ANαliZA

časopis za kritično misel

Letnik 5, številka 4, maj 2002 · ISSN 1408-2969

RAZPRAVE

Damjan Bojadžiev

Negacija in avtoreferenca v paradoksih in izrekih

Institut “Jožef Stefan”, Ljubljana

Citiranje: Damjan Bojadžiev, »Negacija in avtoreferenca v paradoksih in izrekih«, *Analiza: časopis za kritično misel* 5, št. 4 (2001): 24–32.

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.

Damjan Bojadžiev
Institut "Jožef Stefan", Ljubljana

NEGACIJA IN AVTOREFERENCA V PARADOKSIH IN IZREKIH

Avtoreferenca pogosto pelje v protislovje oziroma v paradoks ali vsaj v situacije, ki so videti precej paradoksalne. Razmerja med sestavinami, ki peljejo v takšne situacije (avtoreferenca, negacija), so tu razčlenjena na primeru Gödlovih in sorodnih izrekov. Prikazan je predvsem način, kako se ohranja konsistenca skozi različne oblike in stopnje samozanikanja.

Ključne besede: konsistenca, omega-konsistenca, Gödel, Rosser, Löb.

1. Uvod

Veliko, če ne večina paradoksov, vsebuje neko obliko avtoreference: formulacija paradoksa ali sklepanje, ki izpostavi paradoksalni učinek, se na nek način nanaša nase. Stavek (trditev ali definicija), del stavka ali zaporedje stavkov (argumentacija) lahko skozi nanašanje nase ali na okoliščine lastne formulacije oziroma izjavljanja proizvede paradoksalni učinek – bolj ali manj skrito protislovje, za katerega je videti, da se mu ni mogoče izogniti. Od tod izvira tradicionalno negativna logična podoba avtoreference kot nekakšne postranske, izumetničene oblike reference, ki je vsaj dvomljiva, če že ne pelje nujno v paradoks, zato jo je v spodobnem logičnem jeziku mogoče ali celo priporočljivo prepovedati. Tako nekako bi lahko, z določeno mero karikiranja, povzeli Russellovo stališče, ki je bilo v prvi polovici dvajsetega stoletja še najbolj znan poskus "zatrtja" avtoreference z izgovorom, da bi tako zatrli protislovja oziroma paradokse.

Avtoreferenca ima sicer pogosto negativne učinke v tem smislu, da pripelje do limitativnih rezultatov, ki izpostavljajo meje kakšne spoznavne ali drugačne dejavnosti. Hofstadter celo pravi, da so bili vsi rezultati, ki so bistveno odvisni od zlitja [*fusion*] subjekta in objekta, limitativni (Hofstadter, 1979: 699), pri čemer omenja rezultate matematične logike (Gödel, Church) in fizike (Heisenberg). Toda meje, ki jih odkrivajo ti rezultati, pogosto izražajo nekaj negativnega, neko stvarno in resnično omejitev, le na prvi pogled: če je namreč preko take meje mogoče priti le za ceno protislovja, potem ta meja niti ni prava meja, ker dejansko ničesar ne omejuje ali preprečuje. Osnovni primer take meje in njenega napačnega razumevanja je nedokazljivost stavka, ki trdi ravno to, da je namreč nedokazljiv. Če je logični sistem konsistenten, tak stavek dejansko ni dokazljiv, kar je dovolj blizu

zdravi pameti: če stavek pravi, da ni dokazljiv v določenem sistemu sklepanja, pa bi ga vseeno bilo mogoče dokazati v njem, bi padli v protislovje (med tem, kar stavek trdi, in dejanskim stanjem). V konsistentnem sistemu sklepanja torej takega stavka ni mogoče dokazati. Na tej točki pa se začne bolj problematični del premisleka: če namreč stavek pravi, da ni dokazljiv, in (v konsistentnem sistemu) dejansko ni dokazljiv, potem je stavek resničen in imamo nedokazljivo resnico. Iz tega izvira želja, da bi tudi to resnico dokazali, in vtis, da sistemu, kjer to ni mogoče, nekaj manjka, da je z njim nekaj narobe. Penrose oziroma popularna recepcija njegovih izpeljav je ta presežek resničnosti nad dokazljivostjo celo povzdignila v nekakšno logično optiko, ki naj bi odlikovala človeški um nasproti spoznavnim zmožnostim formalnega sistema sklepanja ali računalnika: mi *vidimo*, da je stavek resničen, formalni sistem pa tega ne more (Penrose, 1989). Seveda pa tu ne gre za kakšno dobesedno videnje, temveč za razumevanje ali sklepanje, ki se ne odvija le v formalnem sistemu samem. Odločilno pa je to, da je nasproti takšnim trditvam treba pač ponoviti že opravljeni premislek: če bi stavek uspeli dokazati v danem sistemu sklepanja, bi padli v protislovje (dokazali bi stavek, ki trdi, da *ni* dokazljiv). To, da stavka ne moremo dokazati, če delamo v konsistentnem sistemu sklepanja, ni kakšna omejitev tega sistema, temveč preprosto varovalka oziroma varnostni signal, ki kaže, da (še) delamo v takšnem, namreč konsistentnem sistemu. Protisloven sistem sklepanja pa običajno ni kaj prida koristen, ker lahko v njem tradicionalno dokažemo karkoli.

Paradoksalni vtis nedosegljive, nedokazljive resnice pa je tudi treba nekoliko ublažiti: ta nedokazljiva resnica namreč ni (splošno) veljavna [*valid*], ker je resničnost stavka odvisna od interpretacije formalnega sistema. Poenostavljeno rečeno, obstajajo interpretacije, v katerih nedokazljivi stavek ni resničen, ker trdi preveč: ne le, da ni dokazljiv v sistemu, v katerem je formuliran, temveč da ni dokazljiv niti v močnejših sistemih: več o tem v Bojadžiev (1999: 46).

2. Paradoksi in Gödlovi izreki

Ugotovitev, da v konsistentnem logičnem sistemu ni mogoče dokazati stavka, ki trdi lastno nedokazljivost, povzema (prvi) Gödlov izrek o nepopolnosti (1931). Ta izrek hkrati kaže, kako produktiven je lahko paradoks, in avtoreferenca v njem, v razvoju formalnih sistemov logike. Stavek, ki trdi lastno nedokazljivost, je namreč prevod paradoksalnega stavka lažnivca, ki trdi lastno neresničnost, v dokaznoteoretični idiom. To zvezo, in analogijo z Richardovim paradoksom (najmanjšega števila, ki ga ni mogoče definirati z manj kot N besedami), je izpostavil tudi Gödel v uvodnem prikazu svojih ugotovitev. Dodatno zvezo Gödlovih rezultatov s paradoksom lažnivca pa lahko najdemo v drugem Gödlovem izreku, ki izhaja iz (dokaza) prvega izreka: drugi izrek o nepopolnosti pravi, da v konsistentnem formalnem sistemu ni mogoče dokazati ravno njegove konsistence, točneje rečeno, stavka, ki trdi njegovo konsistenco. Temu bi lahko, po analogiji s paradoksom lažnivca, rekli paradoks dokazovalca (konsistence): dokler namreč dokaz konsistence *ne* uspeva, je to dober znak, da je sistem dejansko konsistenten, če pa dokaz konsistence uspe, se to lahko zgodi le zato, ker sistem dejansko *ni* konsistenten. Lahko pa premislimo tudi takole: recimo, da bi

izrek trdil obratno, da je v konsistentnem sistemu mogoče dokazati trditev njegove konsistence. Toda trditev konsistence je dokazljiva tudi v nekonsistentnem sistemu, torej nam njen dokaz dejansko ne bi povedal ničesar o konsistenci sistema (konsistenten je ravno sistem, v katerem nečesa *ni* mogoče dokazati). Če pa dvomimo v konsistenco sistema, potem nam dokaz trditve konsistence ne bi nič pomagal, tako kot nam, če dvomimo v resničnost besed nekoga, ne pomaga, če ta nekdo reče, da govori resnico. Pri obeh izrekih o nepopolnosti gre torej le za navidezne omejitve sposobnosti dokazovanja, ki dejansko ne izpostavljajo nečesa, kar bi morali biti sposobni dokazati, pa žal ne moremo, temveč le afirmirajo temelj vsakega dokazovanja, namreč konsistenco logičnega sistema. Pri drugem Gödlovem izreku pa se da dodati še tale premislek, povzet po Franzénu (v tisku): dejstvo, da v konsistentnem logičnem sistemu ni mogoče dokazati trditve njegove konsistence, ne izraža kakšne omejitve zmožnosti dokazovanja, ker trditve konsistence ni mogoče niti konsistentno postulirati. Zato običajna metafora, da ima konsistentna teorija nekakšno "slepo pego", zaradi katere ne more "(u)videti" lastne konsistence, ni povsem ustrezna. Analogija, ki jo predlaga Franzén, je nezmožnost, da bi resnično zatrdili [*truly affirm*], da nikoli ne govorimo o sebi.

Gödlova izreka hkrati kažeta pozitivno vlogo, teoretično pot paradoksov, ki se, po prevodu v formalno govorico matematične logike, prelevijo iz zabavnih intelektualnih ugank ali zadreg v generatorje znanja o formalnih sistemih. Ta zveza med paradoksi in napredovanjem vednosti pa ni niti redka niti pravzaprav presenetljiva: Quine poudarja, da so nekoč nauk, da zemlja kroži okoli sonca, imenovali Kopernikov paradoks, in sicer celo tisti, ki so ta nauk sprejemali (Quine, 1976: 9). Gödlova izreka o nepopolnosti pa pravzaprav izpostavljata, ne nekakšno zaprtost, utesnjenost ali "nepopolnost" logičnih sistemov v tem smislu, da bi jim nekaj določnega manjkalo (do popolnosti), temveč nasprotno, njihovo odprtost ali neizčrpnost, kot poudarja Franzén že v naslovu svoje knjige o Gödlovi nepopolnosti: *Inexhaustibility – A Non-Exhaustive Approach*. Če namreč nek stavek ni dokazljiv, potem je sistemu mogoče dodati negacijo tega stavka, ne da bi porušili konsistenco sistema; če pa ima sistem še nekoliko strožjo obliko konsistence, mu je namesto negacije nedokazljivega stavka mogoče konsistentno dodati ta stavek sam (Feferman, 1995). Toda novi sistem bo spet podložen obema Gödlovima izrekoma in tako naprej.

Stavek, ki zanika lastno dokazljivost, ponazarja še eno, verjetno bolj pomembno in striktno pozitivno obliko avtoreference: stavek ne le govori o sebi (da ni dokazljiv), temveč tudi to počne na avtoreferenčen način. Stavek se namreč naredi tako, da se določen postopek (diagonalizacijo) uporabi na njem samem oziroma na opisu tega postopka. Ta postopek pa je že sam po sebi avtoreferenčen, ker iz odprtega stavka oziroma formule, ki vsebuje neko spremenljivko, proizvede stavek, ki govori o tej formuli, tako da spremenljivko v formuli zamenja z imenom te formule same. Opis tega postopka se nato vstavi v neko formulo, nakar se postopek uporabi še na tej formuli sami. Takšna samoaplikacija, ko se postopek uporabi na opisu samega sebe, je sicer pogosta v

računalništvu in v formalizmih, ki opisujejo računske postopke (*lambda* račun). Več o samoaplikaciji v Bojadžiev (1999).

3. Paradoksi in neodločljivost

Zveza med paradoksi (lažnivca, Richardovim paradoksom) in Gödlovimi ugotovitvami ponazarja splošno tezo, da paradoksi po prevodu v jezik formalnega sklepanja vodijo v neodločljivost: obstoj stavkov, ki niso niti dokazljivi niti ovrgljivi. Gödlova izreka imata namreč tudi (še bolj) "negativno" plat: stavek, ki trdi lastno nedokazljivost in trditev konsistence nista niti ovrgljiva v konsistentnem sistemu, točneje rečeno, nista niti ovrgljiva v sistemu, ki ima nekoliko močnejšo obliko konsistence, o kateri bo še govora v nadaljevanju. Na prvi pogled sicer zveze med paradoksi in neodločljivostjo ni, ker je pri paradoksu hkrati dokazljiva neka trditev A in njeno nasprotje ¬A, pri neodločljivosti pa ni dokazljivo niti A niti ¬A. Toda klasično sta ti situaciji ekvivalentni:

$$\neg A \,\&\, \neg\neg A \leftrightarrow \neg A \,\&\, A$$

Gödlovi in nekateri drugi znani rezultati o neodločljivosti (Churchev izrek) so sicer odvisni od avtoreference, v splošnem pa ni jasno, ali neodločljivost vedno oziroma nujno zadeva avtoreferenco. Če pa avtoreferenca za neodločljivost ni nujna, potem lahko domnevamo, da ni nujna niti za paradoks. Primer neodločljivosti, za katero ni videti, da vsebuje kakšno obliko avtoreference, je Paris-Harringtonov izrek (1977), ki zadeva relacijo vsebovanosti v zaporedjih (zelo visokih) končnih dreves. Kljub temu pa Hofstadter domneva, da so vsi neodločljivi stavki dejansko odvisni od avtoreference, še več, da so odvisni od avtoreference ravno na gödlovski način, da namreč trdijo svojo nedokazljivost v nekem deduktivnem sistemu skozi kakšen sistem aritmetičnega kodiranja dokazov (Hofstadter, 1979: 708). Ta domneva pa je dovolj splošno formulirana in vsebuje dovolj nedoločene eksistenčne kvantifikacije (*nek* deduktiven sistem, *nek* sistem kodiranja), da je lahko konsistentna tudi z izreki tipa Paris-Harrington. V prid tej domnevi govori tudi dejstvo, da so izreki tipa Paris-Harrington ekvivalentni trditvi določene oblike konsistence za sistem, v katerem so neodločljivi (Feferman, 1994), trditev konsistence pa implicira stavek, ki trdi lastno nedokazljivost. Po drugi strani pa mora Hofstadterjeva domneva pokriti zelo veliko oziroma obsežno množico primerov: neodločljivost je zelo pogost pojav že v preprostem smislu, da je neodločljivih stavkov neskončno mnogo, v tehničnem smislu pa, da je množica teh stavkov gosta [*dense*] glede na vsako razumno topologijo (Calude *et al*, 1994).

4. Paradoksi in negacija

Če ostanemo pri stavkih, ki neposredno govorijo o sebi, velja pri omenjeni Russellovi in kasneje Wittgensteinovi prepovedi oziroma ugotovitvi nemožnosti avtoreference (*Traktat* 3.332) dodati, da paradoksalnih učinkov ne kaže pripisovati zgolj avtoreferenci stavka, temveč celoti tega, kar stavek pove o sebi. Logično eksplozivna je namrec šele mešanica treh sestavin:

(1) avtoreferenca,

(2) predikat (resničnost, dokazljivost),

(3) negacija.

Ključna je tu negacija, ker brez nje ne more priti do protislovja, brez protislovja pa tudi pravega paradoksa ni. Toda samoopisovanje ima lahko tudi brez negacije dovolj zanimive, "performativne" učinke, ki so lahko tudi paradoksalni, čeprav protislovje tedaj ni skrito v premisah samih. Najbolj zanimiv primer paradoksalnega pozitivnega samoopisovanja je verjetno Löbov paradoks, ki, tako kot stavek lažnivca, kombinira avtoreferenco s predikatom resničnosti, vendar brez negacije in ne na nivoju atomarnih, temveč na nivoju sestavljenih izjav. Če namreč v "stavek resnicoljubneža"

Ta stavek je resničen,

ki sam ni logično kdo ve kako zanimiv, vpeljemo implikacijo

Če je ta stavek resničen, potem *P*

(kjer se demonstrativ "ta stavek" nanaša na zgornji stavek v celoti), pridobimo nekoliko presenetljivo deduktivno moč: dokažemo lahko ta *P*, tj. karkoli! Sklepanje je sicer precej zavozlano, prikazano pa je, na primer, v Boolos in Jeffrey (1980: 186). Pri tem paradoksu sicer ne gre za notranje protislovje, ki je skrito v paradoksu samem, temveč za zunanje protislovje med siceršnjim stanjem stvari in sklepom *P*, ko je *P* neresničen stavek, na primer "Dedek Mraz obstaja". Bolj pomemben kot paradoks sam pa je verjetno Löbov *izrek*, ki je bil izvirno dokazan na podoben, vendar še nekoliko bolj zavozlan način. Izrek govori o posledicah t.i. principa refleksije, ki pravi, da iz dokazljivosti stavka sledi tudi stavek sam – tj., da je dokazljiv stavek tudi resničen. To sicer po definiciji velja v zanesljivem [*sound*] sistemu, v splošnem pa to ni značilnost predikata dokazljivosti. Löbov izrek pravi, da je stavek posledica lastnega principa refleksije, kar pomeni, da v konsistentnem sistemu princip refleksije ne more veljati za vse stavke, ker bi bil sicer sistem nekonsistenten, saj bi bil v njem dokazljiv vsak stavek. Löbov izrek je sicer ekvivalenten drugemu Gödlovemu izreku o nepopolnosti (Boolos, 1979: 11).

5. Variante samozanikanja

Za paradokse avtoreference je bistveno to, *kaj* avtoreferenčen stavek pravi o samem sebi. Löbov paradoks kaže, da negacija ni nujna sestavina paradoksa, obratno pa negacija oziroma samonegacija tudi ni zadostna za protislovje. Paradoks se pojavi le v situacijah ekstremnega samozanikanja, kakršna je, na primer, izjava lažnivca. Če pa ostanemo pri milejših variantah, kjer je predikat resničnosti zamenjan s predikatom dokazljivosti, dobimo zanimivo gradacijo samozanikanja v zvezi z Rosserjevo izboljšavo Gödlovega izreka oziroma njegovega drugega dela, ki pravi, da stavek, ki trdi lastno nedokazljivost, ni niti ovrgljiv, če ima sistem določeno močnejšo obliko konsistence, imenovano omega-konsistenca. To obliko konsistence je mogoče lažje opredeliti preko njenega nasprotja, omega-nekonsistence ali omega-protislovnosti, ki je nekakšno "skoraj"protislovje: sistem je omega-nekonsistenten,

če je v njem dokazljivo, da ima neko število določeno lastnost, hkrati pa je za vsako posamezno naravno število dokazljivo, da *nima* te lastnosti. Formalnega protislovja med tema dvema trditvama dejansko ni, ker formalno sklepanje lahko uporabi le končno število premis, tu pa imamo le "intuitivno" nasprotje med splošno trditvijo in neskončno množico partikularnih trditev, ki ji nasprotujejo. Če bi situacijo posplošili tako, da bi pozabili, da gre za naravna števila, bi dobili tole logično strukturo:

$\neg P(a), \neg P(b), \neg P(c)$ ...

in hkrati

$(\exists x)\, P(x)$,

kjer so *a*, *b*, *c* ... vsi predmeti, ki so poimenovani v sistemu. Lahko bi rekli, da gre tu za nekakšno "neskončno protislovje" med splošnim in (vsakim) posameznim, recimo med splošnim načelom in sodbo ali ravnanjem v vsakem posameznem primeru. Približen primer iz vsakdanjega življenja bi bila izjava: "Moše je neumen, Izak laže, Rebeka krade ... Ampak jaz že nisem antisemit!" Na nivoju formalne semantike (teorija modelov) se neskončno protislovje sicer razreši skozi obstoj modelov, ki vsebujejo še dodatne elemente, ki v sistemu samem niso poimenovani; v aritmetičnem primeru so to neskončna stevila oziroma serije neskončnih števil (Boolos in Jeffrey, 1980: 195).

Ugotovitev, da stavek, ki trdi lastno nedokazljivost, ni ovrgljiv v omega-konsistentnem sistemu, je bolj jasna v transponirani obliki: če je stavek ovrgljiv, je sistem omega-nekonsistenten. Tedaj je namreč dokazljiva njegova negacija, ki pravi, da stavek *je* dokazljiv, da torej ima *nek* dokaz, po drugi strani pa se da dokazati, da stavek nima nobenega konkretnega dokaza (za vsako posamezno število se da dokazati, da *ni* število dokaza stavka). Rosser je to ugotovitev poenostavil tako, da je skonstruiral drug avtoreferenčen stavek, ki ni ovrgljiv že, če je sistem enostavno konsistenten. To je naredil tako, da je primerno zakompliciral predikat dokazljivosti, ki ga stavek uporablja: strukturo omega-konsistence je tako rekoč prenesel iz premise ovrgljivosti v sklep tako, da jo je vgradil vanj v zanikani, "sprevrnjeni" obliki, po zgledu prehoda iz

$(A \,\&\, B) \rightarrow C$

v

$A \rightarrow (C \vee \neg B)$

(kjer *A* ustreza premisi konsistence, *B* pa omega delu premise omega-konsistence). Rosser je konstruiral stavek, ki tudi zanika lastno dokazljivost, vendar na šibkejši in manj neposreden način. Stavek ne govori le o (svojih) dokazih, temveč tudi o protidokazih (dokazih svoje negacije) in o njihovi dolžini. Rosserjev stavek R pravi, da za vsak njegov dokaz obstaja protidokaz, ki ni daljši od tega dokaza. Stavek $\neg$R tedaj pravi, da obstaja dokaz stavka R, ki je krajši od vsakega protidokaza. Negaciji Gödlovega in Rosserjevega stavka tako obe afirmirata svoji lastni negaciji in se torej posredno zanikata. Toda stavek $\neg$R se močneje zanika kot $\neg$G, tako da je $\neg$R težje dokazati, kar pomeni, da je lažje dokazati, da $\neg$R ni dokazljiv. Kako to konkretno

deluje v formulah in zakaj enostavna konsistenca zadošča tudi za nedokazljivost R, je jasno iz dokaza Rosserjevega izreka (*glej*, na primer, Kleene, 1952: 208).

Poleg primerjave stopnje samozanikanja pri ¬G in ¬R lahko primerjamo tudi stopnjo samozanikanja pri R in ¬R: medtem ko R o sebi pravi, da ima vsak dokaz krajši protidokaz, pa ¬R o sebi pravi, da ima protidokaz, ki je krajši od vsakega dokaza. Če bi hoteli poimenovati to razliko v stopnji samozanikanja, bi lahko rekli, po analogiji s pojmoma zveznosti in enakomerne zveznosti, da gre pri R za navadno, pri ¬R pa za *enakomerno* samozanikanje. Če sedaj pregledamo, kaj obravnavani stavki pravijo o sebi, dobimo naslednjo gradacijo samozanikanja:

R, G, ¬G, ¬R

Čeprav torej samozanikanje prej ali slej pripelje v protislovje ali paradoks, če se dovolj stopnjuje, pa obstajajo sistematične in dosledne variante samozanikanja, ki so združljive s konsistenco. Zmerno interaktivna predstavitev teh variant oziroma ustreznih stopenj konsistence je postavljena na http://nl.ijs.si/~damjan/theorems.html.

6. Pragmatika samozanikanja

Paradoksi ali navidezno paradoksalni rezultati gödlovske sorte izkazujejo značilno trdovratnost, če jih skušamo nekako razrešiti ali eliminirati: na take poskuse radi reagirajo tako, da se obnovijo ali prenesejo na drugo raven. Če se, na primer, paradoksu lažnivca skušamo izogniti z opustitvijo načela bivalence, tako da bi poleg resničnosti in neresničnosti dopustili še tretjo možnost, paradoksalnost, se paradoks reproducira v obliki stavka *Ta stavek je neresničen ali paradoksalen* (okrepljeni lažnivec (Haack, 1978: 140)). Pri Gödlovih izrekih pa je reprodukcija paradoksalnosti še bolj sistematična: če namreč sistemu, ki mu "manjka" očitna, nedokazljiva resnica (na primer, da je konsistenten), to resnico dodamo, dobimo razširjen sistem, ki ima enako "napako". Tudi v novem sistemu bo veljalo, da bo trditev njegove konsistence ostala nedokazljiva in tako naprej v ordinalni nedogled. Vsak poskus, da bi odpravili "primanjkljaj" sistema, le reproducira ta primanjkljaj (Bojadžiev, 1999: 41). Mitološka paralela tej situaciji bi bila Tantalovo poseganje za hrano in vodo ali Sizifovo valjenje skale; ni čudno, da zgodovina paradoksov in avtoreference beleži tudi smrtne primere (Filetas) in primere večkratne hospitalizacije (Cantor, Gödel).

Paradoksalna avtoreferenca pa ima tudi bolj vsakdanje učinke: Watzlavick in soavtorja povezujejo Gödlov izrek s paradoksom nepričakovane tihe vaje [*unexpected examination*] v kontekstu komunikacijske teorije o dvojni vezavi [*double bind*] (Watzlavick *et al*, 1967). Ko učitelj napove nepričakovano tiho vajo v naslednjem tednu, se vzpostavi paradoksalna situacija, ki nekoliko spominja na nekakšno končno varianto omega-protislovja: vaja ne more biti konec tedna, ker tedaj ne bi bila več nepričakovana, torej ne more biti niti dan prej, torej ... pa vseeno bo enkrat naslednji teden (Hughes in Brecht, 1978: 35). Paradoksalna situacija te vrste, ki se ji ni mogoče izmakniti, je ugodna za

nastanek različnih patologij, po drugi strani pa ima lahko terapevtski učinek (tehnika predpisovanja simptoma, "nalaganja užitka").

V analogijah zgornje vrste natančnost paralele s "paradoksalnimi" izreki o avtoreferenci sicer pada, zato pa so primerjave lahko toliko bolj daljnosežne: tako Wilden pravi, da je Gödlov izrek edina teorija dovolj visokega logičnega tipa, da pojasni, zakaj v človeški komunikaciji ni mogoče dokončno prekoračiti [*ultimately transcend*] paradoksa (Wilden, 1980: 122). Podobno nenatančna, vendar nekoliko bolj prepričljiva je Cherniakova zgodba o paralizirajočem efektu nekakšne avtoreferenčne kognitivne strukture, primerljive z Gödlovim stavkom (Cherniak, 1978). Verjetno sicer res ni plavzibilno govoriti o "Gödlovem stavku človeka", kot ugotavlja (Webb, 1980), lahko pa bi, po zgledu Smullyanove primerjave med Löbovim izrekom in samouresničujočimi se [*self-fulfilling*] prepričanji (Smullyan, 1987: 123), primerjali Gödlov izrek in samofalsificirajoča se, "avtodestruktivna" prepričanja, ki spodnesejo same sebe. S primerno žlico soli pa bi lahko, kljub Webbovemu svarilu, to avtodestruktivnost vzeli tudi bolj dobesedno: Gödlov stavek za človeka bi lahko deloval kot nekakšen recept za avtodestrukcijo, ki se aktivira v procesu njegove artikulacije ali razumevanja; po analogiji s terminologijo performativov bi mu lahko rekli konvulziv, asfiksiativ, kombustiv ...

damjan.bojadziev@ijs.si
http://nl.ijs.si/~damjan/me.html

NEGATION AND SELF-REFERENCE IN PARADOXES AND THEOREMS

Self-reference frequently leads to contradiction, paradox or at least situations which appear rather paradoxical. Relations between the ingredients, which lead to such situations (self-reference, negation), are analyzed here for the case of Gödel's and related theorems. The emphasis is on the way in which consistency is preserved through various forms and levels of self-denial.

Keywords: consistency, omega-consistency, Gödel, Rosser, Löb.

Literatura

Bojadžiev, D. (1999). “Refleksivno širjenje refleksivnega sistema”. *Analiza*, 2-3, str. 39-47.

Boolos, G. (1979). *The Unprovability of Consistency*. Cambridge: Cambridge Univ. Press.

Boolos, G. in Jeffrey, R. (1980). *Computability and Logic*. Camridge: Cambridge Univ. Press.

Calude, C., Jürgensen, H. in Zimand, M. (1994). “Is Independence an Exception?”. *Applied Mathematics and Computation*, 66, št. 1, str. 63-76.

Cherniak, C. (1981). “The Riddle of the Universe and Its Solution”. V Hofstadter, D. in Dennett, D., *The Mind's I: Reflections on Self and Soul*, Harmondsworth: Penguin, str. 269-276.

Feferman, S. (1995). “Penrose’s Gödelian argument”. *Psyche*, 2 (7); HTML verzija na psyche.cs.monash.edu.au.

Feferman, S. (1994). “Deciding the undecidable”. V Feferman, S., *In the Light of Logic*, Oxford: Oxford Univ. Press.

Franzén, T. (v tisku). *Inexhaustibility – A Non-Exhaustive Approach*. Lecture Notes in Logic. Association for Symbolic Logic.

Haack, S. (1978). *Philosophy of Logics*. Cambridge: Cambridge Univ. Press.

Hofstadter, D. R. (1979). *Gödel, Escher, Bach – An Eternal Golden Braid*. New York: Basic Books.

Hughes, P., Brecht, G. (1978). *Vicious Circles and Infinity – An Anthology of Paradoxes*. Harmondsworth: Penguin.

Kleene, S. (1952). *Introduction to Metamathematics*. New Jersey: D. Van Nostrand.

Penrose, R. (1989). *The Emperor's New Mind*. Oxford: Oxford Univ. Press.

Quine, W. v. O. (1977). “The Ways of Paradox”. V Quine, W. v. O., *Ways of Paradox and Other Essays*, Cambridge, Mass.: Harvard Univ. Press.

Smullyan, R. (1987). *Forever Undecided – A Puzzle Guide to Gödel*. Oxford: Oxford Univ. Press.

Watzlawick, P., Beavin, J. in Jackson, D. (1967). *The Pragmatics of Human Communication*. W. W. Norton & Co.

Webb, J. (1980). *Mechanism, Mentalism and Metamathematics – An Essay on Finitism*. Dordrecht: D. Reidel Publ. Co.

Wilden, A. (1980). *System and Structure*. London: Tavistock Publ.