

Znano je, da mora vsaka kognitivna psihološka teorija, ki ima poleg pojasnjevalnih tudi prediktabilne lastnosti, subjektivemu vedenju pripisati neko vrsto ali stopnjo racionalnosti. Da bi bilo dejanje racionalno, mora biti posledica primerno povezanih vsebin delovalčevih mentalnih stanj, prepričanja in želja. Opazovanje odnosov med množico vsebin prepričanja in želja in z njimi primerno povezane dejavnosti omogoča predvidevanje človekovega vedenja in v skladu s tem pripisovanje racionalnosti.

Očitno je kot pogoj racionalnosti potrebno predpostaviti, da delovalec poseduje določene logične sposobnosti, oziroma sposobnosti rezoniranja. Nekatera od teh rezoniranj morajo biti deduktivna. Delovalec mora namreč biti v stanju, da glede na dano množico prepričanj opazi in izbere pravilna rezoniranja ali inference, ki so za njega primerne. Inferenco v tem primeru pojmuje kot poseben primer dejavnosti. Pomembno je, da je lahko inferenca ali sposobnost rezoniranja, če jo pojmuje kot dejavnost, nezavedna ali nereflektirana (na primer z evolucijo "vgrajeni" mehanizem).

Vprašanje, ki neposredno sledi iz teh uvodnih opomb, je vprašanje o tem, kakšno inferencialnost oziroma točneje, kakšno deduktivno sposobnost mora imeti delovalec, da lahko njegovo vedenjsko aktivnost smatramo za racionalno. Sodobna literatura nudi dve okvirni možnosti¹. Prvo stališče, ki delovalca pojmuje kot idealnega racionalnega delovalca, bomo imenovali stališče idealne racionalnosti. Drugo stališče na delovalca gleda kot na končno bitje, ki v končnem času razpolaga z omejenimi deduktivnimi sposobnostmi. Cilj tega stališča je, da najde minimalno inferencialno osnovo, na podlagi katere je še vedno možno pripisati vedenju racionalnost. Teorijo minimalne racionalnosti je na najpopolnejši način razvil Christopher Cherniak.²

Začrtajmo osnovne razlike med tema dvema stališčema glede na deduktivne pogoje, ki jih vsako od njih postavlja pri racionalnosti.

Najstrožji in najbolj idealiziran pogoj stališča idealne racionalnosti je pogoj *deduktivne sklenjenosti*. Po njem racionalni delovalec, ki poseduje množico prepričanj, verjame vse konsenkvence vsakega člana množice. Takšen pogoj racionalnosti najdemo v nekaterih epistemoloških logikah, na primer pri Hintikki.³ Čeprav je ta pogoj upravičen glede na potrebno aksiomatizacije logike prepričanja, je očitno nesprejemljiv kot del

¹ Tretje možnosti, po kateri za pripisovanje racionalnosti ni potrebna nikakršna inferencialnost, ne bomo upoštevali.

² Cherniac, C., *Minimal Rationality*, A Bradford Book, The MIT Press, Cambridge, Mass., 1986.

³ Hintikka, J., *Knowledge and Belief*, Itaca, Cornell University Press, 1962.

psihološkoteoretične razlage racionalnosti. Če predpostavimo, da so vsebine prepričanja propozicije strukturirane kot stavki, je jasno, da niti enemu končnemu bitju ne moremo pripisati sposobnosti zavestne kontrole nad neskončno polo konsekvenc takih stavkov.

Oslabitev pogoja deduktivne sklenjenosti se lahko doseže na ta način, da se od idealnega delovalca zahteva, da iz dane množice prepričanj naredi vse pravilne inference, ki so očitno primerne. Pri tej oslabitvi so vse konsekvence prepričanja, ki se jih mora delovalec zavedati, zožene na očitno primerne inference. Ta pogoj, ki ga po Cherniakju lahko imenujemo pogoj *idealne inference*, je široko sprejet v teorijah odločanja in v teorijah iger kot osnova za maksimiziranja koristnosti. Pogoj idealne inference je še vedno preoster za namen psihološke razlage in predvidevanja vedenja. Možno je namreč, da je delovalec v stanju, da opazi, kakšna inferencialnost mu je potrebna za zadovoljitev določene želje ali cilja, vendar obenem ni v stanju, da zahtevano inferenco izvede.

Nadaljna oslabitev, ki še pripada stališču idealne racionalnosti, je pogoj *idealne konsistence*. Po njem bo delovalec z določeno množico prepričanj in želja eliminiral vsako nekonsistenco znotraj te množice. Takšna podoba idealne racionalnosti je značilna za Quinov in Davidsonov princip naklonjenosti (principle of charity), ki vodi razumevanje izjav in interpretacijo vedenja. Če ga apliciramo na preference, pogoj idealne konsistence zahteva tranzitivnost preferenc delovalca v določenem trenutku (če delovalec preferira a nad b in b nad c, tedaj bo preferiral tudi a nad c). Tversky in Kahneman sta v nizu člankov pokazala praktično nevzdržnost pogoja idealne konsistence.⁴

Vsi navedeni pogoji idealne racionalnosti vsebujejo tudi normativno komponento. V primeru pogoja deduktivne sklenjenosti (isto velja tudi za ostale pogoje) normativnost zahteva, da kadar delovalec verjame p in ima p konsekvenco q, mora delovalec verovati q.

Alternative idealni racionalnosti bodo težile k minimalizaciji idealne in normativne komponente. Naloga razlage vedenja na osnovi minimalne racionalnosti je več kot delikatna. Po eni strani z oslabitvijo pogojev inference, ki so neizogibni za pripisovanje racionalnosti, zapademo v nevarnost, da izgubimo vsakršno inferencialno osnovo; po drugi strani pa se s trditvijo, da je določena stopnja oslabitve pogoja pravzaprav zahtevana stopnja minimalne inference, soočimo s problemom nedoločenega ali ad hoc kriterija za določitev minimalnih pogojev. Teorija minimalne racionalnosti C. Cherniakja išče osnovo minimalne inference v, za delovalca v določenem trenutku, možnih ali uresničljivih deduktivnih korakih. Normativnost zahtev inferencialnosti se poskuša eliminirati tako, da se smatrajo za deskriptivne, ne pa tudi za nujne pogoje za posameznega delovalca. Menimo, da je osnovni problem pri takšni rešitvi minimalne racionalnosti v nedoločenosti kriterijev, ki definirajo potrebno stopnjo inferencialnosti. Tako ima Cherniakova teorija precejšnjo heuristično vrednost, medtem ko so njene razlagalne in prediktabilne komponente vsekakor take, da je o njih potrebno razpravljati.

V tem tekstu se ne bomo zadrževali pri diskusiji in eventualni kritiki Cherniakove pozicije. Namesto tega bomo poskušali razložiti in braniti neko drugačno strategijo razlage minimalne racionalnosti. Teorija minimalne racionalnosti, ki jo želimo

⁴ Napotujemo na novejši zbornik del te usmeritve, *Judgment under Uncertainty: Heuristics and Biases*, edited by Amos Tversky, Daniel Kahneman, Paul Slovic, Cambridge University Press, 1984.

prikazovati, je teorija Roberta Stalnakerja.⁵ Stalnakerjeva teorija obljublja, da bo dala razlago racionalnega vedenja končnih in nepopolnih bitij, ki izgleda naravno in zadovoljuje naše običajne intuicije o tej vrsti racionalnosti. Nadalje ta teorija, ki temelji na drugačnih predpostavkah kot večina teorij idealne racionalnosti, uspešno rešuje problem normativnosti in istočasno nima težav z nedoločeno kriterija za pripisovanje minimalne racionalnosti.

Specifičnost Stalnakerjevega izhodišča je v njegovi pragmatično-dispozicijski strukturi, ki je izražena v terminih semantike možnih svetov. Pragmatično-dispozicijska podoba opisuje stanje prepričanja in vsebino prepričanja. Dispozicijska komponenta te podobe določa stanje prepričanja. Po njej,

verjeti da P pomeni, delovati na način, ki bo težil k zadovoljitvi želje v svetu (svetovih), v katerem je P resničen.

Takšna opredelitev prepričanja izloči razred primernih dejavnosti (primernih za delovalca), ki težijo k zadovoljitvi njegovih interesov in želja v situacijah, v katerih je P resničen. Druga komponenta pragmatične podobe se nanaša na določitev vsebine ali objekta prepričanja.

Vsebina ali objekt prepričanja (imenovali ga bomo S-propozicija) je funkcija iz množice možnih svetov v resnične vrednosti. Ti dve komponenti definirata semantični okvir za določanje stanja prepričanja. Sestavljen je iz množice svetov in relacije napram S-propozicijam, ali iz tistih možnih svetov, v katerih je neka propozicija resnična.

Da bi se izognili terminološki zmedbi, je potreba poudariti naslednje: z S-propozicijami razumemo funkcijo, ki razločuje med množico možnih svetov, v katerih je neka izjava resnična, in množico, v katerih je lažna. S terminom izjava ali stavek označujemo tisto, kar se v večini drugih psiholoških teorij imenuje "propozicija", ali tisto, o čemer je prepričanje.

V skladu s pragmatično podobo imeti prepričanje pomeni biti v relaciji napram S-propoziciji ali množici možnih svetov, v katerih je izjava s pomočjo katere se da opisati tisto, o čemer je prepričanje, resnična. Namesto o možnih svetovih, lahko govorimo tudi o logično možnih situacijah, si čimer se na tej stopnji razlage nič bistvenega ne spremeni. Tedaj stanje prepričanja lahko definiramo kot neprazno množico možnih situacij ali, v skladu z dispozicijsko komponento, kot dispozicijo, ki se izraža v različnih vrstah situacij.

Opis stanja prepričanja kot relacije napram S-propozicijam omogoča, da se razlaga racionalnosti razširi na veliko obsežnejši razred bitij kot v primeru teorij, ki stanje prepričanja obravnavajo kot množico sprejetih stavkov. Tako na primer, če pripisujemo našemu psu Bobiju prepričanje, da je kost zakopana na vrtu, trdimo, da je Bobi disponiran, da deluje na način, ki bo težil k temu, da bi bil zadovoljiv v primeru, da je njegova "misel", da je kost zakopana na vrtu, resnična. To, da Bobi ve ali razume, da je prepričanje, ki se lahko izrazi s stavkom "kost je zakopana na vrtu", resnično, pomeni kratko malo, da ima določen mehanizem razlikovanja možnih situacij, v katerih je ta misel resnična, od tistih, v katerih ni.

Vsebina prepričanja torej ni stavek in stanje prepričanja ni izraženo s pomočjo ujemanja s stavkom ali s sprejemanjem stavka. Nasprotno, kakšna inferencialnost ali

⁵ Stalnaker, R., *Inquiry*, A Bradford Book, The MIT Press, 1984.

kakšna deduktivna sposobnost je potrebna, da bi delovalcu pripisali racionalnost, če je vsebina prepričanja S-propozicija.

Vprašanje že sugerira odgovor.

V kolikor je vsebina stanja prepričanja določena kot množica (možnih svetov ali alternativnih možnih situacij), tedaj bo inferencialnost določena s strukturalnimi relacijami med elementi množice. Hartly Field je predlagal,⁶ da naj bi te relacije tvorile Booleovo algebro (kompletno in atomarno). Mi ne bomo vztrajali na tem aspektu. Pogledjmo, katere relacije med člani množice sam Stalnaker vidi kot nujne deduktivne pogoje, ki so postavljeni za množico propozicij, da bi predstavljali stanje prepričanja (ali mentalno stanje nasplošno). Stalnaker postavlja tri pogoje:

1. Če je P član množice propozicij, v katere subjekt verjame, in ima P za posledico Q, tedaj je tudi Q član te množice.
2. Če sta P in Q vsak zase člana množice, tedaj je tudi njuna konjunkcija, P & Q, član množice.
3. Če je P član množice, tedaj negacija od P ni član množice.

Opazimo, da so ti trije navedeni deduktivni pogoji ravno najstrožji pogoji idealnih teorij racionalnosti. Gre za pogoje deduktivne sklenjenosti, konjunkcije in konsistence. Ali ni s tem Stalnakerjeva teorija izgubila vsako upanje, da ponudi konkurenčno razlago minimalne racionalnosti? Odgovor je negativen. Ne samo, da ni, temveč to predstavlja njeno prednost pred drugimi teorijami.

Strategija teorije je sledeča. Naravno stališče, ki ni v nasprotju z našimi intuicijami, je, da je delovalec lahko v istem trenutku v več kot enem stanju prepričanja. V skladu s pragmatično-dispozicijske podobo je prepričanje dispozicijsko stanje, ki je odvisno od situacij in določa primernost dejavnosti. Torej: "... mnoga ločena stanja prepričanja pomenijo mnoge ločene dispozicije, od katerih ima vsaka svoje lastno področje izražanja. To je lahko plavzibilno samo v kolikor obstaja naraven način povezovanja ločenih prepričanj z dejavnostmi...".⁷

Nadaljnji koraki te strategije so sestavljeni iz tega, da se pokaže, kako idealni pogoji dedukcije ne normirajo inferencialnosti posameznih delovalcev. To je možno ob pomoči navedene predpostavke.

Sedaj imamo možnost pokazati drugo pomembno prednost Stalnakerjeve teorije racionalnosti. Njegova teorija je v stanju določiti idealne pogoje racionalnosti, ki istočasno nimajo normativne vrednosti, s katero bi zanikali racionalnost bitjem, ki teh idealnih pogojev ne izpolnjujejo. Idealna racionalnost je določena, v skladu z nakazano podobo, kot edinstveno integrirano stanje prepričanja, ki zadovoljuje zgornje pogoje. Končna, neidealna bitja so racionalna, kadar njihova posamezna ločena in neintegrirana stanja prepričanja zadovoljujejo te pogoje.

Na tej stopnji je možno potegovati analogijo med Stalnakerjevo podobo racionalnosti in epistemološkimi teorijami, ki opisujejo rast znanstvenega vedenja. Naša individualna racionalna kompetenca bo naravno težila k temu, da pod vplivom novih informacij integrira različna ločena stanja prepričanja. Prav iz tega manevra integriranja je sestavljena sposobnost deduktivnega raziskovanja. Podobno se dogaja tudi v

⁶ Glej: Field, H., Stalnaker on Intentionality, *Pacific Philosophical Quarterly*, 67, 1986.

⁷ Stalnaker, navedeno delo, str. 86.

znanstvenem procesu zamenjevanja starih teorij z novimi. Znanstvenik je v stanju, da pod vplivom novih informacij formulira in sprejme hipotezo, ki bo šele kasneje verificirana. Na nivoju psihološke teorije želi Stalnaker uvesti podobno možnost za posameznega delovalca. To mu uspeva s pomočjo pojma stanja sprejemanja.

Stanje sprejemanja tvori širok razred propozicionalnih opredelitev, katerih sestavni del je stanje prepričanja. Poleg prepričanja stanje sprejemanja vsebuje tudi stanja, kot so predpostavlanje, postuliranje, domnevanje itd. Stanje prepričanja je najtemelnejše stanje znotraj pojma sprejemanja. Za razliko od prepričanja delovalec lahko sprejme neko propozicijo samo za trenutek, pri tem pa ve, da ne bo postala njegovo trajno prepričanje. Prav tako lahko sprejme nekaj kot resnično, na primer z namenom argumentiranja, pri tem pa ve, da bo ovrženo.

Očitno je, da je pojem sprejemanja šibkejši od pojma prepričanja. Ne ravnamo napak, če ga identificiramo s kondicionalnimi prepričanji. S tem, da formulira pojem sprejemanja, se Stalnaker lahko loti določanja strategije integriranja ločenih prepričanj. V tem tekstu se bomo zadovoljili s tako površno določitvijo te strategije. Naša naloga je, da pokažemo, da je koncepcija o ločenih stanjih prepričanja sposobna zadovoljiti deduktivne pogoje.

Pojdimo po vrsti. Pogoj deduktivne sklenjenosti zahteva, da delovalec verjame vse nujne kosekvence vsake propozicije, ki jo verjame znotraj posameznega stanja prepričanja. V skladu z elementi strategije, ki smo jo prikazali, bo obramba opredelitve potekala na sledeči način:

a) prepričanje je dispozicijsko stanje, ki je povezano z dejavnostjo in ki določa primernost dejavnosti,

b) primernost dejavnosti je definirana z upoštevanjem vseh propozicij, ki jih delovalec verjame znotraj posameznega stanja prepričanja,

c) kosekvence stanja prepričanja so določene kot vse ostalo, kar se tiče stanja prepričanja, z upoštevanjem njegove povezanosti z dejavnostjo.

Če je torej delovalec disponiran tako, da deluje na način, ki teži k temu, da bi bil uspešen, če je P resničen, tedaj bo to veljalo tudi za kosekvence P-ja. Temu je mogoče ugovarjati, da v kolikor je delovalčeva dejavnost umeščena v situaciji, v kateri je resnično njegovo prepričanje, da je Večernica planet, to še ne pomeni, da mora verjeti, da je Danica planet. Spomnimo se, da je stanje prepričanja definirano z množico možnih situacij, v katerih je propozicija resnična. Razumeti propozicijo, pomeni opaziti, katere možnosti so izključene z njeno resničnostjo. Propozicija "Danica je planet" ni izključena z resničnostjo propozicije "Večernica je planet". S tem, da subjekt ne ve, da je druga propozicija nujno ekvivalentna prvi, se pokaže samo to, da se moti glede na nekatera lingvistična dejstva, ki konstituirajo izjavo "Večernica je planet".

Na ta način lahko pokažemo, da je upravičeno trditi, da je opredelitev "Večernica je planet" prisotna kot "aktivno" prepričanje v delovalčevem stanju prepričanja, da pa s tem iz njegovega stanja prepričanja ni izključena kosekvence te opredelitve "Danica je planet". Ta druga opredelitev bo vključena v širši in šibkejši pojem stanja sprejemanja kot tiha, nezavedna presupozicija.

Preidimo na drugi deduktivni pogoj. Ta zahteva, da mora delovalec, ki verjame P in verjame Q, verjeti tudi njuno konjukcijo. Pragmatično-dispozicijska podoba racionalnosti zahteva dvojje: da mora idealno racionalni delovalec zadovoljevati ta pogoj, in kot drugo, da nezadovoljevanje tega pogoja s strani končnega delovalca nima za

posledico njegove neracionalnosti. Oglejmo si najprej drugo zahtevo. Ta naravno sledi iz tega, da smo sprejeli pogoj deduktivne sklenjenosti. Njeno plavzibilnost bomo pokazali na primeru, ki izvira od Hartya Fielda. Zamislimo si filozofa, ki sprejema eksistenco neke entitete, recimo propozicij, ko razmišlja o filozofiji uma. Zavrača pa eksistenco te entitete, ko razmišlja v terminih ontologije. Razlaga je naravna, če predpostavimo, da je ista entiteta kot vsebina prepričanja vsebovana v dveh ločenih delovalčevih stanjih prepričanja. V enem kontekstu, kontekstu filozofije psihologije, bo prišlo v razlago delovalčevih misli in dejavnosti eno stanje prepričanja. V kontekstu ontologije bo na področje razlage prišlo drugo stanje prepričanja. Isto vsebino stanja prepričanja bomo razložili glede na dva različna situacijska konteksta in sprejeli, da se delovalec glede na različne kontekste nahaja v dveh različnih stanjih prepričanja.

nahaja v dveh različnih stanjih prepričanja.

Resnejši problem je s prvo zahtevo. Paradoks loterije (lottery paradox) je resen kandidat za blokiranje zahteve, da mora idealno racionalni delovalec zadovoljevati drugi pogoj. Paradoks loterije, ki je poznan v teoriji odločanja, kaže na sledečo težavo. Predpostavimo, da je danih v prodajo milijon loterijskih srečk, od katerih samo ena dobi nagrado. S tem da kupim srečko, je moja možnost, da dobim nagrado, samo ena proti milijon. Naravno je, da bom sprejel hipotezo, da moja srečka ne bo dobila nagrade. Tako hipotezo lahko sprejem tudi za vsako drugo od teh milijon srečk. Sigurno pa je, da bo ena srečka dobila nagrado. Če sprejem stališče, da katerakoli posamezna srečka ne bo dobila nagrade, ne morem racionalno sprejeti tudi konjunkcije teh stališč.

Primer s paradoksom loterije uporabi Kyburg, da bi argumentiral proti tezi, po kateri so deduktivni pogoji in s tem tudi pogoj konjunkcije nujni pogoju racionalnega rezoniranja (s čimer argumentira proti deduktivnemu modelu na splošno). Če bi bilo rezoniranje določeno s pogojem konjunkcije, trdi Kyburg, bi bilo stanje prepričanja determinirano z eno samo propozicijo, ki jo opisuje kot "monumentalno kompleksno konjunktivno stališče". To bi tudi v resnici bilo tako, če bi propozicije razumeli kot stavke ali opredelitve. Razumevanje propozicije kot množice možnih svetov pa otopi ostrino te kritike.

Tretjega Stalnakerjevega deduktivnega pogoja, pogoja konsistence, ne bomo posebej razlagali, ker je jasno, da sledi iz tega, da sprejmemo upravičenost prvih dveh pogojev.

Povzemimo na kratko značilnosti prikazane strategije. Po njej so mentalna stanja razumljena kot dispozicijske sposobnosti. Kot dispozicijska stanja so direktno povezana z dejavnostmi. Področje realizacije dispozicij so alternativni možni rezultati dejavnosti. Vsebina mentalnega stanja, konkretno prepričanje, je tedaj množica možnih svetov ali množica alternativnih možnih situacij ali S-propozicij. Poudarek v strategiji je na tem, da je delovalec lahko istočasno v nekaj različnih stanjih prepričanja, od katerih je vsako predstavljeno z množico možnih situacij.

Idealni racionalni delovalec bo opisan s pomočjo enega samega stanja prepričanja, medtem ko bodo manj popolni delovalci opisani s pomočjo več stanj prepričanja ali več stanj sprejemanja, kar je manj zahteven pojem in predstavlja širšo skupino propozicionalnih opredelitev, ima pa isto strukturo kot stanja prepričanja. Uvajanje stanja sprejemanja med drugim omogoča razlikovanje "aktivnih" in zavestnih prepričanj od nezavednih in presuponiranih. Tako idealno stanje prepričanja kot tudi tisto manj popolno zahtevata zadovoljitev deduktivnih pogojev inferencialnosti.

Menimo, da takšna strategija podaja sprejemljivo in naravno podobo minimalne racionalnosti, osvobojeno težav konkurenčnih teorij.