

# Funkcionalistička teorija i njezini nedostaci

---

**Andabak, Mateja**

**Master's thesis / Diplomski rad**

**2018**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Zadar / Sveučilište u Zadru**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:162:492423>

*Rights / Prava:* [In copyright](#) / [Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2024-04-19**



**Sveučilište u Zadru**  
Universitas Studiorum  
Jadertina | 1396 | 2002 |

*Repository / Repozitorij:*

[University of Zadar Institutional Repository](#)



Sveučilište u Zadru

Odjel za filozofiju

Diplomski sveučilišni filozofije; smjer: nastavnički (dvopredmetni)

**Mateja Andabak**

# **Funkcionalistička teorija i njezini nedostaci**

**Diplomski rad**

Zadar, 2018.

Sveučilište u Zadru

Odjel za filozofiju

Diplomski sveučilišni studij filozofije; smjer: nastavnički (dvopredmetni)

## Funkcionalistička teorija i njezini nedostaci

Diplomski rad

Student/ica:

Mateja Andabak

Mentor/ica:

Doc. dr. sc. Nives Delija Treščec

Zadar, 2018.



## Izjava o akademskoj čestitosti

Ja, **Mateja Andabak**, ovime izjavljujem da je moj **diplomski** rad pod naslovom **Funkcionalistička teorija i njezini nedostaci** rezultat mogea vlastitog rada, da se temelji na mojim istraživanjima te da se oslanja na izvore i radove navedene u bilješkama i popisu literature. Ni jedan dio mogea rada nije napisan na nedopušten način, odnosno nije prepisan iz necitiranih radova i ne krši bilo čija autorska prava.

Izjavljujem da ni jedan dio ovoga rada nije iskorišten u kojem drugom radu pri bilo kojoj drugoj visokoškolskoj, znanstvenoj, obrazovnoj ili inoj ustanovi.

Sadržaj mogea rada u potpunosti odgovara sadržaju obranjenoga i nakon obrane uređenoga rada.

Zadar, 26. listopada 2018.

## Sadržaj

|                                                                     |    |
|---------------------------------------------------------------------|----|
| Uvod .....                                                          | 1  |
| 1. Povijesni razvoj problema duha i tijela .....                    | 3  |
| 2. Razlike između bihejviorizma, fizikalizma i funkcionalizma ..... | 6  |
| 3. Vrste funkcionalizma i njegovi predstavnici .....                | 9  |
| 4. Psihologija i problem redukcije .....                            | 12 |
| 4.1. Putnamova teorija interpretacije .....                         | 15 |
| 4.2. Kvalitativni i intencionalni sadržaj u funkcionalizmu .....    | 17 |
| 4.2.1. Teorija semantičkih svojstava i teorija sličnosti .....      | 19 |
| 5. Funkcionalizam i Turingov stroj .....                            | 20 |
| 6. Umjetna inteligencija .....                                      | 23 |
| 7. Drugi autori o odnosu mentalnog i fizičkog .....                 | 26 |
| 8. Nedostaci i prigovori funkcionalizmu .....                       | 28 |
| 8.1. Prigovori Williama G. Lycana .....                             | 28 |
| 8.2. Kalkeovi prigovori .....                                       | 30 |
| 8.3. Problemi s funkcionalizmom po Nedu Blocku .....                | 31 |
| 9. Zaključak .....                                                  | 34 |
| Popis literature .....                                              | 36 |
| Sažetak .....                                                       | 37 |
| Abstract .....                                                      | 37 |

## Uvod

Tema ovog diplomskog rada jest funkcionalistička teorija uma i njezini nedostaci. No, da bismo mogli objasniti što je to funkcionalizam moramo se barem osvrnuti na teorije koje su prethodile funkcionalizmu. Tako ćemo ukratko objasniti dualizam, bihejviorizam, fizikalizam te na koncu i funkcionalizam. Zatim ćemo se detaljnije baviti samim funkcionalizam te ćemo obraditi vrste funkcionalizma, njegove predstavnike i njihove teorije. Osim toga, dotaknut ćemo se i psihologije te problema redukcije, odnosno pozabavit ćemo se pitanjem može li se psihologija svesti na fiziku. Ova tema će nam poslužiti kao temelj za obradu Fodorovih i Putnamovih stajališta po pitanju funkcionalizma. Također, u ovom radu će se objasniti veza funkcionalizma i Turingovog stroja, odnosno strojne tablice. Samim time što se dotičemo Turingovog stroja otvara nam mogućnost da objasnimo i fizikalizam vrsta te fizikalizam instanci. Potonji je bitan upravo za funkcionalizam te će biti objašnjeno zašto je to tako. Često se neopravdano u pretjerano usku vezu dovode bihejviorizam i funkcionalizam pa ćemo pokušati u potpunosti razlučiti te dvije pozicije u filozofiji uma. No, glavni cilj nam je objasniti što je funkcionalizam u filozofiji uma, zatim koje su vrste funkcionalizma a isto tako i pokazati koji su nedostaci funkcionalizma. U tu svrhu prikazat ćemo najpoznatije prigovore funkcionalizmu koji su uglavnom formulirani u djelima Neda Blocka.

„Funkcionalizam je, dakle, teorija uma koja nastoji objasniti mentalna stanja i mentalne događaje funkcionalno, tj. pozivanjem na funkcije ili zadaće koje ispunjavaju svjesna stanja i svjesni događaji, kao i ostala mentalna stanja i mentalni događaji. Te su funkcije određene kauzalnim vezama između mentalnih stanja, tj. mjestima koja pojedina stanja zauzimaju unutar razgranate uzročno-posljedične mreže.“<sup>1</sup> Funkcionalizam se javlja tijekom šezdesetih godina prošlog stoljeća zahvaljujući upravo Putnamu i to kao reakcija na fizikalizam. Tada je već postojala teorija identiteta koja je bila suprotstavljena dualizmu i bihejviorizmu. Funkcionalisti poriču fizikalističku tezu da prava priroda mentalnih stanja jest u njihovoj neurološkoj osnovi. Funkcionalisti smatraju da prava priroda mentalnih stanja nije u njihovoj neurološkoj osnovi, već u njihovoj funkcionalnoj osnovi.<sup>2</sup> Iako funkcionalizam „službeno“ nastaje tek u 20. stoljeću, može se smatrati da je postojao još u antičkoj filozofiji zahvaljujući Aristotelovoj teoriji duše (duha). Aristotel je smatrao da je duša neodvojiva od tijela, dok je npr. Platon tvrdio da može postojati neovisno o tijelu, odnosno da može opstati i

---

<sup>1</sup> Janović Tomislav, Pećnjak Davor, *Nefunkcionalnost funkcionalizma*, In: Maja Hudoletnjak Grgić, Davor Pećnjak, Filip Grgić (ur.): *Aspekti uma*. Institut za filozofiju, Zagreb, 2011, str. 2

<sup>2</sup> Berčić Boran, *Filozofija – svezak drugi*, Ibis grafika d.o.o., Zagreb, 2012, str. 190

izvan njega. Za razliku od Platonove tvrdnje da duša može postojati izvan tijela, Aristotel je tvrdio da je (ljudska) duša oblik prirodnog, organiziranog ljudskog tijela - skup moći ili sposobnosti koje mu omogućuju da izraze "bitnu bitnost" koja je za Aristotela stvar ispunjenja funkcije ili svrhe, koja ga definira kao onakvog kakav jest. Osim toga, naznake funkcionalizma javljaju se u 17. stoljeću u djelima Hobbesa. On smatra da je razmišljanje zapravo neka vrsta računanja koje proizlazi iz mehanicističkih načela koje pak možemo usporediti s pravilim aritmetike. Ukratno rečeno, razmišljanje se može izvoditi putem sustava različitih fizikalnih tipova.<sup>3</sup>

---

<sup>3</sup> <https://plato.stanford.edu/entries/functionalism/>

## 1. Povijesni razvoj problema duha i tijela

Kako bi prikazali povijesni razvoj ovog problema poslužit ćemo se člankom *Problem um – tijelo* Williama G. Lycana. Lycan započinje svoj prikaz objašnjavanjem dualizma duha i tijela. Rješenje ovog problema putem dualizma prvi navodi Descartes. Descartes ga naziva Kartezijanskim dualizmom prema kojemu je um čisto duhovan. Ljudsko biće je dualitet uma i tijela. Isto kao što tijelo može postojati bez uma, tako i um može preživjeti eventualno uništenje tijela. Premda se čini da se poprilično razlikuju, um i tijelo se međusobno uvjetuju. Procesi koji se događaju unutar tijela uzrokuju osjete, a isto tako mentalna aktivnost uzrokuje tjelesne akcije ili pokrete. Sve do polovice prošlog stoljeća, ova teorija je dominirala filozofijim uma. Zatim se sve promijenilo iz dva razloga; akumulirani utjecaj logičkog pozitivizma i verifikacijska teorije značenja te brojni prigovori od strane drugih filozofa. Neki od prigovora su npr. da nam nematerijalni kartezijanski um nije zapravo potreban za objašnjenje već ionako vidljivih činjenica te da fizičko kretanje ovih nematerijalnih entiteta nije u skladu sa zakonima fizike.<sup>4</sup>

Prva teorija koja se suprotstavila dualizmu bio je bihejviorizam. Već su Carnap i Ryle uočili da su uvjeti i testovi za mentalna pripisivanja zapravo bili bihejvioralni. U filozofiji je bihejviorizam (naravno) poprimio metodološki oblik: uglavnom onaj analitičkog bihejviorizma, izraženog u tvrdnji da su mentalna svojstva jednostavno stvari koje se tiču bihejvioralnih odgovora na utjecaje iz okruženja.<sup>5</sup> Kad kažemo npr. nekoga nešto boli, to ne znači da je to nešto što se događa unutar te osobe nego da se ta osoba zaista ponaša na način koji iskazuje bol. Da bi se odgovorilo na pitanje može li stroj misliti, obavlja se Turingov test kojeg prihvaća svaki bihejviorist. Riječ je o testu koji je razvio Alan Turing još 1964. godine, odnosno o manipulaciji putem verbalnih testova. Posebno poglavlje bit će posvećeno upravo ovome. Bihejviorizam se, za razliku od dualizma, riješio problema odnosa duha i tijela tako što je prihvatio nematerijalne uzroke ponašanja. Međutim, ni bihejviorizam nije mogao proći bez kritika pa mu mnogi zamjeraju kako izostavljajući ono unutrašnje i subjektivno, zapravo izostavlja nešto vrlo stvarno i nadasve bitno. Prigovori su, ipak, u nekim točkama bili održivi. Primjerice ljudi koji su iskreni su zapravo savršeno svjesni svojih mentalnih pojava; zatim moguće je da se dva čovjeka, iako istog ponašanja, psihološki potpuno razlikuju; te na koncu,

---

<sup>4</sup> Lycan William G., *Problem um – tijelo*, ZINK Znanstveno-istraživački inkubator, Filozofski fakultet u Sarajevu FFSA, Sophos Volume 1, 2008, str. 108

<sup>5</sup> Ibid., str. 109



analize ponašanja mentalnih svojstava koje iznosi bihejviorizam čine se odgovarajućima samo ako pretpostavimo ostatak mentalnosti.

Zatim dolazimo do teorije identiteta koju pronalazimo kod Placea i Smarta. Oni smatraju da unutrašnja mentalna stanja i procese ne treba izjednačavati s vanjskim ponašanjem te da su mentalne pojave neurofiziološke, a ne duhovne kako su tvrdili dualisti. „One su identične sa stanjima i procesima koji se pojavljuju u centralnom živčanom sistemu onoga tko posjeduje ta mentalna stanja i procese; preciznije rečeno, svako metalno stanje ili proces je brojno identičan s nekim takvim neourofiziološkim stanjem ili procesom.“<sup>6</sup> Dakle, prema teoriji identiteta, mentalni stanja i procesi se uistinu pojavljuju u čovjekovom centralnom živčanom sustavu. Osim toga, ova teorija dokazuje da se organizmi mogu psihološki razlikovati bez obzira na sličnost njihovih vanjskih ponašanja. David Lewis i D.M. Armstrong su iznijeli direktan argument u korist teorije identiteta. Naime, oni su smatrali da se termini mentalnog definiraju kauzalno, u terminima uzroka i posljedice. Ovdje se ponovno dotičemo primjera boli za koju smatraju da zadržava svoju pojamovnu vezu s vanjskim ponašanjem, ali isto tako se podvrgava empirijskoj identifikaciji s unutrašnjim stanjima. Slijedi, dakle, da je bol ono neurofiziološko stanje koje ispunjava određeni prostor.

Nadalje, Hilary Putnam i Jerry Fodor su ukazali na to da postoji vjerojatnost da se teorija identiteta razumijeva kao vrsta mentalnog procesa; odnosno da neko mentalno stanje uvijek ima neurofiziološku karakterizaciju. Međutim, teoretičari identiteta su napravili grešku jer su se ograničavali na biologiju nekog bića koje je npr. sposobno osjetiti bol. Putnam i Fodor ukazuju upravo na potrebu ispravku ove greške. Prilikom osjeta boli nije važno postojanje „C-vlakana (per se)“ koji šalju impulse nego je važno ono što su ti impulsi proizveli, odnosno na koji način su doprinijeli da organizam funkcionira kao cjelina. Međutim, tipovi mentalnih stanja su zapravo više identificirani s „apstraktnim funkcionalnim ulogama nego s neurofiziološkim tipovima“. Stoga je funkcionalizam doktrina po kojoj ono što čini mentalno stanje stanjem određenog tipa, kao što je bol, miris ljubičice, vjerovanjem da su koale otrovne itd., jest njegov poseban skup funkcionalnih odnosa, njegova uloga u bihejvioralnoj ekonomiji njegovog subjekta.<sup>7</sup> Putnam uspoređuje mentalno stanje s funkcionalnim odnosno logičkim stanjem kompjutera. Kompjuterski programi mogu biti realizirani bilo kojom konfiguracijom hardwarea. Isto tako, psihološki programi mogu biti realizirani od strane različitih organizama koji pak imaju različite psihokemijske sastave. Ono

---

<sup>6</sup> Ibid., str. 111

<sup>7</sup> Ibid., str. 114

što funkcionalizam želi pokazati jest to da fizičko stanje u nečijem umu ima neurofiziološku deskripciju, ali da može imati i funkcionalnu deskripciju, kao i mentalni opis. Ukratko rečeno, želi pokazati da um ne mora biti strogo fizički.

Na koncu, dolazimo do funkcionalizma homunkulusa. Funkcionalizam strojeva je naišao na mnoge kritike te je dovelo do toga da je bilo potrebno vratiti funkciju u funkcionalizam. Ovo „vraćanje“ je započelo s Fodorom, Dennettom te Cumminsom koji su dali drugačiju perspektivu psihološkog objašnjenja. Oni su smatrali da se bihevioralni podaci trebaju promatrati kao manifestacije psihološkog kapaciteta subjekta, a ti se kapaciteti dalje objašnjavaju shvaćanjem subjekta kao sustava kojeg čine međusobno povezane komponente. Ove komponente su homunkulus svaka za sebe. Homunkulus je mali čovjek koji se nalazi unutra nekog subjekta i izvršava određene zadatke, odnosno identičan je s funkcijom koju obavlja. Homunkulusi se smatraju neurofiziološkim strukturama. Kriterij realizacije koji su koristili funkcionalisti, nije bio zadovoljavajuć. Naime, oni su smatrali da je veza između nekog individualnog fizičkog organizma i apstraktnog programa jednostavna stvar, odnosno da se događa jedan na jedan. Neki su predlagali da se ova greška ispravi uvođenjem teleološkog dodatka čime bi neko fizičko stanje organizma moglo realizirati funkcije ukoliko bi imalo istinsko organsko jedinstvo. Jednostavnije rečeno, činilo bi ono što mu je biološki određeno. Treći razlog navode Van Gulick, Milikan, Dretske, Fodor i mnogi drugi. Smatralo se da je teleologija ključ rješavanja problema te da se ona mora uvesti u svaku analizu intencionalnosti i to na način da se ukaže na psihobiološku funkciju mentalnih stanja. Franz Brentano, nadalje, uvodi razliku između psiholoških fenomena usmjerenih na objekte i fizičkih fenomena koji nisu usmjereni na takav način. Međutim, ovdje se postavlja pitanje kako neki čisti fizički entitet može uopće biti usmjeren na nešto ako su mentalna stanja fizička? Problem je u tome što čisti fizički entiteti ovo svojstvo ne mogu imati.

## 2. Razlike između biheviorizma, fizikalizma i funkcionalizma

Jerry Fodor u svom članku *Problem duha i tijela* daje odredbe biheviorizma i fizikalizma. Tradicionalne filozofije duha dijelimo na dualističke i materijalističke. Prve kažu da je duh nefizikalna supstancija, dok druge tvrde da ono mentalno nije različito od fizičkoga te da su sva mentalna stanja, procesi i operacije u načelu istovjetni. Bihevioristi tvrde da se sav govor o mentalnim uzrocima može eliminirati iz jezika psihologije u korist govora o okolinskim podražajima i ponašajskim odgovorima. S druge strane, filozofija duha koju nazivamo funkcionalizmom, a koja je nastala iz filozofskog razmišljanja u razvoju umjetne inteligencije, teorije kompjutacije, lingvistike, kibernetike i psihologije, priznaje mogućnost da bi tako različiti sustavi kao što su ljudi, računarski strojevi i bestjelesni duhovi mogli imati mentalna stanja. Međutim, da bismo uopće mogli razumjeti funkcionalizam, Fodor smatra da se prvo moraju sagledati nedostaci dualizma koji su i doveli do razvoja drugih teorija. Najveći nedostatak dualizma jest njegova nemoć da primjereno objasni kauzalnost. Naime, „kod dualizma nije evidentno kako bi nefizički duh dao povoda za bilo kakvu fizičku posljedicu bez povrede očuvanja zakona mase, energije i trenutka.“<sup>8</sup> Drugi nedostatak je taj što je dualizam nespojiv s praksom psihologije. Tako je primjerice John B. Watson dao radikalni prijedlog da ponašanje nema mentalne uzroke; on je promatrao ponašanje nekog organizma kojeg je proučavao kao opažljiv odgovor na podražaje, za koje je pak smatrao da su upravo uzroci tog ponašanja. Njegove misli dalje razvija B. F. Skinner za kojeg se smatra da je utemeljitelj pravca zvanog biheviorizam. Skinner smatra da je uloga psihologije bila ta da katalogizira zakone koji određuju uzročne odnose između podražaja i odgovora. Stvar je u tome da interakcije ovdje nema. Time dolazimo do radikalnog biheviorizma. Riječ je o filozofiji duha koja poriče postojanje duha i mentalnih stanja, procesa, svojstava i radnji. Argument protiv biheviorizma koji Fodor iznosi je sljedeći; psihologija ne ide putem biheviorizma nego je upravo suprotno. Početkom 60ih godina prošlog stoljeća nastaju dvije nove materijalističke filozofije duha koje ipak dopuštaju mentalne uzroke. Nastaju zbog sumnje u to da dualizam i radikalni biheviorizam iscrpljuju moguće pristupe filozofiji duha. Riječ je o logičkom biheviorizmu te teoriji istovjetnosti centralnog stanja. Logički biheviorizam je semantička teorija o tome što znače mentalni termini. Temeljna ideja ove teorije jest: pripisati mentalno stanje (npr. žeđ) organizmu isto je što i reći da je organizam sklon da se ponaša na određeni način (npr. da pije vodu ako mu je dostupna). Iz ovoga slijedi da svako mentalno pripisivanje po značenju je ekvivalentno nekoj ako-onda tvrdnji koja

---

<sup>8</sup> Fodor J.A., *Problem duha i tijela*, Računala, mozak i ljudski um, Izdavački centar Rijeka, Rijeka, 2001, str. 64

izražava dispoziciju za ponašanje. Po definiciji bihevioralni hipotetički sud ne uključuje nikakve mentalne termine. Antecedens hipotetičkog suda govori samo o podražajima, a konzekvens samo o ponašajskim odgovorima.<sup>9</sup> Logički biheviorizam zapravo samo inzistira na tome da se u načelu značenja bilo kojeg mentalnog termina može dati pomoću bihevioralnih hipotetičkih sudova. On postavlja mentalne termine na istu razinu s neponašajskim dispozicijama fizičkih znanosti. Na području mentalnog mnogi primjeri uzrokovanja događaja događajem sadrže slučaj kad neko mentalno stanje uzrokuje neko drugo, a logički biheviorizam ne nudi nikakvu analizu ove vrste uzrokovanja što rezultira time da se logički biheviorizam oslanja na pretpostavku da psihologija traži manje negoli fizikalne znanosti po pitanju uzročnosti. S druge strane, imamo teoretičara istovjetnosti koji tvrdi da mentalni termini referiraju na neurofiziološka stanja. Glavna im je prednost što uzimaju doslovno objašnjavateljske konstrukte psihologijske teorije. Ova teorija pokazuje kako mentalistička objašnjenja u psihologiji mogu biti doslovni izvještaji o uzročnoj povijesti ponašanja.<sup>10</sup>

Zatim dolazimo do fizikalizma. Općenito, fizikalizam bi bio filozofska pozicija koja tvrdi da je sve što postoji fizičke prirode. Odnosno, fizikalizam tvrdi da postoje samo oni entiteti koji su priznati od strane suvremene znanosti. Fizikalisti, međutim, ne tvrde da mentalno nikako ne postoji. Naprotiv, oni smatraju da mentalna stanja postoje, ali su samo vrsta fizičkih stanja. Mentalna stanja su prema fizikalistima jedna vrsta fizičkih stanja, odnosno riječ je o neurološkim stanjima naših mozgova. Mentalna stanja su u krajnjoj liniji samo fizička stanja. Fizičko i mentalno jedno su te isto. Relacija između mentalnog i fizičkog jest relacija identiteta. Za fizikalista stanja mozga ujedno jesu stanja svijesti. Jedna fizička stanja uzrokuju druga fizička stanja, budući da su neka fizička stanja ujedno i mentalna, može se reći da i mentalna stanja imaju uzročnu moć.<sup>11</sup> Ova teorija se upravo zbog ovoga naziva i teorijom identiteta. Prema fizikalizmu, sve što postoji može biti svedeno na fizičko, pa tako i ono mentalno. Fizikalisti također smatraju da je otkriće prave prirode mentalnih stanja u rukama znanosti te da je ona i odgovorna za takva otkrića. Znanstvena otkrića kao što su primjerice što je zapravo grmljavina ili H<sub>2</sub>O se nazivaju teorijske identifikacije ili aposteriorne identifikacije. Ovakva otkrića su ona u kojima znanost otkriva što neke pojave u svijetu oko nas doista jesu. To su sve samo činjenična pitanja koja zahtijevaju empirijska istraživanja. Međutim, fizikalist ipak nije spreman automatski isključiti dualizam jer fizikalist

---

<sup>9</sup> Ibid., str. 67

<sup>10</sup> Ibid., str. 70

<sup>11</sup> Berčić Boran, *Filozofija – svezak drugi*, Ibis grafika d.o.o., Zagreb, 2012, str. 163

smatra da je dualizam logički moguć, ali da je de facto neistinit. On smatra da je logički moguće da um bude nematerijalne prirode, ali je s druge strane uvjeren da mentalna stanja de facto jesu neurološka stanja. Ono što zaključujemo iz ovoga jest to da fizikalist smatra da je identitet mentalnog i fizičkog kontigentan i aposterioran, a ne nužan i aprioran. Postoje dvije vrste fizikalizma, a to su fizikalizam vrsta i fizikalizam instanci. Fizikalizam instanci tvrdi samo da su sve mentalne pojedinačnosti koje postoje neurofiziološke, a fizikalizam vrsta donosi mnogo općenitiju tvrdnju da su sve mentalne pojedinačnosti koje uopće mogu postojati neurofiziološke. Fizikalizam instanci ne isključuje logičku mogućnost da strojevi i bestjelesni duhovi imaju mentalna stanja dok fizikalizam vrsta odbacuje tu mogućnost budući da ni jedni ni drugi nemaju neurone.<sup>12</sup>

Na koncu dolazimo do funkcionalizma. Funkcionalizam je filozofija duha koja je temeljena na razlici koju kompjuterska znanost povlači između sistemskog hardwarea ili fizičkog ustrojstva i softwarea ili programa. Funkcionalizam ne isključuje mogućnost da mehanički i eterski sustavi imaju mentalna stanja i procese. S jedne strane, kaže Fodor, tj. kad se radi o uzročnoj naravi međudjelovanja duha i tijela, u pravu je teoretičar identiteta; a s druge strane kad je riječ o relacijskoj naravi mentalnih svojstava, u pravu je logički bihejviorist. Funkcionalizam može riješiti ovu dilemu naglašavajući razliku koju računarstvo prevodi između hardwarea i softwarea; dakle, on može osmisliti i uzročnu i relacijsku narav mentalnoga. Pojam uzročne veze omogućuje definiranje mentalnih stanja pomoću njihova uzročnog odnosa prema drugim mentalnim stanjima. Funkcionalizam tolerira materijalističku soluciju problema duha i tijela koju nudi teorija istovjetnosti, odnosno prihvaća da su mentalna svojstva tipično definirana u terminima njihova odnosa i to da je interakcija duha i tijela tipično uzročna. Strojevi npr. pružaju dobre primjere za dva pojma koji su središnji u funkcionalizmu, shvaćanje da se mentalna stanja međusobno definiraju i shvaćanje da ona mogu biti realizirana mnogim sustavima.<sup>13</sup>

Ukratko rečeno, razlike između ovih pravaca su sljedeće: radikalni bihejviorizam tvrdi da nema duha te da ne postoje mentalna stanja dok logički bihejviorizam dopušta mentalne uzroke tj. pripisivanje mentalnih stanja organizmima. Nadalje, teorija istovjetnosti tvrdi da mentalni termini referiraju na neurofiziološka stanja odnosno da mentalna objašnjenja mogu biti izvještaji o uzročnom ponašanju. Fizikalizam izjednačava mentalno s fizičkim, dok funkcionalizam smatra da postoji veza, neki određeni odnos između mentalnog i fizičkog.

---

<sup>12</sup> Fodor J.A., *Problem duha i tijela*, Računala, mozak i ljudski um, Izdavački centar Rijeka, Rijeka, 2001, 72-73

<sup>13</sup> Ibid., str. 74

Glavna razlika između biheviorizma i funkcionalizma je u tomu što prvi smatra da mentalno uopće ne postoji, a drugi tvrdi ne samo da postoji nego da postoji i veza između mentalnog i fizičkog. Razlika između fizikalizma i funkcionalizma jest ta da fizikalizam smatra da su mentalno i fizičko jednaki, odnosno jedna te ista stvar, dok ih funkcionalizam odvaja, ali smatra da su povezani određenim odnosom.

### 3. Vrste funkcionalizma i njegovi predstavnici

Od nastanka funkcionalizma kao pravca razvilo se nekoliko vrsta; i to zdravorazumski ili analitički funkcionalizam, referencijalni funkcionalizam i strojni funkcionalizam. Svaku od ovih vrsta ćemo pojedinačno objasniti te vidjeti daje li neka od njih prihvatljivo rješenje pitanja mentalnosti. Zdravorazumski funkcionalizam temelji se na činjenici da stvari u svijetu prepoznajemo i razlikujemo jedne od drugih upravo prema tome što one obavljaju odnosno prema tome kako to obavljaju i za što su namijenjene da obavljaju.<sup>14</sup> Na ovaj način bi se, smatraju zagovornici zdravorazumskog funkcionalizma, moglo objašnjavati i sam rad uma, odnosno područje mentalnosti. Da bismo shvatili kako bi ovaj način trebao funkcionirati, Janović i Pećnjak nam u svom članku *Nefunkcionalnost funkcionalizma*, daju primjer termostata. Termostat je, kao što svi znamo, naprava koja mjeri i regulira temperaturu u određenom prostoru, te funkcionira na način da isključuje ili uključuje neki drugi stroj s kojim je povezan. Opis ove namjene, odnosno ova njegova funkcija je ono što termostat razlikuje od ostalih stvari ili drugačije rečeno, čini termostat upravo termostatom. Upravo prema kriteriju namjene, mi možemo razumjeti mentalna, ali i općenito svjesna stanja. Budući da znamo što je npr. bol, žeđ, želja, itd., zahvaljujući našem zdravom razumu ili znanju, mi smo u stanju ova mentalna stanja shvatiti i kao funkcionalna. Isti autori ponovno objašnjavaju ovu tvrdnju na konkretnom primjeru. Kada traje dugo toplo ljeto, čovjek ožedni. Rečenicu koja izražava ulazni podatak za mentalno stanje žeđi možemo ovako specificirati: "vrućina i suhoća u okolini uzrokuje žeđ u čovjeka". Rečenicu koja izražava unutrašnja stanja specificiramo pak ovako: "žeđ uzrokuje želju da je utažimo" — ovo unutarnje stanje povezano je s vjerovanjem "žeđ se može utažiti uzimanjem osvježavajućeg pića". Općenita želja za utažavanjem žeđi i vjerovanje da osvježavajuće piće utažava žeđ, uzrokuju daljnju želju uzimanjem određenog

---

<sup>14</sup> Janović Tomislav, Pećnjak Davor, *Nefunkcionalnost funkcionalizma*, In: Maja Hudoletnjak Grgić, Davor Pećnjak, Filip Grgić (ur.): *Aspekti uma*. Institut za filozofiju, Zagreb, 2011, str. 2

osvježavajućeg pića.<sup>15</sup> No, smatraju spomenuti autori, ovdje dolazi do problema za funkcionaliste. Kad pogledamo gore navedeni primjer, vidimo da je cijeli primjer definiran međusobnim odnosima komponenata od kojih se sastoji; dakle, radi se o cirkularnosti koja možda i nije prihvatljiva. Nije prihvatljiva iz razloga što se teorijama koje sadrže cirkularnost ne može objasniti, odnosno one ne mogu biti temeljem za objašnjavanje samog načina funkcioniranja uma. Cirkularnost se, međutim, može i otkloniti i to na način da se preformuliraju definicije i međusobne veze. Preformulacija se izvodi tako da se sve implicitne definicije preformuliraju u eksplicitne. No, ovo samo otvara drugi problem funkcionalistima jer preformulacija može samo dodatno zakomplicirati objašnjenje uma i svijesti. S druge strane, potrebno je uvesti i logički jezik uz već postojeći jezik psihologije. Čak i ako bi uspjeli preformulirati rečenice, odnosno definicije i međusobne veze, i dalje ostaje neodgovoreno pitanje kako se neka određena vrsta intencionalnih djelovanja mogu formulirati bez pozivanja na mentalistički vokabular. Naime, ako želimo opisati određeno ponašanje kao izlazni podatak nekog sustava, onda to ponašanje u pravilu nije slučajno nego je racionalno, tj. usmjereno nekom cilju. Taj je cilj tipično mentalno reprezentiran odgovarajućom željom ili namjerom — željom da se ponašanjem postigne to-i-to — i vjerovanjem — vjerovanjem da će se čineći to-i-to postići to-i-to. Prema teoriji djelovanja, djelovanje je djelovanje upravo i samo u svjetlu navedenih intencionalnih stanja. To je ono što djelovanje razlikuje od pukih radnji, npr. nesvjesnih tjelesnih pokreta.<sup>16</sup> Pravi problem nastaju kad bi se pokušalo objasniti iskustvena stanja. Ovdje se kao primjer uzima sljedeći Fosterov misaoni eksperiment. Zamišljamo osobu koja je slijepa od rođenja, ali osobu koja je svjesna da postoje osobe koje posjeduju osjet vida i imaju mogućnost vidjeti boje. Ona, dakle, može razumjeti, da postoji doživljavanje fizičkog spektra boja, ali ona to nikada nije sama iskusila pa joj samim time nedostaje psihološki doživljaj. Foster smatra, da bi osoba slijepa od rođenja mogla, prema funkcionalizmu, znati kako izgleda biti u stanju vidjeti boje ukoliko bi otkrila uzročnu ulogu vizualnog iskustva u uzročno-posljedičnom sustavu. Osim toga, Foster smatra da se znanje može usvojiti na dva različita načina: jedan način bi bio da osoba koja vidi riječima dočara osobi koja je slijepa, znanje o vizualnosti; a drugi način bi bio da se slijepo osobi prenese znanstveno utvrđeno znanje. Međutim, osoba slijepa od rođenja ne bi mogla usvojiti iskustveno znanje što znači da funkcionalizam ne bi mogao objasniti psihološki karakter tog iskustva. Prema tome, funkcionalizam nije potpuna psihološka teorija.

---

<sup>15</sup> Ibid., str. 3

<sup>16</sup> Ibid., str. 6

Nadalje, navest ćemo primjer iskustva ili doživljaja boje. Osjetilno iskustvo zelene boje je, naravno, različito od iskustva crvene boje. Možemo zatim zamisliti da postoje dvije osobe koje su potpuno istovjetne u svim psihološkim funkcionalnim karakteristikama, ali imaju drugačija kvalitativna iskustva boja; kad jedna osoba ima kvalitativna iskustva crvenog druga ima zelenog i obrnuto. Pritom možemo zamisliti i da doživljaj boja odgovara doživljaju istovjetnih objekata. Na primjer, možemo zamisliti da i jedna i druga osoba vide zalazak Sunca i zastavu Narodne Republike Kine kao kromatski istovjetne prizore, a da su pritom unutarnje kvalitete (qualia) njihovih doživljaja međusobno potpuno različite. Nadalje, možemo zamisliti i da su obje osobe naučile zvati doživljaj boje zalaska Sunca i doživljaj boje kineske zastave "crveno", a doživljaj trave "zeleno".<sup>17</sup> Na temelju ovoga zaključujemo da se neće moći uočiti nikakva razlika koja bi nam pokazala da su iskustva boja ove dvije osobe i ovih dvaju subjekata zapravo različita. To nam govori da funkcionalizam nije sposoban adekvatno opisati osjetilna iskustva.

Zatim dolazimo do strojnog funkcionalizma. Ljudski um, odnosno rad ljudskog uma, se često znao uspoređivati s radom stroja. Ova teorija se razvila kad su se počela upotrebljavati računala. Računala rade po programima koji se mogu opisati prema tzv. dijagramu tijeka. Ova ideja o dijagramu tijeka potekla je od Dennetta. Dijagram tijeka sadrži ulazne i izlazne podatke, ali i unutrašnja stanja koja predstavljaju strojne obrade podataka po određenim pravilima. Analitički funkcionalizam se suočavao s nekoliko problema, no strojni funkcionalizam nailazi na još jedan po pitanju propozicijskih stavova. Kada propozicijske stavove prikažemo kao niz formalnih simbola koji postaju predmet formalnih operacija, ostaje nejasno na koji bi način simboli trebali izražavati njihov specifičan sadržaj. Budući da je dovoljno da se proces odvija prema formalnim karakteristikama simbola te formalnim uputama za njihovu transformaciju, sadržaj postaje izlišan — za njega u dijagramu tijeka kao da nema mjesta.<sup>18</sup> Propozicijski stavovi, dakle, imaju i neka određena semantička svojstva, dok se strojna analogija bavi samo sintaktičkim. Funkcionalisti su ovaj problem pokušali riješiti Turingovim strojem, odnosno strojnim tablicama. Funkcioniranje Turingovog stroja će kasnije biti dodatno pojašnjeno, no ovdje ćemo samo ukratko navesti do kakvog se zaključka došlo korištenjem Turingovog stroja. Turingov stroj se može realizirati na mnoštvo različitih načina i ovdje priroda realizacije uopće ne igra nikakvu ulogu; dakle, može biti realiziran i materijalno i duhovno. Samim time se dolazi do zaključka da se onda vječno pitanje o tome kako je naš um realiziran, postavljalo krivo; a pravo pitanje postaje pitanje autonomije uma.

---

<sup>17</sup> Ibid., str. 8

<sup>18</sup> Ibid., str. 12



No, ovo pitanje samo proširuje problem jer se dalje otvara pitanje izomorfности dvaju ili više sustava, koji funkcionalisti pokušavaju riješiti tako da žele pronaći način realizacije koji bi odgovarao opisu uma i svijesti te ničemu drugom. Logičan izbor bio je ljudski mozak. Prema toj pretpostavci, funkcionalni opis rada mozga onaj je specifični opis koji odgovara danm ograničenju. Međutim, ni to se rješenje nije pokazalo naročito kreativnim. Naime, pokazalo se da je moguće pronaći interpretaciju koja je istovjetna funkcioniranju ljudskog mozga, a kojoj ne odgovara nikakav um niti bilo što što ostvaruje svijest.<sup>19</sup> Ovaj protuprimjer daje Ned Block koji će biti detaljnije obrađen u poglavlju gdje će se raspravljati o nedostacima i prigovorima funkcionalizmu.

#### 4. Psihologija i problem redukcije

Prije nego što prijedemo na redukciju u psihologiji i na kraju krajeva, na funkcionalizam, ukratko ćemo prikazati problem same redukcije. Neven Sesardić u svojoj knjizi *Fizikalizam* pojašnjava da se još od vremena Talesa i Heraklita raspravljalo o tome može li se složeni svijet prikazati i objasniti jednom jednostavnom teorijom. Ova teza o postojanju jedne jedine fundamentalne znanosti povezana je s materijalističkom filozofijom, jer je u toj tradiciji fizika bila redovito smatrana ujedinjujućom teorijom na koju bi trebale biti svedene i u kojoj bi trebale „nestati“ posebne znanosti. Ujedinjene znanosti se najčešće zamišlja kako redukcija ostalih empirijskih disciplina na fiziku.<sup>20</sup> Neki od paradigmatskih primjera redukcije su redukcija fenomenološke termodinamike na statističku mehaniku ili redukcija optike na Maxwellovu elektromagnetsku teoriju. Neki filozofi smatraju da se u ovim slučajevima uopće ne može govoriti o redukciji, dok neki drugi pak smatraju da se takva redukcija ne može očekivati u biologiji ili u psihologiji. Do neslaganja po ovom pitanju dolazi zbog teškoća pri samom prikazivanju redukcije. Te teškoće se tiču dva formalna uvjeta redukcije, koje iznosi Ernest Nagel u svom djelu *Struktura nauke*. Ovdje je slučaj takav da u teoriji koju treba reducirati (T2) pronalazimo neke izraze koje nemamo u primarnoj teoriji (T1), odnosno bazi same redukcije. Da bi se redukcija izvršila, potrebno je te izraze koji se javljaju samo u T2 ili prevesti u jezik T1, ili pokazati da se oni u stvari odnose na iste predmete kao i neki

---

<sup>19</sup> Ibid., str 14

<sup>20</sup> Sesardić Neven, *Fizikalizam*, GRO Kultura, Beograd, 1984, str. 78

izrazi iz T1. To Nagel naziva uvjetom povezanosti.<sup>21</sup> Međutim, Jerry Fodor i Hilary Putnam su smatrali da se na ovaj način nikako ne mogu povezati termini psihologije i fizike. Drugi uvjet koji zahtijeva heterogena redukcija, nakon što su termini iz sekundarne teorije zamijenjeni terminima iz primarne teorije, jest taj da mora biti moguće sekundarnu teoriju logički izvesti iz primarne. Ovaj uvjet je nazvao uvjetom izvodivosti. Ovdje također nastaje problem jer je sekundarna teorija često samo dijelom istinita pa zapravo ne može biti logički spojena s primarnom teorijom. Jedino što možemo izvesti je ispravljena sekundarna teorija, koja više ne bi ni bila teorija nego samo nova konstrukcija. Sesardić se ovdje ukratko dotiče pitanja redukcije i emergencije u kemiji i biologiji, no mi ćemo se odmah prebaciti na to isto pitanje u psihologiji kako bi mogli doći do funkcionalizma. Glavno pitanje koje se ovdje postavlja jest je li uopće racionalno očekivati da se psihologija može reducirati na fiziku da bi se znanost ujedinila? Mnogi suvremeni filozofi smatraju da je nerealno očekivati takvo nešto te da se zapravo psihološki zakoni ne mogu svesti na zakone fizike. Već spominjani Fodor i Putnam iznose svoje funkcionalističke teze u prilog ovoj tvrdnji, ali i Donald Davidson sa svojim anomalističkim monizmom. Da bi se psihologija reducirala na fiziku, potrebno je uspostaviti korelaciju između psiholoških i neurofizioloških predikata. Za to nam je potreban i fizikalizam, odnosno fizikalizam instanci jer funkcionalizam smatra da se fizikalizam vrsta ne može održati. Zašto je to tako vidjet ćemo u sljedećem Fodorovom argumentu. Fodor, naime, smatra da psihologija ne razlikuje svoja stanja po strukturi ili materijalu u kojem se odigravaju već po funkciji koju izvršavaju: npr. pamćenje je pohranjivanje informacija, strah je stanje koje proizvodi tendenciju izbjegavanja opasnosti, itd. Tu nema riječi o materijalnom mediju tih stanja i o njihovim strukturalnim obilježjima.<sup>22</sup> Riječ je o tome da različita neurofiziološka stanja mogu izvoditi istu funkciju te je zbog toga besmisleno očekivati mogućnost redukcije.

Nadalje, Fodor i Putnam smatraju da se psihologija bavi ispitivanjem određenih funkcija; odnosno psihologiju zanima kako se neki organizam ponaša u određenim situacijama, a ne od čega je taj organizam napravljen. Psihologija se, prema tome, ne može svesti na neurofiziologiju jer je ona puno više općenitija od neurofiziologije. Psihologija se odnosi na funkcionalno izomorfne sisteme, bez obzira na to koliko su slični odnosno različiti po svojoj fizičkoj građi. Razliku između funkcionalnog i strukturalnog opisa možemo usporediti s razlikom između primjerice software i hardwarea. Da bi nam što jednostavnije to prikazao te na koncu i objasnio, Fodor nam daje jedan vrlo jednostavan primjer. Riječ je principu rada

---

<sup>21</sup> Ibid., str. 78-79

<sup>22</sup> Ibid., str. 84-85

automata za coca-colu. Njegova je funkcija da izbacuje konzervu coca-cola kada primi 10 centi. U njega se može ubaciti novčić od 5 i novčić od 10 centi. Automat se može nalaziti u dva različita stanja  $q_1$  i  $q_2$ , koja su definirana isključivo po tome kako reagiraju na dva moguća ulaza (inputs) – dvije vrste novčića. Evo implicitnih definicija tih dvaju stanja:

( $q_1$ ) Ako je ubačeno 5 centi i automat se nalazi u stanju  $q_1$  nema izlaza (output), a automat prelazi u stanje  $q_2$ . ako je ubačeno 10 centi i automat se nalazi u stanju  $q_1$ , on izbacuje coca-colu i ostaje u stanju  $q_1$ .

( $q_2$ ) Ako je ubačeno 5 centi i automat se nalazi u stanju  $q_2$ , on izbacuje coca-colu i prelazi u stanje  $q_1$ . ako je ubačeno 10 centi i automat se nalazi u stanju  $q_2$ , on izbacuje coca-colu i 5 centi i prelazi u stanje  $q_1$ .<sup>23</sup> Ovdje je, dakle, prikazano kako automat za coca-colu funkcionira, odnosno njegov software. Mi ne možemo znati kako on izgleda iznutra jer nigdje nije napisano, odnosno nije nam poznat izgled njegovog hardwarea. Ovaj program prikazuje sve moguće „fizičke instancijacije“ i upravo iz tog razloga ne može biti reduciran.<sup>24</sup> Cijela se zapravo funkcionalistička teza zasniva na Fodorovim i Putnamovim tvrdnjama da psihološkim vrstama odgovaraju fizički predikati koji pak nisu vrste. Upravo je ova tvrdnja ona koja onemogućava redukciju.

Donald Davidson u svom članku *Mental Events* brani fizikalizam instanci, ali na nešto drugačiji način nego što su to radili funkcionalisti. On svoj dokaz bazira na tri premise: principu kauzalne interakcije mentalnog i fizičkog; principu nomološke prirode kauzalnosti te anomalizmu mentalnog. Prva premisa nam govori da ipak postoje slučajevi kad se mentalni i fizički događaji nalaze u uzročno-posljedičnom odnosu. Druga tvrdi da ukoliko se dva događaja i nalaze u takvom odnosu, onda oni moraju pod nekim svojim opisom potpadati pod neki deterministički zakon. Davidson se najviše posvetio ovoj trećoj premisi gdje mu je cilj bio pokazati da nema takvih determinističkih zakona na osnovi kojih bi se mentalni događaji mogli predvidjeti ili objasniti. Davidson smatra da kauzalna veza mora postojati između određenih događaja bez obzira na to na koji način su ti događaji opisani. S druge strane imamo zakone koji pak imaju lingvističku narav i za njih je bitno upravo kako je nešto i opisano. Ono što je važno ovdje uočiti jest to da dva događaja koji su uzročno povezani mogu instancirati zakon ako su opisani na određeni način, a ako su opisani na neki drugi način to se ne mora nužno dogoditi. Budući da su događaji po Davidsonu mentalni po tome kako su opisani, tj. u terminologiji propozicionalnih stavova: vjeorvanja, želja, namjera, nadanja itd.,

---

<sup>23</sup> Ibid., str. 86-87

<sup>24</sup> Ibid., str. 87

logički je moguće da zakoni na kojima se zasnivaju pojedinačni slučajevi interakcije mentalnog i fizičkog uvijek uzimaju mentalne događaje ne pod njihovim mentalnim, nego pod nekim drugim opisom.<sup>25</sup> Davidson, nadalje, smatra da u psihologiji možemo naći samo statističke i grube generalizacije koje nikako ne mogu postati striktnim zakonom. U mentalnoj sferi vlada holizam gdje su uvjerenja i želje u ovisnosti od ostalih mentalnih stanja. Upravo ovaj holizam uspostavlja anomalizam mentalnog. Vratimo se ukratko na Davidsonovu drugu premisu. Ukoliko se ona prihvati, onda slijedi da mentalni događaji jednostavno moraju imati nekakav opis pod kojim će se instancirati striktni zakoni. To bi, naravno, morao biti njihov fizički opis i time se dalje uspostavlja anomalizam – tvrdnja da je svaki pojedinačni mentalni događaj identičan s nekim fizičkim događajem, bez obzira na činjenicu da ne postoji nikakav zakon koji bi ove dvije sfere uopće mogao dovesti u vezu.

#### 4.1. Putnamova teorija interpretacije

Ovdje nije riječ samo o Putnamovoj teoriji interpretacije nego također i o računarskoj psihologiji. Putnam smatra da se psihologija bavi ispitivanjem određenih funkcija, apstrahirajući potpuno od fizičkih i strukturalnih osobina onih stanja kroz koja organizam prolazi vršeći razne operacije. Jedna misao kao što je na primjer „trava je uokolo“ je opća misao o običnim fizičkim stvarima. Prema računalnom modelu uma, um posjeduje „program“ ili skup pravila analognih pravilima koje upotrebljava računalo. Mišljenje tada uključuje manipulaciju riječima i drugim znakovima.<sup>26</sup> Razlog zašto to ne može biti kako je i opisano je taj što takav program nije specifičan. Putnamova teza je da um misli uz pomoć reprezentacija. Ova teza za sobom povlači još dvije ideje. Prva ideja je da mišljenje ne uključuje samo manipulaciju simbolima ili arbitrarnim predmetima nego zahtijeva manipulaciju simbolima koji imaju vrlo specifičnu strukturu i to onu formalnog jezika. Druga ideja je takva da smatra da ljudski um misli konstruirajući neku vrstu „modela“ svoje okoline koji ne mora biti jednak stvarnom svijetu. Um dakle, upotrebljava formalni jezik istovremeno kao medij kompjutacije i medij reprezentacije. Međutim, um može upotrebljavati više od jednog formalnog jezika jer različiti dijelovi mozga mogu kompjutirati u različitim „medijima“. Tzv. najbolji slučaj je kada um svu kompjutaciju vrši u jednom formalnom jeziku. Razumijevanje uma njegovog vlastitog „medija kompjutacije i reprezentacije“ sastoji se u njegovom posjedovanju

---

<sup>25</sup> Ibid., str. 98

<sup>26</sup> Putnam Hilary, *Računarska psihologija i teorija interpretacije*, Računala, mozak i ljudski um, Izdavački centar Rijeka, Rijeka, 2001, str. 154

verifikacijske semantike za medij. Ova verifikacija zapravo predstavlja kompjutabilne predikate koji mogu onda predstavljati prihvatljivost, zajamčenu potvrđljivost ili vjerodostojnost.

Putnam smatra da se jezik interpretira preko skupa pravila koja označavaju stupnjeve konfirmacije, a ta pravila moraju biti kompjutabilna. Ovdje se postavlja pitanje zašto se koristi verifikacijska semantika, a ne na primjer teorija značenja? Razlozi zašto se ne koristi teorija značenja su ti što ona nije psihologija; zato što ne možemo uspoređivati reprezentacije i nekonceptualiziranu stvarnost te taj što ne možemo dati bilo kakvu teoriju o tome u čemu se sastoji upotreba neke rečenice jer je to posao upravo verifikacijske semantike. Verifikacijska semantika sadržava pojam vjerovanja te je to vjerovanje ono što ju čini kognitivnom. Istovremeno je ona i računalna što ju opet čini funkcionalističkom. Nadalje, model semantike mora uključiti i funkciju korisnosti koja mora biti kompjutabilna, ali isto tako mora uključivati i funkciju stupnja konfirmacije.

Putnam ovo objašnjava na primjeru dvoje ruritanske djece. Zamislimo da negdje na svijetu postoji zemlja zvana Ruritanijska. Zamislimo da u toj zemlji postoje male razlike između dijalekata koji se govore na sjeveru i jugu. Jedna je od tih razlika značenje riječi „grag“: na sjevernom dijalektu znači srebro, a na južnom aluminij. Zamislimo, također, dvoje djece, Oskara i Elmera, koji odrastaju u Ruritanijskoj. Oskar odrasta na sjeveru, a Elmer na jugu. Zamislimo da su na sjeveru zdjele i tave normalno napravljene od srebra dok su na jugu normalno izrađene od aluminija. Možemo pretpostaviti da Oskar i Elmer imaju istu mentalnu reprezentaciju „graga“, da imaju ista vjerovanja povezana s „gragom“ itd. (...) Razlika u ekstenziji konstituira razliku u značenju te ako je ekstenzija fiksirana kolektivno, a ne individualno, proizlazi da značenje riječi „grag“ u dva ideolekta, nije isto čak i ako ne postoji ništa „psihološko“, ništa „u njihovim glavama“ što čini razliku u značenju. Značenje nije u glavi.<sup>27</sup> Da bismo mogli odrediti jesu li određene stvari sinonimi, moramo obratiti pozornost na dvije stvari; ekstenziju i korelirane stereotipe. Pod stereotipima se misli na određeni skup vjerovanja koje bi trebali imati svi govornici u vezi neke određene riječi (npr. tigar ima pruge). Ovdje Putnam daje usporedbu s ruritanskom djecom. Ovoga puta zamislimo da Oskar živi u eksperimentalnoj istraživačkoj stanici vrlo blizu Sjevernog pola. Zamislimo da u ovoj zajednici, iz nekog razloga, nema zelenih biljki, ali zato postoje umjetne biljke i čak umjetna trava. Razmotrimo sada riječ „trava“ u Oskarovu ideolektu i ideolektu običnog američkog

---

<sup>27</sup> Ibid., 159

djeteta koje ćemo zvati Elmer. Ako uzmemo da stereotip trave jednostavno perceptivni stereotip, tada će taj stereotip biti potpuno isti u Oskarovu kao i u Elmerovu umu.<sup>28</sup>

Sami stereotipi su vjerovanja izražena riječima te nam teorija stereotipa može poslužiti kao pomoćno sredstvo pri prevođenju nekih drugih riječi ili reprezentacija. Ono pri čemu nam ona ne može pomoći je definiranje pojma istovjetnosti sadržaja za riječi ili mentalne reprezentacije. Značenje riječi je funkcija različitih perceptivnih prototipova, ali također i više ili manje apstraktnih vjerovanja. Dva govornika mogu imati isto značenje za neku riječ, ali različite perceptivne prototipove koji su povezani s tom riječju.

Vratimo se sad na odluku o sinonimima. Na temelju svega dosad napisanog, ta odluka je postala nemogućom. Zašto? Zbog gledanja na ekstenziju, odnosno zbog „izdvajanja“ razlika u samoj ekstenziji. Kako onda možemo nešto uopće interpretirati? Spomenut ćemo sada neke druge autore koji su se bavili ovim pitanjem. Donald Davidson i Willard Quine su smatrali da je interpretacija bitni holistički pothvat. Za njih interpretirati jezik pretpostavlja pronalaženje prijevodne sheme koja se da naučiti i najznačajnije, koja je takva da smo kod interpretacije govornika stranog jezika sposobni „razumjeti“ njihove namjere, vjerovanja i ponašanje.

#### 4.2. Kvalitativni i intencionalni sadržaj u funkcionalizmu

Ranije smo spominjali kako je Fodor na vrlo jednostavan način htio prikazati kako se razlikuju funkcionalni i strukturalni opis. Fodor tu razliku prikazuje na primjeru automata za coca – colu. No, u ovom poglavlju se nećemo baviti tom razlikom nego ćemo objasniti još dva pojma koja su jako bitna za funkcionalizam, a koje Fodor u svom članku *Problem duha i tijela* također obrađuje. Riječ je o kvalitativnom i intencionalnom sadržaju mentalnih stanja. Gledište da mentalna stanja imaju ova dva sadržaja je zapravo tradicionalno gledište. Fodor u članku objašnjava što su to kvalitativni i intencionalni sadržaj, kakav oni značaj imaju za funkcionalizam ili pak predstavljaju li neki problem za funkcionalističku poziciju.

Kvalitativni sadržaj ne može biti spoznat pomoću opisa nego samo iskustvom. Zbog toga nam je zapravo vrlo teško reći i samim time zaključiti što bi ovaj sadržaj trebao predstavljati i na koncu, što bi on uopće trebao biti. Ovim primjerom Fodor ga pokušava opisati. Zamislimo da gledamo u prazni zid kroz crveni filter. Zamijenimo sada filter zelenim, a sve ostalo ostavimo isto kao što je i bilo. Nešto će se u naravi vašeg iskustva primijeniti kad se filter

---

<sup>28</sup> Ibid., 162

promijeni, a to je upravo ona vrsta stvari koju filozofi zovu kvalitativnim sadržajem.<sup>29</sup> Funkcionalizam je obavezan definirati mentalna stanja u terminima njihovih uzroka i njihovih posljedica. Na temelju toga, čini se da možemo zaključiti da dva mentalna stanja mogu imati iste uzročne relacije, a istodobno se razlikovati upravo po ovom kvalitativnom sadržaju. Zašto je ovaj sadržaj problem za funkcionalizam, Fodor objašnjava na primjeru obrnutog spektra. Moguće je zamisliti dva promatrača koja su slična u svim relevantnim psihološkim pogledima osim što je iskustvo koje za jednog ima kvalitativni sadržaj crvenog za drugog zeleno. Ništa u njihovu ponašanju ne treba pokazivati razlike jer oba vide zrelu rajčicu i jarkocrveni izlazak sunca kao slične po boji i oba tu boju zovu „crveno“. Štoviše, kauzalne veze između njihovih iskustava i drugih mentalnih stanja bit će također identične. Moguće je da oba misle o Crvenkapici kad vide zrelu rajčicu i imaju osjećaj tuge kad vide zelenu boju. Čini se da će oni dijeliti sve ono što je zahvaćeno pojmom kauzalne uloge njihovih iskustava, a da će se ipak kvalitativni sadržaj njihovih iskustava razlikovati. Ako je to moguće, tada se funkcionalističko objašnjenje neće moći primijeniti na mentalna stanja koja imaju kvalitativni sadržaj.<sup>30</sup> Možemo, dakle, pretpostaviti da je posjedovanje kvalitativnog sadržaja glavni faktor onog što mentalna stanja zapravo i čini svjesnima. Problem je jedino u tome što se to ne može objasniti.

S druge strane imamo intencionalni sadržaj. Intencionalni sadržaj nam pruža određena semantička svojstva. Ovdje imamo primjer Galilea i Enricoa. Ako netko kaže „Galileo je bio Talijan“, tada njegov iskaz, kao Enricovo vjerovanje, izražava propoziciju o Galileu koja je istinita ili lažna, ovisno o Galileovoj domovini. Ova paralela između simboličkog i mentalnog ističe tradicionalni zahtjev o jedinstvenom tretiranju jezika i uma.<sup>31</sup> Ono što nam je važno jest to da moramo razumjeti ova svojstva jer se u teorijama kognitivne znanosti najčešće radi o vjerovanjima koje jedan organizam posjeduje ili može posjedovati. Međutim, postoje i neke vrste stvari koje nisu mentalna stanja, a ipak imaju intencionalni sadržaj. Riječ je, naravno, o simbolima. Ovi simboli su o stvarima. Pretpostavka je da postoje mentalni simboli i da oni imaju semantička svojstva. Posjedovanje vjerovanja uključuje stanje relacije prema mentalnom simbolu i ta vjerovanja nasljeđuju svoja semantička svojstva od mentalnih simbola. Udruživanje semantičkih svojstava mentalnih stanja s onima mentalnih simbola je kompatibilno s računarskom metaforom. Misliti o nekom sustavu kao o računalu ili

---

<sup>29</sup> Fodor Jerry, *Problem duha i tijela*, Računala, mozak i ljudski um, Izdavački centar Rijeka, Rijeka, 2001, str. 79

<sup>30</sup> Ibid., str. 79-80

<sup>31</sup> Ibid., str. 81

kompjutoru znači postaviti pitanje o prirodi koda u kojem on kompjutira i o semantičkim svojstvima simbola u tom istom kodu. Analogija između uma i računala implicira postuliranje mentalnih simbola.

#### 4.2.1. Teorija semantičkih svojstava i teorija sličnosti

No, nije sve ipak toliko jednostavno. Naime, mnogi filozofi su smatrali da nešto mora biti pogrešno u ovoj tvrdnji da je upravo semantičko ono što čini relaciju između misli i onoga o čemu je misao relacija sličnosti. Ovdje imamo za primjer rečenicu *Ivan je visok*. Misao je istinita ako i samo ako je Ivan zaista visok. Time dolazimo do dvije teorije; teorije semantičkih svojstava i teorije sličnosti. Prva teorija tvrdi da bi onda bilo potrebno objasniti kako je jedna misao povezana s određenim, posebnim stanjem stvari. Druga pak teorija tvrdi da mišljenje misli sadrži posjedovanje mentalne slike koja pokazuje Ivana koji je visok. Ovdje, međutim, nije jasno što misao o njegovoj visini razlikuje od npr. misli o položaju njegovog tijela ( ne znamo stoji li, sjedi li, leži li, itd). Sama ideja da su semantička svojstva mentalnih reprezentacija određena aspektima svojih funkcionalnih uloga jest dovoljan uvjet za posjedovanje ovih semantičkih svojstava te se ona može specificirati u uzročnim terminima. Funkcionalist prepoznaje tri tipa uzročnih odnosa između psiholoških stanja koja uključuju prije navedene mentalne reprezentacije: - Ivan je visok. → vjerovanje je ovdje normalna posljedica određenih stimulacija kao što je viđenje Ivana u okolnostima koje nam jasno pokazuju njegovu visinu; - vjerovanje je normalni uzrok određenih posljedica u ponašanju kao što je samo izricanje rečenice „Ivan je visok“; - vjerovanje je normalni uzrok, ali isto tako i normalna posljedica nekih drugih vjerovanja. Sve ovo nas dovodi do toga da propozicija izražena danom mentalnom reprezentacijom ovisi o kauzalnim svojstvima mentalnih stanja u kojima se upravo ta mentalna reprezentacija uopće i javlja.



## 5. Funkcionalizam i Turingov stroj

Funkcionalizam je gledište da se razumijevanje mentalnog sastoji upravo u razumijevanju njegovog funkcionalnog ustrojstva. Osim toga, funkcionalizam u filozofiji uma smatra da mentalna stanja treba shvatiti kao funkcionalna stanja organizama. Da bi uopće znali nešto više o mentalnim stanjima ili o nekom određenom mentalnom stanju, moramo znati koja je njegova funkcija koja sudjeluje u funkcioniranju cjeline malog organizma. Dakle, za mentalno stanje je bitno koju funkciju ono obavlja u cjelini te njegove uzročne veze s vanjskim podražajima koji dovode do njega; s drugim mentalnim stanjima te sa tjelesnim ponašanjem organizma do kojeg ono dovodi.

Kao što već znamo, funkcionalizam se često uspoređuje s biheviorizmom iako su to dva različita pravca. Prema funkcionalizmu predmet proučavanja jesu i unutrašnji psihološki mehanizmi koji za biheviorizam predstavljaju irelevantni ili čak nepostojeći sadržaj „crne kutije“. Funkcionalistička analiza uspjela je zahvatiti ono što bihevioristička nije, a to je intuicija da naše reakcije ne ovise samo o podražajima koje primamo već i o mentalnim stanjima koje imamo – željama, vjerovanjima, preferencijama, itd.<sup>32</sup> Ono što funkcionaliste zapravo zanima, jest ustrojstvo našeg softwarea, a ne izgled našeg hardwarea. Funkcionalisti smatraju da su funkcionalna stanja našeg uma fizički realizirana u neurološkim stanjima našeg uma, dok se s druge strane ne slažu s tezom da vrste neuroloških stanja moraju odgovarati vrstama mentalnih odnosno funkcionalnih stanja. Po ovome možemo zaključiti da funkcionalisti odbacuju fizikalizam vrsta. „Fizikalizam vrsta tvrdi sljedeće: za svaki psihološki predikat F postoji neki neurofiziološki predikat G, takav da se neki organizam nalazi u stanju F ako i samo ako se nalazi u stanju G. Fizikalizam instanci za razliku tvrdi: za svaki psihološki predikat F, neki organizam se nalazi u stanju F ako i samo ako se nalazi u neurofiziološkom stanju ove ili one vrste. U prvom slučaju se tvrdi identičnost psiholoških i neurofizioloških svojstava, a u drugom samo identičnost pojedinačnih psiholoških stanja s pojedinačnim neurofiziološkim stanjima. Fizikalizam vrsta je očigledno mnogo jača verzija i on povlači istinitost fizikalizma instanci, dok obratno ne biva.“<sup>33</sup> Funkcionalisti, dakle, prihvaćaju slabiju verziju fizikalizma i zbog toga se često govori o „višestrukoj mogućnosti realizacije“. Riječ je o ideji da jedno te isto mentalno stanje može biti realizirano na više različitih načina. Fizikalizam instanci je za funkcionalizam važan upravo iz ovog razloga, ali i

---

<sup>32</sup> Berčić Boran, *Filozofija – svezak drugi*, Ibis grafika d.o.o., Zagreb, 2012, str. 189

<sup>33</sup> Sesardić Neven, *Fizikalizam*, GRO Kultura, Beograd, 1984, str. 84

zato što otvara put ka psihologiji. Ukoliko bi se neko određeno mentalno stanje uvijek moralo realizirati u nekom određenom neurološkom stanju, onda bi se zakoni psihologije zaista svodili na zakone u neurologiji, i to u potpunosti.

Nadalje, funkcionalisti pomoću Turingovog stroja pokušavaju objasniti funkcioniranje ljudskog uma. „Turingov stroj je apstraktni stroj čije je ponašanje u potpunosti određeno i to na sljedeći način: Ako je stroj u stanju S1, i ako primi input I, onda će dati output1 i prijeći u stanje S2. U svom najjednostavnijem obliku Turingov stroj sastoji se od čitača položenog na traci koja je podijeljena na polja, a neka su polja označena znakovima. Čitač se kreće jedno mjesto ulijevo ili udesno, ovisno o tome u kakov se stanju nalazio i o tome kakav je znak pročitao. Čitač može upisati znak u prazno polje ili obrisati upisani znak. Nakon obavljanje funkcije čitač može ostati u istom stanju ili prijeći u neko drugo stanje. I tako u nedogled.“<sup>34</sup> Iz navedenog možemo vidjeti da se funkcionalni ustroj stroja sastoji od tri parametra: ulazne informacije; stanja stroja te izlazne informacije. Iste ove parametre možemo usporediti i sa ljudskim ponašanjem: ulazna informacija = vanjski podražaj; stanje stroja = unutrašnje mentalno stanje; izlazna informacija = tjelesno ponašanje. Između funkcionalnih ustroja Turingovog stroja te ljudskog uma vlada potpuni izomorfizam. Ponašanje ljudskog uma možemo opisivati na isti način na koji smo gore opisali ponašanje Turingovog stroja. Potpun popis ponašanja za neku osobu predstavlja potpuni opis njenog funkcionalnog ustroja. Upravo bi taj opis trebao i biti cilj psihološkog istraživanja.<sup>35</sup>

Turing je također vjerovao u postojanje umjetne inteligencije te je predložio uvjet koji bi stroj morao zadovoljiti kako bi mu se mogla pripisati umjetna inteligencija. Ovo se naziva Turingovim testom. „Zamislamo da preko ekrana ili pisanih poruka komuniciramo s nekim ili nečim što se nalazi u susjednoj sobi. Mi ne znamo je li to kompjuter ili živ čovjek. Ako bismo na osnovi razmijenjenih poruka zaključili da komuniciramo s inteligentnim bićem, onda bismo tome što se nalazi u susjednoj sobi, ma što to bilo, trebali pripisati inteligenciju. Ako je stroj dovoljno složen da kod nas uspije stvoriti dojam da je inteligentan, onda on doista jest inteligentan.“<sup>36</sup> Na ovaj način funkcionalizam omogućava pripisivanje inteligencije strojevima. Iako možda izgleda komplicirano, sama ideja je zapravo vrlo jednostavna. Posjedovanje mentalnih stanja i svjesnih iskustava se svodi potpuno na posjedovanje određene vrste funkcionalnog ustrojstva. Prema tome, bilo što što ima jednak funkcionalni ustroj ima i jednaka mentalna stanja te također i svjesna iskustva. Ovaj stav se naziva tzv.

---

<sup>34</sup> Berčić Boran, *Filozofija – svezak drugi*, Ibis grafika d.o.o., Zagreb, 2012, str. 192

<sup>35</sup> Ibid., str. 192

<sup>36</sup> Ibid., str. 192-193

„principom funkcionalne invarijantnosti“ koji glasi ovako: Funkcionalni ustroj u potpunosti određuje svjesno iskustvo.<sup>37</sup> Međutim, ovdje nastaje problem. Mi znamo da mehanički sistemi koji su funkcionalno izomorfni našoj psihologiji ne mogu imati mentalna stanja. Ovaj problem se naziva problem odsutnih qualia. Qualia su subjektivne kvalitete naših mentalnih stanja te se one ne mogu kvantificirati. Qualia je nemoguće verbalno opisati; njih se jednostavno treba doživjeti.

Na koncu, kao primjer Turingovog testa možemo uzeti Searlov misaoni eksperiment koji on naziva kineska soba. “. Searle upita nekoga da zamisli nekoga tko ne razumije kineski i tako manipulira kineskim simbolima u skladu s pravilima iskazanim jezikom koji razumije. Zamislite pravila dovoljno spretno sastavljena tako da je ta osoba u stanju nastaviti korespondenciju na kineskom – odgovarajući na (pisana) kineska pitanja (pisanim) kineskim odgovorima na način koji se ne može razlikovati od izvedbe govornika čiji je to materinski jezik. Jasno, premda korespondent ne mora biti u stanju otkriti tu činjenicu, sam manipulator simbola ne razumije kineski. Razumjeti kineski nije samo stvar manipuliranja značenjskim simbolima na neki prikladan način. Ti simboli moraju nešto značiti za sustav koji izvodi operacije.<sup>38</sup> Ono što se ovim primjerom želi pokazati je da strojevi ne rade doslovno isto što i mi radimo. Tek kada stroj bude mogao znati značenje simbola kojim manipulira, moći će mu se pridodati oznaka inteligencije; tada će strojevi moći izvoditi radnje poput nas.

---

<sup>37</sup> Ibid., str. 193

<sup>38</sup> Searle John, *Umovi, mozgovi i programi*, Računala, mozak i ljudski um, Izdavački centar Rijeka, Rijeka, 2001, str. 91

## 6. Umjetna inteligencija

Budući da smo se u prethodnom poglavlju dotaknuli umjetne inteligencije, odnosno pitanja može li se stroju pripisati inteligencija, obradit ćemo članak Johna R. Searle o ovoj temi. Može li stroj misliti? John R. Searle se u svom članku „Umovi, mozgovi i programi“ pita upravo to. Iz ovog članka proizlazi da samo strojevi i mogu misliti. Riječ je o posebnoj vrsti strojeva koja posjeduje unutarnju uzročnu moć koja je jednaka onima u mozgu. Searle razlikuje „jaku“ i „slabu“ umjetnu inteligenciju kako bi mogao pripisati psihološki ili filozofski značaj računalnim simulacijama ljudskih spoznajnih sposobnosti. Prema slaboj umjetnoj inteligenciji (kasnije u tekstu UI), računalo je zapravo moćan alat koji nam omogućuje da što preciznije formuliramo i testiramo određene hipoteze. S druge strane imamo jaku UI prema kojoj računalo nije tek puki alat u proučavanju uma. Dobro programirano računalo jest um u smislu da kad ugradimo i podesimo primjereni program u to isto računalo, ono doslovno može razumjeti i posjedovati neka druga kognitivna stanja. Ovaj članak se zapravo više odnosi na jaku UI jer Searle ništa ne prigovara onoj slaboj. Usredotočuje se na tvrdnju da prikladno programirano računalo može doslovno imati kognitivna stanja te da programi ugrađeni unutar njega uistinu objašnjavaju samu ljudsku spoznaju.

Searle se dalje bavi radom Rogera Schanka i njegovih suradnika. Cilj Schankovog programa je bio simulirati ljudsku sposobnost razumijevanja priča. Za ljudsko razumijevanje priča je karakteristično da su ljudi zapravo sposobni odgovoriti na pitanja, iako ponekad neke informacije i nisu izričito dane u samoj toj priči. Schank nam daje sljedeći primjer:

Jedan je čovjek ušao u restoran i naručio hamburger. Kada ga je dobio, hamburger je bio spržen te je čovjek ljutito napustio restoran ne plativši račun i ne ostavivši napojnicu. Ako vam netko postavi pitanje je li taj čovjek pojeo hamburger, vaš odgovor će vjerojatni biti negativan. U drugoj priči imamo drugačiju situaciju. Čovjek je također ušao u restoran te naručio hamburger. Kada ga je dobio, zadovoljno je napustio restoran i platio račun te ostavio napojnicu. Na pitanje je li drugi čovjek pojeo hamburger, vjerojatno ćemo odgovoriti potvrdno. Ono što Schank želi dokazati jest da njegovi strojevi na ovaj isti način mogu davati odgovore na ista odnosno slična pitanja. Međutim, da bi uopće mogli odgovoriti na bilo kakvo pitanje, oni moraju imati „reprezentaciju“ o onoj vrsti informacije kojom ljudi raspolažu, tj. na određeni input (koji je u ovom slučaju bio priča) računala su u stanju dati odgovore, odnosno output na priču na isti onaj način na koji bi to učinili sami ljudi. Zagovornici jake UI

smatraju da, pored toga da računalo simulira ljudsku sposobnost, i to da stroj doslovno razumije priču te da nudi odgovore na postavljena pitanja te da stroj svojim programom uistinu objašnjava ljudsku sposobnost razumijevanja priča i odgovaranja pitanja o njima. Ovaj program možemo provjeriti već prije spomenutim eksperimentom kineske sobe. Na svaki upit koji smo dobili na kineskom jeziku mi pravilno obavljamo operaciju odgovaranja na iste onako kako nam upute nalažu. Nitko iz tih naših odgovora ne može zaključiti da mi zapravo nismo govornici kineskog jezika, odnosno da ne znamo pisati na kineskom. Upravo na tom principu, smatra Searle, funkcioniraju naša računala, a samim time i Schankov program. Oni, dakle, provode informacije koje zapravo ne razumiju, tj. na određeni input prema danim uputama (programima) daju određeni output. S druge strane, zagovornici jake UI smatraju da programirana računala razumiju priče i raznjašnjavaju ljudsko razumijevanje. Istina je da takav računalni program funkcionira ispravno i dobro, ali mi na temelju onoga što nam je poznato, možemo samo zaključiti da taj program manipulira simbolima koje ne razumije. Sama operacija nije dovoljna za razumijevanje.

Searle želi također objasniti što to „razumijevanje“ zapravo znači. Razumijevanje znači posjedovanje mentalnih (intencionalnih) stanja te njihovu istinitost (valjanost i uspješnost). Kada npr. kažemo „Vrata znaju kada se otvore zahvaljujući fotoelektričnoj ćeliji“, mi zapravo ne mislimo da vrata zaista posjeduju neko razumijevanje već im samo produžujemo vlastitu intencionalnost. Predikat „razumjeti“ koristimo samo u metaforičkom smislu jer razumjeti mogu smo ona bića koja zaista posjeduju mentalna stanja, odnosno intencionalnost.

Naravno, na sva ova njegova opažanja, morali su se pojaviti i neki odgovori. Tako ovdje spadaju sistemski odgovor, robot-odgovor, odgovor simulatora mozga, kombinacijski odgovor, odgovor drugih umova te odgovor mnogih tvoraca. Mi ćemo obraditi samo prva tri.

Sistemski odgovor je stigao s Berkleya. Berkleyevci smatraju da je činjenica da pojedinac koji je zaključan u sobu i ne razumije priču (ukoliko je to istina) jest dio sistema, a sam sistem razumije priču. Dakle, oni razumijevanje pripisuju čitavom sistemu, a ne samom pojedincu koji je dio tog sistema. Searle im na to odgovara da bez obzira gdje se taj pojedinac nalazio, bilo to u sobi ili negdje vani, i dalje ostaje to da on kineski ne razumije. A ako ne razumije, tada ne postoji nikakav način na koji bi on razumio sistem jer je sistem dio njega. Osim toga, ukoliko spojimo neku osobu i manipulaciju, nije nužno da će to dovesti i do razumijevanja. Drugi odgovor, odnosno robot-odgovor stiže s Yalea. Oni pretpostavljaju da je unutar robota ugrađeno računalo koje ne bi uzimalo samo formalne simbole kao input i izbacivao iste kao output već bi taj isti robot mogao također i zamjećivati stvari oko sebe, jesti, hodati i izvoditi

slične radnje. Taj robot bi osim toga imao i ugrađen ekran s kojim bi mogao vidjeti svijet oko sebe, a sa svim njegovim dijelovima bi upravljao računalni mozak. Budući da bi imao mozak, imao bi onda i istinsko razumijevanje i uvid u mentalno. Searl s druge strane smatra da dodatak motoričkih vještina također ne jamči razumijevanje ili intencionalnost jer takav robot i dalje samo slijedi dane formalne upute i može manipulirati, ali ne i razumijeti pravila. Na koncu, dolazimo do trećeg odgovora koji ćemo obraditi, a taj je odgovor simulatora mozga s Berkleya i MIT-a. Ovi stručnjaci su zamislili pripremu programa koji simulira stvarnu sekvencu neurona ispaljenih u sinapsu mozga izvornog govornika kineskog kad ovaj razumije i daje odgovor na kineske priče; tj. zamislili su da taj program ili skup programa rade kao stvarni ljudski mozgovi kad obrađuju prirodni jezik. Tad bi se moglo reći da stroj razumije priču. Ono što je neobično kod ovog odgovora je to da se zapravo čitava ideja UI sastoji u tome da nije potrebno znati kako radi mozak da bi se znalo kako radi um. Searl na ovo odgovara na sljedeći način. On kaže da zamislimo čovjeka koji umjesto znanja materinskog jezika i baratanja sa simbolima upravlja složenim nizom vodovodnih cijevi i ventila. Primivši kineske simbole, on u engleskom jeziku traži odgovore koje ventile treba otvoriti odnosno zatvoriti. Svaki spoj odgovara jednoj sinapsi u kineskom mozgu, a kad se naprave točna paljenja, kineski odgovori izlaze kao outputi iz kraja serije cijevi. Međutim, kako ovdje možemo govoriti o razumijevanju? Problem sa simulatorom mozga je u tome da simulacija formalnih struktura sekvence neuronskih paljenja na sinapsi ne simulira istodobno i ono bitno u mozgu, tj. uzročna svojstva i sposobnost stvaranja samih intencionalnih stanja.

Postavljamo si pitanje zašto ljudi mogu razumijevati, a strojevi ne? Searle smatra da nešto čemu pripadaju uzročna svojstva može imati intencionalnost, a to su ljudi koji su određena vrsta organizma s određenom biološkom strukturom koja pod određenim uvjetima, uzročno mogu stvoriti percepciju, djelovanje, razumijevanje, učenje i slične radnje. Glavna točka njegovog argumenta je ta da formalni elementi nemaju intencionalnost niti posjeduju uzročne moći. Searle, dakle, na gore postavljeno pitanje, odgovara da samo strojevi i mogu misliti, i to oni strojevi koji posjeduju uzročne moći isto kao i ljudski mozak. Iz tog razloga nam UI može govoriti jako malo o mišljenju jer nam zapravo ne daje nikakve informacije o strojevima. UI se ne bavi strojevima, nego programima. Što god intencionalnost zapravo bila, nijedan program nije dovoljan da bi se ona „umjetno“ proizvela.

## 7. Drugi autori o odnosu mentalnog i fizičkog

Navest ćemo neke poznate primjere tj. teorije koje tematiziraju i pokušavaju riješiti odnos mentalnog i fizičkog. Obradit ćemo teorije J.J.C. Smarta, Lewisa te D.M. Armstronga iz razloga što se jednim dijelom mogu smatrati funkcionalističkim teorijama, ali će nam također i jasnije predložiti razlike između fizikalizma i funkcionalizma.

Smartova veza sa funkcionalizmom bila bi jedino ta što se jednim od izvora funkcionalizma može smatrati njegova „sadržajno neutralna analiza.“ Smart je pokušao analizirati mentalne koncepte na način koji neće zahtijevati obvezu/povezanost prema mentalnim ili fizičkim stanjima koncepta. Ova analiza specificira mentalna stanja u smislu podražaja koji ih uzrokuju. U ovoj analizi Smart „prekida“ vezu s biheviorizmom po pitanju inzistiranja da su mentalna stanja stvarna s kauzalnom efikasnošću. Smart je koristio ovu analizu samo da bi branio a priori teoriju identiteta. Smatrao je da je apsurdno misliti da postoje savršene korelacije između mentalnih stanja i stanja mozga, a da ta stanja ne budu identična.<sup>39</sup> Ono po čemu se funkcionalizam razlikuje od teorije identiteta (fizikalizam) jest to da funkcionalizam ostaje na apstraktnoj razini opisa kauzalnih uloga tj ne opisuje, ne iskazuje koja je narav medija u kojem se realiziraju kauzalne uloge. Utoliko je Smartova „neutralna analiza“ (apstraktni opis) u određenom smislu i funkcionalistička. Fizikalist neće pristati na postojanje nefizičkih svojstava, ali isto tako ne može reći da mi naša unutrašnja stanja opisujemo na način da im samo pridajemo fizička svojstva. Postoje dva razloga zašto je to tako; nije jasno koja bi to uopće fizička svojstva bila te bi ta strategija zapravo pretvorila fizikalizam u tezu o značenju mentalnih termina. Jedino rješenje za ovaj problem bi bilo neko treće rješenje koje se naziva „sadržajno neutralnom analizom“. <sup>40</sup> Smart iznosi sljedeću formulaciju: „Kada netko kaže Vidim žućkasto-narančastu naknadnu sliku, on hoće reći otprilike ovo: Nešto se događa što je slično onome što se događa kada su mi otvorene oči, budan sam i ispred mene se nalazi dobro osvijetljena naranča, to jest, kada stvarno vidim naranču.“ <sup>41</sup> Ovdje se radi o sadržajno neutralnim riječima što objašnjava kako osjeti mogu biti moždani procesi, a da čovjek koji o njima izvještava zapravo ne mora ništa ni znati o moždanim procesima. Smart je htio pokazati da se opisi unutrašnjih stanja tumače kao

---

<sup>39</sup> Block Ned, *Introduction: What is Functionalism?*, Minnesota Studies in the Philosophy of Science 9:261-325 (1978), str. 179

<sup>40</sup> Sesardić Neven, *Fizikalizam*, GRO Kultura, Beograd, 1984, str. 48

<sup>41</sup> Ibid., str. 48

indirektni opisi nepoznatih stanja pomoću njihovih tipičnih uzroka. Mentalna se stanja, dakle, identificiraju po sličnosti sa stanjima koja se javljaju u određenim situacijama.<sup>42</sup>

Smartova sadržajno neutralna analiza određuje mentalna stanja kao tipičnu posljedicu vanjskih djelovanja na čovjekovo tijelo, dok ih Armstrong s druge strane specificira pomoću njihovih tipičnih posljedica. Ovakvo specificiranje je uvjetovano time što on sadržajno neutralnu analizu predlaže za sva mentalna stanja, a ne samo za naknadne slike, bol i slično. Sadržajno neutralna analiza se svodi na tezu da su mentalna stanja definirana pomoću svojih uzroka ili posljedica ili i uzroka i posljedica. David Lewis obrazlaže da kako prihvatanje kauzalne analize mentalnih termina zapravo vodi ka obrani samog materijalizma. Lewis i Armstrong su zastupnici kauzalne teorije identiteta (uzročna uloga svijesti potvrđuje njezinu fizičku narav) koja ima funkcionalističke elemente te mentalna stanja također definiraju putem kauzalnih veza, međutim oni se ne zaustavljaju na neutralnoj/apstraktnoj razini opisa nego i specificiraju mentalna stanja kao neurofiziološka stanja. Poistovjećuju mentalno i fizičko: mentalna stanja su procesi u mozgu. Njih dvojica dakle zastupaju i materijalizam središnjeg stanja (mentalna stanja identična sa stanjima u mozgu, centralnom nervnom sustavu). Za razliku od njih funkcionalisti nisu nužno materijalisti jer ostaju na onom apstraktnom funkcionalnom opisu (npr. strojna tablica) što im omogućava tvrdnju da različiti supstrati ili stvari mogu biti nositelji kauzalnih uloga (mozgovi, čipovi, duhovi, strojevi) Lewis i Armstrong specificiraju te nositelje kauzalnih uloga, za njih su to procesi u mozgu pa njihova teorija ima funkcionalističke i fizikalističke (teorija identiteta) aspekte.

---

<sup>42</sup> Ibid., str. 48



## 8. Nedostaci i prigovori funkcionalizmu

Mnogi filozofi su pisali o funkcionalizmu, ali i u onom negativnom kontekstu. Neki od filozofa koji su naveli nedostatke i prigovore funkcionalizmu, a koje ćemo obraditi u ovom radu su William G. Lycan, William Kalke te već prije spominjani, Ned Block. Krenut ćemo redom te prvo obraditi članak Williama G. Lycana koji se posebno odnosi na Putnama. Lycan zapravo u svom članku pokušava dati svoj pogled na mentalna stanja te se također osvrće na Putnamovu hipotezu funkcionalizma. Lycan započinje svoj članak navodeći da postoje tri vrste materijalizma; reduktivni materijalizam koji je okarakteriziran teorijom identiteta, eliminativni pogled Richarda Rortya, P. K. Feyerabenda i W. V. O. Quinea, te funkcionalizam kojeg zagovaraju Hilary Putnam i J. A. Fodor. Mnogi zapravo sumnjaju da je ova posljednja vrsta, dakle funkcionalizam, različita od prve dvije. Lycan želi sagledati logičku strukturu funkcionalizma kao opće teorije uma, te zatim provjeriti je li funkcionalizam zaista zasebna, autonomna teorija. Na koncu, Lycan donosi argumente protiv funkcionalizma kao autonomne teorije.

### 8.1. Prigovori Williama G. Lycana

Funkcionalizam se izražava u obliku slogana: „Mentalno stanje je logično (ili funkcionalno) stanje Turingovog stroja“. Turingov stroj, kao što je već prije navedeno, ima konačan raspon diskretnih mogućih ulaznih signala koji se mogu unijeti u njega i na linearnoj traci (apstraktno zamišljeni) kao i sličan raspon diskretnih mogućih izlaza i konačan raspon diskretnih unutarnjih stanja.<sup>43</sup> Poseban Turingov stroj je određen onim što Putnam naziva tablicom stroja. Ova tablica je program po kojem stroj zapravo i radi. Može se izračunati i predvidjeti ponašanje samog stroja upravo pomoću ove tablice ako se zna kakvo je bilo početno stanje stroja te, naravno, ako se pozorno prate sve promjene odnosno svi unosi koje stroj zabilježava. Dakako, neka unutarnja stanja koje stroj navodi su njegova logička ili funkcionalna stanja. Međutim, treba imati na umu da strojevi nisu fizički objekti. Ono što Lycan prigovara Putnamu jest to da Putnam više bavio tradicionalnim pitanjima teorije uma, tj. pokušao je razjasniti neka pitanja tako da je upućivao na to da su analogne zagonetke koje se odnose na Turingov stroj zapravo vrlo lako rješive.

---

<sup>43</sup> Lycan William G., *Mental states and Putnam's functionalist hypothesis*, Australasian Journal of Philosophy, Vol. 52, No. 1, May 1974, str. 48

Lycan zatim povlači paralelu između Armstrongovog i Putnamovog funkcionalizma. Za Armstronga su mentalna stanja identična stanjima živčanog sustava prema tranzitivnosti identiteta. Putnam identificira mentalna stanja s funkcionalnima na teoretskoj osnovi, a potom taj problem ostavlja kibernetici. Kibernetika je ta koja bi trebala dati odgovore na pitanja koja su mentalna stanja identična pojedinim funkcionalnim stanjima pojedinog stroja. Osim što se poziva na kibernetičiste, Putnam također ostavlja neurofiziolozima da odgovore na pitanja koja tjelesna stanja osobe u stvari shvaćaju funkcionalna stanja (ona funkcionalna stanja koja je iz tablice stroja napisao upravo kibernetičist), a sve to pozivajući se na njihove uloge koje imaju prilikom stajanja između podražaja i odgovora. Putnam zatim nastavlja s tim da, ako je funkcionalizam istinit, onda teorija identiteta mora biti lažna. Jer ako su mentalna stanja identična funkcionalnim stanjima, i ako se razlikuju funkcionalna stanja od neuroloških stanja, onda se mentalna stanja također razlikuju od neuroloških stanja.

Lycan zatim iznosi tri argumenta protiv Putnamove hipoteze. Prvi argument se odnosi na to da pretpostavimo da smo uzeli određenu jedinicu stroja koji se nalazi u nekom određenom stanju i pažljivo manipulira tom istom jedinicom. Svaki od njegovih fizičkih dijelova zamjenjujemo dijelovima različite vrste, izrađenih od različitih materijala i različitom fizičkom konfiguracijom. To činimo na takav način da ova jedinica ima istu ulazno-izlaznu funkciju kao i izvorni Turingov stroj, tako da se ova radnja i dalje može izvoditi na isti način kao i kod izvornog stroja. Stroj je sada u drukčijem strukturnom stanju od njegovog izvornog, izrađen od drugačije vrste raspoređenih materija, ali je, *ex hypothesi*, u istom funkcionalnom stanju. Pretpostavka da su funkcionalna stanja identična strukturalnim stanjima koja ih ostvaruju, povezana s ne-identitetom dviju strukturnih stanja navedenih ovdje, dovodi do kontradikcije; pa je pretpostavka identiteta lažna. Drugi argument je sličan Kalkeovom pa ga nećemo opširno opisivati, ali riječ je o tome da funkcionalna stanja ili svojstva npr. PC sustava nikada nisu identična s PC stanjima koja ih ostvaruju. Treći argument kaže da je Turingov stroj, kako je do sada zamišljen, apstraktni objekt raznih vrsta. Teško je zamisliti da takav stroj uopće može imati nekakvo stanje, ali kad bi i imao, ona bi isto tako morala biti apstraktna. Postavlja se, samo po sebi, pitanje kako onda neko fizičko stanje može biti identično s nečim apstraktnim? <sup>44</sup>

---

<sup>44</sup> Ibid., str. 53

## 8.2. Kalkeovi prigovori

Kalke u svom članku iznosi svoja stajališta te pokušava objasniti što to zapravo nije u redu s funkcionalizmom koji donose Fodor i Putnam. Iako se, kako kaže, slaže s dosta toga što obojica iznose u svojim radovima, Kalke poriče da teorija identiteta može biti isključena a priori po funkcionalnoj/strukturalnoj razlici koju oni razvijaju. Fodor smatra da je psihologija sustavni pokušaj objašnjenja i predviđanja ponašanja organizama i tvrdi da u svojoj prvoj fazi, to uključuje razvoj teorije unutarnjih stanja organizma tako da su uvjeti teorije koji se ne odnose na ponašanje funkcionalno okarakterizirani te da je teorija sposobna adekvatno predvidjeti ponašanje organizma s obzirom na znanje trenutne situacije podražaja. No, što znači reći da su ta stanja funkcionalno okarakterizirana?<sup>45</sup> Iako Fodor govori o organizmima kao živim bićima, jasno je da se ova faza ne ograničava samo na sustav živih bića. Fodor želi održati tezu da postoji stroga logička razlika između funkcionalne i strukturalne karakterizacije. Međutim, nikada nije do kraja objasnio zašto je tako mislio.

Putnam s druge strane tvrdi da je na logičkoj osnovi nemoguće da funkcionalna stanja nekog fizičko-kemijskog sustava mogu ikada biti identična fizičko-kemijskim svojstvima koja ih realiziraju. Kalke iznosi tri izjave koje su zapravo reformulacija Putnamovog argumenta za nužnost ne-identiteta funkcionalnog i strukturalnog stanja. Prva kaže da ako su dva sustava različitih PC konstitucija funkcionalno izomorfni, onda odgovarajuća (iako ne-identični) PC stanja (svojstva) ostvaruju isto funkcionalno stanje. Druga izjava se nastavlja na prvu te kaže ako su dva različita PC sustava funkcionalno izomorfna, njihova funkcionalna stanja (svojstva) nisu identična s PC stanjima (svojstvima) koja ih ostvaruju. Na koncu, treća izjava kaže da je bilo koji PC sustav funkcionalno izomorfan do bezbrojnih beskonačnosti PC sustava sa sasvim različitom PC građom. Posljednje dvije zajedno impliciraju da funkcionalna stanja PC sustava nikada nisu identična s PC svojstvima koja ih realiziraju.<sup>46</sup>

Ni Fodor niti Putnam zapravo ne daju nikakve kriterije za primjenu funkcionalne, odnosno strukturalne razlike između stanja i svojstava koja karakteriziraju PC sustave. Fodor sugerira da je unutarnje stanje funkcionalno karakterizirano pod uvjetom da sve što mi znamo jest to da je njegova uloga u proizvodnji ponašanja. Putnam sugerira da ako pogledamo sustav PC kao realizaciju konačnog probabilističkog Turingovog stroja ili računala, njegovu funkcionalnu karakterizaciju daje "tablica stroja". No ova tablica stroja jednostavno je

---

<sup>45</sup> Kalke William, *What is wrong with Fodor and Putnam's functionalism?*, Noûs, vol. 3, No. 1, February 1969, str. 84

<sup>46</sup> Ibid., str. 86 - 87

specifikacija određenih stanja isključivo u smislu njihovog međusobnog odnosa i (vanjskog) ponašanja.

### 8.3. Problemi s funkcionalizmom po Nedu Blocku

Ned Block svoj članak *Troubles with Functionalism* započinje objašnjavanjem što su biheviorizam, fizikalizam te funkcionalizam te potom iznosi prigovore istima. Kao najjednostavniju verziju funkcionalizma navodi Turingov stroj, odnosno kaže da svaki sustav koji posjeduje mentalna stanja je opisan bar jednom tablicom Turingova stroja, te da je svako mentalno stanje sustava identično nekom od stanja tablice stroja. Block daje primjer tijela koja izvana izgledaju kao ljudska, ali su iznutra potpuno drugačija. U našoj glavi, skup malih ljudi izvodi operaciju, zajedno s oglasnim pločama s svjetlima koja ukazuju na podskupove tih istih malih ljudi. Njihov posao se sastoji u tome da provedu stanja unutarnje strojne tablice tako da tijelo shvati ista stanja stroja kao što ih i mi shvaćamo. U funkcionalizmu, aktivnost homunkulus glave pokazuje da ima mentalitet baš kao i mi; međutim, mi očito ne želimo pripisati iskustva svijesti, boli, te svega ostalog homunkulus glavi. Ovim primjerom Block želi pokazati da je funkcionalizam opsjednut problemom liberalizma. Zatim opisuje simulaciju homunkulus glave koja ima najviše šansi da bi bila nomološki moguća. Kao primjer daje sljedeće; pretpostavimo da kinesku vladu pretvaramo u funkcionalizam, te uvjerimo njezine dužnosnike da na jedan sat realiziraju ljudski um. Svakom od milijardu ljudi Kine omogućavamo posebno dizajnirani dvosmjerni radio koji ih na odgovarajući način povezuje s drugim osobama i s umjetnim tijelima koje smo prethodno zamislili. Zamjenjujemo svakog malog čovjeka s građaninom Kine i njegovim radijem. Umjesto oglasne ploče, dogovorimo da imamo slova prikazana na nizu satelita postavljenih tako da ih se može vidjeti s bilo kojeg mjesta u Kini. Ovaj sustav od milijardu ljudi koji komuniciraju jedni s drugima uključujući i satelite, igra ulogu tzv. vanjskog mozga koji je povezan s umjetnim tijelom putem radija.<sup>47</sup> Isto tako, sustav bi mogao na neko kratko vrijeme ostvariti funkcionalnu organizaciju, koju mi ipak ne želimo pripisati nikakvom mentalitetu. Sustav je funkcionalno ekvivalentan nekoj osobi, budući da je sve funkcionalno ekvivalentno drugoj stvari (ima jednak skup input-output unutarnjih stanja). Vrijeme koje je potrebno za ovu funkcionalnu realizaciju uopće nije toliko važno jer je jedino bitno da ta realizacija bude ista kao u nekoj osobi. Block ovdje savjetuje da se zamisli osoba čiji su mentalni procesi

---

<sup>47</sup> Block Ned, *Troubles with Functionalism*, Minnesota Studies in the Philosophy of Science 9:261-325, 1978., str. 71

usporeni jer na taj način i osoba i homunculi glava mogu realizirati funkcionalna stanja istom brzinom.

Oba navedena primjera su usmjeravana na funkcionalnost strojnog stanja jer su zapravo upućivala ni na što drugo nego na stanja tablice stroja samog sustava. Premda je to najjednostavnija i najpoznatija verzija funkcionalizma, Block ovo primjenjuje i na ostale verzije funkcionalizma jer se ekstrapolacija odnosi na kvalitativna stanja, a ne na mentalna stanja općenito, kako na prvi pogled izgleda. Sve verzije funkcionalizma kažu da je kvalitativno stanje upravo funkcionalno stanje. Najjednostavniji primjer je već prije spominjani primjer boli. Međutim, mi također možemo zamisliti bića u funkcionalnom stanju, a koja uopće ne osjećaju bol. Upravo zbog ovoga, funkcionalizam je „kriv“ za liberalizam.

Ovaj članak, odnosno prigovori funkcionalizmu, mogu se jednostavno sumirati na argument odsutnosti kvalija. Block započinje time da iznosi tvrdnju funkcionalizma da su kvalitativna stanja, kao što je npr. bol koju smo već prije uzimali za primjer, zapravo funkcionalna stanja te da su povezana međusobno inputima, outputima te ostalim unutarnjim stanjima. Nadalje, ako možemo zamisliti vjerojatan slučaj sustava koji ostvaruje isti skup funkcionalnih stanja koje pojedina osoba čini, ali gdje intuitivno želimo izbjegavati pripisivanje sposobnosti da uopće možemo doživjeti kvalitativna stanja u sustavu, onda je prima facie, funkcionalizam uznemiren već spomenutim problemom liberalizma. Postoji vrlo važna razlika između sustava s mozgom te homunkulus glave. Budući da mi znamo da smo sustavi s mozgom, samim time znamo da takvi sustavi imaju kvalije. Mi zapravo nemamo nekakvu teoriju o kvalijama koja bi objasnila kako je to moguće, ali zato imamo dovoljan razlog za zanemarivanje bilo kakve prima facie sumnje o samoj kvaliteti sustava s mozgovima. Postoji još jedna razlika između nas, sustava s mozgom i homunkulus glava; oni su sustavi osmišljeni da nas oponašaju, ali mi nismo dizajnirani da oponašamo ništa. Ta činjenica sprječava svaki pokušaj rasprave na temelju zaključka o najboljem objašnjenju za kvalije homunkulus glava. Najbolja objašnjenja vrisaka i strahova u homunkulus glavama nisu njihovi bolovi, već su oni dizajnirani tako da oponašaju naše krikove i strahove. Blockov argument protiv funkcionalizma se sastoji u sljedećem; ako doktrina ima apsurdni zaključak da ne postoji neovisan razlog za vjerovanje, ako nema načina objašnjavanja apsurdnosti ili pokazivanja pogrešnog ili nevažnog, te ako nema razloga za vjerovanje doktrini koja vodi do apsurdnosti na prvom mjestu, onda ne treba prihvatiti doktrinu. Block tvrdi da ne postoji neovisni razlog da vjerujemo u mentalitet homunkulus glave, te da ne znam ni na koji način objasniti apsurdnost zaključka da ima mentalitet. Argumenti za funkcionalizam su ti da se

funkcionalni identiteti mogu pokazati istinitim na temelju analiza značenja mentalne terminologije te da je ovo zapravo najbolje rješenje dostupno mind-body problemu. Ipak, Block priznaje da je moguće definirati bilo koji uvjet mentalnog stanja u smislu plitkosti koje se odnose na druge uvjete mentalnih stanja, te uvjete inputa i outputa. Ovdje je bila riječ o psihofunkcionalizmu. Prema psihofunkcionalizmu ono što treba računati kao inpute i outpute sustava jest zapravo empirijsko pitanje. Brojanje neuronskih impulsa kao inputa i outputa izbjeglo bi probleme koji su skicirani, budući da bi mozgovi u bocama koji su nemoćni (paralizirani), mogli imati ispravne neuronske impulse čak i bez tjelesnih pokreta. Block ovoj tezi prigovara da može doći do paralize koja utječe na živčani sustav, pa se ovaj problem pojavljuje kako u funkcionalizmu tako i u psihofunkcionalizmu. Isto tako daje i odgovor na ovaj prigovor. Bolesti živčanog sustava mogu zapravo promijeniti mentalitet; npr. omogućiti osjet boli ljudima koji nemaju tu sposobnost osjećanja boli. Stoga bi moglo biti istina da je široko rasprostranjena bolest živčanog sustava uzrokovala povremenu paralizu ljudima nesposobnim za određena mentalna stanja. Vratimo se sad na argument odsutnosti kvalija. Ovaj argument iskorištava mogućnost da funkcionalisti ili psihofunkcionalisti žele funkcionalna/psihofunkcionalna stanja identificirati s boli koja se jednostavno može pojaviti bez razloga. Osim toga, čini se da se ovo može dogoditi bez nekog prethodnog događaja. Doista, postoje činjenice koje pružaju vjerodostojnost ovom pogledu. Nakon frontalnih lobotomija, pacijenti obično izvješćuju da još uvijek imaju bolove, iako ih bol više ne smeta. Ovi pacijenti pokazuju sve "senzorske" znakove boli (npr. prepoznaju šiljke kao oštre), ali često imaju malu ili nikakvu želju da se izbjegnu "bolni" podražaji.<sup>48</sup>

Za kraj, Block se dotiče problema inputa i outputa. Ovdje se pokušava izbjeći šovinizam. Jedan od načina da se to i učini bio bi da se inputi i outputi okarakteriziraju samo kao inputi i outputi. Tako bi funkcionalni opis osobe jednostavno navodio listu outputa po brojevima. Na taj način bi sustav mogao biti funkcionalno ekvivalentan nama ako bi imao skup stanja, inputa i outputa koji bi bili povezani jedni s drugima na način na koji su i naša stanja, inputi i outputi povezani, bez obzira na to kakvi oni jesu. Problem s ovom verzijom funkcionalizma jest taj što je previše liberalan. Svaki opis funkcionalizma za Blocka je zapravo prikaz liberalizma ili šovinizma te se zbog toga moraju odbaciti.

---

<sup>48</sup> Ibid., str. 87

## 9. Zaključak

Još od antičkog vremena prisutan je problem duha i tijela. Prvi put se susrećemo s ovim problem već kod Platona i Aristotela. Mnogo kasnije ovaj problem Descartes posebno obrađuje i uspostavlja dualističku teoriju ili tzv. „interakcionistički dualizam“; ljudsko biće jest dualitet uma i tijela. Sljedeći pravac koji nastaje i koji se suprotstavlja dualizmu jest bihejviorizam koji tvrdi da nema duha niti mentalnih stanja. Dalje slijedi fizikalizam prema kojemu su mentalna stanja identična fizičkima. Na koncu dolazimo do funkcionalizma koji je bio i glavna tema ovoga rada. Funkcionalizam je teorija koja tvrdi da ono što čini nešto mentalnim stanjem određenog tipa ne ovisi o njegovom unutarnjem ustroju, već o načinu na koji funkcionira ili ulozi koju igra u sustavu čiji je dio. Pod vrste i glavne predstavnike funkcionalizma naveli smo zdravorazumski; referencijalni te strojni funkcionalizam, dok su glavni predstavnici J. A. Fodor i Hilary Putnam. Zdravorazumski funkcionalizam smatra da stvari u svijetu razlikujemo prema njihovim funkcijama, referencijalni tvrdi da funkcionalizam nije u stanju opisati osjetilna iskustva, a strojni uspoređuje rad ljudskog uma s radom stroja. S ovim potonjim smo se najviše i bavili jer funkcionalisti svoju teoriju vole objašnjavati na primjeru Turingovog stroja. Fodor nam je na primjeru automata za coca – colu pokušao prikazati razliku između funkcionalnog i strukturalnog opisa, detaljnije nam je pojasnio kvalitativni i intencionalni sadržaj te njihovu važnost za funkcionalizam. Putnam se pak više bavio računalnom psihologijom te je tako i nastala njegova teorija interpretacije. Putnam smatra da se jezik interpretira preko skupa pravila koja označavaju stupnjeve konfirmacije, a ta pravila moraju biti kompjutabilna. Prije nego smo krenuli s objašnjavanjem veze funkcionalizma i Turingovog stroja, prisjetili smo se podjele fizikalizma na fizikalizam vrsta i fizikalizam instanci te njihove važnosti za funkcionalizam. Kao što znamo, funkcionalizam prihvaća samo fizikalizam instanci. Funkcionalisti smatraju da pomoću Turingovog stroja mogu objasniti funkcioniranje ljudskog uma iako je Turingov stroj zapravo apstraktan stroj koji se ponaša na točno određen način. Sam Turing je smatrao da se ovaj stroj može dovesti u vezu s umjetnom inteligencijom te je iz tog razloga postavio kao uvjet Turingov test. Želi li se nekom stroju pripisati inteligencija mora proći ovaj test. Međutim, kako smo vidjeli na primjeru kineske sobe, testom se može vrlo lako manipulirati.

Funkcionalistička teorija nije održala svoju plauzibilnost te je doživjela čitav niz opovrgavanja i prigovora od kojih smo i mi odabrali neke i ovdje ih prikazali. Obradili smo tri autora koji su se bavili ovom temom, a to su W. Lycan, W. Kalke te N. Block. Lycan najviše

prigovora upućuje Putnamu po pitanju Turingovog stroja. Lycan smatra da se Putnam previše bavio tradicionalnim pitanjima teorije uma te smatra, čak i slučaju kad bi stroj mogao imati neka mentalna stanja, da bi mentalna stanja bila apstraktna te se kao takva ne bi uopće mogla identificirati s fizičkima. Kalke prigovara i Putnamu i Fodoru po pitanju teorije identiteta. Putnam i Fodor razvijaju funkcionalnu odnosno strukturnu razliku s čime se Kalke ne slaže, te smatra da uopće nema razloga za primjenu iste. Blockov članak donosi nekoliko prigovora, ali najveći naglasak je na argumentu odsutnosti kvalija. Daje nam primjer sustava s mozgom te homunculi glava pokušavajući nam tako prikazati razlike između čovjekovog uma te stroja. Riječ je o tome da homunkulus glave oponašaju ljudsko ponašanje, dok sami ljudi ne oponašaju apsolutno ništa. Block se također dotiče i problema inputa i outputa, odnosno naglašava da ih se treba gledati upravo kao inpute i outpute te im ne dodavati nikakva druga značenja.



## Popis literature

1. Berčić B., *Filozofija – svezak drugi*, Ibis grafika d. o. o., Zagreb, 2012
2. Block N., *Introduction: What is Functionalism?*, In Donald M. Borchert (ed.), [Book Chapter]. MacMillan (1996)
3. Block N., *Troubles with Functionalism*, Minnesota Studies in the Philosophy of Science 9:261-325 (1978)
4. Chalmers D.J., *Philosophy of mind*, Oxford University Press, New York, 2002, str. 55 - 87
5. Fodor J.A., *Problem duha i tijela*, Računala, mozak i ljudski um, Izdavački centar Rijeka, Rijeka, 2001, 63-85 str.
6. Janović Tomislav, Pećnjak Davor, *Nefunkcionalnost funkcionalizma*, In: Maja Hudoletnjak Grgić, Davor Pećnjak, Filip Grgić (ur.): *Aspekti uma*. Institut za filozofiju, Zagreb, 2011, pp. 1-18.
7. Kalke W., *What is wrong with Fodor and Putnam's Functionalism*, Noûs, Vol. 3, No.1, Wiley, 1969, 83-93 str.
8. Lycan W.G., *Mental states and Putnam's functionalist hypothesis*, Australasian Journal of Philosophy, Vol. 52, No. 1, May 1974
9. Lycan William G., *Problem um – tijelo*, ZINK Znanstveno-istraživački inkubator, Filozofski fakultet u Sarajevu FFSA, SOPHOS Volume 1, 2008, 107-127 str.
10. Putnam H., *Računarska psihologija i teorija interpretacije*, Računala, mozak i ljudski um, Izdavački centar Rijeka, Rijeka, 2001, 154-170 str.
11. Searle J.R., *Umovi, mozgovi i programi*, Računala, mozak i ljudski um, Izdavački centar Rijeka, Rijeka, 2001, 134-154 str.
12. Sesardić N., *Fizikalizam*, GRO Kultura, Novi beograd, 1984.
13. <https://plato.stanford.edu/>

## Sažetak

### Funkcionalistička teorija i njezini nedostaci

Funkcionalizam pokušava objasniti vezu između uma i tijela, odnosno bavi se tzv. mind-body problemom. Najjednostavniji oblik funkcionalizma može se opisati Turingovim strojem koji funkcionira na principu tablica putem inputa te outputa. Turingov stroj je također putem inputa i outputa povezan i s umjetnom inteligencijom. Glavni predstavnici funkcionalizma su Hilary Putnam te Jerry Fodor. Osim toga, mnogi filozofi su se bavili problemom mentalnoga i fizičkoga te naglašavali da su zapravo svijest i osjećaji procesi mozga. Naglasak je na tome da se mentalna stanja mogu identificirati s fizičkim stanjima. Također, pojedini filozofi su iznosili i prigovore funkcionalizmu. Prigovori su nastali po pitanju Turingovog stroja, odnosno samog njegovog principa rada. Ovi filozofi smatraju da se mentalna stanja ne mogu usporediti niti identificirati s fizičkima jer bi mentalna stanja stroja zapravo bila apstraktna te samim time neupotrebljiva.

**Ključne riječi:** funkcionalizam, input, output, Turingov stroj, mentalna stanja, svijest, procesi mozga

## Abstract

### Theory of Functionalism and its disadvantages

Functionalism tries to explain the connection between the mind and the body, ie it deals with the so-called mind-body problem. The simplest form of functionalism can be described by a Turing machine that functions on the basis of the table by input and output. Turing's machine is also connected through inputs and outputs with artificial intelligence. The chief representatives of functionalism are Hilary Putnam and Jerry Fodor. In addition, many philosophers have dealt with the problem of mental and physical, and emphasized that consciousness and feelings are actually brain processes. The emphasis is on the fact that mental states can be identified with physical states. Also, some philosophers have argued against functionalism. Objections were raised regarding the Turing machine, or its very principle of work. These philosophers believe that mental states can not be compared or identified with physical states because the mental state of the machine would actually be abstract and thus unusable.

**Key words:** functionalism, input, output, Turing machine, mental states, consciousness, brain processes