

izvorni znanstveni članak/original scientific paper.
primlje/received: 5.5.2020

UDK: 131 Putnam H.

TEORIJA FUNKCIONALNOG IZOMORFIZMA U FILOZOFIJI UMA

Karlo GARDAVSKI

University of Sarajevo, Faculty of Philosophy
Franje Račkog 1, 71000 Sarajevo, B&H
E-mail: karlo.gardavski@gmail.com

ABSTRACT

Tema ovoga rada je prikaz alternativne teorije u području filozofije uma nastale razvojem kognitivnih znanosti i kompjuterskih tehnologija dvadesetoga stoljeća. Posebno je važna Turnigova mašina čije je proučavanje na logičkoj razini pridonijelo novim načinima razumijevanja mentalnoga. Teorija kao što je funkcionalizam (Putnam) ne proučava mentalno na onaj način na koji su to proučavali dualisti ili materijalisti u filozofskoj tradiciji. Mentalno se treba opisati generalno logički ("viši nivo" opisa), tj. opisuje se njegova funkcija (odnos inputa i outputa) koja je nastala povlačenjem paralele između čovjeka i mašine (poznato kao kompjuterska metafora).

Ključne riječi: mentalno, funkcionalizam, um – tijelo, Turingova mašina, mašina, izomorfni funkcionalizam, mentalna stanja

The topic of this paper is a presentation of an alternative theory in the field of philosophy of mind, which was created by the development of cognitive sciences and computer technologies of the twentieth century. Of particular importance is Turnig's machine whose study at the logical level has contributed to new ways of understanding the mental. A theory such as functionalism (Putnam) does not study the mental in the same way as dualists and materialists do in the philosophical tradition. The Mental has to be described in a generally logical manner ("higher level" of description), i.e. its function is described (input-output ratio), and is created by drawing a parallel between man and machine (known as a computer metaphor).

Key words: mental, functionalism, mind-body, Turnig's machine, machine, functional isomorphism, mental states

1.0 Uvod

Jedno od pitanja koje je od samih početaka mučilo filozofiju i filozofe je: Da li čovjek posjeduje određenu nematerijalnu supstancu koja se razlikuje od tijela? Da li je ta nematerijalnost zapravo odraz onog što nas čini ljudima? Da li je moguće da smo više od fizičke materije? Laike je obično strah ulaziti u dublju problematiku ovoga problema, jer postoji strah da će se čovjek svesti na jedan oblik mašine (stroj) bez duha.

Mišljenje da tijelo samo ima probavna i reproduksijska svojstva je materijalno, a duša (um, duh ili nešto drugo u ovome kontekstu) nematerijalna se odnosi na čovjekove psihičke sposobnosti (mišljenje, želje, osjećanja, racio itd.). Platon je, recimo, zagovarao ono što bi danas opisalo dualističku teoriju. Odvojenost uma i tijela, odnosa materijalnog i nematerijalnog, propadljivog i nepropadljivog. Um, duša, duh i sl. mogu preživjeti tijelo jer se nalaze u posebnome svijetu. Ideja te oni čine čovjeka specijalnim bićem koje se razlikuje od svega ostaloga u prirodi. Ovakvo gledište preuzimaju monoteističke religije i još više obogaćuju taj narativ religijskim metaforama itd., jer “govor o duhu i njegovoj egzistenciji kada tijelo umre je dio našega intelektualnog i duhovnog života jako dugo vrijeme, nije iznenađujuće da smislenost hipoteza koje se bave interakcijom između duha i tijela, i pitanjima kao šta se dešava kada su tijela bez duše, trebala da budu savršeno smisljena” (Putnam, 1999, 95). Zbog toga je ovakvo dualističko tumačenje problema prihvaćeno u zapadnoj kulturi (preko Grka, zatim srednjeg vijeka, pa sve do današnjeg dana). Religijsko i filozofsko promatranje problema duha izgleda različito, tj. ima drugu svrhu, ali ipak se ne smije negirati utjecaj nekih religijskih misli koje su utjecale na filozofsko mišljenje, a i obrnuto. Naprimjer, “pod utjecajem sv. Augustina, Platonove ideje da duša nema materijalnu supstancu postala je dominantna kršćanska doktrina nekoliko stoljeća” (Putnam, 1999, 97).

Početkom sedamnaestoga stoljeća teološko-platonička učenja u duhu su se prilagodila razvoju nauke (matematike i fizike) toga vremena i svijesti o kauzalnim dešavanjima u fizičkome svijetu. Duh je postao tada paralela fizičkim dešavanjima (kao što je, npr., rad mozga). Taj model bi se mogao opisati kao “par sinkroniziranih satova¹: tijelo je sat koji je navijen i koji radi, bio sretan ili nesretan, do svoje smrti (...). I mentalna stanja rade sretno ili nesretno (...)” (Putnam, 1981, 76). U ovome periodu

¹ Ovdje Putnam referira na Leibniza i njegovo objašnjenje u kakvome odnosu su duh i tijelo.

tražilo se da se duh i tijelo te njihova interakcija objasni sistematičnije, ali da se striktno i jedna i druga kategorija razlikuju. “Descartes je držao da su um i tijelo esencijalno jedinstvo. Nekako um – tijelo jedinstvo jeste koje misli, osjeća, sjeća se i iskazuje osobnost” (Putnam, 1981, 77). Čovjek je mogao imati tako dvije supstance, duh i tijelo, a obje su radile kao jedno.

Ono što se zapravo krije u tradicionalno-dualističkom učenju jeste to da mentalna svojstva pripadaju duhu koji je različit od tijela. Taj entitet duh je bio nosilac tih svojstava (mentalnih i intelektualnih). Generalno i materijalističko gledište uzima to da je duh entitet, ali fizičkog karaktera, tj. izjednačavao je duh sa fizičkim svojstvima (u filozofiji uma ovaj tip teorije je teorija identiteta). Materijalizam je mentalna svojstva stavio kao fizičko-kemijska svojstva čovjeka, ali nije izbjegao zamku koja je ostala iz dualizma, naime zamku duha.

Filozofi poput Hilaryja Putnama smatraju da je pitanje poznato kao problem um – tijelo krivo postavljeno još u svome začetku. Ono što Putnam smatra da se krije iza njega, a zapravo je i pravo pitanje, jeste: Da li je moguća autonomija našega mentalnoga života? Ovo pitanje nema veza s onim što bi se zvalo čovjekova supstanca (bila materijalna ili spiritualna). Kad bismo pokušali dati odgovor na to pitanje, vidjeli bismo nevažnost pitanja koje li je čovjek supstance, jer ovako postavljanje problema ne može ponuditi nikakve rezultate, a samim tim će i više kočiti progres u mišljenju. Problem duha i tijela izgleda da je nejasno postavljen u svome početku upravo zbog toga što je duh krivo kategoriziran. Vigenštajnski rečeno, na nejasno (postavljeno) pitanje ne može se dati jasan odgovor.

2.0 Funkcionalizam

Teorija funkcionalizma je nastala u filozofiji uma poslije sredine dvadesetoga stoljeća². Ta teorija “nastala je filozofskom refleksijom na razvoj umjetne inteligencije, kompjuterske teorije, lingvistike, kibernetike i psihologije” ili, općenitije rečeno, kognitivne znanosti (Fodor, 2004, 168). “Radna hipoteza te znanosti je bila da postoje mehanizmi koji su

² “Funkcionalizam filozofije i psihologije (...) su usko povezani, jer izgleda da imaju zajedničko porijeklo u djelima Aristotela” (Block, 2004, 183). Važno je to naglasiti jer će sam Putnam u kasnijoj fazi razmišljanja sebe smatrati bliži Aristotelovom učenju o duhu i mentalnom, dok je funkcionalizam teorija s kojom je krenuo prvenstveno raditi. Želim samo naglasiti tu poveznicu.

pozadina naših kognitivnih kapaciteta koji su vrste obrade informacija, naime, računanje koje operira na mentalnim reprezentacija” (Shagrir, 2005, 221). Funkcionalizam ne spada niti u jedan tip tradicionalnih teorija u polju filozofije uma, naime materijalizma, dualizma, ali niti logičkog biheviorizma (teorija nastala i proširila se utjecajem Bečkoga kruga). Ova teorija želi “ponuditi filozofska objašnjenja zbog nivoa apstraktnosti, prepoznati mogućnost da sistem raznolik kao kod ljudskih bića, mašina za računanje i bestjelesnih duha mogu imati mentalna stanja.

U pogledu funkcionalista psihologija³ sistema ne ovisi od stvari od kojih je izgrađena (živih ćelija, metala ili duhovne energije) nego kako su stvari složene” (tj. kako su stvari složene i kakav je odnos između njih) (Fodor, 2004, 168–169). “Funkcionalizam dozvoljava da čisto fizička bića mogu imati um; ali funkcionalizmu ne treba fizičko stanje kao uvjet za posjedovanje uma” (Heil, 2004, 148). Funkcionalisti naginju na materijalizam, ali samo u smislu da im je materijalizam potreban da pruži teoriju kako fizička bića mogu imati um, svijest, mišljenje (proces) itd. Funkcija uma je središnje pitanje ove teorije, a sad da li taj um ima čovjek, marsovac, mašina ili majmun, potpuno je nebitno. Um se promatra kao skup procesa koji manipulira simbolima, odnosno mentalna stanja su bila funkcionalno određena kao operacije sa simbolima. Zato je razvoj kompjuterske tehnologije i artificijelne inteligencije bio uzor za razvoj funkcionalizma, jer je pružio primjer rada jednoga stroja: predstavio se odnos hardwarea i softwarea, input i output stanja kao osnova za usporedbu čovjeka i mašine na logičkoj razini. Tu usporedbu možemo nazvati kompjuterska metafora⁴. Ono što funkcionalizam nudi jeste primjer mašine koji se uzimao kao glavno opisno sredstvo u funkcionalizmu, a to jeste Turingova mašina.

Taj uređaj je

“apstraktan uređaj koji se može realizirati fizički na neodređen broj načina, tako bi se možda trebalo odnositi prema umovima: jednostavnije rečeno, njihova funkcija ih čini onim što oni jesu a ne to od čega su izgrađeni” (Gaynesford, 2006, 85).

³ Odnosi se na mentalna stanja i mentalna dešavanja.

⁴ Drugi naziv koji Putnam koristi za kompjutersku metaforu je “računarski model uma”, koji opisuje kao model u kome “um posjeduje program ili skup pravila analognih pravilima koje upotrebljava računar” (Putnam, 1989, 125).

Teorija ovakvoga tipa radikalno mijenja pitanja tradicije, jer s uključivanjem artificijelne inteligencije u priču ono što se smatralo (samo) ljudskim mijenja svoj status, pa se uviđa da je i kod mašina način racionaliziranja bio sličan. Zbog toga pitanje uma ili duha treba preusmjeriti sa pitanja “Šta je to?” na pitanje “Kako ono radi konceptualno i logički?”.

3.0 Turingova mašina⁵

Mašina koju je Alan Turing originalno osmislio sastoji se od tri dijela: spremišta, izvršne jedinice i dijela za kontrolu. Trebalo je da sva tri dijela vrše slične funkcije kao i čovjek. “Spremište je tu za spremanje informacija i da korespondira sa papirom (...) na koji on stavlja svoje kalkulacije ili na kojem su isprintana uputstva” (Turing, 2004, 215). Kada mašina računa preko svoga spremišta, taj proces je sličan procesima koje čovjek obavlja tokom određenih misaonih funkcija. “Izvršna jedinica je dio koji izvršava različite individualne operacije uključene u računanje”, a “dužnost kontrole jeste da nadgleda da li se uputstva ispunjavaju pravilno i u pravome redu” (Turing, 2004, 215). Tablica uputstava jeste ono što se zove kompjuterskim programom, niz kodova i uputstava koji kazuju kako će mašina da obrađuje podatke. Ti kodovi su zvani input stanja mašine jer se odnose na unutrašnja stanja mašine, tj. na njen program. Output stanja mašine su kada je komanda zadana preko koda izvršena i vidljiva. Tako, recimo, input stanje je kod XYZ, a kada ga mašina registira, onda na izvrši funkciju “upaliti sijalicu” i to je output stanje koje je vidljivo na mašini. Da bi odnos input i output stanja bio ispravan, ona međusobno moraju korespondirati, tj. ako je određena operacija izvršena, ona je ispravna samo ako odgovara skupu simbola (tablica simbola koja

⁵ Alan Turing s idejom ove mašine zapravo je želio ispitati da li mašine mogu biti svjesne, mogu li misliti (kao ljudi). Zatim, želio je usporediti rad kompjutera i rad čovjekova mozga. Naravno, na fizičkoj razini mašina i čovjek se očito razlikuju, ali ako se mašina sakrije i može samo komunicirati sa ljudima preko papira na koji se ispisuju odgovori, može li čovjek znati da je to bila mašina? Pokus kojim je Alan to želio dokazati je poznat kao Turingov test. Taj test je bio takozvana “igra imitacije”, “koja se igrala sa troje ljudi, muškarca (A), žene (B) i ispitivača (C) koji može biti jedan ili drugi spol. Ispitivač se nalazi u sobi odvojenoj od druga dva igrača. Cilj igre za ispitivača je utvrditi koje od to dvoje je muškarac a koje žena. On ih zna po oznakama X i Y i na kraju igre on će reći X je A i Y je B ili X je B i Y je A” (Turing, 2004, 212). Igra se može još odvijati tako da igrač B pokuša pomoći mašini i krene odgovarati točno kako bi skrenuo pažnju mašine da je drugi ispitanik A.

je zapravo uputstvo). “U slučaju ljudskih umova, naprimjer, to je slučaj perceptivnih inputa i bihevijalnih outputa” (Gaynesford, 2006, 86). To znači da preko čula ono što čovjek dobiva kao svoji simbolički alfabet, refleksijom nad tim daje određene naznake koje se poklapaju sa tim što je primio.

“Turingova mašina je uređaj s ograničenim brojem unutrašnjih konfiguracija, a svaka od njih označava da mašina može biti u jednom ograničenom broju stanja, P i trake koju mašina skenira na kojoj se može pojaviti određen znak” (Putnam, 1979, 364). Na traci su iscrtani kvadrati na koje se ispisuju simboli. “Mašina ima i mehanizam za printanje koji može (a) izbrisati simbol koji se pojavljuje na kvadratu što se skenira i (b) ispisati neki drugi simbol (iz mašininog alfabeta) na kvadrat” (Putnam, 1979, 365).

Ono što određuje mašinu jeste njena tablica mašine koja funkcionira tako što redovi tablice korespondiraju sa slovima alfabeta, dok kolone (stubac koji se sastoji od više kvadra) korespondiraju stanjima u kojima se mašina nalazi. Naprimjer, na svakome kvadratu trake nalazi se određeni simboli XYZ koji predstavlja neko uputstvo za rad mašine. Mašina to skenira i obavlja u skladu s tim koju funkciju svaki znak predstavlja, a znak je određen u alfabetu mašine. “Tablica mašine opisuje mašinu ako mašina ima unutarnja stanja koja korespondiraju kolumnama, i ako *ispunjava* instrukcije na tablici” (Putnam, 1979, 265). Kad mašina skenira neki simbol koji označava neko stanje, taj simbol ima uputstvo prema kojem mašina vrši određenu naredbu.

4.0 Izomorfni funkcionalizam

Koncept koji Putnam⁶ zastupa, kao onaj koji će pokušati rasvijetliti problematiku uma – tijela, u filozofiji uma jeste izomorfni⁷ funkcionalizam. On glasi: “dva sistema su izomorfno funkcionalna ako postoji ko-

⁶ “Prva polovica Putnamove karijere dominirana je utjecajem i njegovim odnosom naspram logičkog pozitivizma i verifikacionizma” (Gaynesford, 2006, 76). To nije iznenađujuća činjenica s obzirom na to da je njegov mentor bio Hans Reichenbach, koji je utjecao na njegovo ranu karijeru.

Izomorfizam u ovome kontekstu jeste “jedan prema jedan korespondencija – među sekvencama operacija koje vrši mašina (jednolikost inputa i outputa)” (Heil, 2001, 398).

respodencija između jednog stanja i drugoga koji su u funkcionalnoj relaciji” (Putnam, 1975, 291). Ova teorija nastala je razvojem kompjuterske (računarske) tehnologije dvadesetoga stoljeća. Primjer koji Putnam koristi za objašnjavanje sve zbrke u filozofiji uma jeste poznat kao računarska metafora. Tom metaforom on želi reći da je “ljudski mozak-um sličan kompjuterima, i da možemo doznati bitne stvari uma polazeći od kompjutera” (Nenad Mišćević, 1989, 7). Putnam ne želi reći da je čovjek jednak jednom kompjuteru, nego da razvojem te nauke nailazimo na informacije koje se tiču čovjekove kognitivne moći, kao i na to da “analogija između čovjeka i kompjutera nije: kompjuteri nam pomažu da shvatimo odnos između duha i tijela, nego: ne mogu postojati nikakvi problemi u vezi sa odnosom između duha i tijela” (Rorty, 1990, 251). Primjer koji se uzima u objašnjavanju ove metafore jeste Turingova mašina, čiji se rad uzima kao osnova za rad svake druge kompliciranije mašine. Mašine za Putnama imaju svoju pozitivnu i negativnu stranu. “Pozitivna strana mašina je ta što smo zahvaljujući njima stekli ideju, kompjuterskim mašinama konkretnije, funkcionalne organizacije. Mašine su nas natjerale da pravimo razliku između apstraktne strukture i konkretne realizacije. (...) U slučaju kompjuterskih mašina (...) najbitnija svojstva nisu bila fizičko-kemijska. Mašine su nas natjerale da shvatimo da je ideja funkcionalne organizacije veoma važna. Negativna strana mašine je da nas tjera da previše pojednostavljujemo. Ideja funkcionalne organizacije postala je jasnom kroz veoma ograničen, veoma specifičan sistem funkcionalne organizacije. Tako da “trenutno iskušenje je pretpostaviti da moramo imati veoma ograničenu i specifičnu vrstu funkcionalne organizacije” (Putnam, 1979, 300).

“Funkcionalne relacije su sekvence relacija, tj. stanje A uvijek je praćeno stanjem B, za F da bi bilo funkcionalno izomorfno, mora biti da je stanje A praćeno stanjem B u sistemu I ako i samo ako je stanje F(A) praćeno stanjem F(B) u sistemu II” (Putnam, 1979, 292). Ako želimo, recimo, da na papir (tablicu) isprintamo X, onda će sistem I biti u stanju A i tu će se zadržati sve dok se X ne isprinta na papir, kada se X nađe na papiru, tada je sistem II u stanju F(A). Ako je stanje ovakvo, to znači da je F izomorfno sa sistemom I i II. Drugim riječima, F je “točna teorija funkcije sistema I, na funkcionalnom ili psihološkom nivou, ona izomorfizam između sistema I i sistema II mora mapirati sva svojstva i relacije određene u sistemu II tako da je F istinito ako su sve reference sistema I zamijenjene referencama sistema II, i sva svoj-

stva i simboli za odnos u F su reinterpetirana po uzoru na mapiranje” (Putnam, 1979, 292).

Posebno je važno istaknuti da, kada ovako predstavimo Putnamovu teoriju, tj. da dva sistema mogu biti različitoga sadržaja, prisjetimo se klasičnog problema uma i tijela kao odnosa fizičkog i duhovnog, gdje se ova dva različita stanja slažu iako su izomorfna. Nije bitno na kraju šta je um ili duh ili neko drugo svojstvo, koje god svojstvo imali, bitna je funkcija. Ako bismo, primjera radi, upoznali neke ljude sa druge planete, njihova mentalna svojstva se na jednome materijalnome nivou mogu razlikovati od čovjekovih, npr. neke moždane stanice su materijalno drugačije od naših, ali to ne znači da su i naša mentalna stanja drugačija, jer mogu operirati funkcionalno izomorfno kao što je prethodno navedeno. Mentalna stanja nije potrebno identificirati sa materijalnim niti duhovnim (onom podjelom zaoštrenom od Descartesesa). Putnam želi naglasiti svojom teorijom da sve što nam treba jesu dva modela input i output stanja (simbola koji predstavljaju određene komande i vanjskih stanja koji tu komadu pokazuju kao izvršenu) ili sistema koji su ostvarenje jedne konceptualne i apstrahirajuće strukture.

Ako je zaista Putnam u pravu, zašto se kroz povijest toliko pažnje posvetilo problemu uma – tijela kao problemu odnosa materijalnoga i duhovnog (nematerijalnog)? Općenito u povijesti filozofije filozofi su se zalagali ili za dualističko ili za materijalističko objašnjenje ovoga problema. Putnam smatra da su obje teorije pogrešne zbog toga što, ako smatramo da je naš mentalni život fizičkog ili spiritualnog sastava, on će se morati nekako odraziti fizički u našem ponašanju. Putnamov primjer izgleda ovako: “zamislimo ploču koja ima dvije rupe, jedna je okrugla, a druga je četvrtasta, i neku manju kocku koja treba da prođe kroz neku od rupa. Kocka će proći kroz četvrtastu rupu” (Putnam, 1979, 295).

Sada kada bi trebalo objasniti ovaj događaj, prvo što bismo radili bilo bi da opišemo fizička svojstva kocke i rupe u koju je upala. To bismo uradili zbog toga što je čovjek vođen logikom da, ako je nešto izgrađeno od neke materije, mora imati objašnjenja na fizičkome nivou. To znači da to što je kocka upala u rupu treba da objasnimo time što je kocka takvoga i takvoga stanja atoma i molekula. Međutim, problem se javlja zato što takvim objašnjenjem nismo ništa rekli osim sastava stvari⁸. Da bi teo-

⁸ Rorty je za ovaj Putnamov primjer rekao “kako je Putnam otkrio ontološku prepreku između opisivanja mikrostrukture i makrostrukture fenomena” (para. Rorty, 1990, 38).

rija bila uspješna, ona mora da ocrta objašnjenje. “Pod objašnjenjem jedne zapažene činjenice razumijevamo njeno uklapanje u kakav opći zakon” (Reichenbach, 1964, 36). U ovome primjeru treba dati jedan općenitiji zakon ili objašnjenje sa geometrijske razine, a ne sa atomske. To je općenitije objašnjenje koje će se moći primijeniti na više sličnih slučajeva, jer molekularna razina kocke nam ništa ne govori o tome zašto je mogla proći krug. Ako bismo zamijenili prvu kocku sa drugom, dolazi druga, a ta druga ima i drugu atomsku strukturu. Poenta koju izvlačimo iz Putnamovog primjera jeste da “kako god mentalno funkcioniralo, izgleda da nema ozbiljnoga razloga da se vjeruje da je objašnjivo našim fizičkim i kemijskim svojstvima” (Putnam, 1979, 297).

5.0 Čovjek i mašina

Pitanje koje nastaje, a koje vuče svoje korijene iz filozofske tradicije, jeste: Kako može mašina znati/tvrditi da je ispravno izvršila neku komandu? To je pitanje koje potječe iz klasične novovjekovne teorije znanja (od Descartesa pa do Kanta i na dalje), a odgovor na to uključuje tri bitne informacije: 1. opise sekvenci stanja kroz koja je mašina prošla, 2. opis pravila prema kojima mašina operira (onih u tablici mašine), 3. objašnjenje ovoga procesa u cjelini. Odnos simbola kojima mašina barata jeste logički opis mašine (tablica mašine). Fizički (materijalni) opis mašine ne bio nematerijalan u smislu da su stanja kroz koja mašina prolazi naznačena samo u tablici. Kako može mašina znati neki simbol X? “Samo ako opisuje sekvence stanja kroz koje T^9 (...) prolazi kada tvrdi da je X^{10} , poslušana pravila itd., ali nema sekvence stanja kroz koju T mora da prođe da bi bio u jednom stanju” (Putnam, 1979, 367). To znači da mašina može biti u stanju A ili B, ali ne mora znati da je u tome stanju. Možemo tako opisati i čovjekovo iskustvo prema nekome objektu X. Čovjek ne mora znati stanja ili koje vrste čulne spoznaje on ima prema objektu X, a opet može imati znanje o tome objektu, da je ono ispred njega i da postoji.

Putnam ovaj problem slikovitije objašnjava kada je, npr., riječ o sudu “Ja imam bol”; “Nešto me boli”. Kad je čovjek u stanju boli, on nema svjesnog, reflektivnog, diskurzivog momenta da zaključi da ga nešto boli. Tako i mašina, kada je u nekome stanju, nije potrebno da to

⁹ Turingova mašina

¹⁰ simbol iz tablica mašine

zna ili nije potrebna jedna protokolna propozicija o tome da bi se došlo do nekoga oblika znanja. Sudu X je u boli, ne treba daljnje pitanje kako X zna da zna da li je u boli. Isto se dešava i kada je u pitanju mašina. Trenutak boli je jednostavno jedna verbalna ekspresija toga momenta i nije potrebna dodatna racionalizacija nad njom da bi se znalo na šta se misli.

Kada Turingovu mašinu opisujemo sa pozicije njene tablice mašine, opisuje se onaj dio mašine koji se odnosi na traku, dio za ispisivanje i skeniranje i jedan konačan skup stanja u kojima može mašina biti. Putnam ovo naziva “logičkim opisom” Turingove mašine koji se razlikuje od fizičkoga opisa mašine, tj. svih fizičkih dijelova sve mašine (tu se opisuje ono što Putnam naziva strukturalna stanja). Oba ova opisa govore o načinima pomoću kojih se može opisati stanje jedne mašine¹¹. Logički opis jeste opis stanja mašine na temelju manipulacije simbola, dok fizički ili strukturalni opis jeste opis dijelova iz kojih je mašina izgrađena i kako ti dijelovi međusobno funkcioniraju. Logička stanja mašine (software) su mentalna stanja kod čovjeka, a strukturalna stanja mašine (hardware) su fizička stanja čovjeka.

Ako se zamisli “mašina koja može otkriti neka svoja strukturalna stanja, onda je u poziciji veoma slična čovjeku koji može otkriti neke, ne sve, kvarove u svome tijelu, sa raznim stupnjevima sigurnosti” (Putnam, 1979, 372). Ako se mašini neki dio pokvari, može li ona znati u tome momentu da joj je baš taj dio pokvaren? Odgovor na ovo pitanje za Putnama je da oba načina interpretiranja ovoga problema mogu biti solucije, jer je moguće vidjeti na temelju manipulacije simbola da nešto nije u redu sa mašinom. Ako se ide fizičkim objašnjenjem stanja, isto tako se može vidjeti da je neki dio mašine u kvaru, npr. neka cijev i sl. Ono što je ovdje bitno istaknuti jeste to da logička, odnosno mentalna stanja možemo promatrati ne referirajući se na fizička i strukturalna stanja. Drugo, bitno je da se mentalna stanja ne mogu predstaviti kao identična fizičkim stanjima i obrnuto (nema jedinstva identiteta). Ako se kaže da je stanje mašine (recimo stanje A) jednako tome što mašina u tome stanju ima neke dijelove pokrenute, je neodrživa teza jer stanje A i rad nekih dijelova nisu sinonimi u jeziku mašine. To su samo opisi stanja.

¹¹ “Funkcionalizam ostavlja mjesta za posebne nivoe opisivanja. Ponašanje kompjuterske mašine može se istraživati na temelju njenog fizičkog sastava ili istraživanjem njegova programa” (Heil, 2001, 400). No funkcionalizam se više zanima i njegovi odgovori se kreću u sferi “višeg nivoa” objašnjavanja ili onoga što je kod Putnama logički opis, dok je “niži nivo” objašnjenja opisivanje fizičkih dijelova.

6.0 Mentalna stanja kao funkcionalna stanja

Kada uzmemo u obzir tvrdnju: “tradicionalni materijalizam drži da se mentalna ponašanja riječima mogu odrediti izjednačavanjem tih koncepata sa fizičko-kemijskim svojstvima” (Putnam, 1979, 414), možemo pitati sljedeće: ako je to točno, onda ako X preferira A u odnosu na B, moglo bi se odrediti prema fizičko-kemijskim sastavima čovjeka ili mašine da preferiraju A ili B. To ako zamislimo da je Turingova mašina slična čovjeku. Problem je što za tvrdnju X preferira A u odnosu na B nema validne logičke inferencije iz koje se može izvući da ta tvrdnja korespondira sa fizičko-kemijskim svojstvima. Ta tvrdnja je logički nezavisna zbog toga što je izjava takvoga tipa uvjetovana drugim stvarima, kao npr. vokabularom kojim operira jedna mašina ili čovjek, značenjem ili upotrebom simbola jednoga vokabulara. Ako opet pogledamo primjer Turingove mašine koja ima traku sa simbolima što daju komande prema kojima se mašina ponaša, bez obzira na to što su one izvor vokabulara, logička stanja mašine ne mogu korespondirati sa fizičkim stanjem mašine. Fizičko stanje jeste bitan element opisivanja čovjeka ili mašine, ali jedino kao preduvjet da bi se logička stanja mogla izraziti. Fizičko se ne može usporediti s onim što se nazivalo mentalno ili logičko.

“Većina funkcionalista su spremni dopustiti da je svaka pojedinačna bol” (primjera radi jer je i Putnam taj primjer obradio) “fizičko stanje ili dešavanje, i doista da za svaku bol –osjećaj organizma, tu je (vjerojatno) jedno fizičko stanje koje ostvaruje tu bol u jednoj vrsti organizma. Tamo gdje se funkcionalisti razlikuju od fizikalista¹² (...) je pitanje šta je zajedničko svim bolovima (...). Funkcionalist kaže da to što je zajedničko je funkcija (...). Funkcionalisti mogu biti fizikalisti u dozvoljavanju da su svi entiteti (stvari, stanja, dešavanja itd.) koji postoje fizički entiteti, ne slažući se samo s tim da to što povezuje sve vrste stvari zajedno je fizičko svojstvo” (Block, 2004, 184).

Tako naprimjer, ako bol uzmemo kao mentalno stanje i kažemo da je bol samo stimuliranje nekih nervnih stanica mozga čovjeka, nekih X stanica, onda znači da bol kao stanje jeste ako te stanice stimuliraju. Problem se stvara u tome što životinje mogu osjetiti bol, ali je mapa njihova

¹² Fizikalisti u ovome kontekstu su ekvivalencija materijalistima.

nervnoga sistema potpuno drugačija te se neće moći bol izjednačiti sa stimuliranjem X stanica koje ima čovjek. Jer ako bi se to uradilo, značilo bi da samo čovjek osjeti bol. Drugi problem: ako bismo bol ili bilo koje drugo mentalno stanje izjednačili sa nekim stanjem mozga (stanjem kao fizičko-kemijskim), to bi se moglo analizirati kao “biti u stanju A je biti u stanju B”, tj. da ako netko ima bol, on je u tome trenutku točno jedno stanje mozga. Ova tvrdnja može imati filozofski smisao ili biti filozofski informativna ako se da reducirati. Tako naprimjer, “biti u stanju boli je imati određen neugodan osjećaj je filozofski neinformativno; biti u stanju boli je imati određenu bihevijalnu narav je, ako je to točno, filozofski informativno” (Putnam, 1975, 429). Do ovoga problema dolazi zbog mijenjanja termina svojstvo i koncept.

Svojstvo je termin koji Putnam koristi kao “Blanket term¹³” za stvari kao što su: biti u stanju boli, biti u nekome određenom stanju mozga, imati određenu bihevijalnu narav. Koncept je termin kojim Putnam želi označiti sinonimni skup nekih izraza¹⁴. Koncept kao što je temperatura može se izjednačiti sa sinonimnim skupom riječi temperatura i toplina, ako se uzme da su sinonimi (ili drugi primjer: broj dva se može poistovjetiti sa skupom svih parnih brojeva, broj dva i parni broj mogu biti sinonimi). Ako se pokuša postaviti pitanje koja su svojstva temperature, odgovor se neće kretati u smjeru jednoga fizikalnoga opisa. Primjer, temperatura je kinetička energija može se tumačiti kao da je kinetička energija svojstvo temperature, ali se ne može tumačiti kao koncept za temperaturu. U ovu zamku su zapali filozofi koji su redukcijom probali riješiti problem, a time su izjednačili svojstva i koncepte, što je, prema Putnamu, greška. Ako istu analogiju primijenimo, naprimjer, sa boli kao stanju mozga, ta analogija neće biti ispravna. Jer ako bi bol bila X, bila bilo šta, bol i X ne bi trebali biti sinonimi. Za Putnama bol jedino može biti funkcionalno stanje jednoga organizma koje ima bol.

Za objašnjenje ovoga organizam će biti Turinogva mašina ili jedna vrste te mašine. Samo što će u ovome slučaju biti to mašina čija promjena stanja neće biti unaprijed određena nego je probabilistička. Dalje će imati osjetne inpute i motor outpute, što znači da “tablica mašine specificira svaku moguću kombinaciju stanja i kompletni set osjetnih inputa i uputstva koja određuju vjerojatnost idućega stanja, i također vjerojatnost

¹³ Idiom koji znači da jednim terminom označavamo više stvari ili više grupa stvari koje su međusobno povezane.

¹⁴ Za Putnama koncepti nisu sinonimni skupovi, ali se mogu poistovjetiti sa sinonimnim skupovima.

motor outputa¹⁵” (Putnam, 1975, 434). Za daljnje razumijevanje ove mašine potrebno je ono što Putnam naziva “opis sistema”. “Deskripcija S, gdje S je sistem, je istinita dotle dok S posjeduje posebna stanja S1, S2... Sn koja su međusobno povezana i povezana sa motor outputom i osjetnim inputom sa prijelazom vjerojatnosti koja je tada tako-i-tako u tablici mašine” (Putnam, 1975, 434). Tablica mašine je funkcionalna organizacija jednoga sistema, a taj sistem se sastoji iz više stanja što ga čine jednom cjelinom, koja je opet bitna za “opis sistema”. Znati cjelinu stanja, koja je relevantna jednome opisu, je znati kako će se jedna mašina ponašati na temelju inputa i outputa. Za sve ovo nije potrebno nikakvo poznavanje fizičko-kemijskih svojstava mašine.

Svoju hipotezu o boli kao funkcionalnim stanjima Putnam izlaže u četiri stavke: “(1) Svi organizmi koji su sposobni osjećati bol su probabilističke mašine. (2) Svaki organizam koji može osjećati bol posjeduje jednu vrstu opisa (npr. biti u mogućnosti osjećati bol znači imati odgovarajuću vrstu funkcionalne organizacije). (3) Niti jedan organizam koji može osjećati bol ne može se rastaviti na dijelove koji posjeduju opise referirane pod stavkom (2). (4) Za svaki opis referiran pod (2) postoji podset osjetnih inputa takav da je organizam koji posjeduje opis u boli ako i samo ako su neki od njegovih osjetnih inputa u tome podsetu” (Putnam, 1975, 434).

Implikacija Putnamove četiri stavke je da sva mentalna svojstva, bilo to bol ili nešto drugo, tumačimo kao funkcionalno stanje. To znači da je neko mentalno stanje tako-i-takvo jer ima određen opis. Funkcionalna stanja nisu samo puki korelat mentalnim stanjima nego pokušavaju opisati ponašanje stanja organizma o kome je riječ.

7.0 Zaključak

Cilj rada je bio pokazati kako su tradicionalne teorije unutar problematike kojom se bavi filozofija duha (a to se odnosi na dualizam i materijalizam) u svome početku bile nejasne teorije, te kako kategoriziranje mentalnoga sa materijalnim ili spiritualnim (propadljivim ili nepropadljivim) postaje prazna kategorija. Razvoj kompjuterske tehnologije i kognitivnih znanosti, te posebice kompjuterskih mašina, potaknuo je filozofe

¹⁵ Motor u ovome kontekstu znači stvarati ili proizvoditi, a motor output je ono što nastaje stvaranjem na temelju inputa koji se dobivaju.

poput Putnama na ponovno ispitivanje onoga što se nazivalo mentalnim. Mentalno više nije moglo biti stvar, nego opis funkcija (odnos inputa i outputa, ili odnos percepcija i ponašanja koje čovjek vrši) koje vrši mašina ili čovjek. To je jedno konceptualno/logičko objašnjenje rada mentalnoga koje se razlikuje od tradicionalnih gledišta u filozofiji uma. Funkcionalistička teorija zahtijeva “viši nivo” opisa mentalnoga, tj. opis funkcioniranja ili rada jednoga organizma. Materijalističko gledište funkcionalisti poput Putnama kritikuju sa dozom opreza, dok, kada je riječ o kartezijskom dualizmu, on se pokušava u potpunosti prevladati. Razlog opreza je u tome što materijalizam nudi učenje poput bioloških ili fizioloških stanja jednoga organizma, što može pridonijeti u razumijevanju mentalnog, ali ne može ga objasniti. Zato Putnam, kada koristi računarsku metaforu, uzima je ne radi toga što smatra da fizička stanja jednoga organizma daju odgovor na pitanje “Što je mentalno?”, već zbog toga što postoji paralela prema kojoj različiti organizmi sa različitim fizičko-kemijskim svojstvima mogu imati ono što se karakterizira kao mentalno.

BIBLIOGRAFIJA

1. Block, Ned (2004): What is functionalism? U: Heil John (ur.) (2004): *Philosophy of Mind (A Guide and Anthology)*. Oxford: Oxford University Press.
2. Fodor, Jerry (2004): The mind-body problem. U: Heil John (ur.) (2004): *Philosophy of Mind (A Guide and Anthology)*. Oxford: Oxford University Press. Gaynesford, Maximilian (2006): Hilary Putnam (*Philosophy Now*). Acumen Publishing Limited.
3. Heil, John (2004): Introductio. U: Heil John (ur.) (2004): *Philosophy of Mind (A Guide and Anthology)*. Oxford: Oxford University Press. (str. 139–149)
4. Heil, John: Hilary Putnam (1926–). U: A. P. Martinich i David Sosa (ur.) (2001): *Blackwell Companions to Philosophy: A Companion to Analytic Philosophy*. Blackwell Publishers Ltd.
5. Rajhenbah, Hans (1964): Rađanje naučne filozofije (*The Rise of Scientific Philosophy*). Beograd: Nolit.
6. Mišćević, Nenad (1989): Uvod. U: Nenad Mišćević i Nenad Smokrović (ur.) (1989): *Kompjutori, mozak i ljudski um*. Rijeka: Izdavački centar Rijeka. (str. 7–34)

7. Putnam, Hilary (1979): *Philosophical Papers, Volumen 2: Mind, Language and Reality*. Cambridge: Cambridge University Press.
8. Putnam, Hilary (1981): *Reason, Truth and History*. Cambridge: Cambridge University Press.
9. Putnam, Hilary (1989): *Računarska psihologija i teorija interpretacije*. U: Nenad Miščević i 10. Nenad Smokrović (ur.) (1989): *Kompjutori, mozak i ljudski um*. Rijeka: Izdavački centar Rijeka.
11. Putnam, Hilary (1999): *The threefold cord: mind, body, and the world*. New York: Columbia University Press.
12. Rorty, Richard (1990): *Filozofija i ogledalo prirode (Philosophy and the Mirror of Nature)*. Sarajevo: Veselin Masleša.
13. Shagrir, Oron (2005): *The Raise and Fall of Computational Functionalism*. U: Ben-Menahem, Yemima (ur.) (2005): *Hilary Putnam (Contemporary Philosophy in Focus)*. Cambridge: Cambridge University Press.
14. Turing, Alan M. (2004): *Computing machinery and intelligence*. U: Heil John (ur.) (2004): *Philosophy of Mind (A Guide and Anthology)*. Oxford: Oxford University Press.