ANαliZA

časopis za kritično misel

Letnik 6, številka 1-2, junij 2002 · ISSN 1408-2969

FILOZOFIJA DAVIDA LEWISA

Danilo Šuster

Lewisova teorija protidejstvenikov

Pedagoška fakulteta Univerze v Mariboru

Citiranje: Danilo Šuster, »Lewisova teorija protidejstvenikov«, *Analiza: časopis za kritično misel* 6, št. 1-2 (2002): 95–108.

To delo je objavljeno pod licenco Creative Commons Priznanje avtorstva 4.0 Mednarodna (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/deed.sl.

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.

Danilo Šuster
Pedagoška fakulteta Univerze v Mariboru

LEWISOVA TEORIJA PROTIDEJSTVENIKOV[1]

Protidejstvenik "Če bi bilo res, da *A*, potem bi bilo res, da *B*" je irealni pogojnik v pogojnem naklonu, ki ima pomembno vlogo v analizah filozofsko zanimivih pojmov. David Lewis je (skupaj z R. Stalnakerjem) avtor pojasnila protidejstvenikov v modelu možnih svetov. Lewis pojasni protidejstvenik kot strogi pogojnik, pri katerem se njegova strogost spreminja. Teorija uspešno pojasni semantične intuicije o veljavnosti in neveljavnosti sklepanj, v katerih nastopajo protidejstveniki. Toda v pojasnilu nastopa pojem "podobnosti", ki ga Lewis sprva ni pojasnjeval, ampak se je zanašal na njegovo intuitivno razumevanje. Kritiki dokazujejo, da intuitivna globalna podobnost med svetovi ni ne zadostni ne nujni pogoj za resnico protidejstvenikov. S kritikami se strinjam, vendar menim, da vrednost semantičnega pojasnila s tem ni ogrožena.

Ključne besede: teorija protidejstvenikov, David Lewis, semantika možnih svetov, logika pogojnikov, podobnost.

> "Orr je bil nor in bi ga bilo mogoče zapisati za nesposobnega. Nič drugega mu ne bi bilo treba, kakor da prosi za to. Toda brž ko bi to storil, ne bi bil več nor in bi moral znova v akcijo. Orr bi bil nor, če bi še letel, in normalen, če bi se tega branil, toda brž ko je normalen, mora še leteti. Če še leti, je nor in mu v resnici ne bi bilo treba leteti: toda če noče leteti, je normalen in mora leteti."
>
> – Joseph Heller, Kavelj 22

I.

V razmišljanju o tem, kaj bi se zgodilo, če bi nesrečni pilot Orr zaprosil za razrešitev vojaških dolžnosti, se nam zvrti v glavi. Razlog za to je preplet pogojnikov, ki nas vodijo od ene nemogoče posledice do druge in nazaj. Pričujoča razprava je posvečena Lewisovi teoriji protidejstvenikov, pogojnikov, ki se začno z "če bi bilo resnično to, kar dejansko ni ..." in s tem poskus, da bi se

[1] V članku nastopajo predelani deli drugega in petega poglavja doktorske disertacije (Šuster, 1995), ki je nastala s podporo Ministrstva za znanost republike Slovenije.

s pomočjo logičnega in filozofskega pojasnila znašli v mreži pogojnikov in odvisnosti, v katero smo ujeti.

Protidejstveni pogojniki imajo pomembno vlogo v filozofskih analizah in pojasnilih. Logične in semantične lastnosti razreda pogojnih stavkov v naravnem jeziku so predmet preučevanja od Megarikov dalje. V članku bom predstavil sodobno semantično teorijo protidejstvenikov v modelu možnih svetov, ki jo je razvil David Lewis. Osredotočil se bom na semantično pojasnilo klasičnih sklepanj: hipotetični silogizem, kontrapozicija in okrepitev antecedensa. V Lewisovem pojasnilu neveljavnosti teh sklepanj nastopa pojem "podobnosti" med možnimi svetovi, ki ga Lewis sprva ni pojasnjeval, ampak se je zanašal na njegovo intuitivno razumevanje. Zato bom na koncu navedel nekaj kritik, po katerih intuitivna globalna podobnost med svetovi ni ne zadostni ne nujni pogoj za resnico protidejstvenikov. S kritikami se strinjam, vendar menim, da vrednost elegantnega semantičnega modela s tem ni ogrožena.

Najprej nekaj o pogojnikih. Stavek "Če *A*, potem *B*" je pogojni odvisnik, slovnica večine evropskih jezikov pa loči povedni in pogojni naklon (indikativ in konjunktiv). Zato se je uveljavila ločitev med pogojnikom v povednem naklonu in pogojnikom v pogojnem (veznem, želelnem) naklonu (lat. "modus subjunctivus"). V večini filozofskih razprav iz angleškega govornega področja je predmet obravnav "subjunctive conditional", pogojnik v veznem naklonu. Pogojni naklon podaja dejanje glede na možnost in pogoje njegove uresničitve, zato je primeren za izraz sodb o hipotetičnem, zgolj možnem (pa tudi želenem, zaukazanem, priporočenem ...), ki so pogosto predmet filozofskih raziskav. Protidejstvenik "Če bi bilo res, da *A*, potem bi bilo res, da *B*" (uporabljal bom oznako "*A* > *B*") je irealni pogojnik v pogojnem naklonu, saj ta konstrukcija nakazuje, da je antecedens dejansko napačen (zato PROTIdejstvenik). V slovenščini irealni pogojnik še najbolj natančno označimo s pomočjo besedice *ko*, ki zaznamuje nerealni pogoj, da se dejanje glavnega stavka ne more več uresničiti. "Ko bi prižgal vžigalico, bi ta zagorela" nas opozarja na nerealnost situacije, resničnost ali napačnost te zveze je (za slovničarja) neodvisna od dejstev.

Protidejstveniki imajo pomembno vlogo v filozofskih analizah in pojasnilih. Naj navedem nekaj tipičnih primerov takšnih analiz:

> "*Vzrok* lahko definiramo kot *predmet, ki mu sledi drugi, pri čemer vsem predmetom, ki so podobni prvemu, slede predmeti, podobni drugemu.* Ali z drugimi besedami, *če ne bi bilo prvega predmeta, ne bi bilo niti drugega*." (Hume, 1974: 119)
>
> Posplošitev "Vsi *F* so *G*" je pravi *zakon narave* in ne le univerzalna slučajna resnica, kadar iz nje sledi pogojnik "Če bi ta *a* bil *F*, tedaj bi *a* bil *G*."
>
> Oseba *x* ravna *svobodno*, kadar bi, če bi se tako odločila, lahko ravnala drugače, kot je.
>
> Propozicija *p* je *resnična* pomeni, da če bi obstajala Idealna Racionalna

Znanstvena Skupnost, bi ta sprejela propozicijo, da *p*.

Trditev, da ima predmet *a* (kocka sladkorja) neko *dispozicijsko* lastnost *F* (recimo topljivost) pomeni, da če bi *a* bil v okoliščinah *X* (v vodi), tedaj bi imel lastnost *G* (bi se raztopil).

Stavek v fikciji, recimo "Sherlock Holmes je detektiv", pomeni, da bi v primeru, če bi bile zgodbe Conana Doyla resnične, Sherlock Holmes zares bil detektiv.

Rešitve pomembnih filozofskih vprašanj – problem svobodne volje, status zakonov narave, pojasnilo pojma vzroka, narava dispozicijskih lastnosti, vprašanja ontologije in metafizike temelje na razmisleku o tem, kaj bi bilo res v neki nedejanski situaciji. Protidejstveniki nastopajo v teh pojasnilih, ker z njimi izražamo določene vzorce odvisnosti – vzročne, nomološke, razlagalne, mereološke ... Protidejstvena konstrukcija s tem, da označi hipotetično obnašanje predmetov ali uporabo pojmov ob izpolnitvi nekega hipotetičnega pogoja, zrcali tiste odvisnosti, ki jih želimo v analizi poudariti kot osnovno značilnost analiziranega pojma.

II.

Razcvet modalne semantike v 60-ih letih 20. stoletja je Davida Lewisa in Roberta Stalnakerja vodil do zamisli o pojasnilu protidejstvenikov v modelu možnih svetov. V semantiki možnih svetov je nujna resničnost propozicije razložena kot resnica te propozicije v vseh možnih svetovih. Stavek "$\Box A$" ("Nujno je res, da A") je tedaj resničen na svetu *i*, če in samo če je *A* resničen na vseh svetovih, ki so dostopni (možni) glede na izhodiščni svet *i*. Strogi pogojnik "$\Box(A \supset B)$"[2] pa je resničen na *i*, če in samo če je *B* resničen na vseh *A*-svetovih, ki so dostopni (možni) glede na svet *i* ("*A*-svet" je oznaka za možni svet, na katerem je *A* resnično).

Prva ideja v semantičnem modeliranju protidejstvenikov je kar ta, da protidejstvenik beremo kot neke vrste strogi pogojnik. Katera nujnost pa nastopa pri protidejstvenikih? Najbolj odmevna zamisel je bila, da je nujnost protidejstvenika resnica ustreznega materialnega pogojnika na vseh svetovih, ki so kar najmanj različni od sveta *i*, razlika je le v tem, da je na njih resničen dejansko napačni antecedens. Relaciji dostopnosti tedaj ustreza *podobnost* med svetovi, protidejstvenik pa je resničen, če je konsekvens resničen na vseh svetovih, ki so najmanj različni od dejanskega sveta (najbolj podobni dejanskemu svetu). Recimo, v drami B. Shawa junakinja pravi:

Če ti ne bi bilo ime Ernest, te ne bi ljubila.

Kdaj je to res? Vzemimo možne svetove, ki se od dejanskega sveta razlikujejo samo v tem, da na njem našemu Ernestu ni ime Ernest (ampak, recimo, Edvard, Henrik, Harald ...). Če junakinja na vsakem od teh svetov v tega moža ni

[2] "$A \supset B$" je oznaka za materialno implikacijo.

zaljubljena, potem je navedeni protidejstvenik resničen. Kaj pomeni reči, da je nek svet minimalno različen od dejanskega sveta? Za začetek zadošča, da se opremo na intuitivno razumevanje minimalne razlike, ki temelji na pojmu podobnosti in bližine dejanskemu svetu. Svet, na katerem je možu ime Huang Chi, je tako intuitivno bolj različen od dejanskega sveta kot pa svet, v katerem je junaku ime Edvard.

Naj bo torej "[]*A*" splošna oznaka za "podobnostno" nujnost – propozicija "[]*A*" trdi, da je *A* resničen na vseh svetovih, ki so najbolj *podobni* dejanskemu svetu. "*i* R *j*" pa beremo kot: "svet *j* je najbolj podoben (blizu) svetu *i*". Pogoj za resnico protidejstvenika se potem glasi:

(SP) $A > B$ je resnično, če in samo če je $[](A \supset B)$ resnično na svetu *i*, torej kadar je *B* resnično na vseh *A*-svetovih, ki so v relaciji R s svetom *i*.

Protidejstvenik kot "podobnostno" striktni pogojnik je po (SP) resničen na *i*, če in samo če je ustrezni materialni pogojnik resničen na vseh svetovih, ki so v relaciji R ("najbolj podobni") s svetom *i*. Gre za strogost glede na operator "[]", ki jo v kriteriju (SP) izraža besedica VSI (resnica na vseh svetovih določene vrste).

Toda takoj ko imamo v modalni logiki strogost, so veljavna naslednja sklepanja:

(MON) Iz $A \vdash B$ sledi $A, C \vdash B$

(NEC) Če velja $\vdash A$ / $\vdash B$, tedaj velja tudi $[]A$ / $[]B$

(OA) Iz $[](A \supset B)$ sledi $[]((A \,\&\, C) \supset B)$

Za logiko protidejstvenikov je odločilna okrepitev antecedensa (OA), ki sledi iz (MON) in (NEC) tudi v najšibkejši modalni logiki, ki temelji na "[]". (OA) pa v logiki protidejstvenikov ni mogoče sprejeti. Vzemimo:

(1) Če bi uprasnil vžigalico, bi ta zagorela.

(2) Če bi uprasnil vžigalico in jo namočil v mleko, bi ta zagorela.

Po teoriji strogega pogojnika bi morala biti (1) in (2) resnična, toda drugi pogojnik je intuitivno napačen. Problem je v spreminjanju strogosti kriterija podobnosti. Če postavimo strogo merilo glede na antecedens (1) (v najbolj podobnih svetovih ne namakam vžigalice), tedaj je resnično (1), vendar tudi (2), saj je ob tem kriteriju antecedens NAPAČEN na vseh zelo podobnih svetovih, materialni pogojnik pa zaradi neresničnega antecedensa resničen. Če kriterij podobnosti omilimo (podobni svetovi obsegajo tudi svetove, v katerih vžigalico namakam), potem je (2) sicer napačno, vendar tudi (1), saj ni več res, da na vseh podobnih svetovih uprasnjena vžigalica zagori. Problem sklepanja (OA) je, da če je *B* resnično na vsakem *A*-svetu, ki je v relaciji R z *i*, tedaj je *B* resnično tudi na vsakem (*A* & *C*)-svetu, ki je v relaciji R z *i*!

Najbolj znano pojasnilo protidejstvenikov, ki ohranja idejo nujnosti, toda ovrže sklepanje (OA), je razvil David Lewis. V Lewisovi prvi formulaciji je nujnost označena kot sfera svetov S_i okoli sveta *i*, v kateri so vsi dostopni svetovi, torej svetovi, ki so možni glede na svet *i*. Stavek "$\Box A$" je tedaj resničen na svetu *i*, če

in samo če je A resničen v celotni sferi S_i. Strogi pogojnik "$\Box(A \supset B)$" pa je resničen na *i*, če in samo če je materialni pogojnik "$(A \supset B)$" resničen v celotni sferi S_i, torej je *B* resničen na vseh *A*-svetovih v sferi S_i. Za različne relacije dostopnosti dobimo različne vrste nujnosti – logični nujnosti ustreza množica vseh logično možnih svetov, fizikalni nujnosti množica vseh svetov, kjer veljajo zakoni sveta *i* (Lewis, 1973: 7).

Protidejstvenik definira *podobnostna* nujnost – v sferi S_i so vsi svetovi, ki so do določene fiksne mere podobni svetu *i*. Problem protidejstvenikov pa je, da lahko najdemo naslednja resnična zaporedja:

$A > B$

$(A \,\&\, A') > \neg B$

$(A \,\&\, A' \,\&\, A'') > B$

...

Recimo (primer navaja Bennett, 1974: 384):

(1) Če bi jaz hodil po zaledenelem jezeru (*A*), bi ostal led čvrst (*B*).

(2) Če bi jaz in ti hodila po zaledenelem jezeru ($A \,\&\, A'$), bi se led vdrl ($\neg B$).

(3) Če bi jaz in ti hodila bosa po zaledenelem jezeru ($A \,\&\, A' \,\&\, A''$), bi ostal led čvrst (*B*).

Ko bi bil protidejstvenik strogi podobnostni pogojnik, potem bi bil konsekvens *B* resničen na vseh "podobnostno" dostopnih *A*-svetovih. Ker pa so vsi protidejstveniki v zaporedju resnični, torej hkrati prvi in drugi, ne more biti dejanskemu svetu dostopen noben ($A \,\&\, A'$)-svet: če bi bil, bi bilo na njem hkrati resnično *B* in $\neg B$, torej bi bil tak svet nemogoč. Če je prvi protidejstvenik (1) resničen, potem je v modelu strogih pogojnikov tudi drugi (2) trivialno resničen, saj je po večini teorij protidejstvenik, katerega antecedens je resničen na nemogočem svetu, trivialno resničen. Zato je resničen vsak protidejstvenik s tem antecedensom, torej denimo tudi:

> Če bi jaz in ti hodila po zaledenelem jezeru ($A \,\&\, A'$), potem bi ostal led čvrst (*B*).

Rešitev iz zagate je v ideji, da nas vsak nov antecedens v zaporedju (A; $A \,\&\, A'$; $A \,\&\, A' \,\&\, A''$...) "napoti" v območje svetov, ki so bolj oddaljeni (manj podobni) od dejanskega sveta. Razvrstitev po podobnosti razvršča svetove v sistem sfer okoli sveta *i*. Protidejstvenik je v tem modelu spremenljivi, "podobnostno" strogi pogojnik, kjer je ustrezna sfera S_i množica vseh svetov, ki so vsaj do neke fiksne stopnje podobni svetu *i*. V semantiki zadošča primerjalni pojem podobnosti: če je svet *j* bližji svetu *i* kot *k*, potem je bližje središču (svetu *i*). Možni svetovi, ki so v enaki meri podobni *i*, tvorijo neko podsfero okoli sveta *i*. Sfera okoli sveta *i* je množica svetov S_i, tako da je vsak element v S_i dostopen svetu *i* in bolj podoben svetu *i* kot vsak neelement te sfere.

Vsakemu svetu *i* pripišemo sistem sfer podobnosti $\$_i$, množico množic možnih svetov $\{S_1^i, S_2^i ...\}$. Vsaka sfera S_i^k vsebuje svetove, ki so podobni *i* do določene mere, mera podobnosti pa se spreminja. Manjša ko je sfera okoli sveta *i*, bolj podoben mora biti svet svetu *i*, če naj sodi v to sfero. Če je *j* bližje *i* kot pa *k*, potem vsaj ena sfera okoli *i* vsebuje *j*, ne pa *k*. Celotna množica svetov, ki so kakorkoli podobni svetu *i*, pa je strukturirana v vgnezdeni sistem sfer – za vsaki dve sferi velja, da ena popolnoma vsebuje drugo. Če svet *j* sodi v sfero S_i, potem sodi tudi v vsako večjo sfero, ki vsebuje sfero S_i.

Vzemimo spet začetni primer:

(1) Če bi uprasnil vžigalico (*A*), bi ta zagorela (*B*).

(2) Če bi uprasnil vžigalico in jo namočil v mleko (*A* & *C*), bi ta zagorela (*B*).

Po logični diagnozi je v primeru (1) v sferi S^1 okoli sveta *i* vsak *A*-svet tudi *B*-svet, torej je resnično $A > B$. Toda sfera S^1 je neprimerna za protidejstvenik (2), saj želimo reči, da znotraj neke sfere podobnosti noben (*A* & *C*)-svet *ni B*-svet. Torej drugemu protidejstveniku pripišemo drugačno (širšo) sfero podobnosti S^2, v kateri bo veljalo, da na nobenem (*A* & *C*)-svetu ni resnično, da *B*. Ker pa bi morebitni dodatni pogoj *D* lahko spremenil oceno resnice protidejstvenika, v tem primeru spet razširimo sfero podobnosti in zatrdimo, da so vsi (*A* & *C* & *D*)-svetovi znotraj sfere S^3 tudi *B*-svetovi. Strogi podobnostni pogojnik zahteva, da izberemo in *fiksiramo* eno samo sfero dostopnosti – pri tem pa je bila S^1 izbrana za prvi antecedens, S^2 za drugi, S^3 za tretji ... Za protidejstvenik kot spremenljivo strogi pogojnik pa ni potrebno izbirati, saj obstajajo različne sfere, v katerih so resnični različni pogojniki. *Spremenljivost* je logikova rešitev, kako najti pogojnik med "A ⊃ B" in "□(A ⊃ B)", ki bo ustrezal logičnim lastnostim protidejstvenikov.

Strogemu pogojniku ustreza ena sama sfera možnih svetov, protidejstveniku pa sistem sfer, definiranih z relacijo podobnosti. Strukturo sistema sfer natančneje označimo z operacijami iz teorije množic, osnovno idejo pa izrazijo naslednje definicije (Lewis, 1973: 20):

> □*A* je resnično na *i*, če in samo če je *A* resničen na vsakem svetu v *vsaki* sferi S v $\$_i$.
>
> □(*A* ⊃ *B*) je resnično na *i*, če in samo če je *B* resničen na vsakem *A*-svetu v vsaki sferi S v $\$_i$.
>
> $A > B$[3] je resničen na *i*, glede na sistem sfer $\$_i$, če in samo če:
> (i) nobenega *A*-sveta ni v nobeni sferi S v S_i
> ali
> (ii) neka sfera S v S_i vsebuje vsaj en *A*-svet in *A* ⊃ *B* je resnično na vsakem svetu v S.

V primeru (i) gre za antecedens, ki ni resničen na nobenem svetu, ki je možen (dostopen) glede na *i*, protidejstvenik je tedaj trivialno resničen. Kadar pa je

[3] Lewis za protidejstvenik sicer uporablja oznako "*A* □→ *B*".

antecedens možno resničen, tedaj je protidejstvenik resničen, če in samo če je v neki *A*-dopustni sferi (*A*-dopustna sfera vsebuje nek *A*-svet) *B* resničen na vseh *A*-svetovih. V tem modelu je $A > B$ lahko resnično na *i*, čeprav obstajajo tudi (A & $\neg B$)-svetovi, kolikor so ti "bolj oddaljeni" od *i*, v neki sferi v sistemu $\$_i$, ki zaobsega svetove, ki so manj podobni *i* kot pa (A & B)-svetovi.

Zgodovinsko gledano je osnovne ideje pojasnila protidejstvenikov s pomočjo možnih svetov prvi razvil Robert Stalnaker. Koristno bo, če Lewisovo in Stalnakerjevo teorijo postavim v širši okvir teorij o pogojniku:

	POGOJNIK	**SPREMENLJIVOST**	**STROGOST**
(1)	$A \supset B$ (Filon, Russell)	–	–
(2)	$A > B$ (Stalnaker)	+	–
(3)	$\Box(A \supset B)$ (C. I. Lewis)	–	+
(4)	$A \Box\!\!\rightarrow B$ (D. Lewis)	+	+

Nespremenljivi nestrogi pogojnik (1) je klasični Filonov materialni pogojnik. "$A \supset B$" izraža neko posebno, fiksno propozicijo, ki je resnična ali napačna glede na dejstva v eni sami (dejanski) situaciji. Kot je znano, je C. I. Lewis (3) razumel pogojnik strožje – kot strogo implikacijo. Za ugotavljanje njene resnice v semantiki možnih svetov preverjamo resnico materialnega pogojnika v mnoštvu situacij, ki pa je fiksirano, kot bi danes rekli, samo glede na tip nujnosti.

Med obema "stabilnima" pogojnikoma stoji spremenljivi pogojnik, pri čemer pa imamo spet dve možnosti. Pogojnik je variabilni materialni pogojnik (2) ali pa variabilni strogi pogojnik (4). Robert Stalnaker zagovarja prvo možnost kot splošno teorijo pogojnikov, tako tistih v povednem kot tistih v veznem naklonu. David Lewis zagovarja teorijo o protidejstveniku kot variabilno striktnem pogojniku. Razlika med Stalnakerjem in D. Lewisom pa je mnogo manjša kot, recimo, razlika med Russellom in C. S. Lewisom. Teorija o variabilno materialnem pogojniku je namreč posebni primer teorije variabilne strogega pogojnika.

Variabilnost pogojnika $A > B$ pomeni, da se v semantični oceni protidejstvenikov spreminja izbor *A*-sveta; strogost pomeni, da ne gre samo za en *A*-svet, ampak za *VSE* svetove, ki izpolnijo določen kriterij. Lewis in Stalnaker se strinjata, da je izbor variabilen, toda za Lewisa antecedens izbere, kadar kaj izbere, *vse* svetove določene vrste, za Stalnakerja pa vedno le enega. Zato ima Stalnakerjeva definicija obliko:

> $A > B$ je resnično na svetu *i*, če in samo če je *B* resničen na najbližjem dostopnem *A*-svetu glede na svet *i*, kadar tak svet obstaja.

Če se izognemo finesam, potem lahko standardno pojasnilo protidejstvenikov v modelu možnih svetov zapišemo kot:

> Protidejstvenik $A > B$ je resničen na svetu *i*, če in samo če je *B* resničen na vseh *A*-svetovih, ki so najbolj podobni svetu *i*.

III.

Posebnost protidejstvenikov je intuitivna neveljavnost nekaterih klasičnih vzorcev, ki običajno določajo pogojnike: hipotetični silogizem (HS), kontrapozicija (KP) in okrepitev antecedensa (OA).

(HS) $A > B, B > C / A > C$

(KP) $A > B / \neg B > \neg A$

(OA) $A > B / (A \,\&\, C) > B$

Vzemimo (HS), ki skupaj s sklepanji *modus ponens*, *modus tollens* in (KP) tvori klasično logiko pogojnikov. Gre za samoevidenten vzorec sklepanja, ki ne potrebuje upravičevanja, saj gre za načelo, ki ga predpostavlja vsak postopek racionalnega upravičevanja sklepanj. Toda vzemimo naslednji sklepanji:

(Am) Če bi živel na Aljaski, bi živel v ZDA.

Če bi živel v ZDA, bi živel v New Yorku. (Ker so mi tam ponudili zaposlitev, tam so moji sorodniki ...)

Če bi živel na Aljaski, bi živel v New Yorku.

(K) Če bi se temperatura v sosednji sobi dvignila za dve stopinji Celzija, bi kanarček v kuhinji še vedno prepeval.

Če bi prišlo v kuhinji do eksplozije plina, bi se temperatura v sosednji sobi dvignila za dve stopinji Celzija.

Če bi prišlo v kuhinji do eksplozije plina, bi kanarček v kuhinji še vedno prepeval.

Sklepanji imata obliko: $B > C$ in $A > B$, torej $A > C$. Če so premise resnične, potem v klasični logiki pogojnikov oblika teh sklepanj zagotavlja resnico sklepa. Toda če so v navedenih primerih premise resnične, tedaj protidejstveni sklep ni resničen! Če bi živel na Aljaski, seveda ne bi živel v New Yorku. In če bi prišlo v kuhinji do eksplozije plina, je kanarček zelo verjetno ne bi preživel.

Proti kontrapoziciji govorita naslednja protiprimera:

> Če bi imel več kot dva otroka, ne bi imel več kot deset otrok. Torej, če bi imel več kot deset otrok, ne bi imel več kot dva otroka.
>
> Moj pes je mešanec – oče in mati psa sta mešanca. Če bi oče tega psa bil čistokrven, bi bil pes še vedno mešanec. Torej, če bi bil pes čistokrven, bi bil njegov oče mešanec. (Stalnaker, 1987: 124)

Protidejstvena sklepanja (OA), (HS) in (KP) so sistematično povezana. Pravzaprav lahko vsak protidejstveni protiprimer za (HS) preoblikujemo v protiprimer sklepanju (OA). Recimo, da je res:

> Če Janez ne bi bil izločen na natečaju za zaposlitev, Jože ne bi dobil te zaposlitve.

Kajti Janez je bil resen protikandidat in če ne bi bil izločen, bi on dobil zaposlitev. Res je tudi:

Če bi Janez umrl, ne bi bil izločen na natečaju za zaposlitev.

Mrtvega človeka seveda ni mogoče izločiti iz natečaja. Po (HS) bi potem sledilo:

Če bi Janez umrl, Jože ne bi dobil te zaposlitve.

Sklep pa ni resničen, saj bi Jože še vedno lahko dobil zaposlitev, tudi če Janez zaradi smrti ne bi sodeloval na natečaju.

Iz

Če Janez ne bi bil izločen na natečaju za zaposlitev, Jože ne bi dobil te zaposlitve.

bi po (OA) sledila neresnična propozicija:

Če bi Janez umrl in ne bi bil izločen na natečaju za zaposlitev, Jože ne bi dobil te zaposlitve.

Tudi sklepanji (KP) in (HS) sta med seboj povezani. Primer, ko ne velja (KP), lahko preoblikujemo v protiprimer za (HS). Recimo, da je na pošteni loteriji en glavni in več manjših dobitkov. Potem sprejmemo:

Če bi že zadel na loteriji, potem ne bi zadel glavnega dobitka. ($A > \neg B$)

Po (KP) pa sledi neresnica:

Če bi zadel glavni dobitek, tedaj ne bi zadel na loteriji. ($B > \neg A$)

V tem primeru seveda sprejmemo:

Če bi zadel glavni dobitek, tedaj bi zadel na loteriji. ($B > A$)

Iz obeh protidejstvenikov, ki ju sprejemamo kot resnična, $B > A$ in $A > \neg B$, sledi po (HS) sklep, ki je očitno neresničen:

Če bi zadel glavni dobitek, potem ne bi zadel glavnega dobitka. ($B > \neg B$)

Ko bi bil protidejstvenik materialni pogojnik, bi iz sklepa oblike $B > \neg B$ izpeljali, da ne bom zadel glavnega dobitka. Ko bi bil protidejstvenik strogi pogojnik, bi iz tega izpeljali, da nujno ne bom zadel glavnega dobitka. Oba sklepa se zdita problematična – prvi, ker protidejstvenik ni materialni pogojnik, in drugi, ker ni nujno, da ne zadenem glavnega dobitka!

Kot kažejo zgornji primeri, so sklepanja med seboj v sistematični zvezi. Iz veljavnosti (HS) sledi tudi veljavnost (OA):

(1) $A > B$ / hipoteza

(2) $(A \,\&\, C) > A$ / tavtologija klasične logike, ki jo lahko vedno vpeljem v dokaz

(3) $(A \,\&\, C) > B$ / iz 2 in 1 po (HS)

Iz veljavnosti (KP) prav tako sledi (OA) ob sprejetju oslabitve konsekvensa (OK): "Iz $A > B$ in $B \Rightarrow C$ izpelji $A > C$"; kjer "$B \Rightarrow C$" označuje, da "C" logično sledi iz

"*B*" (prim. Stalnaker, 1987: 125):

(1) $A > B$ / hipoteza

(2) $\neg B > \neg A$ / iz 1 po (KP)

(3) $\neg A \Rightarrow \neg(A \,\&\, C)$ / teza klasične logike, ki jo lahko vedno vpeljem v dokaz

(4) $\neg B > \neg(A \,\&\, C)$ / (OK) iz 3, 2

(5) $(A \,\&\, C) > B$ / iz 4 po (KP)

Povezave med navedenimi vzorci kažejo, da gre za neko globinsko, nemonotono logiko, ki razlikuje protidejstvenik tako od strogega kot od materialnega pogojnika. Semantični model tipa Stalnaker-Lewis (SL) je nastal ravno kot elegantno pojasnilo neveljavnosti sklepanj (HS), (KP) in (OA). Neveljavnost teh vzorcev semantično sledi iz *različnih* izborov najbližjega sveta glede na različne antecedense. Vzemimo začetni primer:

(Am) Če bi živel na Aljaski, bi živel v ZDA. ($A > B$)

Če bi živel v ZDA, bi živel v New Yorku. ($B > C$)

Če bi živel na Aljaski, bi živel v New Yorku. ($A > C$)

Neveljavnost (HS) temelji na tem, da je najbližji *A*-svet bolj različen od dejanskega sveta kot pa najbližji *B*-svet. Svet, v katerem živim na Aljaski, je manj podoben dejanskemu svetu od sveta, v katerem živim v ZDA (v kakem bolj razumnem kraju). "$A > B$" je resnično, torej je *B* resnično na najbližjih *A*-svetovih dejanskemu svetu. Najbližji "živim-na-Aljaski-svet" dejanskemu svetu je zares svet, v katerem živim v ZDA, saj je Aljaska del ZDA. Če je resnično "$B > C$", je "*C*" resnično na najbližjih *B*-svetovih, to pa so svetovi, v katerih živim v ZDA, ker imam tam službo, sorodnike … Čeprav velja "$A > B$", se najbližji *A*-svetovi (Aljaska svetovi) razlikujejo od najbližjih *B*-svetov (ZDA svetovi), zato protidejstvene posledice trditve "*B*" (v tem primeru je to "*C*") niso tudi protidejstvene posledice trditve "*A*". Najbližji *A*-svet smo izbrali tako, da smo se prestavili v hipotetično situacijo, v kateri živim na Aljaski. Gre za precejšen odmik od dejanskosti (morda me zagrabi čudna obsesivna želja po Aljaski ali zlata mrzlica ali kaj podobnega). V izboru najbližjega *B*-sveta pa se opiramo na drugačen in bolj verjeten scenarij, ta svet je bolj podoben dejanskemu svetu kot najbližji *A*-svet.

Podobna diagnoza velja tudi za (OA):

(1) Če bi uprasnil vžigalico, bi ta zagorela. ($A > B$)

(2) Če bi uprasnil vžigalico in jo namočil v mleko, bi ta zagorela. [$(A \,\&\, C) > B$]

Antecedensa v (1) in (2) izbirata različne podobne svetove. Najbližji *A*-svet je (za normalne ljudi) bližji dejanskemu svetu kot najbližji ($A \,\&\, C$)-svet. Svetovi so v (SL)-modelu popolnoma določeni, torej je v tem primeru najbližji *A*-svet, na katerem je res, da "*B*" (kjer uprasnem vžigalico in ta zagori) tudi $\neg C$-svet (svet, v katerem *ne* namočim vžigalice). Zato je resnično "$A > \neg C$", saj je nek ($A \,\&\, \neg C$)-

svet bližji *i* od vsakega (*A* & *C*)-sveta (v Lewisovem izvornem modelu), oziroma vsi najbližji *A*-svetovi so ¯*C*-svetovi. Najbližji (*A* & *C*)-svetovi pa so seveda takšni, da je v njih ¯*C* neresnično.

IV.

Model možnih svetov izjemno uspešno pojasni naše semantične intuicije o modalnih sklepanjih. Toda v njem nastopa pojem "bližine" ali "podobnosti", ki ga nismo podrobneje pojasnjevali. Kritiki modela možnih svetov včasih trdijo, da je analiza protidejstvenikov v semantiki možnih svetov povsem prazna, saj nima napovedne moči, ker ne pove ničesar o tem, kdaj je nek svet bolj ali manj podoben drugemu svetu. Lewis sprva govori o primitivni relaciji podobnosti, ki je nov semantičnotehnični pojem (Lewis, 1986: 163). Toda hkrati vendarle razume podobnost kot neko predhodno razumljeno relacijo globalne podobnosti med svetovi, njegovo pojasnilo pa je analiza protidejstvenikov v terminih te relacije (Lewis, 1973: 92).

Globalna podobnost med svetovi je parameter, ki odslikava siceršnjo nejasnost protidejstvenikov, njihovo odvisnost od konteksta, interesov in pričakovanj udeležencev v komunikaciji, ko izbirajo poteze sveta, ki so pomembne v oceni protidejstvenikov. Gre za mnoštvo podobnosti in razlik v mnoštvu aspektov primerjave glede na relativno pomembnost posameznih aspektov. Omejena nejasnost v podobnosti pojasni omejeno nejasnost protidejstvenikov in analiza na more biti jasnejša, kot je fenomen sam, pravi Lewis. S tem se je želel izogniti znanim kritikam, ki trdijo, da je podobnost neuporaben pojem, saj naj bi bilo aspektov podobnosti toliko, da sodbe oblike "*a* je bolj podoben *b* kot *c*" praktično nimajo nobene vrednosti. Pogoji resnice za protidejstvenik so fiksirani le približno – majejo se skupaj s podobnostjo, dva nejasna pojma sta nejasna na koordiniran način. Lewis se pri tem sklicuje na analogijo z barvo: obstajajo zelo nenatančni pogoji za resnico trditve, da je nekaj zeleno. Meja med modrim in zelenim ni fiksna, toda kljub nejasnosti vemo, da je nebo modro in trava zelena (Lewis, 1973: 92-93).

Težava s pojmom podobnosti je v tem, da intuitivna globalna podobnost ni ne zadostni ne nujni pogoj za resnico protidejstvenikov. Kot kažejo številni primeri, obstaja prepad med napovedno močjo po kriteriju podobnosti in intuitivno resnico protidejstvenikov. Intuitivna globalna sodba o podobnosti svetov daje prednost minimalnim razlikam od dejanskega sveta in ohranjanju potez dejanskosti, kar vodi do napačnih sodb o resnici protidejstvenikov. Recimo:

Če bi bil višji, kot sem, ...

> ... tedaj bi bil v najbolj podobnem svetu višji za en centimeter, ne pa za deset centimetrov.

Če bi bil tehtal med 100 in 120 kilogrami ...

Če bi imel več las, kot jih dejansko imam, ...

> ... tedaj bi tehtal, kot sem si sicer želel, 101 kilogram, ne pa 119 kilo-

gramov in imel bi en las več, ne pa, kot sem upal, precej las več.

"Če bi moj avto izginil iz garaže ..." reče voznik opoldne pred svojim vozilom.

Če bi avto izginil iz garaže, tedaj bi v najbolj podobnem svetu izginil malo pred izreko tega stavka, torej bi veljalo:

Če bi moj avto izginil iz garaže, tedaj bi izginil danes, ne pa včeraj zvečer.

Ampak protidejstvenik je napačen, saj je statistično bolj verjetno, da bi avto izginil ponoči!

Recimo, da imamo v laboratoriju substanco, ki se stali pri 23 stopinjah Celzija in zavre pri 24 stopinjah Celzija. Kaj reči o protidejstveniku:

Če bi bila temperatura višja, tedaj bi bila ta substanca v tekočem stanju (Nute, 1980: 70).

V modelu največje podobnosti je to res, toda intuitivno je napačno, kajti substanca bi se lahko uplinila. Ohranjanje potez dejanskosti v konstrukciji podobnosti vodi do napačnih sodb tudi v tipičnih primerih indeterminističnih procesov. Recimo, da vržem kovanec z levo roko in se pokaže cifra. Če bi vrgel kovanec z desno roko, bi se po modelu intuitivne podobnosti še vedno pokazala cifra, ker je svet, v katerem se pokaže cifra, pač najbolj podoben dejanskemu svetu, v katerem se je zares pokazala cifra (Kvart, 1986: 185). Ampak intuicije nam pravijo, da naslednji protidejstvenik ni resničen:

Če bi vrgel kovanec z desno roko, bi se pokazala cifra.

Če bi vrgel kovanec z desno roko, bi se lahko pokazala cifra ali grb, saj gre za indeterministični proces. Najbolj znan protiprimer Lewisovi analizi razvrščanja svetov po globalni intuitivni podobnosti pa je primer, ko v antecedensu nastopa majhna razlika od dejanskosti, ki vodi do konsekvensa, ki vodi do zelo različne dejanskosti. Gre za naslednji primer (iz časov hladne vojne, grozeče nuklearne nevarnosti in znamenite ovalne sobe, iz katere predsednik ZDA nadzira nuklearni arzenal svoje države):

Če bi Nixon pritisnil na gumb, bi prišlo do svetovne katastrofe (Fine, 1975: 452).

Svet, v katerem pride do katastrofe, se (na srečo in zaenkrat) po intuitivnih kriterijih globalne podobnosti med svetovi zelo razlikuje od dejanskosti, torej ji je manj podoben od sveta, v katerem Nixon pritisne na gumb, potem pa zaradi neke male okvare *ne* pride do sprožitve jedrskega orožja. Torej je protidejstvenik napačen po Lewisovem kriteriju podobnosti, čeprav bi ga v dani situaciji ocenili kot resničnega!

Lewis ima dva odgovora na te kritike. Prvi naj bi spodkopal očitke, da kriterij podobnosti vedno daje prednost čim manjšim razlikam od dejanskosti. Lewis uporabi standardno sklicevanje na kontekst, s katerim se spreminja "grobost" ("coarsness") relacije podobnosti. "Če bi bil tehtal med 100 in 120 kilogrami ..." v določenem kontekstu uvaja takšno relacijo podobnosti, po kateri vse teže med

100 in 120 vodijo do enako podobnih svetov. Torej ne upoštevamo minimalnih razlik, ki jih generira fina relacija podobnosti.

Bolj zanimiv je drugi odgovor, s katerim Lewis rešuje problem protidejstvenikov tipa: "Če bi se zgodila majhna sprememba v dejanskem svetu, tedaj bi ta sprememba vodila do sveta, ki se zelo razlikuje od dejanskega sveta." Lewis namesto intuitivne, globalne podobnosti, ki temelji na splošnih sodbah o podobnosti, postavlja vsebinsko določeno primerjalno podobnost. Merila za podobnost pa izpelje iz resnice protidejstvenikov, ki jih predhodno sprejemamo. Gre za dobro izdelan algoritem, v katerem tehtamo odmike od dejanskosti glede na kršitve zakonov narave in ujemanje v dejstvih, kar z intuitivno podobnostjo nima več prav veliko skupnega.

Z Lewisovo metriko ocenjevanja podobnosti med dejanskim svetom in najbližjim svetom, v katerem je realiziran dejansko neresnični antecedens, se v tej razpravi ne bom ukvarjal. Menim, da ni zadostna, toda to ne zmanjša vrednosti Lewisovega *semantičnega* modela. Podobnost je v semantičnem pojasnilu tehnični pojem, kjer nas zanimajo pogoji resnice protidejstvenikov zgolj v tisti meri, ki je potrebna, da razjasnimo njihove logične in sklepalne lastnosti. Formalna analiza razjasni strukturo problematičnega pojma in loči formalne probleme logike in veljavnosti določenih sklepanj od substancialnih vprašanj. Toda takšna analiza je samo skelet, okvir za postavitev substancialnih vprašanj. Od filozofskega pojasnila protidejstvenikov zares pričakujemo več, pojasnilo o naravi pogojev resnice za protidejstvenike – o tistem, kar naredi, da so protidejstveniki resnični, kadar so resnični. Kot substancialna teorija protidejstvenikov je Lewisova teorija podobnosti morda neuspešna, kot elegantno pojasnilo semantičnih lastnosti pa gre za model, ki je logiko pogojnikov postavil na nove temelje.

LEWIS ON COUNTERFACTUALS

Counterfactuals, conditionals of the form "if it had been the case, that A, then it would have been the case, that B", play an important role in the analyses of philosophically important concepts. David Lewis (and independently R. Stalnaker) developed an elegant semantical explanation in terms of possible worlds. According to Lewis counterfactuals are variably strict conditionals. This approach successfully explains semantical intuitions about inferential patterns valid in the classical logic of material and strict conditionals but invalid in the logic of counterfactuals. However, Lewis's account uses the notion of global similarity which was initially unanalysed. Counterexamples show that intuitive global similarity is neither a sufficient nor a necessary condition for the truth of counterfactuals. This critique might succeed as a refutation of the analysis of counterfactuals, but it does not undermine the value of semantical explanation in terms of possible worlds.

Keywords: David Lewis, theory of counterfactuals, conditional logic, possible worlds semantics, similarity.

Literatura

Bennett, J. (1974). "Counterfactuals and Possible Worlds". *Canadian Journal of Philosophy*, 4, str. 381-402.

Fine, K. (1975). "Critical Notice – Review of *Counterfactuals*". *Mind*, 84, str. 451-58.

Hume, D. (1974). *Raziskovanje človeškega razuma*. Ljubljana: Slovenska Matica.

Kvart, I. (1986). *A Theory of Counterfactuals*. Indianapolis: Hackett Publishing Company.

Lewis, D. (1973, s popravki 1986). *Counterfactuals*. Oxford: Basil Blackwell & Cambridge, Mass.: Harvard University Press.

Lewis, D. (1986). *Philosophical Papers Vol. II*. Oxford: Oxford University Press.

Nute, D. (1980). *Topics in Conditional Logic*. Dordrecht: D. Reidel.

Stalnaker, R. (1968). "A Theory of Conditionals". V Jackson, F. (ur.) (1991), *Conditionals*, Oxford: Basil Blackwell, str. 28-45.

Stalnaker, R. (1984[1], 1987). *Inquiry*. Cambridge, Mass.: The MIT Press.

Šuster D. (1995). *Sodobna teorija modalnosti – protidejstvena odvisnost in vzročnost* (doktorska disertacija). Ljubljana: Filozofska fakulteta.