ANαliZA

časopis za kritično misel

Letnik 15, številka 3, december 2011 · ISSN 1408-2969

VZROČNOST

Danilo Šuster

Lewisova teorija vzročnosti

Filozofska fakulteta Univerze v Mariboru

Citiranje: Danilo Šuster, »Lewisova teorija vzročnosti«, *Analiza: časopis za kritično misel* 15, št. 3 (2011): 5–23.

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 izvirna tiskarska datoteka številke, ki ji je bil leta 2026 dodan ta naslovni list. Strani in prelomi ustrezajo tiskani izdaji.

Danilo Šuster
Filozofska fakulteta Univerze v Mariboru

Lewisova teorija vzročnosti*

Najprej predstavim osnovne značilnosti Lewisove klasične protidejstvene teorije vzročnosti (če ne bi bilo vzroka, ne bi bilo učinka). Glavni razlog za izboljšave in premik v analizi vzročnosti od odvisnosti do vpliva predstavlja problem vzročne »izpraznitve«. V Lewisovi novi teoriji je učinek odvisen od vzroka, če in samo če bi se način nastopa vzroka razlikoval, tedaj bi se ustrezno razlikoval način nastopa učinka. Protidejstvena teorija temelji na pojmu odvisnosti in ideji vzroka kot ustvarjalca razlike, manjka pa ji ideja vzroka kot izvora in »proizvajalca« učinka. Kritično obravnavam dva poskusa analize vzročnosti, ki poskušata z uporabo bolj zapletenih vzorcev protidejstvene odvisnosti zajeti idejo vzročne zadostnosti. Nekateri (Hall) menijo, da gre za nepomirljivo dvojnost v našem pojmovanju vzročnosti, Lewisovo pojasnilo vzročnosti kot vpliva pa je poskus združitve obeh polov pojma vzroka.

Ključne besede: David Lewis, protidejstvena odvisnost, vzročnost, vpliv, dva pojma vzroka.

1.

David Lewis (1941-2001) je imel na univerzi v Harvardu marca leta 1999 »Whitehead Lectures,« ki so bila posvečena protidejstveni analizi vzročnosti, temi, h kateri se je vračal od svojih zgodnjih študijskih let.[1] Izhodišče takšne analize je naslednja Humova definicija: » ... lahko definiramo vzrok kot *predmet, ki mu sledi drugi, pri čemer vsem predmetom, ki so podobni prvemu, sledе predmeti, podobni drugemu*. Ali z drugimi besedami, *če ne bi bilo prvega predmeta, ne bi bilo niti drugega*« (Hume 1974, 119). V tej oznaki je vzrok najprej opredeljen kot nek pogoj, ki nastopa v stalni konjunkciji z učinkom, v drugem delu pa kot pogoj, brez katerega učinka ne bi bilo. Teorijo vzroka kot *pogoja* sta razvili pravilnostna teorija vzročnosti (tudi teorija regularnosti) in protidejstvena teorija vzročnosti. V sodobnih razdelavah Humove prve oznake gre v vzročni zvezi za pravilna, z zakoni povezana zaporedja dogodkov: *C* je vzrok za *E*, če in samo če je *C* del minimalne množice pogojev, ki so glede na zakone skupaj zadostni za *E* (prim. Šuster 2003).

Lewis (1973b, 161) začenja s Humovimi *drugimi besedami*: »O vzroku razmiš-

* V članku nastopajo predelani deli šestega poglavja doktorske disertacije (Šuster 1995), ki je nastala s podporo Ministrstva za znanost republike Slovenije.

[1] Prelomni članek je Lewis (1973b), krajša različica tu prevedene integralne verzije »Whitehead Lectures« pa je bila objavljena kot Lewis (2000).

ljamo kot o nečem, kar nekaj spremeni, pri čemer mora ta sprememba biti sprememba glede na to, kar bi se primerilo brez njega. Če vzroka ne bi bilo, ne bi bilo niti njegovih učinkov, vsaj nekaterih med njimi ne, običajno pa kar vseh«. Na razmisleku o vzroku kot pogoju, brez katerega ne bi bilo učinka, temelji že Millova metoda edine razlike kot preverjanja vzročnih korelacij. Če vzamemo to metodo kot pojasnilo vzročnosti, tedaj bo trditev, da je dogodek *C* vzrok za dogodek *E*, pomenila: nastopila sta dogodka *C* in *E* in če ne bi bilo dogodka *C*, potem ne bi bilo dogodka *E*. Protidejstvena analiza je danes splošno uveljavljena, nastopa v sociooloških in družboslovnih raziskavah, uporabljajo jo celo zgodovinarji (prim. Lebow, Tetlock in Parker, (ur.) 2004), zato je dandanes težko razumeti prelomnost Lewisovega predloga. Toda postaviti ga moramo v zgodovinski kontekst – kaj sploh pomeni pogojnik: če ne *bi* bilo *C*, tedaj ne *bi* bilo *E*? V filozofski analizi nekega pojma ne smemo uporabiti pojmov, ki so sami bolj problematični od analiziranega pojma. Za empiriste pa so protidejstveniki, ki govore o neaktualnih možnostih, celo bolj obskurni od pojma vzroka samega. Šele pojasnilo protidejstvenih pogojnikov, ki sta ga neodvisno predstavila Stalnaker (1968) in Lewis (1973a) je omogočilo, da je danes protidejstvena analiza eden najmočnejših filozofskih programov analize vzročnosti.

Izkazalo pa se je, da je protidejstvena odvisnost – brez vzroka ni učinka – premočna zahteva za vzročnost. Včasih bi brez vzroka učinek sicer še vedno nastopil, vendar drugače. In v tej smeri razmišlja »pozni« Lewis (2000, 2004) – vzrok je tisto, kar *vpliva* na učinek, čeprav nastop učinka ni odvisen od nastopa vzroka. V tem sestavku bom najprej predstavil osnovne značilnosti protidejstvene teorije vzročnosti. Glavni razlog za spremembo teorije in premik v analizi vzročnosti od odvisnosti do vpliva predstavlja problem vzročne »izpraznitve,« ko en vzrok »izprazni« (prepreči) nastop drugega vzroka, ki čaka v rezervi in bi opravil enako delo kot prvi (na primer Lee je ustrelil predsednika, ampak, če ga ne bi on, bi ga drugi strelec). Na koncu bom opozoril na nekaj težav novega predloga in dvojnost v našem pojmovanju: vzrok kot ustvarjalec razlike morda zahteva drugačno pojasnilo kot vzrok »proizvajalec«.

2.

Protidejstveniki so pogojniki oblike »Če bi bilo res, da *A*, potem bi bilo res, da *B*«, v katerih je propozicija *A* praviloma neresnična. Osnovna zamisel modela, ki sta ga predstavila Stalnaker in Lewis, je v tem, da je pogojnik resničen (na dejanskem svetu), kadar je potreben manjši odmik od dejanskosti, da je resničen (dejansko neresnični) antecedens *skupaj* s konsekvensom, kot pa da je uresničen antecedens brez konsekvensa. Svet, v katerem uprasnem vžigalico in ta zagori, je bližje dejanskemu svetu od sveta, v katerem uprasnem vžigalico in ta ne zagori. Zato je protidejstvenik »Če bi uprasnil vžigalico, bi ta zagorela« resničen. Bolj tehnično – pogojnik »Če bi bilo res, da *A*, potem bi bilo res, da *B*« je resničen na svetu *i*, če in samo če je *B* resničen na vseh najbližjih *A*-svetovih svetu *i*. To je nekakšno standardno vmesno stališče med Stalnakerjem in Lewisom – v nasprotju z Lewisom sprejemamo, da za vsak možen *A* obstaja

neprazna množica najbližjih svetov, v katerih je *A* resničen. In v nasprotju s Stalnakerjem zanikamo, da vedno obstaja največ en najbližji *A*-svet (o teh razlikah prim. Šuster, 2002). Ampak kako določiti bližino?

Za Lewisa so najbližji svetovi tisti, ki so najbolj *podobni* dejanskemu svetu, upoštevaje vse vidike podobnosti in razlik, ki jih primerjamo in tehtamo. Vzemimo staro šalo: *- Ta vžigalica se noče in noče vžgati. - Zakaj ne? - Pojma nimam. Prejšnjikrat je pa takoj zagorela.* Ali to ogrozi naš protidejstvenik »Če bi uprasnil vžigalico, bi ta zagorela«? Ne, saj je svet, v katerem še enkrat uprasnem ugasnjeno vžigalico v *danem* kontekstu manj podoben dejanskemu svetu od sveta, v katerem uprasnem vžigalico, katere kapica je dejansko nedotaknjena. Katere vidike upoštevamo? Prevladujoči zakoni narave določajo značaj možnega sveta, torej imajo podobnosti zakonov veliko težo – zakoni narave določajo, da glede na kemični sestav kapice, dane okoliščine (kisik) in silo trenja vžigalica zagori in v najbolj podobnih svetovih so ti zakoni ohranjeni. Lewis govori o primerjalni globalni podobnosti med svetovi, ki je osnovna in tudi odslikava odvisnost protidejstvenikov od konteksta, interesov in pričakovanj udeležencev v komunikaciji, ko izbirajo poteze sveta, ki so pomembne v oceni protidejstvenikov. No, vedno pa pri Lewisu velja, da je dejanski svet najbližji dejanskosti, bolj je podoben samemu sebi kot katerikoli drugi svet.

Lewisovo razvrščanje svetov po globalni intuitivni podobnosti je deležno številnih kritik (prim. Šuster 2002), k njej se bomo še vrnili, a oglejmo si zdaj, na kakšen način protidejstveniki nastopajo v Lewisovem izvornem pojasnilu vzročnosti (Lewis 1973b). Izhodišče je odvisnost, ki jo bom imenoval krepka protidejstvena odvisnost med dvema propozicijama (KPO): »Če ne bi bilo resnično, da *A*, tedaj ne bi bilo resnično, da *B*« in »Če bi bilo resnično, da *A*, tedaj bi bilo resnično, da *B*.« Naj bo '>' oznaka za protidejstveni pogojnik (»če bi …, potem bi … .«).

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

Ker velja odvisnost med propozicijami, ne pa med dogodki, moramo povezati propozicije z dogodki. Če je *C* dogodek, naj bo $O(c)$ propozicija, da je *C* nastopil, $O(e)$ pa propozicija, da je nastopil dogodek *E*. Za poljubna dogodka *C* in *E* potem velja, da je *E* vzročno odvisen od *C*, če in samo če:

VZ.O i. *C* in *E* sta različna dogodka, ki sta dejansko nastopila,
ii. $O(c) > O(e)$,
iii. $\neg O(c) > \neg O(e)$.

Spomnimo se, da je za Lewisa dejanski svet najbolj podoben dejanskemu svetu (*sebi*). To pomeni, da iz dejanske resnice propozicij $O(c)$ in $O(e)$ takoj sledi resnica protidejstvenika »$O(c) > O(e)$« (najbližji $O(c)$-svet je kar dejanski svet, v katerem pa je po pogoju (i) resnično tudi, da $O(e)$). Zato bo ob dejanskem

nastopu dveh različnih dogodkov *C* in *E* takoj izpolnjena zahteva (ii), vzročna zveza pa bo temeljila le na veljavnosti hipotetičnega pogoja (iii): brez *C* ni *E*. Avtomatsko resnico pogoja (ii) v Lewisovi logiki protidejstvenikov zagotavlja teza (CS):

CS $(A \,\&\, B) \quad (A > B)$

Kadar je dogodek *E* vzročno odvisen od dogodka *C*, takrat je *C* vzrok za *E*. Ne velja pa vedno obratno, velikokrat veriga vmesnih dogodkov vodi od *C* do *E*, tako da *E* ni *neposredno* vzročno odvisen od *C*, pa je kljub temu *C* vzrok za *E*. Poklical sem v zdravstveni dom in zdaj je pred vrati zdravnik: moj klic je povzročil zdravnikov prihod, toda vmes je še cel sklop dogodkov, ki so do tega privedli. Vzročnost je za večino teoretikov tranzitivna, vzročna odvisnost pa temelji na protidejstveni odvisnosti, ki *ni* tranzitivna. Zakaj ne?

Naj bodo *C*, *D* in *E* trije dejanski dogodki tako, da *D* ne bi nastopil brez *C* in *E* ne bi nastopil brez *D*. Tedaj je *C* vzrok za *E*, četudi bi *E* lahko nastopil (povzročen na nek drug način) brez *C*. Za zgled vzemimo primer vzročne izpraznitve [*preemption*], ki ima pomembno vlogo Lewisovih popravkih izvorne analize. Recimo, da dogodek (C_1) povzroča drugega (*E*) in obenem »izprazni« vzročno delovanje tretjega (C_2), ki bi sicer imel enak učinek, če prvi ne bi nastopil. Dejansko delujoča vzročna veriga »izniči« delovanje rezervne, alternativne verige. Recimo, računalniški program za pisanje mi omogoča izdelavo rezervnih kopij vseh odprtih dokumentov v vnaprej določenih časovnih razmikih. Glede na mojo izbiro bodo rezervne kopije izbrisane, ko pravilno shranim vsa besedila. Če je normalni proces shranjevanja prekinjen (denimo zaradi izpada elektrike), ne izgubim vsega, saj imam še vedno ohranjene rezervne kopije. Ob 10.00 se odločim, da bom shranil vse, kar sem napisal (C_1), zato pritisnem na tipko »shrani« (*D*) in datoteka se zapiše na trdi disk (*E*). Glede na prejšnjo določitev bi program točno ob 10.00 naredil rezervno kopijo (C_2), vendar dogodek *D* ta ukaz »izprazni«. Brez C_1 ni *D* in brez *D* ni *E*, toda dogodek *E* bi nastopil tudi brez C_1, saj je v rezervi čakal dogodek C_2. Odvisnost ni tranzitivna, vendar je bil prav dogodek C_1 vzrok dogodka *E*.

Za Lewisa je tranzitivnost vgrajena v pojem vzročnosti, zato dopolni VZ.O do definicije vzročne zveze tako, da vpelje *vzročne verige*. Če je *C*, *X*, *E* ..., končno zaporedje dejanskih posameznih dogodkov tako, da je *X* vzročno odvisen od *C* (tj. $\neg C > \neg X$), *E* od *X* (tj. $\neg X > \neg E$) in tako naprej, potem je to zaporedje vzročna veriga. In končno, en dogodek povzroča drugi dogodek, če in samo če obstaja vzročna veriga, ki vodi od prvega do drugega. Dogodek *C* je potem vzrok za dogodek *E*, če in samo če:

i. *C* in *E* sta različna dogodka, ki sta dejansko nastopila.

ii. Med *C* in *E* obstaja vzročna veriga: za neko končno zaporedje dogodkov $X_0, X_1 ..., X_n$, tako, da $X_0 = C$, in $X_n = E$, za vsak $i < n$, velja: $O(X_i) > O(X_{i+1})$ in $\neg O(X_i) > \neg O(X_{i+1})$.

V nadaljevanju bom zapis poenostavil in pisal »C«, »E« včasih kot dogodka, včasih pa kot propoziciji, ki bi ju sicer moral pisati kot O(c) in O(e). Protidejstvena odvisnost ni tranzitivna, relacija »biti prednik protidejstvene odvisnosti« [*ancestral*], naj bo to »R«, pa je in ta zagotavlja tranzitivnost vzročnosti. *C* je v odnosu R do *E*, če in samo če je *E* neposredno protidejstveno odvisen *C ali* pa je *E* prek vmesnih korakov protidejstvene odvisnosti povezan s *C* (zato je odločitev, da bom shranil zapisano, vzrok za zapis datotek).

3.

Kadar sta dva dogodka povezana kot vzrok in učinek, takrat učinek ni povezan z vzrokom na enak način, kot je dogodek vzrok povezan z dogodkom učinkom. To je sorazmerno enostavna ugotovitev o asimetrični zgradbi vzročnosti, za katero pa je težko najti dobro teoretsko pojasnilo, gre za tako imenovani problem »učinkov«. Umrl je zaradi zastrupitve z gobami: če se ne bi zastrupil, ne bi umrl. Ampak, ali ni potem res tudi, da če ne bi umrl, potem ne bi bilo res, da se je zastrupil z gobami? Brez vzroka ni učinka in brez učinka ni vzroka: kako lahko potem protidejstvena odvisnost (enostavno ali po korakih) razloži *asimetrijo* vzročnosti?

Če bi ne bilo učinka, bi vzrok še vedno nastopil, vendar bi nekaj presekalo njegovo »vzročno moč«, pravi Lewis. Problem rešuje tako, da sprejema protidejstvenik: ¬C > ¬E, zanika pa obratni protidejstvenik, ki je v tem primeru (časovni) *oziralec*, ki pravi, da če učinek *E* ne bi nastopil, tedaj njegov vzrok *C* ne bi nastopil. Zakaj ne? »Oziralci« [*back-tracking conditionals*] so protidejstveniki, v katerih je čas nastopa antecedensa *za* časom nastopa konsekvensa, torej govorijo o protidejstveni odvisnosti preteklega od sedanjega. Tipično so resnični protidejstveniki usmerjeni v času naprej, govore o tem, kaj bi se zgodilo *po* tem, ko bi nastopil dogodek ali stanje, ki ga uvaja antecedens. Oziralci so tipično nedoločeni in nejasni. V literaturi običajno nastopa primer prepira med Jackom in Jimom (prim. Bennett 2003: 205). Jack je dan po prepiru še vedno zelo jezen na Jima, torej, če bi ga Jim prosil za pomoč, mu Jack *ne bi* pomagal. Toda Jim je ponosen mož, on nikdar ne bi prosil Jima za pomoč po takšnem prepiru. Torej, če bi Jim prosil Jacka za pomoč, včeraj ne bi bilo prepira, Jack bi bil spet dobrodušni prijatelj in *bi* pomagal Jimu. V tem primeru nihamo med dvema protidejstvenikoma, v prvem protidejstveno sklepamo v času *naprej* (ni pomoči), v drugem v času *nazaj* (ni prepira).

Protidejstveniki so za Lewisa v splošnem nejasni, in nejasnost aspektov podobnosti razrešujemo v različnih kontekstih na različne načine. Trdi pa, da je v običajni razrešitvi nejasnosti protidejstvena odvisnost asimetrična in oziralci vedno napačni. Če bi bila sedanjost drugačna v tem, da bi nekem trenutku nastopil *A* namesto ne-*A*, tedaj bi bila preteklost (več ali manj) nespremenjena. Samo posebni konteksti (potovanje v času, morebitna vzročnost za nazaj) morda zahtevajo nestandardno razrešitev nejasnosti in aspektov podobnosti, v kateri izidejo oziralci kot resnični. Vžigalica pred mano je bila uprasnjena (*C*) in je zagorela (*E*). Če je ne bi uprasnil, potem ne bi zagorela. Ampak, če ne bi zago-

rela, ne moremo nedvomno reči, kaj bi bilo resnično *pred tem*, torej tudi tega ne, da pred tem *ni* bila uprasnjena. Glede na Lewisovo metriko podobnosti je manjši odmik od dejanskosti v tem, da se znebimo dogodka *E* tako, da *ohranimo* njegov vzrok *C*. Dogodek *C* bi še vedno nastopil, vendar ne bi povzročal dogodka *E*.

Če se nam to zdi čudno, razmislimo o količini sprememb, potrebnih, da realiziramo nek hipotetični antecedens. Denimo, da vklopim stikalo in luč zagori, kar privede do tega, da muha vzleti, foton doseže strop ipd. Če stikala ne bi vklopil, luč ne bi zagorela. Ampak, kaj po Lewisu pomeni domneva, da stikala ne bi vklopil? Zgodovina našega dejanskega sveta *i* nekako *skrene* malo pred trenutkom vklopa, toliko, da je dopuščen nek razumen prehod do nastopa hipotetičnega stanja stvari, ki ga opisuje antecedens. Ob predpostavki determinizma so s tem kršeni zakoni narave sveta *i*. V literaturi o protidejstvenikih je za takšno realizacijo antecedensa, ki nastopi v nasprotju z zakoni dejanskega sveta, uveljavljen Lewisov pojem »malega čudeža« [*small miracle*]. Če bi bil resničen oziralec, da če luč ne bi zagorela, tedaj stikalo pred tem *ne* bi bilo vklopljeno, tedaj bi v najbližjem svetu, kjer luč ne zagori, ne nastopila *nobena* od drugih vzročnih *sledi* vklopa. Ne le, da luč ne bi zagorela, ampak bi ne bilo niti zvoka, svetlobe, segretja ozračja, muha ne bi vzletela, ... V tem hipotetičnem svetu bi nastopil velik »čudež«, ki bi izbrisal vse sledi dejanskega vklopa stikala. Primerjajmo to z Lewisovim predlogom najbližjega sveta, v katerem je oziralec »Če bi luč ne zagorela, stikalo ne bi bilo vklopljeno« neresničen. Luč ne zagori, stikalo pa je še vedno vklopljeno in vse vzročne sledi vklopa ostanejo. Da pridemo do tega sveta, je potreben nek razmeroma manjši »čudež,« neka lokalna kršitev zakonov dejanskega sveta tik preden luč zagori, zaradi katere vklop stikala ne vodi do gorenja luči, vse ostalo (zvok, muha, foton, ...) pa je nespremenjeno.

Protidejstvena analiza elegantno rešuje tudi klasični problem *epifenomenov*. V nevihtnem oblaku dobijo vrhnji deli oblakov pozitivni, spodnji deli pa negativni naboj (*C*), kar povzroči blisk, izravnavo naelektritve (*E*) in grom (*D*). Potem velja: $\neg C > \neg E$ in $\neg C > \neg D$, vendar tudi $\neg E > \neg D$ (brez bliska ni groma). Toda blisk (*E*) ni vzrok za grom (*D*), oba sta le *epifenomena* skupnega vzroka (*C*). Rešitev za Lewisa je spet v zanikanju problematičnega protidejstvenika, da brez bliska ni groma. Kot vemo iz rešitve problema učinkov bi brez bliska njegov vzrok, naelektritev vrhnjega dela oblakov še vedno nastopila. In v tem primeru bi tudi grom še vedno sledil. Kršena (»mali čudež«) je, po Lewisovi analizi, samo *zakonita* vez, ki vodi od naelektritve do bliska.

4.

Lewisova analiza vzročnosti je takoj naletela na vrsto ugovorov. Protidejstvena odvisnost ne more biti zadostni pogoj za vzročno odvisnost, saj nastopa med dvema propozicijama tudi takrat, kadar ne gre za pravo vzročno odvisnost. Prve kritike so opozarjale na dejstvo, da je relacija protidejstvene odvisnosti preširoka, gre za *genus*, vzročnost pa je le ena od vrst protidejstvene odvisnosti. Kadar,

denimo, med *C* in *E* nastopa logična in analitično odvisnost, tedaj velja tudi protidejstvena odvisnost. Kim je naštel vrsto primerov, v katerih velja: ¬C > ¬E, pa vendar *C* ni vzrok za *E* (Kim, 1973). Najenostavnejši primer te odvisnosti nastopi, kadar je *E* konjunkcija, *C* pa eden od členov konjunkcije, recimo: ¬B > ¬(A & B). V »ontološkem« branju je konjunkcija sestavljeni dogodek »učinek,« »vzrok« pa njegov konstitutivni del. Kim navaja primere tipa: »Če ne bi napisal črke L, ne bi napisal imena glavnega mesta Slovenije«, toda prvi dogodek ni vzrok za drugega, ampak le njegov del. Problematični so tudi primeri tako imenovane »Cambridge odvisnosti«: »Če Sokrat ne bi umrl, tedaj Ksantipa ne bi postala vdova«. Kim trdi, da gre za dva različna dogodka, ki nastopita sočasno, na različnih krajih, pa vendar prvi ni vzrok za drugega. Drugi znani primeri zadevajo dejanja – dejavnik stori nekaj s tem, da stori nekaj drugega – s tem, da obrnem kljuko, odprem okno. Kimova splošna poanta je, da obstajajo različni vzorci odvisnosti med dogodki, ki izpolnijo pogoje protidejstvene analize. En dogodek lahko (asimetrično) določa drug dogodek, vendar ga ne določa vzročno.

Toda tako kot vsaka druga teorija vzročnosti, tudi protidejstvena teorija temelji na pojasnilu členov vzročne relacije, ki so tradicionalno *dogodki*. Če izključimo obskurne primere *causa sui*, tedaj vzročna relacija standardno nastopa med dvema *različnima* dogodkoma. V navedenih ugovorih ne gre za protidejstveno odvisnost med dvema različnima dogodkoma ali pa gre za primere, ko sploh ne gre za *prave* dogodke.[2] Prav vsaka teorija vzročnosti pa mora pojasniti, kaj sploh so dogodki. Res pa je, da je Lewis prisiljen v posebne koncesije. Po običajnem, aristoteljanskem pojmovanju, je dogodek neka sprememba v substanci, v sodobni terminologiji – predmet pridobi ali izgubi neko lastnost. Vendar pa protidejstvena analiza deluje samo, če obstajajo vzročne verige in s tem vmesni členi. Če pa za vsak vmesni člen zahtevamo, da je neka sprememba, tedaj enostavno ne bo dovolj vzročnih verig in vmesnih dogodkov. Zato je pri Lewisu dogodek v vzročni verigi tudi to, da neki premikajoči se predmet vztraja v svojem gibanju, da je sled bakra navzoča v danem vzorcu, skratka stanje stvari in vztrajanje v nekem stanju (Lewis 1986, str. 216 in 261).

5.

Najtrši oreh za protidejstveno analizo so primeri vzročne izpraznitve. Vzemimo standardni primer strelskega voda, v katerem je vsak strel sam po sebi zadosten za smrt kaznjenca. Vsaj eden (morda pa vsi) od strelov je vzrok smrti, čeprav ni resnično: če ne bi bilo *tega* strela, kaznjenec ne bi umrl. Ampak takšni primeri simetrične naddoločenosti so redki in nenavadni tudi v fiziki (prim. Bunzl, 1979). Tipično (ob hipotezi determinizma) obstaja za vsak dogodek en sam vzročno zadostni pogoj za njegov nastop. V primerih simetrične naddoločenosti (strelski vod) tudi naše intuicije o tem, kaj je pravi vzrok, niso jasne in protidejstvena analiza odslikava to negotovost. Problem pa je *a*simetrična naddo-

[2] Na Kimove ugovore odgovarjajo Swain (1980), Lewis (1986: 263) in Yagisawa (1979).

ločenost, kadar je jasno, da en dogodek povzroči učinek, vendar pri tem izprazni potencialni vzrok, ki čaka v rezervi. Vzemimo klasični primer:

Mož se odpravlja na pot skozi puščavo in si zato pripravi posodo z vodo. Ima pa dva sovražnika – prvi naskrivaj nalije v posodo strup (*S*), ki povzroča smrt z zadušitvijo, drugi pa, ne da bi vedel za dejanje prvega, preluknja posodo (*P*). Mož odpotuje, (zastrupljena) voda odteče, mož dehidrira (*D*) in umre (*U*).

Najbolj plavzibilna se zdi ugotovitev, da je smrt moža povzročilo preluknjanje posode, saj je umrl zaradi pomanjkanja vode. To dejanje je izpraznilo vzročno moč alternativnega vzroka (strupa), ki je čakal »v rezervi«. Prvi problem je v tem, da smrt ni vzročno odvisna od preluknjanja posode: mož bi umrl tudi, če drugi sovražnik ne bi preluknjal posode, zato ni resnično ¬P > ¬U. Kako je tedaj *P* lahko pravi vzrok?

Lewisov odgovor že poznamo. Če je *P* vzrok za *U*, hkrati pa ni resnično, da ¬P > ¬U, to v protidejstveni analizi zaradi *tranzitivnosti* vzročnosti pomeni, da nastopa *veriga*, torej vsaj en *vmesni* člen. V našem primeru je to dehidracija, zato velja: ¬P > ¬D in ¬D > ¬U. Ker protidejstvena odvisnost ni tranzitivna, sicer ni resnično, da ¬P > ¬U, vendar je *P* še vedno vzrok smrti. Ampak smrt moža ni neposredno odvisna niti od preluknjanja niti od strupa. Če mož ne bi dehidriral, bi še vedno umrl, tokrat zaradi delovanja strupa drugega sovražnika. Kako je tedaj smrt lahko vzročno odvisna prav od dehidracije? Zgodnji Lewis zanika vzvratno odvisnost preluknjanja od dehidracije: če mož ne bi dehidriral, bi bila posoda *pred* tem še vedno preluknjana (in s tem izpraznila vzročno delovanje strupa), le da preluknjanje ne bi povzročilo dehidracije.

Lewisova obravnava asimetrije oziralcev je za mnoge sporna. Analiza smrti v puščavi se res zdi nenavadna: če mož ne bi dehidriral, bi bila po Lewisu posoda *pred* tem še vedno preluknjana, strup zato ne bi deloval, le da preluknjanje ne bi povzročilo dehidracije in s tem smrti. Le kako bi bilo to mogoče? Bi mož našel kak skriti puščavski vir vode? Bi mu pomagali nomadi? Bi njegovo telo nekako zdržalo brez vode? Ali ni bolj intuitivno, da rečemo, da, če mož ne bi dehidriral, bi pač umrl zaradi strupa? Morda res, pravi pozni Lewis, vendar bi umrl *drugače* – ne zaradi dehidracije, ampak zaradi zadušitve. Splošno, kadar je dogodek *C* vzrok za dogodek *E*, vendar bi brez *C* dogodek *E* še vedno nastopil (zaradi delovanja »rezervne« vzročne verige), tedaj je osnovna intuicija ta, da bi brez *C* nastopil drugačen dogodek *E*. Dogodek *E* s tem postane zelo »krhek,« že majhna sprememba v času, kraju, vpletenih silah …, bi privedla do *drugačnega* dogodka. Protidejstvena analiza je s tem ohranjena, saj še vedno velja, da brez vzroka ne bi bilo *natanko takšnega* učinka, kot je dejansko nastopil.

Če Klement Jug ne bi leta 1924 padel v Triglavski severni steni, bi vseeno umrl, torej padec ni bil vzrok smrti? Umrl seveda bi, ampak drugače, torej je bil padec vzrok *te* določene smrti. Brez vzroka ne bi bilo učinka *ali* pa bi ta nastopil v drugem trenutku, kot je dejansko, ali pa bi nastopil na drugačen način od načina,

na katerega je dejansko nastopil. Prototip vzročnosti za zgodnjega Lewisa so mreže nevronov, v katerih nevroni so ali niso v vzburjenem stanju in aktivirajo ali pa ne druge nevrone. Binarno vzročno logiko tipa vse ali nič (brez prvega ni drugega) pozni Lewis razširja z analogno logiko vzročnega *vplivanja* in modificiranja, dogodek *C* je vzrok za dogodek *E*, samo če bi brez dogodka *C* ne bilo *E*, ali pa bi ta nastopil kasneje, drugje, drugače. Dobro, če Klement Jug pri plezanju ne bi nosil robca, bi bil njegov padec neznatno drugačen. Torej je bil tudi robec vzrok *te* smrti? Čisto vse, kar vodi do nekakšne razlike, še ne more biti vzrok!

Lewis je v svoji obravnavi vzročnosti metodološko precej ortodoksen in zagovarja klasično pojmovno analizo, vendar je v njegovem celotnem opusu zaznaven vpliv zgodnjih študijskih let, ko je poslušal predavanja Gilberta Rylea, H. P. Gricea, P. F. Strawsona in J. L. Austina, velikih zagovornikov tako imenovane filozofije običajnega jezika. Tako Lewis v svojo analizo nenehno vpleta *jezikovne* norme – kaj smo pripravljeni reči, do česa smo jezikovno upravičeni ipd. In te norme dopuščajo tudi bolj *robustno* istovetnost dogodkov, *isti* dogodek se lahko pripeti kasneje, drugje, malo drugače itd. (koncert se je začel ob 20.00, ampak, če bi se začel ob 20.05 v drugi dvorani, bi šlo še vedno za isti koncert). Prav tako velja, da je razlika v smrti, do katere privede robec, neznatna in jo lahko glede na norme zanemarimo. Norme niso fiksne – včasih dogodek nastopi drugače, vendar gre za isti dogodek; včasih dogodek nastopi drugače pa gre za drugi dogodek; in včasih dogodek zaradi nekega dejavnika nastopi malo drugače, vendar ta dejavnik ne šteje kot vzrok.

V literaturi o »izpraznitvi« največkrat nastopata Billi in Suzi, ki hkrati vržeta vsak svoj kamen v steklenico. Suzijin kamen zadene prvi in razbije steklenico. Brez Suzijinega meta bi tudi udarec Billijevega kamna privedel do raztreščenja steklenice. Bi v tem primeru šlo za drugačen dogodek-učinek? Lewis opozarja, da bi se v danem primeru steklenica raztreščila ob skoraj enakem času kot se je dejansko raztreščila in na skoraj enak način, kot se je. Če menimo, da je raztreščenje modalno krhko, potem že majhna sprememba vodi do drugačnega dogodka. Ampak jezikovna norma ni uniformna: včasih govorimo kot da so dogodki robustni (koncert), včasih tako, kot da so krhki (smrt v puščavi), včasih pa se morda ne moremo odločiti (raztreščenje steklenice). Zaradi neodločenosti glede teh vprašanj Lewis uvaja pojem *alteracije* (spremembe) dogodka.

Alteracija dejanskega raztreščenja steklenice je krhek alternativni dogodek (malo drugačno raztreščenje, ki pa je drugačen dogodek od dejanskega) ali pa krhka različica istega dejanskega dogodka (malo drugačno raztreščenje, ki pa je še vedno isti dogodek kot dejanski) ali celo zelo krhek dejanski dogodek raztreščenja steklenice. V vsakem primeru pa je alteracija *krhka*: že najmanjša sprememba naredi iz nje drugačen dogodek. Zdaj lahko Lewis tudi odvisnost vzročnega *vpliva* izrazi kot odvisnost ali-ali: dogodek *C* je vzrok za dogodek *E*, samo če bi brez dogodka *C* ne bilo dejanske alteracije dogodka *E*. Če bi dejanska alteracija učinka *E* nastopila kasneje, drugje, drugače …, bi zaradi krhkosti alteracij avtomatsko šlo za drugačen dogodek. Ker gre za alteracije, imamo odvisnost »vse ali nič,« vendar že majhna sprememba naredi »nič« (izniči dejanski učinek, ker ga spremeni v drugačen dogodek). V grobem: *C* je vzrok

za *E*, samo če brez *C* ne bi nastopil *E tak*, kot je. To lahko pomeni, da bi nastopil spremenjeni *E* (ampak še vedno *E*), ali pa nek drug dogodek in sploh ne *E*. Lewis trdi, da je odlika tega pojasnila prav v tem, da se nam ni treba vnaprej odločiti med alteracijami, ki so različne različice enega in istega dogodka ter alteracijami, ki so različni, vendar podobni dogodki.

6.

Preden pridemo do končne uradne verzije teorije vzročnosti kot vpliva si moramo ogledati še ugovor, ki spremlja protidejstveno analizo že od vsega začetka. Gre za problem »bližine«, ki nastane zaradi posebnosti v razumevanju najbližjega sveta, v katerem *ne* nastopi dejanski vzrok. Vzemimo primer, s katerim skuša Nute ovreči protidejstveno analizo vzročnosti (Nute, 1980: 96). Kamen je privezan na dolgi, fiksirani vrvici. Naj *C* označuje dogodek, ko odrežemo vrvico na določeni točki vrvice (*p*), *E* pa naj bo padca kamna, ki temu sledi. Če oba dogodka nastopita, je *C* vzrok za *E*, zato po predlagani analizi brez *C* ne bi bilo resnično, da nastopi dogodek *E*. Nute pa trdi, da ta protidejstvenik ni resničen. Kajti, če bi bila vrvica prerezana na neki drugi točki (milimeter više ali niže), bi kamen seveda še vedno padel. Najbližji svetovi, v katerih vrvica ni prerezana na dani točki *p*, so za Nuteja svetovi, v katerih je vrvica še vedno prerezana, vendar na neki drugi točki in kamen zato še vedno pade. V tem primeru protidejstvena odvisnost ni nujni pogoj za nastop vzročne zveze, saj *C* povzroča *E*, čeprav bi brez dogodka *C* dogodek *E* še vedno nastopil.

Če bi bil Nutejev ugovor odločilen, bi zlahka pokazali na popolno neustreznost protidejstvene analize vzročnosti. Kaj je namreč pravi antecedens v zgornjem primeru: »Če vrvice ne bi prerezal, …. « ali »Če vrvice ne bi prerezal na točki *p*, …«? Nute očitno misli, da gre za drugi antecedens. Toda, recimo, da vžigalica zagori, ker jo uprasnem. Torej, če vžigalice ne bi uprasnil, tedaj ne bi zagorela? Toda kateri je najbližji svet, v katerem vžigalice nisem uprasnil? Vžigalico sem dejansko uprasnil z določeno hitrostjo *v*, določeno silo *f*, v določenem trenutku *t*. »Če vžigalice ne bi uprasnil, ... « torej dejansko pomeni, če sledimo Nuteju, »Če vžigalice ne bi uprasnil s hitrostjo *v*, silo *f*, v trenutku *t*, …«. Najbližji svet, v katerem se to zgodi, je svet, v katerem uprasnem vžigalico malo manj silovito, pod drugačnim kotom ali kasneje ali, ... in vžigalica še vedno zagori. Ali bo po tem razmisleku v standardnem modelu protidejstvenikov sploh še ostal kakšen resničen protidejstvenik?

Lewis se tega problema zaveda in pravi, da v presoji o resnici protidejstvenika »Če ne bi bilo *C*, tedaj ne bi bilo *E*« in je dogodek *C* zelo krhek, ne nastopa *najbližji* ne-*C* svet. Kajti v najbližjem ne-*C* svetu nastopi alteracija *C*, ki ima zelo verjetno podobne učinke kot izvorni dogodek *C*. Lewis je sprva predlagal, da si zamislimo, da je *C* nekako »posesan« iz dejanskega sveta, popolnoma in čisto odstranjen iz zgodovine, ne da bi za sabo pustil kak svoj fragment ali približek. To morda deluje v svetu enostavnih dogodkov, ko se dani nevron sproži ali pa ne. V bolj zapletenih primerih pa je odločitev težja. Collins, Hall in Paulova

(2004: 20) navajajo primer pogovora med Suzi in Billijem, ki poteka od poldneva do 12.30. Ali bi še šlo za isti dogodek, če bi se pogovarjala o drugi temi? O isti zadevi, ampak uro kasneje, dan kasneje? Nekaj metrov stran, nekaj kilometrov stran? Če bi se pogovarjali drugi dve osebi? Če sploh ne bi šlo za pogovor, ampak za tekmovanje v teku med njima?

»Če bi Kserkes zmagal pri Salamini, bi še vedno bili barbari«, je zapisal Voltaire in večina zgodovinarjev se strinja glede pomena zmage Grkov v pomorski bitki za razvoj zahodne civilizacije (prim. Lebow, Tetlock in Parker (ur.), 2004: 90). V splošnem zgodovinarji nimajo težav s protidejstveniki in tem, kaj pomeni, da nek dogodek ne bi nastopil. Tudi pozni Lewis se opira na to, da se nekako znajdemo v razumevanju vzročnih protidejstvenikov, čeprav teoretsko ni povsem jasno, kakšne *natanko* so protidejstvene situacije, ki si jih zamišljamo. Tako kot pri učinkih se nam tudi pri vzrokih ni treba odločiti vnaprej. Ko upoštevamo vrsto alteracij vzroka *C*, nam ni treba reči, katere od njih so še različice vzroka *C* in katere so njegove alternative. Vzemimo spet Billija in Suzi, vendar upoštevajmo, kaj se zgodi, če se malo spremeni Suzijin met (kamen je lažji, Suzi vrže malenkost močneje, hitreje ipd.), Billijev met pa ostane takšen, kot je. Raztreščenje steklenice bo drugačno. Zdaj pa si oglejmo alteracije v učinku, v katerih je Billijev met na podoben način malo drugačen, Suzijin met pa ostane takšen, kot je. Raztreščenje bo natanko takšno kot je bilo. Suzijin met je vzrok, ki vpliva na raztreščenje, Billijev pa ne.

V grobem dogodek *C vpliva* na dogodek *E* kadar možne (a ne preveč ekstremne) alteracije vzroka *C* vodijo do pomembnih razlik v učinku *E*. Ali drugače povedano, dogodek *C* je vzrok za dogodek *E*, samo če bi brez dogodka *C* ali pa bi *C* nastopil kasneje, drugje ali drugače, ne bilo dogodka *E* ali pa bi ta nastopil kasneje, drugje ali drugače. Še bolj natančno, *C vpliva* na *E*, če in samo če obstaja vrsta C_1, C_2, ... različnih, ne-preveč-oddaljenih alteracij *C* (vključno z dejansko alteracijo *C*) in vrsta E_1, E_2, ... alteracij *E*, pri čemer se vsaj nekatere med njimi razlikujejo tako, da, če bi nastopil C_1, potem bi nastopil E_1 in če bi nastopil C_2, potem bi nastopil E_2 in tako naprej.

Binarno logiko vse ali nič (brez vzroka ni učinka) pri poznem Lewisu zamenja postopna sprememba: spremembe v vzroku so korelirane s spremembami v učinku. Vzemimo spet Nutejev ugovor. Če odrežemo vrvico na točki *p*, kamen pade. Če odrežemo vrvico na neki drugi točki (milimeter više ali niže), bi kamen še vedno padel, a malo drugače (zaradi drugačnega delovanja sil). Spreminjanje dogodka vzroka vodi do sprememb v dogodku učinku, a pri tem se nam ni treba odločiti, katere alteracije so različne verzije vzroka in katere so že alternativni, drugačni dogodki. Odvisnost »vse ali nič« je zdaj le *mejni* primer vpliva: prerez vrvice vpliva na padec kamna, v mejnem primeru *brez* prereza vrvice ni padca kamna. Bolj splošno: *E* protidejstveno kovariira s *C* samo v primeru, kadar spremembam v načinu nastopa *C* sledijo ustrezne spremembe v nastopu *E*. Zgodnji Lewis je poznal samo en način kovariacije: brez *C* ni *E*. Ampak to je le mejni primer, vzročnost je zdaj prednik relacije vpliva, kjer *C* vpliva na *E* samo, če *E* protidejstveno kovariira s *C* (v zadostni meri, Lewis vedno opozarja, da ne gre za odnos vse ali nič).

7.

Vzrok v protidejstveni analizi je relevantni »ustvarjalec razlike«. A mnogi bi rekli, da tu nekaj manjka – ali ni bistveno za relacijo vzročnosti, da učinki izvirajo, izhajajo, nastanejo iz svojih vzrokov? Elisabeth Anscombe, sama mati sedmih otrok, kot prototip vzročne relacije navaja fizikalno starševstvo, kjer gre prav za materialno izhajanje, z delitvijo (prim. Anscombe, 2003). Protidejstvena analiza pa nam ne pokaže vzroka kot *izvora*, ki je povezan s svojim učinkom. Povedano drugače: zaenkrat imamo pojasnilo vzroka kot nujnega pogoja, manjka pa dimenzija vzroka kot pogoja, ki je vzročno zadosten za nastop učinka.

Vzročno nujni pogoj izraža, v danih okoliščinah, odnos odvisnosti med ne-nastopom dogodka *C* in ne-nastopom dogodka *E*: brez dogodka *C* ne bi bilo dogodka *E*. Zadostni pogoj in s tem odvisnost nastopa dogodka *E* od dogodka *C* izraža drugi protidejstvenik: če bi nastopil dogodek *C*, tedaj bi nastopil dogodek *E*. Toda kadar sta *C* in *E* dogodka, ki sta dejansko nastopila, je pri Lewisu protidejstvenik $C > E$ avtomatsko resničen ne glede na to, ali sta dogodka *C* in *E* vzročno povezana ali ne. Morda bi lahko poskusili idejo zadostnosti izraziti s standardnim obratom v logiki pogojev (če *A*, potem *B* je ekvivalentno če ne-*B*, potem ne-*A*). *C* je potem zadosten za *E*, kadar je res: »Če ne bi nastopil dogodek *E*, tedaj ne bi nastopil dogodek *C*«. S tem pa smo takoj v težavah, ker potem nimamo več odgovora na problem učinkov – vzročna odvisnost postane simetrična. Kot je videti, smo v dilemi: po (enostavni) protidejstveni analizi sprejmemo, da je katerikoli dejanski dogodek vzročno zadosten za nastop kateregakoli drugega dejanskega dogodka ali pa se temu izognemo, vendar potem vzročna odvisnost ne ločuje med vzrokom in učinkom.

Vzorce odvisnosti in neodvisnosti odkrijemo šele v modalnem prostoru, zato v protidejstveni analizi vzročnosti, kadar sta *A* in *B* resnični propoziciji, lahko poskusimo tako, da uvedemo drugo potezo. Če bi *A* bil resničen, ali bi bilo tedaj res, da če *A* ne bi bil resničen, tedaj tudi *B* ne bi bil? Tako dobimo »vgnezdeni« pogojnik: »$A > (\neg A > \neg B)$«. Postopek lahko ponavljamo in zastavljamo bolj zapletena vprašanja. Vzročnost v tem primeru modeliramo z bolj zapletenimi vzorci protidejstvenikov. V iskanju pravega vzročnega pogojnika je nekaj predlogov, ki gradijo na takšnih močnejših, vgnezdenih vzorcih odvisnosti. Mislim pa, da ne prinašajo nič novega v protidejstveno analizo vzročnosti, še vedno gre za odvisnost, ki ji manjka ideja zadostnosti in izvora.

Rasmussen (1982) je, recimo, poskušal zagovarjati idejo, da protidejstvena analiza zajame tudi idejo zadostnosti tako, da je oporekal standardni formulaciji nujnega in zadostnega pogoja:

C je zadostni pogoj za *E*, če in samo če $C > E$.

C je nujni pogoj za *E*, če in samo če $\neg C > \neg E$.

Rasmussen trdi, da idejo vzročne zadostnosti dogodka *C* za dogodek *E*, kadar sta oba dejansko nastopila, izrazi nek sestavljeni pogoj:

C je zadostni pogoj za *E*, če in samo če $C > E$ in $(\neg C > (C > E))$

Kadar sta *C* in *E* dva dejanska dogodka, takrat po Rasmussenu resnica protidejstvenika $\neg C > \neg E$ implicira resnico vgnezdenega protidejstvenika $(\neg C > (C > E))$. Če ne bi nastopil dogodek vzrok, tedaj, če bi nastopil dogodek vzrok, potem bi nastopil tudi učinek (Rasmussen 1982: 209). Toda njegova izpeljava je zmotna:

$1\ \neg C > \neg E$	// hipoteza
$2\ \neg C > \neg C$	// teza v logiki protidejstvenikov
$3\ \neg C > (\neg C \ \& \ \neg E)$	// iz 1 in 2
$4\ \neg C > (\neg C > \neg E)$	// CS iz 3
$5\ \neg C > (\neg E > \neg C)$	// CS iz 3
$6\ \neg C > (C > E)$	// kontrapozicija iz 5 (?)

Kritičen je prehod od vrstice (5) do (6). Rasmussen pravi, da je mogoče ob domnevi, da *C* ne nastopi in protidejstveni ekvivalenci $\neg C$ in $\neg E$ ($\neg C > \neg E$ in $\neg E > \neg C$), uporabiti kontrapozicijo na konsekvensu pete vrstice. Ampak ta korak je zmoten, saj bi sicer, ob domnevi CS, lahko izpeljali vzročno odvisnost med poljubnima dvema dogodkoma (recimo *D* in *F*), ki sta dejansko nastopila. Recimo:

$1D$	// hipoteza
$2F$	// hipoteza
$3\ D > F$	// CS iz 1 in 2
$4\ F > D$	// CS iz 1 in 2
$5\ \neg F > \neg D$	// kontrapozicija zaradi 3 in 4

Rasmussen sicer trdi, da je mogoče izpeljati kontrapozicijo dveh protidejstveno ekvivalentnih propozicij ($\neg C$, $\neg E$) ob hipotetični domnevi, da *C* ne nastopi. Tega pa ne utemelji in nobeno splošno sprejeto načelo v logiki protidejstvenikov ne dopušča izpeljave kontrapozicije protidejstveno ekvivalentnih propozicij. Recimo, da je nastopil dogodek *C* (vklop stikala), ki je neodvisen od dogodkov *D* (Zemlja se vrti okrog Sonca) in *F* (pes zalaja). Oglejmo si izpeljavo:

$1\ C$	// hipoteza
$2\ \neg C > D$	// hipoteza
$3\ \neg C > F$	// hipoteza
$4\ \neg C > (D \ \& \ F)$	// 2, 3 logika pogojnikov
$5\ \neg C > (D > F)$	// CS iz 4
$6\ \neg C > (F > D)$	// CS iz 4
$7\ \neg C > (\neg F > \neg D)$	// kontrapozicija iz 6 (?)

Sklep (7) pravi: »Če ne bi vklopil stikala, tedaj, če pes ne bi zalajal, potem se Zemlja ne bi vrtela okrog Sonca«. Rezultat je absurden, temelji pa na resničnih pogojnikih (2): »Četudi ne bi vklopil stikala, bi se Zemlja še vedno vrtela okrog Sonca« in (3) »Četudi ne bi vklopil stikala, bi pes še vedno zalajal«. Mislim, da je izpeljava teze $(\neg C > (C > E))$ iz $\neg C > \neg E$ napačna.

Toda, če zadostnosti ne moremo izpeljati, jo morda lahko enostavno vgradimo v analizo in okrepimo protidejstveno odvisnost. Kadar sta C in E dogodka, ki sta dejansko nastopila, takrat je dogodek E po Rasmussenu vzročno odvisen od dogodka C, če in samo če: $\neg C > \neg E$ in $\neg C > (C > E)$. Ker je v logiki protidejstvenikov vse tisto, kar je resnično ob dani hipotezi, resnično konjunktivno, dobimo s tem naslednji pogoj odvisnosti: $\neg C > (\neg E \ \& \ (C > E)$. S tem dobimo neko »finejšo« odvisnost, ki se morda bolj približa pogojniškemu modeliranju vzročne odvisnosti. Na ta način skuša protidejstveno analizo vzročnosti okrepiti Kadri Vihvelin (1991). Primerjajmo pogoja:

KPO $(A > B) \ \& \ (\neg A > \neg B)$

DPO $(A > B) \ \& \ \{(\neg A > [\neg B \ \ \& \ (A > B)]\}$

V nekaterih primerih imamo protidejstveno odvisnost dveh dogodkov (KPO), vendar ne gre za pravo vzročno odvisnost. DPO (»dvojna protidejstvena odvisnost«) naj bi po Vihvelinovi izločila pravi vzročni pogoj. Zlahka ugotovimo, da kadar nek sistem pogojev izpolni shemo DPO, takrat izpolni tudi shemo KPO, obratno pa ne velja. Mary je v komi in ne odgovarja na moja vprašanja. Če bi odgovorila, ne bi bila v komi, njen odgovor in njena koma sta v odnosu KPO. Toda, če bi odgovorila, ali bi potem bilo res, da če ne bi odgovorila, tedaj ne bi bila v komi? Ne, če ne bi odgovorila, potem bi to lahko bilo zato, ker morda ne bi slišala mojega vprašanja, ker bi mislila na kaj drugega ali kaj podobnega (Vihvelin, 1991: 176). Njen odgovor in njena koma nista v odnosu DPO. Lahko bi rekli, da je odsotnost kome *nujni* pogoj za njen odgovor, ni pa *zadostni*. Recimo še, da avto v normalnem voznem stanju vozi po cesti s hitrostjo 80 km na uro. Potem je, v danih normalnih okoliščinah, resnično po KPO:

> Če avto ne bi bil vžgan, potem ne bi vozil s hitrostjo 80 km na uro.

Toda, kaj reči o DPO:

> Če avto ne bi bil vžgan, potem (ne bi vozil s hitrostjo 80 km na uro in če bi bil vžgan, tedaj bi vozil s hitrostjo 80 km na uro).

Ta protidejstvenik je za Vihvelinovo napačen, kajti v situaciji, ko avto ni vžgan in ne vozi s hitrostjo 80 km na uro, ni resnično, da če bi bil vžgan, tedaj bi vozil s to hitrostjo. Lahko bi seveda vozil z drugačno hitrostjo ali pa sploh miroval.

Primerjajmo to z znanim primerom: dobro narejena vžigalica v okolju, bogatem s kisikom zagori, ko je uprasnjena. Potem je res KPO:

> Če vžigalica ne bi bila uprasnjena, tedaj ne bi zagorela.

Ampak res je tudi DPO:

> Če vžigalica ne bi bila uprasnjena, potem (ne bi zagorela in če bi bila uprasnjena, tedaj bi vžigalica zagorela).

Privlačno pojasnilo razlike je v tem, da je vžig avtomobila samo vzročno nujni pogoj za vožnjo s hitrostjo 80 km na uro, ne pa tudi vzročno zadostni pogoj, samo z »manipuliranjem« vžiga ne manipuliramo z vožnjo z dano hitrostjo. Toda uprasnitev vžigalice je več kot to, je tudi vzročno zadostna za gorenje: uprasnitev vžigalice je tista, ki privede do gorenja vžigalice.

Seveda pa je shema DPO izpolnjena tudi v primeru, ko v najbližjih svetovih »manipuliramo« samo z, recimo, navzočnostjo kisika, ki je enkrat navzoč pri uprasnitvi vžigalice, drugič pa ne. Torej je navzočnost kisika v enaki meri vzročno zadostni pogoj za gorenje vžigalice kot uprasnitev vžigalice! Kajti res je:

> Če v okolici ne bi bilo kisika, tedaj vžigalica ne bi zagorela.

Ampak zagovarjamo lahko tudi resnico DPO:

> Če v okolici ne bi bilo kisika, potem (vžigalica ne bi zagorela in če bi bilo v okolici dovolj kisika, tedaj bi vžigalica zagorela).

Če dobro narejena vžigalica v okolju, bogatem s kisikom zagori, ker je uprasnjena, potem bi brez kisika res ne zagorela, toda še vedno bi bila uprasnjena. In če bi zdaj v tej hipotetični situaciji (ni kisika, vžigalica je uprasnjena in ne zagori) bilo na voljo dovolj kisika, tedaj bi vžigalica zagorela, saj domnevno menjamo le paramater »kisik« (vžigalica je še vedno uprasnjena ipd.). Res pa je, da bi se nekateri (tudi Lewis) strinjali, da kisik dejansko *je* vzrok v enaki meri kot uprasnitev, le da vsakdanja pragmatika zahteva, da kot vzrok izberemo najbolj relevantni in izstopajoči dejavnik v določeni situaciji.

Kaj reči o primeru z avtomobilom? Najprej sploh ni jasno, ali gre za dva neodvisna dogodka, kar je prvi pogoj za vzročno zvezo. A v nekem ohlapnem smislu, ki ga tudi Lewis dopušča, morda lahko govorimo o stanju vžiga in stanju vožnje s hitrostjo 80 km na uro. Bolj problematična je nejasnost v izbiri najbližjih možnih svetov. Spomnimo se: avto v normalnem voznem stanju vozi po cesti s hitrostjo 80 km na uro. Kaj bi pomenilo, če ne bi bil vžgan? Ali bi potem bil v garaži ali na parkirišču? Ali pa bi nekaj zmotilo električni krogotok

vozečega avtomobila in bi se avto na lepem ustavil? Če prvo, kar se zdi intenca Vihvelinove (primer s komo), tedaj gre bolj za problem sporne odvisnosti preteklosti od sedanjosti kot pa za vprašanje vzročne zadostnosti. Če pa drugo, tedaj ni tako jasno, ali je shema DPO zares neveljavna: če bi nekaj zmotilo električni krogotok vozečega avtomobila in bi se avto na lepem ustavil, potem, če te prekinitve ne bi bilo, bi spet vozil normalno, s takšno hitrostjo kot prej.

Mislim, da je posebnost DPO in vgnezdenih pogojnikov najprej v tem, da z vgnezdenjem pogojev nakaže idejo vzroka kot pogoja, s katerim najlaže »manipuliramo« in s tem delujemo na dogodek učinek. Drugo posebnost najbolje opazimo v primeru, ko so *A*, *B* in *C* dejansko resnične propozicije, tako da je *C* vzročno odvisen od *A*, *B* pa ni v nikakršni zvezi ne z *A*, ne z *C*. V tem primeru je lahko veljavna krepka protidejstvena odvisnost med disjunkcijo »(*A* Ú *B*)« ter *C*, ne pa tudi močnejša dvojna odvisnost DPO. Naj bo dejansko resnično naslednje: *A* = vžigalica je uprasnjena; *B* = v trenutku *t* se na Marsu zakotali kamen; *C* = vžigalica zagori. Potem se lahko prepričamo, da je res KPO:

$$[(A \vee B) > C)] \;\&\; [\neg(A \vee B) > \neg C)]$$

Ni pa res DPO:

$$[(A \vee B) > C)] \;\&\; \{\neg(A \vee B) > [\neg C \;\&\; ((A \vee B) > C))]\}$$

Vgnezdeni pogojnik v shemi DPO tako deluje kot nek filter, kot procedura, s katero preverjamo »čistost« pogojenosti. Torej je tista posebnost pogoja DPO, ki ga razlikuje od KPO v tem, da izloči prave, »atomarne« vzroke od čudnih, sestavljenih pogojev, ki lahko vsebujejo tudi povsem irelevantne informacije. Toda vsaka teorija vzročnosti potrebuje teorijo o tem, kaj so relata v vzročni relaciji. Razumna teorija dogodkov bo odstranila konjunktivne in disjunktivne dogodke. To je Lewisova rešitev, ki se mi zdi v protidejstveni analizi vzročnosti bolj perspektivna kot pa raba zapletenih vgnezdenih shem pogojnikov.

8.

Tudi dvojna, trojna ali še bolj zapletena protidejstvena odvisnost pa je še vedno samo *odvisnost*, ki razglaša kot vzročne zelo nenavadne vzorce. Zanimivi so vzorci dvojne preprečitve (prim. Hall, 2004: 241). Suzi pilotira bombnik na poti do cilja operacije, Billi pa je zdaj pilot zaščitnega lovskega letala, ki sestreli sovražno prestrezno lovsko letalo, ki bi sicer uničilo Suzijin bombnik. Suzi uspešno izpelje nalogo in natančno odvrže bombe. Brez Billijeve sestrelitve ne bi bilo uspešnega bombardiranje cilja. V tem primeru en dogodek (Billijeva sestrelitev sovražnika) prepreči drug dogodek (sovražni lovec uniči Suzi), ki bi preprečil tretji dogodek (bombardiranje cilja). Billijeva sestrelitev je vzrok bombardiranja, čeprav se morda spopad med obema lovcema dogaja stotine

kilometrov stran od Suzi, ki se tega sploh ne zaveda, spopad pa nima nobenega vpliva na njeno delovanje. Lahko se prepričamo, da velja:

> Če Billi ne bi sestrelil sovražnega lovca, potem (ne bi bilo uspešnega bombardiranja cilja in če bi sestrelil sovražnika, potem bi bilo bombardiranje uspešno).

Seveda pa zelo težko rečemo, da uspešno bombardiranje »izvira« iz Billijeve sestrelitve, saj sploh ni nobene fizikalne zveze med dogodki v eni (spopad lovcev) in drugi regiji (bombardiranje). Podobnih primerov je več, poseben problem je vzročnost odsotnosti, s katero se ukvarja tudi Lewis. Odsotnost hrane povzroči lakoto in cepljenje prepreči, da bi zboleli (naredi »odsotnost bolezni«). Protidejstvena odvisnost razvršča takšne vzorce med vzročne, čeprav smo daleč od ideje izhajanja ali vzroka kot izvora.

Hall (2004) po mojem mnenju upravičeno opozarja, da obstaja nepomirljiva napetost med dvema pojmovanjema vzročnosti. Po eni strani je vzrok »proizvajalec« – tisto, kar privede ali vodi do učinka, učinki so prek določenih procesov *povezani* s svojimi vzroki. Po drugi strani je vzrok »ustvarjalec razlike«, učinek je odvisen od vzroka, brez vzroka ni učinka. Po Hallovem mnenju je vsak poskus enotnega pojasnila obsojen na neuspeh, zato mora filozof sprejeti dva različna pojma in dve različni analizi vzročnosti. Vzročnost kot odvisnost je kar protidejstvena odvisnost, za vzročnost kot »produkcijo« pa Hall predlaga izpiljeno različico stare nomološke teorije – vzrok je element množice pogojev, ki so, glede na zakone narave, skupaj zadostni za učinek.

Lewis takšne dvojnosti ne sprejema, vzročnost je zanj ena sama stvar, skupna mnogim vzročnim mehanizmom. No, idejo vzroka kot *conditio sine qua non* zdaj analizira kot idejo *vplivanja* in ali ni v tem vendarle nek poskus združitve obeh polov pojma vzroka? Pri novih igralnih konzolah lahko zgolj s premiki rok v prostoru nekaj metrov pred TV aparatom upravljamo z videoigro, ki teče na ekranu. Premik desne roke in že se na ekranu dvigne palica za golf, poravnava z levo roko in palica se malo obrne, zamah z roko in palica udari žogico. To je, približno, Lewisova nova prispodoba za vzročnost kot vpliv: »premiki« na strani vzroka – kasneje, drugje, drugače ali sploh nič – vodijo do »premikov na ekranu«, na strani učinka – kasneje, drugje, drugače ali sploh nič. Vzrok *vpliva* (in torej na nek način *deluje*, čeprav včasih »na daljavo«) na učinek, kadar je učinek na ustrezen način *odvisen* od vzroka.

Lewis on Causation

Lewis's classical counterfactual theory of causation is presented. According to the simple version C causes E if E depends on C and E depends on C iff, had C not occurred, E would not have occurred. Difficulties concerning preemption have proven to be the biggest bugbear for Lewis's early theory which led to revision. According to late Lewis E depends on C iff, had the manner of C's occurrence differed in any of various ways, the manner of E's occurrence would have correspondingly differed. Counterfactual theory is based on the idea of dependency and the idea that the cause makes the difference; it lacks the idea of the cause as the source or the producer of its effect. Two proposals for a stronger conditional, which would improve counterfactual analysis in the respect of adding causal sufficiency, are criticized. Some have argued (Ned Hall) that the best thing for a philosopher to do is to break the concept of causation into two distinct concepts. We can view Lewis's late analysis of causation as influence as an attempt to provide a univocal analysis of our ordinary notion of causation.

Keywords: David Lewis, counterfactual dependence, causality, influence, two concepts of causation.

Literatura

Anscombe, G. E. (2003). »Vzročnost in determinacija«. *Analiza* 4:7, str. 21-36.

Bennett, J. (2003). *A Philosophical Guide to Conditionals*, Oxford: Oxford University Press.

Bunzl, M. (1979). »Causal Overdetermination«. *The Journal of Philosophy* 76 (1979), str. 134-50.

Collins, Hall in Paul (ur.) (2004). *Counterfactuals and Causation*. Cambridge MA: The MIT Press.

Hall, N. (2004). »Two Concepts of Causation«. V Collins, Hall in Paul (ur.), str. 225-276.

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

Kim, J. (1973). »Causes and Counterfactuals« .*The Journal of Philosophy* 70 (1973), str. 570-72.

Lebow, Tetlock in Parker (ur.) (2004). *Unmaking the West: »What-If?« Scenarios That Rewrite World History*. Ann Arbor: The University of Michigan Press.

Lewis, D. (1973a). *Counterfactuals*. Cambridge Mass.: Harvard University Press.

Lewis, D. (1973b). »Causation«. V *The Journal of Philosophy* 70 (1973), str. 556-567. Slovenski prevod: Lewis, D. (1993). »Vzročnost«. *Anthropos*, št. 3-4 (1993), str. 284-293.

Lewis, D. (1983). *Philosophical Papers Vol. I*. Oxford: Oxford University Press.

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

Lewis, D. (2000). »Causation as Influence«. *The Journal of Philosophy* 97 (2000), str. 182-97.

Lewis, D. (2004). »Causation as Influence«. V Collins, Hall in Paul (ur.), str. 75-106.

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

Rasmussen, S. (1982). »Ruben on Lewis and Causal Sufficency«. *Analysis* 42, str. 207-211.

Stalnaker, R. (1968). »A Theory of Conditionals«. V Rescher, N. (ur.), *Studies in Logical Theory*. Oxford: Blackwell.

Swain, M. (1980). »Causation and Distinct Events«. V van Inwagen, P. (ur.), *Time and Cause*. Dordrecht: D. Reidel.

Šuster, D. (1995). *Sodobna teorija modalnosti – protidejstvena odvisnost in vzročnost*. Doktorsko delo. Ljubljana: Filozofska fakulteta.

Šuster, D. (2002). »Lewisova teorija protidejstvenikov«. *Analiza* 6:1/2, str. 95-108.

Šuster, D. (2003). »Vzročnost«. *Analiza* 7:4, str. 37-53.

Vihvelin, K. (1991). »Freedom, Causation and Counterfactuals«. *Philosophical Studies* 64, str. 161-184.

Yagisawa T. (1979): »Counterfactual Analysis of Causation and Kim's Examples«. *Analysis* 39, str. 85-95.