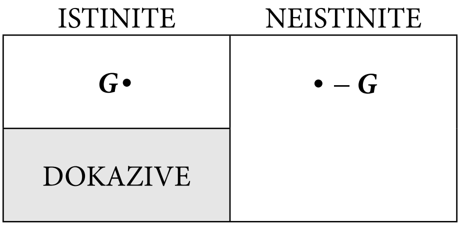
Gödelovi teoremi*

Zvonimir Šikić[1]

1. Što tvrde Gödelovi teoremi o nepotpunosti

Zamislimo skup svih propozicija koje su izrazive u nekoj deduktivnoj teoriji. Neke su od njih istinite, a one preostale su neistinite. Idealna deduktivna teorija bila bi ona u kojoj ne bi bila dokaziva nijedna neistinita propozicija, ali bi bile dokazive sve (u njoj izrazive) istinite propozicije. Gödelov 1. teorem o nepotpunosti tvrdi da nijedna dovoljno jaka deduktivna teorija nije u tom smislu idealna. (Deduktivne teorije koje obuhvaćaju elementarnu teoriju brojeva „dovoljno su jake".) Naime, iz Gödelovog 1. teorema slijedi da svaka korektna[2] i dovoljno jaka deduktivna teorija uvijek sadrži odgovarajuću propoziciju ***G*** (tzv. Gödelovu propoziciju) koja je istinita, ali u toj teoriji nije dokaziva. Naravno, negacija te propozicije –***G***, koja je neistinita, također nije dokaziva u toj teoriji (v. sl. 1.).



Slika 1.

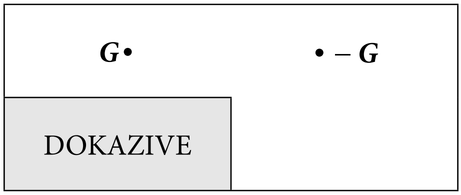
Gödel se u svome izvornom dokazu 1. teorema (Gödel, 1931.) zapravo i ne poziva na pojam istinitosti, nego dokazuje da svaka dovoljno jaka deduktivna teorija, koja je ω–konzistentna[3], uvijek sadrži odgovarajuću propoziciju ***G*** takvu da ni ona ni njezina negacija –***G*** nisu dokazive u toj teoriji (v. sl. 2.).

*Istoimeno predavanje u sklopu Stručno metodičkih večeri nastavne sekcije održano je 5. listopada 2011. i koncipirano na osnovu članka objavljenog u časopisu Rugjer (br. 5 (1996) pp. 30-35). Susretljivošću uredništva časopisa Rugjer prenosimo u cijelosti spomenuti članak.

[1]Zvonimir Šikić, Fakultet strojarstva i brodogradnje Sveučilišta u Zagrebu

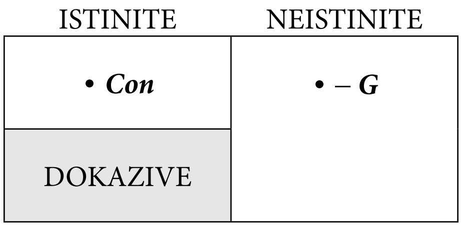
[2]Deduktivna teorija je korektna ako ne dokazuje neistine.

[3]Deduktivna teorija je konzistentna ako nijedna njezina propozicija nije ujedno dokaziva i opovrgljiva u toj teoriji. (Propozicija je opovrgljiva ako je njezina negacija dokaziva.) Dovoljno jaka deduktivna teorija je ω-inkonzistentna ako je u njoj dokazivo da neki prirodni broj ima neko svojstvo, iako je za svaki pojedini prirodni broj opovrgljivo da ima to svojstvo. Naravno, teorija je ω-konzistentna ako nije ω-inkonzistentna. Ako je teorija ω-konzistentna, onda je i konzistentna, ali je moguće da teorija bude konzistentna iako nije ω-konzistentna.



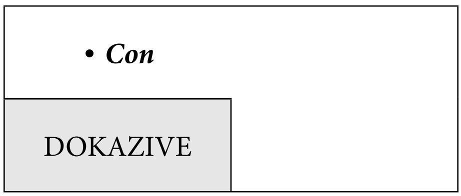
Slika 2.

Gödelov 2. teorem o nepotpunosti tvrdi da svaka korektna dovoljno jaka deduktivna teorija uvijek sadrži odgovarajuću propoziciju ***Con***, kojom se izražava konzistentnost deduktivne teorije u njoj samoj, te da ta propozicija, iako je istinita, u toj teoriji nije dokaziva (v. sl. 3.).



Slika 3.

Ni u dokazu svoga 2. teorema Gödel se ne poziva na pojam istinitosti, nego dokazuje (Gödel, 1931.)[4] da svaka dovoljno jaka deduktivna teorija, koja je ω–konzistentna, uvijek sadrži odgovarajuću propoziciju ***Con*** koja nije dokaziva u toj teoriji (v. sl. 4.).



Slika 4.

Formulacije teorema nepotpunosti koje spominju pojam istinitosti uveo je Tarski. Te su formulacije neposredne posljedice izvornih Gödelovih teorema i definicije pojma istinitosti Tarskog (usp. dalje). Sam Gödel nije se pozivao na pojam istinitosti jer je taj pojam izuzetno problematičan.

[4]U Gödelovom članku iz 1931. nalazi se samo skica dokaza 2. teorema. Teorem je trebao biti detaljno dokazan u drugom dijelu toga članka, koji Gödel nikada nije napisao. Prvi potpuni dokaz 2. teorema objavio je Bernays u (Hilbert-Bernays, 1939.).

2. Što je istina

Što znači da je neka propozicija istinita? To znači da jest tako kako ona tvrdi. Ovaj jednostavni odgovor nalazimo još kod Aristotela, a ponavljali su ga mnogi od Aristotela do Tarskog. (Taj odgovor je u temelju tzv. teorije korespondencije.) Primjerice, propozicija „Snijeg je bijel” istinita je ako i samo ako snijeg jest bijel.

„SNIJEG JE BIJEL” JE ISTINITA $\Leftrightarrow$ SNIJEG JE BIJEL

Istaknuta ekvivalencija je valjana definicija istinitosti propozicije „Snijeg je bijel”. Međutim, ono što tražimo nije definicija istinitosti pojedine propozicije, nego je to definicija općeg pojma istinitosti ***Ist***, koji će se moći valjano primijeniti na svaku propoziciju iz određenog skupa propozicija (npr. skupa propozicija koje su izrazive u nekoj deduktivnoj teoriji). To znači da iz definicije pojma ***Ist*** moraju slijediti ekvivalencije oblika

$$\textbf{\textit{Ist}}\ \text{‘}\boldsymbol{A}\text{’} \Leftrightarrow \boldsymbol{A},$$

za cijeli jedan skup propozicija ***A***. Tarski je definirao takav opći pojam ***Ist*** za skup aritmetičkih propozicija ***A***, ali u okviru jezika teorije skupova koji bitno proširuje jezik aritmetičkih propozicija (Tarski, 1931. i 1935.). To nije bilo slučajno.

Razmotrimo sljedeće propozicije koje izriče Krećanin Epimenid:

(K) ***Krećani uvijek lažu.***

(L) ***Ja sada lažem.***

Ako je propozicija K istinita, onda jest tako kako ona tvrdi, pa je zato neistinita. Dakle, propozicija K nije istinita. To smo dokazali bez ikakvog empirijskog istraživanja konkretnih propozicija koje izriču Krećani, što je paradoksalno. Naime, empirijska situacija u kojoj su sve ostale izjave Krećana doista lažne, odmah dovodi do paradoksa, jer tada iz neistinitosti propozicije K slijedi njezina istinitost. (Zamislimo li, osim toga, empirijsku situaciju u kojoj je Epimenid jedini Krećanin, a propozicija K jedina propozicija koju je on izrekao, dolazimo do paradoksa koji je iste vrste kao onaj koji generira propozicija L; usp. dalje.)

Paradoksalnost propozicije L još je očitija. Ako je ona istinita, onda jest tako kako ona tvrdi, pa je zato neistinita. Ako je ona neistinita, onda nije tako kako ona tvrdi, pa je zato istinita.

Russell, Tarski i mnogi drugi smatrali su da je problem ovakvih propozicija u samoreferiranju, tj. u tome da one govore o sebi samima. Primjerice, paradoksalnost propozicije L proizlazi iz činjenica da ona sama o sebi govori da nije istinita. Dakle, za nju vrijedi

$$\boldsymbol{L} \Leftrightarrow \textbf{-}\ \textbf{\textit{Ist}}\ \text{‘}\boldsymbol{L}\text{’}$$

Osim toga vrijedi i

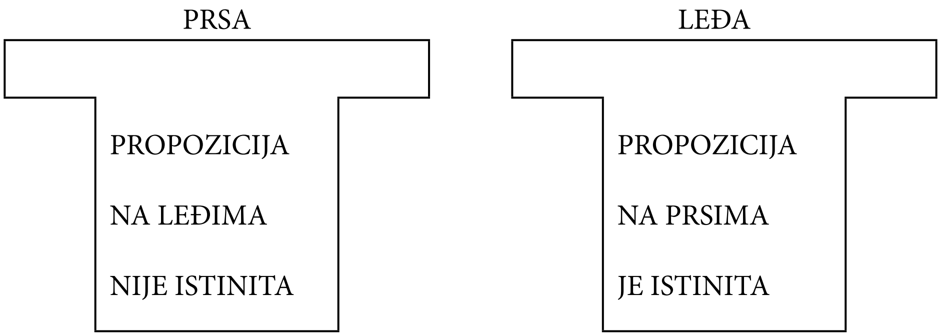
$$\boldsymbol{L} \Leftrightarrow \boldsymbol{Ist}\ \text{‘}\boldsymbol{L}\text{’}$$

jer to vrijedi za svaku propoziciju (usp. gore). No, uokvirene ekvivalencije evidentno su kontradiktorne.

Tarski je zato odbacio mogućnost definiranja općeg pojma istine za propozicije nekog jezika u samom tom jeziku. Istinitost propozicija objektnog jezika $\boldsymbol{J_0}$ definira se tek u proširenom metajeziku $\boldsymbol{J_1}$. Dakle, jezik $\boldsymbol{J_1}$ sadrži pojam $\boldsymbol{Ist_0}$ koji se odnosi (samo) na propozicije jezika $\boldsymbol{J_0}$. Ako želimo definirati pojam $\boldsymbol{Ist_1}$, koji se odnosi na propozicije metajezika $\boldsymbol{J_1}$, moramo prijeći u još obuhvatniji metametajezik $\boldsymbol{J_2}$ itd.

OBJEKTNI JEZIK	META JEZIK	METAMETA JEZIK	
J_0	J_1	J_2	. . .
	($\boldsymbol{Ist_0}$)	($\boldsymbol{Ist_1}$)	

U ovakvoj hijerarhiji jezika ne može doći do samoreferiranja kakvo nalazimo u propoziciji L, pa ni do složenijih „kružnih” samoreferiranja kakva nalazimo na T-majicama što ih prodaju neka američka sveučilišta:



Lako je utvrditi da su propozicije na prsima (P) i leđima (L), uzete zajedno, paradoksalne. Naime, to se neposredno vidi iz sljedećeg kruga evidentnih implikacija:

$$\begin{tikzcd}
& \boldsymbol{Ist}\ \text{‘}\boldsymbol{P}\text{’} \arrow[dr, Rightarrow] & \\
\boldsymbol{Ist}\ \text{‘}\boldsymbol{L}\text{’} \arrow[ur, Rightarrow] & & -\boldsymbol{Ist}\ \text{‘}\boldsymbol{L}\text{’} \arrow[dl, Rightarrow] \\
& -\boldsymbol{Ist}\ \text{‘}\boldsymbol{P}\text{’} \arrow[ul, Rightarrow] &
\end{tikzcd}$$

Primijetimo, međutim, da je teorija istine koju nam nudi Tarski prilično radikalna u svojim zabranama. Uobičajena uporaba pojma istinitosti u prirodnim jezicima gotovo nikada ne udovoljava njegovim zabranama. Primjerice, g. Jones ne bi smio izjaviti da većina Nixonovih izjava o Watergateu niti jest niti će ikada biti istinita jer se može dogoditi da će Nixon u jednoj od svojih izjava o Watergateu spomenuti i tu njegovu izjavu, čime se stvara „zli krug" samoreferiranja koji ruši hijerarhiju Tarskog. Svakodnevni govor o istini i laži „á la Tarski" gotovo je nemoguć. Stoga se čini da je odbacivanje svakog samoreferiranja i svakog kružnog samoreferiranja prestroga zabrana. Možda bi se mogao naći i neki blaži kriterij koji bi još uvijek isključivao paradoksalne propozicije.

Međutim, Kripke je jasno upozorio (Kripke, 1972.) da je prirodni govor o istini i laži gotovo uvijek u opasnosti da bude paradoksalan, te da to ne ovisi samo o strukturi propozicija koje izričemo, nego još češće o empirijskim situacijama u kojima ih izričemo. Razmotrimo, primjerice, sljedeće propozicije koje izriču g. Jones i predsjednik Nixon:

Jones: Većina Nixonovih izjava o Watergateu je neistinita.

Nixon: Sve Jonesove izjave o Watergateu su istinite.

Ako su sve Jonesove izjave o Watergateu istinite, te ako je omjer Nixonovih istina i laži o Watergateu točno pola-pola, onda su te dvije propozicije paradoksalne. (Situacija je tada analogna onoj na T-majicama.) U svakoj drugoj empirijskoj situaciji te su propozicije potpuno benigne. Ovakvi problemi pojam istinitosti i dalje čine teško razumljivim pojmom.[5]

Ipak, na razini pojedinih deduktivnih teorija hijerarhizirana teorija istine Alfreda Tarskog sasvim dobro funkcionira. On je primjerom pokazao kako se pojam aritmetičke istine, dakle istine u jeziku aritmetike J_0, može precizno definirati u obuhvatnijem jeziku teorije skupova J_1. No, da bismo bolje razumjeli daljnja razmatranja, potrebno je nešto detaljnije objasniti što je deduktivna teorija.

3. Što je deduktivna teorija

Svaka deduktivna teorija formulirana je u odgovarajućem jeziku. Na primjer, aritmetika (elementarna teorija brojeva) formulirana je u jeziku aritmetike.

Nelogički elementi jezika su:

(1) Imena objekata kojima se teorija bavi; u slučaju aritmetike to su brojevi 0, 1, 2, 3, itd.

[5]Od pojave Kripkeove teorije (Kripke, 1975.), koja je unijela sasvim novo svjetlo u razmišljanje o pojmu istine, ponuđene su još neke teorije. Po našem mišljenju, uz već spomenutu Kripkeovu, najuspješnije su Guptina (Gupta, 1982) i Barwise-Etchemendyjeva (Barwise & Etchemendy, 1987.).

(2) Imena operacija (unarnih, binarnih itd.) koje objekte prevode u objekte; u slučaju aritmetike to su npr. unarna operacija generiranja sljedbenika (ʻ), binarna operacija zbrajanja (+), binarna operacija množenja (·), itd.

(3) Predikati (jednomjesni, dvomjesni, itd.) kojima se izriču svojstva objekata i odnosi među objektima; u slučaju aritmetike to su npr. biti prost (***Pr***), biti identičan sa[6] (=), biti manje od (<), biti djelitelj od (|), itd.

Nelogički elementi jezika omogućuju izgradnju jednostavnih rečenica toga jezika. Takve su, na primjer, sljedeće rečenice jezika aritmetike:

$$\boldsymbol{Pr}\,(5 \cdot 4 + 3) \qquad 5 + 3 = 4 \cdot 2$$

$$7 \mid 13 \qquad 4 + 9 > 8 \cdot 6$$

(Primijetimo, usput, da prve dvije rečenice izriču istinite aritmetičke propozicije, dok druge dvije izriču neistine.)

Logički elementi jezika su:

(4) Negacija (–), te logički veznici „i" (&), „ili" (∨), „ako onda" (→), itd.

(5) Univerzalni kvantifikator „za svaki" (∀), partikularni kvantifikator „za neki" (∃), uz varijable (tj. zamjenice) ***x, y, z***, itd.

Logički elementi jezika omogućuju izgradnju složenih rečenica tog jezika. Takve su, na primjer, sljedeće rečenice jezika aritmetike:

$$-\boldsymbol{Pr}\,(3) \vee 3 < 0 \qquad \forall x\,(x < 0)$$

$$\forall x\, \exists y\,(x < y) \qquad \forall x\, \forall y\, \exists z\,(x \mid z \,\&\, y \mid z)$$

(Primijetimo, usput, da prve dvije rečenice izriču neistinite aritmetičke propozicije, dok druge dvije izriču istine.)[6a]

Naravno, jezik[7] deduktivne teorije još ne čini samu tu teoriju. Za nju su još potrebni aksiomi i logička pravila zaključivanja pomoću kojih se iz aksioma izvode (i tako dokazuju) teoremi te teorije.

Na primjer, neki od aksioma aritmetike su:

$$\forall x\, \exists y\,(x < y) \qquad -\exists x\,(x < 0) \qquad \text{itd.}$$

[6]Identitet zapravo spada u logičke elemente jezika, ali smo ga zbog jednostavnosti prikaza ovdje uvrstili u nelogičke elemente.
[6a]Lako je dokazati da se logički elementi mogu reducirati na samo četiri: –, &, ∀ i = . Naime, svi se ostali (uz odgovarajuću specifikaciju "svih ostalih") mogu definirati pomoću ta četiri. Nelogički elementi jezika aritmetike mogu se reducirati na samo dva: + i · .
[7]Tip jezika koji smo upravo opisali i u kojem je moguće formulirati svaku deduktivnu teoriju zove se jezikom 1. reda.

Neka od logičkih pravila zaključivanja su:

$$\frac{A \quad A\rightarrow B}{B} \qquad\qquad \frac{\forall x\, A(x)}{A(t)} \qquad \text{itd.}$$

Dakle, deduktivna teorija je potpuno određena kada su određeni njezin jezik, njezini aksiomi i njezina pravila zaključivanja. Sve one propozicije koje se u jeziku teorije mogu izvesti iz njezinih aksioma, pomoću njezinih pravila zaključivanja, dokazive su propozicije te teorije. Kraće ih zovemo teoremima te teorije.

Među deduktivnim teorijama razlikujemo dvije osnovne vrste. U prvu vrstu spadaju teorije koje opisuju jednu strukturu i njihov je krajnji cilj dokazivanje svih istina o toj jednoj strukturi. Takva je, na primjer, deduktivna aritmetika koja opisuje strukturu prirodnih brojeva i čiji je krajnji cilj dokazivanje svih istina o toj strukturi (tj. svih istina o prirodnim brojevima)[8]. U drugu vrstu spadaju teorije koje opisuju više struktura (cijelu klasu struktura) i čiji je krajnji cilj dokazivanje svih onih istina koje vrijede u svim strukturama te klase. Takva je, na primjer, teorija Abelovih grupa koja opisuje cijelu klasu grupa i čiji je krajnji cilj dokazivanje svih onih istina koje vrijede u svim grupama te klase. Prirodoslovne teorije, ukoliko su uopće formulirane kao deduktivne teorije, spadaju u prvu vrstu jer je njihov predmet jedna jedina struktura, u ovom slučaju sama priroda. Nas će zanimati deduktivne teorije prve vrste.

4. Što tvrdi Gödelov teorem o potpunosti

Deduktivna teorija, kako smo je definirali u prethodnom odjeljku, može promašiti svoj krajnji cilj zbog nedovoljnosti svojih aksioma ili pak zbog nedovoljnosti svojih pravila zaključivanja. Povijest matematike, od Euklidovih *Elemenata* do Dedekindove[9] konačne aksiomatizacije aritmetike (Dedekind, 1888.), pokazuje koliko je teška bila potraga za dovoljnom aksiomatizacijom temeljne matematičke strukture prirodnih brojeva. Povijest logike, od Aristotelovog *Organona* do Booleovih *Zakona mišljenja* (Boole, 1854.), pokazuje kako se teško dolazi do dovoljnog sustava logičkih pravila zaključivanja. Booleov sustav logičkih pravila, iako neizmjerno bogatiji od Aristotelovog, još je uvijek bio nedovoljan za izvođenje teorema većine matematičkih teorija.

Frege je bio prvi koji je sustav logičkih pravila zaključivanja pokušao destilirati iz matematičke prakse[10]. Zahvaljujući tome uspio je doći do dovoljnog sustava logičkih pravila (Frege, 1879.), koji je ugradio u svoj deduktivni sustav elementarne i više aritmetike (Frege, 1893.-1903.).

[8]Gödel je, svojim 1. teoremom o nepotpunosti dokazao da se takav krajnji cilj ne može postići.

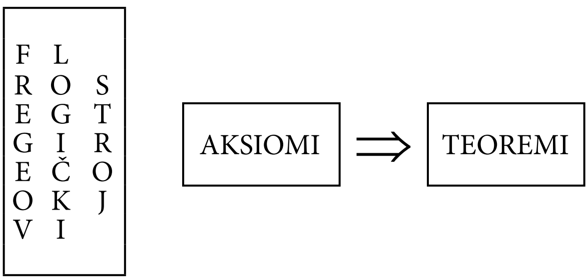
[9]Peanova aksiomatizacija prirodnih brojeva koju nalazimo u (Peano, 1889.) zapravo je preuzeta od Dedekinda (Dedekind, 1888.), što Peano i navodi na samom početku svojega članka. Više o cijeloj problematici aksiomatizacije prirodnih brojeva može se naći u (Šikić, 1989.), pod naslovom *Kako su i zašto aksiomatizirani prirodni brojevi*.

[10]Više o značaju takvog pristupa logici može se naći u (Šikić, 1989.), pod naslovom *Boole i Frege - matematika logike vs. logika matematike.*

Taj se deduktivni sustav pokazao kontradiktornim u Fregeovom dodatnom pokušaju da ga učini čisto logičkim, tj. u pokušaju da sve aritmetičke pojmove definira pomoću logičkih pojmova, te da sve aritmetičke aksiome izvede iz logičkih pravila. Bez obzira na taj specifični neuspjeh, Fregeov logički sustav (Frege, 1879.) najznačajniji je korak logike od njezina utemeljenja u *Organonu*. Naime, Russell i Whitehead su primjerom pokazali (Whithead & Russell, 1910.-1913.) kako se uz pomoć Fregeovih pravila zaključivanja sva poznata matematika može izvesti iz svega nekoliko aksioma. Time je empirijski potvrđena dovoljnost Fregovih pravila zaključivanja.

Ta empirijska potvrda još nije dokaz da se svaki (poznati ili nepoznati) logički valjani izvod može realizirati pomoću Fregeovih pravila zaključivanja. To je dokazao tek Gödel svojim teoremom o potpunosti (Gödel, 1929. i 1930.). Teorem tvrdi da Fregeova logička pravila zaključivanja[11] iz zadanih aksioma uvijek generiraju **sve** njihove logičke posljedice. Naravno, Gödelov teorem o potpunosti ima smisla tek onda ako je pojam logičke posljedice definiran neovisno o pravilima zaključivanja; inače bi bio puka tautologija. Zametak takve neovisne definicije nalazimo kod Bolzana, a njezina novija povijest uključuje imena Bernaysa, Hilberta, Gödela (1929. i 1930.), Tarskog (1936.), Kemenyja i Henkina. Jednostavno je možemo formulirati na sljedeći način: Propozicija $\boldsymbol{P}$ je logička posljedica propozicija (npr. aksioma) $\boldsymbol{A_1,A_2,A_3,\ldots}$ ako u svakoj strukturi u kojoj vrijede $\boldsymbol{A_1,A_2,A_3,\ldots}$ nužno vrijedi i $\boldsymbol{P}$; drugim riječima, ako svaka struktura koja modelira propozicije $\boldsymbol{A_1,A_2,A_3,\ldots}$ nužno modelira i propoziciju $\boldsymbol{P}$.

Dakle, ponovimo to još jednom, Fregeova logička pravila su potpuna jer, primijenjena na bilo koji skup aksioma, sigurno generiraju **sve** njihove logičke posljedice. Zamislimo li Fregeova pravila kao fiksni logički stroj u koji se umeću aksiomi (v. sl. 5.), onda svaku deduktivnu teoriju možemo zamišljati kao strojni generator svih teorema koji su logičke posljedice umetnutih aksioma. (Činjenicu da je propozicija $\boldsymbol{T}$ teorem određene deduktivne teorije označavat ćemo sa $\vdash \boldsymbol{T}$, u skladu sa sl. 5.)



Slika 5.

[11]Teorem vrijedi za sustave logičkih pravila zaključivanja sasvim određene vrste, koje u literaturi nalazimo pod raznim imenima: kvantifikacijska logika, predikatska logika, logika 1. reda, funkcijski račun 1. reda itd. No, svaki sustav te vrste potječe od Fregeovog "prasustava".

Povežemo li Gödelov teorem o potpunosti s njegovim 1. teoremom o nepotpunosti, lako ćemo zaključiti da je nepotpunost korektnih i dovoljno jakih deduktivnih teorija uzrokovana nepotpunošću njihovih aksioma. Kratko bismo mogli reći da su korektne i dovoljno jake deduktivne teorije logički potpune i faktički nepotpune. (Tu činjenicu još jasnije ističe Skolem-Löwenheimov teorem (Skolem, 1919. i Löwenheim, 1915.). Taj teorem tvrdi da svaki skup aksioma kojim želimo karakterizirati strukturu s beskonačno mnogo objekata (kakva je, na primjer, temeljna matematička struktura prirodnih brojeva) nužno vrijedi i u mnogim drugim strukturama koje su bitno različite od strukture koju želimo karakterizirati. Drugim riječima, deduktivne teorije prve vrste, koje bi trebale karakterizirati jednu jedinu strukturu (usp. kraj 3. odjeljka), zapravo nisu moguće.

5. Kako se dokazuje nepotpunost

Svaka deduktivna teorija zapravo je stroj za generiranje teorema. Ako je takav stroj korektan i dovoljno jak, onda je on faktički nepotpun, iako je logički potpun. Zašto je to tako pokazat će nam jedan jednostavni primjer faktički nepotpunog stroja, koji je tipičan (usp. Quine, 1946. i Smullyan, 1992.).

Zamislimo stroj koji printa nizove simbola, –, ***P*** i ***d***, koje ćemo zvati izrazima. Na primjer, ***Pd***–, ***d***, – – ***d*** i ***P***– su izrazi. Stroj ih printa u diskretnim vremenskim razmacima koje ćemo zvati trenucima. Dakle, ako stroj u 1. trenutku printa izraz ***Pd***–, u 2. trenutku izraz ***d***, u 3. trenutku izraz – – ***d*** i u 4. izraz ***P***–, onda rezultat njegovoga rada nakon prva četiri trenutka izgleda ovako:

	4	3	2	1
	P–	– – ***d***	***d***	***Pd***–

Izrazi oblika

(1)	***PX,***	(2)	–***PX,***
(3)	***PdX,***	(4)	–***PdX,***

gdje je ***X*** bilo koji izraz, imaju sljedeća značenja[12]:

(1) Izraz oblika ***PX*** znači da stroj printa (printao je, printat će) izraz ***X***.

(2) Izraz oblika –***PX*** znači da stroj ne printa izraz ***X***.

[12]Značenja ostalih izraza nisu bitna za daljnju argumentaciju. Oni čak mogu biti bez značenja.

Prije no što definiramo značenja izraza koji su oblika (3) i (4), potrebno je definirati značenje dijagonalizacije ***d***. Dijagonalizacija ***dX*** izraza ***X*** je izraz ***XX***. Na primjer, dijagonalizacija ***d–Pd*** izraza ***–Pd*** je izraz ***–Pd–Pd***.

(3) Izraz oblika ***PdX*** znači da stroj printa dijagonalizaciju izraza ***X***.

(4) Izraz oblika ***–PdX*** znači da stroj ne printa dijagonalizaciju izraza ***X***.

Izraze oblika (1) - (4) zvat ćemo P-rečenicama. Dakle, stroj (između ostalog) printa P-rečenice koje govore o tome što on printa, tj. stroj je samoreferirajući. Postavlja se pitanje je li moguće konstruirati stroj koji bi bio **korektan**, tj. ne bi printao neistinite P-rečenice, i koji bi, uz to, bio **faktički potpun**, tj. prije ili kasnije otprintao bi sve istinite P-rečenice. Odgovor je negativan:

AKO JE STROJ KOREKTAN, ONDA JE NEPOTPUN.

Razmotrimo, naime, P-rečenicu ***–Pd–Pd***. Budući da je ona dijagonalizacija od ***–Pd***, lako se vidi da vrijede sljedeće ekvivalencije:

Rečenica ***–Pd–Pd*** je istinita. ↔ Stroj ne printa dijagonalizaciju od ***–Pd***. ↔

↔ Stroj ne printa rečenicu ***–Pd–Pd***.

Dakle, imamo dvije mogućnosti:

1. Rečenica ***–Pd–Pd*** je istinita i stroj je ne printa.
2. Rečenica ***–Pd–Pd*** nije istinita i stroj je printa.

Ako je stroj korektan, onda druga mogućnost ne dolazi u obzir, pa preostaje samo prva mogućnost, tj. stroj je faktički nepotpun. (Uočimo da ključna P-rečenica ovoga dokaza, ***–Pd–Pd***, sama o sebi govori da nije printana.)

Ovakav dokaz faktičke nepotpunosti možemo ponoviti za svaku deduktivnu teoriju (stroj), ako je ona korektna i dovoljno jaka u sljedećem smislu:

I. Teorija sadrži rečenice koje govore o dokazivosti njezinih rečenica.

II. Teorija sadrži funkciju dijagonaliziranja koja omogućuje izgradnju konkretne Gödelove rečenice ***G*** koja sama o sebi govori da nije dokaziva.

Naime, sada «printan" postaje «dokaziv", P-rečenice postaju rečenice o dokazivosti, konkretna rečenica ***–Pd–Pd*** (koja sama o sebi govori da nije printana) postaje konkretna Gödelova rečenica ***G*** (koja sama o sebi govori da nije dokaziva).

Gödel je zapravo dokazao da deduktivna teorija koja sadrži elementarnu aritmetiku ima svojstva I. i II. Najprije je svakom izrazu ***J***, izgrađenom od logičkih i nelogičkih elemenata teorije, pridružio njegov broj ***'J'*** koji je ime tog izraza u samoj teoriji. To se pridruženje zove Gödelovom numeracijom. Budući da teorija obuhvaća elementarnu aritmetiku, ona sadrži rečenice koje izriču tvrdnje o brojevima koje pomoću Gödelove numeracije postaju tvrdnje o izrazima. Neke od tih rečenica izriču da nešto jest rečenica, neke da nešto jest aksiom,..., neke da je nešto dokaz neke rečenice. Gödel je pokazao kako se svaki od navedenih predikata: ***Rx*** (***x*** je rečenica), ***Ax*** (***x*** je aksiom),..., ***yDx*** (***y*** je dokaz od ***x***), može definirati unutar deduktivne teorije u tom smislu da je

$$\vdash \boldsymbol{Rn} \;\leftrightarrow\; \textbf{\textit{n je Gödelov broj rečenice}}$$

$$\vdash \boldsymbol{An} \;\leftrightarrow\; \textbf{\textit{n je Gödelov broj aksioma}}$$

$$\vdots$$

$$\vdash \boldsymbol{mDn} \;\leftrightarrow\; \textbf{\textit{m je Gödelov broj niza rečenica koje čine dokaz rečenice čiji je Gödelov broj n}}$$

To je bio najsloženiji dio Gödelovog dokaza nepotpunosti[13]. Kada je definiran dvomjesni predikat ***D***, lako je definirati jednomjesni predikat dokazivosti ***Dok*** na sljedeći način:

$$\boldsymbol{Dok\ x} \;\leftrightarrow\; \exists \boldsymbol{y(yDx)}$$

Time je potvrđeno I. svojstvo deduktivne teorije.

Gödel je potom aritmetički definirao funkciju dijagonaliziranja ***d***, da bi pomoću nje, na sljedeći način, izgradio rečenicu ***G***:

$$\boldsymbol{G = -Dok\ d('-Dok\ d(x)')}$$

Nije bilo teško dokazati[14] da za tu rečenicu vrijedi

$$\vdash \boldsymbol{(G \;\leftrightarrow\; -Dok\ 'G')},$$

[13]Gödel je razvio cijelu teoriju primitivno rekurzivnih funkcija da bi došao do tog rezultata.

[14]Jednostavan dokaz te činjenice može se naći u (Šikić, 1992.), gdje je, osim toga, pokazano u kojoj je on vezi s Cantorovim postupkom dijagonalizacije. Ključnu ulogu te činjenice, koja nije eksplicitno istaknuta u Gödelovom radu, uočio je Carnap, pokazavši da iz Gödelovih razmatranja zapravo slijedi da za svaki predikat ***Pr***, koji je definabilan u teoriji, postoji rečenica ***C*** za koju vrijedi $\vdash$ ***(C*** $\leftrightarrow$ ***Pr 'C')*** (Carnap, 1934.). Odavde neposredno slijedi rezultat Tarskog o nedefinabilnosti pojma istinitosti propozicija neke teorije u samoj toj teoriji. Naime, pojam istinitosti ***Ist*** morao bi zadovoljavati $\vdash$ ***(R*** $\leftrightarrow$ ***Ist 'R')*** za svaku rečenicu ***R*** te teorije. Kada bi takav pojam ***Ist*** bio definabilan u samoj teoriji, onda bi u njoj bio definabilan i pojam – ***Ist***. Za taj bi pojam (prema Carnapu, usp. prethodnu napomenu) postojala rečenica ***C*** sa svojstvom $\vdash$ ***(C*** $\leftrightarrow$ – ***Ist 'C')***, što je u kontradikciji s prethodnim zahtjevom. Dakle, pojam ***Ist*** nije definabilan u samoj teoriji.

tj. da ona sama o sebi tvrdi da nije dokaziva, te da je to dokazivo u samoj teoriji (uočite znak $\vdash$ u gornjoj formuli). Time je potvrđeno II. svojstvo deduktivne teorije.

Nakon toga se dokaz nepotpunosti može provesti analogno dokazu faktičke nepotpunosti ranije opisanog stroja. Gödelov je dokaz bio nešto složeniji jer se on nije htio pozvati na pojam istinitosti[15], kojim smo se mi koristili u gornjem dokazu. (Zato u izvornom Gödelovom teoremu ne nailazimo na korektnost kao jedan od uvjeta nepotpunosti, nego je taj uvjet konzistentnost.)

Gödelov 2. teorem o nepotpunosti može se dokazati tako da se dokaz 1. teorema, koji je dokaz jedne tvrdnje o deduktivnoj teoriji, reproducira u samoj toj teoriji. Dakle, može se dokazati da je u samoj teoriji dokazivo da iz konzistentnosti teorije slijedi njezina nepotpunost, (tj. nedokazivost Gödelove rečenice $\boldsymbol{G}$):

$$\vdash (\boldsymbol{Con} \rightarrow -\boldsymbol{Dok}\ '\boldsymbol{G}')$$

S druge strane, ako bi se konzistentnost teorije mogla dokazati u samoj teoriji, vrijedilo bi

$$\vdash \boldsymbol{Con}$$

Iz te bi dvije činjenice slijedilo $\vdash -\boldsymbol{Dok}\ '\boldsymbol{G}'$, tj. $\vdash \boldsymbol{G}$ (jer je $\boldsymbol{G} = -\boldsymbol{Dok}\ '\boldsymbol{G}'$), što je u suprotnosti s 1. teoremom o nepotpunosti, koji dokazuje da ne vrijedi $\vdash \boldsymbol{G}$. To znači da je pretpostavka o dokazivosti konzistentnosti bila pogrešna. Dakle, konzistentnost teorije nije dokaziva u samoj teoriji.[16]

$$\boxed{\nvdash \boldsymbol{Con}}$$

Naravno, pretpostavka je svakog dokaza 2. teorema da konzistentnost teorije možemo izraziti u samoj teoriji (rečenicom koju smo označili s $\boldsymbol{Con}$). To je neposredna posljedica izrazivosti predikata dokazivosti $\boldsymbol{Dok}$, jer konzistentnost znači nedokazivost kontradikcije oblika $\boldsymbol{A}\ \&\ -\boldsymbol{A}$, koju kraće označavamo sa $\perp$. Dakle,

$$\boldsymbol{Con} = -\boldsymbol{Dok}\ '\perp'.$$

[15] Gödel je svojim teoremima o nepotpunosti (između ostalog) htio pokazati da nije moguće realizirati Hilbertov program opravdavanja klasične nefinitne matematike finitnim sredstvima. Njegovi teoremi to pokazuju tek ako su finitno dokazani. Budući da pojam istinitosti nije finitan, trebalo ga je izbjeći.

[16] Na ovakav dokaz 2. teorema upućuje Gödel u I. dijelu svojega članka. Međutim, detaljna provedba takvog dokaza bila bi doista mukotrpna, pa možda zbog toga Gödel nikada nije napisao II. dio svojeg članka (usp. 3. napomenu). Bernaysov se dokaz zato temelji na uvjetima izvodljivosti, tzv. Bernaysovim uvjetima, koji su formalizacija nekoliko osnovnih svojstava predikata ***Dok*** i iz kojih relativno brzo slijedi redukcija 2. teorema na 1. teorem. (Bernaysovi uvjeti još su uvijek presloženi da bismo ih ovdje mogli formulirati. Bitno su jednostavniji Löbovi uvjeti iz kojih ne slijedi samo 2. teorem nego i svi ostali rezultati o dokazivosti; usp. dalje.)

6. Što još znamo o dokazivosti

Vidjeli smo da ključnu ulogu u dokazu nepotpunosti neke dovoljno jake teorije ima Gödelova rečenica ***G*** koja sama o sebi tvrdi da nije dokaziva u toj teoriji. Gödel je dokazao da ona u toj teoriji nije dokaziva, što znači da je istinita.

Henkin je 1952. postavio pitanje što je s rečenicom ***H*** koja, u dovoljno jakoj teoriji, sama o sebi tvrdi da je dokaziva u toj teoriji[17]:

$$\vdash \boldsymbol{(H \leftrightarrow Dok\ 'H')}.$$

Iz Löbovog teorema slijedi da je takva rečenica uvijek dokaziva u toj teoriji (Löb, 1955). Naime, Löb je rješavajući Henkinov problem dokazao da za svaku rečenicu ***A*** vrijedi

(L) $$\vdash \boldsymbol{(Dok\ `A' \rightarrow A)} \quad \Rightarrow \quad \vdash \mathbf{A}.$$

Kreisel je 1965. dokazao da je Gödelov 2. teorem o nepotpunosti neposredna posljedica Löbovog teorema. Naime, po Löbovom teoremu vrijedi implikacija:

$$\vdash \boldsymbol{(Dok\ `\bot' \rightarrow \bot)} \ \Rightarrow \ \vdash \bot,$$

čija je konzekventa u suprotnosti s konzistentnošću teorije (jer konzistentnost znači da nije $\vdash \bot$). Dakle, njezina antecedenta ne vrijedi:

$$\nvdash \boldsymbol{(Dok\ `\bot' \rightarrow \bot)} \ \text{ tj. } \nvdash -\boldsymbol{Dok}\ '\bot',$$

jer $\boldsymbol{Dok}\ `\bot' \rightarrow \bot$ znači $-\boldsymbol{Dok}\ '\bot'$. (Činjenica da iz ***A*** slijedi kontradikcija znači da nije ***A***, tj. $\boldsymbol{A} \rightarrow \bot$ znači $-\boldsymbol{A}$.) Budući da je $-\boldsymbol{Dok}\ '\bot' = \boldsymbol{Con}$, dobili smo 2. Gödelov teorem:

$$\nvdash \boldsymbol{Con}.$$

Kripke je 1965. dokazao da vrijedi i obrat, tj. da je Löbov teorem posljedica 2. Gödelovog teorema. S vremenom su dokazana i mnoga druga svojstva **dokazivosti u dovoljno jakim deduktivnim teorijama**. Sva ta svojstva uvijek su se uspjela reducirati na nekoliko temeljnih svojstava, koja je formulirao Löb i koja se stoga zovu Löbovim uvjetima (Löb, 1952.). To su:

(D1) $\vdash \boldsymbol{A} \ \Rightarrow \vdash \boldsymbol{Dok\ `A'}$,

(D2) $\vdash \boldsymbol{(Dok\ `A' \ \&\ Dok\ `A \rightarrow B' \rightarrow Dok\ 'B')}$,

(D3) $\vdash \boldsymbol{(Dok\ `A' \rightarrow Dok\ 'Dok\ 'A'\ ')}$.

17Prema (Carnap, 1934) takva rečenica uvijek postoji; usp. 14. napomenu.

To nije bilo slučajno. Soloway je 1976. dokazao da se svi rezultati o dokazivosti u dovoljno jakim deduktivnim teorijama mogu izvesti iz Löbovih uvjeta (D1) - (D3) i Löbovog teorema (L). Time je dokazano da postoji potpuna teorija dokazivosti.

Napominjemo, na kraju, da je cijelo područje koje je 30-ih godina otvoreno Gödelovim otkrićem nepotpunosti - i dalje predmetom intenzivnih logičkih istraživanja.

Literatura:

Barwise, J. & Etchemendy, J. 1987. The Liar. *An Essay on Truth and Circularity*. Oxford University Press, Oxford.

Boole, G. 1854. *An Investigation of the Laws of Thought*, Walton & Moberly, London.

Carnap, R. 1934. *Logische Syntax der Sprache*, Springer, Beč.

Dedekind, R. 1888. *Was sind und was sollen die Zahlen*, Vieweg, Brunswick.

Frege, G. 1879. *Begriffschrift, eine Formelsprache des reinen Denkens*, Nebert, Halle.

Frege, G. 1893 - 1903. *Grundgesetze der Arithmetik, begriffschriftlich abgeleitet, Vol.1. i 2.* Pohle, Jena.

Gödel, K. 1929. *Über die Vollständigkeit des Logikkalküls*, neobjavljena disertacija, Sveučilište u Beču.

Gödel, K. 1930. *„Die Vollständigkeit der Axiome des logischen Funktionenkalküls"*, Monatshefte für Mathematik und Physik, vol.37.

Gödel, K. 1931. *„Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I."* Monatshefte für Mathematik und Physik, vol. 38.

Gupta, A. 1982. *„Truth and Paradox"*. Journal of Philosophical Logic, vol. 11.

Hilbert, D. & Bernays, P. 1934 - 1939. *Grundlagen der Mathematik*, Vol. 1. i 2, Springer, Berlin.

Kripke, S. 1975. *„Otuline of a theory of Truth"*. Journal of Philosophy, vol. 72.

Löb, M. 1955. *Solution of a problem of Leon Henkin*, Journal of Symbolic Logic, vol. 20.

Löwenheim, L. 1915. *„Über Möglichkeiten im Relativkalkül*, Mathematische Annalen, vol. 76.

Peano, G. 1889. *Arithmetices principia*, nova methodo exposita, Bocca, Torino.

Quine, W. O. 1946. *„Concatenation as a basis for arithmetic"*, Journal of Symbolic Logic, vol. 11.

Skolem, T. 1919. *„Untersuchungen über die Axiome des Klassenkalküls und über Produktions - und Summationsprobleme welche gewisse Klassen von Aussagen betreffen"*, Skrifter utgit av Videnskapsselskapet i Kristiania, I, no. 3.

Smullyan, R. 1992. *Gödelžs Incompleteness Theorems*, Oxford University Press, Oxford.

Šikić, Z. 1989. *Kako je stvarana novovjekovna matematika*, Školska knjiga, Zagreb.

Šikić, Z. 1992. *„The diagonal argument - a study of cases"* International Studies in the Philosophy of Science, vol.6, No.3.

Tarski, A. 1931. *„O pojeciju prawdy w odniesieniu do sformalizowanych nauk dedukcyjnych"* Ruch Filozoficzny, vol.12. Prevedeno na njemački u (Tarski, 1935).

Tarski, A. 1935. *Der Warheitsbegriff in den formalisierten Sprachen*. Studia Philosophica, Vol. 1.

Tarski, A. 1936. *„O pojciu wynikania logicznego"*, Przeglad Filozoficzny, vol 39. Prevedeno na njemački u (Tarski, 1936).

Tarski, A. 1936. *„Über den Begriff der logischen Folgerung"*, Actes du Congres International de Philosophie Scientifique, vol. 7.

Whitehead, A.N. & Russell, B. 1910-13. *Principia Mathematica,* Vol.1, 2. i 3. Cambridge University Press, Cambridge.