ANαliZA

časopis za kritično misel

Letnik 6, številka 4, april 2003 · ISSN 1408-2969

FILOZOFIJA SAULA KRIPKEJA

Saul Kripke

Identiteta in nujnost

prevedel Borut Cerkovnik

Citiranje: Saul Kripke, »Identiteta in nujnost«, *Analiza: časopis za kritično misel* 6, št. 4 (2002): 51–76.

Prevod. Licenca Creative Commons, ki velja za arhiv revije, za prevode ne velja; pravice izvirnika pripadajo njegovim imetnikom.

Izdajatelj: Društvo za analitično filozofijo in filozofijo znanosti, Ljubljana.

Elektronska različica je bila leta 2026 rekonstruirana iz izvirnih stavnih datotek in preverjena s tiskano izdajo. Številke strani ustrezajo tiskani izdaji; prelomi strani znotraj prispevka se lahko nekoliko razlikujejo.

Saul Kripke
prevedel Borut Cerkovnik

IDENTITETA IN NUJNOST[*]

Problem, ki se je pogosto pojavil v sodobni filozofiji, je: "Kako so možne *kontingentne* identitetne izjave?" To vprašanje je izraženo po analogiji s tem, kako je Kant izrazil svoje vprašanje: "Kako so možne sintetične sodbe *a priori*?" V obeh primerih je običajno veljalo za samoumevno: pri Kantu, da so sintetične sodbe *a priori* možne, v sodobni filozofski literaturi pa, da so kontingentne izjave o identiteti možne. Razen tega, da sem omenil to analogijo, se ne nameravam ukvarjati s Kantovim vprašanjem: Potem ko je bila napisana precej debela knjiga, ki je poskušala odgovoriti na vprašanje, kako so možne sintetične sodbe *a priori*, so kasneje prišli drugi, ki so trdili, da je rešitev tega problema v tem, da so sintetične sodbe *a priori* seveda nemogoče in da je bila knjiga, ki je poskušala pokazati nasprotno, napisana zaman. Ne bom razpravljal o tem, kdo je imel prav glede možnosti sintetičnih sodb *a priori*. V primeru kontingentnih izjav o identiteti pa je večina filozofov imela občutek, da pojem kontingentne identitetne izjave vodi v nekaj, kar je podobno sledečemu paradoksu. Proti možnosti kontingentnih identitetnih izjav lahko podamo argument, ki je podoben sledečemu:[1]

Prvič, zakon o zamenljivosti identitete pravi, da za poljubna predmeta x in y velja, da če je x identičen z y, potem če ima x določeno lastnost F, jo ima tudi y:

(1) $(x)(y)[(x = y) \supset (Fx \supset Fy)]$

Po drugi strani je vsak predmet gotovo nujno samoidentičen:

(2) $(x)\Box(x = x)$

---

[*] Članek "Identity and Necessity" je izvirno izšel v Milton K. Munitz (ur.), *Identity and Individuation*, New York University Press, New York 1971, str. 135-164, preveden pa je po ponatisu v Jaegwon Kim & Ernest Sosa (ur.), *Metaphysics* (*An Anthology*), Blackwell Publishers, Malden, Mass. & Oxford 1999, str. 72-89. Prevajalske in uredniške opombe so označene z zvezdico ali pa navedene v oglatih oklepajih.

[1] Ta razprava je bila predstavljena ustno, brez napisanega teksta, za predavateljski niz *New York University* o identiteti, ki je tvoril zvezek *Identity and Individuation*. Predavanje je bilo posneto, pričujoča razprava pa predstavlja prepis teh posnetkov, ki so bili urejeni le neznatno, da ne bi bil spremenjen slog izvirnika. Če si bralec zamisli stavke te razprave tako, kot so bili povedani, improvizirano, s pravimi premori in poudarki, lahko to olajša njegovo dojemanje. Kljub temu pa morda še vedno ostajajo odlomki, ki jim je težko slediti, odmerjen čas pa je zahteval zgoščeno predstavitev argumenta. (Daljša različica nekaterih od teh stališč, ki je kljub temu zgoščena in ki kljub temu predstavlja prepis ustnih opazk, je izšla v Donald Davidson in Gilbert Harman (ur.), *Semantics of Natural Language* (Dordrecht: Reidel 1972) [Omenjeno delo *Naming and Necessity* je s predgovorom iz leta 1980 na voljo tudi v slovenskem prevodu: Saul Kripke, *Imenovanje in nujnost* (Ljubljana: Krtina 2000)].) Tu in tam sem moral obrzdati pridržke, dodelave in potrditve mojih opazk, še posebej v razpravi o teoretski identifikaciji in o problemu duh-telo. Če to ne bilo storjeno, bi opombe, ki so bile dodane izvirniku, postale celo bolj nepripravne.

(3) $(x)(y)(x = y) \supset [\Box(x = x) \supset \Box(x = y)]$

pa je substitucijski primerek teze (1) – substitucijskega zakona. Iz (2) in (3) lahko sklepamo, da za vsak *x* in *y* velja, da če je *x* enak *y*, potem je nujno, da je *x* enak *y*:

(4) $(x)(y)((x = y) \supset \Box(x = y))$

To je zato, ker del pogojnika $\Box(x = x)$ odpade, ker se zanj ve, da je resničen.

To je argument, ki je bil velikokrat naveden v nedavni filozofiji. Njegov sklep pa so pogosto imeli za zelo paradoksalen. David Wiggins, na primer, v svoji razpravi "Identity-Statements" pravi:

> Kontingentne identitetne izjave nedvomno eksistirajo. Naj bo $a = b$ ena od njih. Iz njene enostavne resnice in (5) [=(4) zgoraj] lahko izpeljemo "$\Box(a = b)$". Kako pa lahko potem sploh obstajajo kakšne kontingentne identitetne izjave?[2]

Potem pravi, da je na ta argument možnih pet različnih odzivov, zavrne vse te odzive in se odzove sam. Ne nameravam razpravljati o vseh možnih odzivih na to izjavo, omenil bom samo drugega od tistih, ki jih Wiggins zavrne. Ta pravi:

> Morda bi lahko sprejeli ta rezultat in zagovarjali stališče, da ni – pod pogojem, da sta "*a*" in "*b*" lastni imeni – nič narobe. Posledica tega je, da z lastnimi imeni ne moremo tvoriti kontingentnih identitetnih izjav.

Potem pa pravi, da je nezadovoljen s to rešitvijo in da so bili mnogi drugi filozofi tudi nezadovoljni s to rešitvijo, drugi pa so jo kljub temu zagovarjali.

Zaradi česa se zdi izjava (4) presenetljiva? Pravi, da za poljubna predmeta *x* in *y* velja, da če *x* je *y*, potem je nujno, da *x* je *y*. Omenil sem že, da bi lahko kdo ugovarjal temu argumentu zaradi tega, ker je neresnična že premisa (2), to je, da ni tako, da je vsaka stvar nujno istovetna sama s sabo. Dobro, na primer, sem jaz sam nujno samoidentičen? Našel bi se kdo, ki bi morda argumentiral, da v nekaterih situacijah, ki si jih lahko zamislimo, celo ne bi eksistiral in bi bila zato izjava "Saul Kripke je Saul Kripke" neresnična oziroma ne bi bilo tako, da sem samoidentičen. Nemara v nekem takšnem svetu ne bi bilo niti resnično niti neresnično reči, da je Saul Kripke samoidentičen. Dobro, morda je tako, vendar je to v resnici odvisno od filozofskega stališča, ki ga ima kdo o témi, o kateri ne bom razpravljal: to je, kaj naj rečemo o resničnostni vrednosti izjav, ki omenjajo predmete, ki ne eksistirajo niti v aktualnem svetu niti v kateremkoli danem možnem svetu ali protidejstveni situaciji. Na tem mestu se bomo odločili za šibko tolmačenje nujnosti. Izjavo lahko štejemo za nujno, če bi bila, kadar predmeti, ki so omenjeni v njej, eksistirajo, ta izjava resnična. Če bi hoteli biti zelo previdni glede tega, bi se morali spustiti v vprašanje o eksistenci kot predikatu in se

[2] R. J. Butler (ur.), *Analytical Philosophy, Second Series* (Oxford: Blackwell, 1965), str. 41.

vprašati, ali lahko to izjavo preoblikujemo v sledečo obliko: Za vsak *x* je nujno, da če *x* eksistira, je *x* samoidentičen. Tukaj se ne bom spuščal v to posebno obliko subtilnosti, ker ne bo relevantna za mojo glavno témo. Prav tako v resnici ne nameravam obravnavati formule (4). Vsakdo, ki verjame formuli (2), je, po mojem mnenju, zavezan formuli (4). Če sta *x* in *y* isti stvari in če sploh lahko govorimo o modalnih lastnostih kakšnega predmeta, to je, običajno rečeno, če lahko govorimo o modalnosti *de re* in o tem, da ima predmet *nujno* določene lastnosti kot takšne, potem formula (1), mislim, mora veljati. Kjer je *F* kakršnakoli lastnost[*], vključujoč lastnost, ki vsebuje modalne operatorje, in če sta *x* in *y* isti predmet in ima *x* neko določeno lastnost *F*, potem mora *y* imeti isto lastnost *F*. In tako je tudi, če je lastnost *F* sama takšne oblike, da nujno ima neko drugo lastnost *G*, v posebnem primeru lastnost biti nujno identičen z nekim določenim predmetom. Dobro, ne bom razpravljal o sami formuli (4), ker sama po sebi za nobeno posamezno resnično izjavo o identiteti ne trdi, da je nujna. Saj sploh ne pove ničesar o *izjavah*. Za vsak *predmet x* in *predmet y* pravi, da če sta *x* in *y* isti predmet, potem je nujno, da sta *x* in *y* isti predmet. In to – mislim, če razmislimo o tem (vendar če kdo ne razmišlja tako, temu v prid tukaj ne bom argumentiral) – v resnici pomeni nekaj, kar se skoraj ne razlikuje od izjave (2). Ker je *x* po definiciji identitete edini predmet, ki je identičen z *x*, se mi zdi "$(y)(y = x \supset Fy)$" samo malo bolj gostobeseden način, kako povemo "Fx", in tako $(x)(y)(y = x \supset Fx)$ pove isto kot $(x)Fx$, ne glede na to, kaj je "*F*" – v posebnem primeru celo če "*F*" zastopa lastnost nujne identitete z *x*. Torej, če ima *x* to lastnost (nujne identitetne z *x*), jo trivialno ima vse, kar je identično z *x*, kakor trdi (4). Iz izjave (4) pa očitno lahko izpeljemo, da morajo biti različne partikularne izjave o identiteti nujne, kar pa potem velja za zelo paradoksalno posledico.

Wiginns pravi: "Kontingentne identitetne izjave nedvomno eksistirajo." Primer kontingentno identitetne izjave je, da je prvi poštni minister Združenih držav identičen z izumiteljem očal z dvojnimi lečami oziroma da sta oba identična s človekom, za katerega je *Saturday Evening Post* trdil, da je njegov ustanovitelj (mimogrede: sodim, da to zadnje *ni res*). Nekatere takšne izjave so očitno kontingentne. Da je en in isti človek izumil očala z dvojnimi lečami in delal kot poštni minister Združenih držav, je očitno kontingentno dejstvo. Kako lahko to uskladimo z resnico izjave (4)? Dobro, tudi to je zadeva, v katero se ne nameravam podrobno spuščati. O njej bom zelo dogmatičen. Mislim, da jo je dobro uredil Bertrand Russell s svojim pojmom doseg opisa. V skladu z Russellom lahko nekdo, na primer, pravilno reče, da avtor *Hamleta* morda ne bi napisal *Hamleta*, ali celo, da avtor *Hamleta* morda ne bi bil avtor *Hamleta*. Tukaj pa sedaj seveda ne zanikamo nujnosti identitete predmeta s samim seboj; ampak rečemo, da je za nekega določenega človeka res, da je dejansko bil edinstvena oseba, ki je napisala *Hamleta*, in, drugič, da ta človek, ki je dejansko bil človek, ki je napisal *Hamleta*, morda ne bi napisal *Hamleta*. Z drugimi besedami: če bi se

[**] V izvirniku sicer piše "Where *x* is any property at all", vendar Kripke na tem mestu govori o lastnosti.

Shakespeare odločil, da ne piše tragedij, morda ne bi napisal *Hamleta*. V teh okoliščinah ta človek, ki je dejansko napisal *Hamleta*, ne bi napisal *Hamleta*. Russell to vpelje s tem, da reče, da ima v takšni izjavi prvi nastop opisa "ta avtor *Hamleta*" širok doseg.[3] To je, da rečemo: "Ta avtor *Hamleta* ima sledečo lastnost: da morda ne bi napisal *Hamleta*." *Ne* trdimo, da velja sledeče, namreč, da avtor *Hamleta* ni napisal *Hamleta*, ker to ni res. To bi pomenilo, da bi se lahko zgodilo, da je nekdo napisal *Hamleta* in vendar ni napisal *Hamleta*, kar bi bilo protislovno. Nadalje, če pustimo ob strani podrobnosti Russellove formulacije, ki je odvisna od njegove teorije opisov, se zdi, da mora to razlikovanje vsebovati vsaka teorija opisov. Na primer, če bi kdo srečal harvardskega rektorja in ga imel za asistenta, bi morda rekel: "Harvardskega rektorja sem imel za asistenta." S tem pa ne meni, da ima za resnično sledečo propozicijo: "Harvardski rektor je asistent." To bi lahko, na primer, menil, če bi verjel, da je neke vrste demokratični sistem na Harvardu prišel tako daleč, da je njegov rektor sprejel delo asistenta. Verjetno pa to ni tisto, kar meni. Kar res meni, je, kakor je pokazal Russell, sledeče: "Nekdo je harvardski rektor, jaz pa sem ga imel za asistenta." V enem od Russellovih primerov nekdo reče: "Mislim sem, da je tvoja jahta dosti večja, kot je." Drugi pa mu odgovori: "Ne, moja jahta ni dosti večja, kot je."

Pod pogojem, da pojem modalnosti *de re*, in s tem pojem kvantificiranja v modalnih kontekstih, sploh ima kakšen smisel, imamo čisto primerno rešitev, s katero se izognemo paradoksom, če univerzalna kvantifikatorja v (4) zamenjamo z opisi, ker lahko iz tega izpeljemo samo to,[4] na primer v primeru očal z dvojnimi lečami, da obstaja človek, ki se mu je zgodilo oboje – da je izumil očala z dvojnimi lečami in da je bil prvi poštni minister Združenih držav –, in je nujno samoidentičen. Obstaja predmet *x*, ki je izumil očala z dvojnimi lečami, in po naključju tudi predmet *y*, ki je prvi poštni minister Združenih držav, in končno je nujno, da *x* je *y*. Kaj sta tukaj *x* in *y*? Tukaj sta tako *x* kot *y* Benjamin Franklin in gotovo je lahko nujno, da je ta identičen sam s sabo. Če torej sprejmemo Russellov pojem dosega, z opisi ni nobenih težav.[5] Jaz pa hočem tukaj samo

[3] Drugi nastop tega opisa ima ozek doseg.

[4] V Russellovi teoriji sledi $F(\iota x Gx)$ iz $(x)Fx$ in $(\exists!x)Gx$, pod pogojem, da ima opis v $F(\iota x Gx)$ za svoj doseg celoten kontekst (v Russellovi terminologiji iz leta 1905 ima "primarni nastop"). Samo v tem primeru je $F(\iota x Gx)$ "o" denotaciji določnega opisa "$\iota x Gx$". Če to pravilo apliciramo na (4), dobimo rezultat, ki je naveden v tekstu. Upoštevaj, da če ima v dvoumni obliki $\Box(\iota x Gx = \iota x Hx)$ vsaj eden od obeh opisov "primarni nastop", ta formula ne bo zatrdila nujnosti formule $\iota x Gx = \iota x Hx$, če imata oba opisa sekundarni nastop, pa jo bo. Tako morajo biti v jeziku brez eksplicitnih indikatorjev dosega opisi oblikovani z najožjim možnim dosegom – samo v tem primeru bo $\sim A$ negacija A, $\Box A$ necesizacija A in podobno.

[5] Prvotnejše razlikovanje z istim namenom je bilo seveda srednjeveško razlikovanje *de dicto – de re*. Da Russellovo razlikovanje dosega odstrani modalne paradokse, so pokazali številni logiki, še posebej Smullyan.

Da bi se izognil nesporazumu, naj poudarim, da seveda ne trdim, da Russellov pojem dosega reši Quinov problem "esencializma", kar pa pokaže, še posebej v povezavi s sodobnimi modelnoteoretskimi pristopi, je, da kvantificirani modalni logiki – zato da bi se izognila temu, da ima to, da je en in isti človek izumil očala z dvojnimi lečami in vodil prvo poštno ministrstvo, za nujno resnico – ni treba zanikati vseh primerkov oblike $(x)(y)(x = y \bullet \supset \bullet Fx \supset Fy)$ in tudi ne vseh primerkov oblike "$(x)(Gx \supset Ga)$" (kjer naj bo "a" zamenjan z nepraznim določnim opisom, čigar doseg je celota "Ga"). Zato da bi zagotovili te rezultate, ni treba privzeti Russellove kontekstualne definicije opisa; vendar morajo druge logične teorije, fregejanske ali drugačne, ki

dogmatično prekiniti s tem vprašanjem in preiti na vprašanje o imenih, ki ga postavi Wiggins. In Wiggins pravi, da bi morda sprejel rezultat in navaja, da ni, pod pogojem, da sta *a* in *b* lastni imeni, nič narobe. Potem pa to zavrne.

Kaj pa je poseben problem glede lastnih imen? Vsaj če kdo glede tega ni domač s filozofsko literaturo, o lastnih imenih naivno čuti nekaj, kar je podobno sledečemu. Prvič, če kdo reče "Cicero je bil govornik", potem ime "Cicero" v tej izjavi uporabi enostavno za to, da bi izbral nek določen predmet in da bi potem temu predmetu pripisal neko določeno lastnost, v tem primeru namreč nekemu določenemu človeku pripiše lastnost, da je bil govornik. Če kdo drug uporablja drugo ime, kakršno je, recimo, "Tulij", še vedno govori o istem človeku. Če kdo reče "Tulij je bil govorec", pripiše isto lastnost istemu človeku. Lahko rečemo tudi tako: dejstvo oziroma stanje stvari, ki je prikazano s to izjavo, je isto, če nekdo reče "Cicero je bil govorec" ali pa če nekdo reče "Tulij je bil govorec". Zdi se torej, da je funkcija imen *enostavno* v tem, da referirajo, ne pa da opišejo predmete s takšnimi lastnostmi kot "biti izumitelj očal z dvojnimi lečami" ali "biti prvi poštni minister". Zdi se, da Leibnizov zakon in zakon (1) ne velja samo v univerzalno kvantificirani obliki, ampak tudi v obliki "če $a = b$ in Fa, potem Fb", kjerkoli "*a*" in "*b*" stojita na mestu imen, "*F*" pa na mestu predikata, ki izraža neko pristno lastnost tega predmeta:

$$(a = b \cdot Fa) \supset Fb$$

Isti argument lahko ponovimo in sklepamo, kjer "*a*" in "b" zastopata katerikoli imeni, da "Če $a = b$, potem nujno $a = b$". In tako lahko tvegamo sklep: da – kadarkoli sta "*a*" in "*b*" lastni imeni – če *a* je *b*, potem je nujno, da *a* je *b*. Identitetne izjave med lastnima imenoma morajo biti nujne, če so resnične. To stališče je dejansko zagovarjala, na primer, Ruth Barcan Marcus v svojem članku o filozofski interpretaciji modalne logike.[6] V skladu s tem stališčem velja, da kadarkoli, na primer, nekdo pravilno izjavi identiteto med dvema imenoma, kot je, na primer, ta, da Cicero je Tulij, mora njegova izjava biti nujna, če je resnična. Takšen sklep pa *se zdi* očitno neresničen. (Sam imam, tako kot drugi filozofi,

jemljejo opise kot primitivne, nekako izraziti ista logična dejstva. Frege je pokazal, da je nek enostaven, neiteriran kontekst, ki vsebuje nek določni opis z ozkim dosegom, ki ne more biti interpretiran tako, kot da je "o" denotaciji tega opisa, lahko intepretiran tako, kot da je o njegovem "smislu". Nekateri logiki so se zanimali za vprašanje o pogojih, pod katerimi je, v nekem intenzionalnem kontekstu, nek opis z ozkim dosegom ekvivalenten opisu s širokim dosegom. Ena od odlik russellovske obravnave opisov v modalni logiki je, da odgovor na navedeno vprašanje (v grobem je odgovor ta, da je ta opis nek "togi označevalec" v smislu tega predavanja) potem pogosto sledi iz drugih postulatov kvantificirane modalne logike: nobeni posebni postulati niso potrebni – tako kot v Hintikkini obravnavi. Celo če jemljemo opise kot primitivne, so lahko posebne zahteve glede tega, kdaj je doseg nerelevanten, postavljene kot bolj osnovni zakoni.

[6] R. B. Marcus, "Modalities and intensional languages," v *Boston Studies in the Philosophy of Science*, zvezek 1 (New York: Humanities Press, 1963), str. 71 in dalje. Glej tudi Quinove "Comments" in razpravo, ki sledi. [Navedeni članek Marcusove je skupaj z omenjeno razpravo kot dodatkom objavljen tudi v zbirki njenih člankov *Modalities: Philosophical Essays* (Oxford: Oxford University Press, 1993), str. 3-35. Kot drugi dodatek k temu članku v njeni zbirki nastopa še njen komentar Smullyanovega članka "Modality and Description", na katerega se v prejšnji opombi nanaša tudi Kripke.]

navado, da ko rečem, da "se zdi očitno neresnično", to pomeni, da "očitno je neresnično". Dejansko mislim, da je to stališče resnično, čeprav ne čisto v tej obliki, ki jo zagovarja ga. Marcus.) V vsakem primeru se zdi očitno neresnično. En primer je podal profesor Quine v svojem odgovoru profesorici Marcus na navedenem simpoziju: "Mislim, da kljub temu nasprotje med lastnimi imeni in opisi, kakor ga potegne profesorica Marcus, prinaša težave. Paradigma pripisovanja lastnih imen je etiketiranje. Planet Venero lahko nekega lepega večera etiketiramo z lastnim imenom 'Hesperus'. Isti planet lahko znova nekega dne pred sončnim vzhodom etiketiramo z lastnim imenom 'Phosphorus'." (Quine misli, da se je nekaj takšnega nekoč dejansko zgodilo.) "Ko končno odkrijemo, da smo dvakrat etiketirali isti planet, je naše odkritje empirično, in to ne zaradi tega, ker bi bila lastna imena opisi." Glede na povedano, se je za planet Venero, kakor so ga videli zjutraj, mislilo, da je zvezda in je bila imenovana "Danica" oziroma (da odpravimo vsako vprašanje o rabi opisa) je bila imenovana "Phosphorus". Za en in isti planet se je, če so ga videli zvečer, mislilo, da je druga zvezda, Večernica, imenovana pa je bila "Hesperus". Kasneje so astronomi odkrili, da sta bili Phosphorus in Hesperus eno in isto. Gotovo ne zaradi kakšnega apriornega sklepanja, ki bi jim morda omogočalo, da sklepajo, da Phosphorus je Hesperus. Dejansko bi se lahko na podlagi informacije, ki bi jo imeli, morda izkazalo tudi drugače. Zato, tako se sklepa, mora biti izjava "Hesperus je Phosphorus" običajna kontingentna, empirična resnica, ki bi lahko morda bila tudi drugačna, zato mora biti stališče, da so resnične izjave o identiteti med imenoma nujne, neresnično. Drug primer, ki ga Quine poda v knjigi *Word and Object*, je vzet od profesorja Schrödingerja, slavnega začetnika kvantne mehanike. Določeno goro lahko vidimo tako s Tibeta kot z Nepala. Ko so jo videli z ene strani, je bila imenovana "Gaurisanker", ko so jo videli z druge strani, je bila imenovana "Everest"; in kasneje je potem prišlo do empiričnega odkritja, da Gaurisanker *je* Everest. (Quine v nadaljevanju pravi, da sodi, da je ta primer dejansko geografsko napačen. Domnevam, da se glede geografske informacije ne bi smeli zanašati na fizike.)

Seveda, en možen odziv na ta argument je, da zanikamo, da so imena kot "Cicero", "Tulij", "Gaurisanker" in "Everest" v resnici lastna imena. Poglejte, lahko bi rekli (nekdo je to tudi rekel: njegovo ime je bilo "Bertrand Russell"), da ravno zato, ker so izjave kot "Hesperus je Phosphorus" in "Gaurisanker je Everest" kontingentne, lahko vidimo, da omenjena imena v resnici niso čisto referenčna. Vi niste, če uporabim frazo ga. Marcus, zgolj "etiketirali" nekega predmeta, dejansko ste ga opisali. Kaj kontingentno dejstvo, da Hesperus je Phosphorus, pomeni? Dobro, pomeni, da *ta* zvezda na določenem delu neba zvečer je *ta* zvezda na določenem delu neba zjutraj. Podobno kontingentno dejstvo, da Guarisanker je Everest, pomeni, da je ta gora, ki jo gledamo s takšnega in takšnega kota v Nepalu, ta gora, ki jo gledamo z drugega takšnega in takšnega kota v Tibetu. Zato so lahko takšna imena kot "Hesperus" in "Phosphorus" samo okrajšave za opise. Termin "Phosphorus" *mora* pomeniti "ta

zvezda, ki jo vidimo ...” oziroma (bodimo pazljivi, ker se je dejansko izkazalo, da ni zvezda) “to *nebesno telo*, ki ga vidimo s takšne in takšne pozicije ob tem in tem času zjutraj”, ime “Hesperus” pa mora pomeniti “to nebesno telo, ki ga vidimo na takšni in takšni poziciji ob tem in tem času zvečer”. Torej, sklepa Russell, če hočemo rezervirati termin “ime” za stvari, ki v resnici zgolj imenujejo nek predmet, ne da bi ga opisale, so edina prava lastna imena, ki jih lahko imamo, imena naših lastnih neposrednih čutnih podatkov, predmetov naše lastne “neposredne seznanjenosti”. Edina takšna imena, ki nastopajo v jeziku, so demonstrativi, kot sta “to” in “tisto”. In zlahka vidimo, da je ta zahteva po nujnosti identitete, ki je razumljena tako, da izvzame identitete med imeni iz vsakega predstavljivega dvoma, lahko zagotovljena samo za demonstrativna imena neposrednih čutnih podatkov; kajti samo v takšnih primerih je lahko neka izjava o identiteti med dvema različnima imenoma v splošnem imuna pred kartezijanskim dvomom. Obstajajo tudi druge stvari, ki jih je Russell včasih dopuščal kot predmete seznanjenosti, na primer sebstvo; tu se nam ni treba spuščati v podrobnosti. Drugi filozofi (na primer ga. Marcus v svojem odgovoru, vsaj v javni razpravi, kakor se je spominjam – ne vem, če je to šlo v tisk, tako da to nemara ne bi smelo biti “etiketirano” njej[7]) so rekli: “Če so imena v resnici samo etikete, pristne etikete, potem bi nam dober slovar lahko povedal, da so imena istega predmeta.” Imate nek predmet *a* in nek predmet *b* z imenoma “John” in “Joe”. Potem bi vam v skladu z ga. Marcusovo nek slovar lahko povedal, ali sta ali nista “John” in “Joe” imeni istega predmeta. Seveda ne vem, kaj bi lahko naredili idealni slovarji, zdi pa se, da običajna lastna imena ne izpolnijo te zahteve. V primeru običajnih lastnih imen gotovo lahko pridete do čisto empiričnih odkritij, recimo, da Hesperus je Phosphorus, čeprav smo mislili drugače. Lahko smo v dvomih, ali Gaurisanker je Everest in ali je Cicero dejansko Tulij. Sedaj bi lahko celo odkrili, da smo se pri predpostavki, da Hesperus je Phosphorus, zmotili. Mogoče so se astronomi zmotili. Tako se zdi, da je to stališče napačno in če z imenom ne mislimo kakšnega umetnega pojma imen kot Russell, ampak lastno ime v običajnem smislu, potem lahko obstajajo kontingentne identitetne izjave, v katerih so uporabljena lastna imena, nasprotno stališče pa se zdi očitno napačno.

V nedavni filozofiji so kot primere kontingentnih identitetnih izjav izpostavili veliko drugih identitetnih izjav, ki se nemara razlikujejo od vseh vrst, ki sem jih omenil. Ena od njih je, na primer, izjava “Toplota je gibanje molekul”. Prvič, predpostavlja se, da je to odkrila znanost. Predpostavilo se je, da so empirični znanstveniki v svojih raziskavah odkrili (in, predpostavljam, da so), da je zunanji pojav, ki ga imenujemo “toplota”, dejansko molekularno nihanje. Drug primer takšnega odkritja je, da je voda H_2O, še nadaljnji primeri pa so, da je zlato

[7] Smelo bi biti. Glej njeno opazko v *Boston Studies in the Philosophy of Science*, zvezek 1, str. 115, v razpravi, ki sledi omenjenemu članku. [Glej navedeni odlomek tudi v ponatisu te razprave “Modalities and Intensional Language: Discussion” v Marcusove knjigi *Modalities: Philosophical Essays* (Oxford: Oxford University Press, 1993), str. 33-34.]

element s takšnim in takšnim atomskim številom, da je svetloba tok fotonov in tako dalje. Vse to so – v nekem smislu izraza "identitetna izjava" – identitetne izjave. Drugič, mislili so, da so to očitno kontingentne identitetne izjave, samo zato ker so to bila znanstvena odkritja. Navsezadnje bi se lahko za toploto izkazalo, da ni gibanje molekul. Predlagane so bile tudi alternativne teorije toplote, na primer kalorična teorija toplote. Če bi bile te teorije toplote pravilne, potem toplota ne bi bila gibanje molekul, ampak neka substanca, ki bi preplavljala vroče telo, imenovana "toplotnost". In stvar smeri znanosti in ne kakšne logične nujnosti je bila, da se je za eno teorijo izkazalo, da je pravilna, in da se je za drugo teorijo izkazalo, da je nepravilna.

Tako, tukaj imamo znova navidez drug preprost primer kontingentne identitetne izjave. Zanj se je predpostavljalo, da je zelo pomemben primer zaradi njegove povezave s problemom duh-telo. Bilo je mnogo filozofov, ki so hoteli biti materialisti in materialisti v neki posebni obliki, ki je danes poznana kot "identitetna teorija". V skladu s to teorijo je neko določeno mentalno stanje, kot je to, da je neka oseba v stanju bolečine, identično z nekim določenim stanjem njenih možganov (oziroma nemara, v skladu z nekaterimi teoretiki, s stanjem njenega celotnega telesa), v vsakem primeru pa z nekim določenim materialnim oziroma živčnim stanjem njenih možganov oziroma telesa. In tako bi bilo v skladu s to teorijo to, da sem, če sem, v tem trenutku v stanju bolečine, identično s tem, da je v nekem določenem stanju moje telo oziroma moji možgani. Drugi so ugovarjali, da to ne more biti res, ker si navsezadnje lahko zamislimo, da bi moja bolečina eksistirala, celo če to stanje mojega telesa ne bi. Nemara bi si lahko zamislili, da sploh ne bi bil utelešen in bi še vedno bil v stanju bolečine, ali pa bi si, obratno, lahko zamislili, da bi moje telo eksistiralo in bi bilo v prav istem stanju, celo če ne bi bilo nobene bolečine. Dejansko bi si lahko zamislili, da bi bilo v tem stanju, celo če tako rekoč sploh ne bi bilo nobenega duha "za njim". Običajni odgovor je priznal, da bi se vse to lahko zgodilo, vendar je to nerelevantno za identiteto med mentalnim in fizikalnim stanjem. Rečeno je bilo, da je ta identiteta zgolj še ena kontingentna znanstvena identifikacija, ki je podobna identifikaciji toplote z gibanjem molekul ali pa vode s H_20. Ravno tako, kakor si lahko zamislimo toploto brez vsakega molekularnega gibanja, si lahko zamislimo neko mentalno stanje brez vsakega ustreznega možganskega stanja. Toda: kakor prvo dejstvo ne škoduje identifikaciji toplote in gibanja molekul, tako drugo dejstvo sploh ne škoduje identifikaciji nekega mentalnega stanja z ustreznim možganskim stanjem. In tako so mnogi filozofi zadnjega časa temu, da lahko obstajajo kontingentne identitetne izjave te oblike, pripisovali velik pomen za naše teoretsko razumevanje problema duh-telo.

Da končno pojasnim, kaj *sam* mislim – v nasprotju s tem, kar se zdi, in s tem, kar mislijo drugi –, mislim, da so v obeh primerih, v primeru imen in v primeru teoretskih identifikacij, identitetne izjave nujne in ne kontingentne. To se pravi, da so te izjave, če so *resnične*, nujne; seveda neresnične identitetne izjave niso

nujne. Kako je lahko možno, da kdo zagovarja takšno stališče? Nemara mi ne bo uspelo podati celotnega odgovora na to vprašanje, čeprav sem prepričan, da je to stališče resnično. Da pa bi začel z odgovorom, naj naredim nekaj razlikovanj, ki jih nameravam uporabiti. Prvo je razlikovanje med *togim* in *netogim označevalcem*. Kaj pomenita ta termina? Kot primer nekega netogega označevalca lahko podam izraz, kot je "izumitelj očal z dvojnimi lečami". Predpostavimo, da je bil Benjamin Franklin tisti, ki je izumil očala z dvojnimi lečami, tako da "izumitelj očal z dvojnimi lečami" označuje oziroma referira nekega določenega človeka, namreč Benjamina Franklina. Vendar si lahko zlahka predstavljamo, da bi ta svet lahko bil drugačen, da bi v drugih okoliščinah nekdo drug še pred Benjaminom Franklinom prišel do tega izuma, v tem primeru bi bil *on* izumitelj očal z dvojnimi lečami. Tako je v tem smislu izraz "izumitelj očal z dvojnimi lečami" netog: V določenih okoliščinah bi bil izumitelj očal z dvojnimi lečami en človek; v drugih okoliščinah bi bil izumitelj očal z dvojnimi lečami drug človek. V nasprotju s tem obravnavajmo izraz "kvadratni koren števila 25". Neodvisno od empiričnih dejstev lahko podamo aritmetični dokaz, da je kvadratni koren števila 25 število 5, ker pa smo to dokazali matematično, je to, kar smo dokazali, nujno. Če sploh razmišljamo o številih kot bitnostih – naj predpostavimo, vsaj za namen tega predavanja, da to naredimo –, potem izraz "kvadratni koren števila 25" nujno označuje določeno število, namreč 5. Takšen izraz imenujem "*togi* označevalec". Nekateri filozofi mislijo, da kdorkoli, ki sploh uporablja pojma togi oziroma netogi označevalec, s tem že pokaže, da je žrtev določene zmede oziroma ni pozoren na določena dejstva. Kaj menim z izrazom "togi označevalec"? Menim nek termin, ki označuje isti predmet v vseh možnih svetovih. Da preprečim določeno pomoto, ki gotovo ni moja, ne uporabljam izraza "bi lahko označeval kak drug predmet", da bi se nanašal na dejstvo, da bi lahko ta jezik uporabljali drugače. Na primer, izraz "izumitelj očal z dvojnimi lečami" bi lahko prebivalci tega planeta vedno uporabljali za nanašanje na tistega človeka, ki je podkupil Hadleyburg.[*] Tako bi bilo, če, prvič, ljudje na tem planetu ne bi govorili slovenščine[**], ampak kakšen drug jezik, ki fonetično sovpada s slovenščino[***], in če bi, drugič, v tem jeziku izraz "izumitelj očal z dvojnimi lečami" pomenil "človek, ki je podkupil Hadleyburg". Potem bi se seveda v njihovem jeziku nanašal na kogarkoli, ki bi dejansko podkupil Hadleyburg v tej protidejstveni situaciji. To ni tisto, kar menim. Kar menim s tem, ko rečem, da bi se lahko nek opis nanašal na nekaj drugega, je, da bi lahko v *našem* jeziku, kakor ga *uporabljamo* pri opisovanju neke protidejstvene situacije, obstajal nek drug predmet, ki bi izpolnil opisne pogoje, ki jih *mi* podajamo za referenco. Tako, na primer, če govorimo o drugem možnem svetu oziroma o neki protidejstveni situaciji, uporabljamo frazo "izumitelj očal z dvojnimi lečami" za nanašanje na kogarkoli, ki bi v tej protidejstveni situaciji izumil očala z dvojnimi

[**] *The Man That Corrupted Hadleyburg* je naslov ene od knjig Marka Twaina.
[***] V izvirniku "English".
[***] V izvirniku "with English".

lečami, ne pa na osebo, ki bi jo ljudje *v* tej protidejstveni situaciji imenovali "izumitelj očal z dvojnimi lečami". *Ti ljudje* bi morda lahko govorili nek drug jezik, ki fonetično sovpada s slovenščino[*], v katerem je "izumitelj očal z dvojnimi lečami" uporabljen na drugačen način. Tukaj se *ne* ukvarjam s tem vprašanjem. Kar se tega tiče, bi lahko bili gluhonemi ali pa sploh ne bi bili ljudje. (Tudi če ljudje ne bi obstajali, bi še vedno lahko obstajal izumitelj očal z dvojnimi lečami – bog ali hudič bi zadostoval.)

Drugič, ko govorim o pojmu togega označevalca, s tem ne mislim, da iz tega pojma sledi, da mora predmet, na katerega se nanaša, eksistirati v vseh možnih svetovih, to je, da mora nujno obstajati. Nekatere stvari, nemara matematične bitnosti, kot so pozitivna cela števila, če sploh eksistirajo, nujno eksistirajo. Nekateri ljudje so menili, da bog eksistira in da nujno eksistira; drugi, da kontingentno eksistira; spet drugi, da kontingentno ne eksistira; in drugi, da nujno ne eksistira:[8] v igri so bile vse štiri izbire. V vsakem primeru pa, ko uporabim pojem togega označevalca, s tem ne impliciram, da predmet, na katerega se nanaša, nujno eksistira. Vse, kar menim, je, da v vsakem možnem svetu, kjer omenjeni predmet *res* eksistira, v vsaki situaciji, kjer *bi* ta predmet eksistiral, uporabimo omenjeni označevalec za to, da označimo ta predmet. V situaciji, kjer ta predmet ne obstaja, bomo rekli, da ta označevalec nima referenta in da omenjeni predmet, ki je tako označen, ne eksistira.

Kot sem rekel, za mnoge filozofe je prav pojem togega označevalca problematičen *per se*. In ugovor teh ljudi bi lahko navedli tako, kakor sledi: Poglej, govoriš o situacijah, ki so protidejstvene, to se pravi, da govoriš o drugih možnih svetovih. Ti svetovi pa so navsezadnje popolnoma nepovezani z aktualnim svetom, ki ni samo še eden možni svet; je aktualni svet. Tako moraš – preden sploh govoriš o, recimo, nekem takšen predmetu, kot je Richard Nixon, v drugem možnem svetu – povedati, kateri predmet v tem drugem možnem svetu bi *bil* Richard Nixon. Pogovorimo se o situaciji, v kateri bi bil, kakor bi *vi* rekli, Richard Nixon član SDS[**]. Gotovo je član SDS, o katerem govorite, nekdo, ki ima mnogo svojih lastnostih zelo drugačnih od Nixona. Preden sploh lahko rečemo, ali bi ta človek bil Richard Nixon ali ne, moramo postaviti kriterije identitete preko možnih svetov. Tukaj so ti drugi možni svetovi. V njih obstajajo vseh vrst predmeti z drugačnimi lastnosti, kot jih ima katerikoli aktualni predmet. Nekateri od teh predmetov so Nixonu podobni na te načine, nekateri od njih na druge načine. No, kateri od teh predmetov je Nixon? Podati moramo kakšen kriterij istovetnosti. In to kaže, kako prav pojem togega označevalca vodi v nek

[**] V izvirniku "with English".

[8] Če ne obstaja nikakršno božanstvo, in še posebej, če je neeksistenca nekega božanstva *nujna*, je dvomljivo, ali lahko uporabimo "on", da bi se nanašali na neko božanstvo. Uporabo v tekstu moramo vzeti nedobesedno.

[**]* Gre za levičarsko organizacijo *Students for a Democratic Society* (SDS), ustanovljeno v začetku šestdesetih let, ki je, med drugim, aprila 1965 sklicala pohod na Washington proti vojni v Vietnamu. *Študenti za demokratično družbo* je odigrala tudi ključno vlogo v študentskih protestih na *Columbia University* leta 1968.

krog. Predpostavimo, da neko določeno število označimo kot število planetov. Potem – če je to, tako rekoč, naš favoriziran način označevanja tega števila – bomo v vsakem drugem možnem svetu morali vsako število, ki je število planetov, identificirati s številom 9, ki je število planetov v aktualnem svetu. Tako – to so dokazovali številni filozofi, implicitno, na primer, Quine in eksplicitno mnogi drugi na njegovi sledi – v resnici ne moremo vprašati, ali je nek označevalec tog ali netog, ker najprej potrebujemo nek kriterij identitete preko možnih svetov. Po najbolj skrajnem stališču – ker so možni svetovi tako nepovezani z našim lastnim – v resnici sploh ne moremo reči, da je kakšen predmet v njih *isti* kot nek predmet, ki eksistira sedaj, ampak samo, da obstajajo predmeti, ki so bolj ali manj podobni stvarem v aktualnem svetu. V resnici naj torej ne bi govorili o tem, kar naj bi bilo v drugem možnem svetu resnično o Nixonu, ampak samo o tem, kakšni Nixonovi "dvojniki" bi obstajali (ta termin uporablja David Lewis[9]). Nekateri ljudje v drugih možnih svetovih imajo pse, ki jih imenujejo "Checkers"[*]. Drugi podpirajo ABM[**], nimajo pa psa, ki ga imenujejo Checkers. Obstajajo različni ljudje, ki so bolj ali manj podobni Nixonu, za nobenega od njih pa v resnici ne moremo reči, da je Nixon; so samo Nixonovi *dvojniki*, vi pa – glede na to, kateri je v skladu z vašim najljubšim kriterijem najbolj podoben Nixonu – izberete, kateri je najboljši dvojnik. Takšna stališča so na široko razširjena tako med zagovorniki kvantificirane modalne logike kot med njenimi nasprotniki.

Za vsa ta govorjenja se mi zdi, da jemljejo metaforo o možnih svetovih na nek način mnogo preresno. Zdi se, kot da bi bil "možni svet" podoben tuji deželi ali pa oddaljenemu planetu tam nekje. Zdi se, kot da skozi teleskop nejasno vidimo različne igralce na tem oddaljenem planetu. Dejansko se zdi stališče Davida Lewisa najbolj sprejemljivo, če kdo jemlje to sliko dobesedno. Nihče daleč stran, na drugem planetu, ne more biti absolutno identičen z nekom tukaj. Toda celo če imamo kakšne čarobne metode transportacije za prestavljanje ene in iste osebe s planeta na planet, v resnici potrebujemo nek epistemološki kriterij identitete, da bi bili sposobni reči, ali je nekdo na tem oddaljenem planetu ista oseba kot nekdo tukaj.

Vse to se mi zdi nek popolnoma zavajajoč način gledanja na stvari. Gre za stališče, da morajo biti protidejstvene situacije opisane čisto kvalitativno. Tako, na primer, ne moremo reči: "Če bi le Nixon dal zadostno podkupnino senatorju X, bi spravil Carswella[***] skozi," kajti to se nanaša na določena človeka, na Nixona in Carswella, in govori o tem, katere stvari bi bile resnične zanju v neki

[9] David K. Lewis, "Counterpart theory and quantified modal logic", *Journal of Philosophy* 65 (1968), str. 113 in dalje.

[**] Checkers je bil koker španjel Richarda Nixona.

[**]* Sporazum o protibalističnih izstrelkih *Anti-Ballistic Missile Treaty* (ABM) sta Sovjetska zveza in Združene države Amerike podpisali leta 1972, ko je bil predsednik ZDA Richard Nixon.

[**]** Richard Nixon je leta 1970 predlagal Georga Harrolda Carswella za člana vrhovnega sodišča Združenih držav, senat pa je njegovo nominacijo zavrnil s 45 glasovi za in 51 proti.

protidejstveni situaciji. Namesto tega moramo reči: "Če bi nek človek, ki ima takšne in takšne lase in ima takšne in takšne politične nazore, dal podkupnino nekemu človeku, ki je bil senator in je imel takšne in takšne druge kvalitete, potem bi bil potrjen nek človek, ki je bil sodnik na jugu in je imel mnogo drugih kvalitet, enakih kot Carswell." Z drugimi besedami, protidejstvene situacije moramo opisati čisto kvalitativno in potem vprašati: "Če ta situacija vsebuje ljudi in stvari s takšnimi in takšnimi kvalitetami, kateri od teh ljudi je Nixon (oziroma njegov dvojnik), kateri Carswell (oziroma njegov dvojnik) in tako naprej?" To se mi zdi napačno. Kdo nas bo odvrnil od tega, da rečemo "Nixon bi morda spravil Carswella skozi, če bi naredil določene stvari"? Govorimo *o Nixonu* in sprašujemo, kaj bi bilo v določenih protidejstvenih situacijah resnično *zanj*. Rečemo lahko, da če bi Nixon počel to in to, bi bil namesto njega za predsednika ZDA izvoljen Humphrey. Tisti, ki jim nasprotujem, bi argumentirali: "Ja, toda kako ugotoviš, ali je ta človek, o katerem govoriš, dejansko Nixon?" To bi tudi zares težko ugotovili, če bi celotno situacijo opazovali skozi teleskop, tega pa tu ne počnemo. Možni svetovi niso nekaj, na kar bi aplicirali takšno epistemološko vprašanje. In če je fraza "možni svetovi" tista, ki komurkoli da misliti, da bi takšno vprašanje apliciral, bi moral samo opustiti to frazo in uporabiti kakšen drug izraz, recimo "protidejstvena situacija", ki bi morda bil manj zavajajoč. Če rečemo: "Če bi Nixon podkupil takšnega in takšnega senatorja, bi Nixon spravil Carswella skozi," je to, kar *je dano* prav v tem opisu te situacije, to, da je to neka situacija, v kateri govorimo o Nixonu in o Carswellu in o nekem takšnem in takšnem senatorju. In zdi se, da lahko prav tako, kakor ugovarjamo *postuliranju* [*stipulating*] tega, da govorimo o določenih ljudeh, ugovarjamo postuliranju tega, da govorimo o določenih *lastnostih*. Zagovorniki drugega stališča jemljejo govor o določenih kvalitetah kot neproblematičen. Ne rečejo: "Kako vemo, da je ta kvaliteta (v drugem možnem svetu) kvaliteta rdečosti?" Govorjenje o določenih *ljudeh* pa je za njih problematično. Sam pa ne vidim, da bi imeli kaj več razlogov za ugovarjanje v enem, ne pa tudi v drugem primeru. Menim, da tovrstni ugovori v resnici izvirajo iz ideje o možnih svetovih, ki naj bi eksistirali tam nekje, opazujemo pa jih lahko skozi poseben teleskop. Še bolj problematično je stališče Davida Lewisa. V skladu z Lewisom velja, da ko rečemo: "V določenih okoliščinah bi Nixon spravil Carswella skozi," v resnici menimo: "Nek človek, ki ni Nixon, mu je pa zelo podoben, bi spravil nekega sodnika, ki ni Carswell, mu je pa zelo podoben, skozi." Morda je tako, da bi lahko nek človek, ki je zelo podoben Nixonu, spravil skozi nekega človeka, ki je zelo podoben Carswellu. Toda *to* ne bi poskrbelo niti za Nixona niti za Carswella, prav tako ne bi prignalo Nixona samega do tega, da bi rekel: "*Jaz* naj bi naredil to in to, da bi spravil Carswella skozi." Vprašanje je, ali bi v določenih okoliščina *sam* Nixon lahko spravil *Carswella* skozi. In mislim, da je navedeni ugovor osnovan na zavajajoči sliki.

Namesto tega lahko čisto brez težav govorimo o togih in netogih označevalcih. Še več, zanje imamo enostaven intuitiven test. Rečemo, na primer, lahko, da bi

število planetov morda bilo drugačno od dejanskega števila planetov. Morda bi jih bilo, na primer, samo sedem. Rečemo lahko, da bi morda bil izumitelj očal z dvojnimi lečami nekdo drug, ne pa človek, ki je *dejansko* izumil očala z dvojnimi lečami.[10] Ne moremo pa reči, da bi kvadratni koren od 81 morda bilo neko drugo število, kot dejansko je, kajti to število mora biti ravno 9. Če ta intuitivni test apliciramo na lastna imena, kakršno je, na primer, "Richard Nixon", se zdi, da se bo zanje izkazalo, da so togi označevalci. Prvič, celo če govorimo o protidejstveni situaciji, v kateri za Nixona predpostavljamo, da je naredil drugačne stvari, domnevamo, da še vedno govorimo prav o Nixonu. Rečemo: "Če bi Nixon podkupil določenega senatorja, bi spravil Carswella skozi," in domnevamo, da se z "Nixon" in "Carswell" nanašamo na prav ista človeka kot v aktualnem svetu. In zdi se, da ne moremo reči "Nixon bi morda bil nek drug človek, kot dejansko je", razen, seveda, če bi to menili metaforično. Morda bi bil neke druge *vrste* oseba (če verjamete v svobodno voljo in v to, da ljudje niso podkupljivi po naravi). Morda bi si mislili, da je ta izjava v tem smislu resnična, toda Nixon v drugem, dobesednem smislu ne bi mogel biti neka druga oseba, kot dejansko je, četudi bi bil Humphrey sedemintrideseti predsednik Združenih držav. Tako je fraza "ta sedemintrideseti predsednik" netoga, "Nixon" pa, tako se zdi, tog.

Naj, preden se vrnem k vprašanju identitetnih izjav, izpostavim še eno razlikovanje. To razlikovanje je zelo temeljno in ga je tudi težko razbrati. V nedavni razpravi je veliko filozofov, ki so razpravljali o smiselnosti različnih kategorij resnic, obravnavalo te kategorije resnic kot identične. Nekateri od tistih, ki jih imajo za identične, jih hrupno zagovarjajo, drugi pa, kot Quine, pravijo, da so vse identično brez pomena. Običajno pa se jih ne razlikuje. To so kategorije kot "analitično", "nujno", "*a priori*" in včasih celo "gotovo". Ne bom govoril o vseh, ampak samo o pojmih apriornosti in nujnosti. Zelo pogosto ju imajo za sinonima. (Mnogo filozofov ne bi smeli opisati kot tistih, ki jih imajo za sinonima; ti jih enostavno *uporabljajo* izmenjaje.) Sam jih hočem razlikovati. Kaj menimo s tem, ko za neko izjavo rečemo, da je *nujna*? Enostavno menimo, da je omenjena izjava, prvič, resnična, in, drugič, da ne bi moglo biti drugače. Ko

[10] Nekateri filozofi mislijo, da so določni opisi v slovenščini [*in English*] dvoumni, da včasih "ta izumitelj očal z dvojnimi lečami" togo označuje človeka, ki je dejansko izumil očala z dvojnimi lečami. Ne povsem prepričan se nagibam k zavrnitvi tega stališča, ki je oblikovano kot teza o slovenščini [*about English*] (v nasprotju z nekim možnim hipotetičnim jezikom), vendar o tem tukaj ne bom argumentiral.
Kar bi rad omenil, je, da, v nasprotju z nekaterimi mnenji, ta domnevna dvoumnost ne more nadomestiti russellovskega pojma doseg opisa. Obravnavajmo naslednji stavek: "Število planetov bi lahko nujno bilo sodo." Ta stavek gotovo lahko beremo tako, kakor da izraža neko resnico, da če bi obstajalo osem planetov, bi bilo število planetov nujno sodo. Vendar bo ta izjava brez razlikovanj dosega neresnična tako ob "referencialnem" (togem) kot ob netogem branju tega opisa. (Ker je število planetov 9, togo branje privede do neresničnosti tega, da bi 9 lahko bilo nujno sodo.)
"Togo" branje je ekvivalentno russellovskemu primarnemu nastopu, netogo branje pa najbolj notranjemu dosegu – nekateri so, sledeč Donnellanu, nemara površno, to branje imenovali "atributivna" raba. S tem pa so zanemarili možnost vmesnega dosega, v navedenem primeru nameravano branje $\Diamond \Box$ (število planetov je sodo) ne naredi opisa $\Box$ (število planetov je sodo) niti za najširšega niti za najožjega možnega.

rečemo, da je nekaj *kontingentno* resnično, menimo, da čeprav je dejansko tako, bi se lahko zgodilo, da bi bile stvari drugačne. Če hočemo to razliko pripisati kakšni filozofski panogi, bi jo morali pripisati metafiziki. V nasprotju s tem obstaja pojem *apriorne resnice*. Za apriorno resnico velja tista, za katere resničnost lahko *vemo* neodvisno od vsega izkustva. Upoštevajte, da to samo po sebi še ne pove ničesar o vseh možnih svetovih, razen če to ne postane del definicije. Vse, kar pove, je, da lahko neodvisno od vsega izkustva vemo, da je resnična o aktualnem svetu. Morda po kakšnem filozofskem argumentu iz našega védenja – neodvisno od izkustva – sledi, da če je nekaj resnično o aktualnem svetu, moramo vedeti, da je resnično tudi o vseh možnih svetovih. Če pa hočemo to vpeljati, potrebujemo kakšen filozofski argument. *Ta* pojem, če bi ga pripisali kakšni filozofski panogi, pa ne pripada metafiziki, ampak epistemologiji. Opraviti ima z načinom, kako lahko za določene stvari vemo, da so dejansko resnične. Vendar bi se seveda lahko zgodilo, da je vse, kar je nujno, nekaj, kar lahko vemo *a priori*. (Mimogrede: upoštevajte, da ima tako definiran pojem apriorne resnice v sebi še *drugo* modalnost: Zanjo *lahko* vemo neodvisno od vsega izkustva. Je nekoliko kompliciran, ker tukaj obstaja nekakšna dvojna modalnost.) Ne bom imel časa, da bi ta pojma tukaj nadrobno raziskal, toda ena stvar, ki jo lahko takoj opazimo, je, da ta dva pojma nikakor nista trivialno enaka. Če imata enak obseg, lahko to vpeljemo samo s kakšnim filozofskim argumentom. Kakor je navedeno, pripadata različnima področjema filozofije. Eden ima nekaj opraviti z *védenjem*, s tem, kaj lahko na določen način vemo o *aktualnem* svetu. Drugi ima opravka z *metafiziko*, s tem, kako *bi lahko* bil ta svet dan; če je tako, kot je, bi lahko bilo še kako drugače? Jaz menim, da pravzaprav noben razred izjav ni vsebovan v drugem. Tukaj pa se moramo pogovoriti o sledečem: Je vse, kar je nujno, nekaj, kar lahko vemo *a priori* oziroma vemo *a priori*? Obravnavajmo sledeči primer: Goldbachova domneva. Ta pravi, da je vsako sodo število vsota dveh praštevil. To je matematična izjava, in če je resnična, mora biti nujna. Gotovo ne bi mogli reči, da čeprav je dejansko vsako sodo število vsota dveh praštevil, bi lahko obstajalo kakšno posebno število, ki bi bilo sodo in ne bi bilo vsota dveh praštevil. Kaj bi to pomenilo? Po drugi strani pa ne poznamo odgovora na vprašanje, ali dejansko *je* vsako sodo število vsota dveh praštevil, in trenutno nimamo nikakršne metode, da bi odločili o tem. Tako da gotovo ne vemo, niti *a priori* in celo niti *a posteriori*, da je vsako sodo število vsota dveh praštevil. (Dobro, nekaj evidence nam nemara nudi to, da še nismo našli protiprimera.) Vendar gotovo na noben način ne vemo *a priori*, da je vsako sodo število dejansko vsota dveh praštevil. Seveda pa navedena definicija pravi ravno "*lahko* vemo neodvisno od izkustva" in kdo bi lahko rekel, da je res, da *bi* to *lahko* vedeli neodvisno od izkustva. Težko je natančno ugotoviti, kaj ta trditev pomeni. Lahko bi bilo tako. To bi po eni strani morda pomenilo, da če bi to bilo resnično, bi to lahko *dokazali*. Ta trditev je gotovo napačna, če jo na splošno apliciramo na matematične izjave, delovati pa moramo znotraj nekega fiksnega sistema. To je tisto, kar je dokazal Gödel. In celo če s tem menimo nek

"intuitiven dokaz v splošnem", bi se morda zgodilo (to stališče je vsaj tako jasno in tako verjetno kot nasprotno), da čeprav je izjava resnična, človeški um tega na noben način ne bi mogel nikoli dokazati. Seveda je en način, kako bi bil to sposoben dokazati nek *neskončni* um – tako da bi drugega za drugim pregledal vsako naravno število in jih preveril. V tem smislu se za to seveda nemara ve *a priori*, vendar pa lahko to ve samo nek neskončni um, kar pa potem vodi v druga zapletena vprašanja. Nočem razpravljati o vprašanjih glede možnosti tega, da bi kdo izvedel neko neskončno število dejanj, s katerimi bi drugo za drugo pregledal vsako število. O tem je bilo napisano ogromno filozofske literature: Nekateri so navajali, da je to logično nemogoče; nekateri, da je logično možno; nekateri pa ne vedo. Najpomembnejše je, da ni trivialno, da lahko kdo neko takšno izjavo, ravno zato ker je nujna, ve *a priori*. Preden odločimo, da kdo zanjo lahko ve tako, potrebujemo kakšno upoštevanja vredno pojasnitev. In tako vidimo, da celo če je vse, kar je nujno, v nekem smislu *a priori*, tega ne bi smeli vzeti kot trivialno zadevo definicije. To je vsebinska filozofska teza, ki potrebuje nekaj dela.

Drug primer, ki bi ga lahko podali, je povezan s problemom esencializma. Tukaj je pult. Vprašanje, ki je bilo v filozofiji pogosto postavljeno, je: Katere so njegove bistvene lastnosti? Katere lastnosti – razen trivialnih, kot je samoidentičnost – so takšne, da jih ta predmet, če eksistira, mora imeti,[11] oziroma so takšne, da če jih nek predmet ne bi imel, ne bi bil ta predmet?[12] Na primer: to, da je narejen iz lesa in ne iz ledu, bi morda bila neka bistvena lastnost tega pulta. Vzemimo šibkejšo izjavo, da ni narejen iz ledu. To nam omogoča tako močno argumentacijo, kot jo potrebujemo, nemara celo tako dramatično. Predpostavimo, da je ta pult dejansko narejen iz lesa, in se vprašajmo, ali bi lahko bil prav ta pult prav od začetka svoje eksistence narejen iz ledu, recimo zamrznjen iz vode iz Temze? Imamo močan občutek, da *ne* bi mogel biti, čeprav bi gotovo kdo lahko dejansko naredil neko mizo iz vode iz Temze, tako da bi jo s pomočjo kakšnega procesa zamrznil in jo postavil prav na to mesto, kjer stoji ta stvar. Če bi kdo naredil tako, bi seveda naredil nek *drug* predmet. To ne bi bil *prav ta pult*, in tako ne bi bilo primera, v katerem bi bil iz ledu oziroma iz vode iz Temze narejen prav

[11] Ta definicija je običajna formulacija pojma bistvena lastnost, vendar moramo narediti izjemo za eksistenco samo: po dani definiciji bi bila eksistenca trivialno bistvena. Eksistenca je za nek predmet nujna, samo če ta predmet nujno eksistira. Nemara obstajajo še druge izbrane lastnosti, ki vsebujejo eksistenco, za katere je ta definicija prav tako sporna. (Za to pripombo se zahvaljujem Michaelu Slotu.)

[12] Omenjena stavka podajata ekvivalentni definiciji pojma nujna lastnost, ker je $\Box((\exists x)(x = a) \supset Fa)$ ekvivvalentno z $\Box(x)(\sim Fx \supset x \neq a)$. Druga formulacija je vsekakor bila močna vzpodbuda za teorije o "identifikaciji preko možnih svetov". Ker predlaga, da obravnavamo "nek predmet *b* v drugem možnem svetu" in preverimo, ali ga lahko identificiramo z *a*, tako da se vprašamo, ali mu manjka kakšna bistvena lastnost *a*-ja. Naj torej poudarim, da čeprav je neka bistvena lastnost (trivialno) neka lastnost, brez katere nek predmet ne more biti *a*, iz tega nikakor ne sledi, da bistvene, čisto kvalitativne lastnosti *a*-ja skupaj tvorijo zadostni pogoj za biti *a*, niti to, da so *sploh kakšni* čisto kvalitativni pogoji zadostni za to, da je nek predmet *a*. Nadalje, celo če bi nujni in zadostni kvalitativni pogoji za to, da je nek predmet Nixon, morda eksistirali, še vedno ne bi bili kaj dosti upravičeni zahtevati, da morajo biti čisto kvalitativno opisane vse protidejstvene situacije. Vprašamo lahko, ali bi Nixon morda lahko bil član demokratske stranke, ne da bi se spuščali v te podrobnosti.

ta pult tukaj. Vprašanje, ali bi se lahko kasneje, recimo čez eno minuto, spremenil v led, je nekaj drugega. Tako se zdi, da če je takšen primer pravilen – in to so menili zagovorniki esencializma –, da ta pult ne bi mogel biti narejen iz ledu, to je: v vsaki protidejstveni situaciji, v kateri bi rekli, da ta pult sploh eksistira, bi morali reči tudi, da ni bil narejen iz zamrznjene vode iz Temze. Nekateri so seveda zavrnili vsakršen pojem bistvene lastnosti kot pojem brez pomena. Običajno zaradi tega (in mislim, da bi to, recimo, rekel Quine), ker so mislili, da je odvisen od pojma identitete preko možnih svetov, ta pa tudi nima pomena. Ker sem že zavrnil to stališče, se ne bom znova ukvarjal z njim. Lahko govorimo o *prav tem predmetu* in o tem, ali bi lahko imel določene lastnosti, ki jih dejansko nima. Na primer, lahko bi bil v drugi sobi in ne v tej, v kateri dejansko je, morda celo ravno v tem trenutku, ne bi pa mogel od vsega začetka biti narejen iz vode, zamrznjene v led.

Če je esencialistično stališče pravilno, je lahko pravilno samo v primeru, če ostro razlikujemo med pojmoma *a posteriori* in *a priori* resnice na eni in pojmoma kontingentne in nujne resnice na drugi strani, ker čeprav je izjava, da ta miza, če sploh eksistira, ni bila narejena iz ledu, nujna, to gotovo ni nekaj, kar vemo *a priori*. Kar vemo, je, prvič, da pulti običajno niso narejeni iz ledu, običajno so narejeni iz lesa. To zgleda kot les. Čutim, da ni hladen – če bi bil narejen iz ledu, pa bi verjetno bil. Zato sklepam, da verjetno ni narejen iz ledu. Tukaj je moja celotna sodba *a posteriori*. Lahko bi ugotovil, da sem žrtev duhovitega trika in da je ta pult dejansko narejen iz ledu; kar pa pravim, je, da si ob predpostavki, da dejansko ni narejen iz ledu – dejansko je narejen iz lesa –, nihče ne more zamisliti, da bi v določenih okoliščinah lahko bil narejen iz ledu. Tako moramo reči, da čeprav ne moremo vedeti *a priori*, ali je ta miza narejena iz ledu ali ne, ob predpostavki, da ni narejena iz ledu, *nujno* ni narejena iz ledu. Z drugimi besedami, če je *P* izjava, da ta pult ni narejen iz ledu, z *a priori* filozofsko analizo vemo, da "Če *P*, potem nujno *P*". Če ta miza ni narejena iz ledu, nujno ni narejena iz ledu. Po drugi strani potem z empiričnim raziskovanjem vemo, da je *P*, antecendens navedenega pogojnika, resničen – da ta miza ni narejena iz ledu. Sklepamo lahko po *modus ponensu*:

$P \supset \Box P$

P

$\Box P$

Sklep – "$\Box P$" – pravi, da je nujno, da ta miza ni narejena iz ledu, za ta sklep pa vemo *a posteriori*, ker je ena od premis, na katerih je utemeljen, *a posteriori*. Tako lahko pojem bistvenih lastnosti zagovarjamo samo z razlikovanjem med pojmoma *a priori* resnica in nujna resnica, kar pa zagovarjam tudi sam.

Vrnimo se k vprašanju identitet. Če obravnavamo izjavo "Hesperus je Phosphorus" ali izjavo "Cicero je Tulij", vse to lahko odkrijemo z empiričnim

raziskovanjem, za naša empirična prepričanja pa bi se lahko izkazalo, da so napačna. Tako kakor se običajno argumentira, morajo biti takšne izjave torej kontingentne. Nekateri so sprejeli drugo stran tega kovanca in so menili: "Zaradi tega argumenta o nujnosti morajo biti izjave o identiteti med imenoma spoznatne *a priori*, tako da samo zelo posebna kategorija imen morda v resnici deluje kot imena; druge stvari so neprava imena, prikriti opisi ali nekaj takega. Vsekakor za nek določen omejen razred izjav o identiteti vemo *a priori*, to pa so tiste izjave, ki vsebujejo pristna imena." Če sprejmemo razlikovanja, ki sem jih naredil, se nam ni treba zaleteti v nobenega od obeh sklepov. Lahko menimo, da so določene izjave o identiteti med imenoma – čeprav zanje pogosto vemo *a posteriori* in morda niso spoznatne *a priori* – dejansko nujne, če so resnične. Torej imamo nekaj prostora, da to menimo. Seveda pa imeti nekaj prostora, da tako menimo, ne pomeni, da bi tako tudi morali meniti. Oglejmo si torej evidenco. Prvič, če se spomnimo moje opazke, da se zdi, da so imena togi označevalci – kakor takrat, ko uporabimo ime Nixon –, da bi, celo v protidejstvenih situacijah, govorili o nekem določenem človeku. Če rečemo: "Če Nixon ne bi napisal pisma Saxbeju, bi morda spravil Carswella skozi,"[**] v tej izjavi govorimo o Nixonu, Saxbeju in Carswellu, prav o teh ljudeh iz aktualnega sveta in o tem, kaj bi se jim lahko zgodilo v določenih protidejstvenih okoliščinah. Če so imena togi označevalci, potem do vprašanja o tem, ali so te identitete nujne, sploh ne more priti, ker bosta "*a*" in "*b*" toga označevalca nekega določenega človeka oziroma stvari *x*. Potem se bosta *a* in *b* – celo v vsakem možnem svetu – nanašala na isti predmet *x* in na nobenega drugega, tako da do situacije, v kateri *a* morda ne bi bil *b*, sploh more priti. To bi bila namreč situacija, v kateri predmet, ki ga tudi sedaj imenujemo "*x*", ne bi bil identičen sam s seboj. Torej situacija, v kateri Cicero ne bi bil Tulij ali pa Hesperus ne bi bil Phosphorus, ni možna.[13]

Zakaj, če pustimo ob strani enačenje nujnosti z apriornostjo, so ljudje mislili drugače? Sta dve stvari, zaradi katerih so ljudje mislili drugače.[14] Nekateri ljudje

[**] 31. marca 1970 je takratni predsednik Združenih držav Amerike Richard Nixon odgovoril na pismo republikanskega senatorja Williama B. Saxbeja iz Ohia z dne 30. marca. Nixon je v njem vztrajal pri nominaciji sodnika G. Harrolda Carswella za člana vrhovnega sodišča ZDA. Saxbe mu je v pismu, ki je bilo skupaj s predsednikovim objavljeno 1. aprila tega leta, odgovoril, da Carswell ni nepristranski, kar kaže "njegovo pričanje o njegovi vlogi pri ustanovitvi rasno segregacijskega podeželskega kluba. Prepričan sem, da to vprašanje, poleg obtožb, da Carswell ni samo rasist, ampak tudi človek povprečne inteligence in nadarjenosti, skrbi mnogo članov senata."

[13] Strinjam se s Quinom, da "Hesperus is Phosphorus" je (oziroma je lahko) empirično odkritje; z Marcusovo pa, da je nujno tako. Tako Quine kot Marcusova se glede na predstavljeno stališče motita v tem, da enačita epistemološke in metafizične zadeve.

[14] Ti dve navedeni zamenjavi, še posebej druga, se obe nanašata na zamenjavo metafizičnega vprašanja o nujnosti izjave "Hesperus je Phosphorus" z epistemološkim vprašanjem o njeni apriornosti. Kajti če je Hesperus identificiran z njegovim položajem na nebu zvečer, Phosphorus pa z njegovim položajem zjutraj, lahko raziskovalec prav dobro ve, še pred empirično raziskavo, da Hesperus je Phosphorus, če in samo če eno in isto telo zavzema položaj *x* zvečer in položaj *y* zjutraj. *A priori* materialna ekvivalenca teh dveh izjav pa vendar ne implicira njune stroge (nujne) ekvivalence. (Iste opazke veljajo za primer toplote in gibanja molekul.) Podobne opazke do neke mere veljajo za odnos med "Hesperus je Phosphorus" in "'Hesperus' in

se nagibajo k temu, da obravnavajo identitetne izjave kot metajezikovne izjave, k temu, da identificirajo izjavo "Hesperus je Phosphorus" z metajezikovno izjavo "'Hesperus' in 'Phosphorus' sta imeni istega nebesnega telesa". To pa bi seveda lahko bilo neresnično. Termina "Hesperus" in "Phosphorus" bi morda lahko uporabljali kot imeni *dveh* različnih nebesnih teles. Seveda pa to nima nobene zveze z nujnostjo identitete. V istem smislu bi lahko bilo neresnična tudi izjava "2+2 = 4". Frazi "2 + 2" in "4" bi lahko uporabljali tako, da bi se nanašali na dve različni števili. Na primer, zamislimo si lahko jezik, v katerem bi "+," "2," in "=" uporabljali na standarden način, "4" pa bi uporabljali za, recimo, kvadratni koren od minus 1, ki ga sicer imenujemo "*i*". Potem bi bilo "2 + 2 = 4" neresnično, ker 2 plus 2 ni enako kvadratnemu korenu od minus 1. To pa ni to, kar hočemo. Nočemo reči samo tega, da bi lahko neka izjava, ki jo dejansko uporabljamo za izražanje nečesa resničnega, izražala nekaj neresničnega. To izjavo hočemo uporabljati na *naš* način in videti, ali bi lahko bila neresnična. Naredimo torej to. Kako ljudje razmišljajo? Pravijo: "Poglej, Hesperus bi morda ne bil Phosphorus. Tukaj imamo nek določen planet, ki je bil viden zjutraj in je bil viden zvečer, kasneje pa se je dejansko empirično izkazalo, da sta bila en in isti planet. Če bi se izkazalo drugače, bi bila to dva različna planeta oziroma dve različni nebesni telesi, kako torej lahko rečeš, da je neka takšna izjava nujna?"

Sta dve stvari, ki jih lahko imajo v mislih ti ljudje. Prvič, lahko menijo, da ne vemo *a priori*, ali Hesperus je Phosphorus. To sem že dopustil. Drugič, morda menijo, da si lahko dejansko zamislijo okoliščine, ki bi jih imeli za okoliščine, v katerih Hesperus ne bi bil Phosphorus. Naj pomislimo, kakšne bi bile te okoliščine, tako da tukaj uporabimo ta termina kot *imeni* nekega planeta. Na primer, lahko bi se zgodilo, da Venera zjutraj vzide na prav tistem mestu, kjer smo jo videli, da pa po drugi strani na mestu, ki ga Venera dejansko zaseda zvečer, ni Venere, njeno mesto pa zasede Mars. To vse je protidejstveno, ker je Venera dejansko tam. Nadalje, lahko si tudi zamislimo, da bi v tem protidejstvenem svetu Zemljo naseljevali ljudje, ki bi uporabljali ime "Phosphorus" za Venero zjutraj in "Hesperus" za Mars zvečer. Vse to je seveda v redu, vprašamo pa se lahko, ali bi to bila situacija, v kateri Hesperus ni Phosphorus? Seveda je to situacija, v kateri bi ljudje lahko resnično *rekli* "Hesperus ni Phosphorus"; vendar naj bi stvari opisali v našem jeziku, ne v njihovem. Opišimo torej to situacijo v našem jeziku. Dobro, kako bi se lahko dejansko zgodilo, da Venera zvečer ne bi bila na tem mestu? Na primer, recimo, da obstaja nek komet, ki pride naokoli vsak večer in malo premakne stvari. (To bi si na znanstven način lahko zelo enostavno zamislili: v resnici ne na preveč enostaven način – dejansko si to zelo težko zamislimo.) Zgodi se torej, da pride naokrog vsak večer, stvari pa se malo premaknejo. Mars se premakne prav na

'Phosphorus' imenujeta isto stvar". Zamenjava, ki je tudi na delu, je seveda zamenjava med tem, kar *mi* rečemo o neki protidejstveni situaciji, in tem, kako bi jo opisali ljudje, ki so *v* tej situaciji; tudi ta zamenjava se verjetno nanaša na zamenjavo apriornosti in nujnosti.

mesto, kjer je Venera, potem komet zjutraj premakne stvari nazaj v njihov normalen položaj. Če razmišljamo o tem planetu, ki ga sedaj imenujemo "Phosphorus", kaj naj rečemo? Dobro, lahko rečemo, da ta komet prečka nebo in premakne Phosphorus, tako da tega ni na mestu, ki ga zvečer normalno zavzema Phosphorus. Če res to rečemo in v resnici uporabimo "Phosphorus" kot ime nekega planeta, potem moramo reči, da v teh okoliščinah Phosphorus ne bi bil na mestu, kjer ga dejansko vidimo; oziroma drugače, Hesperus zvečer ne bi bil na mestu, na katerem ga dejansko vidimo. Morda bi lahko rekli, da v teh okoliščinah Hesperus ne bi imenovali "Hesperus", ker bi bil Hesperus na drugem mestu. Zaradi tega pa se Phosphorus še vedno ne bi razlikoval od Hesperusa; kar bi se zgodilo, je, da bi bil Hesperus na drugem mestu, ne pa na mestu, kjer dejansko je in nemara ne bi bil v takšnem položaju, da bi ga ljudje imenovali "Hesperus". To pa ne bi bila situacija, v kateri Phosphorus ne bi bil Hesperus.

Vzemimo drug primer, ki bo morda jasnejši. Predpostavimo, da nekdo uporablja "Tulij" za nanašanje na rimskega govorca, ki je ovadil Katilino, ime "Cicero" pa za nanašanje na človeka, čigar dela je moral študirati pri latinščini v tretjem letniku gimnazije. Seveda morda ne ve vnaprej, da je ta dela napisal prav tisti človek, ki je ovadil Katilino, in da je to neka kontingentna izjava. Toda zaradi dejstva, da je ta izjava kontingentna, ne bi smeli misliti, da je izjava, da Cicero je Tulij – če je resnična, dejansko pa je resnična –, kontingentna. Predpostavimo, na primer, da je Cicero dejansko ovadil Katilino, vendar je mislil, da je bil ta politični dosežek tako velik, da mu ni treba napisati nobene knjige. Bi rekli, da bi to bile okoliščine, v katerih ne bi bil Cicero? Zdi se mi, da je odgovor negativen – v teh okoliščinah bi pač rekli, da Cicero ni napisal nobenega literarnega dela. To, da bi moral napisati določena dela, ni nujna lastnost Cicerona – na tak način, kakor senca sledi človeku; zlahka si lahko zamislimo situacijo, v kateri Shakespeare ne bi napisal Shakespearovih del, ali pa situacijo, v kateri Cicero ne bi napisal Ciceronovih del. Kar se morda zgodi, je, da fiksiramo referenco termina "Cicero" z uporabo neke opisne fraze, kakršna je "avtor teh del". Ko pa smo enkrat fiksirali to referenco, uporabljamo ime "Cicero" *togo*, da bi označili človeka, ki smo ga dejansko identificirali z njegovim avtorstvom teh del. Ne uporabljamo ga, da bi označili kogarkoli, ki bi napisal ta dela namesto Cicerona, če jih je napisal kdo drug. Morda bi se zgodilo, da človek, ki je napisal ta dela, ne bi bil človek, ki je ovadil Katilino. Morda bi ta dela napisal Kasij. Toda potem ne bi rekli, da Cicero je Kasij, razen če bi govorili na nek zelo ohlapen in metaforičen način. Rekli bi, da Cicero teh del, s kateri smo ga morda identificirali in s katerimi je postal znan, ni napisal, namesto njega pa jih je napisal nekdo drug, recimo Kasij.

Na osnovi takšnih primerov ne moremo misliti, da so identitetne izjave kontingentne. Če bi na tej osnovi mislili tako, bi napačno razložili relacijo med *imenom* in *opisom, uporabljenim za fiksiranje njegove reference* – imeli bi jih za *sinonima*. Celo če referenco nekega takšnega imena, kot je "Cicero", fiksiramo s

človekom, ki je napisal ta in ta dela, potem, kadar govorimo o protidejstvenih situacijah, ko govorimo o Ciceronu, ne govorimo o komerkoli, ki *bi* v takšnih protidejstvenih situacijah napisal ta in ta dela, ampak o Ciceronu, ki smo ga identificirali s kontingentno lastnostjo, da je človek, ki je dejansko, to je v aktualnem svetu, napisal določena dela.[15]

Upam, da je to v glavnem jasno. Vendar sem dejansko predpostavil nekaj, za kar v resnici ne verjamem, da je v splošnem resnično. Predpostavimo, da referenco nekega imena res fiksiramo z nekim opisom. Celo če to naredimo, s tem to ime ne bo *sinonimno* s tem opisom, ampak to ime uporabljamo *togo* za nanašanje na predmet, ki smo ga tako imenovali – celo če govorimo o protidejstvenih situacijah, kjer imenovana stvar ne bi izpolnila omenjenega opisa. Po mojem mnenju je to dejansko resnično za tiste primere imenovanja, kjer je referenca fiksirana z opisom. Vendar pa tudi mislim, v nasprotju z večino novejših teoretikov, da je referenca imen redko fiksirana s pomočjo opisa oziroma da se to skoraj nikoli ne zgodi. S tem pa ne menim tega, kar je rekel Searle: "Reference ne fiksira nek posamezni opis, ampak prej nek grozd, neka družina lastnosti." Menim, da se lastnosti v tem smislu *sploh* ne uporabljajo. Vendar nimam časa da bi se tukaj spuščal v to. Predpostavimo torej, da je resnična vsaj ena polovica prevladujočih stališč o imenovanju: da je referenca fiksirana z opisi. Celo če bi to bilo resnično, ime ne bi bilo sinonimno z opisom, ampak bi bilo uporabljeno za to, da *imenuje* nek predmet, ki ga izberemo s kontingentnim dejstvom, da izpolni nek določen opis. In tako, celo če si lahko zamislimo nek primer, kjer človek, ki je napisal ta dela, ne bi bil človek, ki je ovadil Katilino, ne bi smeli reči, da bi to bil primer, v katerem Cicero ne bi bil Tulij. Morali bi reči, da je to nek primer, v katerem Cicero ni napisal teh del, ampak jih je napisal Kasij. Identiteta med Ciceronom in Tulijem pa še vedno velja.

Naj se vrnem na primer toplote in gibanja molekul. To je gotovo primer kontingentne identitete! Novejša filozofija je to vedno znova poudarjala. Če je to torej primer kontingentne identitete, potem si dajmo zamisliti, v katerih okoliščinah bi bila neresnična. Glede te izjave pa sam menim, da okoliščine, ki so jih filozofi očitno imeli v mislih kot okoliščine, v katerih bi bila ta izjava neresnična, dejansko niso takšne okoliščine. Najprej se seveda dokazuje, da je

[15] Če bi glede tega pulta kdo protestiral, da *bi se* navsezadnje *lahko izkazalo*, da je bil narejen iz ledu, in bi zato lahko bil narejen iz ledu, bi mu odgovoril, da v resnici meni, da bi lahko *nek pult* izgledal ravno tako kot tale in bil postavljen na isto mesto kot tale, vendar pa bi bil narejen iz ledu. Na kratko, glede *nekega pulta, narejenega iz ledu*, bi lahko bil v *isti epistemološki situaciji*, v kateri glede *tega* pulta dejansko sem. V glavnem tekstu sem argumentiral, da bi morali enako odgovoriti na ugovore, ki pravijo, da bi se lahko izkazalo, da Hesperus ni Phosphorus ali da Cicero ni Tulij. Tukaj potem pride do izraza pojem "dvojnika". Ker iz ledu ni bila izklesana ta miza, ampak nek epistemski "dvojnik"; ne Hesperus-Phosphorus-Venera, ampak dva različna dvojnika tega – v dveh vlogah, v katerih Venera dejansko nastopa (kot zvezda Danica in kot zvezda Večernica), ki sta različni. Ravno zaradi tega dejstva *ta miza* ne bi mogla biti narejena iz ledu. Izjave o modalnih lastnostih te mize se nikoli ne nanašajo na dvojnike. Če pa kdo meša epistemološke in metafizične probleme, je na dobri poti k teoriji dvojnikov, ki so jo zagovarjali Lewis in drugi.

"Toplota je gibanje molekul" aposteriorna sodba; z znanstvenim raziskovanjem bi se morda izkazalo drugače. Kot sem povedal že prej, to ne pove nič proti stališču, da je ta sodba nujna – vsaj v primeru, če imam prav. Vendar pa so tukaj ljudje gotovo imeli v mislih zelo specifične okoliščine, v katerih, tako so mislili, bi bila sodba, da je toplota gibanje molekul, neresnična. Kakšne bi bile te okoliščine? Lahko bi jih potegnili iz dejstva, da smo to, da je toplota gibanje molekul, odkrili empirično. Kako je bilo s tem? Kaj smo odkrili najprej, ko smo odkrili, da je toplota gibanje molekul? Obstaja nek določen zunanji pojav, ki ga lahko občutimo s čutilom za dotik in ki proizvede nek občutek, ki ga imenujemo "občutek toplote". Potem odkrijemo, da je ta zunanji pojav – ki proizvede občutek, ki ga občutimo z našim čutilom za dotik – dejansko molekularno nihanje v stvari, ki se je dotikamo, neka zelo velika stopnja molekularnega nihanja. Torej, bi morda lahko mislili, si moramo – zato da si zamislimo situacijo, v kateri toplota ne bi bila gibanje molekul – samo zamisliti situacijo, v kateri bi imeli prav isti občutek, ta občutek pa bi proizvedlo nekaj drugega, ne pa gibanje molekul. Podobno bi si lahko, če bi si hoteli zamisliti situacijo, v kateri svetloba ni tok fotonov, zamislili situacijo, v kateri bi bili na čisto isti način občutljivi na nekaj drugega, ki bi proizvedlo – vendar ne na podlagi toka fotonov – to, čemur pravijo vidno izkustvo. Da bi primer okrepili oziroma da pogledamo na drugo stran kovanca, bi lahko obravnavali tudi situacijo, v kateri *smo* v stiku z gibanjem molekul, vendar nam v tej situaciji takšno gibanje ne da občutka toplote. In morda bi se lahko tudi zgodilo, da bi bili mi oziroma bitja, ki bi naseljevala ta planet, narejeni tako, da bi nam, recimo, povečanje v gibanju molekul ne dalo tega občutka, ampak bi nam nasprotno prav ta občutek dala upočasnitev molekul. To bi bila situacija, tako bi morda lahko mislili, v kateri toplota ne bi bila gibanje molekul oziroma, natančneje, situacija, v kateri temperatura ne bi bila povprečna molekularna kinetična energija.

Jaz pa mislim, da ne bi bilo tako. Razmislimo znova o tej situaciji. Prvič, razmislimo o njej v aktualnem svetu. Takoj si zamislimo naš svet, ki so ga zasedli številni marsovci, ki – ko se dotaknejo ledu, ki ima počasno molekularno gibanje – dobijo prav ta občutek, ki ga mi imenujemo "občutek toplote", in ki ne dobijo občutka toplote – dejansko morda ravno nasprotni občutek –, ko dajo roko v bližino ognja, ki povzroča veliko molekularnega nihanja. Bi rekli "Oh, to meče senco dvoma na to, da je toplota gibanje molekul, ker obstajajo ti drugi ljudje, ki ne dobijo tega istega občutka"? Očitno, da ne, in nihče ne bi mislil tako. Namesto tega bi rekli, da marsovci, ko občutijo mraz, nekako dobijo prav ta občutek, ki ga dobimo mi, ko občutimo toploto, in da ne dobijo občutka toplote, ko občutijo toploto. Zdaj pa razmislimo o protidejstveni situaciji.[16] Predpostavimo, da bi bila

[16] Ali ni situacija, ki sem jo pravkar opisal, tudi protidejstvena? Lahko bi pač bila vsaj v primeru, če takšni marsovci dejansko nikoli ne napadejo. Strogo rečeno, razlikovati hočem med tem, kako *bi* govorili v neki (mogoče protidejstveni) situaciji, *če* bi prišlo do nje, in kako *res* govorimo *o* protidejstveni situaciji, za katero vemo, da se ne bo zgodila – tj., razlikovati hočem med jezikom, ki bi ga uporabljali v neki situaciji, in jezikom, ki ga *res* uporabljamo za to, da bi to situacijo opisali. (Obravnavaj opis: "Predpostavimo, da vsi

zemlja že od vsega začetka poseljena s takšnimi bitji. Najprej si zamislimo, da sploh ne bi bila poseljena: potem ni nikogar, ki bi imel kakršenkoli občutek toplote. Vendar pa ne bi rekli, da bi bilo v takšnih okoliščinah nujno, da toplota ne eksistira; rekli bi, da bi toplota morda lahko eksistirala, če bi, na primer, obstajali ognji, ki bi segrevali ozračje.

Predpostavimo, da fizikalni zakoni ne bi bili dosti drugačni: Ognji res segrevajo ozračje. Potem bi toplota obstajala, četudi tam naokrog ne bi bilo nobenih bitij, da bi jo občutila. Zdaj pa predpostavimo, da pride do evolucije in da nastane življenje in da so tam naokrog neka bitja. Ta pa niso podobna nam – bolj so podobna marsovcem. Bi sedaj rekli, da se je toplota nenadoma spremenila v mraz, ker jo bitja tega planeta občutijo tako? Ne, mislim, da bi to situacijo morali opisati kot situacijo, v kateri, čeprav bi občutila toploto, ta bitja na tem planetu toplote ne bi občutila takrat, ko bi bila izpostavljena toploti. Občutila bi jo, ko bi bila izpostavljena mrazu. In to je nekaj, kar si lahko gotovo brez težav zamislimo. To si lahko zamislimo ravno tako, kot si lahko zamislimo naš planet, ki so ga zavzela te vrste bitja. Razmislimo o tem v dveh korakih. Najprej obstaja neko stanje, v katerem bitij sploh ni, in gotovo si lahko zamislimo, da ima ta planet še vedno toploto in mraz, čeprav ni tam naokrog nikogar, da bi to občutil. Potem ta planet skozi evolucijski proces obljudijo bitja, ki imajo drugačno živčevje kot mi. Potem bi lahko bila ta bitja neobčutljiva na toploto; ne bi je občutila na naš način; po drugi strani pa bi občutila mraz čisto tako, kakor mi občutimo toploto. Vendar bi toplota še vedno bila toplota, mraz pa mraz. In zlasti to potem na noben način ne nasprotuje temu, da rečemo, da bi v tej protidejstveni situaciji toplota še vedno *bila* molekularno gibanje, *bila* tisto, kar proizvajajo ognji in tako dalje – ravno tako, kakor če na tem planetu sploh ne bi bilo nikakršnih bitij. Podobno bi si lahko zamislili, da bi ta planet naseljevala bitja, ki dobijo vidne občutke, ko so v zraku zvočni valovi. Zaradi tega ne bi smeli reči: "V takšnih okoliščinah bi zvok bil svetloba." Namesto tega bi morali reči: "Ta planet naseljujejo bitja, ki so v nekem smislu vizualno občutljiva na zvok, morda pa so celo vizualno občutljiva na svetlobo." Če je to pravilno, potem to, da je toplota gibanje molekul in da je svetloba tok fotonov, še vedno lahko je in še vedno bo nujna resnica.

Na kratko rečeno: termina "toplota" in "gibanje molekul" uporabljamo kot toga označevalca za nek določen zunanji pojav. Ker je toplota dejansko gibanje molekul in sta ta označevalca toga, na podlagi argumenta, ki sem ga podal tukaj, velja, da je *nujno*, da je toplota gibanje molekul. Videz kontingentnosti nam ustvari dejstvo, da smo toploto identificirali s kontingentnim dejstvom, da na tem planetu po naključju obstajajo bitja (namreč mi sami), ki so na določen način občutljiva nanjo, to je, ki so občutljiva na gibanje molekul oziroma na toploto – oboje sta ena in ista stvar. In to je kontingentno. Da bi identificirali toploto, torej uporabljamo opis "to, kar povzroča te in te občutke, oziroma to, kar občutimo na

govorimo nemško." Ta opis je v slovenščini [*in English*].) Prvi primer (kako bi govorili v protidejstveni situaciji) postane bolj nazoren, če si zamislimo, da je protidejstvena situacija dejanska.

ta in ta način". Toda pri uporabi tega dejstva uporabljamo neko kontingentno lastnost toplote, ravno tako kot – da bi ga identificirali – uporabljamo kontingentno dejstvo o Ciceronu, da je napisal ta in ta dela. Potem termin "toplota" v enem in termin "Cicero" v drugem primeru uporabljamo *togo* – da bi označili predmeta, ki ju zastopata. In seveda je termin "gibanje molekul" tog; vedno zastopa gibanje molekul, nikoli kakšen drugi pojav. Torej, kakor je rekel škof Butler: "Vsaka stvar je tisto, kar je, in ne nekaj drugega."** Zato bo "Toplota je gibanje molekul" nujno, ne kontingentno, *iluzijo* kontingentnosti pa imamo na isti način, kakor bi lahko imeli iluzijo kontingentnosti v razmišljanju, da bi bila ta miza morda narejena iz ledu. Morda bi mislili, da si lahko to kdo zamisli, če pa poskušamo, lahko na podlagi premisleka vidimo, da je to, kar smo si v resnici zamišljali, samo to, da je prav na tem mestu drug pult, ki je bil narejen iz ledu. Dejstvo, da lahko ta pult identificiramo s tem, da je predmet, ki ga vidimo in se ga dotikamo, na tem in tem mestu, je nekaj drugega.

Kako pa se to sedaj nanaša na problem duha in telesa? Običajno se to razume tako, da je to neka kontingentna identitetna izjava, ravno taka kot "Toplota je gibanje molekul". Vendar je ne more biti. Kontingentna identitetna izjava, kakršna je "Toplota je gibanje molekul", ne more obstajati, ker, če imam prav, "Toplota je gibanje molekul" ni kontingentna identitetna izjava. Poglejmo naslednjo izjavo. Na primer: "To, da čutim bolečino ob tem in tem času, je isto kot to, da sem ob tem in tem času v takšnem in takšnem možganskem stanju," oziroma: "Bolečina je v splošnem takšno in takšno živčno (možgansko) stanje."

To izjavo imajo za kontingentno na sledečih osnovah. Prvič, lahko si zamislimo, da to možgansko stanje eksistira, čeprav sploh ni nobene bolečine. Samo znanstveno dejstvo je, da takrat, ko smo v nekem določenem možganskem stanju, občutimo bolečino. Drugič, nekdo bi si morda zamislil neko bitje, ki občuti bolečino, ki pa sploh ni v nobenem določenem možganskem stanju, morda celo nima možganov. Ljudje celo mislijo, vsaj *prima facie*, čeprav se morda motijo, da si lahko zamislijo popolnoma raz(u)telešena [*disembodied*] bitja, vsekakor pa bitja, katerih telesa našim nikakor ne bi bila podobna. Tako se zdi, da si lahko zamislimo določene okoliščine, v katerih bi bil ta odnos neresničen. Vendar če so te okoliščine res okoliščine, si moramo zapomniti, da z njimi ne moremo enostavno opraviti tako, da rečemo, da je to samo iluzija, nekaj, kar si na videz lahko zamislimo, dejansko pa tega ne moremo narediti – na enak način smo napačno mislili, da bi si lahko zamislili situacijo, v kateri toplota ne bi bila gibanje molekul. Ker čeprav lahko rečemo, da smo toploto izbrali kontingentno, s kontingentno lastnostjo, da vpliva na nas na tak in tak način, ne moremo na enak način reči, da bolečino izberemo kontingentno – z dejstvom, da vpliva na nas na tak in tak način. Po tem stališču bi obstajalo možgansko stanje, mi pa ga izberemo s kontingentnim dejstvom, da vpliva na nas kot bolečina. Vendar bi to morda bilo res za možgansko stanje, ne more pa biti res za bolečino. To izkustvo

** *Fifteen Sermons*, Predgovor § 39.

samo mora biti *to izkustvo*, in ne morem reči, da je to, da je bolečina, kontingentno dejstvo o bolečini, ki jo sedaj občutim.[17] Dejansko je videti, da sta v prvi vrsti oba termina "moja bolečina" in "to, da sem v takšnem in takšnem možganskem stanju" toga označevalca. To je, kadarkoli je karkoli takšna in takšna bolečina, je bistveno prav ta predmet, namreč takšna in takšna bolečina, in kadarkoli je karkoli takšno in takšno možgansko stanje, je bistveno prav ta predmet, namreč takšno in takšno možgansko stanje. Torej sta oba termina toga označevalca. Ne moremo reči, da bi ta bolečina morda bila nekaj drugega, morda kakšno drugo stanje. Oba sta toga označevalca.

Drugič: način, kako bi razmišljali o tem, kako ju izberemo – namreč bolečino, s tem ko je izkustvo neke določene vrste, in to možgansko stanje, s tem ko je stanje nekega določenega materialnega predmeta, ko je iz takšne in takšne molekularne konfiguracije –; oba izbereta svoja predmeta bistveno in ne naključno, to je, izbereta ju z bistvenimi lastnostmi. Kadarkoli *so* te molekule v tej konfiguraciji, mi *res* imamo takšno in takšno možgansko stanje. Kadarkoli občutiš *to*, imaš res neko bolečino. Tako se zdi, da je identitetni teoretik v težavah, kajti – ker imamo dva toga označevalca – omenjena identitetna izjava je nujna. Ker izbereta svoja predmeta bistveno, ne moremo reči, da je ta primer, kjer se zdi, da si zamišljate identitetno izjavo kot neresnično, v resnici neka iluzija, takšna kot iluzija, ki jo ima kdo v primeru toplote in gibanja molekul, ker je ta iluzija odvisna od dejstva, da toploto izberemo z neko določeno kontingentno lastnostjo.[18] Tako za

[17] Najpopularnejše identitetne teorije, ki jih zagovarjajo danes, eksplicitno ne izpolnijo te preproste zahteve. Kajti te teorije običajno menijo, da je neko mentalno stanje neko možgansko stanje in da je to, kar naredi možgansko stanje za mentalno, njegova "vzročna vloga", dejstvo, da stremi k temu, da proizvede določeno obnašanje (kot nameni proizvedejo akcijo ali pa bolečina bolečinsko obnašanje) in je rezultat določenega dražljaja (na primer bolečine pri vbodih). Če relacije med možganskim stanjem in njegovimi vzroki in posledicami obravnavamo kot kontingentne, potem je *biti takšno-in-takšno-mentalno-stanje* kontingentna lastnost možganskega stanja. Naj bo *X* neka bolečina. Teoretiki vzročne vloge identitete menijo, (1) da je *X* neko možgansko stanje, (2) da moramo dejstvo, da je *X* neka bolečina, analizirati (grobo) kot dejstvo, da je *X* rezultat določenega dražljaja in proizvaja določeno obnašanje. Dejstvo, omenjeno v (2), seveda obravnavajo kot kontingentno, možgansko stanje *X* bi morda eksistiralo in ne bi stremelo k temu, da v odsotnosti drugih pogojev proizvede primerno obnašanje. Tako (1) in (2) trdita, da bi neka določena bolečina morda obstajala, vendar ne bi bila bolečina. To se mi zdi samoevidentno absurdno. Zamislite si katerokoli bolečino: je možno, da bi lahko eksistirala *sama na sebi*, pa vendar ne bi bila bolečina?
Če *X* = *Y*, potem *X* in *Y* delita vse lastnosti, vključno z modalnimi lastnosti. Če je *X* bolečina in *Y* ustrezno možgansko stanje, potem je *biti neka bolečina* neka bistvena lastnost *X*-a in *biti neko možgansko stanje* neka bistvena lastnost *Y*-a. Če je ta korespondenčna relacija dejansko identiteta, potem mora biti za *Y nujno*, da ustreza neki bolečini, in za *X nujno*, da ustreza nekemu možganskemu stanju, kajpada temu posameznemu možganskemu stanju *Y*. Obe trditvi se zdita neresnični; *zdi se*, da je čisto možno, da bi *X* lahko obstajal brez ustreznega možganskega stanja; oziroma da bi to možgansko stanje lahko obstajalo, ne da bi bilo občuteno kot bolečina. Identitetni teoretiki, v nasprotju z njihovo skoraj splošno sedanjo prakso, ne morejo sprejeti teh intuicij; morajo jih zanikati, hkrati pa poskušati upravičiti. Tega pa sploh ni prav lahko narediti.

[18] Morda bi bilo tukaj koristno, da nanovo formuliram ta argument. Če sta "bolečina" in "stimulacija C-vlakna" toga označevalca, mora tisti, ki ju enači, obravnavati to identiteto kot nujno. Kako lahko to uskladimo z očitnim dejstvom, da bi se morda lahko izkazalo, da stimulacija C-vlakna sploh nima nobene zveze z bolečino? Morda bi poskusili odgovoriti z analogijo s primerom toplote in molekularnega gibanja; ta zadnja identiteta je tudi nujna, vendar pa kdo morda verjame, da bi se – preden je znanstvena raziskava pokazala

manevriranje ostane bore malo prostora; nemara nič prostora. Identitetni teoretik, ki sprejema, da je bolečina možgansko stanje, mora tudi sprejeti, da nujno je možgansko stanje. Zato ne more dopustiti, ampak mora zanikati, da bi lahko obstajale situacije, v katerih bi kdo imel bolečino, ne pa tudi ustreznega možganskega stanja. Vendar običajno v argumentih identitetne teorije to še zdaleč ni zanikano. Dejansko to že od vsega začetka dopuščata tako materialist kot njegov nasprotnik. Ta pravi: "Seveda bi lahko obstajal primer, da bi imeli bolečine brez možganskih stanj. To je kontingentna identiteta." Tega pa ne more biti. Sprejeti mora, da smo pod vplivom kakšne iluzije, ko mislimo, da si lahko zamislimo, da bi lahko bolečine obstajale brez stanj možganov. Edini model tega, kaj bi morda bila ta iluzija, o katerem lahko razmišljam – oziroma vsaj model, podan z analogijo, ki jo predlagajo sami materialisti, namreč model toplote in molekularnega gibanja –, v tem primeru enostavno ne deluje. Tako je materialist pred zelo žilavim izzivom. Pokazati mora, da te stvari, za katere mislimo, da jih lahko imamo za možne, dejansko niso možne. Pokazati mora, da te stvari, ki si jih lahko zamislimo, dejansko niso stvari, ki si jih lahko zamislimo. To pa zahteva nek filozofski argument, ki se zelo razlikuje od argumenta tiste vrste, ki je bil podan v primeru toplote in molekularnega gibanja. In moral bi biti globlji in subtilnejši od argumenta, ki ga lahko sam doženem, in subtilnejši od vseh, ki so se kdajkoli pojavili v kakršnikoli materialistični literaturi, ki sem jo prebral. Tako bi bil sklep te raziskave ta, da analitična orodja, ki jih uporabljamo, delujejo proti identitetni tezi in tako delujejo proti splošni tezi, da so mentalna stanja prav fizikalna stanja.[19]

drugače – morda lahko izkazalo, da molekularno gibanje ni toplota. Odgovor je seveda ta, da je v resnici možno samo to, da bi ljudje (oziroma kakšna racionalna ali čuteča bitja) lahko bili v *isti epistemski situaciji*, v kateri smo dejansko mi, in identificirali *nek fenomen* na isti način, kakor mi identificiramo toploto, namreč tako, da jo občutijo z občutkom, ki ga mi imenujemo "občutek toplote", ne da bi ta fenomen bil molekularno gibanje. Nadalje, ta bitja bi morda s prav nobenim živčnim mehanizmom ne bila občutljiva na molekularno gibanje (tj. na toploto). Nemogoče je razložiti, kako očitna možnost stimulacij C-vlakna na isti način ne bi bila bolečina. Tudi tukaj bi morali predpostaviti, da bi lahko bili v isti epistemološki situaciji in bi lahko identificirali nekaj na isti način, kakor identificiramo bolečino, ne da bi to ustrezalo stimulaciji C-vlakna. Toda bolečino identificiramo s tem, da jo občutimo, in če bi se stimulacija C-vlakna lahko pojavila, ne da bi občutili kakršnekoli bolečine, potem bi se – v nasprotju z nujnostjo te identitete – stimulacija C-vlakna pojavila, čeprav bolečina ne *bi obstajala*. Težava je v tem, da čeprav je "toplota" togi označevalec, je toplota izbrana s kontingentno lastnostjo, da se jo občuti na določen način; bolečina pa je na drugi strani izbrana z neko bistveno (pravzaprav nujno in zadostno) lastnostjo. Ker za občutek velja, da če je *občuten* kot bolečina, potem *je* bolečina.

[19] Vse argumente proti identitetni teoriji, ki se opirajo na nujnost identitete ali na pojem bistvene lastnosti, je seveda navdahnil Descartesov argument za njegov dualizem. Zgodnejši argumenti, ki so bili površno zavrnjeni z analogijami s toploto in molekularnim gibanjem ter z izumiteljem očal z dvojnimi lečami, ki je bil tudi poštni minister, so imeli takšen navdih: in takega ima tudi moj argument tukaj. R. Albritton in M. Slote sta me obvestila, da sta neodvisno poskušala podati esencialistična argumenta proti identitetni teoriji, verjetno pa so to naredili tudi drugi.

Najenostavnejši kartezijanski argument lahko nemara nanovo formuliramo na sledeči način: Naj bo "*A*" neko ime (togi označevalec) Descartesovega telesa. Potem Descartes argumentira, da ker bi on lahko eksistiral, celo če *A* ne bi, ◊ (Descartes ≠ *A*), torej Descartes ≠ *A*. Tisti, ki so ga obdolžili, da je naredil modalno napako, so pozabili, da je "*A*" tog. Njegov argument je veljaven, sklep pa pravilen, če sprejmemo

Naslednja téma bi bila moja lastna rešitev problema duh-telo, te pa nimam.

njegovo (morda dvomljivo) premiso. Po drugi strani pa – če menimo, da Descartes po njegovi smrti ne bi več eksistiral – lahko pridemo do "Descartes ≠ *A*" brez uporabe modalnega argumenta; ker če bi bilo tako, bi *A* brez dvoma, če bi bil truplo, preživel Descartesa. Tako bi imel *A* neko lastnost (obstajati v določenem času), ki je Descartes ne bi imel. Z istim argumentom lahko pridemo do tega, da kip ni kos kamna oziroma mešanica molekul, iz katerih je sestavljen. Če bi sklepali samo na neidentiteto, bi bil to potem morda šibak sklep. (Glej D. Wiggins, *Philosophical Review* 77 (1968), str. 90 in dalje.) Kartezijanski modalni argument pa lahko gotovo razširimo tako, da bi nam priskrbel tudi močnejše relevantne sklepe.