnje ugotovitve, da je boljša tista teorija, ki zagotavlja dejanske vzročne

³ *Deferent je polmer glavnega kroga.*

⁴ Glej Posner and Raichle (1997).

⁵ Citirano po Churchland (1986, str. 460).

processe, lahko izpeljemo glavni sklep tega argumenta:

Boljša je tista teorija, ki pri realizaciji predpostavljenih procesov upošteva omejitve iz delovanja možganov.

Naj dodam še kratek komentar. Verjetno bi le redki kognitivni znanstveniki oporekali trditvi, da je osnovna funkcionalna arhitektura človekovega psihološkega sistema determinirana z možgani. Z argumentom sem hotela pokazati, da, če enkrat sprejmemo to trditev, potem ne moremo več zagovarjati popolne avtonomije raziskovanja na kognitivni ravni. Nevrofiziološki podatki o možganih so tisti, ki imajo pravico zavrniti kognitivno teorijo. Če kognitivna teorija ne zadovoljuje nevrofizioloških omejitev in tako nima verodostojne realizacije v možganih, potem ni kandidatka za resnično teorijo.

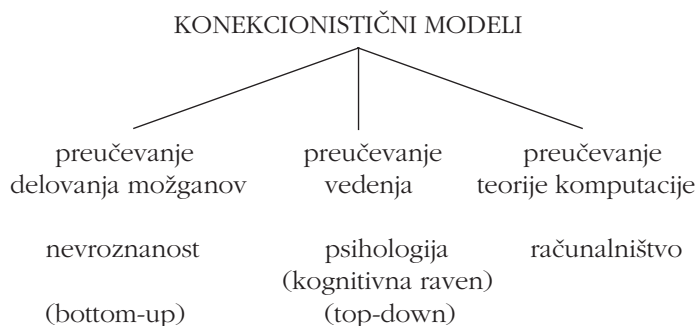
Nevroznanstveniki so v zadnjem času odkrili mnoge podrobnosti o delovanju posameznih nevronov in njihovem povezovanju v nevronske mreže. S sodobnimi metodami (PET, MRI, ERP)⁴ ugotavljajo, kateri deli možganov so aktivni pri posameznih kognitivnih funkcijah. Ti izsledki predstavljajo osnovne podatke, mimo katerih nobena realistična kognitivna teorija ne more. Naj kot zgled omenim znano "pravilo stotih korakov", ki sta ga prva formulirala Feldman in Ballard leta 1982:

"Nevroni, katerih osnovna komputacijska hitrost je nekaj milisekund, morajo omogočiti kompleksno vedenje, ki se izvede v nekaj sto milisekundah. To pomeni, da je CELOTNO KOMPLEKSNO VEDENJE IZVEDENO V MANJ KOT STO ČASOVNIH KORAKIH".⁵

"Pravilo stotih korakov" predstavlja omejitev pri konstrukciji kognitivne teorije. Če je za kompleksno vedenje, npr. prepoznavanje besede, potrebnih nekaj sto milisekund, teorija pa za isto nalogo zahteva več milijonov časovnih korakov, potem taka teorija ni nevrološko verodostojna. Boljša je teorija, ki lahko pojasni izbrani kognitivni pojav tako, da hkrati upošteva omejitve stotih časovnih korakov.

Klasični in konekcionistični kognitivni modeli

Kognitivna teorija mora ustrezno pojasniti vedenje v določeni situaciji in biti hkrati nevrološko verodostojna. To pa nadalje pomeni, da bo boljša tista kognitivna metodologija, ki ne bo metodološko avtonomna glede na nevroznanost. Če v luči teh ugotovitev pogledamo dva najbolj razširjena pristopa h kognitivnemu modeliranju, klasične simbolne modele in konekcionizem, potem lahko ugotovimo, da le konekcionizem ob matematično-računalniških raziskavah črpa tako iz raziskovanj na kognitivni ravni kot iz nevroznanosti. To mu daje prednost, nadaljnja raziskovanja pa bodo pokazala, v kolikšni meri bo tak pristop uspešen pri pojasnjevanju kognitivnih procesov.



⁶ Glej npr. Fodor & Pylyshyn (1988).

⁷ Glej npr. Bechtel and Abrahamsen (1991), Butler (1994).

Z argumentom iz izbire teorij smo pokazali, da je boljša tista teorija, ki je nevrološko bolj verodostojna. Za "klasike" kot zagovornike popolne avtonomije psihologije je bilo vprašanje, kako je funkcionalna arhitektura dejansko realizirana, vprašanje "gole" implementacije, zato se za nevrološko verodostojnost svojih modelov niti niso posebno trudili.⁶ Drugače pa je v konekcionizmu, kjer so modeli nastali po zgledu nevronov in njihovih povezav in podobnost z delovanjem možgan vseskozi igra pomembno vlogo. Znanstvenice in znanstveniki, predvsem tisti, ki prihajajo s področja nevroznanosti, pa vseeno opozarjajo, da so modeli preveč preprosti in vsebujejo nekatere značilnosti, ki jih pri možganih ne najdemo. Naj na kratko povzamem razloge za (predvsem v primerjavi s klasičnimi modeli) in proti nevrološki verodostojnosti konekcionističnih modelov.⁷

Razlogi za:

- sistem medsebojno povezanih enot (nevronska mreža)
zvezno spreminjanje vrednosti
izhod je nelinearna funkcija vhodnih podatkov
kompleksnost povezav
- vzporedno delovanje (omogoča hitro izvajanje osnovnih operacij - pravilo 100 korakov)
- plastičnost - povezave se spreminjajo
- nenadzorovano učenje
- vsebinsko naslavljanje spomina (asociativni spomin)
- postopno slabšanje izvedbe ob poškodbah manjšega dela enot
- modeli določenih poškodb možganov ustrezajo vedenju ljudi s takimi poškodbami (npr. disleksija).

Razlogi proti:

- veliko dejstev o delovanju celice in nevrotansmitterjev, ki jih modeli ne zajamejo
- priklic iz spomina je pri človeku hitrejši, kot nakazujejo modeli
- učenje s pravilom "backpropagation"

⁸ O primerjavi s klasičnimi modeli glej Markič (1996).

- enote in vezi v modelu niso prostorsko umeščene
- modeli so izjemno majhni v primerjavi z dejanskimi nevronskimi mrežami.

Naj povem nekaj besed o razlogih proti.⁸ Prav gotovo ne smemo pozabiti, da so konekcionistični modeli v splošnem zelo velika poenostavitev glede na dejanske nevronske sisteme. Toda vprašanje, ki se postavlja na tem mestu, ni, ali gre za dovolj dobre realistične modele, ampak, ali je poenostavitev taka, da jo še lahko sprejmemo kot dovolj verodostojno za kognitivne modele. Pri tem je treba imeti v mislih dejstvo, da bi imelo natančno modeliranje posameznega nevrona ali upoštevanje delovanja različnih neurotransmiterjev za posledico tako kompleksen model, da ne bi mogel zaobseči posplošitev, ki so bistvene za kognitivno raven. Zato je za kognitivni model nujno, da abstrahira nekatere podrobnosti. Raven opisa kognitivnega modela je v veliki meri odvisna od funkcije, ki naj jo model izvaja, in od pragmatične odločitve, katera raven bi lahko zagotovila zadovoljivo razlago. Sodobni konekcionistični modeli so zato med seboj zelo različni kar se tiče upoštevanja podrobnosti in obstaja cela serija modelov, od najbolj realističnih do zelo poenostavljenih. Gre za nekakšen "daj-dam" med preglednostjo na eni strani ter upoštevanjem nevroloških podrobnosti na drugi. Toda tudi tisti zelo poenostavljeni modeli (npr. za izvajanje določenih jezikovnih procesov) so omejeni glede na funkcionalne lastnosti možganov in skušajo biti kolikor mogoče nevrološko prekinjeni, pač glede na praktične možnosti. Verjetno se bodo nekatere podrobnosti današnjih konekcionističnih modelov izkazale za napačne, toda to ne zmanjšuje koristnosti splošnih principov.

Tako lahko sklenemo, da so konekcionistični modeli nedvomno nevrološko bolj verodostojni kot klasični modeli. Ker so idealizacija, so v nevarnosti, da ostajajo modeli igračke, ki ne bodo mogli nikoli zajeti pravih dimenzij (zadnji razlog). Vendar to ni specifičen problem konekcionističnih modelov, ampak se tiče vprašanja idealizacije v znanosti in grozi vsem oblikam kognitivnega modeliranja.

LITERATURA

- BECHTEL, W. and ABRAHAMSEN, A. (1991): **Connectionism and the Mind**, Basil Blackwell, Oxford.
- BUTLER, K. (1994): "Neural Constraints in Cognitive Science", **Minds and Machines**, Vol. 4., No. 2, 129-162.
- COHEN, G., KISS, G. & LeVoi, M. (1993): **Memory**, Open University Press, Buckingham.
- FODOR, J.A. and PYLYSHYN, Z.W. (1988): "Connectionism and Cognitive Architecture: A Critical Analysis", **Cognition** 28, 3-71.

MARKIČ, O. (1996): "Kognitivizem in konekcijonizem - dva pristopa v kognitivni znanosti", *Časopis za kritiko znanosti*, 176, 101-115.

PAPINEAU, D. (ur.) (1996): *The Philosophy of Science*, Oxford University Press, Oxford.

POSNER, M., RAICHLE, M. (1997): *Images of Mind*, Scientific American Library, New York.