Univerzitet u Beogradu
Filozofski fakultet

Jovana B. Kostić

SAMOREFERENCIJA I TEORIJA POJMOVA

doktorska disertacija

ФИЛОЗОФСКИ
ФАКУЛТЕТ
1838

Beograd, 2021

University of Belgrade
Faculty of Philosophy

Jovana B. Kostić

SELF-REFERENCE AND THE THEORY OF CONCEPTS

Doctoral Dissertation

ФИЛОЗОФСКИ
ФАКУЛТЕТ
1838

Belgrade, 2021

PODACI O MENTORU I ČLANOVIMA KOMISIJE

MENTOR:

Dr Miloš Adžić, docent
Filozofski fakultet
Univerzitet u Beogradu

ČLANOVI KOMISIJE:

Dr Živan Lazović, redovni profesor
Filozofski fakultet
Univerzitet u Beogradu

Dr Zoran Petrić, naučni savetnik
Matematički institut
SANU

Datum odbrane:

SAMOREFERENCIJA I TEORIJA POJMOVA

REZIME

Ova doktorska disertacija bavi se *samoreferencijom* kao formalnim svojstvom *intenzionalnog značenja* jezičkih izraza, pre svega predikata. Formalna svojstva intenzionalnog značenja predikata trebalo bi da budu predmet nove logičke teorije koju je zamislio i u kratkim crtama opisao jedan od najznačajnijih logičara i matematičara prošlog veka, ali i uopšte - Kurt Gedel (Kurt Gödel). Pošto je koristio termin *pojam* za svojstva ili relacije koje čine intenzionalno značenje predikata, Gedel je tu teoriju nazvao *teorija pojmova*. Brojne, ali kratke i često zagonetne, Gedelove napomene svedoče o njegovom očekivanju da ta teorija postane za logiku centralna.

Osobina pojmova koja dobija poseban značaj u kontekstu zasnivanja formalne teorije koja se njima bavi jeste upravo njihova samoreferencija, tj. mogućnost primene pojma na samog sebe, kao i učestvovanja pojma u građenju sopstvenog značenja. Glavni zadatak ovog rada je bolje razumevanje uloge ta dva oblika samoreferencije u zasnivanju i razvoju teorije pojmova. Videćemo da ona može biti trostruka. Pre svega, samoreferencija u odnosu na određeni objekat moguća je upravo zahvaljujući intenzionalnom značenju tog objekta. Ona tako ukazuje na formalne karakteristike intenzionalnog značenja na kojima se zasniva njena mogućnost i koje ga suštinski razlikuju od *ekstenzionalnog značenja*. Osim toga, samoreferencija može imati i značajnu ulogu u definisanju pojmova i dokazivanju njihovih svojstava. Konačno, određene instance pojmovne samoreferencije vode paradoksima koji mogu činiti teoriju u kojoj su formulisani protivrečnom. Rešavanje tih paradoksa bi trebalo da omogući postavljanje osnova teorije pojmova.

Ova disertacija podeljena je u četiri dela. U uvodnom delu razjasnićemo osnovne termine koje ćemo koristiti u disertaciji. Takođe ćemo u kratkim crtama opisati istoriju samoreferencije i uticaj koji je imala na razvoj matematike. Taj uticaj je u najvećoj meri ostvaren preko njene uloge u formulisanju paradoksa unutar matematičkih teorija. Potreba da se ti paradoksi izbegnu dovela je do promene u razumevanju osnovnih matematičkih pojmova, kojom je samoreferencija stavljena izvan polja matematičkih istraživanja.

U drugom delu rada srešćemo se sa drugačijim gledanjem na samoreferenciju i njene potencijalno paradoksalne posledice. U tom delu bavimo se teorijama u kojima je samoreferencija postala standardno metodološko sredstvo, a njene paradoksalne posledice osnova za dolaženje do važnih, tipično negativnih, rezultata. Videćemo da različiti oblici samoreferencije koji se u tim oblastima javljaju imaju zajedničko jezgro koje objašnjava njihovu ulogu u dolaženju do tih rezultata ali i u formulisanju paradoksa.

Treći deo rada predstavlja pokušaj da se približimo razumevanju odnosa samoreferencije i intenzionalnog značenja objekta u odnosu na koji se ona javlja. Time bi trebalo da se približimo i razumevanju formalnih osobina pojmova. U ovom delu rada takođe ćemo detaljnije opisati paradokse kojima pojmovna samoreferencija vodi. Branićemo rešenje tih paradoksa koje se zasniva na ograničavanju smislene primenljivosti pojma i redefinisanju komplementa pojma. Pokušaćemo da pokažemo da to rešenje, ako je formulisano u okviru intuicionističke logike, predstavlja solidnu osnovu teorije pojmova.

Zaključni deo rada sumira rezultate istraživanja i objašnjava Gedelovu filozofsku poziciju i njenu vezu sa teorijom pojmova.

KLJUČNE REČI: Kurt Gedel, samoreferencija, intenzionalno značenje, pojam, teorija pojmova

NAUČNA OBLAST: Filozofija

UŽA NAUČNA OBLAST: Filozofija logike

SELF-REFERENCE AND THE THEORY OF CONCEPTS

ABSTRACT

This doctoral dissertation is a study of *self-reference* as a formal property that characterizes the *intensional meaning* of some linguistic expressions, such as predicates. Formal properties of intensional meanings of predicates are supposed to be the subject matter of a new logical theory that is envisioned and briefly described by one of the most important logicians and mathematicians of the last century, and in general - Kurt Gödel. Since he used the term *concept* to denote the properties or relations that form the intensional meaning of predicates, Gödel termed such a theory the *theory of concepts*. His numerous, but short and often enigmatic, remarks show that he expected this theory to become central to logic.

A property of concepts that assumes a special importance in the context of establishing a formal theory that studies them, is just their self-reference, i.e. the possibility of an application of a concept to itself, or of its containment in its own meaning. Better understanding of the role that both kinds of self-reference have in establishing and developing the theory of concepts will be the main goal of this work. We will see that this role can be threefold. First of all, the possibility of self-reference with respect to some object depends on its intensional meaning. It can thus expose some formal properties of intensional meaning on which it rests and which crucially distinguish it from the *extensional meaning*. Besides that, the self-reference can have a significant methodological use in the future theory, where it can participate in the definitions of concepts and in the proofs of their properties. Finally, some instances of self-reference lead to paradoxes, that can make the theory in which they are formulated inconsistent. An appropriate solution to these paradoxes should set the ground for the theory of concepts.

This dissertation is divided into four parts. In the first part we explain the terms we will be using throughout the work. We also briefly describe the history of self-reference and the influence it had on the development of mathematics. This influence is accomplished mainly through the role that self-reference had in the formulation of paradoxes inside mathematical theories. The need to avoid these paradoxes has led to a change in the way some basic mathematical notions are understood, which put the considerations of self-reference outside the field of mathematical investigation.

In the second part we see quite a different attitude towards the self-reference and its consequences, that became a standardly used tool in particular areas of logic. The apparently paradoxical consequences of self-reference are inside these theories transformed into some important, mostly negative, results. We will see that different forms of self-reference that appear in these areas share a common core which explains their role in achieving these results, but also in deriving different paradoxes.

The third part is an attempt at approaching the question of the connection between the self-reference with respect to some entity and its intensional understanding. This should bring us closer to understanding the formal properties of concepts that embody this intensional meaning. This part also tries to trace an adequate solution to the intensional paradoxes, guided by the remarks Gödel made on the subject. We argue in favor of one particular solution accomplished by restricting the meaningful applicability of concepts, and redefining the notion of a complement of some concept. We will try to show that, if developed inside intuitionistic logic, this solution makes for a solid basis of the future theory.

In the final part we review the results of the investigation and describe Gödel's philosophical position which explains his interest in the theory of concepts.

KEYWORDS: Kurt Gödel, self-reference, intensional meaning, concept, theory of concepts

SCIENTIFIC FIELD: Philosophy

SCIENTIFIC SUBFIELD: Philosophy of logic

SADRŽAJ

SPISAK SLIKA

1 UVOD

Glavni deo ovog rada predstavlja istraživanje *samoreferencije* i njenih posledica, pre svega iz perspektive njihove uloge u razumevanju formalnih karakteristika *intenzionalnog značenja*. Pod ovim terminom podrazumevamo značenje koje, bar prvenstveno, nije skupovno-teorijsko. Tu drugu vrstu značenja zvaćemo *ekstenzionalno značenje*. Semantika formalnih jezika logičkih i matematičkih teorija nastala na skupovno-teorijskoj osnovi, kao osnovnu relaciju u određivanju značenja lingvističkih objekata, kao što su predikati i drugi logički i matematički simboli (veznici, funkcije, relacije), uzima relaciju denotacije u kojoj oni stoje prema skupovima i drugim skupovno-teorijskim objektima. Takva je standardna semantika klasične logike, kao i različitih matematičkih teorija koje su na njoj zasnovane. U ovom radu ćemo, sa druge strane, pokušati da utvrdimo kakvo bi dodatno razumevanje jezičkih mehanizama mogao da donese alternativni pristup koji uzima u obzir i intenzionalne aspekte značenja. Time ni u kom slučaju ne pokušavamo da odbacimo ekstenzionalni pristup ili da umanjimo njegov značaj. Standardna skupovno-teorijska semantika koja proizilazi iz radova Tarskog (Alfred Tarski) i drugih, pokazala se izuzetno uspešnom. Ona je u velikoj meri doprinela razjašnjenju osnova logike, a posebno razvoju jedne njene grane - teorije modela. Osim toga, ta semantika imala je veliki uticaj i na filozofiju, posebno u drugoj polovini prošlog veka. Međutim, njena prekomerna upotreba dovela je do zanemarivanja alternativnih pristupa značenju i do ignorisanja činjenice da je intenzionalno značenje i dalje prisutno i relevantno, kako u običnom jeziku, tako i u formalnim jezicima logike i matematike. Neki aspekti intenzionalnog značenja logičkih i matematičkih simbola sadržani su, na primer, u načinu konstruisanja objekata koje oni denotiraju, u strukturi tih objekata ili u strukturi koju oni zajedno čine. Ti aspekti mogu imati određeni matematički i logički značaj, o čemu, kao što ćemo videti, svedoči i sama istorija matematike.

Među razlozima za uspeh skupovno-teorijske semantike može biti i činjenica da je razvoj matematike u velikoj meri unapredio naše razumevanje skupova koji u toj semantici predstavljaju značenje predikata, tj. njihovu *ekstenziju*. Nasuprot tome, naše razumevanje objekata koji bi trebalo da imaju analognu ulogu u kontekstu intenzionalnog značenja predikata i drugih lingvističkih izraza i dalje je prilično nerazvijeno. Te objekte zvaćemo, prateći Gedela, *pojmovima*. Pod pojmovima bismo tako podrazumevali svojstva ili relacije koje čine intenzionalno značenje predikata. Trenutno ne postoji teorija koja bi obezbedila solidnu osnovu za razumevanje njihovih formalnih osobina, kojima bi i određene osobine lingvističkih objekata mogle biti objašnjene. Napredovanje u tom razumevanju trebalo bi da bude omogućeno budućom *teorijom pojmova*, koja bi predstavljala intenzionalni pandan teorije skupova. Tu buduću teoriju Gedel je zamislio kao teoriju koja se bavi formalnim osobinama pojmova i relacija u kojima oni stoje, među kojima je osnovna *relacija primene pojma* (ili njoj inverzna *relacija potpadanja pod pojam*). Važno je naglasiti da zamisao o toj teoriji nije podređena nikakvom posebnom stanovištu u pogledu prirode i ontološkog statusa pojmova. Svojstva pojmova kojima bi se ta teorija bavila ne zavise od toga da li su pojmovi zavisni od jezika ili od naših lingvističkih navika i mentalnih sposobnosti, ili su nasuprot tome, nezavisno postojeći. Buduća teorija bi trebalo da pruži bolje razumevanje onih formalnih osobina pojmova koje su sugerisane određenim lingvističkim fenomenima, a ne bilo kakvom metafizičkom teorijom. Dakle, jedino što bi bilo pretpostavljeno o pojmovima jeste da su, za razliku od skupova čije se osobine čine u potpunosti nezavisnim od jezika na kojem su opisane, pojmovi usko vezani za jezik. Iz tog razloga bi i određeni lingvistički fenomeni i mehanizmi mogli imati važnu ulogu u razumevanju pojmova i potencijalno otkriti njihove osobine. To bi se posebno ticalo onih osobina predikata i drugih lingvističkih objekata čije razumevanje izmiče njihovoj skupovno-teorijskoj interpretaciji.

Jedna od takvih osobina karakteriše one predikate, formule i funkcijske simbole koji se primenjuju na sebe ili čije definicije referiraju na njih same. Ta osobina lingvističkih objekata, koju ćemo zvati njihovom samoreferencijom, ispoljava se kako u običnom jeziku, tako i u formalnim jezicima određenih matematičkih i logičkih teorija. Brojni značajni rezultati koji su omogućeni njenom upotrebom svedoče o sveopštoj prisutnosti i metodološkoj vrednosti samoreferencije. Neke od tih rezultata predstavićemo u ovom radu. Objašnjenje i opravdanje samoreferencije u njihovoj osnovi nalazi se, kako ćemo pokušati da pokažemo, u intenzionalnom razumevanju objekata u odnosu na koje se ona javlja. Iz toga bi sledilo da je samoreferencija koja karakteriše te objekte zasnovana na samoreferenciji koja se tiče pojmova koji čine njihovo intenzionalno značenje.

Istraživanje odnosa između samoreferencije i intenzionalnog značenja moglo bi da ukaže na značaj koji samoreferencija, u vidu primene pojma na samog sebe i učestvovanja pojma u građenju sopstvenog značenja, ima za razumevanje formalnih osobina pojmova i zasnivanje teorije koja se njima bavi. Mogućnost te *pojmovne samoreferencije* tiče se oba aspekta buduće teorije - njene semantike, kao i sintakse njenog jezika. Ona se, sa semantičke strane, tiče nameravanog značenja njenih terama, koje predstavljaju pojmovi, kao i relacije njihove primene, čije bi razumevanje bilo među glavnim ciljevima te teorije. Mogućnost primene pojma na samog sebe, pored toga, određuje i sintaktičku osobinu koja toj relaciji treba da bude pripisana u budućoj teoriji. Jedna od glavnih karakteristika teorije pojmova, koja bi je razlikovala od drugih logičkih i semantičkih teorija, bilo bi upravo to što ona ne bi zazirala od pojmovne samoreferencije. Naprotiv, ona bi je koristila za razjašnjenje nekih važnih formalnih osobina pojmova. Te osobine mogle bi činiti osnovu razlike između pojmova i njima analognih ekstenzionalnih objekata, tj. skupova onako kako ih razumemo u kontekstu teorije *ZFC*. U odnosu na te ekstenzionalne objekte, mogućnost samoreferencije se ne javlja. Formalizacijom svojstava pojmova koji omogućavaju njihovu samoreferenciju, jedan važan mehanizam primene i definisanja predikata dobio bi semantičko opravdanje. Istovremeno bi i teorija pojmova u samoreferenciji dobila jedno snažno metodološko sredstvo.

Sa druge strane, u kontekstu zasnivanja formalne teorije u kojoj bi samoreferencija trebalo da bude prisutna, neophodno je razmotriti i njene potencijalno negativne posledice u obliku paradoksa kojima ona vodi, a koji mogu tu teoriju učiniti protivrečnom. Način na koji je izvođenje kontradikcija izbegnuto u oblastima koje ćemo razmotriti, a paradoksalne posledice samoreferencije koja se u njima javlja transformisane u važne negativne rezultate, sugerisaće način na koji isto može da se postigne i u vezi sa pojmovnom samoreferencijom. To bi moglo da usmeri našu potragu za adekvatnim rešenjem *intenzionalnih paradoksa* koja bi trebalo da dovede do postavljanja osnove teorije pojmova.

U pokušaju da utvrdimo koja vrsta rešenja intenzionalnih paradoksa može voditi zasnivanju teorije pojmova, pratićemo i Gedelove napomene, posebno one sadržane u nedavno objavljenim delovima njegovog *Nachlass*-a koji čine neke od *Max Phil* svezaka. Mislimo, pre svega, na *Max Phil IX* (Gödel, 2020a), *Max Phil X* (Gödel, 2017b), *Max Phil XI* (Gödel, 2020b) i *Max Phil XII* (Gödel, 2021). Gedel je ostavio značajnu pisanu zaostavštinu koju čine rukopisi, napomene i tekstovi koje nije objavio za života. Deo tog materijala posthumno je objavljen u trećem tomu Gedelovih sabranih dela (Gödel, 1995). Tamo objavljeni materijal ne uključuje *Max Phil* sveske koje sadrže beleške i napomene koje je Gedel pisao za sebe o različitim filozofskim, matematičkim, lingvističkim, teološkim i drugim temama. Postoji ukupno petnaest takvih svezaka (Gedel ih je napisao šesnaest ali je jedna izgubljena). Datumi koji se u njima mogu naći svedoče da su pisane u periodu između 1934. i 1955. godine. *Max Phil* sveske su u vlasništvu Univerziteta u Prinstonu, odnosno, njegovog *Institute for Advanced Study* i čuvaju se u njegovoj biblioteci na odeljku *Department of Rare Books and Special Collections*. Gedelove napomene u ovim sveskama pisane su stenografskim pismom *Gabelsberger*, što ih je učinilo teško pristupačnim. Određeni delovi tih zapisa transkribovani su zahvaljujući grupi koja radi na istraživačkom projektu *Kurt Gödel*

Philosopher: From Logic to Cosmology pod rukovodstvom Gabrijele Kroko (Gabriella Crocco). Mi ćemo se u radu oslanjati na nemačke verzije četiri pomenute sveske iz *Max Phil* ciklusa, koje su nastale kao rezultat rada na ovom projektu i tako postale jedine *Max Phil* sveske čiji su transkripti u celini objavljeni. Osim na njih, oslanjaćemo se na još jedan nedavno priređen i objavljen deo Gedelovog *Nachlass*-a koji čine beleške za njegov kurs iz elementarne logike održan 1939. godine na Univerzitetu Notr Dejm (Gödel, 2017a). Kako je najveći deo Gedelovih napomena o teoriji pojmova sakupljen i objavljen od strane njegovog intelektualnog biografa - Hao Vanga (Hao Wang), značajan izvor materijala za našu temu predstavljaće i njegova dela, među kojima posebno (Wang, 1996)[1]. Kroz rad ćemo nastojati da adekvatno prikažemo Gedelova gledišta u vezi sa pojmovima i njihovom samoreferencijom, čiji nagoveštaji se mogu naći na tim različitim mestima, kao i u njegovim objavljenim delima. Pokušaćemo da njegove napomene iz tih raznorodnih izvora međusobno povežemo u jedno koherentno stanovište koje bi moglo da usmeri našu potragu za teorijom pojmova. Pojmovi, njihove osobine i teorija koja bi se njima bavila, dugo su bili predmet Gedelovog razmišljanja - od najmanje kraja tridesetih, pa sve do kraja njegovog života (1978). S obzirom na to, prirodno je očekivati da se njegovo mišljenje o tim pitanjima menjalo. Ipak, sudeći po tome šta je Gedel o njima govorio Vangu i kako ih je on sam u retrospektivi video i interpretirao, njegova gledišta o toj temi su tokom celog tog vremenskog perioda ostala suštinski ista. Iako je teorija pojmova u njegovim delima ostala na programatskom nivou, čini se da Gedel ipak ni u kom periodu nije gajio sumnje u pogledu mogućnosti njenog zasnivanja i značaja koji bi ono imalo za logiku i matematiku.

Pre nego što pređemo na formalnije teme, razmotrićemo par primera koji bi trebalo da ilustruju šta je to samoreferencija, način na koji ona funkcioniše, kao i neočekivane posledice koje njena upotreba može da proizvede.

1.1 ODGOVOR JE U NASLOVU OVOG POGLAVLJA

Kako sam naslov sugeriše, samoreferencija se sastoji u referiranju nekog objekta na samog sebe.[2] Pošto postoje različite vrste objekata koje imaju referencijalnu funkciju, kao i različiti načini na koje ta funkcija može biti ostvarena, samoreferencija se takođe može javiti u različitim oblicima. Ona može biti sadržana u referiranju neke reči, izraza ili rečenice na samu sebe, u slici koja prikazuje samu sebe ili priči koja se pojavljuje u sopstvenom zapletu, itd. U primerima kao što su: 'ova fraza', 'Šta je tačan odgovor na ovo pitanje?', 'Ova rečenica ima pet reči', '„Formira istinitu rečenicu kada se nadoveže na svoj citat" formira istinitu rečenicu kada se nadoveže na svoj citat', samoreferencija se postiže referiranjem nekog izraza - fraze ili rečenice na samu sebe. Cilj te samoreferencije može biti konstruisanje rečenice koja tvrdi nešto o samoj sebi, kao što je slučaj u poslednja dva primera. To se postiže pripisivanjem određenog predikata izrazu koji referira na samu rečenicu koja je rezultat te predikacije. Ako je ta rečenica istinita, onda se kaže da se dati predikat primenjuje na nju, odnosno da ona ima svojstvo koje taj predikat izražava.

Predikat koji izražava određeno svojstvo ili pojam može i sam posedovati to svojstvo, kao što je slučaj sa predikatima 'srpski' koji je i sam srpska reč, ili 'kratak' koji je i sam kratka reč. Ta vrsta samoreferencije sastoji se u primeni svojstva, pojma ili funkcije na izraz koji označava ili opisuje samo to svojstvo, pojam ili funkciju. Ako u potpunosti iskoračimo van jezika, nailazimo

1 Ova knjiga sadrži zapise stavova i mišljenja koja je Gedel delio sa Vangom u njihovim mnogobrojnim diskusijama, zajedno sa Vangovim komentarima kojima su oni stavljeni u odgovarajući kontekst. I sam logičar i matematičar, Vang je bio Gedelov blizak prijatelj i sagovornik, posebno u poznim godinama Gedelovog života. Njemu dugujemo značajna saznanja o Gedelovom životu i delu koja u velikoj meri proširuju ona do kojih možemo doći na osnovu Gedelovih objavljenih radova.

2 Naslov je inspirisan delima Rejmonda Smalijana (Raymond Smullyan) poput *Kako se zove ova knjiga?* (u originalu: *What is the name of this book?*) ili *Ovoj knjizi nije potreban naslov* (u originalu: *This book needs no title*).

na primere samoreferencije koji se tiču primene pojma ili funkcije ne na njihova imena ili opise, nego na njih same. Na primer, *pojam pojma*, koji je i sam pojam, primenjuje se na samog sebe. Slično je slučaj sa identičkom funkcijom koja se može primeniti na svaki objekat, uključujući i samu sebe, da bi kao vrednost dala sam taj objekat. To su primeri vrste samoreferencije koja bi se prikladnije mogla nazvati *samoprimenljivošću* ili *samoprimenom* (funkcije, pojma, predikata ili nečeg sličnog). Druga vrsta samoreferencije, koja je naveliko korišćena u logičkim i matematičkim oblastima kojima ćemo se ovde baviti, tiče se karakterizacije nekog entiteta koja na neki način referira na sam taj entitet ili na samu sebe. Primer takve karakterizacije je standardna definicija sabiranja prema kojoj je *zbir dva pozitivna broja jednak nasledniku zbira prvog i prethodnika drugog broja*. Tu vrstu samoreferencije zvaćemo *samokarakterizacija*.

Kako će biti naglašeno kasnije u radu, samoreferencija se pokazala kao veoma korisno sredstvo, pre svega u odgovarajućim kontekstima, kao što su oni koje pružaju logika i matematika. Ipak, to sredstvo je neophodno pažljivo koristiti. Naime, paradigmatski primeri upotrebe samoreferencije su oni koji stoje u osnovama različitih paradoksa - naizgled korektnog rasuđivanja koje počiva na ubedljivim premisama ali vodi neprihvatljivim zaključcima. U većini slučajeva paradoksi se svode na dedukovanje kontradikcije iz određenih pretpostavki. Jedan od najpoznatijih paradoksa je takozvani PARADOKS LAŽLJIVCA koji se tiče rečenice 'Ova rečenica je lažna'. Ta rečenica bi po pretpostavci trebalo da bude ili istinita ili lažna. Međutim, ako uzmemo u obzir uslove za njenu istinitost i lažnost videćemo da ne može biti nijedno od ta dva. Ako je istinita, onda je ono što tvrdi slučaj, što znači da je lažna. Sa druge strane, ako je lažna, onda tvrdi nešto što jeste slučaj, što znači da je istinita. Pretpostavka da je data rečenica istinita ili lažna tako vodi kontradiktornom zaključku da mora biti i jedno i drugo. Time smo naizgled prinuđeni da odbacimo pretpostavku da rečenica u osnovi PARADOKSA LAŽLJIVCA uopšte ima istinosnu vrednost.

Posebnu grupu paradoksa, koji takođe imaju veze sa samoreferencijalnim jezičkim izrazima, čine paradoksi koji se tiču pojma definicije. Najjednostavniji od njih je takozvani BERIJEV PARADOKS koji se tiče formulacija poput sledeće: 'Najmanji prirodni broj koji ne može biti definisan sa manje od 15 reči'. Ovaj izraz naizgled uspeva da izdvoji jedinstveni prirodni broj sa datom osobinom i time ga ipak definiše koristeći manje od 15 reči. Definicija do koje se na taj način dolazi je samoreferencijalna u smislu da referira na sve definicije koje sadrže manje od 15 reči, među kojima je i ona sama. Ona istovremeno pruža primer samokarakterizacije datog broja koji je opisan pomoću totaliteta (svih brojeva koji su definisani sa manje od 15 reči) kojem i on sam pripada. Paradoksi te vrste (kao što je i poznati RIŠAROV PARADOKS) tipično se javljaju u kontekstu istraživanja o kontinuumu i sa njim povezanih pitanja definabilnosti. Pošto se tiču definicija kao jezičkih izraza, takvi paradoksi se nazivaju *semantičkim* ili *lingvističkim paradoksima*.

Među dobro poznatim paradoksima iz iste grupe je i GRELING-NELSONOV PARADOKS koji se tiče mogućnosti primene predikata na samog sebe. Predikati koji se ne primenjuju na sebe, tj. koji nemaju njima izraženo svojstvo, zovu se *heterološki* predikati. Takvi su, na primer, predikati 'engleski', 'ljudski', 'plavi', pošto ti predikati nisu redom niti engleski, niti ljudski, niti plavi. Paradoks nastaje pri pokušaju da utvrdimo da li se predikat 'heterološki' primenjuje na sebe ili ne. Ako je prvo slučaj, tj. ako se predikat 'heterološki' primenjuje na sebe, onda on ima svojstvo koje izražava, što znači da se ne primenjuje na sebe. Sa druge strane, ako se ne primenjuje na sebe, samim tim je to jedan heterološki predikat što znači da ima svojstvo koje izražava, pa se tako ipak primenjuje na samog sebe. Pošto je opet slučaj da obe pretpostavke vode kontradikciji, primorani smo da zaključimo da nijedna od njih ne može važiti. Jedan od najpoznatijih paradoksa u matematici - RASELOV PARADOKS, može se posmatrati kao verzija prethodno opisanog paradoksa u kojem su predikati interpretirani kao skupovi koji čine njihove ekstenzije. Paradoks se tiče ekstenzije predikata 'heterološki' koju bi trebalo da čini skup svih skupova koji nisu sopstveni elementi. Ako je taj skup sopstveni element, onda on poseduje svojstvo koje karakteriše sve njegove elemente što znači da ipak nije sopstveni element. Sa druge strane, ako taj skup nije sopstveni element, onda samim tim poseduje svojstvo na osnovu kojeg su njegovi elementi izabrani, pa je zato i on sam sopstveni element. Kako bi formulacija ovog paradoksa bila izbegnuta unutar aksiomatske teorije skupova, uvedena su određena ograničenja koja se tiču formiranja skupova. Ta ograničenja tiču se *aksiome neograničene komprehenzije* prema kojoj skup može biti formiran kao ekstenzija proizvoljnog predikata. Videćemo nešto kasnije u radu da bi uvedena ograničenja trebalo da budu opravdana ekstenzionalnom prirodom skupova koja ih čini nezavisnim od predikata pomoću kojih su navodno definisani.

Međutim, sličan paradoks preti svetu intenzionalnih objekata, u čijoj prirodi se ne može naći opravdanje za slična ograničenja kojima bi taj paradoks bio izbegnut. On se takođe može posmatrati kao verzija GRELING-NELSONOVOG PARADOKSA, u kojoj su umesto skupovima, predikati interpretirani pojmovima koje izražavaju. U toj intenzionalnoj verziji, paradoks se tiče *pojma neprimenljivosti na samog sebe*. Obe pretpostavke, da se dati pojam primenjuje kao i da se ne primenjuje na samog sebe, vode kontradikciji na potpuno analogan način kao u slučaju RASELOVOG PARADOKSA, što pokazuje da nijedna od njih ne može biti tačna.

Određene samoreferencijalne konstrukcije, slične onima o kojima je do sada bilo reči, impliciraju neka proizvoljna tvrđenja umesto kontradikcija. Na primer, Smalijan pokušava da dokaže da Deda Mraz postoji tako što formuliše sledeće tvrđenje: 'Ako se ne varam, Deda Mraz postoji' (Smullyan, 1978, pp. 202-203). Pretpostavimo da se Smalijan ne vara. Onda implikacija koju on tvrdi jeste istinita, i iz nje i pretpostavke da se ne vara, koja je njen antecedens, eliminacijom implikacije zaključujemo da Deda Mraz postoji. Time smo zapravo dokazali da je implikacija koju Smalijan tvrdi zaista istinita, jer se pokazalo da njen konsekvens sledi iz njenog antecedensa. Dakle, Smalijan se zapravo ne vara, iz čega, ponovnom primenom navedenog argumenta, možemo zaključiti da Deda Mraz zaista postoji. Ovaj argument ne zavisi od konkretne rečenice koju dokazuje, pa bi isto tako mogao biti korišćen u prilog bilo koje druge rečenice koja bi imala ulogu konsekvensa u samoreferencijalnom tvrđenju od kojeg se polazi. Drugim rečima, argument je zasnovan na konstrukciji rečenice A koja je ekvivalentna rečenici 'Ako je A istinito, onda B', pri čemu B može biti proizvoljna rečenica. Njegovim uopštenjem može se pokazati da je svaka takva rečenica A dokaziva, što kao posledicu ima dokazivost proizvoljnog konsekvensa implikacije kojoj je ta rečenica ekvivalentna. Rasuđivanje koje vodi tom paradoksalnom zaključku poznato je pod imenom KARIJEV PARADOKS.

Smalijan daje i druge primere korišćenja samoreferencije u dokazu proizvoljnih tvrđenja. Jedan od njih tiče se rečenice 'Ova rečenica je lažna i Deda Mraz ne postoji' (Smullyan, 1978, p. 203). Pretpostavimo da je data rečenica istinita. Tada bi i njen prvi konjunkt morao biti istinit, što ne bi bilo slučaj jer bi on tada pogrešno tvrdio da je rečenica lažna. Dakle, pošto ne može biti istinita, data rečenica bi trebalo da bude lažna. Razlog zbog kog je lažna ne može biti taj što je njen prvi konjunkt lažan jer ako je rečenica lažna, on tvrdi istinu. Zato jedino može biti slučaj da je rečenica lažna zbog toga što je njen drugi konjunkt lažan, tj. zbog toga što je lažno da Deda Mraz ne postoji. Na taj način smo ponovo došli do dokaza da Deda Mraz postoji, samo na osnovu pretpostavke da je samoreferencijalna rečenica od koje smo pošli istinita ili lažna.

Pošto u svakom od navedenih paradoksalnih argumenata neka samoreferencijalna konstrukcija igra ključnu ulogu, pokušaji da se takvi argumenti izbegnu uglavnom su fokusirani na preispitivanje legitimnosti tih konstrukcija ili načina na koji o njima rasuđujemo. Argument kojim bi trebalo izbeći paradokse je često taj da su pojmovi u osnovi samoreferencijalnih konstrukcija koje se u njima koriste neprikladno upotrebljeni. Neki od spomenutih primera samoreferencije tiču se matematičkih pojmova kao što su pojam skupa, funkcije ili broja. Videćemo da mogućnost samoreferencije u odnosu na te pojmove suštinski zavisi od načina na koji se oni razumeju. Naime, čini se da samo određeni intenzionalni način razumevanja tih pojmova može opravdati mogućnost samoreferencije u pogledu njima označenih objekata. Razumevanje tih pojmova i način na koji se oni koriste menjali su se tokom istorije matematike, delimično pod uticajem paradoksa kojima je određeni način njihovog razumevanja vodio. U nastavku ovog dela rada razmotrićemo neke od tih promena i videti kako su one uticale na mogućnost formulisanja matematički relevantnih instanci samoreferencije.

1.2 EKSTENZIONALNI ZAOKRET U MATEMATICI

Pitanje značenja matematičkog jezika jedno je od centralnih pitanja u filozofiji matematike. Ono je usko povezano sa pitanjem prirode matematičkog znanja i ontološkog statusa matematičkih objekata, u vezi sa kojima postoje različite, suprotstavljene teorije. Ali i bez ulaženja u filozofsku raspravu, moguće je utvrditi izvesne činjenice o tome kako matematičari razumeju značenje matematičkih simbola koje koriste u svom radu. Ono što će nas prvenstveno zanimati jeste kako su se uloga i značaj intenzionalnog i ekstenzionalnog značenja matematičkih pojmova preplitale i menjale kroz istoriju matematike. Prema uobičajenoj interpretaciji, uzimanje u obzir intenzionalnih aspekata matematičkog jezika najbolje se vidi u značaju koji se pripisuje definicijama i karakterizaciji matematičkih objekata u jeziku (cf. Feferman, 1985; Halbach & Visser, 2014). To se često dovodi u vezu sa konstruktivističkom pozicijom u filozofiji matematike, jer se pretpostavlja da se definicije matematičkih objekata tiču načina na koji oni mogu biti konstruisani ili izračunati od strane čoveka. Kasnije u radu videćemo da je Gedelovo stanovište po pitanju intenzionalnog značenja matematičkog jezika i njegove uloge u matematici, koje je bar delom motivisalo njegovo insistiranje na potrebi za teorijom pojmova, bilo radikalno drugačije. Prema njegovom stanovištu, intenzionalno značenje matematičkog jezika, sadržano u matematičkim pojmovima, ima ulogu ne samo u njihovom konstruisanju i izračunavanju, nego prvenstveno u razumevanju sveta matematičkih objekata. Pri kraju rada vratićemo se ovoj temi i nešto detaljnije razmotriti Gedelovo stanovište u filozofiji matematike, odnosno, njegov *pojmovni realizam*. Ono čime ćemo se do tada baviti ne zavisi od te Gedelove pozicije.

Pre sredine devetnaestog veka, nekim od najbitnijih matematičkih objekata, kao što su funkcije i skupovi, bilo je uobičajeno pripisati određenu intenzionalnu dimenziju izraženu u jeziku. Ta intenzionalna dimenzija ticala se načina na koji su matematički objekti konstruisani ili pravila njihove primene. Na primer, bilo je uobičajeno pretpostaviti da je pravilo na osnovu kojeg se neka funkcija izračunava suštinski deo te funkcije. Uz funkciju je tako nužno išla i njena definicija formulisana pomoću matematičkih operacija koje omogućavaju njeno izračunavanje. Prema Devlinovoj rekonstrukciji (Keith Devlin), takvo razumevanje funkcija bilo je uslovljeno interesima matematičara koji su bili fokusirani na njihovo korišćenje u računanju (Devlin, 2003).

Promena u razumevanju pojma funkcije u matematici desila se sredinom devetnaestog veka, kada je interes matematičara sa računanja pomoću neke funkcije preusmeren na razumevanje toga šta funkcija zapravo radi. Taj zaokret, koji su inicirali neki poznati matematičari kao što su Dirihle (Peter Gustav Lejeune Dirichlet), Riman (Bernhard Riemann) i Dedekind (Richard Dedekind), pokazao se izuzetno plodotvornim, pre svega u analizi gde je omogućio uspešno ispitivanje neprekidnih i diferencijabilnih funkcija (Devlin, 2003). Ono što je postalo posebno bitno jeste koju vrednost neka funkcija pripisuje svojim argumentima i kakav je međusobni odnos njenih argumenata i vrednosti, jer se smatralo da od toga zavise njena relevantna svojstva. Samim tim, više nije bilo ključno da način na koji vrednost funkcije zavisi od njenog argumenta bude izražen nekom formulom koja bi pokazivala kako se ta vrednost može izračunati. To je rezultiralo nezavisnošću funkcije od njene definicije ili načina izračunavanja. Funkcija je umesto toga izjednačena sa skupom uređenih parova njenih argumenata i vrednosti, koji se naziva *grafom funkcije*.

Trend definisanja matematičkih pojmova pomoću pojma skupa pojavio se i u drugim oblastima matematike. Tako je, na primer, u približno istom periodu Gaus (Carl Friedrich Gauss) inicirao novi pristup matematičkim strukturama, koji je potom razvijen od strane Dedekinda, a u kom je struktura shvaćena kao skup na kojem su definisane operacije sa odgovarajućim svojstvima

(Devlin, 2003). Nešto kasnije, Frege (Gottlob Frege) je ponudio definiciju prirodnih brojeva, prema kojoj prirodni brojevi predstavljaju svojstva skupova određena njihovom kardinalnošću. Ta definicija može se shvatiti kao da izjednačava prirodne brojeve sa skupovima koji predstavljaju ekstenziju pomenutog svojstva skupa: broj n bi tako bio izjednačen sa skupom svih skupova koji imaju n elemenata. Pritom se pretpostavlja da dva skupa imaju isti broj elemenata ako i samo ako između njih postoji bijekcija, što datu definiciju čini necirkularnom.

U ovom kontekstu, značenje pojma skupa postaje posebno bitno. Baš kao u slučaju funkcija, skupovi su u početku bili tretirani kao delimično intenzionalni objekti - pod skupom se podrazumevala kolekcija nekih objekata zajedno sa kriterijumom koji ti objekti treba da zadovolje da bi pripadali toj kolekciji. Kriterijum bi pritom bio da elementi datog skupa imaju određeno svojstvo ili stoje u određenoj relaciji, tj. da potpadaju pod određeni pojam. Drugim rečima, skupovi su bili tretirani kao ekstenzije pojmova pomoću kojih su definisani. Tako je pitanje pripadnosti nekom skupu bilo usko povezano sa pitanjem potpadanja pod odgovarajući pojam. Kako su pojmovi i u njima sadržana svojstva izraženi pomoću predikata ili formula određenog formalnog jezika, veza između pripadnosti skupu i posedovanja odgovarajućeg svojstva formulisana je pomoću aksiome komprehenzije:

$$\exists x \forall y \left(y \in x \leftrightarrow P\left(y\right)\right).$$

Tom aksiomom se za proizvoljan predikat P koji izražava određeno svojstvo tvrdi da postoji skup svih i samo takvih objekata koji poseduju to svojstvo. Ovaj princip je u potpunosti usvojen od strane Fregea, koji je, takođe, smatrao da predikati i formule matematičkog jezika izražavaju objektivno intenzionalno značenje (*smisao* ili *Sinn* u njegovoj terminologiji)[3].

Intenzionalno razumevanje skupova prisutno je i kod Bolcana (Bernard Bolzano), Kantorovog preteče u istraživanju skupova i beskonačnosti. Prema Bolcanovom stanovištu, dva skupa su jednaka samo ako je njihov način određivanja ili građenja isti (Bolzano, 2014, p. 98). Osnivač teorije skupova, Kantor (Georg Cantor), u početku je zauzimao slično stanovište o vezi između pojmova i skupova, ali je bio oprezniji od Fregea po pitanju aksiome komprehenzije čiju je primenu ograničio na određeni način. Naime, Kantor je pretpostavljao da samo pojmovi definisani za objekte prethodno specifikovanog domena mogu odrediti skup kao svoju ekstenziju u tom domenu (Tait, 2000, p. 272). Njegovo shvatanje skupova je kasnije izmenjeno, verovatno pod uticajem paradoksa koji su i pored tako ograničene komprehenzije pretili u kontekstu njegove teorije transfinitnih brojeva. U novoj karakterizaciji koju je Kantor koristio, skup nije određen kao ekstenzija pojma, nego kao „bilo koji totalitet o kojem se može misliti kao o jednom" (Tait, 2000, pp. 280-281). Paradoksi su imali još izraženiji uticaj na druge logičare i matematičare koji je u krajnjoj liniji doveo do napuštanja takozvane *naivne teorije skupova* u čijoj osnovi je bio opisani pojam skupa, i do prihvatanja u potpunosti ekstenzionalnog pojma skupa. U tom novom shvatanju, skupovi su izjednačeni sa kolekcijama objekata koje su u potpunosti nezavisne od svojstava tih objekata, relacija u kojima oni stoje ili struktura koju čine.

Pre nego što detaljnije razmotrimo kako je do te promene došlo i kakvu su ulogu paradoksi u tome imali, primetimo još i da otprilike u isto vreme vidimo začetke moderne logike. Tome je u najvećoj meri doprineo Frege (a zatim i Hilbert i Bernajs (Paul Bernays)), koji je nastojao da formira solidnu logičku osnovu iz koje bi, po njegovom mišljenju, celokupna aritmetika trebalo da bude izvodiva. Moderna logika, čiji je razvoj time započet, obezbedila je adekvatno okruženje unutar kojeg tvrđenja i dokazi ekstenzionalnih matematičkih teorija, kakva će i teorija skupova postati, mogu na adekvatan način biti formulisani i sistematizovani.

3 To Fregeovo gledište posebno dolazi do izražaja u njegovoj prepisci sa Hilbertom (David Hilbert) i Tomeom (Johannes Thomae) i kritici njihovog formalističkog stanovišta koje se najjasnije ispoljava u njihovom viđenju definicija matematičkih objekata (vidi, na primer, Resnik, 1974).

1.2.1 *Skupovno-teorijski paradoksi*

Paradoksi koji se tiču skupova i drugih skupovno-teorijskih pojmova imali su naizgled veliki uticaj na razvoj matematike. Osim RASELOVOM PARADOKSU, takva uloga pripisana je i paradoksima koji nose imena Kantora i Burali-Fortija (Cesare Burali-Forti). KANTOROV PARADOKS tiče se pojma kardinalnog broja skupa i nastaje u kontekstu razmatranja kardinalnog broja totaliteta svih skupova (ili svih kardinalnih brojeva). Ako je taj totalitet i sam skup, onda je njegova kardinalnost veća od kardinalnosti bilo kog drugog skupa. Međutim, na osnovu Kantorove teoreme znamo da bi kardinalnost partitivnog skupa tog skupa bila strogo veća od njegove kardinalnosti. Skup svih skupova tako ipak ne bi imao najveću kardinalnost, što protivreči prethodno rečenom. Sličan paradoks, koji se ovog puta tiče ordinala, je BURALI-FORTIJEV PARADOKS. Taj paradoks nastaje pri razmatranju totaliteta svih ordinala koji je, pod pretpostavkom da su svaka dva ordinala uporediva na osnovu njihovog standardnog uređenja, dobro uređen. Iz toga sledi da bi, da je skup, i sam taj totalitet bio jedan ordinal. Kao takav, on bi morao biti element samog sebe, iz čega sledi da bi bio striktno manji od samog sebe.

Prema usvojenom viđenju istorije matematike, ti i njima slični paradoksi, pre svega RASELOV PARADOKS koji se tiče bazičnog mehanizma primene predikata i relacije pripadnosti skupu kojom se ona interpretira, imali su neposredan i značajan uticaj na matematičku zajednicu. Paradoksi su navodno pokazali da su osnovni logički i matematički pojmovi protivrečni što je ozbiljno uzdrmalo osnove matematike i zahtevalo radikalnu reviziju njenih pojmova. Sa druge strane, novije istorijske rekonstrukcije pokazuju da uticaj paradoksa ipak nije bio tako dramatičan, bar ne u krugu matematičara, koji njima nisu bili iznenađeni (Peckhaus, 2004; Moore & Garciadiego, 1981).

Izvođenje kontradikcija iz određenih pretpostavki deo je uobičajene prakse matematičara koja omogućava dokazivanje tvrđenja metodom svođenja na apsurd. Upravo u tom kontekstu su se kontradikcije u osnovi Kantorovog i Burali-Fortijevog paradoksa prvi put i pojavile. Njihovo izvođenje nije za cilj imalo da se pokaže protivrečnost osnovnih pojmova teorije skupova, nego da se dođe do indirektnog dokaza određenih činjenica koje se tiču skupova, kardinala ili ordinala. Kantor je tako na osnovu kontradikcije koja je u osnovi njegovog paradoksa došao do zaključka da totalitet svih skupova (ili svih kardinalnih brojeva) nije skup, nego *inkonzistentna kolekcija* (*inconsistent multiplicity*), što bi odgovaralo *pravoj klasi* u modernoj terminologiji. Na osnovu argumenta sličnog Burali-Fortijevom, Kantor je došao do istog zaključka s obzirom na totalitet svih ordinala. Opšti zaključak na koji ta dva argumenta ukazuju bio bi, po njegovom mišljenju, da je neophodno razlikovati kolekcije koje se mogu konzistentno smatrati jedinstvom - to bi bili skupovi, i inkonzistentne kolekcije definisane pomoću određenih pojmova, koje su suviše velike da bi bile skupovi. Pripisivanje takvih osobina nekoj inkonzistentnoj kolekciji koje samo skupovi mogu imati (kao što su kardinalni i ordinalni broj), može voditi kontradikcijama. Ovaj Kantorov stav prema paradoksima odražava i stav Bolcana koji je takozvane paradokse koji prete analizi beskonačnosti smatrao bezopasnim. Na primer, umesto da ga smatra problemom koji ukazuje na neispravnost našeg pojma beskonačnosti, Bolcano je rezultat da je svaki beskonačni skup u bijekciji sa nekim svojim pravim podskupom usvojio kao rezultat koji otkriva karakteristično svojstvo beskonačnih skupova (Bolzano, 2014, pp. 95-97). Sličan pogled na paradokse delio je sa njima i Burali-Forti. On je kontradikciju do koje je došao video kao dokaz da pretpostavka da su svaka dva ordinala uporediva ne može biti ispravna (iako, kako se ispostavilo, nije bio u pravu). Dakle, kontradikcije u osnovi Kantorovog i Burali-Fortijevog paradoksa bile su poznate i korišćene u matematičkim argumentima i pre nego što su, navodno, postale ozbiljan problem za teoriju skupova.

Da otkriće pomenutih i njima sličnih paradoksa nije izazvalo nikakvu katastrofu u matematičkom svetu pokazuje i činjenica da je verzija RASELOVOG PARADOKSA bila poznata u Getingenu pod nazivom CERMELOV PARADOKS dugo pre nego što je Rasel objavio svoje otkriće, bez da je to imalo ikakav značajan uticaj na rad matematičara u teoriji skupova. Verzija KANTOROVOG PARADOKSA je takođe u to vreme bila formulisana od strane Hilberta. Hilbert je opisao svoju verziju kao čisto matematički paradoks pošto se u njegovom izvođenju ne javljaju pojmovi transfinitne teorije skupova, nego samo dva principa formiranja skupova koji se koriste u svakodnevnoj matematičkoj praksi: unija skupova i formiranje skupa svih funkcija na nekom skupu. Skupovi su definisani kao kolekcije koje se dobijaju iterativnom primenom te dve operacije počevši od skupa prirodnih brojeva. Neka je U unija svih tako izgrađenih skupova. Na osnovu pretpostavljene definicije skupa, i samo U je skup. Tada je i totalitet svih unarnih funkcija definisanih na U takođe skup koji se standardno označava sa U^U. Kao takav, on bi trebalo da bude podskup unije svih skupova, tj. trebalo bi da važi $U^U \subseteq U$. Međutim, kako je Hilbert pokazao, ta pretpostavka omogućava definisanje unarne funkcije na skupu U koja, na osnovu njene definicije, ne pripada skupu U^U. Naime, ako važi $U^U \subseteq U$, onda postoji funkcija iz skupa U^U u skup U koja svakoj funkciji iz prvog skupa pripisuje tačno jedan element drugog skupa i to različite elemente različitim funkcijama (takva funkcija zove se *injekcija*). Svaka funkcija iz skupa U^U može tada biti indeksirana elementom skupa U koji joj je pripisan na taj način. Definišimo funkciju f na skupu U tako da za svaku funkciju f_x iz skupa U^U, važi $f(x) \neq f_x(x)$. Pretpostavimo da je tako definisana funkcija f element skupa U^U i da je samim tim indeksirana nekim elementom a skupa U. Tada će za svaki element x skupa U, važiti $f_a(x) \neq f_x(x)$. Pošto je a element skupa U, možemo ga supstituisati mesto x, čime dobijamo da je $f_a(a) \neq f_a(a)$. Pošto smo na taj način dedukovali kontradikciju, prinuđeni smo na zaključak da funkcija f ipak ne pripada totalitetu svih funkcija na skupu U. Ključni korak u izvođenju ovog paradoksa jeste definisanje funkcije f na način koji garantuje da će se ona razlikovati od svake funkcije u datom skupu. Ono što omogućava takvu definiciju jeste u ovom slučaju *metoda dijagonalizacije* koja se, kao što ćemo videti, koristi u formulisanju brojnih drugih paradoksa, ali i u dolaženju do bitnih logičkih i matematičkih rezultata.

Neograničena aksioma komprehenzije, koja se obično smatra glavnim uzročnikom skupovno-teorijskih paradoksa, ne koristi se u izvođenju ovog paradoksa. Ipak, sličan rezultat se u ovom slučaju postiže operacijom unije i njenom primenom na totalitet svih skupova, kojom bi skup U trebalo da bude formiran. Definicija skupa U je samoreferencijalna zato što referira na sve skupove uključujući i skup U kao i svaki drugi skup koji je formiran pomoću njega (Peckhaus & Kahle, 2002).

Iako je uspeo da dedukuje kontradikciju koristeći samo osnovne skupovno-teorijske operacije, Hilbert nije smatrao da to pokazuje neodrživost standardne matematičke prakse. Razlog zbog kog Hilbertov i Cermelov paradoks nisu smatrani ozbiljnim problemom među matematičarima iz Getingena, i nisu čak odmah ni objavljeni, jeste taj što je njihova uloga bila da pokažu da je neka konkretna pretpostavka koja se tiče skupova ili načina na koji se o njima rasuđuje pogrešna. Paradoksi su shvaćeni kao poteškoća koju treba prevazići razvojem discipline i promenom pristupa njenom predmetu, a ne kao nešto što pokazuje da je sama disciplina osuđena na propast. Hilbert je bio naročito eksplicitan u svom stavu da paradoksi nisu nešto od čega treba zazirati, nego nešto što treba očekivati u razvoju svake mlade discipline, kakva je teorija skupova bila u to vreme (cf. Peckhaus & Kahle, 2002, p. 9).

Ono što može pretvoriti navedene paradokse i kontradikcije u ozbiljan problem za osnove matematike jeste odbijanje da se odbaci ijedna pretpostavka koja se tiče usvojenog pojma skupa i pravila rasuđivanja o njemu. Umesto toga može se tvrditi da paradoksi otkrivaju neizbežnu protivrečnost sadržanu u pojmu skupa koja onemogućava zasnivanje konzistentne teorije kojoj bi taj pojam bio u osnovi. Među istoričarima matematike postoji saglasnost da je osoba odgovorna za takav preokret u shvatanju paradoksa Rasel (Bertrand Russell). Raselov stav prema paradoksima je najverovatnije bio uslovljen njegovim filozofskim obrazovanjem i načinom na koji su takozvane antinomije percipirane u određenoj filozofskoj tradiciji, pre svega od strane Kanta i Hegela (Moore & Garciadiego, 1981). Takav stav, prema kom paradoksi pokazuju protivrečnost pojmova korišćenih u njihovom izvođenju, umesto preispitivanju određenih pretpostavki o naivnom pojmu skupa, vodio je kompletnoj reviziji tog pojma.

Kao što je već naglašeno, pojam skupa u osnovi paradoksa predstavlja određenu mešavinu intenzionalnog i ekstenzionalnog pojma. Ta karakteristika pretpostavljenog pojma skupa implicira da bi uslov pripadnosti nekom skupu trebalo da bude određen na osnovu intuitivnog razumevanja toga da li nešto potpada pod pojam kojim je taj skup definisan ili ne. Razumevanje relacije potpadanja pod pojam, ili primene pojma, pokazuje se međutim dosta složenijim (ispravno razumevanje te relacije biće jedan od glavnih zadataka za buduću teoriju pojmova). Moguća reakcija na pojavu skupovno-teorijskih paradoksa bi tako mogao biti zahtev za boljim razumevanjem te relacije i njenog odnosa prema relaciji pripadnosti skupu. Ovaj pristup je zastupao i sam Hilbert koji je smatrao da se do rešenja paradoksa može doći preispitivanjem pretpostavki o naivnoj koncepciji skupa, pre svega o svojstvima ili pojmovima koji su u nju uključeni (Peckhaus & Kahle, 2002, p. 2). Hilbert se dakle zalagao za zadržavanje originalne koncepcije skupa i otkrivanje pogrešnih pretpostavki o njoj za koje je mislio da potiču od tradicionalne logike i vode paradoksima. Ta pozicija međutim nije preovladala, i umesto toga, rešenje paradoksa tipično je traženo u reviziji pojma skupa i njegovoj zameni nekim manje problematičnim pojmom. U nastavku ćemo ukratko predstaviti glavne teorije kojima je vodila potraga za takvim rešenjem paradoksa. Videćemo da svi potencijalni odgovori na paradokse zauzimaju određeno stanovište po pitanju legitimnosti samoreferencije iz perspektive nove koncepcije skupa koju oni pretpostavljaju.

Jedna moguća revizija pojma skupa sastojala bi se u eliminaciji svih njegovih intenzionalnih aspekata i razumevanju skupova kao pukih kolekcija objekata, u čijem građenju svojstva tih objekata i relacije u kojima oni stoje ne igraju nikakvu značajnu ulogu. Takvo viđenje skupova je u osnovi takozvane *iterativne koncepcije skupa* za koju se može uzeti da je formalizovana u danas najšire prihvaćenoj i primenljivoj teoriji skupova - *Cermelo-Frenkel teoriji skupova* (*Zermelo-Fraenkel set theory*, skraćeno *ZF* ili *ZFC* ako uključuje aksiomu izbora). Prema toj koncepciji, skupovi su izgrađeni „odozdo-nagore”, primenom skupovnih operacija na već izgrađene skupove ili druge objekte. Iterativnom primenom tih operacija gradi se hijerarhija skupova u kojoj svaki skup može sadržati samo one skupove koji su izgrađeni pre njega. Dakle, da bi neki skup pripadao toj hijerarhiji, svi njegovi elementi moraju se javiti na nekom nižem nivou hijerarhije. Takvo shvatanje građenja skupova implicira određena njihova svojstva. Tako nijedan skup izgrađen na opisan način ne može biti element samog sebe, niti može sadržati kao elemente skupove u čijem građenju on učestvuje, jer ti skupovi ne mogu postojati pre nego što je on sam izgrađen. Samim tim, odbačena je i mogućnost postojanja skupova svih kardinala ili ordinala, jer bi ti skupovi morali da sadrže same sebe kao i sve kardinale i ordinale koji bi bili izgrađeni pomoću njih. Skup koji bi trebalo da sadrži sve skupove koji nisu sopstveni elementi, i koji bi prema ovoj koncepciji bio univerzalni skup, ne može postojati iz istog razloga. Unija svih skupova u ovoj hijerarhiji rezultirala bi novim skupom, što bi onemogućilo njeno korišćenje u izvođenju HILBERTOVOG PARADOKSA. Ispostavlja se, dakle, da se u osnovi svakog skupovno-teorijskog paradoksa nalazi pretpostavka o postojanju nekog skupa koji bi trebalo da sadrži samog sebe i koji prema novoj koncepciji skupa ne postoji. Na taj način izbegnuto je izvođenje skupovno-teorijskih paradoksa u ovoj teoriji.

Teorija *ZFC* pokazala se veoma uspešnom u matematičkom smislu, što pokazuju i njeni brojni, značajni rezultati. Osim toga, rešenje paradoksa koje ona nudi, a koje se zasniva na poricanju mogućnosti da neki skup bude sopstveni element, u potpunosti je opravdano iterativnom koncepcijom skupa koju ta teorija pretpostavlja.

Slično rešenje paradoksa nudi i teorija tipova koja je formulisana od strane Rasela, a zatim modifikovana i korišćena od strane drugih logičara, među kojima su najpoznatiji Čerč (Alonzo Church) i Gedel. U Gedelovoj verziji te teorije, *klase* (tj. totaliteti objekata koji formiraju ekstenziju nekog pojma) podeljene su u tipove prema vrsti objekata koje sadrže (Gödel, 1931).

Objekti prvog tipa su individue, tj. objekti koji sami nemaju elemente. Drugi tip sastoji se od klasa koje kao elemente sadrže individue, treći tip od klasa koje sadrže klase individua, itd. Klasa bilo kog tipa će kao elemente sadržati isključivo objekte nekog nižeg tipa. Tako ona ne može sadržati samu sebe, ili neku drugu klasu istog ili višeg tipa (kao što je bilo koja klasa koja nju sadrži). Hijerarhija tipova uključena je u jezik teorije tipova time što su imenima klasa pridružene oznake, poput indeksa, koje pokazuju njihov tip. Samo ona tvrđenja koja zadovoljavaju tipska ograničenja smatrana su smislenim formulama te teorije. Iz tog razloga, paradoksi koji se tiču skupova ili klasa u ovoj formalizaciji, ne mogu biti formulisani unutar teorije tipova, čime su i kontradikcije kojima oni potencijalno vode izbegnute.

Rasel se time ipak nije zadovoljio, već je pokušao da modifikuje teoriju kako bi je učinio adekvatnom za rešavanje ne samo skupovno-teorijskih, već i semantičkih paradoksa, kao što su oni koji se tiču problema definabilnosti. U to vreme, njegovo stanovište je bilo da je za izvođenje paradoksa odgovorna upotreba takozvanih *impredikativnih definicija* - definicija koje referiraju na totalitet kojem pripada i sam objekat koji je njima definisan (Russell, 1907). U tom tekstu, Rasel takve definicije zove *nepredikativnim* (*non-predicative*), ali se do danas zadržao samo termin koji smo prvi naveli. Najpoznatiji zastupnik tog stanovišta bio je Poenkare (Henri Poincaré). U skladu sa njim, kako bi izbegao paradokse, Rasel je formulisao *princip cirkularnosti* (*vicious circle principle*) prema kojem totalitet koji bi sadržao objekte koji se mogu definisati samo referiranjem na taj isti totalitet, ne može postojati. Uključivanje tog principa u teoriju trebalo bi da blokira izvođenje ne samo skupovno-teorijskih, nego i semantičkih paradoksa. Rasel je to postigao uvođenjem paralelne hijerarhije *redova* u svoju teoriju tipova. Svakoj klasi u hijerarhiji je pored tipa pripisan i red koji zavisi od tipa objekata preko kojih definicija te klase kvantifikuje. Ako definicija klase ne referira ni na kakav totalitet, ta klasa je nultog reda; ako kvantifikuje samo preko individua, onda je prvog reda; ako kvantifikuje preko klasa individua, onda je drugog reda, itd. Klasa ne može pripadati domenu kvantifikacije sopstvene definicije, jer njena definicija može kvantifikovati samo preko klasa nižeg reda. Ova teorija, koja se naziva *razgranatom teorijom tipova* (*ramified theory of types*), isključuje ne samo mogućnost samopripadanja ili samoprimene, već i drugih vrsta samoreferencije koje se javljaju u karakterizaciji objekata koja sadrži (indirektnu) referenciju na njega samog.

Postignuto rešenje semantičkih paradoksa nalikuje njihovom standardnom rešenju uvođenjem hijerarhije jezika opisane od strane Tarskog. Prema tom rešenju, za određene osobine rečenica jednog jezika važi da se one ne mogu opisati, niti se o njima može nešto tvrditi, unutar istog tog jezika. Takva osobina bi, na primer, bila njihova istinosna vrednost. Tvrđenja o istinosnoj vrednosti neke rečenice moguće je formulisati samo unutar jezika višeg nivoa. Ovu ideju moguće je prilagoditi tako da se odnosi i na mogućnost referiranja na definicije formulisane unutar jednoj jezika koja bi zahtevala upotrebu sredstava koja pripadaju jeziku višeg nivoa (to bi se moglo opravdati činjenicom da je definicija nekih objekata opis koji je *istinit* za sve i samo te objekte).

Zabrana impredikativnih definicija čini razgranatu teoriju tipova neodgovarajućom osnovom matematike, kako je naglašeno pre svega od strane Peana (Giuseppe Peano) i Cermela (Ernst Zermelo) (cf. Cantini, 2009, p. 18; 22). Razlog je to što su impredikativne definicije u upotrebi širom matematike i predstavljaju neizostavni deo mnogih njenih oblasti. Kako bi kompenzovao taj nedostatak svoje teorije, Rasel je uveo *aksiomu reducibilnosti* (*axiom of reducibility*) koja bi trebalo da opravda korišćenje impredikativnih definicija time što će garantovati da je svaka takva definicija ekstenzionalno ekvivalentna nekoj predikativnoj definiciji (tj. definiciji istih objekata koja kvantifikuje samo preko objekata nižeg reda). Rezultat je bio prekomplikovan i neintuitivan sistem. Originalna *prosta teorija tipova* bez redova pokazala se kao mnogo uspešnija i primenljivija. Ta teorija, koja je ovde formulisana tako da se odnosi na klase, može se umesto toga odnositi na pojmove ili na iskazne funkcije, o čemu će više reči biti kasnije u radu.

U ovom kontekstu treba pomenuti još jednu teoriju, nešto manje značajnu od prethodne dve. To je teorija *Nove osnove* (*New Foundation*, kraće *NF*) koju je formulisao Kvajn (W. V. O.

Quine). Ta teorija se može shvatiti kao pokušaj ponovnog uvođenja varijante naivne koncepcije skupa s obzirom na to da ona pretpostavlja donekle ograničenu verziju aksiome komprehenzije. U instancama aksiome komprehenzije koja je pretpostavljena u toj teoriji, mogu se javiti samo takozvane *stratifikovane* formule. Formula je stratifikovana ako se njenim slobodnim promenljivima mogu pripisati tipovi koji je čine smislenom iz perspektive teorije tipova. Time je implicirano da skupovi mogu biti definisani samo određenim formulama i njima izraženim svojstvima. Iako je samoreferencija u vidu pripadnosti skupa samom sebi dozvoljena u ovoj teoriji, paradoksi su ipak izbegnuti, pošto su definicije skupova koje se javljaju u njihovom izvođenju tipično nestratifikovane. Ipak, pretpostavljeno ograničenje komprehenzije čini se nemotivisanim. Pored toga, Kvajnova teorija ima i ozbiljne matematičke nedostatke (vidi: Fraenkel et al., 1973, pp. 161-167).

Dakle, najuspešnije teorije nastale kao reakcija na skupovno-teorijske paradokse, *ZFC* i prosta teorija tipova, izbegavaju paradokse tako što odbacuju mogućnosti samoreferencije u vidu pripadnosti skupa samom sebi. To se opravdava novom koncepcijom skupa koju one pretpostavljaju, a prema kojoj skup nije izgrađen pomoću nekog svojstva koje postaje njegov intenzionalni deo. Umesto toga, građenje skupova započinje primenom skupovnih operacija na neke osnovne, proste objekte koji sami nemaju elemente, i nastavlja se njihovom primenom na skupove koji su na taj način prethodno izgrađeni. Takvo shvatanje građenja skupova opravdava takozvanu *aksiomu zasnivanja* (*axiom of foundation*) teorije *ZFC*, prema kojoj svaki neprazan skup ima neki element sa kojim je disjunktan. Ta aksioma onemogućava da skup sadrži kao element sebe ili neki drugi skup čiji je on sam element. Nemogućnost samoreferencije u vezi skupova ima za posledicu nemogućnost samoreferencije u vezi funkcija i struktura koje su definisane kao skupovi određene vrste. Ekstenzionalno shvaćena funkcija ne može se primenjivati na sebe, pošto skup koji čini tu funkciju ne može biti element sopstvenog domena. Slično tome, struktura koju čini neki skup sa određenim operacijama definisanim na njemu, ne može biti sopstveni element. Takvim skupovima se ne mogu predstaviti ni određeni procesi ili lingvistički entiteti koje odlikuje samoreferencija, kao što su izračunavanja koja sadrže petlju ili cirkularni iskazi.

Odbacivanje ideje o induktivnom građenju skupova koja je u osnovi teorije *ZFC* i teorije tipova, vodi odbacivanju aksiome zasnivanja ili njoj ekvivalentne aksiome. Tako nastaju neke alternativne teorije skupova (takozvane *Non-well-founded set theories*). Umesto te aksiome, ove teorije usvajaju neku verziju *aksiome antizasnivanja* (*anti-foundation axiom*) koja implicira postojanje skupova koji su sopstveni elementi (za pregled tih teorija, vidi: Aczel, 1988).

Razumevanje skupova kao ekstenzija pojmova takođe vodi odbacivanju aksiome zasnivanja i prihvatanju njihove samoreferencije. Ako je neki skup definisan tako da sadrži sve objekte sa određenim svojstvom, i on sam ima to svojstvo, onda taj skup treba da sadrži i samog sebe. Pojam svojstva i posedovanja svojstva ili potpadanja pod pojam se tako pokazuju suštinskim za razmatranja koja vode samoreferenciji, a time i paradoksima, u naivnoj teoriji skupova. Ako učinimo skupove nezavisnim od svojstava njihovih elemenata ili pravila na osnovu kojih su izgrađeni, paradoksi u vezi sa njima se više ne javljaju. Sa druge strane, razdvajanjem skupova od svojstava i pojmova, otvoren je prostor za nezavisno istraživanje tih intenzionalnih entiteta kojima prete analogni paradoksi, poput paradoksa *pojma neprimenljivosti na sebe*. Te intenzionalne verzije skupovno-teorijskih paradoksa i dalje nisu rešene na odgovarajući način.

1.3 POVRATAK INTENZIONALNOSTI

Gedel je očekivao adekvatno rešenje intenzionalnih paradoksa od buduće teorije pojmova. Ona bi trebalo da rasvetli pogrešne pretpostavke o pojmovima koje su dovele do formulisanja paradoksa u vezi sa njima. Ta teorija zamišljena je kao logička teorija koja se bavi formalnim svojstvima pojmova koja otkrivaju njihovu intenzionalnu prirodu. Kakva je tačno priroda pojmova i relacije primene pojma je nešto o čemu imamo određeno intuitivno razumevanje. To razumevanje je, međutim, ograničeno i neprecizno, što se pokazuje i u našem nepoznavanju ispravnog rešenja intenzionalnih paradoksa. Teorija pojmova bi trebalo da rasvetli ta i srodna pitanja tako što će unaprediti naše razumevanje pojmova i učiniti ga preciznijim. Iako je u više navrata naglašavao značaj teorije pojmova za buduću logiku (neke od referenci su: Gödel, 1944; Wang, 1996, poglavlje 8), Gedel ipak nije puno otkrio o tome kako bi ta teorija trebalo da izgleda. Ipak, na osnovu njegovih malobrojnih i kratkih opisa te teorije, može se zaključiti da se Gedel nadao teoriji koja bi precizno opisala principe građenja složenih pojmova od jednostavnijih pomoću logičkih veznika, i utvrdila neophodna ograničenja koja se tiču primene tih principa. Ta ograničenja bi trebalo da vode i rešenju paradoksa koje bi pokazalo da su oni posledica neispravnog korišćenja principa formiranja pojmova. Ideja zasnivanja ove teorije se tako može shvatiti kao sleđenje Hilbertovog predloga da se rešenje paradoksa naivne teorije skupova traži u preispitivanju načina formiranja pojmova u njihovoj osnovi. Pošto od načina na koji je neki pojam formiran treba da zavisi i to da li će on biti (smisleno) primenljiv na neki objekat, principi građenja pojmova bi trebalo da imaju posledice i po razumevanje relacije primene ili smislene primenljivosti pojma. Ispravno razumevanje formalnih osobina te relacije i njenih razlika u odnosu na relaciju pripadnosti skupu bi trebalo da bude jedan od glavnih zadataka buduće teorije pojmova.

1.3.1 *Rešavanje intenzionalnih paradoksa*

Pronalaženje adekvatnog rešenja intenzionalnih paradoksa bi trebalo da unapredi naše razumevanje pojmova i usmeri dalji razvoj teorije koja se njima bavi. Rešavanje paradoksa koji se tiču pojmova po ugledu na skupovne paradokse, tj. zabranom pojmovne samoreferencije, bilo bi nemotivisano. Razlog je to što se u prirodi pojma ne može naći ništa što bi onemogućilo njegovu primenu na sebe ili njegovo definisanje korišćenjem njega samog ili pojmova u čijem značenju on na neki način učestvuje (cf. Wang, 1996, 8.6.20; 8.6.21)[4]. Upravo ta činjenica bi mogla biti razlog zbog kog je potraga za rešenjem skupovno-teorijskih paradoksa brzo preusmerena na potragu za alternativnim pojmom skupa. Novi ekstenzionalni pojam skupa, time što zabranjuje samoreferenciju u vezi sa skupovima, onemogućava izvođenje paradoksa koji počivaju na pretpostavci o mogućnosti te samoreferencije. Brzo je primećeno da to rešenje ne može biti adekvatno sve dok je pojam skupa shvaćen tako da sadrži svojstva i pojmove s obzirom na koje mogućnost samoreferencije ne može biti tako lako odbačena. Gedelovu potragu za rešenjem intenzionalnih paradoksa možemo videti kao povratak početnom problemu pronalaženja adekvatnog rešenja paradoksa koji nastaju u vezi sa izvornim pojmom skupa. Adekvatno rešenje tih paradoksa trebalo bi da bude fokusirano na iznalaženje načina na koji pojmovna samoreferencija može biti formalizovana u budućoj teoriji, a da je pritom dedukovanje kontradikcija iz određenih njenih instanci sprečeno. Ono što može usmeriti potragu za takvim rešenjem bili bi, dakle, primeri formalnih teorija koje se bave objektima u vezi sa kojima se javlja neka vrsta samoreferencije koja omogućava izvođenje paradoksalnih posledica, ali koje u tim teorijama ne vode kontradikcijama. Kao što ćemo videti u narednom delu rada, postoje logičke teorije u kojima je samoreferencija naširoko prisutna i korišćena za definisanje njihovih objekata i dokazivanje različitih rezultata. Upotreba samoreferencije se u tim teorijama pokazala ne samo bezopasnom, već i izuzetno korisnom u metodološkom smislu.

To se prvenstveno odnosi na teoriju izračunljivosti, oblast logike koja se bavi funkcijama izračunljivim pomoću određenog algoritma. Glavne metode te teorije koriste neku vrstu samoreferencije koja se tiče funkcija i u velikoj su meri motivisane paradoksima i konstrukcijama u njihovoj osnovi. Slična logička sredstva za koja se može reći da su izvedena iz paradoksa, tj. iz načina rasuđivanja koje njima vodi, javljaju se u još jednoj oblasti logike gde najistaknutiju pri-

4 Na Gedelove napomene dokumentovane u Vangovoj knjizi referiraćemo pomoću brojeva koji su im tamo pripisani.

menu imaju u dokazu Gedelovih teorema o nepotpunosti (za kompletniji pregled uloge paradoksa unutar logike vidi: Cantini, 2009). Gedelove teoreme pripadaju teoriji dokaza - logičkoj oblasti koja se bavi dokazima formulisanim unutar nekog formalnog sistema. Može se reći da je u ovim oblastima preovladao matematički pristup paradoksima koji je rezultirao njihovim pretvaranjem u važne metode definisanja objekata i dokazivanja njihovih svojstava. Ono što je presudilo u prilog takvom pristupu može biti činjenica da su intenzionalni aspekti određenih objekata, na kojima se zasniva mogućnost njihove samoreferencije, ključni za predmet kojim se te oblasti bave. Iz tog razloga, ti intenzionalni aspekti i osobine objekata koje se na njima zasnivaju ne mogu biti ignorisani.

U nastavku ćemo predstaviti neke od načina na koji se samoreferencija i ekvivalenti paradoksa kojima ona vodi koriste za dokazivanje fundamentalnih rezultata ovih teorija, koji su većinom negativne prirode. Time bi trebalo da postane jasno da samoreferencija može biti izuzetno značajno i primenljivo metodološko sredstvo, što bi i sa pragmatičkog stanovišta trebalo da motiviše napore da se ona uključi u buduću teoriju pojmova. Istovremeno će i način na koji su kontradikcije kojima samoreferencija vodi izbegnute u tim teorijama sugerisati način na koji se one mogu izbeći i u budućoj teoriji pojmova.

2 SAMOREFERENCIJA U LOGICI

Samoreferencija se u logici, kao i u matematici uopšte, može javiti u vidu primene nekog logičkog ili matematičkog objekta (funkcije ili nečeg drugog) na sebe ili sopstvenu definiciju, u vidu sadržavanja nekog objekta (npr. strukture, klase ili pojma) u samom sebi, ili u vidu definicije nekog objekta koja referira na sam taj objekat ili totalitet kojem i on pripada. Samoreferencija u svim tim oblicima ima vrlo značajnu metodološku funkciju. Ona je u osnovi različitih načina konstruisanja ili definisanja logičkih i matematičkih objekata, kao i dokazivanja različitih tvrđenja o njima ili njihovim formalizacijama. Zahvaljujući toj raznolikoj primeni, samoreferencija se pokazala kao veoma korisno logičko sredstvo. Posledice njene upotrebe nisu, međutim, uvek pozitivne. Zapravo, jedna od najznačajnijih uloga samoreferencije jeste prepoznavanje i dokazivanje specifičnih ograničenja određenih formalnih teorija u kojima se ona javlja. Videćemo da je ta uloga samoreferencije u potpunosti analogna njenoj ulozi u izvođenju paradoksa.

U prošlom poglavlju videli smo da su paradoksi u velikoj meri motivisali zasnivanje nekih značajnih formalnih teorija, kao što su *ZFC* i teorija tipova. Zapravo, jedan od ciljeva formalizacije matematičkih teorija bilo je upravo otkrivanje izvora relevantnih paradoksa. Formalizacija je u isto vreme omogućila analizu paradoksa koja je otkrila njihovu najopasniju posledicu - činjenicu da oni mogu učiniti teoriju u kojoj su formulisani protivrečnom. Time se pokazalo da je suočavanje sa paradoksima i njihovo rešavanje posebno važno u kontekstu zasnivanja formalne teorije koja, ako je protivrečna, u potpunosti gubi svoju funkciju i značaj. Samoreferencija koja je u osnovi paradoksa se unutar formalne teorije mora, dakle, pažljivo koristiti kako bi njena metodološka vrednost bila iskorišćena, dok je istovremeno izbegnuta opasnost od protivrečnosti stvorena njenom upotrebom.

Sa druge strane, zahvaljujući njenoj formalnoj prirodi, logika pruža preciznu analizu samoreferencije koja se u njoj javlja, što omogućava bolje razumevanje toga kako ona funkcioniše, kao i uslova neophodnih da bi se ona uopšte mogla formulisati. Razlikovaćemo dva glavna oblika u kojima se samoreferencija javlja u logici: kao *funkcionalna samoreferencija* i kao *lingvistička samoreferencija*. Svaka od te dve vrste samoreferencije odgovara na neki način *pojmovnoj samoreferenciji*. To potiče od bliske veze koja postoji između nekog pojma i odgovarajuće iskazne funkcije koja definiše njegovu ekstenziju, kao i između pojma i formule ili predikata pomoću kojeg je on izražen u jeziku. Zapravo, kako ćemo pokušati da pokažemo, upravo je veza funkcija i formula sa pojmovima, odnosno činjenica da one imaju neko intenzionalno značenje sadržano u tim pojmovima, ono što motiviše njihovu samoreferenciju. Iz tog razloga bi pojmovna samoreferencija, koja bi trebalo da bude analizirana u budućoj teoriji pojmova, mogla da predstavlja zajednički osnov i opravdanje druge dve vrste samoreferencije.

U nastavku ćemo predstaviti i analizirati, redom, funkcionalnu i lingvističku samoreferenciju kao i njihovu primenu i posledice kojima ona u logici vodi. Posle toga ćemo predstaviti i jedan rezultat koji može da se razume kao da daje njihovu objedinjenu analizu (vidi odeljak 2.3). On će nam omogućiti da izdvojimo zajedničke formalne karakteristike te dve vrste samoreferencije koje, kako ćemo videti, one dele i sa onim tipom samoreferencije koji je u osnovi paradoksa opisanih na početku ovog rada. Razlog zbog kog instance samoreferencije u logičkim teorijama ne vode posledicama sličnim onima koje su u osnovi paradoksa, a koje bi učinile te teorije protivrečnim, sugerisaće način na koji bi takve posledice mogle biti izbegnute i u teoriji koja koristi pojmovnu samoreferenciju. Drugim rečima, naša potraga za adekvatnim rešenjem intenzionalnih paradoksa biće usmerena načinom na koji je protivrečnost izbegnuta u oblastima u kojima se funkcionalna i lingvistička samoreferencija javljaju kao važna metodološka sredstva. Takvo rešenje neće zahtevati odricanje od metodološki korisnih aspekata samoreferencije, već bi trebalo da omogući njihovu konzistentnu primenu. Videćemo na primeru teorija koje ćemo razmatrati, da bi ta primena mogla imati veliki značaj i za buduću teoriju pojmova.

2.1 FUNKCIONALNA SAMOREFERENCIJA

Funkcionalna samoreferencija se pre svega javlja u *teoriji izračunljivosti*. To je oblast logike čiji je jedan od ciljeva, koji je motivisao njeno zasnivanje, formalizacija pojma *izračunljive funkcije*. Ta formalizacija trebalo je da posluži tome da se intuitivni pojam izračunljivosti učini matematički preciznim, kako bi se u njegovom istraživanju mogla koristiti matematička sredstva. U tom smislu, formalizacija bi trebalo da pruži analizu pojma izračunljive funkcije i sa njim povezanih pojmova.

Priroda te oblasti i njenog zadatka zahteva uzimanje u obzir intenzionalnih aspekata funkcije sadržanih u pravilima koja čine *algoritam* pomoću kojeg vrednost te funkcije za određeni argument može biti izračunata. Samim tim, ekstenzionalni pristup koji funkciju poistovećuje sa njenim grafom nije u potpunosti usvojen u ovoj oblasti. Pored grafa funkcije, i algoritam pomoću kojeg je taj graf definisan takođe je smatran njenim bitnim delom.

U ovom kontekstu, samoreferencija se može javiti u obliku funkcije koja se primenjuje na sebe ili u čijoj definiciji se javlja ona sama, tj. čiji algoritam poziva samog sebe u procesu izračunavanja. Te vrste samoreferencije u osnovi su nekih od najznačajnijih metoda konstruisanja ili definisanja funkcija, čiji su najistaknutiji primeri *rekurzija* i *dijagonalizacija*. Poslednja metoda je u toj meri prisutna u ovoj teoriji da se ona nekad naziva i „teorija dijagonalizacije" ("the theory of diagonalization", cf. Rogers, 1987, p. 31).

U nastavku će biti predstavljena određena formalizacija pojma izračunljive funkcije. To je samo jedna od velikog broja međusobno ekvivalentnih formalizacija istog pojma. Ona će nam obezbediti formalnu osnovu i sredstva za analizu različitih oblika samoreferencije koji se javljaju u metodama i rezultatima te i drugih formalnih teorija koje se bave izračunljivim funkcijama. Pre toga, daćemo kratak neformalni opis predmeta i glavnih problema teorije izračunljivosti.

2.1.1 *Algoritmi i izračunljive funkcije*

Pod uticajem skupovno-teorijske orijentacije a pre svega zarad preciznosti, u matematici se pojam funkcije često koristi u ekstenzionalnom smislu. To podrazumeva da se pod funkcijom ne misli na pravilo po kojem se svakom elementu *domena* funkcije pridružuje neki element njenog *kodomena*, već na skup uređenih parova objekata koji na taj način nastaje (potencijalno zajedno sa skupovima kojima ti objekti pripadaju - domenom i kodomenom funkcije). Članovi tih uređenih parova interpretirani su kao *argumenti* funkcije - objekti na koje se ta funkcija primenjuje, i kao njene *vrednosti* za date argumente dobijene kao jedinstveni rezultat te primene. Da bi neki skup uređenih parova bio tako shvaćena funkcija, ne mora biti jasno kako je on napravljen. Može takođe biti slučaj da ne postoji nikakva pravilnost koja bi objasnila pripadnost baš određenih uređenih parova tom skupu. Ako je takva funkcija beskonačni skup, onda je posledica nedostatka te pravilnosti nepostojanje konačne procedure kojom bi njena vrednost za određeni argument mogla biti izračunata. Takva funkcija će biti *neizračunljiva* (*uncomputable*). Sa druge strane, funkcija za čije građenje postoji konačna *efektivna procedura* ili *algoritam*, zvaće se *izračunljiva funkcija* (*computable function*). Tako će izračunljiva biti svaka konačna funkcija, kao i ona beskonačna funkcija čiji su elementi određeni u skladu sa nekim pravilom. Ta pravila mogu biti formulisana u obliku instrukcija koje formiraju algoritam za izračunavanje te funkcije. Kada odgovarajući algoritam dobije kao *ulaz* (*input*) neki argument funkcije, izvode se određene operacije prema u njemu sadržanim instrukcijama i kao *izlaz* (*output*) algoritma dobija se vrednost te funkcije za dati argument. Jedan algoritam može činiti samo konačno mnogo instrukcija i one moraju biti precizne i dovoljno jednostavne kako bi mogle da se slede mehanički. Instrukcije treba da odrede izračunavanje funkcije korak po korak i tako da rezultat svakog koraka u izračunavanju bude predvidiv. Ti, i možda neki dodatni, uslovi moraju da budu zadovoljeni kako bi izračunavanje funkcije pomoću određenog algoritma bilo efektivno (cf. Enderton, 1977, p. 529).

Funkcije kojima se teorija izračunljivosti prvenstveno bavi su one koje se primenjuju na prirodne brojeve ili njihove n-torke i kao vrednosti daju takođe prirodne brojeve.[1] Time ova teorija ne gubi na opštosti jer su druge funkcije predstavljive funkcijama na prirodnim brojevima zahvaljujući tome što su njihovi nenumerički argumenti i vrednosti predstavljivi tim brojevima pomoću neke *procedure kodiranja*. Na njih se zato takođe odnose pojmovi teorije izračunljivosti, kao i tvrđenja o funkcijama do kojih ona dolazi. Skup argumenata funkcije za koje ona daje neku vrednost naziva se njenim *domenom*, a skup njenih vrednosti njenom *slikom*. Kaže se da je funkcija *definisana* za objekte iz njenog domena, a da je *nedefinisana* za objekte van tog domena. Ako $\mathbb{N}$ stoji za skup prirodnih brojeva, a $\mathbb{N}^k$ je Dekartov proizvod $\mathbb{N} \times \mathbb{N} \times ... \times \mathbb{N}$ sa k faktora (za $k \geq 1$), onda su funkcije koje nas zanimaju one čiji je domen neki podskup skupa $\mathbb{N}^k$, a slika neki podskup skupa $\mathbb{N}$. Nadalje ćemo često govoriti o funkcijama jednog argumenta kada ono što je za njih rečeno može lako da se uopšti tako da važi i za druge funkcije. Ako je domen funkcije ceo skup $\mathbb{N}^k$, kaže se da je ta funkcija *totalna* na tom skupu, dok ako je njen domen neki podskup skupa $\mathbb{N}^k$, onda je funkcija *parcijalna* na skupu $\mathbb{N}^k$. Skup parcijalnih funkcija tako uključuje i totalne funkcije kao one parcijalne funkcije koje su definisane na celom skupu.

Kada je neki ulaz dat algoritmu, započinje izračunavanje koje se u nekom trenutku završava i daje određeni izlaz, ili se neodređeno produžava bez davanja bilo kakvog izlaza (to se dešava, na primer, kada izračunavanje sadrži neku beskonačnu petlju). Ako se izračunavanje prema nekom algoritmu završava i daje izlaz za svaki element datog skupa, onda taj algoritam određuje jednu totalnu funkciju na tom skupu. Ako se za neki element skupa izračunavanje ne završava i ne daje izlaz, onda algoritam određuje neku parcijalnu funkciju na tom skupu koja nije definisana za taj element. Za parcijalne funkcije f i g, jednakost $f(x_1, ..., x_n) = g(x_1, ..., x_n)$ je interpretirana kao tvrđenje da ako je jedna od te dve funkcije definisana za argumente $x_1, ..., x_n$ onda je i druga definisana za te argumente i njihove vrednosti se slažu.

Posebnu vrstu funkcija predstavljaju one čiji je kodomen skup $\{0, 1\}$. Vrednosti takvih funkcija, koje se nazivaju *iskaznim funkcijama*, interpretirane su kao istinosne vrednosti - istina i laž. Svaka iskazna funkcija je *karakteristična funkcija* skupa koji čine tačno oni brojevi ili njihove n-torke za koje funkcija daje vrednost 0 koja stoji za istinu (običaj da se 0 interpretira kao istina, a 1 kao laž ne slaže se sa onim u logici, ali je u ovoj oblasti čvrsto ustanovljen). Nešto formalnije rečeno, funkcija c_A je karakteristična funkcija skupa A ako važi: $c_A(x) = 0$ ako $x \in A$, i $c_A(x) = 1$ ako $x \notin A$. Ako je karakteristična funkcija skupa A izračunljiva, onda se za skup A kaže da je *odlučiv* ili *izračunljiv*, što znači da postoji efektivna procedura koja za svaki objekat odlučuje da li je on u skupu ili ne. Ta procedura sastoji se u izračunavanju vrednosti funkcije c_A za taj objekat i utvrđivanju toga da li je rezultat izračunavanja 0, u kom slučaju objekat pripada datom skupu ili 1, u kom slučaju mu ne pripada.

Postoji i jedno slabije svojstvo skupova koje se tiče njihove izračunljivosti. To svojstvo imaju oni skupovi za koje postoji efektivna procedura pomoću koje je moguće za svaki element skupa proveriti da je on u skupu, ali koja ne može u opštem slučaju reći za objekte van skupa da nisu njegovi elementi. Skupovi sa tim svojstvom biće tačno oni koji predstavljaju sliku neke izračunljive funkcije. Takvi skupovi zovu se *semi-odlučivi* kao i *efektivno nabrojivi* skupovi. Izračunljiva funkcija čija je slika određeni skup može služiti za nabrajanje elemenata tog skupa kojim se pravi njihova lista. Za svaki element takvog skupa moguće je tada utvrditi da on jeste u skupu pronalaženjem tog elementa na datoj listi. Tom procedurom nije moguće u opštem slučaju

1 Nadalje, kada govorimo o *brojevima* kao argumentima i vrednostima izračunljivih funkcija, pod tim mislimo na *prirodne brojeve*.

utvrditi da neki objekat nije u skupu jer ako je lista beskonačna, procedura traženja objekta koji se u njoj ne nalazi se nikada neće završiti. Pošto je svaka *relacija* na skupu $\mathbb{N}$ neki skup brojeva ili njihovih uređenih n-torki, pojam odlučivosti ili semi-odlučivosti se u istom značenju primenjuje i na relacije. Isti pojam se može primeniti i na predikate, pri čemu bi neki predikat bio odlučiv ili semi-odlučiv ako i samo ako je takva relacija koju on označava. Napomenimo još da su semi-odlučivi skupovi zatvoreni za primenu unije, preseka i Dekartovog proizvoda, dok su odlučivi skupovi osim toga zatvoreni i za primenu komplementacije.

Vezu između dva svojstva skupa opisuje teorema koja tvrdi da je neki skup odlučiv ako i samo ako su i on i njegov komplement semi-odlučivi skupovi (vidi, na primer, Rogers, 1987, p. 58). Veoma značajan rezultat teorije izračunljivosti koji ima važne posledice, nekima od kojih ćemo se ovde baviti, jeste da postoje semi-odlučivi skupovi koji nisu odlučivi. Pitanje odlučivosti nekog skupa ili relacije blisko je povezano sa pitanjem *rešivosti* nekog *problema odlučivosti* (*decision problem*). Takav problem tiče se pitanja o tome da li neki objekat poseduje određeno svojstvo (tj. pripada nekom skupu) ili stoji u određenoj relaciji sa drugim objektima. Ako su skup ili relacija o kojima je reč odlučivi, onda je problem koji se njih tiče *rešiv* (*solvable*); inače je *nerešiv* (*unsolvable*). Na primer, problem koji se tiče toga da li je neki broj prost je rešiv pošto je skup prostih brojeva odlučiv. To sledi iz činjenice da postoje algoritmi (takozvani *primality tests*) pomoću kojih se za svaki broj može utvrditi da li je on prost ili ne. Kako u tom slučaju, tako i uopšte, za problem koji se tiče pripadnosti nekom skupu možemo dokazati da je rešiv pronalaženjem algoritma koji izračunava taj skup. Sa druge strane, kako bismo utvrdili da je neki problem nerešiv, trebalo bi da pokažemo da takav algoritam ne postoji, tj. da karakteristična funkcija datog skupa ne može biti izračunljiva. To je, u opštem slučaju, dosta teže postići.

Ono što, međutim, unapred znamo jeste da većina funkcija zapravo nije izračunljiva. To sledi iz toga što postoji neprebrojivo mnogo funkcija, a samo prebrojivo mnogo konačnih instrukcija koje mogu formirati algoritme za njihovo izračunavanje. Skup totalnih *unarnih* funkcija na skupu $\mathbb{N}$, čija je kardinalnost $Card\,(\mathbb{N})^{Card(\mathbb{N})}$, već je reda veličine kontinuuma, dok skup svih algoritama može biti najviše reda veličine skupa prirodnih brojeva. Neke karakteristične funkcije skupova prirodnih brojeva takođe će biti neizračunljive (zapravo neprebrojivo mnogo njih, jer toliko ima i takvih funkcija: $2^{Card(\mathbb{N})}$), što znači da će problemi odlučivosti koji se tiču svih tih skupova biti nerešivi.

Različiti problemi su često međusobno uporedivi, i *svodivi* jedni na druge. To da se problem P_1 svodi na problem P_2 znači da rešenje problema P_2 omogućava rešenje problema P_1. Samim tim, ako je P_1 nerešiv problem, P_2 mora takođe biti nerešiv. Ako se problem P_1 svodi na problem P_2, ali ne i obrnuto, onda kažemo da je P_2 problem višeg *stepena nerešivosti*. Ako se svaki od dva problema svodi na drugi, onda su istog stepena nerešivosti. Nerešivost mnogih problema odlučivosti tako sledi iz nerešivosti *halting problema* koji ćemo uskoro predstaviti. Ta činjenica motivisala je jedno od pitanja koje je iniciralo istraživanje o nerešivim problemima i njihovom međusobnom odnosu: da li svi nerešivi problemi na neki način odslikavaju nerešivost halting problema (vidi, na primer: Post, 1944). Ono što je posebno zanimljivo u vezi sa tim problemom iz naše perspektive, jeste da u dokazu njegove nerešivosti glavnu ulogu ima jedna vrsta samoreferencije.

Iako već znamo da mnoge funkcije nisu izračunljive, kako bismo mogli da takve funkcije izdvojimo i da ih razlikujemo od izračunljivih funkcija, proces izračunavanja i instrukcije koje čine algoritam prema kojem se ono izvodi moraju biti precizno određeni. To se postiže različitim formalizacijama pojma izračunljive funkcije koje rezultiraju matematičkim modelima intuitivnog pojma izračunljivosti.

2.1.2 *Formalizacija pojma izračunljive funkcije*

Da bi formalno ispitivanje izračunljivih funkcija bilo moguće, potreban je precizan matematički pojam koji odgovara intuitivnom pojmu algoritma, kao i odgovarajući matematički model izračunavanja koji nam omogućava da precizno razlikujemo funkcije izračunljive pomoću nekog algoritma od onih za koje takav algoritam ne postoji. Pošto to može da se postigne na različite načine, kao rezultat dobijamo različite formalne modele izračunljivosti. Neki od njih opisuju izračunavanje kao proces koji se odvija na apstraktnoj mašini praćenjem određenih jednostavnih instrukcija za manipulisanje simbolima koji predstavljaju ulaz algoritma. Izračunljive funkcije su u toj formalizaciji predstavljene onim funkcijama za koje postoji mašina koja ih izračunava. Najznačajnije formalizacije te vrste su *Tjuringove mašine* (*Turing machines*) i *Postove mašine* (*Post machines*). U drugim formalizacijama, izračunljive funkcije su one koje mogu biti definisane unutar određenog formalnog sistema, kao što je Čerčov *λ-račun*. Algoritmi za izračunavanje neke funkcije mogu, osim toga, biti predstavljeni i jednakostima određene vrste koje pokazuju kako se vrednost te funkcije za date argumente može dobiti od vrednosti nekih jednostavnijih funkcija za iste argumente, ili od vrednosti iste funkcije za argumente koji na neki način *prethode* tim argumentima. Tako definisane funkcije, koje odgovaraju izračunljivim funkcijama u ovoj formalizaciji, nazivaju se *parcijalno rekurzivnim funkcijama*.

Jedan od najznačajnijih rezultata teorije izračunljivosti jeste dokaz da su sve navedene formalizacije intuitivnog pojma izračunljive funkcije međusobno ekvivalentne. To znači da njima ponuđeni kriterijumi izračunljivosti određuju istu klasu aritmetičkih funkcija kao izračunljivih. Na koji god od ponuđenih načina pokušamo da preciziramo taj intuitivni pojam, iste funkcije će ispasti izračunljive. Osim toga, takođe važi i da se za reprezentaciju proizvoljne izračunljive funkcije u jednom modelu izračunljivosti može efektivno naći reprezentacija iste funkcije u bilo kom od njemu ekvivalentnih modela izračunljivosti. Zato se za te modele kaže da su *algoritmički ekvivalentni*. Ovi rezultati govore u prilog tome da navedene formalizacije nisu proizvoljne, već da zaista uspevaju da odslikaju neki deo prirode izračunljivih funkcija. To predstavlja značajno svedočanstvo u prilog *Čerč-Tjuringovoj tezi* prema kojoj su te formalizacije formalni ekvivalent intuitivnom pojmu izračunljivosti. Prema toj tezi, svaka funkcija koja je intuitivno izračunljiva takođe je izračunljiva na osnovu neke od ekvivalentnih formalizacija pojma izračunljive funkcije. U nastavku ćemo detaljnije opisani jednu od pomenutih formalizacija - teoriju parcijalno rekurzivnih funkcija.

2.1.2.1 *Parcijalno rekurzivne funkcije*

U ovoj formalizaciji, izračunljive funkcije predstavljene su takozvanim *rekurzivnim funkcijama* zadatim jednakostima koje opisuju način izračunavanja njihovih vrednosti za određene argumente korišćenjem njihovih drugih vrednosti ili vrednosti drugih, jednostavnijih funkcija. U karakterizaciji tih funkcija, *rekurzivne definicije* imaju centralnu ulogu. Rekurzivna definicija opisuje način na koji se vrednost funkcije za neki argument može odrediti na osnovu vrednosti iste funkcije za argumente koji mu na neki način prethode. Takve definicije se, dakle, mogu odnositi samo na funkcije čiji su argumenti uređeni na način koji omogućava da neki od tih argumenata prethode drugima. Drugim rečima, argumenti takve funkcije moraju pripadati skupu na kojem je definisano neko *dobro uređenje*. Skup A je dobro uređen relacijom R ako i samo ako je R neko striktno linearno uređenje[2] na skupu A i svaki neprazan podskup skupa A ima R-najmanji ili R-minimalan element (odnosno, element koji je u relaciji R prema svakom drugom elementu istog podskupa, dok nijedan element tog podskupa nije u toj relaciji prema njemu). Elementi skupa A su linearno uređeni počevši od R-najmanjeg elementa, tako da elementi koji

2 Relacija je striktno uređenje ne nekom skupu ako je irefleksivna i tranzitivna; takva relacija je linearno uređenje ako za svaka dva elementa tog skupa važi da je ili prvi u toj relaciji sa drugim ili drugi u toj relaciji sa prvim.

su u relaciji R prema drugim elementima skupa prethode tim elementima u uređenju. Skup prirodnih brojeva je dobro uređen relacijom $<$ u njenoj standardnoj interpretaciji, zahvaljujući čemu su i rekurzivne definicije funkcija na tom skupu moguće.

Pojam rekurzivne funkcije kao formalnog ekvivalenta pojma izračunljive funkcije predložen je od strane Gedela (Gödel, 1934), koji je pratio sugestiju Herbranda (Jacques Herbrand), dok je svoj najrazvijeniji oblik dobio zahvaljujući Kliniju (Stephen Cole Kleene). U toj formalizaciji, definicije izračunljivih funkcija date su u obliku jednakosti koje služe kao algoritmi za njihovo izračunavanje. U Klinijevoj formulaciji (Kleene, 1936) sve definicije rekurzivnih funkcija dobijene su od nekih osnovnih definicija koje određuju *inicijalne rekurzivne funkcije*, primenom određenih metoda građenja novih rekurzivnih funkcija od već postojećih. Obično se inicijalnim ili osnovnim rekurzivnim funkcijama smatraju sledeće funkcije:

- *Nula funkcija Z:* unarna funkcija koja za svaki argument daje vrednost 0, tj. za svako x važi $Z(x) = 0$;

- *Funkcija sledbenika* S: unarna funkcija koja za svaki argument daje njegovog sledbenika kao vrednost, tj. za svako x važi $S(x) = x + 1$;

- *Funkcije projekcije* p^n_k: za svako $n \geq 1$ i k takvo da $1 \leq k \leq n$, postoji funkcija p^n_k koja, kada se primeni na niz od n brojeva, daje broj koji je na k-toj poziciji u tom nizu kao vrednost, tj. za sve brojeve $x_1, ..., x_n$, važi $p^n_k(x_1, ..., x_n) = x_k$.

Poslednja klauzula određuje beskonačno mnogo rekurzivnih funkcija dobijenih supstitucijom konkretnih prirodnih brojeva mesto n i k.

Osim inicijalnih funkcija, rekurzivne funkcije su i sve one koje se od njih mogu dobiti primenom sledećih metoda:

- *Kompozicija:* ako je f neka n-arna funkcija, a $g_1, ..., g_n$ su sve m-arne funkcije, onda se pomoću njih može definisati nova m-arna funkcija h takva da: $h(x_1, ..., x_m) = f(g_1(x_1, ..., x_m), ..., g_n(x_1, ..., x_m))$;

- *Rekurzija:* ako je f neka n-arna funkcija i g neka $(n+2)$-arna funkcija, onda se pomoću njih može definisati nova $(n+1)$-arna funkcija h takva da:
  $h(0, x_1, ..., x_n) = f(x_1, ..., x_n)$
  $h(y+1, x_1, ..., x_n) = g(y, h(y, x_1, ..., x_n), x_1, ..., x_n)$.

Definicija neke funkcije rekurzijom pokazuje kako se njena vrednost za određeni argument može odrediti pomoću njene vrednosti za prethodnika tog argumenta (na koji tačno način, zavisi od funkcije g). Funkcija h definisana pomoću rekurzije se ponovo javlja (*recurs*) unutar sopstvene definicije, odakle i potiče naziv tih funkcija. Proces izračunavanja vrednosti funkcije h za neki argument sastoji se od izračunavanja njene vrednosti za sve manje argumente počevši od najmanjeg i primenjujući funkciju g na prethodno dobijenu vrednost, sve dok se time ne dođe do datog argumenta. Vrednost funkcije h za najmanji argument mora biti nezavisno određena prvom klauzulom definicije, kako bi proces izračunavanja uopšte mogao da započne. Algoritam za izračunavanje funkcije definisane rekurzijom poziva samog sebe u procesu izračunavanja, što ga čini samoreferencijalnim. Pored navedene vrste rekurzije, koja se naziva i *primitivnom rekurzijom*, postoje i druge vrste rekurzije koje sadrže istu vrstu samoreferencije. To su *rekurzija po nizu vrednosti* (*course of values recursion*) kojom se vrednost funkcije za dati argument određuje pomoću njenih vrednosti za više argumenata koji mu prethode ili čak sve takve argumente; *dvostruka rekurzija* (*double recursion*) koja sadrži rekurziju po dva argumenta istovremeno, i slično. Svaka funkcija definisana pomoću tih metoda može se dobiti pomoću primitivne rekurzije i ostalih metoda teorije rekurzivnih funkcija.

Funkcije koje se mogu konstruisati od inicijalnih rekurzivnih funkcija primenom navedenih metoda zovu se *primitivno rekurzivne funkcije*. Sve primitivno rekurzivne funkcije su totalne. To sledi iz činjenice da su inicijalne funkcije totalne i da operacije kompozicije i rekurzije kojima su ostale primitivno rekurzivne funkcije definisane čuvaju to svojstvo. Gotovo sve intuitivno izračunljive funkcije jesu primitivno rekurzivne. Međutim, poznati su i određeni izuzeci kao što je *Akermanova funkcija* (*Ackermann function*), prvobitno definisana dvostrukom rekurzijom, kojoj ne odgovara nijedna primitivno rekurzivna funkcija. Ti izuzeci pokazuju da primitivno rekurzivne funkcije ne predstavljaju sasvim odgovarajuću formalizaciju izračunljivih funkcija, i da bi, kako bi ona postala odgovarajuća, ta formalizacija morala da bude proširena nekim dodatnim načinom definisanja funkcija. Najčešće se to postiže uvođenjem metode *minimizacije* čijom se primenom grade funkcije koje proširuju klasu primitivno rekurzivnih funkcija i koje, takođe, nisu nužno totalne.

- *Minimizacija*: ako je f neka $(n+1)$-arna funkcija, onda nova n-arna funkcija h može biti definisana na sledeći način:

$$h(x_1,...,x_n) = \mu y\,[f(x_1,...,x_n,y)=0] = \begin{cases} \text{najmanje } y \text{ za koje važi} \\ f(x_1,...,x_n,y)=0 \text{ i} \\ \text{za svako } y' < y, \\ f(x_1,...,x_n,y') \text{ je definisano i} \\ \text{različito od } 0, \\ \text{ako postoji takvo } y; \\ \\ \text{nedefinisano, inače.} \end{cases}$$

Nova metoda definisanja rekurzivnih funkcija koristi takozvani *operator minimizacije* - μ. Procedura izračunavanja funkcije h definisane minimizacijom sastoji se od izračunavanja vrednosti $f(x_1,...,x_n,0)$, $f(x_1,...,x_n,1)$, $f(x_1,...,x_n,2)$ itd., koje se, u slučaju da je svaka od njih izračunljiva, nastavlja sve dok se ne stigne do broja y za koji važi $f(x_1,...,x_n,y)=0$. Kada se stigne do prvog takvog y, izračunavanje funkcije h se završava i daje to y kao njenu vrednost. Funkcija h neće biti definisana za brojeve $x_1,...,x_n$ ako takvo y ne postoji ili ako je za neko $y' < y$, $f(x_1,...,x_n,y')$ nedefinisano.

U slučaju da je funkcija f koja se koristi u definiciji funkcije h minimizacijom karakteristična funkcija neke relacije R, definiciju funkcije h ćemo pisati na sledeći način:

$$h(x_1,...,x_n) = \mu y R(x_1,...,x_n,y) = \begin{cases} \text{najmanje } y \text{ za koje važi} \\ R(x_1,...,x_n,y)\,, \text{ ako takvo } y \text{ postoji;} \\ \\ \text{nedefinisano, inače} \end{cases}$$

gde $R(x_1,...,x_n,y)$ pišemo umesto $(x_1,...,x_n,y) \in R$.

Sve funkcije koje se mogu dobiti od inicijalnih rekurzivnih funkcija primenom kompozicije, rekurzije i minimizacije nazivaju se *parcijalno rekurzivnim*, *μ-rekurzivnim* ili prosto *rekurzivnim funkcijama*. Osim za funkcije, za skupove i relacije se takođe može reći da su primitivno ili parcijalno rekurzivni pod čime se podrazumeva da su takve njihove karakteristične funkcije. Rekurzivni skupovi i relacije bi trebalo da predstavljaju formalni ekvivalent odlučivih ili izračunljivih skupova i relacija. Problemi odlučivosti koji se njih tiču biće *rekurzivno rešivi*. Skupovi i relacije koji predstavljaju sliku neke rekurzivne funkcije, a koji se nazivaju *rekurzivno nabrojivim* skupovima i relacijama, predstavljaju, sa druge strane, formalizaciju semi-odlučivih skupova i relacija.

Metode konstruisanja rekurzivnih funkcija mogu se takođe predstaviti kao *funkcionali*, odnosno, kao funkcije drugog reda koje se primenjuju na rekurzivne funkcije i kao rezultat daju novu rekurzivnu funkciju. U toj prezentaciji, inicijalne funkcije bili bi nularni funkcionali S, Z and p_k^n koji su i sami rekurzivne funkcije. Kompozicija bi bila $(n+1)$-arni funkcional koji, kada se primeni na n-arnu funkciju f i n-torku m-arnih funkcija g_1,...,g_n, daje novu m-arnu funkciju $C_m^n\left(f, g_1, ..., g_n\right)$. Rekurzija bi bila binarni funkcional R^n koji kada se primeni na n-arnu funkciju f i na $(n+2)$-arnu funkciju g, daje novu $(n+1)$-arnu funkciju $R^n\left(f, g\right)$. Konačno, minimizacija bi bila unarni funkcional M^n koji kada se primeni na $(n+1)$-arnu funkciju f daje novu n-arnu funkciju $M^n\left(f\right)$. Klasa rekurzivnih funkcija definisana je tada kao najmanja klasa koja sadrži inicijalne funkcije i koja je zatvorena za primenu funkcionala C_m^n, R^n i M^n. Drugim rečima, funkcija f takva da je $f = F\left(f_1, ..., f_n\right)$, gde su $f_1, ..., f_n$ neke rekurzivne funkcije, je rekurzivna ako i samo ako je F rekurzivni funkcional, što znači da ga je moguće definisati pomoću inicijalnih funkcija S, Z, p_k^n i funkcionala C_m^n, R^n i M^n. Vrednost funkcije definisane pomoću n-arnog funkcionala F za argumente x_1,...,x_k označavaćemo sa $F\left(f_1, ..., f_n; x_1, ..., x_k\right)$.

Terminološka napomena: Nadalje ćemo koristiti mala slova grčkog alfabeta kao što su φ, ψ i χ da označimo proizvoljnu rekurzivnu funkciju, dok ćemo za funkcije koje nisu nužno rekurzivne nastaviti da koristimo slova f, g i h.

2.1.2.2 *Formalna teorija rekurzivnih funkcija*

Definicije rekurzivnih funkcija služe kao algoritmi za izračunavanje tih funkcija: one određuju, korak po korak, proces izračunavanja vrednosti definisane funkcije za date argumente. Način na koji su rekurzivne funkcije definisane i način na koji njihova definicija određuje proceduru njihovog izračunavanja mogu biti predstavljeni unutar formalne teorije (Kleene, 1936). Jezik te formalne teorije bi od nelogičkih simbola sadržao numerale i konstante za inicijalne funkcije. Njeni termi bili bi interpretirani kao imena za prirodne brojeve, dok bi jednakosti između tih terama činile njene atomske formule. Ta teorija bi takođe trebalo da sadrži pravila na osnovu kojih neka jednakost sledi iz drugih, čijom primenom na pretpostavljene jednakosti bi bila građena izvođenja.

Definicija neke m-arne rekurzivne funkcije h može se predstaviti kao niz funkcija koje nastaju u procesu njenog građenja počevši od inicijalnih funkcija primenom kompozicije, rekurzije i minimizacije na već izgrađene funkcije (cf. Kleene, 1971, p. 221). Funkcije iz tog niza mogu biti zadate njihovim rekurzivnim definicijama, tj. jednakostima koje pokazuju kako se njihove vrednosti za proizvoljne argumente mogu izračunati. Te jednakosti će činiti *jednakosni sistem E* koji definiše funkciju h i sadrži instrukcije za njeno izračunavanje. Taj jednakosni sistem može biti predstavljen unutar formalne teorije prevođenjem jednakosti koje ga čine na formalni jezik. Izračunavanje vrednosti funkcije h za argumente $n_1, ..., n_m$ određeno njenom definicijom bilo bi tada predstavljivo unutar iste teorije nekim nizom jednakosti koje formiraju izvođenje. To izvođenje se završava jednakošću oblika $\mathbf{h}\left(\mathbf{n_1}, ..., \mathbf{n_m}\right) = \mathbf{k}$, gde su $\mathbf{n_1}, ..., \mathbf{n_m}$ i $\mathbf{k}$ numerali koji označavaju prirodne brojeve n_1,...,n_m, k, a $\mathbf{h}$ je funkcijski simbol kojim se funkcija h označava[3]. Ono počinje jednakostima iz E koje definišu funkciju h, a svaka naredna jednakost dobijena je od prethodnih na neki od sledeća dva načina:

1. instancijacijom neke jednakosti, tj. supstitucijom određenih numerala mesto njenih slobodnih promenljivih;

3 Podebljana slova stoje za simbole nekog formalnog sistema, dok slova u kurzivu stoje za njima označene objekte.

2. zamenom nekog javljanja terma u datoj jednakosti njemu jednakim termom (pod uslovom da je njihova jednakost prethodno dokazana).

Funkcija $h : \mathbb{N}^m \to \mathbb{N}$ je rekurzivna prema ovoj formalnoj teoriji ako i samo ako postoji neki jednakosni sistem te teorije koji je definiše, odnosno, iz kog je jednakost $\mathbf{h}\left(\mathbf{n_1}, ..., \mathbf{n_m}\right) = \mathbf{k}$ izvodiva pomoću navedenih pravila ako i samo ako važi $h\left(n_1, ..., n_m\right) = k$. Ako je teorija potpuna u odnosu na klasu rekurzivnih funkcija, onda za svaku rekurzivnu funkciju postoji takav niz jednakosti koji je definiše unutar te teorije.

2.1.2.3 *Kodiranje jezika formalne teorije rekurzivnih funkcija*

Gedel je pokazao da jezik svake formalne teorije može biti predstavljen nekim skupom prirodnih brojeva pomoću metode koja se naziva *kodiranjem* ili *gedelizacijom*. Kodiranje jezika neke teorije sastoji se u kodiranju njenih sintaktičkih objekata - simbola, nizova simbola koji čine terme i formule, i nizova formula koji čine izvođenja. To se postiže ustanovljavanjem precizne korespondencije između tih sintaktičkih objekata i određenih prirodnih brojeva. Na taj način je svakom sintaktičkom objektu datog sistema pripisan jedinstveni prirodni broj i to različiti prirodni brojevi različitim sintaktičkim objektima. Broj pripisan sintaktičkom objektu naziva se njegovim *kôdom* ili *Gedelovim brojem*. Kodiranje može biti izvedeno na različite načine sve dok je funkcija u njegovoj osnovi, ona koja određuje korespondenciju između jezika i prirodnih brojeva, izračunljiva. To znači da mora postojati algoritam za izračunavanje kôda proizvoljnog izraza ili izvođenja u sistemu, kao i algoritam koji za svaki prirodan broj može odlučiti da li je on kôd nekog sintaktičkog objekta tog sistema, i ako jeste, kog.

Kao što smo već rekli, algoritam koji odgovara nekoj rekurzivnoj funkciji zadat je unutar formalne teorije nekim sistemom jednakosti koje predstavljaju formule te teorije. Izračunavanje vrednosti funkcije na osnovu tog algoritma predstavljeno je nekim izvođenjem koje započinje jednakostima iz tog sistema, a završava se jednakošću koja pokazuje vrednost funkcije za date argumente. Iz toga sledi da kodiranjem formula i izvođenja datog formalnog sistema istovremeno dobijamo i način na koji možemo kodirati algoritme rekurzivnih funkcija i njihovo izračunavanje predstavljeno u tom sistemu. Kodiranje algoritama i izračunavanja funkcija otvara vrata mnogim značajnim rezultatima teorije rekurzija. Slično se može postići i u drugim formalizacijama izračunljivih funkcija, što čini te rezultate nezavisnim od pretpostavljene formalizacije izračunljivosti. Neki od najznačajnijih rezultata koji se neposredno tiču mogućnosti samoreferencije u vezi rekurzivnih funkcija biće predstavljeni u nastavku.

2.1.2.4 *Klinijeve teoreme o normalnoj formi i o enumeraciji*

Iz prethodno rečenog sledi da svaka rekurzivna funkcija može biti definisana unutar odgovarajuće teorije rekurzivnih funkcija nekim jednakosnim sistemom koji omogućava njeno izračunavanje (tj. izvođenje jednakosti iz tog sistema koje daju njene vrednosti za odgovarajuće argumente). Kodiranje jednakosti i izvođenja datog formalnog sistema omogućava formulisanje tog tvrđenja u obliku teoreme kojom se tvrdi postojanje normalne forme rekurzivnih funkcija. Da bi tu teoremu formulisao, Klini je uveo poseban metamatematički predikat T takav da $T\left(e, x_1, ..., x_n, y\right)$ važi ako i samo ako je y kôd izvođenja koje započinje jednakosnim sistemom kodiranim brojem e, za brojeve $x_1, ..., x_n$. Predikat T može biti interpretiran kao ternarna relacija između algoritma rekurzivne funkcije, nekog niza brojeva i procesa izračunavanja vrednosti te funkcije za date brojeve. Pošto bi svaku rekurzivnu funkciju trebalo da bude moguće definisati unutar odgovarajuće formalne teorije, takođe bi trebalo da važi da za svaku rekurzivnu funkciju postoji kôd njene definicije u toj teoriji koja određuje izračunavanje njene vrednosti za date argumente. Drugim rečima, za svaku n-arnu rekurzivnu funkciju φ, trebalo bi da postoji neko e tako da $\exists y T\left(e, x_1, ..., x_n, y\right)$ važi za sve i samo one brojeve $x_1, ..., x_n$ za koje je funkcija φ definisana.

Izračunavanje funkcije određeno njenom rekurzivnom definicijom nije nužno jedinstveno. Različita izračunavanja iste funkcije bila bi predstavljena unutar formalne teorije izvođenjima koja počinju istim jednakosnim sistemom i završavaju se istom jednakošću dok se proces izvođenja te jednakosti iz datog sistema razlikuje (na primer, po redosledu primene pravila izvođenja). Različita izvođenja imaće različite kôdove. Ako želimo da izaberemo jedan od njih, možemo to učiniti primenom operatora minimizacije. Na taj način bio bi formiran term $\mu y T\left(e, x_1, ..., x_n, y\right)$ koji označava najmanji kôd izračunavanja određenog definicijom funkcije za brojeve $x_1, ..., x_n$.

Da bi neka rekurzivna funkcija bila adekvatno predstavljena unutar formalne teorije, nije dovoljno da postoji neko izračunavanje njene vrednosti za sve brojeve za koje je definisana. Osim toga, potrebno je i da to izračunavanje bude korektno u smislu da daje sve i samo vrednosti te funkcije za date argumente. Kako bi formalno izrazio taj uslov, Klini je uveo funkciju u koja, kada joj je dat kôd nekog konačnog izračunavanja, vraća njegov rezultat. Taj rezultat izračunavanja trebalo bi da predstavlja vrednost funkcije za date argumente.

Pomoću uvedenih sredstava možemo formulisati Klinijevu TEOREMU O NORMALNOJ FORMI koja, može se reći, u formalnim terminima objašnjava šta znači biti rekurzivna funkcija. Teorema tvrdi da za svaku n-arnu rekurzivnu funkciju φ, postoji broj e, takav da:

$$\varphi\left(x_1, ..., x_n\right) = u\left(\mu y T\left(e, x_1, ..., x_n, y\right)\right)$$

za svako $x_1, ..., x_n$ (Kleene, 1936). Ova teorema u suštini tvrdi da svaka rekurzivna funkcija poseduje rekurzivnu definiciju unutar odgovarajuće formalne teorije koja služi kao algoritam za njeno izračunavanje. Kaže se da kôd te definicije, broj e, *rekurzivno definiše* funkciju φ. Rekurzivna funkcija može imati beskonačno mnogo rekurzivnih definicija pošto svaka od njih može biti modifikovana na način koji ne utiče na njene vrednosti. Na primer, svaka od njih može biti proširena kompozicijom sa identičkom funkcijom ili instrukcijom prema kojoj dobijenoj vrednosti treba dodati a zatim oduzeti isti broj, i na slične načine.

Pošto su relacija označena predikatom T i funkcija u rekurzivne, $u\left(\mu y T\left(e, x_1, ..., x_n, y\right)\right)$ će takođe biti rekurzivna funkcija sa $n+1$ argumenata. Ta funkcija se obično označava simbolom Φ_n. Ako preformulišemo TEOREMU O NORMALNOJ FORMI koristeći taj simbol, dobićemo da se njom tvrdi da za svaku n-arnu rekurzivnu funkciju φ, postoji broj e, takav da:

$$\varphi\left(x_1, ..., x_n\right) = \Phi_n\left(e, x_1, ..., x_n\right)$$

za svako $x_1, ..., x_n$. Pomoću kôda algoritma neke rekurzivne funkcije i njenih argumenata, Φ_n izračunava vrednost te funkcije za date argumente. Pošto na taj način može biti korišćena za izračunavanje proizvoljne rekurzivne funkcije, funkcija Φ_n se naziva *univerzalnom funkcijom* (ona odgovara *univerzalnoj Tjuringovoj mašini*). Primenom univerzalne funkcije na brojeve $0, 1, 2, 3, ...$ supstituisane mesto e, može se formirati lista rekurzivnih funkcija čiji su algoritmi kodirani tim brojevima. Naime, $\Phi_n\left(0, x_1, ..., x_n\right)$ će stajati za rekurzivnu funkciju čiji je algoritam kodiran brojem 0; $\Phi_n\left(1, x_1, ..., x_n\right)$ za rekurzivnu funkciju čiji je algoritam kodiran brojem 1; i slično za ostale prirodne brojeve. Na taj način, TEOREMA O NORMALNOJ FORMI implicira i Klinijevu TEOREMU O ENUMERACIJI koja tvrdi postojanje funkcije koja pobrojava sve rekurzivne funkcije.

Brojevi koje rekurzivne funkcije dobijaju u tom pobrojavanju mogu se koristiti kao indeksi funkcijskih simbola koji označavaju te funkcije. Tako bi funkcija φ korišćena u formulaciji teoreme, koja bi se javila na e-tom mestu u listi funkcija, dobila indeks e, pa bismo je umesto sa φ označavali sa φ_e.

2.1.2.5 *Operacije na indeksima rekurzivnih funkcija*

Indeksi rekurzivnih funkcija dobijeni na gore opisan način pokazuju njihov konstruktivni karakter samim tim što kodiraju način na koji se one mogu izračunati. Iz tog razloga bi efektivne operacije na tim funkcijama trebalo da budu predstavljive operacijama na njihovim indeksima (Enderton, 1977, p. 561). Takođe bi trebalo da je moguće naći indeks funkcije dobijene kao rezultat primene tih operacija, na osnovu indekasa funkcija na koje se te operacije primenjuju. Jedna važna teorema kojom se tvrdi takva mogućnost u slučaju određene operacije na rekurzivnim funkcijama je takozvana S-M-N TEOREMA ili TEOREMA O PARAMETRIZACIJI. Ako je neka $(m+n)$-arna funkcija φ rekurzivna, onda će takva biti i svaka n-arna funkcija dobijena fiksiranjem m argumenata funkcije φ, tj. njihovim tretiranjem kao parametara. S-M-N TEOREMA tvrdi da za svaka dva prirodna broja m i n, postoji rekurzivna funkcija S_n^m, koja primenom na indeks $(m+n)$-arne funkcije φ i nekih m brojeva, daje kao vrednost indeks funkcije dobijene supstitucijom tih brojeva na određena mesta u funkciji φ. Drugim rečima, ako e rekurzivno definiše $(m+n)$-arnu funkciju $\varphi(y_1, ..., y_m, x_1, ..., x_n)$, onda $S_n^m(e, p_1, ..., p_m)$ rekurzivno definiše n-arnu funkciju $\varphi(p_1, ..., p_m, x_1, ..., x_n)$.

Ova teorema je blisko povezana sa problemom supstitucije terama mesto slobodnih promenljivih u nekoj formuli, odnosno sa funkcijom koja predstavlja tu supstituciju. Bliska veza između te dve vrste supstitucije i njihove uloge u formulisanju različitih oblika samoreferencije biće detaljnije istražena kasnije u radu.

Indeksiranje rekurzivnih funkcija brojevima omogućava predstavljanje funkcija koje se primenjuju na druge funkcije onim funkcijama koje se primenjuju na njihove indekse. To omogućava formulisanje jedne vrste samoreferencije unutar teorije rekurzija - one koja se sastoji u primeni funkcije na samu sebe, odnosno, na sopstveni indeks. Ova vrsta samoreferencije u osnovi je jedne od najznačajnijih metoda konstruisanja funkcija koja ima veoma široku upotrebu i omogućava dolaženje do nekih značajnih rezultata. Ta metoda, poznata kao *dijagonalizacija*, biće predstavljena u nastavku zajedno sa njenim najznačajnijim primenama.

2.1.3 *Dijagonalizacija*

Metoda *dijagonalizacije* (takođe poznata kao *dijagonalni argument*) koristi se u različite svrhe širom teorije izračunljivosti, i omogućava dolaženje do nekih njenih važnih rezultata. Njena upotreba, međutim, nije ograničena na tu teoriju. Zapravo, ta metoda je najpre korišćena u nekim drugim oblastima matematike u kojima je takođe donela značajne rezultate.

Dijagonalizaciju je kao metodu dokazivanja uveo Kantor koji ju je koristio u dokazu da je skup realnih brojeva neprebrojiv, odnosno, da je strogo veći (u smislu kardinalnosti) od skupa prirodnih brojeva. Kantor je to dokazao svođenjem na apsurd pretpostavke da je skup realnih bojeva prebrojiv. Pretpostavimo da jeste. Tada sve njegove elemente možemo da pobrojimo kao $r_1, r_2, r_3, ...$ tako da se svaki realan broj javi u toj listi kao neko r_n, gde je n prirodan broj veći od 0. Pretpostavimo da je za svaki broj r_n u listi, njegov decimalni zapis fiksiran. Služeći se tom beskonačnom listom definisaćemo realan broj r oblika $0.d_1d_2d_3...$ na sledeći način: uzećemo da je njegova n-ta decimala d_n jednaka 1, osim ako je n-ta decimala broja r_n u listi baš 1. U tom slučaju, neka je decimala d_n jednaka 7. Za svaki broj n, n-ta decimala tako definisanog broja r razlikovaće se od n-te decimale broja r_n u listi. To znači da će realan broj r biti različit od svakog broja iz date liste. To protivreči pretpostavci da su u datoj listi pobrojani svi realni brojevi. Dakle, ipak ne mogu svi realni brojevi biti pobrojani, što znači da je skup realnih brojeva ipak neprebrojiv.

Decimale koje su na n-tom mestu n-tih realnih brojeva u listi, formiraju dijagonalu matrice koju ta lista pravi, pa otud i naziv ove metode. Metoda dijagonalizacije se u ovom slučaju sastoji u definisanju novog realnog broja modifikacijom te dijagonale, koja garantuje da tako definisani broj neće biti jednak nijednom broju u toj listi.

Blisko povezan sa ovim Kantorovim dokazom je i RIŠAROV PARADOKS koji se umesto realnih brojeva tiče njihovih definicija. Skup definicija realnih brojeva na srpskom (ili bilo kom drugom) jeziku je prebrojiv budući da može postojati samo prebrojivo mnogo kombinacija simbola tog jezika koje čine takve definicije. Razlog je to što je alfabet tog jezika prebrojiv i što je, u opštem slučaju, skup svih reči na prebrojivom alfabetu takođe prebrojiv. Sve te definicije mogu se zato pobrojati tako da čine listu definicija realnih brojeva. Tada dijagonalizacijom možemo definisati realan broj za koji važi da se njegova n-ta decimala razlikuje za jedan od n-te decimale broja definisanog n-tom definicijom u datoj listi. Dobijena definicija jeste definicija realnog broja na srpskom jeziku pa bi zato i sama trebalo da pripada datoj listi. Međutim, pretpostavka da se ona javlja na bilo kom mestu te liste vodi protivrečnosti.

U oblasti teorije skupova, Kantor je koristio verziju istog argumenta kako bi dokazao teoremu (poznatu kao KANTOROVA TEOREMA) prema kojoj za svaki skup važi da je kardinalnost njegovog partitivnog skupa strogo veća od njegove kardinalnosti. Drugim rečima, Kantor je pokazao da za proizvoljan skup S važi da ne postoji *surjekcija* iz tog skupa u njegov partitivni skup $\mathcal{P}(S)$, tj. funkcija čija je slika jednaka skupu $\mathcal{P}(S)$. Da bismo to videli, uzmimo proizvoljnu funkciju $f : S \to \mathcal{P}(S)$ i posmatrajmo sledeći podskup skupa S: $T = \{x \in S : x \notin f(x)\}$. Neki element x skupa S pripadaće skupu T ako i samo ako ne pripada skupu $f(x)$. Pretpostavimo da za neko $t \in S$ važi $f(t) = T$. Tada će važiti da je $t \in T$ ako i samo ako $t \notin f(t)$, odnosno, ako i samo ako $t \notin T$, što je kontradikcija. Dakle, skup T ne može biti jedna od vrednosti funkcije f, pa ta funkcija nije surjekcija. Ovaj Kantorov dokaz, tj. način na koji je dijagonalizacija u njemu korišćena za definiciju skupa T, inspirisao je Raselovu formulaciju njegovog paradoksa, kao i druge paradokse poput Cermelovog i Hilbertovog.

U nastavku ovog dela rada videćemo kako je metoda dijagonalizacije prilagođena predmetu teorije izračunljivosti i, umesto za formulisanje paradoksa, korišćena za dokazivanje nekih od njenih najznačajnijih rezultata.

Indeksiranje rekurzivnih funkcija omogućava uniforman način njihovog izlistavanja prema indeksima. Na primer, lista unarnih rekurzivnih funkcija može biti zadata kao prva kolona sledeće matrice:

$$\begin{array}{ccccc}
 & 0 & 1 & 2 & \ldots \\
\varphi_0 & \varphi_0(0) & \varphi_0(1) & \varphi_0(2) & \ldots \\
\varphi_1 & \varphi_1(0) & \varphi_1(1) & \varphi_1(2) & \ldots \\
\varphi_2 & \varphi_2(0) & \varphi_2(1) & \varphi_2(2) & \ldots \\
\vdots & \vdots & \vdots & \vdots &
\end{array}$$

U ovoj matrici, pored unarnih rekurzivnih funkcija, prikazana je i njihova primena na proizvoljne brojeve. Slična matrica može biti konstruisana za neki podskup skupa rekurzivnih funkcija koji čine funkcije sa određenim svojstvom. Dijagonalu te matrice činiće primena tih funkcija na brojeve koji su jednaki njihovim indeksima. Tu dijagonalu možemo koristiti, na način sličan onom u gore navedenim primerima, kako bismo definisali funkciju iste vrste koja se razlikuje od svih funkcija u datoj listi na osnovu njene vrednosti za bar jedan argument. Time bismo pokazali da data lista ne sadrži sve funkcije te vrste. U tome se ukratko sastoji metoda dijagonalizacije unutar teorije rekurzija. U nastavku ćemo videti neke od njenih primena.

2.1.3.1 *Dokaz postojanja rekurzivnih funkcija koje nisu totalne*

Jedna od primena dijagonalizacije unutar teorije izračunljivosti ima za cilj dokaz neadekvatnosti svake formalizacije izračunljivih funkcija prema kojoj su samo totalne funkcije izračunljive. Među njima je i ona koja izračunljive funkcije poistovećuje sa primitivno rekurzivnim funkcijama, za koju se pokazalo da je neadekvatna već na osnovu postojanja konkretnih primera intuitivno izračunljivih funkcija koje nisu primitivno rekurzivne. Dijagonalizacija pruža još jedan dokaz te neadekvatnosti, odnosno preterane isključivosti formalizacije koja se ograničava na pojam primitivne rekurzivnosti.

Sve primitivno rekurzivne funkcije mogu formirati listu napravljenu na osnovu njihovih indekasa. Uzmimo da gore navedena matrica sadrži listu upravo takvih unarnih funkcija $\varphi_0, \varphi_1, \varphi_2, ...$ Tada dijagonalizacijom možemo definisati funkciju g, za koju važi da za svako x, $g(x) = \varphi_x(x) + 1$. Pretpostavimo da se funkcija g javlja negde u datoj listi, tj. da je primitivno rekurzivna. U tom slučaju, postoji broj x_0 takav da $g = \varphi_{x_0}$. Ali tada bi važilo da:

$$\varphi_{x_0}(x_0) = g(x_0) = \varphi_{x_0}(x_0) + 1$$

što je kontradikcija. Dakle, funkcija g ne može biti jednaka nijednoj funkciji koja se javlja u datoj listi, što znači da nije primitivno rekurzivna. Međutim, ako su sve primitivno rekurzivne funkcije izračunljive, tada će i funkcija g biti izračunljiva. To sledi zato što njenu vrednost za svaki argument možemo dobiti kompozicijom funkcije sledbenika i odgovarajuće funkcije iz date liste. Na osnovu toga možemo zaključiti da skup primitivno rekurzivnih funkcija nije potpun jer ne sadrži sve izračunljive funkcije. Dodavanje gore definisane funkcije g tom skupu neće ga učiniti potpunim jer i na tako prošireni skup funkcija možemo primeniti dijagonalizaciju i time pokazati njegovu nepotpunost u odnosu na skup izračunljivih funkcija. Sličan argument može se primeniti i na svaki drugi skup totalnih izračunljivih funkcija, što znači da nijedan od njih neće sadržati sve izračunljive funkcije.

Sa druge strane, dijagonalni argument ne vodi istom zaključku kada je primenjen na skup svih parcijalno rekurzivnih funkcija. Lista svih unarnih parcijalno rekurzivnih funkcija: $\psi_0, \psi_1, \psi_2, ...$ takođe može biti formirana na osnovu indekasa tih funkcija. Na tu listu onda možemo primeniti metodu dijagonalizacije kako bismo definisali funkciju h, takvu da $h(x) = \psi_x(x) + 1$. U ovom slučaju, međutim, ne možemo pokazati da tako definisana funkcija ne pripada datoj listi. Pretpostavimo da joj pripada, tj. da postoji neki broj x_0 takav da $h = \psi_{x_0}$. Tada će važiti:

$$\psi_{x_0}(x_0) = h(x_0) = \psi_{x_0}(x_0) + 1$$

što, međutim, nije kontradikcija ako funkcija ψ_{x_0} nije definisana za argument x_0. Ta mogućnost u ovom slučaju nije isključena jer nismo pretpostavili da su sve funkcije u datoj listi totalne. Dakle, jedino što možemo zaključiti na osnovu navedenog argumenta je da funkcija h, ako je rekurzivna, ne može biti totalna jer, kao što smo videli, ne može biti definisana za sopstveni indeks.

Ova funkcija, takođe, ne može imati totalno rekurzivno zatvorenje. Drugim rečima, ne može postojati totalna rekurzivna funkcija koja proširuje funkciju h, tj. čije se vrednosti slažu sa vrednostima funkcije h za sve argumente za koje je h definisana. Pretpostavimo da postoji takvo totalno rekurzivno zatvorenje funkcije h. Neka to bude funkcija ψ_{x_1}. Pošto je funkcija ψ_{x_1} totalna, $\psi_{x_1}(x_1)$ je definisano, pa isto važi i za $h(x_1)$ koje je po definiciji jednako $\psi_{x_1}(x_1) + 1$. Pošto je funkcija h definisana za x_1, vrednost njenog zatvorenja za taj argument mora se slagati sa njenom vrednošću, odnosno, mora važiti da $\psi_{x_1}(x_1) = h(x_1)$. Međutim, tada će na osnovu tranzitivnosti da važi i $\psi_{x_1}(x_1) = \psi_{x_1}(x_1) + 1$, što je opet kontradikcija.

Tako nas primena dijagonalnog argumenta primorava na zaključak da funkcije koje nisu totalne moraju biti pripuštene među rekurzivne funkcije kako bi formalizacija izračunljivih funkcija kao rekurzivnih bila adekvatna. Osim toga, ovaj argument takođe pokazuje da skup totalnih rekurzivnih funkcija ne može biti rekurzivan. Zapravo, taj skup ne može biti čak ni rekurzivno nabrojiv jer bismo u tom slučaju mogli da izlistamo njegove elemente i da primenom dijagonalnog argumenta definišemo totalnu rekurzivnu funkciju koja ne pripada tom skupu. Drugim rečima, problem totalnosti rekurzivnih funkcija ne može biti rekurzivno rešiv.

Ovaj rezultat primene dijagonalizacije ilustruje ulogu te metode u dokazivanju rekurzivne nerešivosti različitih problema teorije izračunljivosti. Najznačajnija takva primena dijagonalizacije jeste u dokazu rekurzivne nerešivosti *halting problema* koji ćemo predstaviti u nastavku (taj problem bi na srpskom bilo pravilnije nazvati *problemom zaustavljanja*, ali se originalan naziv toliko ustalio da je postao gotovo tehnički termin, pa ćemo ga tako i ovde koristiti).

2.1.3.2 *Nerešivost halting problema*

Halting problem tiče se pitanja da li je neka rekurzivna funkcija definisana za određeni argument, odnosno, da li se izračunavanje te funkcije za dati argument završava (*halts*) i daje određenu vrednost. Možemo reći da se halting problem tiče izračunljivosti relacije *HP* koja označava relaciju $\{(x, y) : \varphi_x(y) \text{ je definisano}\}$. Karakteristična funkcija te relacije bila bi definisana na sledeći način:

$$f(x, y) = \begin{cases} 0, & \text{ako je } \varphi_x(y) \text{ definisano} \\ 1, & \text{ako je } \varphi_x(y) \text{ nedefinisano.} \end{cases}$$

Halting problem će biti rekurzivno rešiv ako i samo ako je funkcija f rekurzivna. Međutim, argument koji ćemo predstaviti pokazuje da ta funkcija ne može biti rekurzivna.

Pretpostavimo da jeste. Onda bi takva bila i funkcija g definisana pomoću funkcije f na sledeći način:

$$g(x) = \begin{cases} 0, & \text{ako } f(x, x) = 1 \\ \text{nedefinisano}, & \text{ako } f(x, x) = 0. \end{cases}$$

Ako je funkcija g rekurzivna, onda je ona rekurzivno definisana nekim brojem x_0, pa je samim tim jednaka funkciji φ_{x_0}. Na osnovu definicija funkcija f i g važi da je $\varphi_{x_0}(x_0)$ definisano ako i samo ako je $f(x_0, x_0) = 0$, dok je $f(x_0, x_0) = 0$ ako i samo ako je $g(x_0)$ nedefinisano. Na osnovu tranzitivnosti ekvivalencije sledi da je $\varphi_{x_0}(x_0)$ definisano ako i samo ako je $g(x_0)$ nedefinisano. Međutim, funkcija g je po pretpostavci jednaka funkciji φ_{x_0}, što znači da je ista funkcija definisana za argument x_0 ako i samo ako nije definisana za taj argument. Dakle, na osnovu toga što suprotna pretpostavka vodi kontradikciji, možemo zaključiti da funkcija g, pa samim tim i funkcija f, ne mogu biti rekurzivne. Iz toga sledi da halting problem nije rekurzivno rešiv (cf. Rogers, 1987, pp. 24-25).

Konstrukcija funkcije g prati osnovnu ideju metode dijagonalizacije. Naime, njena definicija garantuje da će se ta funkcija razlikovati od funkcije φ_x, za svako x, i to na osnovu rezultata njihove primene na isto to x: jedna od njih će biti definisana za taj argument ako i samo ako druga nije. Upotreba dijagonalizacije u definiciji funkcije g tako omogućava dokaz rekurzivne nerešivosti halting problema.

Isti argument može se koristiti za dokaz rekurzivne nerešivosti jednog specijalnog slučaja prethodnog problema koji se tiče pitanja da li je neka rekurzivna funkcija definisana za broj koji predstavlja njen indeks. Drugim rečima, istim argumentom se pokazuje i da funkcija h definisana na sledeći način:

$$h(x) = \begin{cases} 0, & \text{ako je } \varphi_x(x) \text{ definisano} \\ 1, & \text{ako je } \varphi_x(x) \text{ nedefinisano} \end{cases}$$

nije rekurzivna. To sledi iz činjenice da se tako definisana funkcija h može koristiti umesto f u gore datoj definiciji funkcije g, koja bi zato morala biti rekurzivna ako je i h takva. Budući da je funkcija h karakteristična funkcija skupa K:

$$K = \{x : \varphi_x(x) \text{ je definisano}\},$$

sledi da taj skup takođe nije rekurzivan.

Uslov za pripadnost navedenom skupu možemo izraziti i koristeći Klinijev predikat T. Naime, skupu K bi trebalo da pripadaju svi indeksi funkcija za koje važi da postoji neko konačno izračunavanje njihovih vrednosti za njihove sopstvene indekse kao argumente. Dakle, skup K bi zapravo predstavljao ekstenziju predikata $\exists yT(x, x, y)$, dok bi funkcija h služila za određivanje te ekstenzije. Iz navedenog rezultata tako sledi da predikat $\exists yT(x, x, y)$ i relacija označena njime nisu rekurzivni.

Sa druge strane, skup K jeste rekurzivno nabrojiv pomoću funkcije koja prihvata svako x za koje $\varphi_x(x)$ daje neku vrednost. Iz toga što je K rekurzivno nabrojiv ali ne i rekurzivan skup, sledi da njegov komplement nije rekurzivno nabrojiv. Komplement skupa K je skup $\bar{K}$ definisan na sledeći način:

$$\bar{K} = \{x : \varphi_x(x) \text{ nije definisano}\}.$$

Taj skup sadrži sve i samo indekse onih rekurzivnih funkcija koje nisu definisane za svoje indekse. On takođe može biti definisan pomoću Klinijevog predikata T, tj. tako da stoji za ekstenziju predikata $\forall y \neg T(x, x, y)$ koji tvrdi da ne postoji konačno izračunavanje vrednosti funkcije indeksirane brojem x za isto to x. Skup $\bar{K}$ podseća na Raselov skup svih skupova koji nisu sopstveni elementi, a dokaz da on nije rekurzivno nabrojiv podseća na izvođenje RASELOVOG PARADOKSA. Naime, pretpostavimo da skup $\bar{K}$ jeste rekurzivno nabrojiv nekom rekurzivnom funkcijom φ_{x_0}. Tada će važiti da je $\varphi_{x_0}(x_0)$ definisano i jednako 0 ako i samo ako $x_0 \in \bar{K}$, dok će ovo poslednje važiti ako i samo ako $\varphi_{x_0}(x_0)$ nije definisano. Drugim rečima, važiće da $x_0 \in \bar{K}$ ako i samo ako $x_0 \notin \bar{K}$. Kako smo iz pretpostavke da postoji funkcija koja rekurzivno nabraja elemente skupa $\bar{K}$ na taj način izveli kontradikciju, možemo zaključiti da takva funkcija ipak ne može postojati.

Navedeni primeri pokazuju da samoreferencija u vidu primene funkcije na sopstveni indeks može imati značajnu ulogu u dokazivanju rekurzivne nerešivosti određenih problema. Ta uloga ove vrste samoreferencije ostvarena je njenim korišćenjem u metodi dijagonalizacije. Jedna druga vrsta samoreferencije koju ćemo predstaviti u nastavku sastoji se u karakterizaciji funkcije koja sadrži referenciju na samu sebe ili na njom karakterisanu funkciju. Videćemo da je u nekim slučajevima mogućnost te vrste samoreferencije, odnosno njeno opravdanje, takođe rezultat primene dijagonalnog argumenta. Ta vrsta samoreferencije predmet je dva veoma značajna Klinijeva rezultata - njegove prve i druge teoreme rekurzije.

2.1.4 *Prva teorema rekurzije*

Klinijeva PRVA TEOREMA REKURZIJE tiče se karakterističnog svojstva rekurzivnih funkcija da njihove definicije mogu sadržati referenciju na same te funkcije koje definišu. Ako je ζ promenljiva koja stoji za neku n-arnu parcijalno rekurzivnu funkciju, onda ta teorema tvrdi da:

- za svaki $(n+1)$-arni parcijalno rekurzivni funkcional F, jednačina $\zeta(x_1, ..., x_n) = F(\zeta; x_1, ..., x_n)$ ima najmanje parcijalno rekurzivno rešenje φ.

Ovaj rezultat pruža opravdavanje definicija koje su samoreferencijalne u smislu da referiraju na funkciju koju definišu, time što garantuje postojanje rekurzivnih funkcija koje zadovoljavaju takvu definiciju. Vrednost tako definisane funkcije za određene argumente zavisi od njenih drugih vrednosti na način koji je određen funkcionalom F. Jedna od osnovnih metoda definisanja rekurzivnih funkcija - rekurzija, predstavlja primer upravo takve vrste definicije. Definicija neke funkcije pomoću rekurzije pokazuje kako vrednost te funkcije za neki broj zavisi od njene vrednosti za prethodnika tog broja. Prema PRVOJ TEOREMI REKURZIJE ta međusobna zavisnost vrednosti iste funkcije za različite argumente može biti i druge vrste, sve dok je ona opisiva prethodno utvrđenim metodama teorije izraženim pomoću nekog rekurzivnog funkcionala.

Za dokaz teoreme, ograničene na unarne rekurzivne funkcije (koji se lako može prilagoditi funkcijama sa više argumenata), posmatrajmo niz unarnih rekurzivnih funkcija φ_0, φ_1, φ_2, ... takvih da je φ_0 u potpunosti nedefinisana funkcija, dok za svako $n \geq 0$, važi da $\varphi_{n+1}(x) = F(\varphi_n; x)$. Svaka od funkcija φ_{n+1} biće proširenje funkcije φ_n (vidi: Kleene, 1971, p. 339). Taj niz funkcija imaće zato najmanju gornju granicu u vidu funkcije φ koja je definisana za sve i samo one argumente za koje je neka od funkcija u nizu definisana i ima istu vrednost kao ta funkcija i sva njena proširenja. Tada će važiti da za svako x, $\varphi(x) = F(\varphi; x)$. Pored toga, važiće i da je funkcija φ najmanja parcijalno rekurzivna funkcija koja ima to svojstvo (za detalje dokaza vidi: Kleene, 1971, pp. 348-349). Funkcija φ do koje smo došli na taj način je najmanja rekurzivna funkcija koja ne može biti proširena primenom funkcionala F.

Ova teorema ima ulogu *teoreme o fiksnoj tački* za rekurzivne funkcionale, pošto se njom tvrdi postojanje njihovih fiksnih tačaka, tj. funkcija koje ostaju nepromenjene nakon njihove primene. Način na koji se ta teorema dokazuje, odnosno način na koji se konstruiše fiksna tačka funkcionala - iteracijom njegove primene počevši od određene rekurzivne funkcije, je vrlo tipičan za teoreme te vrste (to ćemo nešto kasnije videti i na drugim primerima).

Dozvoljavanje samoreferencijalnih definicija unutar formalne teorije može se činiti problematičnim jer pomoću njih funkcije sa naizgled protivrečnim osobinama mogu biti definisane. Na primer, instanca PRVE TEOREME REKURZIJE tvrdi da postoji rekurzivna funkcija φ, takva da $\varphi(x) = S(\varphi(x))$. Iako je data definicija funkcije naizgled paradoksalna, ipak se u teoriji rekurzija može naći funkcija koja ima njom opisano svojstvo, ali čije postojanje ne vodi kontradikciji. Jedina takva funkcija biće zapravo ona koja nije nigde definisana. Tako je ponovo, kao u prethodno izloženim primenama dijagonalizacije, protivrečnost uzrokovana primenom neke samoreferencijalne metode izbegnuta prepoznavanjem parcijalnosti rekurzivnih funkcija.

Još jedan način na koji karakterizacija neke funkcije može biti samoreferencijalna jeste da sadrži referenciju ne na definisanu funkciju već na kôd njene rekurzivne definicije. Takve definicije opravdane su drugim Klinijevim rezultatom koji će biti izložen u sledećem odeljku.

2.1.5 *Druga teorema rekurzije*

Zahvaljujući kodiranju rekurzivnih definicija funkcija prirodnim brojevima, kojima su te funkcije indeksirane, još jedna vrsta njihove samoreferencijalne karakterizacije postaje moguća - ona koja sadrži referenciju na indeks njom opisane funkcije. DRUGA TEOREMA REKURZIJE garantuje da rekurzivne funkcije opisane na taj način postoje. U jednoj od njenih formulacija ona tvrdi da:

- za svaku $(n+1)$-arnu parcijalno rekurzivnu funkciju $\psi(z, x_1, ..., x_n)$, postoji broj e koji rekurzivno definiše n-arnu funkciju $\psi(e, x_1, ..., x_n)$.

Dokaz ove teoreme koristi S-M-N TEOREMU za određivanje takvog broja e. Pošto su funkcija ψ i funkcije S_n^m, za svaka dva prirodna broja m i n, parcijalno rekurzivne, njihovom kompozicijom dobijamo novu parcijalno rekurzivnu funkciju:

$$\psi\left(S_n^1(y, y), x_1, ..., x_n\right).$$

Neka je ona rekurzivno definisana brojem a. Onda je traženi broj e jednak $S_n^1(a, a)$. Naime, to e rekurzivno definiše funkciju

$$\psi\left(S_n^1(a, a), x_1, ..., x_n\right),$$

odnosno, funkciju $\psi(e, x_1, ..., x_n)$.

Tako dokaz ove teoreme pokazuje da za svaku rekurzivnu funkciju ψ, broj e koji rekurzivno definiše $\psi(e, x_1, ..., x_n)$ može efektivno da se nađe na osnovu kôda te funkcije. Ovaj rezultat može biti formulisan i u obliku jednakosti, odnosno tako da tvrdi da za svaku parcijalno rekurzivnu funkciju $\psi(z, x_1, ..., x_n)$, postoji broj e, takav da: $\varphi_e(x_1, ..., x_n) = \psi(e, x_1, ..., x_n)$ (Kleene, 1971, pp. 352-353).

Ako posmatramo funkciju ψ kao transformaciju algoritama rekurzivnih funkcija predstavljenih njihovim kôdovima, onda ova teorema tvrdi da mora postojati neki algoritam koji i nakon takve transformacije i dalje izračunava istu funkciju. To značenje teoreme postaje eksplicitnije u njenoj drugoj formulaciji u kojoj ona tvrdi da za svaku funkciju ψ postoji broj e takav da $\varphi_e = \varphi_{\psi(e)}$. Gore opisana konstrukcija kôda e takvog algoritma liči na dijagonalnu konstrukciju. Dijagonala matrice koja se u njoj koristi sastoji se od primena funkcija na sopstvene indekse. Funkcije dobijene tom primenom predstavljene su u ovom slučaju njihovim indeksima koji su oblika $S_n^1(y, y)$. Funkcija ψ koja se primenjuje na te indekse, može se razumeti kao funkcija kojom se modifikuje ta dijagonala. Međutim, za razliku od do sada predstavljenih rezultata primene dijagonalnog argumenta, u ovom slučaju se ispostavlja da tako modifikovana dijagonala matrice i sama opisuje neku funkciju koja se nalazi u toj matrici, gde je indeksirana brojem a. Pošto je to slučaj, funkcija ψ se primenjuje i na kôd njene dijagonalizacije $S_n^1(a, a)$, ali time ne uspeva da ga modifikuje tako da on više ne definiše istu funkciju. Naime, kako se ispostavlja, dobijena funkcija i sama je rekurzivno definisana brojem $S_n^1(a, a)$.

Dakle, DRUGA TEOREMA REKURZIJE pokazuje da se u karakterizaciji neke rekurzivne funkcije može koristiti kôd njene rekurzivne definicije. Uloga tog kôda u definiciji zavisiće od funkcije ψ. Ta funkcija može biti i univerzalna funkcija Φ_n koja kada se primeni na kôd neke funkcije izračunava tu funkciju. Iz toga bi sledilo da se pored kôda rekurzivne funkcije, i sama ta funkcija može koristiti u njenoj karakterizaciji. Ta činjenica može biti formulisana kao korolar navedene teoreme. Njime bi se tvrdilo da za svaki rekurzivni funkcional $F(\zeta; z, x_1, ..., x_n)$, može biti nađena parcijalno rekurzivna funkcija φ sa indeksom e, takva da:

$$\varphi_e(x_1, ..., x_n) = F(\varphi_e; e, x_1, ..., x_n)$$

važi za sve brojeve $x_1, ..., x_n$. Za razliku od PRVE TEOREME REKURZIJE koja tvrdi nešto vrlo slično, definicija funkcije φ_e ovde može zavisiti ne samo od same te funkcije već i od kôda njenog algoritma. Još jedna razlika u odnosu na prvu teoremu jeste to što ova teorema ne garantuje da je parcijalno rekurzivna funkcija dobijena na navedeni način - kao fiksna tačka funkcionala F - najmanja takva parcijalno rekurzivna funkcija (Kleene, 1971, p. 353).

Mogli bismo reći da DRUGA TEOREMA REKURZIJE izražava samoreferencijalno svojstvo indekasa rekurzivnih funkcija koji predstavljaju kôdove njihovih rekurzivnih definicija. Neki od tih kôdova kodiraju definicije koje su dobijene od njih samih. Ova teorema je tako često korišćena za opravdanje samoreferencijalnih definicija. Postoje i druge značajne primene te teoreme od kojih ćemo neke pomenuti u nastavku.

2.1.5.1 *Neke primene druge teoreme rekurzije*

Jedna od najznačajnijih upravo spomenutih primena DRUGE TEOREME REKURZIJE jeste ona koja omogućava dokaz RAJSOVE TEOREME. Prema toj teoremi, za svaki neprazan pravi podskup skupa unarnih rekurzivnih funkcija, važi da skup njihovih indekasa nije rekurzivan. Drugim rečima, problem odlučivosti koji se tiče proizvoljnog netrivijalnog svojstva unarnih rekurzivnih funkcija (tj. svojstva koji barem jedna ali ne i svaka takva funkcija ima) je rekurzivno nerešiv. Postoje različiti dokazi ove teoreme. Jedan od njih zasniva se na činjenici da se halting problem može svesti na bilo koji problem te vrste, iz čega sledi da bi rešivost svakog takvog problema implicirala i rešivost halting problema.

Drugi način dokazivanja ove teoreme jeste korišćenjem DRUGE TEOREME REKURZIJE. Neka je $\mathcal{A}$ skup svih unarnih rekurzivnih funkcija sa određenim netrivijalnim svojstvom, dok je A skup indekasa tih funkcija. Pošto je pretpostavljeno svojstvo netrivijalno, trebalo bi da postoji indeks funkcije koja je u datom skupu, neka to bude a, kao i indeks funkcije koja je van tog skupa, neka to bude b. Definišimo funkciju h na sledeći način:

$$h(x) = \begin{cases} a, & \text{ako } x \notin A \\ b, & \text{ako } x \in A. \end{cases}$$

Na osnovu ove definicije jasno je da za svako x važi da je ono u skupu A ako i samo ako $h(x)$ nije u A. Zahvaljujući DRUGOJ TEOREMI REKURZIJE znamo da ako je h rekurzivna funkcija, onda postoji broj n, za koji važi $\varphi_n = \varphi_{h(n)}$. Pokazaćemo da za takvo n ne može važiti ni da je u skupu A ni da je van tog skupa.

Kako bismo to videli, pretpostavimo da n jeste u A. To znači da su funkcija φ_n i njoj identična funkcija $\varphi_{h(n)}$ u $\mathcal{A}$. Iz toga međutim sledi da je $h(n)$ zajedno sa n u A, što ne može biti slučaj na osnovu definicije funkcije h. Pretpostavimo onda da je n van skupa A. Tada su obe funkcije φ_n i $\varphi_{h(n)}$ van skupa $\mathcal{A}$, što znači da ni $h(n)$ ni n nisu u skupu A, što ponovo protivreči definiciji funkcije h. Kako smo na taj način došli do kontradikcije, možemo zaključiti da funkcija h ipak nije rekurzivna, zbog čega na nju ne možemo primeniti DRUGU TEOREMU REKURZIJE. Iz toga sledi da karakteristična funkcija skupa A takođe ne može biti rekurzivna, jer bi inače funkcija h bila rekurzivna. Pošto je A skup indekasa unarnih rekurzivnih funkcija sa proizvoljnim netrivijalnim svojstvom, na osnovu ovog argumenta sledi da će problem odlučivosti za svako takvo svojstvo unarnih rekurzivnih funkcija biti rekurzivno nerešiv (cf. Yanofsky, 2003, p. 381).

DRUGA TEOREMA REKURZIJE tako ima važnu ulogu u dokazivanju jednog veoma opšteg rezultata - nerešivosti problema koji se tiču netrivijalnih svojstava unarnih rekurzivnih funkcija. Neki od primera takvih problema su oni koji se tiču pitanja da li je funkcija totalna (za koji smo već drugim sredstvima pokazali da nije rešiv), pitanja da li je funkcija prazna ili ima konačni domen, da li izračunava sledbenika ili prethodnika nekog broja, itd.

Još jedna primena DRUGE TEOREME REKURZIJE koju ćemo pomenuti tiče se *samoreprodukujuće funkcije* - funkcije koja za svaki argument kao vrednost daje sopstveni indeks. Koristeći datu teoremu možemo pokazati da takva rekurzivna funkcija postoji, odnosno, da postoji funkcija φ_n, takva da za svako x, $\varphi_n(x) = n$. Kako bismo to videli, uzmimo da je funkcija f prva projekcija, tj. da za svako x i y, važi da je $f(y, x) = y$. Na osnovu S-M-N TEOREME znamo da postoji efektivan način za pronalaženje kôda funkcije dobijene od prve projekcije fiksiranjem njenog prvog argumenta. Iz toga sledi da postoji rekurzivna funkcija S, takva da:

$$\varphi_{S(y)}(x) = f(y, x) = y$$

(ako je e indeks rekurzivne funkcije koja odgovara funkciji f, onda $S(y) = S_1^1(e, y)$). Osim toga na osnovu DRUGE TEOREME REKURZIJE znamo da postoji neko n za koje važi $\varphi_n = \varphi_{S(n)}$. Iz toga sledi da

$$\varphi_n(x) = \varphi_{S(n)}(x) = f(n, x) = n.$$

Funkcija φ_n dobijena na ovaj način primenom DRUGE TEOREME REKURZIJE je samoreprodukujuća funkcija koju smo tražili (cf. Yanofsky, 2003, p. 381).

Obe Klinijeve teoreme rekurzije govore nešto o prirodi rekurzivnih funkcija, a to je da se one na različite načine mogu pojaviti u sopstvenim definicijama. Navedene primene tih teorema, kao i mnoge druge, pokazuju vrednost i primenljivost samoreferencijalnih definicija koje učestvuju u dokazima bitnih tvrđenja teorije rekurzija. Baš kao samoreferencijalna primena funkcija korišćena u metodi dijagonalizacije, i samoreferencijalne definicije funkcija su na taj način pretvorene u veoma snažne metode teorije izračunljivosti. To je rezultat jednog posebnog pristupa samoreferenciji i njenim posledicama u ovoj oblasti. Naime, naizgled paradoksalni rezultati kojima neke njihove instance vode u određenim okolnostima, kao što je postojanje fiksne tačke funkcije sledbenika ili postojanje funkcija sa naizgled protivrečnim osobinama, ne uzimaju se kao rezultati koji pokazuju da samoreferencijalne metode nisu odgovarajuće, ili da teorija koja ih koristi nije korektna. Umesto toga, oni su korišćeni za izvođenje određenih negativnih rezultata koji karakterišu rekurzivne funkcije ili teoriju koja se njima bavi. Zahvaljujući tome, pokazalo se da uključivanje samoreferencije u teoriju, u obliku formalizacije funkcija koje se primenjuju na sebe ili učestvuju na neki način u sopstvenoj definiciji, umesto da predstavlja pretnju po njenu konzistentnost, može biti metodološki veoma zahvalno.

Analogni samoreferencijalni fenomeni, sa sličnim primenama i posledicama, javljaju se i u metamatematici - oblasti matematike koja se bavi samim matematičkim teorijama. U toj oblasti, samoreferencija se javlja u vidu lingvističke samoreferencije, kojom ćemo se baviti u sledećem poglavlju.

2.2 LINGVISTIČKA SAMOREFERENCIJA

Vrsta samoreferencije kojom ćemo se baviti u ovom poglavlju sastoji se u referenciji nekog lingvističkog objekta koji ima funkciju tvrđenja na samog sebe. Takav objekat interpretiran je kao da tvrdi nešto o sebi, odnosno pripisuje samom sebi neko svojstvo ili relaciju u kojoj stoji prema nekim drugim objektima. Ovde ćemo se baviti lingvističkom samoreferencijom koja se javlja u jezicima određenih formalnih sistema i koja kao takva ima poseban metamatematički značaj. Kao i u slučaju funkcionalne samoreferencije, neke instance lingvističke samoreferencije takođe imaju važne posledice za teoriju u kojoj se javljaju. Najpoznatije od njih su GEDELOVE TEOREME O NEPOTPUNOSTI koje predstavljaju neke od najznačajnijih rezultata logike i matematike uopšte. Omogućavanjem tih rezultata, koji su u velikoj meri uticali na projekat formalizacije matematike, lingvistička samoreferencija je zapravo i dobila logički i matematički značaj. Ti rezultati na vrlo precizan način odgovaraju nekima od gore navedenih negativnih rezultata teorije izračunljivosti. Njihova sličnost i međusoban odnos biće ispitani nešto kasnije.

Kako bi formula koja tvrdi nešto o samoj sebi mogla biti formulisana u nekom formalnom sistemu, određena svojstva i relacije između formula i drugih objekata tog formalnog sistema moraju da budu opisiva unutar njega samog. Da to važi za bar neka sintaktička i metateorijska svojstva dovoljno jakih formalnih teorija sledi iz određenih rezultata koje ćemo ovde predstaviti.

Pošto se u njihovoj standardnoj formulaciji ti rezultati tiču formalnih sistema za aritmetiku i njihovih proširenja, upravo su ti sistemi, i lingvistička samoreferencija koja se u njima javlja, oni kojima ćemo se ovde baviti. Videćemo da mogućnost da svojstva nekog aritmetičkog formalnog sistema budu opisana unutar njega samog zavisi od prirode tih svojstava koja se tiče njihove izračunljivosti. Odgovarajuća priroda tih svojstava formalnog sistema biće ključna za postizanje njegove svrhe. Kako bismo to razjasnili, počećemo ovo poglavlje kratkim opisom formalnog sistema za aritmetiku i ciljeva njene formalizacije.

2.2.1 *Formalna aritmetika*

Pod formalnom aritmetikom najčešće se podrazumeva *Peanova aritmetika* (skraćeno *PA*). To je teorija zasnovana na logici prvog reda sa jednakošću, čijem jeziku je pridodata individualna konstanta $\mathbf{0}$, kao i operacijski simboli $\mathbf{S}$, $+$ i $\cdot$. Zatvoreni termi tog formalnog sistema predstavljaju imena prirodnih brojeva, a njegove formule tvrđenja o svojstvima prirodnih brojeva kao i aritmetičkih operacija na njima. Nelogičke aksiome teorije *PA* su sledeće:

$(PA1)$ $\forall \mathbf{x}\,(\mathbf{Sx} \neq \mathbf{0})$

$(PA2)$ $\forall \mathbf{x} \forall \mathbf{y}\,(\mathbf{Sx} = \mathbf{Sy} \rightarrow \mathbf{x} = \mathbf{y})$

$(PA3)$ $\forall \mathbf{x}\,(\mathbf{x} + \mathbf{0} = \mathbf{x})$

$(PA4)$ $\forall \mathbf{x} \forall \mathbf{y}\,(\mathbf{x} + \mathbf{Sy} = \mathbf{S}\,(\mathbf{x} + \mathbf{y}))$

$(PA5)$ $\forall \mathbf{x}\,(\mathbf{x} \cdot \mathbf{0} = \mathbf{0})$

$(PA6)$ $\forall \mathbf{x} \forall \mathbf{y}\,(\mathbf{x} \cdot \mathbf{Sy} = (\mathbf{x} \cdot \mathbf{y}) + \mathbf{x})$

$(PA7)$ $(\mathbf{A}\,(\mathbf{0}) \wedge \forall \mathbf{x}\,(\mathbf{A}\,(\mathbf{x}) \rightarrow \mathbf{A}\,(\mathbf{Sx}))) \rightarrow \forall \mathbf{x} \mathbf{A}\,(\mathbf{x})$, gde $\mathbf{A}\,(\mathbf{x})$ stoji za proizvoljnu formulu ove teorije čija je $\mathbf{x}$ jedina slobodna promenljiva.

Standardni model teorije *PA* čini skup prirodnih brojeva sa nulom kao istaknutim elementom, na kojem su definisane standardne aritmetičke operacije - operacija sledbenika, sabiranje i množenje. Nadalje, kada budemo govorili o rečenicama *PA* i drugih formalnih aritmetičkih teorija kao o istinitim ili lažnim, pod time ćemo misliti na njihovu istinitost ili lažnost u ovom modelu.

Aksiome na kojima se zasnivaju navedene aksiome Peanove aritmetike mogu se, osim kod Đuzepe Peana (Giuseppe Peano), naći i kod Dedekinda u njegovom delu *Šta su i čemu služe brojevi?* (u originalu: *Was sind und was sollen die Zahlen?*). Iz tog razloga bi adekvatniji naziv te teorije bio *Dedekind-Peanova aritmetika*. Među njima su i aksiome koje odgovaraju aksiomama $(PA1)$ - $(PA6)$ i izražavaju definišuće karakteristike odgovarajućih aritmetičkih operacija. Primetimo da su definicije sabiranja i množenja sadržane u tim aksiomama zapravo *rekurzivne definicije*. Osim što je uveo rekurzivne definicije aritmetičkih operacija, Dedekind je pokušao i da opravda njihovo korišćenje u aritmetici dokazom postojanja jedinstvenih funkcija na skupu prirodnih brojeva koje su opisane tim definicijama (Müller-Stach, 2017). Skup prirodnih brojeva $\mathbb{N}$, Dedekind definiše kao najmanji skup koji sadrži broj 0 i koji je zatvoren za primenu operacije sledbenika. Njegova TEOREMA REKURZIJE tvrdi da ako je T skup čiji je jedan od elemenata t i ako postoji na njemu definisana unarna funkcija h, onda postoji i jedinstvena funkcija $f : \mathbb{N} \rightarrow T$, definisana na sledeći način:

$$\begin{aligned} f\,(0) &= t \\ f\,(S\,(x)) &= h\,(f\,(x)) \end{aligned}$$

Definicija funkcija na koju se ova teorema odnosi predstavlja poseban slučaj definicije rekurzijom. Ovde smo se njom poslužili da bismo definisali jednu unarnu funkciju. Opštije verzije teoreme, koje se tiču funkcija sa više argumenata, mogu se takođe naći kod Dedekinda. One prirodno vode formulaciji primitivne rekurzije u njenom modernom obliku. Dedekindovi rezultati tako, osim što imaju ulogu u aksiomatizaciji aritmetike pomoću rekurzivnih definicija aritmetičkih operacija, takođe opravdavaju korišćenje takvih definicija u aritmetici.

Jedan od ciljeva formalne aritmetičke teorije, kao što je PA, jeste sistematizacija aritmetičkog znanja. Potpuna sistematizacija bila bi postignuta kada bi takav formalni sistem bio nađen koji bi sadržao aksiome iz kojih su svi i samo istiniti iskazi o prirodnim brojevima i aritmetičkim operacijama izvodivi pomoću pravila izvođenja. To bi značilo da su svi i samo istiniti aritmetički iskazi *dokazivi* unutar takve formalne teorije. Aritmetički iskazi su u formalnoj teoriji predstavljeni njenim rečenicama, tj. formulama koje ne sadrže slobodne promenljive. Dokaz određenog aritmetičkog iskaza u formalnoj teoriji predstavlja neki konačni niz formula (ili drvo formula u zavisnosti od formalizacije) koji se završava rečenicom koja izražava taj iskaz, a počinje instancama aksioma te teorije. Svaka formula u tom nizu, osim aksioma, izvedena je iz formula koje joj prethode pomoću nekog od pretpostavljenih pravila izvođenja (za takvu formulu reći ćemo da je *neposredna posledica* formula iz kojih je izvedena).

Tom formalizacijom, pojam dokaza u aritmetici postaje izuzetno jasan i precizan. Dokaz je time pretvoren u formalnu relaciju između određene rečenice i nekog niza formula koji se njom završava. To svojstvo ove relacije čini pitanje o tome da li je nešto dokaz određene rečenice unutar formalne teorije nedvosmislenim. Time se otvara i mogućnost postojanja efektivne procedure kojom bi se moglo utvrditi da li neka rečenica stoji u *relaciji dokaza* sa nekim nizom formula koji se njom završava. Ta procedura bi se sastojala u proveri toga da li su formule kojima taj niz počinje instance aksioma date teorije i da li su sve ostale formule istog niza neposredne posledice nekih formula koje im prethode. Ako je provera toga da li je neka formula aksioma ili neposredna posledica određenih formula efektivna, onda bi takva bila i provera toga da li je nešto dokaz određene formule. Time bi bio ostvaren jedan od glavnih ciljeva formalizacije aritmetike koji se sastoji u razjašnjenju pojma dokaza aritmetičkog iskaza njegovim svođenjem na odlučivu relaciju između rečenice koja taj iskaz izražava i niza formula koji predstavlja njen dokaz.

Međutim, da li takva formalizacija aritmetičkog dokaza omogućava dokazivanje svih i samo istinitih rečenica formulisanih na jeziku aritmetičke formalne teorije kao što je PA? Drugim rečima, može li se takvom formalizacijom postići potpuna sistematizacija aritmetičkog znanja? Gedelovi rezultati o nepotpunosti daju negativan odgovor na to pitanje. Iz tih rezultata sledi da će za svaku konzistentnu formalnu teoriju koja sadrži aritmetiku i pruža odlučivu formalizaciju relacije dokaza, postojati neka njena rečenica koja će biti istinita ali nedokaziva u toj teoriji. Glavno sredstvo koje je Gedel koristio u dokazu tog rezultata je određena vrsta lingvističke samoreferencije. Razmotrimo rečenicu, po pretpostavci formulisanu unutar PA, koja za sebe tvrdi da je nedokaziva u PA. Takva rečenica biće istinita ako i samo ako je nedokaziva u PA. Pretpostavimo da je ona dokaziva u PA. Tada na osnovu navedene ekvivalencije sledi da je ona lažna. Međutim, u *valjanim* formalnim teorijama, kao što je PA, nijedna lažna rečenica nije dokaziva. Zato moramo zaključiti da ta rečenica ipak nije dokaziva u PA, i da je samim tim istinita. To je čini primerom istinite rečenice teorije PA koja nije dokaziva u toj teoriji. Na osnovu toga možemo zaključiti da PA ne pruža potpunu formalizaciju aritmetike. Ovaj argument nije usmeren isključivo na PA, već se tiče i svih drugih formalnih teorija koje dele relevantna svojstva sa PA, a koja ćemo u nastavku detaljnije opisati. Za sve takve teorije može se na isti način pokazati da su nepotpune. U tome se u veoma grubim crtama sastoji ideja u osnovi Gedelovog rezultata. U nastavku ćemo nešto formalnije opisati tu ideju i izložiti rezultate koji omogućavaju njenu formalizaciju.

Ono što pre svega treba razjasniti jeste u kom smislu se tvrđenje o nekom sintaktičkom svojstvu formalne teorije, kao što je dokazivost u njoj, može shvatiti kao aritmetičko tvrđenje i biti formulisano unutar aritmetičke formalne teorije? Ako je to objašnjeno, onda još treba pokazati i da aritmetička rečenica koja na osnovu ustanovljene interpretacije samoj sebi pripisuje neko sintaktičko svojstvo koje dobija unutar formalnog sistema može biti formulisana u istom tom sistemu. U nastavku ćemo videti kako je Gedel uspeo da formalno opravda sve te pretpostavke njegovog rezultata.

2.2.2 *Kodiranje jezika formalne aritmetike*

Kako bi tvrđenja o sintaktičkim svojstvima i relacijama između objekata neke formalne teorije bilo moguće formulisati kao tvrđenja koja se odnose na prirodne brojeve, sintaktički objekti čija nas svojstva zanimaju - formule i nizovi formula, treba da budu kodirani tim prirodnim brojevima. Kako je već objašnjeno u kontekstu formalne teorije rekurzivnih funkcija, kodiranje jezika neke formalne teorije sastoji se u uspostavljanju izračunljive relacije između simbola te teorije, njenih formula i nizova formula sa jedne strane, i prirodnih brojeva sa druge strane. Kao rezultat, svakom sintaktičkom objektu date teorije pridružen je njegov jedinstveni kôd ili Gedelov broj (Gedelov broj formule **A** označavaćemo sa $[\mathbf{A}]$). Na toj korespondenciji između sintaktičkih objekata neke formalne teorije i prirodnih brojeva, zasnovana je i korespondencija između svojstava i relacija između tih sintaktičkih objekata, koje ćemo zvati sintaktičkim svojstvima i relacijama, i aritmetičkih svojstava i relacija između njihovih kôdova. Primeri sintaktičkih svojstava i relacija su *biti formula*, *imati oblik negacije*, *biti aksioma datog formalnog sistema*, *biti dokaz neke formule u tom formalnom sistemu*, i slično. Nekom sintaktičkom svojstvu odgovara aritmetičko svojstvo ako i samo ako to aritmetičko svojstvo imaju svi i samo oni brojevi koji su kôdovi objekata sa datim sintaktičkim svojstvom. Slično tome, sintaktičkoj relaciji između određenih objekata odgovaraće ona aritmetička relacija u kojoj stoje svi i samo kôdovi tih sintaktičkih objekata. Tako će, na primer, svojstvu *biti formula* neke formalne teorije odgovarati aritmetičko svojstvo koje čini skup tačno onih prirodnih brojeva koji su kôdovi formula te teorije. Tvrđenja koja se tiču sintakse neke formalne teorije mogu tada biti formulisana kao aritmetička tvrđenja koja se tiču kôdova sintaktičkih objekata te teorije. Na taj način sintaktička svojstva i relacije koje karakterišu neku formalnu aritmetičku teoriju postaju blisko povezana sa predmetom kojim se sama ta teorija bavi.

Ali da li u toj teoriji mogu biti formulisana baš ona aritmetička tvrđenja koja su interpretirana tako da govore o njenoj sintaksi? Pošto termi te teorije referiraju na prirodne brojeve, za njih se može uzeti da time posredno referiraju i na sintaktičke objekte iste teorije, čiji su ti brojevi kôdovi. Međutim, i dalje nije jasno da se unutar iste teorije mogu konstruisati i formule koje pripisuju kôdovima tih objekata baš ona aritmetička svojstva i relacije koja odgovaraju na opisan način sintaktičkim svojstvima i relacijama. Da bi to opravdao, Gedel je ponudio sledeći rezultat koji pokazuje da to jeste moguće pod određenim uslovima.

2.2.3 *Predstavljive relacije*

Za mogućnost formulisanja aritmetičkih tvrđenja koja na opisani način odgovaraju tvrđenjima o sintaktičkim relacijama unutar aritmetičke formalne teorije, ključna je priroda tih relacija[4]. Naime, kako je Gedel pokazao, ako je neka relacija *rekurzivna*, onda je ona *predstavljiva* (*representable* ili *strongly representable*) unutar teorije PA i njenih proširenja.[5] Da je neka rekurzivna

4 Radi jednostavnijeg izlaganja u nastavku ćemo o svojstvima povremeno govoriti kao o unarnim relacijama.

5 To će važiti ne samo za PA nego i za svaku drugu konzistentnu, rekurzivno aksiomatsku teoriju izraženu u jeziku aritmetike koja širi Robinsonovu aritmetiku Q. Teorija Q je fragment teorije PA koji ne sadrži aksiomu ($PA7$). Umesto te aksiome, u teoriji Q imamo sledeću aksiomu: $\forall x\,(x = 0 \vee \exists y\; x = S\,(y))$ koja je dokaziva u PA.

n-arna relacija R predstavljiva u teoriji PA znači da postoji neka formula te teorije $\mathbf{A}(\mathbf{x_1},...,\mathbf{x_n})$ za koju važi da je $\mathbf{A}(\mathbf{x_1},...,\mathbf{x_n})$ dokazivo ako važi $R(x_1,...,x_n)$, dok je $\neg\mathbf{A}(\mathbf{x_1},...,\mathbf{x_n})$ dokazivo ako $R(x_1,...,x_n)$ ne važi. Takođe, svaka totalna rekurzivna n-arna funkcija $f(x_1,...,x_n)$ predstavljiva je u PA formulom $\mathbf{B}(\mathbf{x_1},...,\mathbf{x_n},\mathbf{y})$ koja je dokaziva ako važi $f(x_1,...,x_n)=y$, a čija je negacija dokaziva ako važi $f(x_1,...,x_n)\neq y$.

Relacije koje su rekurzivno nabrojive biće, sa druge strane, *slabo predstavljive* (*weakly representable*) u datoj teoriji. Neka n-arna relacija R' je slabo predstavljiva formulom $\mathbf{A}'(\mathbf{x_1},...,\mathbf{x_n})$ u slučaju da je ova formula dokaziva ako i samo ako važi $R'(x_1,...,x_n)$.

Odnos između predstavljivosti i slabe predstavljivosti neke relacije u teoriji T zavisi od osobina te teorije. Ako je teorija T konzistentna, onda formula koja predstavlja neku relaciju u T takođe slabo predstavlja tu relaciju u T. Sa druge strane, ako je teorija T potpuna, onda formula koja slabo predstavlja neku relaciju u T, takođe i predstavlja tu relaciju u T.

Određeni skupovi, relacije i funkcije koji se kao posledica aritmetizacije javljaju u istraživanju sintakse odgovarajuće aritmetičke formalne teorije jesu rekurzivni, pa su prema navedenom rezultatu i predstavljivi u toj teoriji. Takav je, na primer, skup kôdova svih formula teorije koji stoji za svojstvo *biti formula* te teorije. Pošto je taj skup rekurzivan, formule te teorije biće nabrojive na osnovu njihovih Gedelovih brojeva i, osim toga, postojaće efektivna procedura za određivanje toga da li je neki broj kôd neke formule ili ne. Standardno se zahteva da skup (kôdova) aksioma formalne teorije i njena pravila izvođenja takođe budu efektivno zadati. Inače, uloga formalne teorije da rasvetli relaciju dokaza i učini je odlučivom ne bi bila ostvariva. Ako ne možemo efektivno da odredimo da li su formule koje se javljaju u nekom nizu formula aksiome ili da li su one neposredne posledice nekih drugih formula istog niza, onda ne možemo ni efektivno odrediti da li je taj niz formula dokaz rečenice kojom se završava. Da bi to bilo moguće, skup kôdova aksioma date teorije treba da bude rekurzivan, i pored toga, svakom pravilu izvođenja treba da odgovara neka rekurzivna operacija na skupu kôdova formula te teorije. Ako je ovo poslednje slučaj onda će i relacija neposredne posledice u datoj teoriji biti rekurzivna. Zahvaljujući tome biće rekurzivna i relacija dokaza u istoj aritmetičkoj formalnoj teoriji, koja će samim tim biti i predstavljiva u toj teoriji. Na osnovu toga što je relacija dokaza u takvoj teoriji rekurzivna, znamo da će svojstvo dokazivosti u njoj biti rekurzivno nabrojivo. Naime, formula je dokaziva u nekoj teoriji ako i samo ako postoji njen dokaz u toj teoriji. Procedura utvrđivanja toga da li je ona dokaziva sastojala bi se tada u potrazi za njenim dokazom, koja će se, ako formula jeste dokaziva, u nekom trenutku završiti njegovim pronalaskom.

Svojstva i relacije pomenute u prethodnom odeljku su složene pa u dokazu njihove rekurzivnosti, odnosno, predstavljivosti u datoj teoriji, učestvuje i predstavljivost mnogih drugih svojstava i relacija. To su svojstva biti simbol, formula ili niz formula datog sistema ili imati određenu logičku formu, zatim relacija u kojoj neka formula, term i rezultat supstitucije tog terma mesto slobodne promenljive u toj formuli stoje itd. Gedel je pokazao da za svako od navedenih svojstava i relacija postoji primitivno rekurzivna funkcija koja ih izračunava. Na osnovu toga znamo da će za svako od njih postojati i formula koja ih predstavlja u datoj teoriji (cf. Gödel, 1931, pp. 163-170). Pomoću tih formula mogu biti konstruisane i formule koje predstavljaju neka složenija svojstva i relacije kao što je relacija dokaza.

Formula koja predstavlja relaciju dokaza u teoriji PA, i koja se obično označava sa $\mathbf{Prf}(\mathbf{x},\mathbf{y})$, dokaziva je u toj teoriji ako je y kôd dokaza formule kodirane sa x, dok je njena negacija dokaziva ako x i y ne stoje u toj relaciji (za konstrukciju te formule vidi: Gödel, 1931, p. 170). Formula koja slabo predstavlja dokazivost u PA bi tada, na osnovu njenog značenja, trebalo da bude dobijena egzistencijalnom generalizacijom ove formule. To je dakle formula $\exists\mathbf{y}\mathbf{Prf}(\mathbf{x},\mathbf{y})$ koju ćemo kraće označavati sa $\mathbf{Prov}(\mathbf{x})$. Ta formula dokaziva je u teoriji PA ako i samo ako je x kôd formule koja je dokaziva u toj teoriji.

Formula $\mathbf{Prov}(\mathbf{x})$ do koje smo došli na opisan način predstavlja egzistencijalnu generalizaciju formule koja ne sadrži neograničene kvantifikatore (vidi: Gödel, 1931, p. 170). Takva formula pripada posebnom nivou klasifikacije formula prema njihovoj složenosti, koji je deo takozvane *aritmetičke hijerarhije* skupova definisanih tim formulama (tu hijerarhiju uveli su Klini i Mostovski (Andrzej Mostowski)). Ta klasifikacija odnosi se na formule u preneksnoj normalnoj formi (ili njima ekvivalentne) u kojima se svi neograničeni kvantifikatori nalaze na početku formule. Složenost formule određuje broj smenjivanja blokova univerzalnog i blokova egzistencijalnog kvantifikatora u njenom kvantifikatorskom prefiksu. Ako je neka formula ekvivalentna formuli u kojoj se nijedan neograničeni kvantifikator ne javlja, onda je ona najnižeg stepena složenosti koji se označava sa Δ_0 (takođe sa Σ_0 i Π_0). Formula je Σ_n, za $n > 0$, ako počinje nekim blokom egzistencijalnih kvantifikatora koji je praćen Π_{n-1}- formulom. Sa druge strane formula je Π_n, za $n > 0$, ako počinje blokom univerzalnih kvantifikatora praćenih Σ_{n-1}- formulom.

Formula $\mathbf{Prov}\,(\mathbf{x})$ je prema toj klasifikaciji Σ_1-formula, pošto počinje egzistencijalnim kvantifikatorom primenjenim na formulu u kojoj se ne javlja nijedan neograničeni kvantifikator (pa je ta formula između ostalog i Π_0-formula). Formule te složenosti zauzimaju posebno mesto u kontekstu ispitivanja odnosa između teorije izračunljivosti i logike. Naime, kako se ispostavlja, svaka rekurzivno nabrojiva relacija je slabo predstavljiva nekom Σ_1-formulom odgovarajuće aritmetičke teorije. Svaka formula koja slabo predstavlja neku rekurzivno nabrojivu relaciju biće zapravo egzistencijalna generalizacija Δ_0-formule koja predstavlja odgovarajuću rekurzivnu relaciju. Teorija *PA* je Σ_1-potpuna teorija što znači da je svaka istinita Σ_1-rečenica formulisana na njenom jeziku dokaziva u njoj. Iz tog njenog svojstva i njene valjanosti, sledi da su u njoj dokazive sve i samo istinite Σ_1-rečenice. Upravo iz toga sledi da će svaka rekurzivno nabrojiva relacija biti slabo predstavljiva u *PA*. Osim toga, činjenica da je ona Σ_1-potpuna teorija biće takođe dokaziva u *PA*. Drugim rečima, za svaku Σ_1-rečenicu $\mathbf{A}$, važiće da je rečenica $\mathbf{A} \to \mathbf{Prov}\,([\mathbf{A}])$ dokaziva u *PA* (Rautenberg, 2010, p. 277). Ova svojstva teorije *PA* ključna su za primenu Gedelovih rezultata na tu teoriju.

Do sada smo ustanovili da postoje rečenice aritmetičke formalne teorije kao što je *PA*, koje predstavljaju tvrđenja o svojstvima i relacijama između sintaktičkih objekata iste te teorije. Ali da li postoje rečenice koje mogu predstavljati takva tvrđenja o samoj sebi? Da takve rečenice mogu biti formulisane u određenim teorijama koje sadrže aritmetiku posledica je rezultata koji ćemo predstaviti u nastavku.

2.2.4 *Lema o dijagonalizaciji*

LEMA O DIJAGONALIZACIJI (DIAGONAL LEMMA) dokazuje postojanje samoreferencijalnih rečenica unutar formalnih teorija o kojima je do sada bilo reči. Gedel navodi Karnapovu (Rudolf Carnap) *Logičku sintaksu jezika* (1934) kao delo u kojem se taj opšti rezultat prvi put pominje (Gödel, 1934, p. 363). LEMA O DIJAGONALIZACIJI tvrdi da u tim formalnim teorijama, za svaku formulu $\mathbf{A}\,(\mathbf{x})$, čija je $\mathbf{x}$ jedina slobodna promenljiva, postoji rečenica $\mathbf{B}$ koja je dokazivo ekvivalentna rezultatu supstitucije kôda rečenice $\mathbf{B}$ mesto $\mathbf{x}$ u $\mathbf{A}$. Drugim rečima, lema tvrdi da:

- za svaku formulu $\mathbf{A}\,(\mathbf{x})$, sa jednom slobodnom promenljivom $\mathbf{x}$, postoji rečenica $\mathbf{B}$ koja je dokazivo ekvivalentna sa $\mathbf{A}\,([\mathbf{B}])$.

Navedeno tvrđenje ne zahteva da $\mathbf{B}$ bude rečenica koja zapravo kaže za sebe da ona poseduje svojstvo predstavljeno formulom $\mathbf{A}\,(\mathbf{x})$. Jedino što se tvrdi jeste materijalna ekvivalencija dve rečenice, koja podrazumeva slaganje njihovih istinosnih vrednosti. Ipak, dokaz leme koji sadrži konstrukciju takve rečenice $\mathbf{B}$ pokazuje da ona zaista može biti samoreferencijalna.

Konstrukcija samoreferencijalne rečenice koristi *dijagonalizaciju* koja u kontekstu formalnog jezika predstavlja rezultat supstitucije kôda formule mesto njene slobodne promenljive. Operacija supstitucije, budući da je rekurzivna, predstavljiva je u teoriji kao što je *PA* formulom $\mathbf{Sub}\,(\mathbf{x}, \mathbf{y}, \mathbf{z})$ koja je dokaziva ako je z kôd rezultata supstitucije terma kodiranog sa y mesto slobodne promenljive u formuli kodiranoj sa x, a čija je negacija dokaziva ako ta relacija između brojeva x, y i z ne važi. Rečenica dobijena dijagonalizacijom formule kodirane sa x je formula kodirana sa y za koju je $\mathbf{Sub}(\mathbf{x},\mathbf{x},\mathbf{y})$ dokazivo. Umesto $\mathbf{Sub}(\mathbf{x},\mathbf{x},\mathbf{y})$ pisaćemo kraće $\mathbf{Diag}(\mathbf{x},\mathbf{y})$, gde ta formula stoji za 'y je kôd dijagonalizacije formule kodirane sa x'. Jedinstvenost dijagonalizacije neke formule, koja sledi iz jedinstvenosti njenog kôda, dokaziva je unutar iste teorije.

Dokaz leme sadrži metodu konstruisanja različitih samoreferencijalnih rečenica. Za svaku formulu $\mathbf{A}(\mathbf{x})$, možemo da definišemo formulu $\mathbf{A}'(\mathbf{x})$ koja za formulu čiji je kôd supstituisan mesto x tvrdi da njena dijagonalizacija ima svojstvo predstavljeno formulom $\mathbf{A}(\mathbf{x})$. Takva formula definisana je na sledeći način:

$$\mathbf{A}'(\mathbf{x}) =_{\mathbf{def}} \exists \mathbf{y}\,(\mathbf{A}(\mathbf{y}) \wedge \mathbf{Diag}(\mathbf{x},\mathbf{y})).$$

Neka je rečenica $\mathbf{B}$ dijagonalizacija formule $\mathbf{A}'(\mathbf{x})$, odnosno, neka je $\mathbf{B} =_{\mathbf{def}} \mathbf{A}'([\mathbf{A}'])$. Tada $\mathbf{B}$ tvrdi da dijagonalizacija formule $\mathbf{A}'(\mathbf{x})$ ima svojstvo predstavljeno formulom $\mathbf{A}$, odnosno, $\mathbf{B}$ tvrdi da $\mathbf{A}([\mathbf{A}'([\mathbf{A}'])])$ važi. Pošto je jedinstvena dijagonalizacija formule $\mathbf{A}'(\mathbf{x})$ sama rečenica $\mathbf{B}$, sledi da $\mathbf{B}$ zapravo tvrdi za sebe da ima svojstvo predstavljeno formulom $\mathbf{A}(\mathbf{x})$, odnosno, da $\mathbf{A}([\mathbf{B}])$ važi, što čini rečenicu $\mathbf{B}$ samoreferencijalnom. Za takvu rečenicu $\mathbf{B}$ biće dokazivo u sistemu da je ekvivalentna rečenici $\mathbf{A}([\mathbf{B}])$.

Ovaj dokaz sadrži metodu koja omogućava da za proizvoljno svojstvo predstavljivo u datoj teoriji konstruišemo rečenicu koja za sebe tvrdi da ima to svojstvo. Takvu samoreferencijalnu rečenicu možemo razumeti kao fiksnu tačku nekog predikata (tj. formule sa jednom slobodnom promenljivom) pošto se značenje te rečenice naizgled ne menja primenom tog predikata na nju. LEMA O DIJAGONALIZACIJI tako ima ulogu *teoreme o fiksnoj tački* koja tvrdi postojanje fiksne tačke predikata ili formula neke formalne teorije. Na osnovu njenog dokaza očigledno je da je operacija supstitucije i njena predstavljivost u teoriji ono što omogućava konstruisanje samoreferencijalnih rečenica u njoj. U nastavku ćemo videti najvažnije posledice koje samoreferencijalne rečenice mogu imati u teoriji u kojoj se javljaju.

2.2.5 *Gedelove teoreme o nepotpunosti*

Činjenica da dovoljno jaka formalna teorija sadrži samoreferencijalne rečenice može se koristiti za dokazivanje određenih osobina te teorije. Videli smo u prvom delu rada da je jedna od glavnih posledica primene različitih instanci funkcionalne samoreferencije u logici i matematici dokazivanje određenih negativnih rezultata. Sličnu primenu imaće i lingvistička samoreferencija, pre svega na osnovu njene uloge u dokazu GEDELOVIH TEOREMA O NEPOTPUNOSTI. Ti i slični negativni rezultati pokazuju principijelna ograničenja formalnih teorija u kojima se ta vrsta samoreferencije javlja. U ostatku ovog poglavlja predstavićemo glavne takve rezultate i ispitati njihovu vezu sa negativnim rezultatima teorije izračunljivosti.

2.2.5.1 *Prva teorema o nepotpunosti*

Gedelove teoreme tiču su gore opisanih formalnih teorija - onih koje sadrže aritmetiku (zbog čega su i rekurzivne funkcije i relacije predstavljive u njima) i za koje važi da je relacija dokaza unutar njih odlučiva. Kao što smo videli, Gedel je pokazao da formalna teorija te vrste sadrži formule koje nešto tvrde o sintaktičkim svojstvima iste te teorije, kao na primer, da ona dokazuje, ili ne dokazuje, određenu formulu. To se postiže konstrukcijom formule $\mathbf{Prov}(\mathbf{x})$, i njene negacije $\neg\mathbf{Prov}(\mathbf{x})$, na jeziku date teorije, mesto čije promenljive se može supstituisati kôd formule iste teorije. Na osnovu LEME O DIJAGONALIZACIJI znamo da postoji rečenica te teorije, nazovimo je $\mathbf{G}$, koja je dokazivo ekvivalentna rečenici $\neg\mathbf{Prov}([\mathbf{G}])$. Ta rečenica, koja se naziva *Gedelovom rečenicom*, i koja je konstruisana na način opisan u LEMI O DIJAGONALIZACIJI, kaže za sebe da je nedokaziva u teoriji kojoj pripada. Pitanje koje je Gedel postavio o ovoj rečenici i koje ga je dovelo do njegovih rezultata o nepotpunosti jeste da li su ta rečenica ili njena negacija dokazive unutar te teorije (za rečenicu čija je negacija dokaziva u nekoj teoriji, reći ćemo takođe da je ona *opovrgljiva* u toj teoriji). Odgovor je da, ako teorija ima određena vrlo poželjna svojstva, onda nijedna od te dve rečenice ne može biti dokaziva u njoj. Ako je teorija konzistentna, Gedelova rečenica u njoj nije dokaziva, a ako je osim toga i ω- konzistentna, onda isto važi i za njenu negaciju.

Konzistentnost je svojstvo formalne teorije da ne postoji takva rečenica za koju važi da su i ona i njena negacija dokazive u toj teoriji. Ako ne zadovoljava to svojstvo, formalna teorija ne može da ponudi sistematizaciju određenog znanja, jer u inkonzistentnoj formalnoj teoriji sve njene formule, kako istinite tako i lažne, postaju dokazive. Konzistentnost je tako minimalan uslov koji bi svaka formalna teorija trebalo da zadovolji. Sa druge strane, ω-konzistentnost je jače svojstvo za koje je potrebno da ne postoji formula $\mathbf{A}(\mathbf{x})$ za koju važi da je za svako n, dokazivo $\mathbf{A}(\mathbf{n})$, i da je osim toga dokazivo još i $\exists\mathbf{x}\neg\mathbf{A}(\mathbf{x})$. Valjanost formalne teorije, tj. činjenica da su sve njene teoreme istinite u njenom standardnom modelu, implicira njenu ω-konzistentnost. Zato, ako teorija nije ω-konzistentna, onda je neka lažna rečenica dokaziva u njoj. Za dokaz PRVE TEOREME O NEPOTPUNOSTI dovoljno je da za teoriju o kojoj je reč važi restrikcija ω-konzistentnosti na Σ_1-formule (taj uslov naziva se 1-konzistentnost). Svojstvo 1-konzistentnosti ekvivalentno je svojstvu Σ_1-valjanosti, odnosno, svojstvu da su sve dokazive Σ_1-rečenice istinite.

Dokaz da ni Gedelova rečenica $\mathbf{G}$, dokazivo ekvivalentna rečenici $\neg\mathbf{Prov}([\mathbf{G}])$, ni njena negacija nisu dokazive u ω-konzistentnoj formalnoj teoriji na čijem jeziku su te rečenice formulisane je prilično jednostavan. Da bismo to videli, pretpostavimo da je rečenica $\mathbf{G}$ dokaziva. Pošto formula $\mathbf{Prov}(\mathbf{x})$ slabo predstavlja svojstvo dokazivosti u toj teoriji, tada će i $\mathbf{Prov}([\mathbf{G}])$ biti dokazivo. Sa druge strane, pošto je rečenica $\mathbf{G}$ dokaziva, isto će važiti i za rečenicu koja joj je dokazivo ekvivalentna, što znači da će rečenica $\neg\mathbf{Prov}([\mathbf{G}])$ takođe biti dokaziva. Ona, međutim, protivreči rečenici za koju je već rečeno da je dokaziva u teoriji, što čini teoriju u kojoj je $\mathbf{G}$ dokazivo inkonzistentnom. Zato, ako je teorija konzistentna, rečenica $\mathbf{G}$ ne može biti dokaziva u njoj. Pretpostavimo tada da je rečenica $\neg\mathbf{G}$ dokaziva. Onda je dokaziva i njoj ekvivalentna rečenica $\exists\mathbf{y}\mathbf{Prf}([\mathbf{G}],\mathbf{y})$. Sa druge strane, pošto je rečenica $\neg\mathbf{G}$ dokaziva u datoj teoriji, rečenica $\mathbf{G}$ ne može biti dokaziva u njoj, ako je ta teorija konzistentna. Pošto ne postoji dokaz rečenice $\mathbf{G}$ u toj teoriji, za svaki broj n važiće da on ne kodira dokaz rečenice $\mathbf{G}$. Iz toga sledi, na osnovu jake predstavljivosti relacije dokaza, da je za svaki broj n, rečenica $\neg\mathbf{Prf}([\mathbf{G}],\mathbf{n})$ dokaziva. Pošto u ω-konzistentnoj formalnoj teoriji ne može važiti da je za svaki broj n, rečenica $\neg\mathbf{Prf}([\mathbf{G}],\mathbf{n})$ dokaziva, dok je u isto vreme i rečenica $\exists\mathbf{y}\mathbf{Prf}([\mathbf{G}],\mathbf{y})$ dokaziva, $\neg\mathbf{G}$ ne može biti dokazivo u takvoj formalnoj teoriji. Dakle, u ω-konzistentnoj formalnoj teoriji ni rečenica $\mathbf{G}$ ni njena negacija neće biti dokazive.

Na osnovu ovog dokaza, PRVA TEOREMA O NEPOTPUNOSTI tvrdi da u svakoj ω-konzistentnoj formalnoj teoriji koja sadrži aritmetiku i ima rekurzivni skup aksioma i pravila izvođenja, postoji rečenica za koju važi da ni ona ni njena negacija nisu dokazive u toj teoriji. Za takvu rečenicu kaže se da je *neodlučiva* u toj teoriji. Formalna teorija koja sadrži neku neodlučivu rečenicu je *sintaktički nepotpuna*. Dakle, ovom teoremom se tvrdi sintaktička nepotpunost ω-konzistentnih, rekurzivno aksiomatskih teorija koje sadrže aritmetiku (kao što je teorija PA).

Pošto neodlučiva rečenica korišćena u dokazu sintaktičke nepotpunosti aritmetičke teorije kaže za sebe da je nedokaziva u toj teoriji, i zaista jeste nedokaziva u njoj, ona će biti istinita. Teorija o kojoj je reč će zato biti i *semantički nepotpuna* jer postoji istinita rečenica koju ona ne dokazuje.

Nepotpunost konzistentnih, rekurzivno aksiomatskih teorija dokazana ovom Gedelovom teoremom ne može se prevazići njihovim proširenjem prethodno neodlučivim istinitim rečenicama ili nekim novim aksiomama i pravilima izvođenja iz kojih bi te rečenice trebalo da slede. U svakom proširenju takve formalne teorije, rečenica koja tvrdi za sebe da nije dokaziva u novoj formalnoj teoriji moći će ponovo da bude konstruisana, i ako je i nova teorija ω-konzistentna, ponovo će biti neodlučiva u njoj. Samim tim, ta rečenica će biti istinita, pa će i nova teorija biti semantički nepotpuna.

Navedeni dokaz PRVE TEOREME O NEPOTPUNOSTI često se poredi sa dokazom Tarskog da istinitost rečenica određenih formalnih teorija nije predstavljiva unutar istih tih teorija. Teorema Tarskog može se takođe odnositi na aritmetičku formalnu teoriju za koju je još pretpostavljeno da je valjana i da postoji interpretacija koja svaku od njenih rečenica čini istinitom ili lažnom. Pretpostavimo da je istinitost rečenica u teoriji sa datim osobinama predstavljiva njenim predikatom $\mathbf{T}(\mathbf{x})$. Taj predikat bio bi tada dokaziv za tačno one numerale koji označavaju kôdove istinitih rečenica te teorije, dok bi njegova negacija bila dokaziva za one numerale koji označavaju kôdove neistinitih rečenica ili koji ne označavaju kôdove rečenica te teorije. Pošto je negacija ovog predikata, tj. predikat $\neg\mathbf{T}(\mathbf{x})$, takođe deo teorije, LEMA O DIJAGONALIZACIJI garantuje da on ima fiksnu tačku u toj teoriji. Ta fiksna tačka biće rečenica, nazovimo je $\mathbf{L}$, koja je u datoj teoriji dokazivo ekvivalentna rečenici $\neg\mathbf{T}([\mathbf{L}])$. Pretpostavimo da je rečenica $\mathbf{L}$ istinita. Tada će rečenica $\mathbf{T}([\mathbf{L}])$ biti dokaziva u teoriji. Ta dokaziva rečenica neće, međutim, biti istinita jer će istinita biti njena negacija $\neg\mathbf{T}([\mathbf{L}])$ ekvivalentna rečenici $\mathbf{L}$. Pošto bi to protivrečilo pretpostavci o valjanosti date teorije, moramo zaključiti da rečenica $\mathbf{L}$ ipak nije istinita. Pretpostavimo tada da je $\mathbf{L}$ lažna rečenica. Tada će rečenica $\neg\mathbf{T}([\mathbf{L}])$ biti dokaziva u teoriji, pa će i njoj ekvivalentna rečenica $\mathbf{L}$ takođe biti dokaziva. Kako smo za tu rečenicu pretpostavili da je lažna, njena dokazivost opet protivreči pretpostavci o valjanosti teorije kojoj ona pripada. Dakle, ako je teorija o kojoj je reč valjana, njena rečenica $\mathbf{L}$ koja predstavlja fiksnu tačku predikata $\neg\mathbf{T}(\mathbf{x})$ ne može biti ni istinita ni lažna. To, međutim, protivreči drugoj pretpostavci o teoriji prema kojoj su u nekoj interpretaciji sve njene rečenice istinite ili lažne. Jedini način da izbegnemo tu protivrečnost, ako želimo da zadržimo pomenute pretpostavke, jeste da odbacimo mogućnost predstavljivosti istinitosti, a samim tim i neistinitosti, rečenica takve teorije unutar nje same. Pošto istinitost njenih rečenica nije predstavljiva u opisanoj aritmetičkoj teoriji, sledi i da skup istinitih rečenica formulisanih na jeziku te teorije nije rekurzivan.

Ovaj argument Tarskog može se posmatrati kao formalizacija argumenta u osnovi PARADOKSA LAŽLJIVCA, odnosno njegove verzije u kojoj je 'lažno' zamenjeno sa 'neistinito'. Dokaz Gedelove teoreme je, kao što smo videli, zasnovan na sličnom argumentu u kojem je predikat istinitosti zamenjen sa predikatom dokazivosti. Oba argumenta vode nekim negativnim rezultatima koji tvrde postojanje određenih ograničenja formalne teorije na koju se ti argumenti odnose. Dok rezultat Tarskog otkriva ograničenja *izražajne moći* formalnih teorija određene vrste, Gedelovi rezultati pokazuju da je *deduktivna snaga* tih teorija ograničena.

Gedelov dokaz PRVE TEOREME O NEPOTPUNOSTI pokazuje da dovoljno jaka ω-konzistentna teorija ne može za svaku rečenicu formulisanu na njenom jeziku da dokaže ili opovrgne tvrđenje da je ta rečenica dokaziva u njoj. Neke rečenice koje su instance formule koja slabo predstavlja svojstvo dokazivosti u teoriji neće biti odlučive u toj teoriji. Skup koji je slabo predstavljiv neodlučivom formulom unutar aritmetičke formalne teorije je rekurzivno nabrojiv, ali ne i rekurzivan[6]. Dokaz PRVE TEOREME O NEPOTPUNOSTI tako pokazuje da skup teorema ω-konzistentne formalne teorije koja sadrži aritmetiku nije rekurzivan, pošto formula koja ga slabo predstavlja ne može biti odlučiva u toj formalnoj teoriji. Navedena metamatematička razmatranja tako vode rezultatu koji se tiče pitanja koje pripada teoriji izračunljivosti. Veza između metamatematičkih rezultata i rezultata teorije izračunljivosti može se uspostaviti i u drugom smeru. Naime, videćemo da neodlučivost određenih formula odgovarajuće formalne teorije sledi iz rezultata teorije izračunljivost koji dokazuju rekurzivnu nerešivost određenih problema. Pre toga, predstavićemo još jedan dokaz PRVE TEOREME O NEPOTPUNOSTI koji

6 Za formulu sa slobodnim promenljivima reći ćemo da je neodlučiva ako je neka rečenica dobijena supstitucijom numerala mesto njenih slobodnih promenljivih neodlučiva u toj teoriji.

koristi neke slabije pretpostavke o formalnoj teoriji čiju nepotpunost dokazuje. Taj dokaz se takođe može dovesti u vezu sa odgovarajućim rezultatima teorije izračunljivosti.

2.2.5.2 *Gedel-Roserova teorema*

Gedelov dokaz PRVE TEOREME O NEPOTPUNOSTI modifikovan je od strane Rosera (Barkley Rosser) na način koji čini tu teoremu primenljivom na veći skup formalnih teorija (Rosser, 1936). Roserov dokaz ne koristi pretpostavku da je teorija o kojoj je reč ω-konzistentna (ili 1-konzistentna), već samo slabiju pretpostavku da je ona prosto konzistentna. Kako bi omogućio dokaz teoreme o nepotpunosti koji ne zavisi od pretpostavke o ω-konzistentnosti teorije, Roser je uveo novi predikat dokazivosti, označimo ga sa $\mathbf{Prf^R}(\mathbf{x},\mathbf{y})$, koji je definisan na sledeći način:

$$\mathbf{Prf^R}(\mathbf{x},\mathbf{y}) =_{\mathbf{def}} \mathbf{Prf}(\mathbf{x},\mathbf{y}) \wedge \neg\exists\mathbf{z} \leq \mathbf{y}\mathbf{Prf}(\mathbf{neg}(\mathbf{x}),\mathbf{z})$$

gde $\mathbf{Prf}(\mathbf{x},\mathbf{y})$ stoji za Gedelov predikat dokaza, a $\mathbf{neg}(\mathbf{x})$ stoji za numeričku operaciju koja predstavlja veznik negacije (to je operacija koja kada se primeni na kôd neke formule daje kao rezultat kôd negacije te formule). Formula $\mathbf{Prf^R}(\mathbf{x},\mathbf{y})$ intuitivno kaže da je y kôd dokaza formule kodirane sa x, i da ne postoji kraći dokaz negacije te formule. Roserov predikat dokazivosti $\mathbf{Prov^R}(\mathbf{x})$ definisan je kao egzistencijalna generalizacija formule $\mathbf{Prf^R}(\mathbf{x},\mathbf{y})$, odnosno kao $\exists\mathbf{y}\mathbf{Prf^R}(\mathbf{x},\mathbf{y})$. Ako je teorija o kojoj je reč konzistentna, onda će formula $\mathbf{A}$ biti dokaziva u njoj ako i samo ako je $\mathbf{Prov^R}([\mathbf{A}])$ dokazivo. Iz toga sledi da će u konzistentnom formalnom sistemu, rečenice $\mathbf{Prov^R}([\mathbf{A}])$ i $\mathbf{Prov}([\mathbf{A}])$ biti ekvivalentne. Ipak, pošto njihova ekvivalentnost nije dokaziva u teoriji (jer njena konzistentnost, kao što ćemo videti, nije dokaziva u njoj) ta dva predikata dokazivosti mogu imati različita dokazno-teorijska svojstva.

Na osnovu LEME O DIJAGONALIZACIJI znamo da negacija Roserovog predikata dokazivosti ima fiksnu tačku, odnosno, da postoji rečenica, nazovimo je $\mathbf{R}$, za koju je dokazivo da $\mathbf{R} \leftrightarrow \neg\mathbf{Prov^R}([\mathbf{R}])$. Intuitivno, ta rečenica kaže za sebe da ako je dokaziva, onda postoji kraći dokaz njene negacije. Ova rečenica, koja se naziva *Roserovom rečenicom*, koristi se u dokazu nepotpunosti konzistentne, rekurzivno aksiomatske teorije koja sadrži aritmetiku. U takvoj teoriji neće biti dokazive ni rečenica $\mathbf{R}$ ni njena negacija.

Ekvivalentnost Gedelovog i Roserovog predikata dokazivosti omogućava dokaz, analogan Gedelovom dokazu za rečenicu $\mathbf{G}$, da je rečenica $\mathbf{R}$ nedokaziva u konzistentnoj teoriji. Za dokaz da njena negacija takođe nije dokaziva, koristi se jedna neposredna posledica Roserove definicije dokazivosti. Naime, ta defincija garantuje da za svaku formulu $\mathbf{A}$, važi da ako je dokazivo $\mathbf{Prov^R}([\mathbf{A}])$ onda je dokazivo i $\neg\mathbf{Prov^R}(\mathbf{neg}[\mathbf{A}])$ (inače bi za dva prirodna broja m i n, koji kodiranju dokaze formula $\mathbf{A}$ i $\neg\mathbf{A}$, trebalo da važi da je svaki od njih manji od drugog, što je opovrgljivo u relevantnim teorijama). Ako je rečenica $\neg\mathbf{R}$ dokaziva, onda je takođe dokaziva i rečenica $\mathbf{Prov^R}(\mathbf{neg}[\mathbf{R}])$. Na osnovu pomenute posledice Roserove definicije predikata dokazivosti i zakona dvostruke negacije, sledi da je tada $\neg\mathbf{Prov^R}([\mathbf{R}])$ takođe dokazivo. Ali tada je dokazivo i $\mathbf{R}$ koje je ekvivalentno sa $\neg\mathbf{Prov^R}([\mathbf{R}])$. Dakle, $\neg\mathbf{R}$ je dokazivo u datoj teoriji samo ako su i $\mathbf{R}$ i $\neg\mathbf{R}$ dokazivi, što protivreči pretpostavci o konzistentnosti te teorije. Za rečenicu $\mathbf{R}$ tako važi da ni ona ni njena negacija nisu dokazive u konzistentnoj, rekurzivno aksiomatskoj teoriji koja sadrži aritmetiku, što svaku takvu teoriju čini nepotpunom.

Na taj način, koristeći nešto drugačiji predikat dokazivosti od onog koji je Gedel koristio, Roser je uspeo da formuliše rečenicu koja će biti neodlučiva u svakoj konzistentnoj teoriji na čijem jeziku je izražena. Na taj način je dokazana nepotpunost još veće klase formalnih teorija.

2.2.5.3 *Rekurzivno-teorijska verzija prve teoreme o nepotpunosti*

Kako je usput već spomenuto, postoji bliska veza između rezultata o neodlučivosti određenih formula, na kojima se zasniva teorema o nepotpunosti, i rezultata o nerešivosti određenih problema do kojih se dolazi u teoriji izračunljivosti. Ovde ćemo nešto detaljnije ispitati tu vezu. Ukratko, rekurzivno-teorijski pandan PRVE TEOREME O NEPOTPUNOSTI je dokaz postojanja rekurzivno nabrojivih skupova koji nisu rekurzivni, zbog čega je pitanje pripadnosti tim skupovima rekurzivno nerešivo.

Pojam predstavljivosti u PA (ili sličnoj aritmetičkoj teoriji) nudi jedan alternativni formalni model izračunljivosti ekvivalentan onom predloženom od strane teorije rekurzija (ta ekvivalentnost sledi iz samih uslova predstavljivosti relacija i funkcija). Rekurzivno nabrojivi skupovi i relacije su slabo predstavljivi formulama teorije PA, dok su rekurzivni skupovi i relacije predstavljivi njenim formulama koje su odlučive. Konkretno, rekurzivno nabrojiv skup S je slabo predstavljiv formulom $\mathbf{S}(\mathbf{x})$ takvom da je za svaki prirodan broj n, $\mathbf{S}(\mathbf{n})$ dokazivo ako i samo ako $n \in S$. Sa druge strane, rekurzivan skup S' predstavljiv je formulom $\mathbf{S}'(\mathbf{x})$, takvom da za svaki prirodan broj n važi da ako $n \in S'$, onda je rečenica $\mathbf{S}'(\mathbf{n})$ dokaziva, i ako $n \notin S'$, onda je rečenica $\neg\mathbf{S}'(\mathbf{n})$ dokaziva. U ovoj formalizaciji, procedura kojom bi problem pripadnosti nekog broja n skupu S trebalo da bude rešen, predstavlja potragu za dokazom rečenice $\mathbf{S}(\mathbf{n})$. Ta potraga daće pozitivan odgovor ako se završava pronalaženjem dokaza te rečenice, dok će dati negativan odgovor ako se završava pronalaženjem dokaza njene negacije. Da je prvo slučaj, tj. da je rečenica $\mathbf{S}(\mathbf{n})$ dokaziva u datoj teoriji može se izraziti u istoj toj teoriji rečenicom $\mathbf{Prov}([\mathbf{S}(\mathbf{n})])$. Skup S je rekurzivan ako i samo ako je za svako n, rečenica $\mathbf{S}(\mathbf{n})$ odlučiva, odnosno, ako je za svako n, $\mathbf{Prov}([\mathbf{S}(\mathbf{n})])$ ili $\mathbf{Prov}(\mathbf{neg}[\mathbf{S}(\mathbf{n})])$ dokazivo. Dakle, ako S nije rekurzivan skup, onda za neko n, rečenica $\mathbf{S}(\mathbf{n})$ neće biti odlučiva.

Postojanje neodlučivih formula teorije PA tako sledi iz postojanja skupova koji su rekurzivno nabrojivi ali ne i rekurzivni. Neodlučive formule te teorije biće one koje slabo predstavljaju takve skupove. Konkretni primeri tih formula i njihovih neodlučivih instanci zavisiće od nerekurzivnog skupa o kojem je reč.

U prvom delu ovog poglavlja sreli smo se sa nekim skupovima koji su rekurzivno nabrojivi ali ne i rekurzivni. Jedan od njih bio je i skup

$$K = \{x : \varphi_x(x) \text{ je definisano}\}$$

gde φ_x stoji za proizvoljnu rekurzivnu funkciju. Ako označimo sa W_x skup svih brojeva x za koje je funkcija φ_x definisana, skup K možemo zapisati i kao skup $\{x : x \in W_x\}$. Na osnovu definicije skupa K sledi da je neki broj n njegov element ako i samo ako važi da je $n \in W_n$. Pošto je skup W_n slabo predstavljiv u sistemu formulom kodiranom brojem n, važiće da je $n \in W_n$ ako i samo ako je dijagonalizacija te formule dokaziva, odnosno, ako postoji neko a za koje važi:

$$\mathbf{Prov}(\mathbf{a}) \land \mathbf{Diag}(\mathbf{n}, \mathbf{a}).$$

Uopštavanjem ovog slučaja dolazimo do toga da je skup K u ovoj formalizaciji predstavljen formulom:

$$\exists \mathbf{y}(\mathbf{Prov}(\mathbf{y}) \land \mathbf{Diag}(\mathbf{x}, \mathbf{y}))$$

koja kaže da je dijagonalizacija formule kodirane brojem x dokaziva. Na osnovu opisane veze između dokazivosti formula i rekurzivnosti njima slabo predstavljenih skupova, sledi da ova formula odgovara Klinijevom predikatu $\exists y T(x, x, y)$ koji tvrdi postojanje konačnog izračunavanja dijagonalizacije funkcije kodirane brojem x. Iz toga što skup K, pa samim tim i Klinijev predikat, nisu rekurzivni, sledi da ni formula $\exists \mathbf{y}(\mathbf{Prov}(\mathbf{y}) \land \mathbf{Diag}(\mathbf{x}, \mathbf{y}))$ kojom je taj skup slabo predstavljiv u PA nije odlučiva.

Služeći se argumentom kojim se pokazuje da komplement skupa K, tj. skup $\bar{K}$, nije rekurzivno nabrojiv, možemo naći i konkretan primer rečenice koja će biti neodlučiva u datoj teoriji. Kada bi formula $\exists \mathbf{y}\left(\mathbf{Prov}\left(\mathbf{y}\right) \wedge \mathbf{Diag}\left(\mathbf{x},\mathbf{y}\right)\right)$ bila odlučiva, onda bi skup $\bar{K}$ bio slabo predstavljiv formulom:

$$\forall \mathbf{y} \neg \left(\mathbf{Prov}\left(\mathbf{y}\right) \wedge \mathbf{Diag}\left(\mathbf{x},\mathbf{y}\right)\right)$$

ekvivalentnom njenoj negaciji. Iz toga bi sledilo da je skup $\bar{K}$ rekurzivno nabrojiv. Ta formula kaže da dijagonalizacija formule sa Gedelovim brojem x nije dokaziva. Posmatrajmo dijagonalizaciju te formule. Ako je njen Gedelov broj m, onda je njena dijagonalizacija sledeća rečenica:

$$\forall \mathbf{y} \neg \left(\mathbf{Prov}\left(\mathbf{y}\right) \wedge \mathbf{Diag}\left(\mathbf{m},\mathbf{y}\right)\right).$$

Ova rečenica kaže za sebe da nije dokaziva. Ako formula čija je ona instanca slabo predstavlja skup $\bar{K}$ onda ta rečenica takođe kaže za sebe da je njen Gedelov broj u skupu $\bar{K}$. Kada bi ona bila odlučiva, onda bi bilo odlučivo i pitanje da li je kôd funkcije koja rekurzivno nabraja skup $\bar{K}$ element tog skupa, što smo videli da vodi kontradikciji (sličnoj onoj u osnovi RASELOVOG PARADOKSA). Iz toga sledi da je navedena rečenica neodlučiva. Time smo došli do zaključka da rečenica koja tvrdi za sebe da nije dokaziva, pa tako predstavlja fiksnu tačku predikata $\neg\mathbf{Prov}\left(\mathbf{x}\right)$, nije odlučiva u formalnoj teoriji kao što je PA. Tako smo na osnovu rezultata teorije rekurzija da skup $\bar{K}$ nije rekurzivno nabrojiv došli do Gedelovog rezultata o neodlučivosti fiksne tačke predikata $\neg\mathbf{Prov}\left(\mathbf{x}\right)$.

Formula $\forall \mathbf{y} \neg \left(\mathbf{Prov}\left(\mathbf{y}\right) \wedge \mathbf{Diag}\left(\mathbf{x},\mathbf{y}\right)\right)$ odgovara predikatu $\forall y \neg T\left(x,x,y\right)$, za koji je Klini pokazao da ne može biti formulisan u bilo kojoj potpunoj i konzistentnoj formalnoj teoriji (Kleene, 1943). Dakle, konzistentna formalna teorija u kojoj ovaj predikat može biti formulisan, kao što je PA, ne može biti potpuna. Taj rezultat Klini je smatrao *opštim oblikom* teoreme o nepotpunosti.

Klini je ponudio još jedan rekurzivno-teorijski argument u prilog postojanju neodlučivih formula konzistentne aritmetičke teorije (Kleene, 1936). Taj argument oslanja se na dokaz nerešivosti jednog drugog problema teorije rekurzija - problema koji se tiče toga da li neki jednakosni sistem formalne teorije rekurzivnih funkcija definiše *totalnu* rekurzivnu funkciju. Zahvaljujući kodiranju tih jednakosnih sistema, ovaj problem može biti formulisan i kao aritmetički problem. Naime, jednakosni sistem kodiran brojem e definiše n-arnu totalnu rekurzivnu funkciju ako i samo ako važi da $\forall x_1 ... \forall x_n \exists y T\left(e, x_1, ..., x_n, y\right)$. Taj uslov, kao što smo videli, može biti izražen unutar aritmetičke teorije odgovarajućom formulom. Neke instance te formule moraće onda da budu neodlučive jer bi inače ona mogla biti korišćena za odlučivanje pitanja totalnosti rekurzivnih funkcija. Skup svih totalnih rekurzivnih funkcija bi tada bio i rekurzivno nabrojiv, pa bismo na njega mogli da primenimo dijagonalni argument kojim bismo pokazali da taj skup ipak nije potpun (za detalje vidi: Kleene, 1936, pp. 738-742).

2.2.5.4 *Rekurzivno-teorijska verzija Gedel-Roserove teoreme*

Roserov dokaz PRVE TEOREME O NEPOTPUNOSTI takođe sledi iz dokaza nerešivosti određenog problema teorije rekurzija. Taj problem tiče se *rekurzivne razdvojivosti* parova disjunktnih skupova prirodnih brojeva. Svakom paru disjunktnih skupova S i T može biti pripisana funkcija g definisana na sledeći način:

$$g\left(n\right) = \begin{cases} 0, & \text{ako } n \in S \\ 1, & \text{ako } n \in T \\ \text{nedefinisano}, & \text{inače.} \end{cases}$$

Taj par skupova biće *rekurzivno razdvojiv* ako funkcija g ima totalno rekurzivno proširenje. Ako takvo proširenje funkcije g ne postoji, onda je dati par skupova *rekurzivno nerazdvojiv.*

Sa druge strane, isti par skupova S i T biće *razdvojiv u formalnom sistemu* ako postoji formula tog formalnog sistema $\mathbf{A}(\mathbf{x})$ takva da za svako n:

$$\begin{aligned} &\text{ako } n \in S, \ \ \mathbf{A}(\mathbf{n}) \text{ je dokazivo;} \\ &\text{ako } n \in T, \ \ \neg\mathbf{A}(\mathbf{n}) \text{ je dokazivo.} \end{aligned}$$

Svaki par rekurzivno nabrojivih skupova biće razdvojiv u teoriji PA i sličnim teorijama. To sledi iz činjenice da je svaki rekurzivno nabrojiv skup slabo predstavljiv nekom Σ_1-formulom te teorije. Ako je skup S slabo predstavljiv formulom $\exists\mathbf{x}\mathbf{A_0}(\mathbf{y},\mathbf{x})$, a skup T formulom $\exists\mathbf{x}\mathbf{B_0}(\mathbf{y},\mathbf{x})$, onda je par ta dva skupa razdvojiv u istom formalnom sistemu formulom $\mathbf{A}(\mathbf{y}) =_{\mathbf{def}} \exists\mathbf{x}(\mathbf{A_0}(\mathbf{y},\mathbf{x}) \wedge \neg\exists\mathbf{z} < \mathbf{x}\mathbf{B_0}(\mathbf{y},\mathbf{z}))$.

Na osnovu određenih primera znamo da postoje parovi disjunktnih skupova koji nisu rekurzivno razdvojivi. Iz toga sledi da je formula koja razdvaja takav par skupova u formalnom sistemu neodlučiva. Jer, kada bi ta formula bila odlučiva, mogli bismo da je koristimo za definisanje totalne rekurzivne funkcije koja bi bila pripisana datom paru skupova, što bi protivrečilo njihovoj rekurzivnoj nerazdvojivosti. Ovo je u kratkim crtama sadržaj Klinijeve SIMETRIČNE FORME TEOREME O NEPOTPUNOSTI (Kleene, 1950; vidi takođe: Beklemishev, 2010, pp. 10-11).

Za primer rekurzivno nerazdvojivog para skupova posmatrajmo skup brojeva x takvih da $\varphi_x(x) = 0$ i skup brojeva x takvih da je $\varphi_x(x)$ definisano i različito od 0. Tom paru skupova možemo pripisati funkciju f definisanu na sledeći način:

$$f(x) = \begin{cases} 1, & \text{ako } \varphi_x(x) = 0 \\ 0, & \text{ako je } \varphi_x(x) \text{ definisano i } \varphi_x(x) \neq 0 \\ \text{nedefinisano,} & \text{inače.} \end{cases}$$

Pretpostavimo da ova funkcija ima totalno rekurzivno proširenje g. Funkcija g bi tada bila jednaka nekoj totalnoj rekurzivnoj funkciji φ_n. Pretpostavimo da je $\varphi_n(n) = 0$. Pošto je funkcija g proširenje funkcije f, njene vrednosti se slažu sa vrednostima funkcije f za sve argumente za koje je funkcija f definisana, pa iz toga i date pretpostavke sledi da je $g(n) = 1$. Međutim, tada bi moralo da bude slučaj da je $\varphi_n(n) = 1$, što protivreči našoj pretpostavci. Sa druge strane, ako je $\varphi_n(n) \neq 0$, onda je $g(n) = 0$, odnosno, $\varphi_n(n) = 0$, što je opet protivrečno. Dakle, funkcija f (definisana, inače, pomoću dijagonalnog argumenta) ne može imati totalno rekurzivno proširenje.

Ipak, pošto su oba navedena skupa rekurzivno nabrojiva, ovaj par skupova biće razdvojiv u PA nekom formulom $\mathbf{C}(\mathbf{x})$, konstruisanom na gore opisan način. Koristeći tu formulu, možemo definisati funkciju F koja je rekurzivno proširenje funkcije f:

$$F(x) = \begin{cases} 1, & \text{ako je } \mathbf{C}(\mathbf{x}) \text{ dokazivo} \\ 0, & \text{ako je } \neg\mathbf{C}(\mathbf{x}) \text{ dokazivo} \\ \text{nedefinisano,} & \text{inače.} \end{cases}$$

Kada bi formula $\mathbf{C}(\mathbf{x})$ bila odlučiva, odnosno, kada bi za svako n važilo da je ili $\mathbf{C}(\mathbf{n})$ ili $\neg\mathbf{C}(\mathbf{n})$ dokazivo, funkcija F bila bi totalna rekurzivna funkcija. Znamo da to ne može biti slučaj, jer bi tada funkcija F rekurzivno razdvajala dati par skupova za koji smo videli da je rekurzivno nerazdvojiv. Dakle, formula $\mathbf{C}(\mathbf{x})$ mora imati neke neodlučive instance. Na osnovu prethodnog argumenta znamo i da je jedna od tih instanci rečenica koja je rezultat supstitucije kôda funkcije F mesto x u $\mathbf{C}(\mathbf{x})$.

Kako bismo pokazali da neodlučivost Roserovog predikata dokazivosti sledi iz ovog opšteg rezultata, primetimo da kada bi par skupova dokazivih i opovrgljivih rečenica bio rekurzivno razdvojiv, onda bi takav bio i par skupova iz navedenog primera. Naime, u tom slučaju postojala bi funkcija koja rekurzivno razdvaja skupove dokazivih i opovrgljivih rečenica (uključujući i instance formule $\mathbf{C}(\mathbf{x})$) i služeći se tom funkcijom mogli bismo da definišemo totalno proširenje funkcije F. To znači da skup dokazivih i opovrgljivih rečenica ne može biti rekurzivno razdvojiv, iz čega neposredno sledi neodlučivost Roserovog predikata koji razdvaja taj par skupova u formalnom sistemu.

Tako na osnovu rezultata teorije izračunljivosti o rekurzivnoj nerešivosti problema koji se tiču pripadnosti nekom skupu, ili o rekurzivnoj nerazdvojivosti nekih parova skupova, dolazimo do toga da moraju postojati neodlučive rečenice formalnih teorija u kojima su ti skupovi slabo predstavljivi. Osim toga, na osnovu nerešivosti nekih posebnih problema i nerazdvojivosti posebnih parova skupova, može se pokazati neodlučivost baš onih rečenica koje su Gedel i Roser koristili kako bi dokazali nepotpunost takvih formalnih teorija. Pored toga što, kako se ispostavlja, rezultati o nerešivosti problema i o neodlučivosti formula slede jedni iz drugih, do njih se takođe dolazi na sličan način. Nakon što predstavimo još neke rezultate koji pokazuju ograničenja formalnih teorija, posvetićemo se detaljnijem ispitivanju načina na koji se dolazi do tih negativnih rezultata teorije izračunljivosti i metamatematike, pre svega iz perspektive samoreferencije koja u oba slučaja ima veoma bitnu ulogu.

2.2.5.5 *Druga teorema o nepotpunosti*

Gedelova DRUGA TEOREMA O NEPOTPUNOSTI pruža još jedan primer neodlučive rečenice na jeziku konzistentne formalne teorije koja sadrži aritmetiku. To je rečenica koja tvrdi da je sama ta teorija konzistentna. Zahvaljujući gore opisanim sredstvima rekurzivno aksiomatskih teorija koje sadrže aritmetiku, u njima će biti moguće formulisati rečenicu koja predstavlja prirodnu formalizaciju tog tvrđenja. Na primer, takva je rečenica $\neg\exists\mathbf{x}\,(\mathbf{Prov}\,(\mathbf{x}) \wedge \mathbf{Prov}\,(\mathbf{neg}\ \mathbf{x}))$ koja kaže da ne postoji rečenica date teorije za koju važi da su i ona i njena negacija dokazive, ili rečenica $\neg\mathbf{Prov}\,([\bot])$ koja negira dokazivost apsurda u toj teoriji. Prema DRUGOJ TEOREMI O NEPOTPUNOSTI, nijedna od tih rečenica koje su ekvivalentne tvrđenju o konzistentnosti neke konzistentne formalne teorije neće biti dokaziva u toj teoriji.

Dokaz ove teoreme sledi iz dokaza PRVE TEOREME O NEPOTPUNOSTI. Preciznije, njen dokaz sledi iz mogućnosti formalizacije dokaza prve teoreme unutar sistema na koji se ta teorema odnosi. Gedel je u radu u kom je dokazao PRVU TEOREMU O NEPOTPUNOSTI takođe tvrdio da je takva formalizacija njenog dokaza moguća, i da je posledica toga upravo neodlučivost pitanja konzistentnosti teorije unutar nje same. On, međutim, nije naveo detalje o toj formalizaciji. Dokaz je kasnije razrađen od strane drugih logičara (Hilbert & Bernays, 1968). Jedino svojstvo predikata dokazivosti koje je potrebno za dokaz PRVE TEOREME O NEPOTPUNOSTI jeste da je neka formula $\mathbf{A}$ dokaziva u datoj teoriji ako i samo ako je formula $\mathbf{Prov}\,([\mathbf{A}])$dokaziva u njoj, što znači da formula $\mathbf{Prov}\,(\mathbf{x})$ slabo predstavlja svojstvo dokazivosti u toj teoriji. Za dokaz DRUGE TEOREME O NEPOTPUNOSTI, neka dodatna svojstva predikata dokazivosti moraju biti pretpostavljena. Svojstva tog predikata korišćena u dokazu druge teoreme poznata su kao *Hilbert-Bernajsovi uslovi izvodivosti* (*Hilbert-Bernays Derivability conditions*). Njih je kasnije unapredio Leb (Martin Löb) i ti uslovi u njegovoj verziji postali su standardni. Osim njih, za dokaz je potrebna i pretpostavka o gore pomenutom svojstvu teorije *PA* (i drugih makar jednako jakih teorija) koje se takođe tiče predikata dokazivosti, a to je da je za svaku Σ_1-formulu $\mathbf{A}$, dokazivo u *PA* da $\mathbf{A} \to \mathbf{Prov}\,([\mathbf{A}])$. Ako navedeni uslovi koji se tiču predikata dokazivosti važe u nekoj teoriji, onda u njoj može biti dokazano da ako je ta teorija konzistentna, onda Gedelova rečenica nije dokaziva u njoj (ovo sumira prvi deo dokaza PRVE TEOREME O NEPOTPUNOSTI). Pretpostavimo sada da je konzistentnost teorije dokaziva u toj teoriji. Tada je, zahvaljujući tome i dokazivosti navedene implikacije, dokazivo i da Gedelova rečenica nije dokaziva u njoj. Ali pošto Gedelova rečenica kaže za sebe da nije dokaziva u toj teoriji, to bi značilo da je Gedelova rečenica zapravo dokaziva u njoj. Međutim, znamo da to ne može biti slučaj u konzistentnoj formalnoj teoriji. Time smo primorani na zaključak da konzistentnost formalne teorije koja sadrži aritmetiku i koja jeste konzistentna ne može biti dokaziva u istoj teoriji.

Na taj način DRUGA TEOREMA O NEPOTPUNOSTI pokazuje još jedno ograničenje formalnih teorija za koje PRVA TEOREMA O NEPOTPUNOSTI važi, a to je da one nisu dovoljno jake da bi dokazale sopstvenu konzistentnost, ako zapravo jesu konzistentne. Dokaz njihove konzistentnosti zahteva korišćenje nekih sredstva kojima date teorije ne raspolažu. Tako je, na primer, Gencen (Gerhard Gentzen) dokazao konzistentnost od PA korišćenjem transfinitne indukcije do ordinala ε_0 (koji je karakterističan po tome što predstavlja najmanju fiksnu tačku funkcije f koja za svaki ordinal α, kao vrednost daje ω^α). Transfinitna indukcija do ε_0 prevazilazi snagu teorije PA zbog čega Gencenov dokaz njene konzistentnosti ne može biti formalizovan u toj teoriji.

Rečenica koja tvrdi konzistentnost teorije kojoj pripada je takođe u izvesnom smislu samoreferencijalna. Naime, ona referira na sve dokaze i rečenice te teorije, pa indirektno referira i na samu sebe. Tako opet nailazimo na upotrebu samoreferencije u dokazu nekog negativnog rezultata. Ti rezultati pokazuju da iako dovoljno jake formalne teorije mogu opisivati sopstvena sintaktička svojstva, postoji stroga granica u tome šta se o tim svojstvima može dokazati unutar tih teorija.

Obe teoreme o nepotpunosti postavljaju neka ograničenja teorija u kojima je lingvistička samoreferencija moguća. U narednom odeljku videćemo još jedan primer upotrebe lingvističke samoreferencije sa još jednim ograničavajućim rezultatom kao posledicom. Problem sa kojim ćemo se susresti u ovom slučaju sličan je onom koji vodi PRVOJ TEOREMI O NEPOTPUNOSTI. Taj problem, poznat kao *Henkinov problem* (Leon Henkin), tiče se pitanja da li je rečenica koja za sebe tvrdi da je dokaziva u teoriji u kojoj je formulisana dokaziva u njoj ili ne.

2.2.6 *Lebova teorema*

Teorema koja daje pozitivan odgovor na upravo postavljeno pitanje je LEBOVA TEOREMA. Ta teorema tvrdi i nešto više, a to je da za proizvoljnu formulu $\mathbf{A}$, ako je dokazivo u datoj formalnoj teoriji da $\mathbf{Prov}([\mathbf{A}]) \to \mathbf{A}$, onda je $\mathbf{A}$ takođe dokazivo u njoj. Dokaz ove teoreme koristi fiksnu tačku formule $\mathbf{Prov}(\mathbf{x}) \to \mathbf{A}$, odnosno, rečenicu $\mathbf{B}$ koja je dokazivo ekvivalentna rečenici $\mathbf{Prov}([\mathbf{B}]) \to \mathbf{A}$. Leb konstruiše tu fiksnu tačku koristeći metodu iz dokaza LEME O DIJAGONALIZACIJI, što znači da zapravo konstruiše rečenicu koja za sebe kaže da ako je ona dokaziva, onda $\mathbf{A}$. Razmatranja o toj fiksnoj tački, konkretno formalizacija argumenta koji pokazuje da ako je $\mathbf{B}$ dokazivo, onda je dokazivo i $\mathbf{A}$, zajedno sa Lebovim uslovima vode dokazu da ako je $\mathbf{Prov}([\mathbf{A}]) \to \mathbf{A}$ dokazivo u sistemu, onda je $\mathbf{A}$ takođe dokazivo (za detalje dokaza vidi: Löb, 1955, p. 117). Pozitivan odgovor na Henkinov problem sledi neposredno iz ove teoreme.

Fiksna tačka korišćena u ovom dokazu veoma je slična rečenici u osnovi KARIJEVOG PARADOKSA, tj. rečenici B ekvivalentnoj rečenici 'ako je B istinito, onda A', gde je A proizvoljna rečenica. Nije teško pokazati da takva rečenica B mora biti istinita (što smo objasnili u uvodu). Takođe nije teško videti da ako je ta rečenica istinita, onda je A takođe istinito. Time dolazimo do paradoksalnog zaključka da za svaku rečenicu koja se javlja u datoj fiksnoj tački na mestu rečenice A možemo pokazati da je istinita. Sa druge strane, LEBOVA TEOREMA ne omogućava dokaz proizvoljne rečenice $\mathbf{A}$ koja je deo fiksne tačke koja se koristi u njenom dokazu, već samo takve za koju je $\mathbf{Prov}([\mathbf{A}]) \to \mathbf{A}$ takođe dokazivo.

Kako je Smorinski (Craig Smoryski) primetio, LEBOVA TEOREMA pokazuje dodatno ograničenje formalnih teorija u kojima je lingvistička samoreferencija moguća (Smoryński, 1991, p. 119). Naime, iz te teoreme sledi da nisu sve instance formule koja tvrdi valjanost teorije o kojoj je reč dokazive u toj teoriji. Valjanost neke teorije, tj. činjenica da su sve formule dokazive u njoj istinite, može biti izražena formulom $\mathbf{Prov}([\mathbf{C}]) \to \mathbf{C}$, gde $\mathbf{C}$ stoji za proizvoljnu formulu te teorije. Prema LEBOVOJ TEOREMI, jedine instance te formule koje će biti dokazive u datoj teoriji su one koje se tiču formula zaista dokazivih u njoj. Te instance će, međutim, biti trivijalno dokazive jer njihovi dokazi slede iz dokaza tih formula samo na osnovu pravila za uvođenje implikacije.

LEBOVA TEOREMA tako predstavlja još jedan primer upotrebe LEME O DIJAGONALIZACIJI, odnosno lingvističke samoreferencije čiju mogućnost ona dokazuje, za postavljanje još jednog ograničenja formalnih teorija u kojoj se ona javlja.

Sve teoreme predstavljene u ovom delu rada ilustruju način korišćenja lingvističke samoreferencije za dolaženje do određenih, tipično negativnih, metamatematičkih rezultata. One su po tome slične negativnim rezultatima teorije izračunljivosti do kojih se dolazi korišćenjem određenih instanci funkcionalne samoreferencije. Već smo imali prilike da vidimo koliko su rezultati te dve vrste zapravo slični i međusobno zavisni. Poređenje tih rezultata takođe pokazuje da funkcionalna i lingvistička samoreferencija u njihovoj osnovi dele mnoga svojstva, da se javljaju u sličnim oblicima i imaju slične primene. Ta činjenica motiviše pitanje o postojanju nekog zajedničkog jezgra dve vrste samoreferencije koje bi objasnilo sličnost u načinu njihovog javljanja i njihove primene. Tim pitanjem bavićemo se u sledećem poglavlju. Osim toga, bilo bi korisno ispitati i šta je to što čini samoreferenciju glavnim sredstvom za izvođenje paradoksa i dokazivanje negativnih rezultata u određenim oblastima logike i matematike. Postoji li neka zajednička forma tih negativnih posledica samoreferencijalnih metoda i definicija koja bi nam pomogla da ih bolje razumemo? U nastavku ćemo predstaviti odgovore na bar neka od tih pitanja.

2.3 OPŠTA VERZIJA TEOREME O FIKSNOJ TAČKI

Pošto smo razmotrili različite načine na koje se samoreferencija javlja u logici (kao funkcionalna i lingvistička samoreferencija), u ovom delu rada ćemo ih međusobno uporediti i izdvojiti neke ključne osobine teorija koje omogućavaju opisivanje tih različitih samoreferencijalnih fenomena. U tome će nam pomoći jedan opšti Smalijanov rezultat o fiksnoj tački za koji se može uzeti da jednako dobro opisuje i funkcionalni i lingvistički oblik samoreferencije. Zahvaljujući tome videćemo da te dve vrste samoreferencije imaju puno toga zajedničkog. Njihova zajednička struktura će se takođe pokazati reprezentativnom i u odnosu na različite instance samoreferencije koje se javljaju u drugim oblastima matematike.

Kao što smo videli, i funkcionalna i lingvistička samoreferencija se mogu javiti u različitim oblicima, od kojih je jedan onaj u osnovi metode *dijagonalizacije*. Tu se samoreferencija javlja u obliku primene funkcije na sopstveni indeks (tj. kôd njene definicije) ili u obliku formule koja referira na sopstveni Gedelov broj. Samoreferencija u obliku samoprimene funkcija i formula, koja je u osnovi dijagonalizacije, omogućava i jednu drugu vrstu samoreferencije u vezi istih objekata. Ta druga vrsta samoreferencije, opisana DRUGOM TEOREMOM REKURZIJE i LEMOM O DIJAGONALIZACIJI, tiče se funkcija i formula čije definicije sadrže referenciju na njih same, odnosno, na njima definisane funkcije i formule. Koliko su zapravo ta dva rezultata međusobno slična, postaje naročito jasno ako se uzme u obzir ona verzija LEME O DIJAGONALIZACIJI koja se odnosi na formule sa više od jedne slobodne promenljive. U toj verziji, LEMA O DIJAGONALIZACIJI tvrdi da za svaku formulu $A(y, x_1, ..., x_n)$, postoji formula $B(x_1, ..., x_n)$, takva da je sledeća ekvivalencija dokaziva: $B(x_1, ..., x_n) \leftrightarrow A([B], x_1, ..., x_n)$. Lema u ovoj verziji predstavlja tvrđenje gotovo identično onom sadržanom u Klinijevoj DRUGOJ TEOREMI REKURZIJE prema kojoj za svaku rekurzivnu funkciju ψ, postoji broj e (odnosno funkcija indeksirana tim brojem), za koji važi da $\varphi_e(x_1, ..., x_n) = \psi(e, x_1, ..., x_n)$. Oba rezultata, a osim njih i PRVA TEOREMA REKURZIJE, imaju oblik *teoreme o fiksnoj tački*: njima se tvrdi postojanje fiksne tačke funkcija ili funkcionala, kao i predikata (tj. formula sa slobodnim promenljivima). Značaj fiksnih tačaka za teoriju izračunljivosti pokazuje se i u drugim formalizacijama izračunljivih funkcija, kao na primer u *kombinatornoj logici* i u *λ-računu*. Posebno važnu ulogu u tim formalizacijama ima takozvani *kombinator fiksne tačke* i njemu odgovarajući λ-term koji, kada se primeni na neku funkciju, daje njenu fiksnu tačku. On time omogućava implementaciju rekurzije pomoću koje je ta funkcija definisana u teoriji rekurzija.

Zajednička struktura u osnovi navedenih rezultata o fiksnoj tački biće analizirana u prvom delu ovog poglavlja. Tu će biti predstavljena jedna opšta teorema koja će ukazati na osnovna svojstva teorija koja omogućavaju dokaz odgovarajućeg oblika teoreme o fiksnoj tački. Takođe će biti razjašnjena uloga dijagonalizacije u njenom dokazu, pa samim tim i u dokazima različitih teorema te vrste kojima smo se dosad bavili.

Osim u dolaženju do rezultata o fiksnoj tački, dijagonalizacija takođe ima ulogu u dolaženju do različitih negativnih rezultata. Ona se tako koristi za konstruisanje nerešivih problema i neodlučivih formula, ili za dokazivanje da su oni takvi. Razmatranje rekurzivno-teorijskih verzija teoreme o nepotpunosti pokazalo je da su te dve vrste rezultata veoma blisko povezane. U ovom poglavlju bliže ćemo razmotriti njihove zajedničke karakteristike za koje će se ispostaviti da ih dele i sa paradoksima pomenutim na početku rada. Takođe ćemo pokušati da utvrdimo vezu između uloge dijagonalizacije u dokazivanju teorema o fiksnoj tački i njene uloge u dolaženju do negativnih rezultata teorije izračunljivosti i metamatematike.

2.3.1 *Smalijanova teorema o fiksnoj tački*

U ovom odeljku biće predstavljena jedna opšta teorema o fiksnoj tački koja pokazuje da se do fiksnih tačaka čije je postojanje dokazano DRUGOM TEOREMOM REKURZIJE i LEMOM O DIJAGONALIZACIJI dolazi pomoću iste konstrukcije. Ta konstrukcija opisana je u dokazu Smalijanove teoreme (Smullyan, 1984). Teorema tvrdi postojanje fiksne tačke funkcije definisane na proizvoljnom skupu objekata N, pod određenim uslovima koje klasa funkcija Σ kojoj ona pripada, treba da zadovolji. Kako bi ta teorema bila primenljiva i u oblastima koje se bave objektima različitim od funkcija, pojam funkcionalne fiksne tačke donekle je uopšten. Naime, fiksna tačka unarne funkcije na nekom skupu definisana je u odnosu na proizvoljnu relaciju ekvivalencije na tom skupu $\equiv$ (umesto u odnosu na važan primer takve relacije - identitet). Pod fiksnom tačkom funkcije f podrazumeva se njen argument n za koji važi $f(n) \equiv n$ (teorema može biti uopštena tako da se odnosi i na funkcije sa više argumenata). Smalijanova OPŠTA TEOREMA O FIKSNOJ TAČKI tada tvrdi da svaka funkcija iz Σ ima fiksnu tačku u skupu N pod uslovom da postoji binarna operacija F iz Σ takva da:

1. za svaku unarnu funkciju f iz Σ, postoji neko a iz N takvo da za svako x, $F(a, x) \equiv f(x)$;
2. za svaku unarnu funkciju f iz Σ, postoji funkcija f' iz Σ takva da za svako x, $f'(x) \equiv f(F(x, x))$.

Na osnovu prvog uslova možemo zaključiti da funkcija F ima ulogu sličnu ulozi univerzalne funkcije u odnosu na klasu funkcija Σ: rezultat njene primene na kôd neke funkcije iz te klase i neki argument, biće ekvivalentan (na osnovu pretpostavljene relacije ekvivalencije) rezultatu primene te funkcije na dati argument. Iz toga sledi da $F(x, x)$ stoji u relaciji ekvivalencije sa rezultatom primene funkcije kodirane brojem x na isto to x, odnosno, sa dijagonalizacijom te funkcije. Prema tome, drugi uslov zapravo zahteva da za svaku funkciju f iz Σ postoji u istoj klasi funkcija njen *dijagonalizator* - funkcija koja za svaki argument x daje vrednost ekvivalentnu onoj koju f daje za njegovu dijagonalizaciju $F(x,x)$.

Dokaz ove teoreme sadrži konstrukciju fiksne tačke za proizvoljnu funkciju iz Σ. Kako bismo došli do fiksne tačke takve funkcije f, treba najpre da pronađemo u Σ funkciju f' za koju važi da za svako x, $f'(x) \equiv f(F(x,x))$. Na osnovu drugog od navedenih uslova znamo da takva funkcija postoji. Na osnovu prvog uslova, sa druge strane, znamo da će za isto to f' postojati neko a takvo da za svako x, $F(a,x) \equiv f'(x)$. Iz ove dve ekvivalencije na osnovu tranzitivnosti sledi da za svako x, $F(a,x) \equiv f(F(x,x))$. Ako supstituišemo a mesto x u poslednjoj ekvivalenciji, dobićemo da $F(a,a) \equiv f(F(a,a))$, što znači da je $F(a,a)$ fiksna tačka funkcije f. Fiksna tačka proizvoljne funkcije f iz klase Σ dobijena ovom konstrukcijom jeste zapravo dijagonalizacija njenog dijagonalizatora f' (cf. Smullyan, 1984, pp. 287-288).

Istu ideju za konstruisanje fiksne tačke proizvoljne funkcije je izgleda imao i Kari (Haskell Curry) kada je definisao njegov *paradoksalni kombinator* Y, koji predstavlja implementaciju kombinatora fiksne tačke u λ-računu, kao:

$$\lambda f.\left(\lambda x.f\left(xx\right)\right)\left(\lambda x.f\left(xx\right)\right).$$

Kada je primenjen na proizvoljnu funkciju f, ovaj kombinator daje njenu fiksnu tačku:

$$\left(\lambda x.f\left(xx\right)\right)\left(\lambda x.f\left(xx\right)\right).$$

Do ove fiksne tačke takođe se dolazi konstruisanjem dijagonalizatora funkcije f, tj. funkcije $\lambda x.f(xx)$, a zatim njenom dijagonalizacijom, tj. primenom na sebe.

OPŠTA TEOREMA O FIKSNOJ TAČKI može biti interpretirana na različite načine, pa i tako da se DRUGA TEOREMA REKURZIJE i LEMA O DIJAGONALIZACIJI dobiju kao njene instance (Smullyan, 1984, pp. 288-289). Određene interpretacije te opšte teoreme omogućavaju nam da pokažemo postojanje fiksnih tačaka izračunljivih funkcija ili drugih objekata, kao što su formule, predikati, relacije i slično. Uslovi navedeni u datoj teoremi sadrže samo ono što je neophodno kako bi ona, ili neka njena verzija, bile dokazive u teoriji koja se bavi objektima, poput navedenih, koje karakteriše neka funkcija primene. Na primer, u svakom od tih slučajeva, pretpostavljeno je neko kodiranje ili imenovanje datih objekata, kojim se uvode njihova jedinstvena imena. Najčešće se to postiže pomoću prirodnih brojeva, ali mogući su i drugi načini (kao što je stavljanje pod navodnike) sve dok su njima dobijena takva imena na koja su imenovani objekti primenljivi.

Prvi uslov naveden u teoremi zapravo znači da teorija o kojoj je reč mora formalizovati *operaciju primene* koja karakteriše objekte kojima se ona bavi, odnosno, predstaviti je pomoću tih objekata. Ta operacija može biti shvaćena na različite načine u zavisnosti od vrste objekata o kojima je reč. U slučaju funkcija to bi bila njihova primena na argumente kojom bi trebalo da započne neko izračunavanje; u slučaju formula ta operacija odgovarala bi operaciji supstitucije kojom se mesto slobodne promenljive u formuli supstituiše ime nekog objekta; u slučaju predikata to bi bila njegova primena na ime kojom nastaje rečenica, itd. Takva formalizacija operacije primene postignuta je u teoriji rekurzija pomoću univerzalne funkcije Φ_n koja je takođe rekurzivna. Sa druge strane, u odgovarajućim formalnim teorijama, operacija supstitucije terama mesto slobodnih promenljivih u nekoj formuli predstavljiva je pomoću formule $\mathbf{Sub}(\mathbf{x},\mathbf{y},\mathbf{z})$. U svakom od navedenih slučajeva, pretpostavljena formalizacija primene dopušta i primenu objekta o kojem je reč na sebe, odnosno na sopstveni kôd. Tu primenu zvaćemo *dijagonalizacija u širem smislu*. Koji će tačno oblik ona imati zavisi od pretpostavljenog pojma primene (za metode različite od dijagonalizacije koji se mogu koristiti u istu svrhu, kao što je *normalizacija*, vidi: Smullyan, 1957).

Drugi uslov naveden u teoremi zahteva postojanje *dijagonalizatora* objekata koji treba da imaju fiksnu tačku. Pojam dijagonalizatora razlikuje se u odnosu na vrstu objekta i njegove primene. Dijagonalizator funkcije je tako neka druga funkcija koja se primenjuje na kôdove tačno onih funkcija na čiju se dijagonalizaciju prva funkcija primenjuje, i daje istu vrednost. Postojanje dijagonalizatora neke rekurzivne funkcije posledica je mogućnosti primene te funkcije na kôd dijagonalizacije proizvoljne $(n+1)$-arne funkcije koji je dobijen pomoću funkcije S_n^1. Sa druge strane, dijagonalizator neke formule biće formula koja tvrdi da svojstvo izraženo prvom formulom važi za dijagonalizacije formula. Zahvaljujući predstavljivosti operacije supstitucije u dovoljno jakoj teoriji, u njoj ćemo uvek moći da definišemo takvu formulu. Dakle, u opštem slučaju, dijagonalizator nekog objekta formalizuje njegovu primenu na dijagonalizacije drugih objekata. Mogućnost njegovog definisanja ključna je za dokaz OPŠTE TEOREME O FIKSNOJ TAČKI, pošto je fiksna tačka objekata do koje se dolazi u tom dokazu, kao što smo videli, dijagonalizacija njihovog dijagonalizatora (vidi takođe: Smullyan, 1994, pp. 16-17).

Ova teorema tako pokazuje da je neka vrsta dijagonalizacije u vidu primene objekta na sopstveni kôd i neka vrsta kompozicije tih primena dovoljna kako bi bilo dokazano da ti objekti (funkcije, formule, predikati i slični) imaju fiksnu tačku. Takvo okruženje koje zadovoljava navedene uslove formira jedan *aplikativni sistem*, čije su opšte osobine predmet, pre svega, kombinatorne logike (vidi: Smullyan, 1994, p. 26).

2.3.1.1 *Relaciona verzija Smalijanove teoreme*

Smalijan je dokazao još jednu verziju OPŠTE TEOREME O FIKSNOJ TAČKI koja se ovog puta tiče binarnih relacija umesto funkcija i pokazuje uslove pod kojima će svaka binarna relacija iz određene klase relacija imati fiksnu tačku (Smullyan, 1994). Ta teorema je uopštenje teoreme koja dokazuje postojanje fiksnih tačaka funkcija na istom skupu, pošto se te funkcije mogu shvatiti i kao posebna vrsta binarnih relacija. Fiksna tačka binarne relacije R na nekom skupu je element tog skupa i za koji $R(i,i)$ važi. Uzmimo da je Σ klasa svih binarnih relacija na nekom skupu N. Relacija R iz te klase imaće fiksnu tačku pod uslovom da u istoj klasi postoje relacija R' i funkcija d za koje važi sledeće:

1. za neki element a skupa N, važi $R'(d(a), a)$;
2. za sve elemente x i y iz N, ako $R'(x,y)$ onda $R(x, d(y))$.

Dokaz teoreme je izuzetno jednostavan. Da bismo pronašli fiksnu tačku proizvoljne relacije R iz klase Σ, treba da nađemo takav element skupa N, nazovimo ga a, za koji $R'(d(a), a)$ važi. Na osnovu prvog od navedenih uslova znamo da takav element postoji. Tada iz drugog uslova sledi da $R(d(a), d(a))$ takođe važi, što znači da je $d(a)$ fiksna tačka relacije R.

Ako funkcija d predstavlja uopštenje dijagonalizacije, onda je uloga relacije R' u odnosu na relaciju R slična onoj koju dijagonalizator funkcije iz prethodne verzije teoreme ima u odnosu na tu funkciju. Naime, neko x stajaće u relaciji R' prema nekom objektu y samo ako isto to x stoji u relaciji R prema njegovoj dijagonalizaciji.

Ova teorema izostavlja sve detalje o prirodi i osobinama binarnih relacija o kojima je reč, kao i objekata koji stoje u tim relacijama, a koji nisu relevantni za njen dokaz. Time dolaze do izražaja uslovi koji se tiču odnosa između relacija i koji bi trebalo da garantuju postojanje njihove fiksne tačke. Ova teorema tako predlaže neke jasne i opšte kriterijume na osnovu kojih se može utvrditi postojanje fiksne tačke neke relacije ili funkcije. Poznavanje tih kriterijuma moglo bi u nekim slučajevima biti od velikog značaja, posebno ako u vidu imamo ulogu koju fiksne tačke relacija ili funkcija imaju u dokazu različitih rezultata koji se tiču njih ili teorije koja se njima bavi. Tako, na primer, odgovarajuće interpretacije date teoreme vode rezultatima o fiksnoj tački rekurzivnih funkcija i formula kojima smo se prethodno bavili (Smullyan, 1994, pp. 25-38).

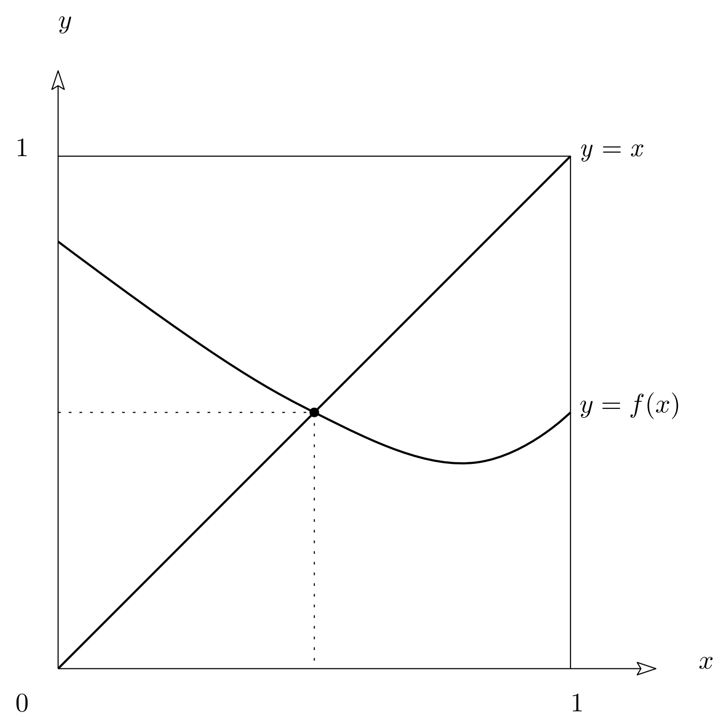
Činjenica da se glavne teoreme teorije izračunljivosti i metamatematike, kojima smo se bavili u prethodnom delu rada, mogu dobiti kao instance OPŠTE TEOREME O FIKSNOJ TAČKI, predstavlja formalno opravdanje našeg tvrđenja da te teoreme imaju ulogu teorema o fiksnoj tački u tim oblastima. Sveopšta prisutnost i široka upotreba teorema o fiksnoj tački nije tipična za te teorije, već karakteriše i mnoge druge matematičke oblasti i discipline. Kako bismo opravdali to tvrđenje, u sledećem odeljku ćemo predstaviti neke od najznačajnijih teorema o fiksnoj tački koje pripadaju teorijama različitim od onih kojima smo se dosad bavili, zajedno sa njihovim glavnim posledicama.

2.3.2 *Fiksne tačke u drugim oblastima matematike*

Fiksne tačke javljaju se svuda u matematici, često sa vrlo značajnim posledicama. Raznolikost rezultata koji dokazuju njihovo postojanje i imaju veliki matematički značaj ilustrovaćemo pomoću nekih od najpoznatijih primera iz oblasti različitih od onih kojima smo se prethodno bavili. Pošto svi rezultati koji tvrde postojanje fiksnih tačaka sadrže neku vrstu samoreferencije, to će pružiti dodatno svedočanstvo o sveprisutnosti i značaju samoreferencijalnih metoda širom matematike.

Izučavanje fiksnih tačaka prvenstveno pripada topologiji, oblasti matematike koja se bavi svojstvima geometrijskih objekata koja ostaju nepromenjena nakon određenih vrsta deformacija tih objekata - onih koje, slikovito govoreći, ne uključuju „kidanje" ili „lepljenje". Preciznije, deformacije geometrijskih objekata kojima se topologija bavi su one u čijoj osnovi je neka neprekidna funkcija. U vrednostima neprekidne funkcije ne postoji nikakva nagla promena ukoliko takve promene nema u njenim argumentima. Drugim rečima, dovoljno mala promena u argumentu te funkcije rezultira proizvoljno malom promenom u njenoj vrednosti. Glavni predmet istraživanja u ovoj oblasti koje se tiče fiksnih tačaka jesu svojstva topološkog prostora koja omogućavaju da neprekidna deformacija geometrijskih objekata u tom prostoru ostavi neke od njegovih tačaka netaknutim. Najpoznatija i najšire primenljiva teorema o fiksnoj tački koja pripada ovoj oblasti jeste BRAUEROVA TEOREMA O FIKSNOJ TAČKI (L. E. J. Brouwer). Ta teorema tvrdi da svaka neprekidna funkcija koja preslikava neki zatvoren, ograničen i konveksan podskup n-dimenzionalnog euklidskog prostora u sebe, ima fiksnu tačku.[7] Na primer, ta teorema se može primeniti na svaku neprekidnu funkciju koja preslikava neki zatvoreni interval u njega samog iz čega sledi da svaka takva funkcija ima fiksnu tačku. Dokaz ovog jednodimenzionalnog slučaja je prilično jednostavan. On se oslanja na činjenicu da graf takve funkcije seče graf identičke ili dijagonalne funkcije na istom intervalu. Graf dijagonalne funkcije sastoji se od tačaka koordinatne ravni oblika (x, x), dok je graf funkcije f napravljen od tačaka $(x, f(x))$. Pošto graf funkcije f seče graf dijagonalne funkcije, mora postojati neki broj x_0, za koji važi da $(x_0, x_0) = (x_0, f(x_0))$. Taj broj predstavlja fiksnu tačku funkcije f. Teorema takođe implicira postojanje fiksnih tačaka neprekidnih funkcija koje preslikavaju disk, kvadrat i druge geometrijske objekte u njih same. Činjenica koja je posebno interesantna za topologiju jeste da svaka topološka transformacija čuva svojstvo podskupa euklidskog prostora koje garantuje da sve neprekidne funkcije na njemu imaju fiksnu tačku (za primene ove teoreme vidi na primer: Casti, 1996, pp. 45-84).

7 Podskup euklidskog prostora je *zatvoren* ako sadrži sve granične tačke; *ograničen* ako je sadržan u kugli; i *konveksan* ako mu pripada svaka tačka koja leži između neke dve njegove tačke.

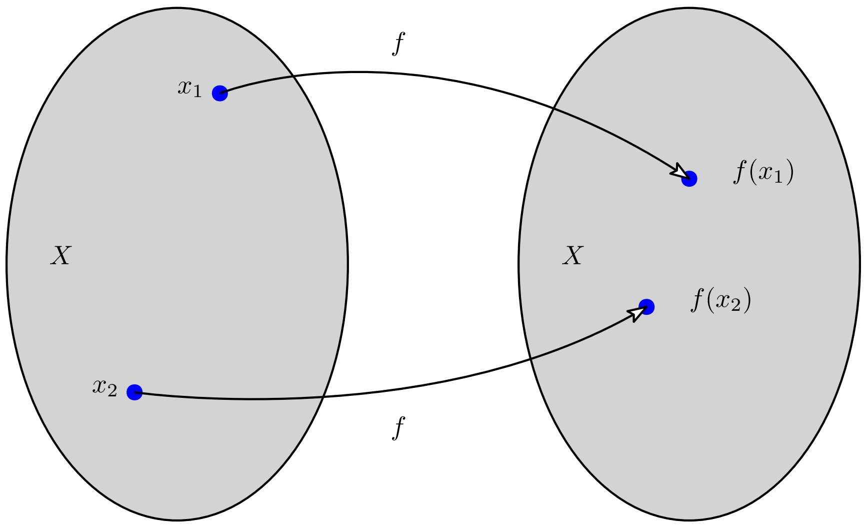


Slika 2.1: Ilustracija BRAUEROVE TEOREME O FIKSNOJ TAČKI na intervalu $[0, 1]$.

Blisko povezan sa Brauerovom teoremom je i BANAHOV PRINCIP KONTRAKCIONOG PRESLIKAVANJA (BANACH'S CONTRACTION MAPPING PRINCIPLE). Taj princip predstavlja verziju teoreme o fiksnoj tački koja pripada teoriji metričkih prostora[8]. Banahov princip tiče se funkcija koje su *striktne kontrakcije* na nekom metričkom prostoru, što znači da je za svake dve tačke tog prostora x_1 i x_2, rastojanje između tačaka $f(x_1)$ i $f(x_2)$ strogo manje od rastojanja između tačaka x_1 i x_2. Teorema tvrdi da svaka striktna kontrakcija na kompletnom metričkom prostoru ima jedinstvenu fiksnu tačku[9]. Dokaz teoreme sadrži konstruktivni metod za pronalaženje te fiksne tačke. On pokazuje da svaki niz tačaka dobijen iterativnom primenom striktne kontrakcije, počevši od neke proizvoljne tačke kompletnog metričkog prostora, konvergira njenoj fiksnoj tački (za dokaz teoreme, vidi: Shapiro, 2016, pp. 28-29).

8 Metrički prostor čini neki skup zajedno sa funkcijom koja svakom paru elemenata tog skupa pripisuje njihovo međusobno rastojanje.

9 Intuitivno, metrički prostor je kompletan ako sadrži sve tačke tog prostora.

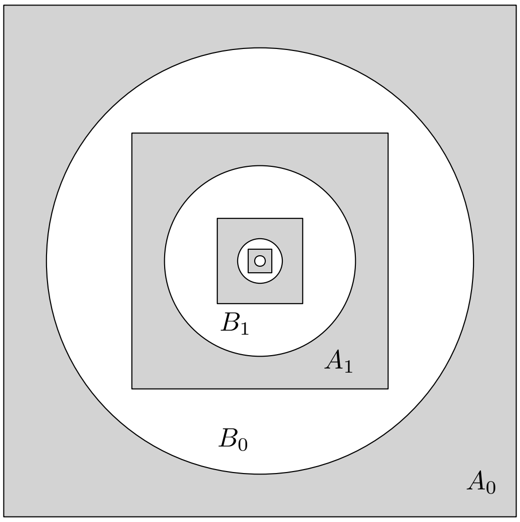


Slika 2.2: Primer kontrakcionog preslikavanja $f : X \to X$ kojim se smanjuje rastojanje između dve tačke metričkog prostora X.

Od teorema o fiksnoj tački van oblasti kojima smo se prvenstveno bavili u radu, pomenućemo još samo jednu. To je TEOREMA KNASTERA I TARSKOG (KNASTER-TARSKI THEOREM) koja, u jednoj od formulacija, tvrdi da svaka monotona funkcija na nekoj kolekciji skupova ima fiksnu tačku. Neka funkcija f na kolekciji skupova parcijalno uređenih relacijom podskupa biće monotona ako i samo ako za svaki par skupova iz te kolekcije A i B važi da ako $A \subseteq B$, onda $f(A) \subseteq f(B)$. Neka je X proizvoljan skup, a $\mathcal{P}(X)$ njegov partitivni skup. Tada na osnovu date teoreme, monotona funkcija f na skupu $\mathcal{P}(X)$ ima fiksnu tačku. Prema dokazu teoreme, do fiksne tačke funkcije f možemo doći konstruisanjem skupa ε svih podskupova E skupa X za koje važi da je $E \subseteq f(E)$. Unija svih elemenata skupa ε biće tada fiksna tačka funkcije f (cf. Shapiro, 2016, p. 9). Primetimo da je dokaz ove teoreme veoma sličan dokazu PRVE TEOREME REKURZIJE. U oba slučaja, fiksna tačka monotone funkcije ili funkcionala može da se nađe kao najmanja gornja granica rastućeg niza skupova ili funkcija, koji predstavljaju njihove argumente. Tako dobijen skup ili funkcija ne može se dalje proširiti primenom te funkcije ili funkcionala, zbog čega će predstavljati njihovu fiksnu tačku.

Ova teorema o fiksnoj tački ima neke važne posledice, među kojima je i BANAHOVA TEOREMA O PRESLIKAVANJU (BANACH'S MAPPING THEOREM) koja tvrdi postojanje fiksne tačke određene monotone funkcije na partitivnom skupu nekog skupa. Konkretno, ta teorema tvrdi da ako je f funkcija iz skupa X u skup Y, a g funkcija iz skupa Y u skup X, onda postoji podskup A skupa X za koji važi $A = X \setminus g(Y \setminus f(A))$. Postojanje ove fiksne tačke može se koristiti u dokazu poznate KANTOR-ŠREDER-BERNŠTAJNOVE TEOREME (CANTOR-SCHRÖDER-BERNSTEIN THEOREM) koja tvrdi da ako postoji injekcija iz skupa X u skup Y i injekcija iz skupa Y u skup X, onda postoji bijekcija između ta dva skupa.

Konstrukcija slična onoj koja omogućava dokaz BANAHOVE TEOREME O PRESLIKAVANJU koristi se i u dokazu jedne jednostavne leme koja za proizvoljne skupove A_1, B i A tvrdi da ako važi $A_1 \subseteq B \subseteq A$ i postoji bijekcija f između skupova A_1 i A, onda postoji i bijekcija između skupova B i A. Do te bijekcije možemo doći konstruisanjem niza podskupova skupa $A = A_0$ i skupa $B = B_0$, dobijenih iterativnom primenom funkcije f počevši od tih skupova, na skup koji se dobija kao slika njene prethodne primene (A_{n+1} je tako jednako $f[A_n]$, a B_{n+1} je jednako $f[B_n]$). Tada će bijekcija između skupova A i B biti konstruisana kao funkcija koja za element x skupa A koji je u B daje kao vrednost to isto x, a za x koje nije u B daje kao vrednost $f(x)$. KANTOR-ŠREDER-BERNŠTAJNOVA TEOREMA se takođe može dobiti kao jednostavna posledica upravo pomenutog rezultata (vidi: Hrbacek & Jech, 1999, pp. 66-67).



Slika 2.3: Ilustracija leme korišćene u dokazu KANTOR-ŠREDER-BERNŠTAJNOVE TEOREME (iz Hrbacek & Jech, 1999, p. 67).

U najopštijoj verziji, TEOREMA KNASTERA I TARSKOG tvrdi da skup fiksnih tačaka monotone funkcije na kompletnoj mreži (koju čini parcijalno uređeni skup čiji svi podskupovi imaju najmanju gornju granicu i najveću donju granicu) i sam predstavlja kompletnu mrežu. Iz toga sledi da je taj skup neprazan, što znači da svaka takva funkcija ima fiksnu tačku, ali takođe i da za svaku od tih funkcija postoji njena najmanja i najveća fiksna tačka.

Teoreme koje dokazuju postojanje fiksnih tačaka određenih funkcija mogu takođe imati značajne posledice koje se tiču postojanja rešenja različitih vrsta jednačina ili sistema jednačina. Na primer, nalaženje fiksne tačke funkcije f ekvivalentno je nalaženju rešenja jednačine $f(x) - x = 0$. Sa druge strane, nalaženje rešenja jednačine $g(x) = 0$ isto je što i nalaženje fiksne tačke funkcije h za koju važi $g(x) = x - h(x)$(Casti, 1996, p. 63). Posledica teorema o fiksnoj tački, da definicije objekata na koje se one odnose mogu sadržati referenciju na njih same, takođe ima upotrebu širom matematike, logike i teorijskog računarstva. To je posebno slučaj sa jednom vrstom takvih definicija - definicija rekurzijom. Te definicije koriste se u programiranju, za definisanje osnovnih aritmetičkih operacija, kao što su sabiranje, množenje i stepenovanje, za definisanje određenih skupova kao što je skup prirodnih brojeva, skup formula nekog formalnog sistema ili njegovih teorema, itd.

Ali pored tih posledica koje omogućavaju razvoj ili čak i samo zasnivanje matematičkih i logičkih teorija, teoreme o fiksnoj tački mogu imati i neke negativne posledice po formalne teorije u kojima su dokazane. S tim u vezi, već smo videli da neki od najvažnijih rezultata koji pokazuju ograničenja određenih logičkih i matematičkih teorija proizilaze upravo iz tih teorema, odnosno iz činjenice da iz njih sledi postojanje fiksnih tačaka nekih funkcija ili formula, koje pod određenim uslovima vodi protivrečnosti. Slične fiksne tačke u osnovi su većine paradoksa koje smo pomenuli u radu. U sledećem odeljku sumiraćemo te negativne posledice samoreferencije na koje smo do sada naišli i pokušati da odredimo zajednički oblik teorema o fiksnoj tački ili njihovih instanci koje vode takvim posledicama.

2.3.3 *Negativne posledice teorema o fiksnoj tački*

Tvrđenje da svaka rekurzivna funkcija ili svaka formula aritmetičkog formalnog sistema ima fiksnu tačku čini se problematičnim samo po sebi. Razlog je to što postoje mnoge rekurzivne funkcije za koje, prema njihovoj definiciji, treba da važi da one menjaju svoje argumente na način koji ne dozvoljava da ijedan od tih argumenata ostane isti i nakon njihove primene. Slično tome, neka formula može izražavati svojstvo za koje se čini da ne može na konzistentan način biti pripisano samoj toj formuli. Teoreme koje tvrde postojanje fiksnih tačaka takvih funkcija i formula zaista imaju naizgled paradoksalne posledice. Tako jedna instanca PRVE TEOREME REKURZIJE tvrdi da postoji rekurzivna funkcija φ za koju važi da $S(\varphi(x)) = \varphi(x)$. Jednako iznenađujuća posledica LEME O DIJAGONALIZACIJI je da pod određenim uslovima postoji formula koja je ekvivalentna svojoj negaciji (na primer, pod uslovom da je istina predstavljiva u nekoj formalnoj teoriji, možemo zahvaljujući toj lemi da pokažemo da postoji rečenica $\mathbf{L}$ za koju će važiti $\mathbf{T}([\mathbf{L}]) \leftrightarrow \neg\mathbf{T}([\mathbf{L}])$). Takve posledice teorema o fiksnoj tački nisu neuobičajene. Ali umesto da proizvedu sumnju u pogledu valjanosti metoda koje su do njih dovele ili same teorije koja se tim metodama služi, te posledice su shvaćene kao pokazatelj da je neophodno uvesti neka ograničenja koja će sprečiti izvođenje kontradikcije iz njih. Naizgled paradoksalni rezultati tako nisu shvaćeni kao neprihvatljivi sve dok ne predstavljaju pretnju po konzistentnost teorije u kojoj se javljaju. Rezultati analogni paradoksima javljaju se u matematici i na druge načine osim kao posledice teorema o fiksnoj tački. Na primer, jedan važan matematički rezultat, BANAH-TARSKI PARADOKS (BANACH-TARSKI PARADOX), sadrži konstrukciju koja, iako nije protivrečna, jeste u suštini paradoksalna. Instanca te teoreme tvrdi da skup tačaka neke lopte u trodimenzionalnom prostoru može biti podeljen u konačan broj podskupova od kojih se mogu sastaviti dve lopte koje su obe po veličini jednake originalnoj lopti. Korisnost tog i sličnih rezultata pokazuje da približavanje paradoksima može u matematici biti nagrađeno neočekivanim i rasvetljujućim otkrićima. O tome svedoči i činjenica da se za najjaču aksiomu velikih kardinala u teoriji skupova uzima sledeća aksioma: $0 = 1$. Potraga za što jačom teorijom skupova predstavlja tako potragu za aksiomama koje se što više približavaju toj aksiomi ali u isto vreme ne čine teoriju protivrečnom.

Da bi bilo moguće približiti se paradoksima u nekoj teoriji, neophodno je najpre ustanoviti određena ograničenja koja će sprečiti da se kao posledica u toj teoriji javi protivrečnost. Ograničavajući rezultati koji sprečavaju izvođenje kontradikcije iz određenih instanci teorema o fiksnoj tački kojima smo se prethodno bavili, tiču se mogućnosti potpune formalizacije neke teorije ili mogućnosti efektivnog nalaženja odgovora na određena pitanja. Tarski koristi fiksnu tačku predikata neistinitosti kako bi pokazao da istinitost rečenica konzistentne i valjane aritmetičke teorije ne može biti predstavljiva u istoj teoriji, što za rezultat ima ograničavanje izražajne moći te teorije. Gedelovi rezultati, sa druge strane, pokazuju da je deduktivna moć još veće klase formalnih teorija ograničena. Kao što smo videli, ti rezultati mogu se shvatiti kao verzije negativnih rezultata teorije izračunljivosti koji se tiču rekurzivne nerešivosti određenih problema. To da negativni rezultati teorije izračunljivosti mogu imati posledice po metamatematička pitanja pokazuje se na još jednom primeru. Taj primer tiče se još jedne vrste ograničenja formalnih teorija koja se sastoji u nepostojanju efektivne metode teorije za proveru valjanosti njenih formula. Za takve teorije kaže se da su *neodlučive*. Čerč je pokazao da je logika prvog reda neodlučiva služeći se rezultatom koji pokazuje da je problem koji se tiče pitanja da li su dva izraza u λ-računu ekvivalentna rekurzivno nerešiv. Nezavisno od njega, Tjuring je pokazao da isto ograničenje logike prvog reda sledi iz nerešivosti halting problema (Fraenkel et al., 1973, pp. 310-320).

Ne samo da su negativni rezultati teorije izračunljivosti i metamatematike međusobno blisko povezani i zavisni jedni od drugih, rezultati obe oblasti su takođe veoma slični paradoksima i često inspirisani njima. U osnovi rezultata koji nameću navedena ograničenja nalaze se iste konstrukcije koje se javljaju u osnovi paradoksa koje smo pominjali u radu. One se zapravo sastoje u *konstruisanju fiksnih tačaka takvih funkcija, pojmova ili predikata, koji intuitivno nemaju fiksnu tačk*u. Te funkcije, pojmovi i predikati najčešće stoje za neko negativno svojstvo (neistinitost, nedokazivost, nedefinljivost, i slično) ili za samu negaciju. Fiksne tačke formule $\mathbf{P}(\mathbf{x}) \to \mathbf{B}$, gde je $\mathbf{B}$ proizvoljna formula, koje se javljaju u KARIJEVOM PARADOKSU i LEBOVOJ TEOREMI, mogu biti shvaćene kao uopštenje fiksne tačke negativnog svojstva $\neg\mathbf{P}(\mathbf{x})$, koja se dobija supstitucijom formule $\perp$ mesto $\mathbf{B}$ (pošto je $\neg\mathbf{P}(\mathbf{x})$ ekvivalentno sa $\mathbf{P}(\mathbf{x}) \to \perp$). Do takvih fiksnih tačaka dolazi se najčešće primenom dijagonalnog argumenta koji je takođe u osnovi Raselovog, Hilbertovog, Cermelovog, Rišarovog, i mnogih drugih paradoksa, kao i negativnih rezultata različitih teorija. Svi načini upotrebe dijagonalnog argumenta koje smo pomenuli do sada imaju dakle iste posledice: oni eksplicitno ili implicitno pokazuju da pod određenim uslovima postoji fiksna tačka neke funkcije ili formule. Postojanje fiksne tačke takvih funkcija i formula koje odgovaraju pomenutim negativnim svojstvima, u nekim slučajevima vodi kontradikciji. To omogućava formulaciju paradoksa ili njihovih formalnih ekvivalenata. Fiksne tačke o kojima je reč su najčešće konstruisane na način opisan u OPŠTOJ TEOREMI O FIKSNOJ TAČKI. Konstrukcija sadržana u toj teoremi se tako može shvatiti kao formalizacija i uopštavanje dijagonalnog argumenta koji je u osnovi kako paradoksa tako i mnogih rezultata teorije izračunljivosti i metamatematike. U tom smislu može se reći da teoreme o fiksnoj tački predstavljaju „matematički pismen način formulisanja većine paradoksalnih argumenata" (Cantini, 2009, p. 112).

Poređenje paradoksa sa tvrđenjima koja vode ograničavajućim rezultatima logičkih i matematičkih teorija treba da sugeriše da se paradoksi koji se mogu formulisati unutar neke teorije ne moraju nužno shvatiti kao pokazatelj da je ta teorija osuđena na propast ili da njen dalji razvoj zahteva temeljnu reviziju njenih osnovnih pojmova. Umesto toga oni mogu biti korišćeni za dolaženje do određenih negativnih rezultata u toj teoriji na isti način na koji su formalni ekvivalenti paradoksa korišćeni za definisanje nerešivih problema i neodlučivih formula. Na taj način nešto korisno može biti izvedeno iz paradoksa (cf. Thomson, 1962, pp. 118-119). Ta lekcija kojoj nas uči istorija logike i matematike, može motivisati i usmeriti novi pristup paradoksima koji se tiču pojmova i njihovih ekstenzija. Sve primene samoreferencije predstavljene u ovom radu svedoče o njenoj snazi i metodološkoj vrednosti koja se ispoljava u teorijama u kojima se samoreferencija koristi na konzistentan način. To bi trebalo da pokaže da rešenje pomenutih paradoksa ne treba tražiti u izbegavanju pojmovne samoreferencije u budućoj teoriji pojmova, već u nekom načinu na koji se paradoksi mogu transformisati u rezultate te teorije koji mogu imati ulogu u rasvetljavanju prirode pojmova i njihovih formalnih svojstava. Sledeći deo ovog rada posvećen je upravo potrazi za takvim rešenjem intenzionalnih paradoksa. Ona će biti usmerena Gedelovim napomenama, kao i lekcijama naučenim prethodnim razmatranjem matematičkog pristupa paradoksima i samoreferenciji uopšte. Jedan od razloga zbog kog naše prethodno razmatranje može biti relevantno za razumevanje pojmovne samoreferencije i intenzionalnih paradoksa jeste činjenica da su različite vrste samoreferencije kojima smo se bavili omogućene intenzionalnim razumevanjem objekata u odnosu na koje se one javljaju. Ako je to slučaj, onda bismo u samoreferenciji koja se tiče pojmova koji sadrže to intenzionalno značenje mogli naći i opravdanje tih različitih vrsta samoreferencije i načina na koji se one koriste.

3

SAMOREFERENCIJA I INTENZIONALNOST

Jedan od ciljeva prethodnog razmatranja različitih vrsta samoreferencije u logici i matematici jeste istraživanje veze između *intenzionalnog razumevanja* objekata o kojima je reč - skupova, funkcija i formula, i prepoznavanja mogućnosti njihove samoreferencije. Pod intenzionalnim razumevanjem nekog objekta mislimo na uzimanje u obzir njegovih intenzionalnih aspekata sadržanih, na primer, u pravilu po kojem je taj objekat formiran, u njegovoj strukturi ili strukturi koju on pravi sa drugim objektima iste vrste. Instance funkcionalne i lingvističke samoreferencije kojima smo se bavili u ovom radu mogu se klasifikovati u instance njihove samoprimene, prisutne u metodi dijagonalizacije, i njihove samokarakterizacije, opravdane različitim teoremama o fiksnoj tački. U ovom delu rada pokušaćemo da razjasnimo ulogu intenzionalnog značenja objekata - funkcija i formula, u motivisanju obe vrste samoreferencije. Time ćemo se približiti našoj glavnoj temi, a to je samoprimena i samokarakterizacija pojmova u kojima je to intenzionalno značenje sadržano.

Uprkos ekstenzionalnosti logike čiji je ona deo, teorija izračunljivosti ne usvaja u potpunosti ekstenzionalno stanovište u pogledu funkcija kojima se bavi. Nasuprot tome, svojstva funkcija koja se u toj oblasti izučavaju suštinski zavise od odgovarajućeg algoritma pomoću kog su one definisane u toj teoriji. Sama priroda istraživanja u ovoj oblasti podrazumeva zauzimanje takve perspektive u odnosu na funkcije, pošto se ono tiče njihove izračunljivosti, pa samim tim i pravila po kojem se to izračunavanje izvodi. Pošto su ta pravila izražena u jeziku u obliku definicija funkcija, intenzionalni pristup se u ovoj oblasti najbolje vidi u interesu za te definicije i njihove kôdove. Slično tome, intenzionalni pristup u metamatematici, u kojoj je lingvistička samoreferencija najpre dobila značaj, pokazuje se u tome što se istraživanja u toj oblasti tiču samih formula i njihovih kôdova, a ne samo njihove skupovno-teorijske interpretacije.

Upravo je taj intenzionalni pristup ono što omogućava formulaciju i korišćenje samoreferencije u teoriji izračunljivosti i metamatematici. Ne postoji razlog zbog kog se neka formula ne bi mogla „primeniti” na samu sebe tako da formira rečenicu koja referira na tu formulu ili neku njenu instancu. To bi posebno trebalo da bude slučaj ako ta formula izražava neko svojstvo koje i ona sama poseduje. Njena samoprimenljivost bi tada odslikavala samoprimenljivost odgovarajućeg pojma, koja bi bila opravdana na isti način. Jednako malo razloga ima za odbacivanje mogućnosti da funkcija, na osnovu odgovarajućeg algoritma, bude definisana i za samu sebe. Ipak, u potpunosti ekstenzionalno, tj. skupovno-teorijsko razumevanje funkcija obeshrabruje svaku ideju o njihovoj samoprimeni. Domen ekstenzionalno shvaćene funkcije ne može sadržati samu tu funkciju jer ona ne može biti formirana pre nego što je taj domen formiran. Slično tome, sve dok rečenice neke formalne teorije razumemo ekstenzionalno, tj. kao imena istinosnih vrednosti, ideja o njihovoj samoreferenciji se ne može ni javiti. Tek ako uzmemo u obzir strukturu neke formule, možemo je proglasiti samoreferencijalnom. Samoprimena funkcija i formula je tako opravdana isključivo njihovim intenzionalnim razumevanjem.

U dokazima određenih teorema o fiksnoj tački, kao i njene opšte verzije, samoprimena objekata kao što su funkcije i formule koristi se za opravdanje jedne druge vrste samoreferencije koja se sastoji u samokarakterizaciji tih objekata, odnosno njihovoj karakterizaciji koja referira na iste te objekte ili totalitet kojem i oni pripadaju. Ta vrsta rezultata je dobrodošla u teoriji koja se bavi intenzionalnim objektima, kao što su strukture, algoritmi i pojmovi. Čini se, naime, da u tom kontekstu mogu postojati objekti koji na neki način *sadrže* same sebe ili totalitet kojem i sami pripadaju. Sa druge strane, u ekstenzionalnom svetu pojam sadržavanja svodi se na sadržavanje elemenata nekog skupa u tom skupu. U tom kontekstu ne može se govoriti o samoreferenciji u smislu *samosadržavanja*, pošto neki skup, kao što smo videli, ne može sadržati kao element samog sebe ili neki drugi skup čiji je on sam element. Razlog za to je način na koji su elementi sadržani u skupu, a koji je sličan načinu na koji su delovi nekog materijalnog objekta sadržani u njemu - kao nezavisno postojeći i unapred formirani objekti. Sledeća Gedelova napomena iz *Max Phil IX* mogla bi da se odnosi baš na tu činjenicu[1]:

> „Kada se *ekstenzionalno* referira na sve moguće konstrukcije [unutar jedne konstrukcije], onda se na njih referira kao na „realizovane" (stvarne), a ne samo moguće [jer one moraju biti jedna po jedna izdvojene i razmotrene, pa zato moraju prethodno biti realizovane], što nasuprot tome, nije slučaj pri *intenzionalnoj* referenciji"[2] (Gödel, 2020a, p. 45).

Ovde pomenuta intenzionalna referencija mogla bi da se odnosi na način na koji neki pojam referira na pojmove, svojstva i relacije koji čine njegovo značenje. Način na koji je složeni pojam izgrađen od drugih pojmova i način na koji on sadrži te pojmove, različit je od načina na koji su elementi nekog skupa sadržani u njemu. Pojmovi ne učestvuju u značenju drugog pojma tako što potpadaju pod taj pojam. Oni na jedan drugi način postaju deo strukture pojma i kao takvi određuju njegovo značenje:

> „Postoje dve vrste ujedinjavanja: 1. čisto spoljašnje formiranje skupa, 2. "strukturalno": stvari + pojmovi = činjenice, više komponovanih pojmova = složeni pojam. U ovom načinu ujedinjavanja nužno je da bar jedna komponenta bude pojam [samo to omogućava formiranje strukture pomoću vezivnog tkiva logičke forme] i da se on u komponovanju javlja "kao pojam" (tj. *predikativno*). Pojam se može javljati takođe čisto kao stvar na mestu subjekta. Jednostavni elementi, sa kojima ovaj proces komponovanja počinje, jesu jednostavne *supstancije*, predikati i relacije"[3] (Gödel, 2021, p. 100).

Da treba razlikovati način građenja nekog pojma od toga kako su materijalni objekti (pa i skupovi) izgrađeni, govori i sledeća Gedelova napomena:

> „Kada se kaže da je pojam crvene ruže struktura koja se sastoji od, na određeni način povezanih, pojmova crvenog i ruže, to znači nešto potpuno različito od toga da je kuća nešto što se sastoji od zidova. Jer ne može se reći da pojmovi grade tu celinu tako što stoje u određenom odnosu (koji je mogao i da ne postoji)"[4] (Gödel, 2021, p. 103).

Ono što omogućava građenje nekog složenog pojma od drugih pojmova jeste njihovo povezivanje u jedinstvenu strukturu (koja je moguća zahvaljujući logičkim pojmovima, kao što su veznici). Pojmovi nisu sadržani u drugom pojmu kao nezavisno postojeći objekti sa unapred određenim,

---

1 U glavnom tekstu dati su moji prevodi Gedelovih napomena iz *Max Phil* svezaka, dok su originalne napomene, odnosno njihovi nemački transkripti, dati u fusnotama koje prate te prevode.

2 „Wenn *ext<ensional>* auf alle möglichen Konstruktionen [in einer Konstruktion] Bezug genommen wird, so wird auf sie „als realisiert" (wirklich), nicht als möglich Bezug genommen [denn sie müssen ja der Reihe nach hergenommen und geprüft, daher vorher realisiert werden], nicht dagegen bei *intens<ionaler>* Bezugnahme."

3 „Es gibt 2 Arten der Zusammenfügung: 1. die rein äusserliche der Mengenbildung, 2. die "strukturelle": Dinge + Begriffe = Sachverhalte, mehrere Begriffe zusammengesetzt = komplizierter Begriff. Bei dieser Art der Zusammenfügung <ist es> wesentlich, dass zumindest ein Bestandteil ein Begriff ist [nur dies gestatt<et> mittels des Kitts der logischen Form die Bildung von Struktur] und dass er in der Zusammensetzung "als Begriff" (d. h. *prädikativ*) auftritt. Ein Begriff kann auch rein dinghaft als Subjekt auftreten. Die einfachen Elemente, mit denen man diesen Prozess der Zusammensetzung beginnt, sind die einfachen *Subst<anzen>*, Prädikate und Relationen."

4 „Wenn man sagt, dass der Begriff rote Rose eine Struktur ist, welche aus den Begriffen rot und Rose in einer bestimmten Verbindung besteht, so bedeutet das etwas ganz Anderes, als dass ein Haus etwas ist, das aus den Mauern besteht. Denn man kann nicht sagen, dass die Begriffe dadurch, dass sie in einer gewissen Beziehung stehen (welche auch nicht bestehen könnte), dieses Ganze bilden."

jedinstvenim, značenjem. Značenje koje oni imaju kao deo nekog pojma, određeno je i drugim pojmova sa kojima se pritom povezuju i strukturom koji na taj način čine. Zbog toga je i način na koji je neki totalitet pojmova sadržan u strukturi ili pojmu različit od ekstenzionalnog koji pretpostavlja referenciju na svaki njegov pojedinačan element. To sugeriše i sledeća Gedelova napomena: „...ne može se govoriti *ekstenzionalno* o svim pojmovima jer bi tako opisana kombinacija bila neizvodljiva *u principu*”[5] (Gödel, 2020a, p. 59). ’Svi’ je u ovom kontekstu pre korišćeno na intenzionalan način, tj. tako da predstavlja „smisao drugog reda koji se odnosi prema uobičajenom smislu na isti način na koji se taj smisao odnosi prema stvarima, naime, tako da su one „obuhvaćene” njime i učinjene pristupačnim mišljenju”[6] (cf. Gödel, 2020a, p. 59). Razlika u načinu na koji neki ekstenzionalni i intenzionalni objekat sadrže određeni totalitet omogućava da intenzionalni objekti, kao što su strukture i pojmovi, ali ne i ekstenzionalni, kao deo tog totaliteta sadrže i same sebe:

> „Iznenađujuća činjenica da u oblasti smisla i strukture nešto može biti sopstveni pravi deo, postaje razumljiva na osnovu drugog slučaja (naizgled manje apsurdnog), koji može biti sadržan kao deo na dva različita načina (kada se isti simbol javlja više puta), što na isti način nije moguće u materijalnom svetu”[7] (Gödel, 2020b, p. 18).

U *Max Phil IX* postoji više Gedelovih napomena o strukturama koje sadrže same sebe, među kojima je sledeća[8]:

> „Postoje čak i konačne strukture koje sadrže prave delove „pod koje” cela struktura i sama potpada [u smislu da je ona posebna realizacija te strukture; svaka stvar bi tako potpadala pod strukturu $\cdot$], što znači da važi $t \supset a$, $a \in t$, pri čemu opisane „intenzije” znače: a uključuje totalitet (*intenzionalno*), naime t, čiji je ona sama deo”[9] (Gödel, 2020a, p. 88).

U *Max Phil X* Gedel daje primer takve konačne strukture: iskazni račun sa jednom aksiomom (Gödel, 2017b, p. 8). Druge primere samosadržavajućih struktura nije teško naći: kolekcija svih parcijalno uređenih skupova i sama čini skup parcijalno uređen relacijom podstrukture; kolekcija svih skupova na kojima je definisana neka asocijativna i komutativna binarna operacija i sama čini takvu strukturu sa Dekartovim proizvodom kao takvom operacijom (modulo relacije bijekcije između skupova), itd. (cf. Feferman, 1984, p.76). Isto tako: „Struktura nizova celih brojeva, na primer, sadrži sebe kao pravi deo i lako je videti da postoje takođe strukture koje sadrže beskonačno mnogo različitih delova, od kojih svaki kao deo sadrži celu strukturu” (Gödel, 1944, p. 130). Mogućnost da i drugi intenzionalni objekti, kao što su pojmovi i algoritmi, sadrže same sebe može se takođe objasniti njihovim posedovanje strukture. Tako, na primer, u kontekstu objašnjenja razloga zbog kog aksioma reducibilnosti ne bi morala da važi za neke pojmove, Gedel daje primer složenih pojmova koji „nisu nešto prosto, već i sami imaju strukturu pa tako mogu

---

5 „Daher kann man nicht *ext<ensional>* von allen Sinnen [in the footnote: D. h. alle Begriffe] sprechen, denn die so beschriebene Kombination ist *princ<ipiell>* undurchführbar.”

6 „...das alle wäre also ein Sinn zweiter Stufe, welcher sich zum gewöhnlichen Sinn ebenso verhält wie dieser zu den Dingen, dass sie nämlich durch ihn „erfasst“, dem Denken zugänglich gemacht werden.”

7 „Die merkwürdige Tatsache, dass im Reich des Sinns und der Struktur etwas sein eigener echter Teil sein kann, wird verständlich durch <die> zweite (anscheinend weniger absurde), dass dasselbe auf 2 verschiedene Weisen als Teil enthalten sein kann (wenn sich das gleiche Zeichen mehrere Male wiederholt), was ebenfalls im Reich der Körper unmöglich ist.”

8 U citatima koji slede, Gedelov simbol ε transkribujemo kao $\in$. Ipak, treba imati u vidu da Gedel možda nije uvek mislio na relaciju pripadnosti skupu, nego na relaciju pripadnosti pojmu, tj. njegovoj ekstenziji.

9 „Es gibt schon endliche Strukturen, die echte Teile enthalten, „unter die“ die ganze Struktur fällt [in dem Sinn, dass sie eine spezielle Realisierung dieser Strukturen ist; jedes Ding würde dann unter die Struktur $\cdot$ fallen], d. h. also $t \supset a$, $a \varepsilon t$, wobei die beschriebenen „Begriffsinhalte“ bezeichnen: a involviert eine Totalität (*intens<ional>*), nämlich t, von der es selbst ein Glied ist.”

„*uključivati*" totalitet pojmova"[10] (Gödel, 2020a, p. 67). Prema Gedelovim rečima, formalni sistem koji bi odgovarao ovom shvatanju pojmova sadržao bi umesto aksiome reducibilnosti pravilo supstitucije za iskazne funkcije koje bi se primenjivalo na promenljive proizvoljnog tipa, kao i nekakve aksiome intenzionalnosti koje bi odgovarale pojmu svojstva (Gödel, 1944, p. 130, fusnota 28).

Mogućnost da određeni objekti sadrže same sebe opravdava njihovo definisanje pomoću impredikativnih definicija koje referiraju na totalitet kojem i sami definisani objekat pripada. Impredikativne definicije su, po Gedelovom mišljenju, neodgovarajuće ukoliko treba da sadrže instrukcije za konstruisanje njima definisanih objekata. Ako ih, nasuprot tome, koristimo samo da identifikujemo ili izdvojimo neki već postojeći objekat, onda „nema ničeg imalo apsurdnog u postojanju totaliteta koji sadrži članove koji mogu biti opisani (tj. okarakterisani na jedinstven način) samo referencijom na isti taj totalitet" (Gödel, 1944., p. 128). Tako bismo, na primer, neku od gore navedenih struktura opisali pomoću referencije na totalitet struktura kojem i ona sama pripada. Pojam, takođe, možemo definisati pomoću nekog totaliteta pojmova čiji je i on sam deo. Na osnovu toga i navedenih Gedelovih napomena može se zaključiti da je Gedel smatrao samokarakterizaciju intenzionalnih objekata, kao što su pojmovi, neproblematičnom, a u nekim slučajevima i adekvatnim pokazateljem strukture tih objekata. Intenzionalno razumevanje funkcija i formula bi tako omogućilo filozofsko opravdanje korišćenja njihovih samoreferencijalnih karakterizacija koje su, u našoj interpretaciji, posledica teorema o fiksnoj tački. Ali u kojoj meri se teoreme o fiksnoj tački uopšte tiču intenzionalnog značenja objekata na koje se odnose i njihovih definicija?

3.1 INTENZIONALNI ASPEKTI FIKSNIH TAČAKA

Teoreme o fiksnoj tački tvrde postojanje objekata koji stoje u određenoj relaciji ekvivalencije prema objektima koji su na neki način dobijeni od njih. Ta relacija ekvivalencije može se razumeti i na potpuno ekstenzionalan način - tako da se njom tvrdi identitet vrednosti dve funkcije od kojih jedna koristi drugu funkciju ili njen indeks, ili identitet istinosnih vrednosti dve rečenice od kojih jedna referira na kôd druge. U tom smislu se za te teoreme ne može reći da tvrde postojanje samoreferencijalnih fenomena. Ono što ih čini tvrđenjima o samoreferenciji jeste njihovo intenzionalno razumevanje prema kom instance tih teorema pružaju definicije algoritama za izračunavanje date funkcije ili opis značenja neke rečenice. Upravo takvo razumevanje je prvenstveno i motivisalo te rezultate i njihovu primenu, iako su oni kasnije postali nezavisni od njega. U nastavku ćemo u nešto više detalja razmotriti ulogu intenzionalnog značenja u formulisanju i primeni lingvističkih fiksnih tačaka, pre svega onih koje se tiču predikata dokazivosti.

Istraživanje fiksnih tačaka predikata dokazivosti i njegove negacije podstaknuto je Gedelovim dokazom teorema o nepotpunosti, kao i rešavanjem Henkinovog problema. Pošto je, po definiciji, fiksna tačka nekog predikata ili formule svaka rečenica koja je dokazivo ekvivalentna onoj koja joj pripisuje svojstvo izraženo tim predikatom ili formulom, isti predikat dokazivosti može imati više različitih fiksnih tačaka. Takođe, različiti predikati i formule mogu biti uzeti da stoje za predikat dokazivosti. Ipak, videćemo da među njima može biti značajnih razlika. Fiksne tačke koje su Henkin i Gedel koristili shvaćene su intenzionalno - kao rečenice koje za sebe kažu da jesu, odnosno nisu, dokazive. Dakle, one treba da sadrže formulu koja zaista izražava svojstvo dokazivosti u nekom formalnom sistemu. Fiksna tačka te formule, osim toga, treba da bude rečenica koja nije samo ekstenzionalno ekvivalentna onoj koja tvrdi njenu dokazivost ili nedokazivost u sistemu, već ima i isto značenje. U slučaju Gedelove rečenice to se vidi, sa jedne

10 „Tritt nicht bei „objektiven" Begriffen wieder die Schwierigkeit des Reduzibilitäts*ax*<*ioms*> auf, wenn diese 1.) nicht etwas Einfaches sind, sondern selbst eine Struktur haben und daher eine Gesamtheit von Begriffen „*involvieren*" können."

strane, u načinu na koji je on definisao predikat dokazivosti - kao egzistencijalnu generalizaciju predikata dokaza u datoj teoriji koji pokazuje strukturu koju dokazi u njoj imaju, i sa druge strane, u načinu na koji je konstruisao fiksnu tačku negacije tog predikata, a koji tu rečenicu zaista čini samoreferencijalnom. Naime, Gedelova rečenica je oblika $\neg\mathbf{Prov}\,(\mathbf{t})$ gde se term $\mathbf{t}$ svodi na numeral koji označava kôd rečenice $\neg\mathbf{Prov}\,(\mathbf{t})$. Henkinova formulacija njegovog problema takođe sugeriše da se on odnosi na formulu koja zaista za sebe kaže da je dokaziva, tj. na formulu kodiranu brojem q koja izražava iskaz da je formula kodirana brojem q dokaziva u datom sistemu (Smoryński, 1991, p. 113). Međutim, iako je interesovanje za fiksne tačke u metamatematičkim istraživanjima u početku bilo motivisano njihovim intenzionalnim razumevanjem, videćemo da je uskoro taj intenzionalni aspekat izgubljen iz vida.

U nekim slučajevima, kao u slučaju Gedelove teoreme, to ne utiče na krajnji rezultat kojem razmatranje relevantnih fiksnih tačaka vodi. Naime, videli smo da je jedino svojstvo predikata dokazivosti koje se koristi u dokazu Gedelove prve teoreme njegova „ekstenzionalna adekvatnost", tj. činjenica da je on dokaziv u teoriji za sve i samo one formule koje su zaista dokazive u toj teoriji. To svojstvo predikata ili formula, poznato kao *Krajselov uslov* (Georg Kreisel) ili slaba predstavljivost, smatra se minimalnim uslovom koji neki predikat ili formula treba da zadovolji da bi na adekvatan način izražavao svojstvo dokazivosti. Takođe se ispostavilo da ulogu Gedelove rečenice u dokazu može imati bilo koja rečenica koja je dokazivo ekvivalentna rečenici koja tvrdi njenu nedokazivost. To pokazuje da ekstenzionalizacija predikata nedokazivosti i njegove fiksne tačke ne utiče na konačan rezultat kojim se pokazuje njena neodlučivost.

U nekim drugim slučajevima, intenzionalni aspekti fiksnih tačaka mogu biti ključni za dolaženje do odgovarajućih rezultata. Tako, na primer, izbor formule koja bi trebalo da stoji za svojstvo dokazivosti, među njoj ekvivalentnim formulama, može odrediti njena svojstva. To se pokazuje na primeru Henkinovog problema za koji je Krajsel pokazao da može biti rešen na različite načine u zavisnosti od toga koja formula je uzeta kao ona koja izražava svojstvo dokazivosti (Smoryński, 1991, pp. 114-116). On sam je predložio dve formule od kojih obe zadovoljavaju njegov uslov, a za koje važi da je fiksna tačka prve dokaziva, a fiksna tačka druge opovrgljiva. Lebova teorema koja, formulisana pomoću Gedelovog predikata dokazivosti, daje pozitivan odgovor na isti problem, takođe ključno zavisi od toga koji je predikat korišćen kao predikat dokazivosti.

Činjenica da fiksne tačke različitih formula, od kojih sve u ekstenzionalnom smislu jednako dobro predstavljaju neko svojstvo, mogu imati različite osobine, pokazuje da se te osobine (kao što je dokazivost u nekom formalnom sistemu) tiču nekih intenzionalnih aspekata formula. Ti intenzionalni aspekti formula sadržani su u njihovoj strukturi ili načinu na koji su one izgrađene. Značenje formule koja stoji za dokazivost postaje relevantno i u kontekstu dokaza DRUGE TEOREME O NEPOTPUNOSTI. To zahteva formulisanje nekih dodatnih uslova koje formule treba da zadovolje da bi izražavale svojstvo dokazivosti. Najuspešnija analiza tih uslova data je u obliku *Lebovih uslova izvodivosti.* Ti uslovi korišćeni su u dokazu njegove teoreme, kao i u dokazu DRUGE TEOREME O NEPOTPUNOSTI. Ako $\vdash A$ stoji za 'A je dokazivo u datoj formalnoj teoriji', a Pr stoji za predikat te teorije koji bi trebalo da izražava to svojstvo, Lebovi uslovi su sledeći:

D1 $\dfrac{\vdash A}{\vdash Pr\,([A])}$

D2 $\vdash (Pr\,([A]) \wedge Pr\,([A \rightarrow B])) \rightarrow Pr\,([B])$

D3 $\vdash Pr\,([A]) \rightarrow Pr\,([Pr\,([A])])$.

Umesto da podstaknu dalja istraživanja o osobinama predikata koji izražava svojstvo dokazivosti, ovi uslovi postali su osnova za formulisanje aksiomatske teorije koja se bavi dokazivošću. Kao odgovarajuća osnova za takvu teoriju uzeta je modalna logika u kojoj je dokazivost predstavljena operatorom koji karakterišu navedene aksiome i pravila. Logika dokazivosti je tako zasnovana kao sistem iskazne modalne logike (vidi: Boolos, 1993). Analiza fiksnih tačaka dokazivosti nadalje se odvijala isključivo unutar te teorije.

Za analizu fiksnih tačaka kao što su Gedelova, Lebova i Henkinova rečenica, potrebno je proširenje navedenih aksioma formalizacijom dijagonalizacije ili neke njene posledice koja je dovoljna za dokazivanje glavnih osobina tih fiksnih tačaka. Obično se za najjaču posledicu dijagonalizacije uzima LEBOVA TEOREMA koja se datoj teoriji dodaje u obliku sledećeg pravila:

$$\text{LT} \quad \frac{\vdash Pr\,([A]) \to A}{\vdash A}\,.$$

U tako proširenoj teoriji, određena važna svojstva fiksnih tačaka konstruisanih pomoću predikata dokazivosti postaju dokaziva. Jedno od njih je da svaka fiksna tačka može biti eksplicitno definisana, što često omogućava i dokaz njenog postojanja. Drugo važno svojstvo je da su sve fiksne tačke neke formule date teorije dokazivo ekvivalentne jedna drugoj. Iz toga sledi da u toj teoriji način na koji je fiksna tačka konstruisana ne utiče na njene osobine, pa se sve fiksne tačke istog predikata ili formule mogu smatrati jednakim. Najznačajniji rezultat postigao je Solovej (Robert Solovay) koji je dokazao, pomoću dve teoreme o potpunosti, da data modalna teorija pruža adekvatnu analizu dokazivosti u aritmetici. Prema tom rezultatu, neka formula je dokaziva u logici dokazivosti ako i samo ako je formula koja predstavlja njen prevod na jezik aritmetike dokaziva u odgovarajućoj aritmetičkoj teoriji (Smoryński, 2002, pp. 471-476). Dalji razvoj ove teorije bio je u velikoj meri usmeren ovim rezultatima i pokušajima da se oni uopšte i dokažu i za druge predikate kao što je Roserov predikat. Ispostavilo se, međutim, da fiksna tačka Roserovog predikata nije nužno jedinstvena. Zavisnost njenih osobina od načina njenog konstruisanja čini tu fiksnu tačku intenzionalnijom od Gedelove i Henkinove rečenice.

Jedino svojstvo rečenica nalik samoreferenciji koje se može izučavati u ovoj oblasti jeste to da su one dokazivo ekvivalentne rečenici koja im pripisuje određeno svojstvo, gde je ta ekvivalencija shvaćena kao slaganje istinosnih vrednosti. Pritom se način konstruisanja fiksne tačke koji bi garantovao da ona zaista jeste samoreferencijalna, kao i način konstruisanja predikata koji bi garantovao da on zaista izražava određeno svojstvo, ne uzimaju u obzir. Intenzionalne odlike predikata dokazivosti tako postaju nevidljive u ovoj oblasti. Analiza dokazivosti i njenih fiksnih tačaka je na taj način u potpunosti ekstenzionalizovana.

Intenzionalni aspekti fiksnih tačaka, sadržani u značenju formula čije su to fiksne tačke i u načinu njihovog konstruisanja, u nekim slučajevima ipak jesu relevantni. Na to ukazuje razlika u osobinama fiksnih tačaka Krajselova dva predikata, kao i različitih fiksnih tačaka Roserovog predikata. Intenzionalno razumevanje fiksnih tačaka, kao samoreferencijalnih rečenica koje sebi pripisuju svojstvo dokazivosti ili nedokazivosti u određenoj teoriji, je i motivisalo istraživanje o fiksnim tačkama u metamatematici, iako je ono kasnije postalo nezavisno od te motivacije. Povratkom na intenzionalno značenje i zahtevanjem da formula čijom se fiksnom tačkom bavimo zaista izražava neko svojstvo, kao i da njena fiksna tačka zaista sebi pripisuje to svojstvo, približavamo se pitanju pojmova i njihove samoreferencije. Ako određeni aritmetički predikat izražava *pojam dokazivosti u aritmetičkoj teoriji*, onda metamatematički rezultati koji se tiču njegovih fiksnih tačaka mogu nešto otkriti i o pojmovnoj samoreferenciji. Ti rezultati mogu biti shvaćeni kao da otkrivaju neizbežna ograničenja izražavanja i analize određenih samoreferencijalnih pojmova unutar formalne teorije.

Pitanja o prirodi pojmovne samoreferencije koja odatle proizilaze su samo neka od onih koja treba da budu rasvetljena u budućoj teoriji pojmova. Ali, kako bi to bilo moguće, prethodno moraju biti rešeni problemi koji se tiču osnovnih pojmova te teorije i paradoksa uzrokovanih njihovom samoreferencijom, a koji su usko povezani sa pomenutim ograničenjem njihove formalne analize.

3.2 SAMOREFERENCIJA U TEORIJI POJMOVA

S obzirom na to da bi, kako je Gedel predviđao, teorija pojmova trebalo da postane centralni deo logike, ona bi trebalo i da se bavi nekim fundamentalnim logičkim pitanjima. Ta pitanja su po Gedelovom mišljenju:

> „...ona koja se tiču osnovnih logičkih pojmova, npr. $\in$, pojam, iskaz, klasa, $\supset$, relacija. Dakle, na primer: da li svakoj iskaznoj funkciji odgovara neki pojam, da li postoje klase koje sadrže same sebe, da li je svaki pojam svuda definisan?"[11] (Gödel, 2020a, p. 62).

Tokom istorije logike dati su različiti odgovori na ta pitanja, sa ciljem da se razjasni pojam pojma i njegove ekstenzije (tj. odgovarajuće klase). Rešenja paradoksa koji se tiču pojmova tražena su u negativnim odgovorima na svako od navedenih pitanja. Videćemo da je, po Gedelovom mišljenju, temeljnije razmatranje poslednjeg pitanja neophodno kako bi naše razumevanje pojmova i njihove primene bilo unapređeno i time otklonjene greške koje vode intenzionalnim paradoksima. Sa druge strane, jedino su pozitivni odgovori na prva dva pitanja, prema kojima pojam može biti definisan proizvoljnom iskaznom funkcijom i primenljiv na sebe, u skladu sa prirodom pojmova. U slučaju samoprimene pojma to se pokazuje i na brojnim primerima kao što su: „pojam pojma, pojam primenljivosti na samo jednu stvar (ili jedan objekat), pojam razlikovanja od skupa svih konačnih matematičkih skupova, pojam pojma sa beskonačnim domenom, itd" (Wang, 1996, 8.6.3).

Upravo bi ta pojmovna samoreferencija, tj. njena mogućnost, mogla da otkrije neke osnovne formalne osobine pojmova koje ih razlikuju od ekstenzionalnih objekata, kao što je njihovo „neinduktivno građenje", pa bi tako imala i značajnu metodološku funkciju u budućoj teoriji. Ako je suditi po teorijama razmatranim u ovom radu, koje se takođe služe samoreferencijom, njena uloga u definisanju pojmova i dokazivanju njihovih svojstava bi takođe mogla biti značajna. Iz tih razloga bi formalizacija pojmovne samoreferencije trebalo da bude neizostavni deo buduće teorije.

Kao što smo videli na početku ovog rada, glavnu prepreku toj formalizaciji predstavljaju paradoksi kojima samoreferencija u određenim slučajevima vodi i koji mogu da učine teoriju u kojoj je ona formalizovana protivrečnom. Gedel, međutim, nije smatrao da je tu prepreku nemoguće prevazići. Baš nasuprot, on je u njoj video značajan podsticaj za razvoj i poboljšanje našeg razumevanja pojmova i njihove primene.

> „Intenzionalni paradoksi se mogu koristiti za dokaz postojanja pojmova. Oni pokazuju da nismo slobodni da uvedemo baš svaki pojam, jer, po definiciji, da smo zaista u potpunosti slobodni, oni [novi pojmovi] ne bi vodili protivrečnostima" (Wang, 1996, 8.5.20).

Odgovarajuće rešenje intenzionalnih paradoksa bi, na prvom mestu, trebalo sugeriše način na koji formiranje pojmova treba da bude regulisano i time učinjeno manje proizvoljnim. Takav pristup intenzionalnim paradoksima odgovarao bi onom koji se u teoriji izračunljivosti i metamatematici zauzima u odnosu na protivrečnosti koje mogu da slede iz određenih instanci funkcionalne i lingvističke samoreferencije. Otkrivene protivrečnosti nisu smatrane pokazateljem da je upotreba tih oblika samoreferencije nelegitimna, nego su umesto toga korišćene za izvođenje značajnih negativnih rezultata. Opšti rezultat do kojeg se na taj način dolazi u teoriji izračunljivosti je da nije svaka izračunljiva funkcija totalna, odnosno definisana za svaki objekat određene vrste. Takav zaključak ističe intenzionalni aspekat izračunljivih funkcija koji potiče od algoritama na

11 „...solche, welche die logischen Grundbegriffe betreffen, z. B. ε, Begriff, Satz, Klasse, $\supset$, Relation. Also z. B.: Gibt es zu jeder Aussagefunktion einen Begriff, gibt es Klassen, die sich selbst enthalten, ist jeder Begriff überall definiert?"

osnovu kojih se one izračunavaju, a koji mogu biti nedefinisani za određene objekte (kao što je, na primer, operacija deljenja sa nulom nedefinisana). Ako zauzmemo sličan stav i po pitanju protivrečnosti koje slede iz nekih instanci pojmovne samoreferencije, videćemo da one takođe mogu otkriti nešto o prirodi pojmova što je zasnovano na njihovoj intenzionalnosti. Štaviše, uvid u prirodu pojmova do kojeg se na taj način dolazi usko je povezan sa rezultatima o prirodi izračunljivih funkcija do kojih se dolazi na veoma sličan način.

U nastavku ćemo nešto detaljnije razmotriti intenzionalne paradokse, čije bi rešavanje predstavljalo početnu tačku u razvoju našeg razumevanja pojmova, i pokušati da utvrdimo kakav uvid u prirodu pojmova oni mogu da pruže.

3.2.1 *Analiza intenzionalnih paradoksa*

U beleškama za kurs iz logike koji je 1939. godine održao na Univerzitetu Notr Dejm, Gedel daje preciznu analizu paradoksa heteroloških, ili kako ih on zove *impredikabilnih predikata*, koji je u uvodu ovog rada predstavljen kao GRELING-NELSONOV PARADOKS (Gödel, 2017a, pp. 106-115). Na drugom mestu, Gedel referira na tu analizu kao na njegovu analizu intenzionalnog paradoksa (Gödel, 2020a, p. 48), što sugeriše da predikate o kojima je reč Gedel interpretira kao pojmove. To čini dati paradoks verzijom intenzionalnog paradoksa koji se tiče *pojma neprimenljivosti na sebe*.

U Gedelovoj analizi, svi pojmovi koji se koriste u izvođenju paradoksa, kao i pretpostavke koje se njih tiču, izdvojeni su i jasno specifikovani. Ti pojmovi su sledeći:

1. pojam objekta - koji obuhvata sve ono što može biti predmet mišljenja,
2. pojam predikata čije je suštinsko svojstvo da je dobro definisan - za svaki objekat je određeno da li se on primenjuje na njega ili ne,
3. pojam primene predikata - koja se za proizvoljan predikat P i objekat x označava sa $P(x)$.

Paradoks onda sledi iz sledećih pretpostavki o navedenim pojmovima:

PRVA je da ako je P predikat, a x proizvoljan objekat, onda je $P(x)$ smisleni iskaz koji je istinit ili lažan;

DRUGA je da ako je za svaki objekat x, $P(x)$ iskaz koji je istinit ili lažan, onda je P predikat;

TREĆA je da je za svaki objekat jedinstveno određeno da li je on predikat ili ne;

ČETVRTA je da je svaki predikat objekat.

Ako navedene pretpostavke važe, onda možemo definisati predikat 'impredikabilan' (u smislu predikata koji se ne primenjuje na sebe), čija bi primena na bilo koji objekat trebalo da daje rečenicu koja je istinita ili lažna. Taj predikat bi takođe trebalo da bude primenljiv i na sebe, jer je prema poslednjoj pretpostavci i on jedan objekat. Međutim, rečenica dobijena tom primenom ne može biti ni istinita ni lažna. Protivrečnost do koje smo na taj način došli pokazuje da postoji neka greška u razumevanju korišćenih pojmova ili navedenih pretpostavki.

Jedna mogućnost je da se neadekvatnim proglasi korišćenje neograničenog kvantifikatora, odnosno *pojma svega* koji bi trebalo da se odnosi na totalitet svih objekata ili svih predikata. Može se tvrditi, kao u teoriji tipova, da takav univerzalni pojam (koji bi se primenjivao na objekte različitih tipova) ne postoji, pa da tako ni predikatu u čijoj definiciji je on korišćen, nijedan pojam ne može odgovarati. Definicija predikata 'impredikabilan' referira na totalitet svih predikata, zbog čega ne definiše odgovarajući pojam.

Gedelove beleške za kurs iz 1939. godine sugerišu da je on u to vreme smatrao navedeno rešenje ispravnim. Ipak, njegov kasniji rad (Gödel, 1944) kao i sveska *Max Phil IX* koju je pisao tokom 1942. i 1943. godine pokazuju da je on zatim promenio mišljenje (cf. Gödel, 1944, p. 138; Gödel, 2020a, p. 48). Umesto navedenog, verovatnijim je smatrao rešenje prema kojem svaki predikat definiše neki pojam, ali nije svaki predikat, kao ni njime definisan pojam, primenljiv na svaki objekat. Moguće je tako da *neke primene pojma, odnosno predikata koji taj pojam izražava, rezultiraju rečenicama koje nisu ni istinite ni lažne.* To bi značilo da je u izvođenju opisanog paradoksa, pojam predikata koji opravdava prvu pretpostavku neadekvatno korišćen. Kako bi ovaj uvid doveo do rešenja paradoksa, on bi trebalo da implicira da je upravo *pojam neprimenljivosti na sebe* jedan od takvih pojmova koji nisu svuda definisani. Konkretno, trebalo bi da važi da taj pojam nije definisan za sebe zbog čega se njegovom samoprimenom ne dobija ni istinit ni lažan iskaz. Takvu primenu pojmova zvaćemo *besmislenom primenom.*

Pretpostavka da neki pojmovi nisu smisleno primenljivi na sve objekte je intuitivno opravdana. Rečenice 'Skup prirodnih brojeva je crven', 'Zgrada Filozofskog fakulteta u Beogradu je mnogobrojna', 'Pojam kuće je paran', izražavaju naizgled ne samo neistinite, već i besmislene primene pojma crvenog, mnogobrojnog i parnog. Shvatanje rečenica koje izražavaju besmislene primene pojma kao smislenih ali lažnih, najčešće ne stvara nikakav problem. Ipak, prethodno razmatranje sugeriše da bi u posebnim okolnostima ono čak moglo voditi paradoksima.

Dakle, moguće rešenje paradoksa *pojma neprimenljivosti na sebe* je da taj pojam nije smisleno primenljiv na sebe, zbog čega ni tvrđenje da se on primenjuje na sebe ni njegova negacija nisu istiniti. To rešenje ipak ne sprečava izvođenje modifikovanog paradoksa koji ne koristi pretpostavku da je *pojam neprimenljivosti na sebe* smisleno primenljiv na sebe. Takav paradoks, odnosno njegovo uopštenje koje se umesto na pojmove odnosi na funkcije, formulisao je takođe Gedel i nazvao ga ČERČOVIM PARADOKSOM zbog toga što se najjednostavnije može predstaviti u Čerčovom λ-računu (Wang, 1996, 8.6.24 - 8.6.26). U nastavku ćemo opisati taj paradoks kao i interpretaciju koja ga čini paradoksom koji se tiče pojmova.

Za neku funkciju reći ćemo da je *regularna* ako je primenljiva na svaki objekat. Regularne funkcije tako odgovaraju pojmovima koji su svuda definisani. Neka je funkcija d definisana na sledeći način:

$$d(F,x) = \begin{cases} F(x), & \text{ako je } F \text{ regularna funkcija} \\ 0, & \text{inače.} \end{cases}$$

Funkcija d bi za svaku datu funkciju trebalo da odluči da li je ona regularna, i ako jeste da je primeni na dati argument, a ako nije da da 0 kao vrednost. Neka je funkcija E definisana na način koji garantuje da za svako x, $E(x) \neq x$. Na primer, njena definicija može biti sledeća:

$$E(x) = \begin{cases} 0, & \text{ako } x \neq 0 \\ 1, & \text{ako } x = 0. \end{cases}$$

Tada možemo definisati funkciju H za koju važi da je za svako x, $H(x) = E(d(x,x))$. Ako je funkcija d regularna, onda je takva i funkcija H. U tom slučaju važiće da je za svako x, $d(H,x) = H(x) = E(d(x,x))$. Ako supstituišemo H mesto x u poslednjoj jednakosti, dobićemo jednakost $d(H,H) = E(d(H,H))$. Argument $d(H,H)$ biće tada fiksna tačka funkcije E, čije postojanje protivreči definiciji te funkcije.

Konstrukcija fiksne tačke funkcije E na vrlo precizan način odgovara konstrukciji korišćenoj u dokazu OPŠTE TEOREME O FIKSNOJ TAČKI (s tom razlikom što je kodiranje funkcija ovde izostavljeno). Kao i u drugim paradoksima, i ovde je ta konstrukcija korišćena za dolaženje do fiksne tačke takve funkcije čija definicija treba da isključi mogućnost postojanja te fiksne tačke.

Ovaj paradoks se može razumeti kao uopštenje i formalno izvođenje paradoksa *pojma pojma koji nije smisleno primenljiv na sebe*. Funkcija $d\,(x,x)$ može biti shvaćena tako da stoji za *pojam pojma koji je smisleno primenljiv na sebe*, a funkcija E kao uopštenje negacije. U tom slučaju bi za *pojam pojma koji nije smisleno primenljiv na sebe* stajala funkcija H. Primena tog pojma na samog sebe vodila bi protivrečnosti na opisani način.

Čini se da je jedini način da se izbegne ovaj paradoks odbacivanje pretpostavke da je funkcija d regularna, tj. da je svuda definisana. To bi značilo da je moguće da ta funkcija nije definisana za argument (H,H), što bi izvedenu jednakost između $d\,(H,H)$ i $E\,(d\,(H,H))$ učinilo bezazlenom (ona više ne bi protivrečila definiciji funkcije E). Iz toga što funkcija d nije definisana za svaki argument oblika (x,x), sledilo bi da postoji neka funkcija, kao što je H, za koju funkcija d ne može da odredi da li je definisana za samu sebe ili ne. Formulisan tako da se odnosi na pojmove umesto funkcija, ovaj zaključak implicira da *ne može uvek biti određeno da li je neki pojam smisleno primenljiv na određeni objekat (kao što je on sam) ili ne.*

Tako se ispostavlja da za izbegavanje intenzionalnih paradoksa nije dovoljno prihvatiti da neki pojmovi, kao što je *pojam neprimenljivosti na sebe*, nisu svuda definisani, pa su zato neke njihove primene besmislene. Osim toga, moramo prihvatiti još i da ne može uvek biti određeno da li je primena pojma na neki objekat smislena ili ne. Na primer, moguće je da se ne može odrediti da li je rečenica '*Pojam crvenog* je smisleno primenljiv na skup prirodnih brojeva' istinita ili lažna. Kada bismo mogli da u svakom slučaju razlikujemo besmislenu primenu pojma od njegove smislene primene, mogli bismo takođe da umesto besmislenim proglasimo te primene lažnim, što bi omogućilo ponovno javljanje paradoksa (nešto slično je gore postignuto korišćenjem funkcije d). Čini se da je to motivisalo Gedelov zaključak da *pojam smislene primenljivosti* ne može svuda biti definisan (cf. Gödel, 2020a, p. 48). Prema Gedelovim rečima:

> „Očigledna primedba da svaki pojam može biti proširen na sve argumente definisanjem drugog pojma koji daje lažne iskaze gde god je početni pojam bio besmislen, može se lako odbaciti isticanjem činjenice da pojam 'smisleno primenljiv' ne mora uvek biti smisleno primenljiv" (Gödel, 1944, p. 137).

Dakle, ne samo da je moguće da *pojam primenljivosti pojma* nije svuda definisan, nego bi isto moglo da važi i za *pojam smislene primenljivosti pojma*.

Ova vrsta rešenja primenljiva je i na druge paradokse koji bi se mogli javiti u vezi sa pojmovima i njihovom primenom, kao što su intenzionalne verzije semantičkih paradoksa. Gedel je verovao da su u osnovi semantičkih paradoksa koji referiraju na neki jezik, intenzionalni paradoksi koji se tiču značenja izraženog u tom jeziku (Wang, 1996, pp. 271-272). U njima se, umesto na sebe, pojam primenjuje na svoj opis ili na rezultat svoje primene na neki drugi objekat. Gedelove napomene koje slede nakon njegove analize paradoksa impredikabilnih predikata u beleškama za kurs na Notr Dejmu, sugerišu da je on nameravao da opiše i intenzionalnu verziju PARADOKSA LAŽLJIVCA, koja bi se ticala *pojma istine* koji se primenjuje na iskaze. Nažalost, tekst se prekida baš na tom mestu (Gödel, 2017a, p. 115). U svakom slučaju, predloženo rešenje intenzionalnih paradoksa bilo bi primenljivo i na taj paradoks. Ono bi impliciralo da *pojam istine* koji se koristi u njegovom izvođenju nije svuda definisan i da, konkretno, nije definisan za iskaz koji negira sopstvenu istinitost.

Možemo razmotriti i intenzionalnu verziju RIŠAROVOG PARADOKSA u njegovoj formulaciji datoj u (Thomson, 1962, pp. 115-118). Tu je ovaj paradoks predstavljen kao verzija heterološkog paradoksa koja se tiče predikata koji se primenjuju na prirodne brojeve. Pretpostavimo da se svi takvi predikati nalaze na jednoj listi i definišimo predikat koji se primenjuje na sve i samo one brojeve n za koje važi da se n-ti predikat u listi ne primenjuje na njih. Tako definisan predikat i sam se primenjuje na prirodne brojeve zbog čega bi trebalo da pripada datoj listi predikata.

Do paradoksa dolazimo razmatranjem toga da li se taj predikat primenjuje na sopstveni broj u listi ili ne. Slično kao u paradoksu heteroloških predikata, obe opcije vodiće kontradikciji. Standardno rešenje ovog paradoksa sastoji se u odbacivanju date definicije kao nelegitimne na osnovu toga što se ona odnosi na sve predikate prirodnih brojeva pa tako i na sam predikat koji treba da je njom definisan. Alternativno rešenje, koje bi u većoj meri odgovaralo intenzionalnoj verziji paradoksa, bilo bi da iako datom definicijom jeste određen pojam koji se primenjuje na prirodne brojeve, taj pojam ne mora nužno biti smisleno primenljiv i na broj koji njegova definicija dobija u datoj listi. Mogućnost da se besmislena primena tog pojma na neke brojeve proglasi lažnom ponovo bi vodila reformulaciji paradoksa.

Dva opšta zaključka o pojmovima i njihovoj primeni proizilaze iz prethodne diskusije: prvo, da neki pojmovi nisu definisani za svaki objekat, zbog čega neke njihove primene mogu biti besmislene; i drugo, da ne može uvek biti određeno da li je pojam definisan za neki objekat, odnosno, da li je smisleno primenljiv na njega. Ta situacija je jako slična onoj koju smo zatekli u teoriji izračunljivosti. Kako smo se prisetili na početku ovog poglavlja, primena dijagonalnog argumenta vodila je zaključku da moraju postojati neke izračunljive funkcije koje nisu totalne, odnosno, čija primena na neki objekat ne daje nikakvu vrednost kao rezultat. Takve funkcije ne mogu uvek biti proširene do totalne izračunljive funkcije dodeljivanjem proizvoljnih vrednosti argumentima za koje one nisu definisane, jer bismo tako bili suočeni sa kontradikcijom. Osim toga, takođe je utvrđeno da problem koji se tiče pitanja da li je neka funkcija totalna ne može biti odlučiv. Da jeste, mogli bismo da rekonstruišemo originalni dijagonalni argument i da njegovom primenom pokažemo da postoji totalna izračunljiva funkcija koja ne pripada listi svih totalnih izračunljivih funkcija napravljenoj pomoću funkcije koja navodno odlučuje to pitanje (vidi odeljak 2.1.3.1). Tako se ispostavlja da rezultati o pojmovima koji se mogu izvesti iz intenzionalnih paradoksa predstavljaju pandan rezultatima koji se tiču izračunljivih funkcija, a pomoću kojih su kontradikcije koje slede iz nekih instanci funkcionalne samoreferencije izbegnute.

Za te rezultate može se uzeti da otkrivaju karakteristike koje se tiču intenzionalnosti objekata o kojima je reč - algoritama i pojmova. Paradoksi koji se tiču tih objekata mogu se shvatiti kao posledica ekstenzionalnog pristupa koji se pokazuje u fokusiranju na ekstenzije pojmova i kriterijume za njihovu primenu sa jedne strane, i sa druge strane, na funkcije koje su izračunljive pomoću nekog algoritma umesto na same te algoritme. Takav pristup nameće pretpostavku da ti intenzionalni objekti nužno određuju neki ekstenzionalni objekat - skup koji čini ekstenziju pojma, i graf funkcije definisane pomoću algoritma. Ispostavlja se, međutim, da ta pretpostavka omogućava formulisanje paradoksa. Kako bismo izbegli te paradokse, izgleda da moramo prihvatiti da su dati intenzionalni objekti *parcijalni* u smislu da nisu definisani za svaki objekat, zbog čega ne određuju uvek neki skup objekata kao njihovu ekstenziju ili odgovarajuću funkciju. Ta parcijalnost intenzionalnih objekata pokazuje u isto vreme da su oni u velikoj meri nezavisni od ekstenzionalnih objekata koje mogu određivati. Neki algoritam funkcioniše na osnovu njegove relacije u odnosu na druge algoritme i operacija pomoću kojih je konstruisan, nezavisno od toga koji skup uređenih parova određuje i da li uopšte određuje takav skup. Slično tome, ekstenzija i domen definicije nekog pojma nisu a priori određeni. Razumevanje pojma ne zahteva poznavanje svakog objekta na koji je taj pojam (smisleno) primenljiv. Kako je Hilbert naglašavao, pojam je pre fiksiran u mreži sa drugim pojmovima nego na osnovu kriterijuma njegove primene (Cantini, 2009, p. 8). Za razliku od skupova koji su u potpunosti određeni svojim elementima, bez kojih ne mogu ni postojati, pojmovi ne zavise od objekata na koje se primenjuju. Ta činjenica mogla bi da odredi i njihove karakteristične formalne osobine. Na prvom mestu, ona objašnjava mogućnost njihove samoprimene i samokarakterizacije i istovremeno omogućava rešenje paradoksa koji su posledica te samoreferencije.

Kako bismo dalje sledili ideju o rešenju paradoksa koji se tiču pojmova prepoznavanjem mogućnosti njihove besmislene primene, moramo pokazati da je moguće zasnovati konzistentnu formalnu teoriju u kojoj su paradoksi izbegnuti na taj način. Ta teorija bi trebalo da precizira koje primene pojmova treba da budu proglašene besmislenim, i na kojoj osnovi. Analiza parcijalnosti pojmova koja bi trebalo da dovede do rešenja intenzionalnih paradoksa mogla bi tako da predstavlja početnu tačku u razvoju buduće teorije pojmova. Slične ideje o rešenju paradoksa naivne teorije skupova ograničenjem smislene primenljivosti pojmova kojima su ti skupovi definisani, već su predlagane. Međutim, ti predlozi najčešće nisu bili dovoljno razrađeni i, osim toga, oni ipak nisu pružali adekvatan pogled na prirodu pojmova. Do kraja ovog poglavlja predstavićemo ukratko neke od tih predloga i pokušati da objasnimo zbog čega su oni neuspešni. U poglavlju nakon toga bavićemo se nekim adekvatnijim rešenjima paradoksa koja su u skladu sa ovde predstavljenim idejama, ali su izneta u kontekstu formalne teorije.

3.3 PARADOKSI KAO SINGULARNE TAČKE PRIMENE POJMA

Ideja da su intenzionalni paradoksi zasnovani na instancama besmislenih primena pojmova sledi i iz Raselovog stanovišta prema kojem pojmovi imaju ograničeni *domen smislenosti* (*range of significance*). Prema tom stanovištu, svaki pojam je smisleno primenljiv samo na objekte iz određene klase objekata. Ideja o ograničenom domenu smislenosti u osnovi je Raselove teorije tipova, koja se odnosi na pojmove ili iskazne funkcije. Pojmovi, podeljeni po tipovima prema njihovom domenu smislenosti, formiraju hijerarhiju pojmova počevši od onih koji se primenjuju na objekte koji sami nisu pojmovi i koji pripadaju prvom tipu, preko pojmova pojmova prvog tipa koji pripadaju drugom tipu, pojmova pojmova drugog tipa koji pripadaju trećem tipu, i tako dalje. Nijedan od pojmova u ovoj hijerarhiji ne može biti smisleno primenljiv na sebe niti na bilo koji drugi pojam koji se smisleno primenjuje na njega. Iz toga sledi da su u osnovama paradoksa neke nedozvoljene, samoreferencijalne primene pojmova opisane besmislenim rečenicama, čije razumevanje kao smislenih vodi kontradikciji.

Iako ova teorija nudi rešenje intenzionalnih paradoksa (bar onih koji se tiču samoprimene pojma), cena tog rešenja je značajno odstupanje od uobičajenih načina upotrebe jezika i intuicija koje se tiču pojmova. Ta teorija umnožava pojmove koji su intuitivno jedinstveni, kao što je *pojam pojma*, tako da postoji po jedan takav pojam na svakom nivou hijerarhije. Nijedan od njih ne može pritom biti smisleno primenjen na sebe ili na druge pojmove istog ili višeg tipa. Ta tipska ograničenja uvedena su u jezik teorije u vidu pravila građenja njenih formula. Posledica toga jeste da se mnoge intuitivno smislene, pa i istinite primene pojmova ne mogu formulisati u toj teoriji. Tako se, na primer, u njoj ne može reći da se *pojam pojma* i *pojam apstraktnog* primenjuju na sebe, ili jedan na drugog. Ono što, osim toga, ograničava izražajnu moć teorije jeste i zabrana referiranja na objekte različitih tipova unutar istog iskaza, čime je onemogućeno formulisanje univerzalnih tvrđenja o njima. U toj teoriji ne možemo formulisati tvrđenje o tome da su svi pojmovi jednaki samima sebi ili da nijedan pojam nije fizička stvar, pa čak ni ona koja su u osnovi same te teorije, kao na primer, da se svaki pojam primenjuje na objekte nižeg tipa ili da tvrđenja koja referiraju na sve pojmove nisu smislena. Tako sama pretpostavka da teorija tipova važi čini principe na kojima je ona zasnovana besmislenim (cf. Gödel, 1944, p. 138). Slična kritika zbog neslaganja sa običnim jezikom može se uputiti i rešenju semantičkih paradoksa uvođenjem jezičkih nivoa koje nudi Tarski. Opisivanje svojstava jezika određenog nivoa (kao što je istinitost njegovih rečenica) moguće je prema toj teoriji samo na jeziku nekog višeg nivoa. Time se takođe ograničava izražajna moć jezika i uvode određene distinkcije u jezik koje intuitivno ne postoje.

Kao alternativu, Gedel je sugerisao da se ideja o ograničenom domenu smislenosti može slediti i bez tako drastičnog odstupanja od uobičajenog načina korišćenja pojmova i jezika uopšte. Adekvatnije rešenje paradoksa, koje bi bilo bliže istini, jeste ono koje većinu primena pojma prihvata kao smislene, i na taj način se u najvećoj mogućoj meri slaže sa običnim jezikom. Rešenje intenzionalnih paradoksa ne zahteva, po Gedelovom mišljenju, značajnu izmenu uobičajenog načina korišćenja pojmova. Umesto toga:

> „Možda se čak ispostavi da je moguće pretpostaviti da je svaki pojam smislen svuda osim za određene 'singularne tačke' ili 'granične tačke', tako da bi paradoksi izgledali kao nešto analogno deljenju sa nulom" (Gödel, 1944, p. 138).

Umesto uvođenja čitavog arsenala teorije tipova, Gedel je smatrao da rešenje intenzionalnih paradoksa zahteva izuzimanje samo određenih objekata iz domena smislene primenljivosti pojmova. U tom smislu, njegovo predloženo rešenje može se uporediti sa Fregeovim predlogom za rešenje RASELOVOG PARADOKSA koje se sastoji u odbacivanju mogućnosti primene predikata ili formula na skupove (tj. ekstenzije pojmova) koji su njima definisani. Aksioma komprehenzije i princip ekstenzionalnosti ograničeni su u skladu sa tim. Ipak, i tako ograničena, aksioma komprehenzije vodi kontradikcijama (vidi: Quine, 1955, za jedan od dokaza). Osim toga, kako je Gedel primetio, to ne može biti jedino ograničenje primene predikata (Gödel, 2017a, p. 113). Izuzimanje datog predikata, odnosno njime definisanog pojma ili skupa, iz domena njegove smislene primenljivosti, povlači i druge izuzetke za sobom. Da bismo to videli, uzmimo za primer predikat Q definisan pomoću predikata P na sledeći način: za svako x, $Q(x) =_{def} P(x) \wedge R(x)$. Ako P nije smisleno primenljiv na sebe, onda ni Q ne može biti smisleno primenljiv na njega jer $Q(P) = P(P) \wedge R(P)$ (isto važi i za druge načine građenja predikata).

Da bi rešenje paradoksa zabranom jedino samoreferencijalnih primena pojma bilo prihvatljivo, morale bi da budu uvedene neke restrikcije koje se tiču definicija pojmova i kojima bi situacije kao malopre pomenuta bile izbegnute. Te restrikcije bi istovremeno trebalo da učine Fregeovo rešenje konzistentnim. Takvim rešenjem bavio se, između ostalih, i Hintika (Jakko Hintikka, 1956; 1957). On je došao do zaključka da se do rešenja RASELOVOG PARADOKSA zasnovanog na Fregeovom predlogu može doći samo ako se pretpostavi određena jaka verzija *principa cirkularnosti* (Hintikka, 1957). Naime, Hintika je pokazao da je moguće izvesti kontradikciju u skladu sa uobičajenom formulacijom principa cirkularnosti prema kojoj definicija nekog predikata ne može referirati na totalitet kojem i taj predikat pripada (Hintikka, 1957). Jedini izlaz iz te situacije je, po njegovom mišljenju, uvođenje jače verzije principa cirkularnosti koja zabranjuje definisanje predikata referiranjem na totalitet kojem taj predikat, ili bilo koji drugi koji se javlja u njegovoj definiciji, pripadaju. To bi u navedenom primeru isključilo mogućnost da P pripada totalitetu objekata za koje je predikat Q definisan jer je P korišćeno u definiciji tog predikata. Hintika je smatrao da se slučajevi koje njegov princip cirkularnosti isključuje mogu smatrati Gedelovim 'singularnim tačkama' pošto nisu deo neke previše restriktivne teorije poput teorije redova (Hintikka, 1956, p. 241).

Njegovo rešenje ipak nije u skladu sa Gedelovim shvatanjem pojmova i njihovih definicija. Kao što je već rečeno, čini se da je opravdano Gedelu pripisati stanovište prema kojem bi impredikativne definicije pojmova bile neizostavni deo teorije pojmova, ukoliko ona treba da rasvetli strukturu pojma koji može da se javi kao deo sopstvenog značenja. Takođe, i sam Fregeov predlog za rešenje paradoksa na koji se Hintika oslanja nije u skladu sa Gedelovom idejom da bi samoprimena pojma trebalo da bude dozvoljena u svim slučajevima osim kada vodi kontradikciji.

Kao što smo videli na početku ovog rada, razvoj teorije skupova koji je zatim usledio, vodio je razdvajanju skupova od svojstava i pojmova čiju ekstenziju bi oni, po početnoj pretpostavci, trebalo da predstavljaju. Pokušaji nalaženja izvora paradoksa u pojmovima izraženim formulama koje se javljaju u aksiomi komprehenzije nisu dalje slеđeni. Umesto toga, teorija o skupovima kao ekstenzijama pojmova zamenjena je teorijom koja pretpostavlja u potpunosti ekstenzionalan pojam skupa za koji neograničena aksioma komprehenzije ne važi.

Povratkom na intenzionalni pojam skupa, vraćamo se i pitanju o tome zašto neograničena aksioma komprehenzije ne važi za ekstenzije pojmova (tj. za klase). Čini se da je Gedelova ideja da ta aksioma u njenoj klasičnoj interpretaciji ne važi ne zbog toga što predikati i formule koje se u njoj javljaju ne uspevaju da definišu odgovarajući pojam, već zbog toga što njima definisan pojam ne mora biti smisleno primenljiv na svaki objekat koji pripada domenu na koji ta aksioma referira. U tom slučaju bi rešenje paradoksa koje će pokazati koji oblik aksiome komprehenzije jeste adekvatan za pojmove, trebalo da omogući teorija koja će se baviti njihovom smislenom primenljivošću.

3.4 FORMALIZACIJA SMISLENE PRIMENLJIVOSTI POJMA

Prvi korak ka zasnivanju teorije pojmova bio bi napravljen dolaskom do boljeg razumevanja relacije primene pojma čime bi bili izbegnuti paradoksi koji su, po pretpostavci, uslovljeni pogrešnim razumevanjem te relacije. To sugeriše i sledeća Gedelova napomena:

> „Krajnji uzrok paradoksa leži tome što mi ne vidimo šta je zaista $\in$-relacija (u pojmovnom svetu), nego vidimo njenu zamenu u onome što smo mi sami konstruisali. Jednako malo vidimo šta pojam "pojam" zapravo jeste"[12] (Gödel, 2020a, p. 48).

Iz prethodne rasprave sledi da bi osnovna greška u razumevanju relacije primene pojma mogla biti u tome što ne vidimo da su neke primene pojmova besmislene. To znači da odgovarajuće rešenje paradoksa koje bi istovremeno unapredilo razumevanje pojmova i postavilo osnove teoriji koja se njima bavi treba da prepozna besmislenost određenih primena pojmova. Odatle bi bila izvedena i odgovarajuća ograničenja osnovnih principa te teorije koji se tiču primene pojma. Potreba za prepoznavanjem besmislenih primena pojma unutar teorije može biti deo razloga zbog kog teorija tipova ne pruža odgovarajuću osnovu za analizu pojmova i njihove primene. Naime, opisi primena pojma koje su besmislene prema toj teoriji, tj. koje nisu u skladu sa tipskim ograničenjima, nisu smatrani dobro formiranim formulama te teorije. Iz tog razloga, principi koji ograničavaju smislenu primenljivost pojmova ne mogu ni biti formulisani unutar nje. Sa druge strane, nepostojanje sličnih ograničenja u budućoj teoriji vodilo bi tome da neke njene rečenice izražavaju besmislene primene pojmova. Takve rečenice bi takođe bile besmislene, što pre svega znači da nisu ni istinite ni lažne. Osobine besmislenih rečenica koje ih razlikuju od drugih vrsta rečenica koje takođe nemaju istinosnu vrednost trebalo bi da se pokažu u kontekstu, na primer u njihovoj interakciji sa drugim formulama ili u statusu koji dobijaju u okviru nekih osnovnih principa teorije koji se tiču značenja njenih rečenica (kao što je aksioma komprehenzije).

Jedna od osnovnih vrsta tvrđenja u vezi sa primenom pojma, koju možemo formulisati unutar teorije koja se njom bavi, jeste ona koja povezuje opise primena pojmova sa rečenicama koje nastaju primenom odgovarajućeg predikata. Ta veza potiče od toga što, kao što smo rekli na početku ovog rada, pojmovi predstavljaju intenzionalno značenje predikata. Kao nužan i dovoljan uslov da se pojam c primenjuje na neki objekat zahtevali bismo tada da se odgovarajući predikat C primenjuje na ime tog objekta. To bi moglo biti izraženo formulom oblika $c\boldsymbol{\eta}x \leftrightarrow Cx$, gde bi $\boldsymbol{\eta}$ stajalo za relaciju primene pojma. Instanca te formule bila bi, na primer, rečenica koja kaže da se *pojam beskonačnog* primenjuje na skup prirodnih brojeva ako i samo ako je skup $\mathbb{N}$ beskonačan. Međutim, ako su neke primene određenog pojma besmislene, onda bi neke rečenice o primeni tog pojma formulisane unutar teorije trebalo da predstavljaju opise tih besmislenih primena, pa tako ne bi mogle biti ni istinite ni lažne. Ako, sa druge strane, svaka primena predikata u toj teoriji daje rečenicu koja je istinita ili lažna, onda nijedna od tako dobijenih rečenica ne može biti ekvivalentna onoj koja tvrdi besmislenu primenu pojma. Na primer, pod pretpostavkom da *pojam crvenog* nije smisleno primenljiv na skup prirodnih brojeva, to da se taj pojam primenjuje na skup prirodnih brojeva ako i samo ako je skup $\mathbb{N}$ crven, ne može da važi sve dok je poslednja rečenica shvaćena kao istinita ili lažna.

12 „Letzter Grund für *Antin<omien>* ist, dass wir nicht sehen, was die ε-Relation eigentlich ist [im Reich der Begriffe], sondern wir sehen einen Ersatz in dem, was wir konstruiert haben. Ebensowenig sehen wir, was der Begriff "Begriff" ist."

Rečenice koje bi odgovarale besmislenim primenama pojma bile bi one izgrađene primenom predikata na ime za koje taj predikat nije definisan, zbog čega ne bi bile ni istinite ni lažne. Predikati koji nisu definisani za sve objekte domena nazivaju se *parcijalnim predikatima*. Čini se, na osnovu rečenog, da bi upravo pripuštanje takvih predikata u buduću teoriju trebalo da omogući formulisanje veze između primene pojma i primene njemu odgovarajućeg predikata.

3.4.1 *Uključivanje parcijalnih predikata u teoriju*

Kako bismo prepoznali mogućnost besmislenih primena pojma i na taj način uspeli da izbegnemo paradokse u budućoj teoriji, neke predikate te teorije (kao što je i sam predikat koji stoji za primenu pojma) moramo razumeti kao parcijalne, odnosno, kao predikate čija primena na određena imena ne daje nužno rečenice koje su istinite ili lažne. Pripuštanje parcijalnih predikata u teoriju zahteva njeno prilagođavanje mogućnosti da neke njene rečenice izgrađene pomoću tih predikata nisu smislene. Postoje dva načina na koje teorija može biti prilagođena toj mogućnosti:

1. ona može biti zasnovana na nekoj logici koja, za razliku od klasične, dvovrednosne logike, omogućava interpretaciju njenih rečenica kao besmislenih, odnosno, ni istinitih ni lažnih;

2. ili klasična logička osnova teorije može biti proširena na način koji omogućava karakterizaciju rečenica kao smislenih ili besmislenih unutar te teorije.

Prva opcija implicira da bi besmislenost nekih primena pojmova trebalo da se pokaže i u samom jeziku teorije koja se njima bavi, dok druga opcija pretpostavlja zauzimanje određene „metaperspektive” u odnosu na rečenice koje opisuju te besmislene primene.

U nastavku ćemo predstaviti neke značajne teorije koje su u skladu sa jednom ili drugom od navedenih opcija. Ono što ćemo pritom pokušati da utvrdimo jeste koju od njih treba da sledimo kako bismo uspeli da zasnujemo teoriju koja u najvećoj mogućoj meri odgovara teoriji kakvu je Gedel zamislio. Jedno od glavnih pitanja koje u vezi sa tim treba razmotriti jeste koja od navedenih opcija omogućava korišćenje aksiome komprehenzije $\exists y \forall x\,(y \boldsymbol{\eta} x \leftrightarrow A\,(x))$ u onom obliku u kom ona najviše odgovara pojmovima, tj. u kom predikat $\boldsymbol{\eta}$ može najprirodnije biti interpretiran tako da stoji za relaciju primene pojma. Kako bismo zadovoljili Gedelov zahtev da svaki predikat definiše neki pojam, nikakva opšta ograničenja u vezi sa tim koja formula može da stoji na mestu formule A u aksiomi ne bi bila dozvoljena. Takođe, ono što bi odgovaralo intuitivnom razumevanju pojmova prema kojem ne postoje nikakva a priori ograničenja njihove primene, bilo bi odsustvo bilo kakvih opštih ograničenja univerzalnog kvantifikatora koji se javlja u aksiomi. Jedina opravdana ograničenja bila bi ona koja se tiču takozvanih 'singularnih tačaka' i koje nameće potreba za izbegavanjem paradoksa. Zadovoljavanje ta dva zahteva u budućoj teoriji pojmova bi trebalo da garantuje njenu saglasnost sa prirodnim jezikom i intuitivnim razumevanjem pojmova.

Teorije koje će ovde biti predstavljene nisu eksplicitno predložene kao teorije koje se tiču pojmova i relacije primene pojma. Međutim, nijednu od njih takođe ne karakteriše primena principa ekstenzionalnosti, što ostavlja mogućnost njihove intenzionalne interpretacije. Pod tim principom mislimo na aksiomu $\forall x\,(x \in a \leftrightarrow x \in b) \rightarrow a = b$, kao i bilo koju drugu aksiomu ili pravilo koje implicira da su dva entiteta koja sadrže iste elemente (tj. primenjuju se na iste objekte) jednaka. U nastavku ćemo ukratko opisati te teorije i razmotriti koja od njih može najbolje biti prilagođena tako da formira osnovu buduće teorije pojmova.

3.5 PARCIJALNI PREDIKATI U TROVREDNOSNOJ LOGICI

Predikati su u klasičnoj logici interpretirani ekstenzionalno, odnosno, tako da stoje za skupove koji čine njihovu ekstenziju. Pošto svaki element domena pripada ili skupu kojim je određeni predikat interpretiran ili njegovom komplementu kojim je interpretirana negacija tog predikata, svaka rečenica klasične logike formirana primenom predikata na neko ime biće istinita ili lažna. U skladu sa tim interpretirani su i n-arni klasični veznici koji se primenjuju na te rečenice: kao totalne operacije na n-arnom Dekartovom proizvodu skupa dve istinosne vrednosti. Klasična logika se iz tog razloga može činiti neodgovarajućom za razmatranje parcijalnih predikata. Umesto nje, neka alternativna logička teorija bi mogla biti predložena, koja uzima u obzir mogućnost da rečenice formulisane na njenom jeziku nisu ni istinite ni lažne. Takva logička teorija pripadala bi takozvanim *viševrednosnim logikama* u kojima rečenici može biti pripisana i neka druga vrednost osim istine i laži.

Viševrednosne logike međusobno se razlikuju po broju dodatnih vrednosti koje pretpostavljaju i po načinu na koji reinterpretiraju veznike kako bi ih prilagodile njima. One su razvijene u različite svrhe. Jedna od njih, koja je najbliža našem predmetu istraživanja, jeste opis logičkih osobina neodlučivih predikata. Kao što smo videli, neki od najznačajnijih rezultata teorije izračunljivosti su negativni rezultati koji tvrde rekurzivnu nerešivost određenih problema koji se tiču pripadnosti nekom skupu. Ti problemi tiču se takvih skupova za koje ne postoji rekurzivna funkcija koja za svaki objekat može da odluči da li je on u tom skupu ili ne. Za bar neke objekte takva funkcija neće dati odgovor. Umesto da parcijalno određuje neki skup, ta funkcija može biti shvaćena kao da određuje neki parcijalni predikat. Istinosna vrednost rečenice formirane pomoću tog predikata može tako biti neizračunljiva pomoću te funkcije. Kako bi prilagodio logiku teorije toj mogućnosti, Klini ju je zasnovao na sistemu *trovrednosne logike* u kojoj treća vrednost ne predstavlja dodatnu istinosnu vrednost, već nedostatak istinosne vrednosti. U skladu sa kontekstom u kom je ta teorija razvijena, treća vrednost može se razumeti kao *neodređeno* na osnovu datog izračunavanja. Klini je predložio dve mogućnosti za takvu interpretaciju logičkih veznika koja uzima u obzir postojanje formula bez određene istinosne vrednosti (Kleene, 1971, pp. 332-340). U oba slučaja, veznici odgovaraju klasičnim veznicima kada su primenjeni na rečenice sa istinosnom vrednošću, ali se njihovo ponašanje u prisustvu neodređenih formula razlikuje.

Prema prvoj interpretaciji, ako ijedna od formula na koje je veznik primenjen nema istinosnu vrednost, onda formula koja tom primenom nastaje takođe nema istinosnu vrednost. Na toj interpretaciji zasnovana je Klinijeva *slaba trovrednosna logika* (*weak three-valued logic*). Prema drugoj, formula izgrađena od neke formule bez istinosne vrednosti, i dalje može imati istinosnu vrednost, ako ona može biti određena u skladu sa klasičnim istinosnim tablicama na osnovu istinosne vrednosti njenih ostalih potformula.

Na primer, rečenica $P(a) \vee Q(a)$ je istinita čim je nekoj od rečenica $P(a)$ ili $Q(a)$ pripisana ta istinosna vrednost, nezavisno od istinosne vrednosti druge rečenice ili činjenice da ona uopšte može biti određena. Slično tome, rečenica $P(a) \wedge Q(a)$ je lažna čim je neki od njenih konjunkata lažan, dok je $\neg P(a)$ neodređena samo ako je i $P(a)$ takva. Implikacija $P(a) \to Q(a)$ je definisana kao $\neg P(a) \vee Q(a)$ a ekvivalencija $P(a) \leftrightarrow Q(a)$ kao $(P(a) \to Q(a)) \wedge (Q(a) \to P(a))$. Na ovoj interpretaciji veznika zasnovana je Klinijeva *jaka trovrednosna logika* (*strong three-valued logic*).

Te dve različite interpretacije veznika mogu biti opravdane različitim vrstama izračunavanja istinosne vrednosti neke složene formule (Feferman, 1984, p. 88). Jaka interpretacija odgovara *paralelnom* izračunavanju, u kojem se vrednosti potformula neke formule izračunavaju istovremeno, dok slaba interpretacija odgovara *sekvencijalnom* izračunavanju u kojem se vrednosti tih potformula izračunavaju jedna za drugom.

Postoji još jedan način na koji slaba interpretacija veznika može biti opravdana. Problemi odlučivanja u vezi nekog predikata mogli bi u nekim slučajevima biti posledica činjenice da taj predikat nije smisleno primenljiv na dati objekat. To bi bio dobar razlog da se izračunavanje istinosne vrednosti rečenice formirane tom primenom, koje bi trebalo da se zasniva na značenju tog predikata, ne završava dolaženjem do određene istinosne vrednosti. U ovom kontekstu, treća vrednost bila bi interpretirana kao *besmislenost*. Ako je neka rečenica besmislena, onda bi i svaka druga rečenica izgrađena pomoću nje primenom veznika trebalo takođe da bude besmislena, što bi bilo u skladu sa slabom interpretacijom veznika.

Još jedna važna trovrednosna logika predložena je od strane Lukašieviča (Jan Łukasiewicz, 1970, pp. 87-89). Ona se razlikuje od Klinijeve jake trovrednosne logike samo u interpretaciji implikacije i ekvivalencije. Na osnovu njene definicije u Klinijevoj logici, sledi da će implikacija čiji antecedens i konsekvens nemaju istinosnu vrednost, i sama biti bez istinosne vrednosti. Takođe, ako nijedna formula u ekvivalenciji nema istinosnu vrednost, onda je neće imati ni ta ekvivalencija. Nasuprot tome, u Lukašievičevom sistemu će implikacija i ekvivalencija biti istinite ako nijedna od formula od kojih su one izgrađene nije ni istinita ni lažna. Za ta dva veznika koristićemo različite simbole: $\supset$ za Lukašievičevu implikaciju, i $\equiv$ za njegovu ekvivalenciju. Treća vrednost formula u Lukašievičevom sistemu interpretirana je kao *moguće*. Jedna od njenih primena bila bi formalizacija rasuđivanja o budućim događajima opisanim rečenicama koje u datom trenutku nisu ni istinite ni lažne.

Svaki od pomenutih logičkih sistema korišćen je na neki način u cilju postizanja alternativnog rešenja skupovno-teorijskih i semantičkih paradoksa koje se ne zasniva na uvođenju tipova ili nivoa jezika. Trovrednosna logika čini se odgovarajućom formalnom osnovom koja omogućava razmatranje primena predikata koje nisu ni istinite ni lažne. Ona eksplicitno prepoznaje neke primene kao takve, pripušta ih u jezik i reinterpretira veznike kako bi ih prilagodila njima. To naizgled omogućava da u teoriji zasnovanoj na takvoj logičkoj osnovi svaka primena pojma ipak bude u korelaciji sa odgovarajućom primenom predikata, jer bi način na koji je ta korelacija formulisana bio drugačiji od klasičnog. Naime, ekvivalencija koja bi bila korišćena za formulisanje te korelacije, a samim tim i za formulisanje aksiome komprehenzije, uzimala bi u obzir i rečenice koje nemaju istinosnu vrednost jer izražavaju neku besmislenu primenu pojma ili primenu predikata na ime za koje taj predikat nije definisan. To bi potencijalno omogućilo konzistentno korišćenje neograničene aksiome komprehenzije. Važno pitanje koje tada treba razmotriti jeste *koja od navedenih trovrednosnih logika, ako ijedna, nudi takvu interpretaciju ekvivalencije koja bi odgovarala njenoj ulozi u aksiomi komprehenzije?*

3.5.1 *Interpretacija implikacije i ekvivalencije*

Problemom interpretacije implikacije i ekvivalencije u kontekstu trovrednosne logike bavi se Vang u (Wang, 1961), gde formuliše račun parcijalnih predikata zasnovan na određenom trovrednosnom logičkom sistemu. Na drugom mestu, komentarišući Gedelove ideje o teoriji pojmova, Vang pominje taj račun kao logičku teoriju koja bi mogla da obezbedi sredstva za bavljenje smislenom primenljivošću, a time i rešenje intenzionalnih paradoksa (Wang, 1987, p. 310). Taj račun zasnovan je na jakoj trovrednosnoj interpretaciji svih veznika osim implikacije i ekvivalencije. Razlog zbog kog Vang smatra Klinijevu jaku interpretaciju ovih veznika neodgovarajućom je taj što ona čini neke osnovne logičke zakone kao što je $A \to A$, nevaljanim. Osim toga, aksioma komprehenzije ne bi mogla da bude formulisana korišćenjem ekvivalencije definisane pomoću te implikacije jer takva ekvivalencija ne bi bila istinita za sve valuacije formula koje povezuje. Ako i za jednu valuaciju neka od tih formula nema istinosnu vrednost, ekvivalencija takođe neće imati istinosnu vrednost. Alternativa koja se sastoji u usvajanju Lukašievičeve interpretacije prema kojoj bi implikacija bila istinita u slučaju da njen antecedens i konsekvens nemaju istinosnu vrednost, ne predstavlja značajno bolju opciju. U toj interpretaciji bi, na primer, $A \supset \neg A$ bilo istinito i kada A nema istinosnu vrednost. Po Vangovom mišljenju, to pokazuje da razumevanje implikacije u trovrednosnoj logici kao funkcije vrednosti njenog antecedensa i konsekvensa nije odgovarajuće. Naša intuicija ostavlja nas bez pomoći u odabiru odgovarajuće funkcije između različitih opcija od kojih nijedna nije bez naizgled neprihvatljivih posledica. Umesto ovih nepouzdanih načina razumevanja implikacije, Vang predlaže interpretaciju prema kojoj implikacija stoji za *relaciju semantičke posledice.* Tako shvaćena implikacija biće istinita samo ako svaka interpretacija koja čini njen antecedens istinitim, čini takvim i njen konsekvens. Zbog prisustva dodatne vrednosti, dovoljan uslov za istinitost ove implikacije je nešto stroži: osim navedenog, potrebno je još i da u svakoj interpretaciji u kojoj antecedens ne dobija istinosnu vrednost, konsekvens nije lažan. Tako shvaćena implikacija je veznik na višem nivou u odnosu na druge veznike u sistemu, i može biti korišćena za opisivanje odnosa između istinosnih vrednosti formula izgrađenih pomoću njih.

Kako je naznačeno u samom tekstu (Wang, 1961, p. 283), nastavak rada je bio planiran, koji bi sadržao opis teorije skupova zasnovane na računu parcijalnih predikata, ali nikada nije objavljen. U tom cilju, Vang bi morao da definiše ekvivalenciju koja bi služila za formulisanje neke verzije aksiome komprehenzije. Ta ekvivalencija bi verovatno bila shvaćena na sličan način kao implikacija, tj. tako da izražava simetričnu verziju relacije semantičke posledice. Upravu tu relaciju između formula, koju označava sa $\vdash\dashv$, koristi Skot za opisivanje odnosa između pripadnosti skupu (interpretiranom λ-apstrakcijom) i instance odgovarajuće formule (Dana Scott, 1975):

$$y \in \lambda x.A \vdash\dashv A^x_y.$$

Veznik $\vdash\dashv$ može biti shvaćen kao da tvrdi jednaku snagu formula koje povezuje, odnosno jednakost njihovih semantičkih posledica. S obzirom na to da rečenice koje izražavaju neku primenu pojma i primenu odgovarajućeg predikata mogu biti shvaćene kao da imaju isto značenje, to bi zapravo mogao biti odgovarajući opis njihovog odnosa.

Najvećim nedostatkom opisane interpretacije implikacije i ekvivalencije smatrana je činjenica da se tako shvaćeni veznici ne mogu iterirati (cf. Feferman, 1984, p. 94). Konkretno, oni ne bi mogli da budu sadržani u formuli koja se javlja u aksiomi komprehenzije.

Potreba za takvom interpretacijom ekvivalencije u kojoj bi ona bila istinita i kada su formule koje povezuje bez klasične istinosne vrednosti, je verovatno ono što je navelo Fefermana (Solomon Feferman) i Aksela (Peter Aczel) da izaberu Lukašievičevu trovrednosnu logiku kao osnovu za njihovo razmatranje rešenja skupovno-teorijskih paradoksa bez uvođenja tipova (Feferman, 1984, pp. 87-93). Pokazalo se da je aksioma komprehenzije formulisana pomoću Lukašievičeve ekvivalencije konzistentna u ovoj teoriji, ali samo ako je ograničena na formule u kojima se ne javljaju implikacija i ekvivalencija (jedini nemonotoni veznici ove logike)[13]

Tako se isto ograničenje u odnosu na ta dva veznika javlja i ovde kao i u formulaciji aksiome komprehenzije pomoću Vangove i Skotove implikacije i ekvivalencije. Posebno mesto koje je na taj način pripisano ovim veznicima moglo bi, međutim, odgovarati ulozi koju bi oni imali u budućoj teoriji pojmova. U kontekstu teorije pojmova, implikacija i ekvivalencija bi prvenstveno mogle biti korišćene za formulisanje osnovnih principa koji se tiču primene pojmova i relacija u kojima oni stoje. U tom smislu bi se njihova uloga razlikovala od uloge drugih veznika (konjunkcije, disjunkcije i negacije) koji bi bili korišćeni u opisu građenja pojmova. U prilog takvoj interpretaciji može se tvrditi da je jedino veznike koji učestvuju u građenju pojmova neophodno razumeti u trovrednosnom smislu i dozvoliti njihovo javljanje u formulama koje mogu biti supstituisane u aksiomi komprehenzije. Pošto implikacija i ekvivalencija naizgled ne učestvuju u građenju pojmova, ti zahtevi se tada njih ne bi ticali.

Rešenje paradoksa tako može biti postignuto u teoriji zasnovanoj na trovrednosnoj logici ali samo ako je implikaciji i ekvivalenciji pripisan poseban status u odnosu na druge veznike,

13 Intuitivno, veznik je monoton ako se ne može desiti da formula napravljena pomoću njega „izgubi” istinosnu vrednost kada neka od njenih potformula „postane” istinita ili lažna. Lukaševičeva implikacija nije monotona jer je formula $A \supset B$ istinita ako su obe formule A i B bez istinosne vrednosti, ali ako je A istinito dok B nema istinosnu vrednost, ili ako A nema istinosnu vrednost a B je lažno, ta implikacija neće imati istinosnu vrednost. Slično tome, $A \equiv B$ je istinito ako su i A i B bez istinosne vrednosti, ali nema istinosnu vrednost ako je samo jedna od te dve formule istinita ili lažna.

koji može biti opravdan s obzirom na posebnu ulogu koje one mogu imati u teoriji pojmova. Paradoksi su izbegnuti zahvaljujući mogućnosti da rečenice koje tvrde primenu određenog pojma ili njemu odgovarajućeg predikata, a koje bi u dvovrednosnoj logici vodile kontradikcijama, budu proglašene besmislenim.

Feferman je ipak bio protiv zasnivanja teorije bez tipova na bilo kojoj trovrednosnoj logici. Razlog za to je činjenica da neki logički zakoni koji se koriste u svakodnevnoj matematičkoj praksi ne važe u tim logičkim sistemima (posebno ako implikacija nije modifikovana na određeni način). Već smo spomenuli da formula $A \to A$ nije valjana u Klinijevom sistemu. Isto važi i za zakon univerzalne instancijacije, tj. za formulu $\forall x F(x) \to F(a)$. Sa druge strane, u Lukašievičevoj logici ne važi modus ponens $(A \wedge (A \supset B)) \supset B$, kao ni slaba redukcija $(A \supset \neg A) \supset \neg A$. Zakon isključenja trećeg $A \vee \neg A$ ne važi ni u jednom od ta dva sistema, što je u tom kontekstu i očekivano, ali isto važi i za zakon kontradikcije $\neg (A \wedge \neg A)$ (nezavisno od interpretacije implikacije). Po Fefermanovom mišljenju, to pokazuje da „ništa nalik uobičajenom rasuđivanju ne može biti izvedeno u bilo kojem od tih logičkih sistema” (Feferman, 1984, p. 95). Umesto analize parcijalnih predikata unutar neke trovrednosne logike, Feferman predlaže njihovo uključivanje u klasičnu logiku. Način na koji je klasična logika modifikovana tako da dozvoli i potencijalno besmislene primene predikata je, kao što ćemo videti, inspirisan načinom na koji su parcijalni predikati i njihova veza sa pojmovima tretirani unutar trovrednosnih logika.

Pre nego što se okrenemo klasičnim teorijama, ukratko ćemo se osvrnuti na jedan hibridan pristup koji se, korišćenjem klasičnih principa rasuđivanja, bavi paradoksalnim rečenicama kao rečenicama formulisanim u kontekstu trovrednosne logike. U toj teoriji, predloženoj od strane Bočvara (Dmitry Anatol'evich Bochvar), paradoksalne rečenice su interpretirane kao besmislene, ali je logika koja se njima bavi klasična, dvovrednosna logika.

3.5.2 *Bočvarov račun iskaza*

Prema Bočvaru, Klinijeva slaba trovrednosna logika predstavlja odgovarajuće okruženje u kojem besmislene primene predikata (tj. iskaznih funkcija) mogu biti izražene. Njegov cilj je da formalno dokaže besmislenost takvih rečenica koje se, po njegovom shvatanju, javljaju i u osnovama paradoksa. Sa tim ciljem on formuliše takozvani *račun iskaza* (*statement calculus*). Taj račun bi trebalo da sadrži rečenice koje izražavaju primenu određenih predikata, ali i one pomoću kojih su takve rečenice opisane kao istinite, lažne ili besmislene (Bochvar, 1981).

U skladu sa tim, Bočvar razlikuje dve vrste iskaza koje naziva *unutrašnjim* i *spoljašnjim iskazima* i zajedno sa njima dva sistema veznika i kvantifikatora. Unutrašnji iskazi su formule koje opisuju objekte koji pripadaju određenom domenu. Tim iskazima bi trebalo da pripadaju i oni pomoću kojih su formulisani semantički i skupovno-teorijski paradoksi. Logika unutrašnjih iskaza je Klinijeva slaba trovrednosna logika. Sa druge strane, spoljašnji iskazi su interpretirani kao da se odnose na druge iskaze istog računa. Osnovnu vrstu spoljašnjih iskaza čine oni koji opisuju unutrašnje iskaze kao istinite ili lažne. Ti iskazi su formirani primenom spoljašnjih veznika na proizvoljni iskaz A čime su formirani iskazi $\vDash A$, koji se čita kao *A je istinito*, i $\sim A$, koji se čita kao *A je lažno*[14]. Navedeni spoljašnji veznici nazivaju se *spoljašnjim tvrđenjem* (*assertion*) i *poricanjem* (*denial*). Ostali spoljašnji veznici odgovaraju unutrašnjim veznicima ali primenjenim na spoljašnje iskaze. Pomoću njih može biti definisan i veznik koji stoji za besmislenost iskaza kao:

$$\downarrow A =_{def} \neg (\vDash A \vee \sim A) .$$

14 Kako bi bila konzistentna sa ostatkom rada, naša upotreba simbola za veznike se u nekoj meri razlikuje od Bočvarove.

Svi spoljašnji iskazi, uključujući one koji se odnose na besmislene unutrašnje iskaze, shvaćeni su kao smisleni. Taj pristup sledi intuiciju da ako je A besmislen iskaz, onda su oba iskaza *A je istinito* i *A je lažno* smislena i lažna, dok je iskaz *A je besmisleno* smislen i istinit. Logika spoljašnjih iskaza je klasična logika, što omogućava primenu klasičnih principa pri rasuđivanju o unutrašnjim iskazima od kojih su neki besmisleni. Razlikovanjem logike koja se odnosi na dva sistema iskaza, Bočvar uvodi u svoju teoriju razliku između objekt-jezika, koji čini sistem unutrašnjih iskaza, i metajezika, koji čini sistem spoljašnjih iskaza. Među formulama objekt jezika ne postoje tautologije i kontradikcije, jer je svaka od njih interpretirana kao besmislena za svaku valuaciju koja čini besmislenom neku od njenih potformula. Među spoljašnjim formulama, sa druge strane, tautologije i kontradikcije postoje. Posebno zanimljive među njima su one koje sadrže veznik $\downarrow$ i koje mogu biti korišćene u dokazu besmislenosti određenih unutrašnjih iskaza.

Na ovom računu iskaza zasnovan je račun iskaznih funkcija ili predikata bez tipova unutar kojeg su paradoksi analizirani i razrešeni dokazom besmislenosti iskaza koji su u njihovoj osnovi. Kako bi paradoksi bili formalizovani u ovom računu, domen individualnih promenljivih shvaćen je kao da uključuje unutrašnje funkcije i iskaze. To omogućava opisivanje samoprimenljivih funkcija i samoreferencijalnih rečenica objekt-jezika. Spoljašnje funkcije i iskazi su, sa druge strane, isključeni iz tog domena. Tu se, na vrlo bitan način, ispoljava pomenuta razlika između objekt i metajezika, što će doći do izražaja prilikom razmatranja paradoksa.

Time je obezbeđena formalna osnova za analizu paradoksa kao što su Raselov i Greling-Nelsonov paradoks. Ekvivalencije koje su u osnovama tih paradoksa, kao što je ona koja tvrdi da se Raselova funkcija, izražena predikatom neprimenljivosti na sebe, primenjuje na sebe ako i samo ako se ne primenjuje na sebe, nisu izvodive u unutrašnjem računu zahvaljujući njegovoj trovrednosnoj interpretaciji. Sa druge strane, pandan toj ekvivalenciji može biti formulisan u spoljašnjem računu. Kao deo spoljašnjeg računa, ta ekvivalencija zapravo predstavlja opis paradoksalne posledice kojoj primena Raselove funkcije na samu sebe vodi u okviru dvovrednosne logike. Ona tvrdi da bi rezultat primene te funkcije na samu sebe imao istu istinosnu vrednost kao njegova negacija. Taj iskaz je oblika $A \Leftrightarrow \neg A$[15]. On međutim *nije kontradikcija*. Jedino što sledi iz njega jeste to da iskaz koji tvrdi da se Raselova funkcija primenjuje na sebe nije ni istinit ni lažan. Argument koji vodi do tog zaključka može biti formalizovan unutar iste teorije pomoću aksiome $(A \Leftrightarrow \neg A) \Leftrightarrow\downarrow A$, koja implicira da ako bi neki unutrašnji iskaz trebalo da ima istu istinosnu vrednost kao i njegova negacija, onda je on besmislen. Pomoću druge aksiome $\downarrow A \Leftrightarrow\downarrow \neg A$, može se pokazati da je i negacija takvog iskaza takođe besmislena. Sa druge strane, spoljašnji iskaz koji tvrdi da je samoprimena Raselove funkcije besmislena je smislen i istinit. Njegova istinitost dokazana je prethodnim argumentom. Na sličan način se dokazuje besmislenost iskaza koji su u osnovi drugih paradoksa. Kao što smo videli u prethodnom delu rada, to će takođe biti iskazi ekvivalentni sopstvenim negacijama, pa će njihova besmislenost biti dokaziva pomoću istih aksioma.

Razlikovanjem objekt i metajezika i pripuštanjem unutrašnjih, ali ne i spoljašnjih funkcija i iskaza unutar domena individualnih promenljivih, Bočvar izbegava opasnost od određenih verzija paradoksa koje bi inače mogle biti formulisane unutar njegove teorije. Jedan takav paradoks ticao bi se iskazne funkcije koja se primenjuje na sve i samo one funkcije za koje je lažno da se one primenjuju na same sebe. Ta funkcija bi se primenjivala na sebe ako i samo ako bi bilo lažno da se primenjuje na sebe. Na osnovu toga bila bi izvodiva formula oblika $A \Leftrightarrow\sim A$ koja jeste kontradikcija jer formula desno od ekvivalencije mora biti ili istinita ili lažna. Ovaj paradoks je izbegnut zahvaljujući tome što opisana funkcija, pošto je spoljašnja, ne može biti sopstveni argument. Iz istog razloga ni paradoks *pojma pojma koji nije smisleno primenljiv na sebe* ne može biti formulisan unutar Bočvarove teorije. Spoljašnja funkcija koja bi stajala za *pojam*

15 Veznik $\Leftrightarrow$ stoji za spoljašnju ekvivalenciju kojom se tvrdi identitet istinosnih vrednosti dva iskaza. Vidi fusnotu 11.

pojma koji nije smisleno primenljiv na sebe ne može da bude argument iste te funkcije. Na taj način blokirano je izvođenje kontradikcije u obliku iskaza $A \Leftrightarrow\downarrow A$.

Bočvarov izbor da sve iskaze o smislenosti i besmislenosti drugih iskaza smatra istinitim ili lažnim, može biti objašnjen njegovim ciljem da zasnuje teoriju koja uključuje rešenje paradoksa dokazom besmislenosti iskaza u njihovoj osnovi, ali bez slabljenja njene logike. Bočvar je uspeo da pokaže da se klasične tautologije mogu utopiti u tautologije njegovog spoljašnjeg računa, pa time i da je slabljenje logike zaista izbegnuto (Bochvar, 1981, pp. 95-97). Njegov račun bi tako, što se logičke snage tiče, mogao da obezbedi solidnu osnovu teorije pojmova.

Ipak, Bočvarovo razlikovanje dva nivoa jezika - jednog u kojem je formalizovana primena iskaznih funkcija, i drugog u kojem je ona opisana, može predstavljati problem u vezi sa formalizacijom odgovarajuće aksiome komprehenzije u njegovoj teoriji. Ta aksioma, čini se, ne može biti formulisana kao unutrašnji iskaz jer bi besmislenost proizvoljne primene neke funkcije učinilo i nju besmislenom. Sa druge strane, kada bi bila formulisana kao spoljašnji iskaz, vodila bi istim paradoksima kojima vodi klasična formulacija te aksiome (kao što su paradoksi spomenuti u prethodnom odeljku).

Zapravo, čini se da konzistentna formulacija aksiome komprehenzije u teoriji koja se zasniva na klasičnoj logici zahteva mešanje dva jezička nivoa koja su strogo razdvojena u ovoj teoriji. To mešanje dva nivoa bilo bi omogućeno uvođenjem pojmova u objekt-jezik koji omogućavaju adekvatnu karakterizaciju onih primena predikata koje treba da odgovaraju smislenim primenama pojmova. Intuicija je da bismo na taj način mogli da dođemo do konzistentne aksiome komprehenzije koja bi povezivala tvrđenja o primenama pojma sa odgovarajućim primenama predikata koje su smislene i istinite. U nastavku ćemo videti da je formulisanje i konzistentno korišćenje takve aksiome komprehenzije moguće u odgovarajućim proširenjima klasične logike.

3.6 PARCIJALNI PREDIKATI U PROŠIRENJIMA KLASIČNE LOGIKE

Parcijalnost predikata teorije zasnovane na nekoj trovrednosnoj logici posledica je njene interpretacije prema kojoj neke rečenice nastale primenom tih predikata na određena imena nisu ni istinite ni lažne. Sve rečenice klasične logike su, sa druge strane, interpretirane kao istinite ili lažne, pa je potrebno naći neki drugi način za uključivanje parcijalnih predikata u teoriju zasnovanu na toj logici, i objasniti u čemu bi se njihova parcijalnost u tom kontekstu sadržala. Drugim rečima, potrebno je naći način za razlikovanje takvih primena predikata koje bi trebalo da su ekvivalentne opisima smislenih primena odgovarajućeg pojma, od onih kojima nikakva smislena primena pojma ne odgovara. Primene predikata te druge vrste predstavljale bi opise besmislenih primena pojma.

To bi se moglo postići uvođenjem nekog *pojma istine* u jezik klasične logike, koji se ne bi u potpunosti slagao sa pojmom istine teorije zasnovane na toj logici u modelu. Razlika bi bila u tome što bi prvi pojam istine bio parcijalan u smislu da pomoću njega ne bismo mogli da unutar teorije opišemo svaku primenu predikata kao istinitu ili lažnu. Za smislene primene pojmova bi tada stajale samo takve rečenice koje u teoriji mogu biti opisane kao istinite ili lažne. Parcijalnost pojma istine bi tako omogućila prepoznavanje parcijalnosti drugih predikata date teorije koja bi se sastojala u tome što njihovom primenom nije nužno dobijena rečenica koja je istinita ili lažna na osnovu pojma istine koji ta teorija koristi.

Ovaj pristup slaže se sa Bočvarovim u zauzimanju metaperspektive u odnosu na rečenice koje bi trebalo da odgovaraju primeni pojma, i u korišćenju pojma istine i smislenosti (zajedno sa njihovim suprotnostima) za karakterizaciju tih rečenica. Jedna bitna razlika jeste ta što su ti pojmovi ovde parcijalni. U nastavku ćemo ukratko opisati dve teorije koje se mogu shvatiti kao formalizacija ovog pristupa istraživanju smislene primenljivosti. Ideja u njihovoj osnovi vodi teoriji sa sličnim rezultatima koja se, međutim, ne služi nikakvim neklasičnim sredstvima.

Pokušaćemo da pokažemo do kraja ovog poglavlja da bi određene osobine te teorije mogle baš nju činiti adekvatnom osnovnom buduće teorije pojmova.

3.6.1 *Parcijalni predikat istinitosti*

Prva teorija koju ćemo razmotriti predstavlja proširenje klasične logike parcijalnim predikatom istinitosti i aksiomama koje ga karakterišu. Predikat koji stoji za smislenost (odnosno za posedovanje istinosne vrednosti), bio bi definisan pomoću tog predikata i takođe bi bio parcijalan. Parcijalnost ovih predikata ne bi zavisila od njihove interpretacije, pošto je ona *klasična*, već od postojanja takvih rečenica teorije formiranih njihovom primenom koje bi bile *neodlučive* u toj teoriji. U tom smislu bi i Gedelov predikat dokazivosti, tj. njegova negacija, mogao biti nazvan parcijalnim predikatom formalnih teorija kojima pripada na osnovu toga što postoje, kao što znamo, neodlučive rečenice tih teorija koje su formirane pomoću njega.

Slično tome, iako će sve rečenice formirane pomoću predikata istinitosti biti istinite ili lažne u modelu, aksiome koje karakterišu taj predikat neće nam nužno dozvoliti da to dokažemo za svaku takvu rečenicu. To znači, iz perspektive semantike te teorije, da bi se moglo desiti da neka rečenica, iako istinita u njenom modelu, ne pripada skupu rečenica kojim je interpretiran predikat istinitosti te teorije.

To je ideja u osnovi Kripkeovog „Nacrta teorije istine" (Kripke, 1975). Ona sledi intuiciju, suprotnu Bočvarovoj, da su rečenice kojima se tvrdi istinitost ili lažnost neke besmislene rečenice takođe besmislene. Primena predikata istinitosti na besmislenu rečenicu bi tako rezultirala rečenicom koja nije ni istinita ni lažna, zbog čega bi i taj predikat bio parcijalan. Kripke je u navedenom radu opisao model teorije koja sadrži takav parcijalni predikat istinitosti, dok je Feferman predložio njenu aksiomatizaciju u predavanjima održanim tokom 1979. godine (Feferman, 2011). Cilj te teorije je izbegavanje semantičkih paradoksa koji se tiču pojma istinitosti bez uvođenja jezičkih nivoa ili bilo kakvog umnožavanja pojma istine koje bi onemogućilo njegovu samoreferencijalnu primenu. Kripke je pokazao da razumevanje predikata istinitosti kao parcijalnog predikata omogućava zasnivanje konzistentne teorije u kojoj bi njegova primena na rečenice u kojima se isti taj predikat javlja bila dozvoljena.

Teorija je formulisana kao proširenje aritmetike predikatom istinitosti i aksiomama koje ga opisuju. Logika predložene teorije je klasična, ali je njen predikat istinitosti, zvaćemo ga T, neklasičan. Umesto klasičnog, Kripke predlaže *trovrednosni model* predikata T. U tom modelu, predikat T interpretiran je parom disjunktnih skupova (S_1, S_2) gde je S_1 neki skup rečenica istinitih u modelu koje čine *ekstenziju* tog predikata, dok je S_2 skup lažnih rečenica koje predstavljaju njegovu *antiekstenziju.* Antiekstenziju predikata T zvaćemo i ekstenzijom predikata F koji stoji za lažnost definisanu kao istinitost negacije date rečenice. Unija ta dva skupa ne iscrpljuje nužno skup svih rečenica teorije kojoj predikat T pripada. Rečenice koje ne pripadaju toj uniju su one za koje taj predikat nije definisan. Drugim rečima, interpretacija predikata F se neće nužno slagati sa interpretacijom predikata $\neg T$, u čemu se i sastoji neklasičnost predikata T.

Kripke je pokazao da za ovako interpretiran predikat istinitosti može važiti *T-shema* Tarskog prema kojoj je rečenica A istinita ako i samo ako je $T([A])$ istinito. Drugim rečima, Kripke je pokazao da će teorija koja sadrži takav predikat istinitosti biti konzistentna. Za njegov dokaz ključna je određena *teorema o fiksnoj tački* (zapravo transfinitna verzija teoreme čije su instance PRVA TEOREMA REKURZIJE i TEOREMA KNASTERA I TARSKOG). Ono što omogućava primenu te teoreme u ovom kontekstu jeste Kripkeova induktivna definicija predikata istinitosti, tj. njegov opis načina na koji je istinosna vrednost pripisana rečenicama aritmetike proširene predikatima T i F (Kripke, 1975, pp. 702-705; Burgess, 2013, pp. 165-170).

Ideja je da istinosna vrednost rečenica koje tvrde istinitost ili lažnost drugih rečenica iste teorije može biti određena tek nakon što je tim rečenicama pripisana istinosna vrednost. Iz tog razloga, Kripke opisuje interpretaciju rečenica teorije koja sadrži predikat T kao proces koji se odvija u fazama. Najpre je istinosna vrednost na standardan način pripisana aritmetičkim rečenicama koje ne sadrže predikate T i F i rečenicama izgrađenim od njih pomoću logičkih veznika. U toj fazi, predikat T je u potpunosti neinterpretiran. Rečenicama koje sadrže predikate T i F istinosna vrednost pripisana je u narednim fazama prema sledećim pravilima: ako je A istinito, onda je i $T([A])$ istinito, dok je $F([A])$ lažno; i ako je A lažno, onda je $T([A])$ lažno, a $F([A])$ istinito. U prvoj od tih faza istinosna vrednost pripisana je rečenicama koje tvrde istinitost ili lažnost aritmetičkih rečenica i rečenicama izgrađenim od njih pomoću veznika. Na taj način, predikat T dobija određenu interpretaciju (S_1, S_2). U drugoj fazi, istinosna vrednost pripisana je i rečenicama koje tvrde istinitost ili lažnost rečenica interpretiranih u prvoj fazi i rečenicama formiranim od njih pomoću veznika. Na taj način postignuta je nova interpretacija predikata T u obliku para skupova (S_1', S_2'). Ovaj proces se nastavlja na isti način tako da su u svakoj sledećoj fazi interpretirane one rečenice koje tvrde istinitost ili lažnost rečenica interpretiranih u prethodnoj fazi. Time je i interpretacija predikata T proširena svim rečenicama koje su prethodno interpretirane kao istinite ili lažne. Prelazak sa jedne na drugu interpretaciju predikata T može biti predstavljen funkcijom ϕ, koja paru skupova (S_1, S_2) pripisuje par skupova (S_1', S_2'). Pošto bi tako dobijeni skupovi S_1' i S_2' trebalo da proširuju interpretaciju predikata T svim onim rečenicama kojima je u prethodnoj fazi interpretacije pripisana istinosna vrednost, trebalo bi da važi da je $S_1 \subseteq S_1'$ i $S_2 \subseteq S_2'$. To znači da se primenom funkcije ϕ čuvaju već ustanovljene istine i laži i potencijalno im se dodaju neke nove. To svojstvo čini funkciju ϕ *monotonom.*

Na opisan način, tj. postepenim proširivanjem interpretacije predikata T, gradi se hijerarhija nivoa koji slede jedan za drugim. Nivoi te hijerarhije mogu biti indeksirani ordinalnim brojevima. Konačni nivoi čije je građenje do sada opisano, bili bi indeksirani konačnim ordinalima: početni nivo u kojem je predikat T neinterpretiran bio bi indeksiran ordinalom 0, a dalji nivoi redom ordinalima $1, 2, 3, ...$

Nijedan od do sada opisanih nivoa hijerarhije ne može, međutim, sadržati istinitu rečenicu koja bi proglasila sve rečenice te hijerarhije istinitim. Kako bi objasnio istinitost te i sličnih univerzalnih rečenica, Kripke uvodi transfinitni nivo koji sledi nakon svih konačnih nivoa hijerarhije. Taj nivo indeksiran je prvim graničnim ordinalom ω koji dolazi nakon svih konačnih ordinala. Na tom prvom transfinitnom nivou hijerarhije, interpretacija predikata T dobijena je unijom svih interpretacija koje su tom predikatu pripisane na konačnim nivoima. Rečenice koje tvrde istinitost ili lažnost svih rečenica konačnih nivoa hijerarhije mogu biti interpretirane kao istinite ili lažne na ovom transfinitnom nivou jer je svakoj od tih rečenica već pripisana istinosna vrednost na nekom nižem nivou.

Kako bismo one rečenice koje dobijaju istinosnu vrednost na prvom transfinitnom nivou proglasili istinitim ili lažnim, moramo da se popnemo na sledeći nivo koji bi bio indeksiran sledbenikom prvog graničnog ordinala - $\omega + 1$. Građenje hijerarhije nastavili bismo nivoima indeksiranim ordinalima $\omega + 2$, $\omega + 3$, itd., koristeći za njihovo građenje funkciju ϕ na način na koji smo je koristili u građenju konačnih nivoa. Nakon svih tih nivoa, dolazimo do sledećeg graničnog nivoa indeksiranog graničnim ordinalom $\omega + \omega$, koji je napravljen na isti način kao i prvi transfinitni nivo - unijom prethodno dobijenih interpretacija predikata T. Građenje hijerarhije se nastavlja formiranjem nivoa koji slede jedan za drugim korišćenjem funkcije ϕ, a graničnih nivoa uniranjem svih interpretacija dobijenih na prethodnim nivoima hijerarhije.

Navedeni Kripkeov opis induktivnog građenja interpretacije predikata T obezbeđuje odgovarajuće uslove za primenu teoreme iz koje sledi da funkcija ϕ korišćena u tom građenju ima fiksnu tačku. Mogućnost primene te teoreme suštinski zavisi od monotonosti funkcije ϕ, odnosno činjenice da se njenom primenom proširuje interpretacija predikata T. Ta teorema garantuje da ćemo iteracijom funkcije ϕ dovoljan broj puta, i uniranjem njenih vrednosti na graničnim nivoima, u nekom trenutku doći do fiksne tačke te funkcije, odnosno, do interpretacije predikata T koju čini par skupova (S_{0_1}, S_{0_2}) za koji važi: $\phi((S_{0_1}, S_{0_2})) = (S_{0_1}, S_{0_2})$. Iz toga sledi da će postojati neki nivo opisane hijerarhije na kom neće biti moguće doći do novih istina primenom predikata T. To će biti nivo na kojem je $T([A])$ istinito ako i samo ako je A istinito, odnosno, na kojem važi T-shema. Teorema o fiksnoj tački koja omogućava ovaj rezultat takođe garantuje da će fiksna tačka dostignuta na opisani način biti najmanja fiksna tačka funkcije ϕ, što znači da će sve druge fiksne tačke kompatibilne sa njom pružati širu interpretaciju predikata T.

Intuitivan razlog zbog kog je fiksna tačka funkcije ϕ dobijena na opisan način *minimalna* je taj što proces koji do nje vodi ne uključuje pripisivanje istinosne vrednosti rečenicama formiranim pomoću T i F na proizvoljan način, odnosno na način koji se ne može podvesti pod navedena pravila. Rečenice kojima u tom procesu nije pripisana istinosna vrednost nazivaju se *nezasnovanim* (*ungrounded*) rečenicama. Za te rečenice se ispostavlja da se istinosna vrednost rečenica od kojih njihova istinosna vrednost zavisi, ne može nezavisno odrediti. Takva će biti svaka rečenica koja samoj sebi pripisuje neki od predikata T i F ili, u opštem slučaju, svaka rečenica koja je deo nekog beskonačnog niza rečenica takvih da istinosna vrednost svake rečenice u tom nizu zavisi od istinosne vrednosti sledeće rečenice. Nekim nezasnovanim rečenicama, kao što je 'Ova rečenica je istinita' može na konzistentan način biti pripisana proizvoljna istinosna vrednost. Time bismo proširili minimalnu fiksnu tačku koja predstavlja interpretaciju predikata T. Sa druge strane, paradoksalnim rečenicama, kao što je ona koja za sebe kaže da je lažna ili neistinita, nijedna istinosna vrednost ne može biti pripisana. Da jeste, iz toga bi sledila kontradikcija samo na osnovu pretpostavljenih pravila interpretacije. Iako vidimo da takve rečenice neće biti ni istinite ni lažne u minimalnoj ili bilo kojoj drugoj fiksnoj tački, rečenica koja to tvrdi takođe neće biti istinita (niti lažna) u toj fiksnoj tački. To je takođe posledica parcijalnosti predikata T.

Što se tiče aksiomatizacije opisane teorije, do nje se dolazi proširenjem formalnog jezika aritmetike predikatima T i F (pri čemu je F definisano na sledeći način: $F([A]) =_{def} T([\neg A])$), zajedno sa aksiomama koje ih opisuju. Te aksiome daju rekurzivnu definiciju predikata T i F, odnosno, uslove za njihovu primenu koji odgovaraju kumulativnom građenju njihove interpretacije. Među njima bi bile sledeće aksiome:

- $T([T([A])])$ ako i samo ako $T([A])$, i $F([T([A])])$ ako i samo ako $F([A])$;
- $T([\neg A])$ ako i samo ako $F([A])$, i $F([\neg A])$ ako i samo ako $T([A])$;
- $T([A \vee B])$ ako i samo ako $T([A])$ ili $T([B])$, i $F([A \vee B])$ ako i samo ako $F([A])$ i $F([B])$.

Ako su aksiome pažljivo odabrane, tako da odgovaraju Kripkeovom modelsko-teorijskom opisu teorije, dobijena aksiomatska teorija bi trebalo da bude konzistentna. Predikat za *smislenost*, ukoliko stoji za posedovanje istinosne vrednosti, može biti definisan na sledeći način:

$$S(x) =_{def} T(x) \vee F(x).$$

Parcijalnost predikata istinitosti i njegova razlika u odnosu na istinu u modelu implicira da iako za svaku rečenicu A, zakon isključenja trećeg $A \vee \neg A$ jeste istinit u modelu teorije,

$$T([A]) \vee F([A]), \text{ odnosno, } S([A])$$

u toj teoriji ne mora važiti, kao što na primer ne važi za rečenicu koja za sebe kaže da je lažna. Za sve rečenice za koje to važi, takođe će važiti da su one istinite prema teoriji ako i samo ako su istinite u njenom modelu, odnosno, važiće da:

$$S([A]) \rightarrow (T([A]) \leftrightarrow A).$$

Shema Tarskog tako može biti konzistentno instancirana svakom rečenicom teorije koja je dokazano istinita ili lažna, uključujući i one rečenice koje i same sadrže predikat T.

U literaturi se mogu naći različite verzije *aksiomatske teorije istine*. Ono što čini tu teoriju privlačnom jeste to što ona uspeva da se konzistentno bavi pojmom istinitosti ne poričući njegovu jedinstvenost i samoreferencijalnu primenljivost, koja u nekim slučajevima omogućava formulisanje paradoksa. Aksiomatske teorije izbegavaju te paradokse na taj način što izdvajaju neke nesumnjive karakteristike pojma istine i uzimaju ih za njegove definišuće karakteristike. Te karakteristike pojma istine tipično ga čine *parcijalnim* pojmom, tj. pojmom koji nije primenljiv na svaku rečenicu ili njenu negaciju. Time se postiže rešavanje paradoksa koje ne zahteva zamenu intuitivnog pojma istine nekim drugim, manje intuitivnim pojmom, ili različitim pojmovima istine primenljivim na različite rečenice. Na taj način, ove teorije postižu za pojam istine upravo ono što nastojimo da postignemo za pojam pojma: da izbegnemo paradokse koji se tiču tog pojma a da ga pritom ne izmenimo u velikoj meri, time što bismo, na primer, zabranili njegovu samoreferencijalnu primenu.

Jedna odlika klasične teorije koja aksiomatizuje parcijalni predikat istinitosti može se, međutim, smatrati problematičnom. To je činjenica da istina u modelu te teorije ne odgovara istini opisanoj predikatom istinitosti koji ona koristi. U modelsko-teorijskom smislu, svi predikati Kripkeove teorije su totalni - oni su interpretirani skupovima kojima neki objekat iz domena ili pripada ili ne pripada. Isto važi i za predikat istinitosti: za svaku rečenicu teorije važiće da ona (tj. njen kôd) pripada skupu kojim je taj predikat interpretiran ili njegovom komplementu, iako ta činjenica ne mora nužno biti prepoznata u teoriji. Takva interpretacija može se kritikovati sa stanovišta da ako je pojam istinitosti zaista parcijalan, onda bi njegova parcijalnost trebalo da se pokaže i u modelu.

U cilju izbegavanja ovog razilaženja dva pojma istine, Rajnhard predlaže reinterpretaciju Kripkeove teorije kojom bi parcijalnost predikata istinitosti i smislenosti trebalo da bude vernije predstavljena (William N. Reinhardt, 1986). Ideja je da rečenice teorije ne treba da budu interpretirane kao istinite samo na osnovu njihove dokazivosti u toj teoriji. Umesto toga, da bi bile interpretirane kao istinite ili uopšte smislene, njihova istinitost ili smislenost treba takođe da budu dokazive. Ti stroži uslovi interpretacije čine istinitost u modelu kompatibilnom sa istinitošću prema teoriji. Parcijalnost istinitosti bi na taj način trebalo da bude predstavljiva u modelu teorije koji, u standardnoj interpretaciji, sve njene predikate predstavlja kao klasične. Iz toga sledi da mogu postojati rečenice dokazive u teoriji koje nisu interpretirane kao istinite ili uopšte smislene (sve one za koje u teoriji nije dokazivo da su takve). Rajnhard predlaže formalističku interpretaciju takvih rečenica prema kojoj bi one bile samo sredstva za dolaženje do istina u datoj teoriji. Ta formalistička interpretacija je po duhu slična Hilbertovoj interpretaciji infinitarnih iskaza i metoda u matematici prema kojoj oni nisu iskazi koji opisuju matematičke činjenice, ali mogu olakšati dolaženje do takvih iskaza i njihovo dokazivanje (Reinhardt, 1986, pp. 224-225). Jedna vrsta rečenica koje bi imale taj status su i one koje tvrde da neka rečenica nije ni istinita ni lažna. Iako u teoriji možemo dokazati rečenicu koja to tvrdi, na primer, o nekoj paradoksalnoj rečenici, istinitost te rečenice neće biti dokaziva u teoriji.

Parcijalni predikat istinitosti opisan ovom aksiomatskom teorijom može biti korišćen u analizi smislene primenljivosti pojma. Kako je već rečeno, jedan način da se omogući analiza smislene primenljivosti u klasičnoj teoriji čiji su predikati interpretirani kao totalni jeste da se tvrđenja o smislenim primenama odgovarajućeg pojma pridruže samo takvim primenama predikata koje mogu biti opisane u teoriji kao istinite ili lažne. Sledeće ekvivalencije bi tada izražavale vezu između primene nekog pojma i istinitosti ili lažnosti odgovarajuće rečenice: pojam c primenjuje se na x ako i samo ako $T\left([C\left(x\right)]\right)$ važi, gde je C predikat koji izražava pojam c; isti pojam se ne primenjuje na x (ali jeste smisleno primenljiv na x) ako i samo ako $F\left([C\left(x\right)]\right)$ važi; dok je pojam c smisleno primenljiv na x ako i samo ako važi ili $T([C(x)])$ ili $F([C(x)])$, u kom slučaju će i $S([C(x)])$ važiti. Aksioma komprehenzije može tada biti izražena formulom:

$$\exists x \forall y\,(S([C(y)]) \rightarrow (x \boldsymbol{\eta} y \leftrightarrow C(y)))$$

(cf. Reinhardt, 1986, p. 231). Kao što za neku rečenicu teorije koja nije istinita ne važi nužno da je lažna, tako ni za objekat koji ne potpada pod neki pojam neće nužno važiti da potpada pod komplement tog pojma. Komplement nekog pojma bio bi pojam koji se primenjuje na neki objekat ako i samo ako se pojam čiji je on komplement ne primenjuje na taj objekat, ali jeste smisleno primenljiv na njega. Rečenice koje nisu smislene ne opisuju niti primenu nekog pojma niti primenu njegovog komplementa. Umesto toga, one mogu biti shvaćene kao da referiraju na objekat na koji dati pojam nije smisleno primenljiv.

Glavni parcijalni predikati ove teorije - predikat istinitosti i smislenosti, kada su primenjeni na neku rečenicu izgrađenu pomoću određenog predikata, mogu se shvatiti tako da tvrde, redom, da se pojam izražen tim predikatom primenjuje na neki objekat čime se formira istinit iskaz, i da se taj pojam smisleno primenjuje na neki objekat pa je iskaz formiran tom primenom smislen (istinit ili lažan). Tvrđenje da se neki pojam primenjuje na određeni objekat bilo bi tada ekvivalentno tvrđenju da je formula koja izražava taj pojam istinita za ime tog objekta, dok bi tvrđenje da se neki pojam smisleno primenjuje na objekat bilo ekvivalentno tvrđenju da je odgovarajuća formula istinita ili lažna za njegovo ime. Parcijalnost predikata istinitosti i smislenosti rečenica bi tada odgovarala rezultatu Gedelove analize intenzionalnih paradoksa prema kojoj *pojam primene pojma* i *pojam smislene primenljivosti pojma* nisu svuda definisani, odnosno, smisleno primenljivi.

Analiza smislene primenljivosti do koje se na taj način dolazi ističe blisku vezu između *pojma istine* i *pojma primene pojma*, koja izgleda omogućava korišćenje prvog pojma u analizi drugog. Njihova veza je još izraženija u Fefermanovoj teoriji koju ćemo sledeću predstaviti, a u kojoj su dva predikata koja stoje za te pojmove shvaćena kao posebne instance iste predikatske sheme koja izražava zadovoljivost neke formule uređenom n-torkom terama (za $n \geq 0$). Ponovo će se ispostaviti da razumevanje ovog predikata kao parcijalnog omogućava konzistentno korišćenje neograničenih principa koji se tiču istinitosti i primene pojma.

3.6.2 *Parcijalni predikat zadovoljivosti*

Prema Fefermanu, glavni cilj zasnivanja teorije bez tipova koja se bavi relacijom pripadnosti klasi (tj. potpadanja pod pojam) jeste da se obezbedi okruženje, adekvatnije od *ZFC* i od teorije tipova, za analizu nekih intenzionalnih objekata, kao što je kategorija svih grupa, kategorija svih kategorija, kategorija svih funktora između neke dve kategorije, itd. (Feferman, 1984, p. 79). Adekvatna teorija koja se bavi tim objektima treba da dozvoli njihovu pripadnosti samima sebi u slučaju da ti objekti imaju istu strukturu kao i njihovi elementi. Zbog toga je bitno da ona ne sadrži tipove, niti eksplicitno - kao teorija tipova, niti implicitno - kao *ZFC*.

Feferman formuliše teoriju koja predstavlja proširenje klasične logike parcijalnim predikatima T_m, za $m \geq 0$, čije posebne instance predstavljaju predikat istinitosti i pripadnosti klasi (Feferman, 1984). Razumevanje tih predikata kao parcijalnih trebalo bi da omogući konzistentno korišćenje neograničenih principa koji karakterišu te predikate - T-sheme i aksiome komprehenzije. Zahvaljujući tome bi i samoreferencijalna primena tih predikata takođe bila formalizovana u toj teoriji[16].

Parcijalni predikat je u ovoj teoriji izjednačen sa parom predikata koji određuju njegovu ekstenziju i antiekstenziju. Objekti za koje nijedan predikat iz tog para nije istinit su oni za koje taj parcijalni predikat nije definisan. Uniforman način konstruisanja predikata koji stoje za

16 Za pregled teorija zasnovanih na sličnim idejama vidi (Cantini, 1996).

ekstenziju i antiekstenziju nekog parcijalnog predikata omogućen je uvođenjem veznika *neklasične ekvivalencije*. Ta ekvivalencija bi trebalo da bude *intenzionalna* u smislu da nije određena samo na osnovu istinosnih vrednosti rečenica koje povezuje. Intenzionalnost te ekvivalencije se pokazuje u činjenici da se njena istinitost ne čuva nužno supstitucijom formula koje u njoj stoje njima ekstenzionalno ekvivalentnim formulama.

Klasična osnova teorije je, u skladu sa tom idejom, proširena simbolom za intenzionalnu ekvivalenciju $\equiv$ i aksiomama koje je karakterišu (Feferman, 1984, pp. 96-101). Korišćenjem te ekvivalencije unutar iste teorije mogu biti formulisana tvrđenja o tome da je neka njena formula A istinita ili da je lažna, i to pomoću, redom, formula $A \equiv t$ i $A \equiv f$, gde je t neka istinita, a f neka lažna rečenica teorije. Svojstvo formule A da je *određena* (*determinate*) definisano je na sledeći način:

$$D([A]) =_{def} ((A \equiv t) \vee (A \equiv f)).$$

Među aksiomama koje karakterišu veznik $\equiv$ je i sledeći par aksioma koje određuju pod kojim uslovima formula izgrađena pomoću tog veznika može biti proglašena istinitom ili lažnom unutar teorije:

(I) $((A \equiv B) \equiv t) \leftrightarrow (A \equiv B)$
i
$((A \equiv B) \equiv f) \leftrightarrow (D([A]) \wedge D([B]) \wedge \neg (A \equiv B))$.

Navedeni par aksioma sugeriše da intenzionalna ekvivalencija može biti shvaćena kao Lukašievičeva ekvivalencija koja je istinita ako i samo ako rečenice koje povezuje imaju istu vrednost, uključujući i slučaj kada nijedna nije ni istinita ni lažna. Umesto tog para aksioma može se takođe pretpostaviti i onaj koji bi više odgovarao Klinijevoj ekvivalenciji koja je istinita samo ako su obe rečenice koje povezuje istinite ili obe lažne:

(II) $((A \equiv B) \equiv t) \leftrightarrow (D([A]) \wedge D([B]) \wedge (A \equiv B))$
i
$((A \equiv B) \equiv f) \leftrightarrow (D([A]) \wedge D([B]) \wedge \neg (A \equiv B))$.

Aksiome koje karakterišu intenzionalnu ekvivalenciju ostavljaju mogućnost da neka rečenica A bude neodređena, odnosno, da za nju ne važi ni $A \equiv t$, ni $A \equiv f$. Par predikata koji određuje ekstenziju i antiekstenziju nekog parcijalnog predikata $P(x)$ može tada biti definisan pomoću, redom, predikata $P(x) \equiv t$ i predikata $P(x) \equiv f$. Pošto se može desiti da za neko x nijedan od ta dva predikata ne važi, predikat $P(x)$ neće nužno biti definisan za sve objekte datog domena.

Pretpostavljeno je da postoji uniforman način na koji se imena formula i njihovih apstrakata (klasa i relacija) mogu dobiti od tih formula. Ako formula A sadrži $x_1, ..., x_m, y_1, ...y_n$ kao slobodne promenljive, onda je vezivanjem njenih promenljivih $x_1, ..., x_m$ i tretiranjem $y_1, ..., y_n$ kao parametara, formirano određeno ime. Ako je $m = 0$, ime dobijeno na taj način biće Gedelov broj te formule zajedno sa njenim slobodnim promenljivima: $([A], y_1, ..., y_n)$. Ako je $m = 1$, dobija se ime klase ili pojma $\{x : A(x, y_1, ..., y_n)\}$, a ako je $m > 1$, dobija se ime m-arne relacije $A(\hat{x}_1, ..., \hat{x}_m, y_1, ..., y_n)$ gde $\hat{x}_i$, za $1 \leq i \leq m$, označava vezanu promenljivu. Principi koji karakterišu tako dobijena imena, predstavljaju instance sheme aksioma formulisane pomoću $(m+1)$-arnog predikata T_m, takvog da se $T_m(x_1, ..., x_m, z)$ može čitati kao: uređena m-torka terama $(x_1, ..., x_m)$ zadovoljava z. Opšta shema aksioma tiče se primene tog predikata na neko ime formirano na gore opisan način koje se javlja na mestu promenljive z:

$$T_m(x_1, ..., x_m, A(\hat{u}_1, ..., \hat{u}_m, y_1, ..., y_n)) \equiv A(x_1, ..., x_m, y_1, ...y_n).$$

T-shema i aksioma komprehenzije su sledeće instance ove sheme:

ZA $m = 0$: $T_0([A], y_1, ..., y_n) \equiv A(y_1, ...y_n)$

ZA $m = 1$: $T_1(x, \{u : A(u, y_1, ...y_n)\}) \equiv A(x, y_1, ...y_n)$.

Formulisanjem aksioma pomoću intenzionalne umesto klasične, ekstenzionalne ekvivalencije, ostavlja se mogućnost da T_m-predikati koji se u njima javljaju budu parcijalni. U slučaju predikata pripadnosti klasi važi da ako je $A(x, y_1, ...y_n)$ karakterisano u teoriji kao istinito, odnosno, ako važi $A(x, y_1, ...y_n) \equiv t$, onda se uzima da x pripada klasi $\{u : A(u, y_1, ...y_n)\}$, dok ako važi $A(x, y_1, ...y_n) \equiv f$, onda se uzima da x ne pripada istoj klasi. U ostalim slučajevima, kada nijedna od dve ekvivalencije ne važi za A (što može biti slučaj posebno ako i samo A sadrži neki T_m-predikat), ne može biti određeno da li x pripada datoj klasi ili ne. U tom slučaju, $T_1(x, \{u : A(u, y_1, ...y_n)\})$ neće biti intenzionalno ekvivalentno ni rečenici t ni rečenici f (inače bi isto važilo i za A pošto je data ekvivalencija tranzitivna). Drugim rečima, $T_1(x, \{u : A(u, y_1, ...y_n)\})$ neće u tom slučaju biti određeno. To implicira da za neku klasu c, može postojati objekat x, za koji ne važi ni $T_1(x, c) \equiv t$, ni $T_1(x, c) \equiv f$. Za taj objekat bi važilo da ne pripada ni klasi c ni njenom komplementu. Drugim rečima, važilo bi da pojam čiju ekstenziju predstavlja ta klasa nije smisleno primenljiv na taj objekat.

Aksel i Feferman su pokazali da je teorija koja koristi datu aksiomu komprehenzije konzervativno proširenje osnovne teorije čija je konzistentnost pretpostavljena (cf. Feferman, 1984, pp. 97-98). Primenom sheme T_1 na klasu R svih klasa koje nisu sopstveni elementi, ekvivalencija u osnovi RASELOVOG PARADOKSA dobija se u sledećem obliku: $T_1(R, R) \equiv \neg T_1(R, R)$. Pravila ove teorije ne dopuštaju nam da iz bilo koje od formula koje stoje u intenzionalnoj ekvivalenciji izvedemo drugu formulu, zbog čega standardno izvođenje kontradikcije iz te ekvivalencije nije moguće. Ta ekvivalencija tako *ne vodi protivrečnosti*, iako pokazuje da postoji fiksna tačka negacije u odnosu na intenzionalnu ekvivalenciju. Zajedno sa ostalim aksiomama koje karakterišu veznik $\equiv$, ona vodi zaključku da rečenica $T_1(R, R)$ nije određena. Zaključak će biti isti koji god od parova aksioma (I) ili (II) da izaberemo. Jedina razlika je u tome što ako je to par aksioma pod (I), dobijena ekvivalencija smatra se istinitom, odnosno, intenzionalno ekvivalentnom rečenici t, što neće biti slučaj ako je taj par aksioma zamenjen parom aksioma pod (II).

Uslovi za pripadnost nekoj klasi, odnosno za primenu nekog pojma, formulisani u ovoj teoriji, veoma su slični onim uslovima do kojih se dolazi pomoću parcijalnog predikata istinitosti (samo pišemo $T([A])$ umesto $A \equiv t$; $F([A])$ umesto $A \equiv f$; i $S([A])$ umesto $D([A])$). Ta dva pristupa predstavljaju odgovor na isti zahtev: *naći način za razlikovanje rečenica koje mogu biti konzistentno smatrane istinitim ili lažnim u prisustvu neograničene aksiome komprehenzije, od onih čija karakterizacija kao istinitih ili lažnih u tom kontekstu vodi kontradikciji.* Jedino su one supstitucione instance formule $C(x)$ koje mogu biti opisane u teoriji kao istinite ili lažne, shvaćene tako da odgovaraju primeni pojma c, odnosno, supstitucionim instancama formule $c\boldsymbol{\eta}x$. One koje ne mogu biti tako opisane u teoriji mogu se shvatiti kao da izražavaju besmislenu primenu pojma c. Kao što smo videli, u oba slučaja je klasična osnova teorije proširena određenim sredstvima koja bi trebalo da omoguće takvu karakterizaciju: predikatom istinitosti koji ne bi bio pripisan svakoj formuli ili njenoj negaciji, ili intenzionalnom ekvivalencijom koja ne bi povezivala svaku formulu sa istinitom ili lažnom rečenicom teorije. Aksiome koje opisuju parcijalni predikat istinitosti i intenzionalnu ekvivalenciju garantuju da nijedna rečenica neće biti dovedena u vezu sa tvrđenjem o primeni pojma ako bi to vodilo kontradikciji.

Ta sredstva, međutim, ne bi bila potrebna ako bi postojao neki drugi način da identifikujemo one instance neke formule koje ne mogu biti konzistentno smatrane istinitim ili lažnim u prisustvu neograničene komprehenzije i T-sheme, i onemogućimo instanciranje tih principa samo tim instancama formula. Na taj način bismo mogli da dođemo do istih rezultata bez uvođenja bilo kakvih neklasičnih elemenata u raspravu. Istovremeno može biti rasvetljeno i koje to rečenice postaju problematične u kontekstu neograničene komprehenzije. U nastavku ćemo videti jedan od načina na koje to može biti postignuto.

3.6.3 *Parcijalnost i pozitivne formule*

Parcijalnost predikata istinitosti i pripadnosti klasi (tj. potpadanja pod pojam) pokazuje se u tome što nije svaka rečenica klasične teorije shvaćena kao da određuje njihovu istinitu ili lažnu primenu. Prethodno opisane teorije nude način na koji tako shvaćena parcijalnost predikata može biti prepoznata u klasičnoj teoriji. Prepoznavanje parcijalnosti tih predikata omogućava izbegavanje paradoksa kojima vode neke rečenice u kojima se ti predikati ili njihove negacije javljaju. Na osnovu poznatih paradoksa znamo da su takve rečenice one koje *negiraju primenu određenog pojma (kao što je pojam istine), ali ne mogu biti shvaćene kao da time tvrde primenu njemu komplementarnog pojma (kao što je pojam laži)*. Kada bismo našli način da izbegnemo instanciranje bazičnih principa upravo takvim rečenicama, mogli bismo da izbegnemo paradokse ne izlazeći van okvira klasične logike.

Na taj način se može opravdati sledeća verzija Fefermanove teorije predikata T_m (Feferman, 1984, pp. 101-105). U toj teoriji, predikat T_m zamenjen je parom disjunktnih predikata $\left(T_m, \bar{T}_m\right)$ koji stoje za njegovu ekstenziju i antiekstenziju. Svaki predikat iz tog para karakteriše posebna aksioma koja izražava uslove za njegovu primenu na ime formule ili njenog apstrakta (koji stoji za odgovarajuću klasu ili relaciju) i odgovarajući broj imena. Zbog parcijalnosti predikata T_m , uslovi za njegovu primenu ne mogu, kao što je standardno, biti zadati pomoći neke formule A, dok su uslovi za primenu njegovog komplementa, predikata $\bar{T}_m$, zadati pomoću njene negacije $\neg A$. Razlog je to što će u klasičnoj logici uvek jedna od te dve formule biti istinita. Feferman predlaže da, umesto toga, uslovi za primenu predikata T_m i $\bar{T}_m$ budu zadati pomoću *pozitivnih aproksimacija* instanci formule A i njene negacije. U pozitivnoj formuli, negacija se javlja samo ispred njenih atomskih potformula na jeziku osnovne teorije, koja je najčešće aritmetika. Za svaku formulu i njenu negaciju, formule koje predstavljaju njihove pozitivne aproksimacije mogu biti nađene pomoću nekih dobro poznatih logičkih zakona (kao što su De Morganovi zakoni, zakon dvostruke negacije, i slični). U tom procesu, *negacije predikata T_m i $\bar{T}_m$ zamenjene su, redom, predikatima $\bar{T}_m$ i T_m.*

Uzmimo za primer formulu $\neg\left(Px \vee \neg Qx\right) \wedge \neg T_1\left(x, \{y : By\}\right)$. Ako su P i Q predikati osnovne teorije, onda je njena pozitivna aproksimacija formula $\left(\neg Px \wedge Qx\right) \wedge \bar{T}_1\left(x, \{y : By\}\right)$, dok je pozitivna aproksimacija njene negacije formula $\left(Px \vee \neg Qx\right) \vee T_1\left(x, \{y : By\}\right)$.

Pozitivna aproksimacija formule $A\left(x_1, ..., x_m, y_1, ..., y_n\right)$ označava se sa $A^+\left(x_1, ..., x_m, y_1, ..., y_n\right)$, dok se pozitivna aproksimacija njene negacije označava sa $A^-\left(x_1, ..., x_m, y_1, ..., y_n\right)$. Sheme aksioma koje karakterišu parcijalne predikate T_m, formulisane pomoću klasične ekvivalencije, su sledeće:

ZA T_m: $T_m\left(x_1, ..., x_m, A\left(\hat{u}_1, ..., \hat{u}_m, y_1, ..., y_n\right)\right) \leftrightarrow A^+\left(x_1, ..., x_m, y_1, ..., y_n\right)$

ZA $\bar{T}_m$: $\bar{T}_m\left(x_1, ..., x_m, A\left(\hat{u}_1, ..., \hat{u}_m, y_1, ..., y_n\right)\right) \leftrightarrow A^-\left(x_1, ..., x_m, y_1, ..., y_n\right)$.

T-shema i aksioma komprehenzije dobijaju se kao posebne instance ovih aksioma:

ZA $m = 0$: $$\begin{array}{l} T_0\left([A], y_1, ..., y_n\right) \leftrightarrow A^+\left(y_1, ...y_n\right) \\ \bar{T}_0\left([A], y_1, ..., y_n\right) \leftrightarrow A^-\left(y_1, ...y_n\right) \end{array}$$

ZA $m = 1$: $$\begin{array}{l} T_1\left(x, \{u : A\left(u, y_1, ...y_n\right)\}\right) \leftrightarrow A^+\left(x, y_1, ...y_n\right) \\ \bar{T}_1\left(x, \{u : A\left(u, y_1, ...y_n\right)\}\right) \leftrightarrow A^-\left(x, y_1, ...y_n\right). \end{array}$$

Nije nužno da se neka formula smatra ekvivalentnom svojoj pozitivnoj aproksimaciji. Razlog za to jeste zamena predikata $\neg T_m$ predikatom $\bar{T}_m$, ili predikata $\neg\bar{T}_m$ predikatom T_m u procesu dolaženja do te pozitivne aproksimacije. To je, dakle, slučaj samo ako neki od negiranih T_m predikata koji se javljaju u formuli ne odgovara predikatu sa kojim je zamenjen. U tom slučaju, izvorna formula nije shvaćena kao da implicira svoju pozitivnu aproksimaciju. Dakle, rečenice koje nisu ekvivalentne svojim pozitivnim aproksimacijama, i koje zbog toga ne mogu biti korišćene za instanciranje shema aksioma, jesu one koje negiraju primenu predikata T_m ili $\bar{T}_m$, ali time ne tvrde primenu, redom, predikata $\bar{T}_m$ i predikata T_m. Intuitivno, takve rečenice bi stajale za primenu T_m predikata na objekte za koje oni nisu definisani.

Za $m = 1$, dve aksiome daju uslove za pripadnost klasi koja odgovara formuli A, i njenom komplementu (koji je definisan pomoću formule A^-, a ne formule $\neg A$). Pripadnost klasi i njenom komplementu zavisiće tako od istinitosti određenih pozitivnih formula teorije. Pošto formule A^+ i A^- nisu nužno ekvivalentne formulama A i $\neg A$, može se desiti da nijedna od njih ne bude istinita za određenu supstituciju nekog terma mesto njihovih slobodnih promenljivih. Objekat označen tim termom neće, prema tome, pripadati ni datoj klasi ni njenom komplementu.

Formulacija uslova za istinitost i pripadnost klasi pomoću pozitivnih formula blokira izvođenje kontradikcije. Naime, pošto te formule ne sadrže negacije rečenica koje tvrde istinitost ili pripadnost klasi, aksiome koje karakterišu te predikate neće omogućiti izvođenje ekvivalencija koje opisuju fiksnu tačku negacije. Umesto toga, biće izvodiva samo ekvivalencija oblika $T_1(R, R) \leftrightarrow \bar{T}_1(R, R)$ koja ne vodi kontradikciji. Ta ekvivalencija uspeva da na konzistentan način izrazi našu intuiciju da postoji klasa za koju važi da je tvrđenje da ona pripada samoj sebi ekvivalentno tvrđenju da ona pripada sopstvenom komplementu (takva klasa bila bi ekstenzija *pojma neprimenljivosti na sebe*). Slično tome, što se tiče rečenice u osnovi PARADOKSA LAŽLJIVCA, jedino će nekontradiktorna ekvivalencija $T_0([L]) \leftrightarrow \bar{T}_0([L])$ biti izvodiva, koja tvrdi da je ta rečenica istinita ako i samo ako je lažna. Tako bi opet intuicija da je data rečenica u jednom smislu istinita, a u drugom lažna, bila prepoznata, bez da to vodi protivrečnosti. Sledeća Gedelova napomena govori u prilog tome da bi ta intuicija mogla biti ispravna:

> „U osnovi *paradoksa* lažljivca može biti činjenica da relacija mapiranja jezika (prema strukturalnoj jednakosti) nije uvek jednoznačna, iako uglavnom jeste. Verovanje o jednoznačnosti zbog toga što pred očima imamo samo “tipične” slučajeve (ili možda slučajeve koji nisu tako “netipični” kao kod *paradoksa* (tj. pri formiranju svedočanstva za jednoznačnost previđamo određene po sebi malo verovatne i posebne mogućnosti)). To bi tako bila jedna vrsta dvostruke istine: ista rečenica je u jednom smislu istinita i u jednom drugom lažna, dakle jedna samo jezikom uslovljena dvostruka istina”[17] (Gödel, 2021, p. 80).

Fefermanova teorija u ovoj verziji odgovara onoj koju je predložio Gilmor kao teoriju takozvanih *parcijalnih skupova* i njima odgovarajuće relacije pripadnosti (Paul Gilmore, 1967). Ta teorija zasnovana je na logici prvog reda proširenoj sa dva predikata: $\in$ - koji stoji za pripadnost nekom skupu, i $\notin$ - koji stoji za pripadnost njegovom komplementu. Uslovi za njihovu primenu formulisani su pomoću shema aksioma koje su instancirane jedino pozitivnim formulama. Parcijalni skup i njegov komplement definisani su nekim parom pozitivnih formula koje se mogu dobiti kao pozitivne aproksimacije neke formule i njene negacije. Uslov da neki objekat pripada datom skupu jeste da za njega važi prva od tih formula i negacija druge, dok je uslov da objekat pripada komplementu tog skupa taj da za njega važi druga formula i negacija prve. Komplement nekog skupa tako nije definisan u odnosu na ceo univerzum, što ostavlja mogućnost da postoje objekti koji ne pripadaju ni nekom skupu ni njegovom komplementu.

17 „Der *Antinomie* des Lügners könnte der Sachverhalt zu Grunde liegen, dass die Abbildungsrelation der Sprache (vermöge Strukturgleichung) nicht eindeutig ist, obwohl sie im Allgemeinen eindeutig ist. Das Glauben an die Eindeutigkeit weil wir bloss “typische” Fälle vor Augen haben (oder vielleicht Fälle, die nicht so “untypisch” sind wie bei der *Antin*<*omien*> (d. h. wir übersehen bei der Erzeugung der Evidenz der Eindeutigkeit gewisse an sich unwahrscheinliche und spezielle Möglichkeiten)). Das wäre also eine Art doppelte Wahrheit: ein Satz ist in einem Sinn wahr und in einem anderen falsch, also eine nur sprachlich bedingte doppelte Wahrheit.”

Gilmor je dokazao konzistentnost ove teorije konstruisanjem njenog modela na način veoma sličan Kripkeovoj konstrukciji modela teorije parcijalnog predikata istinitosti (glavnu ulogu i ovde ima određena teorema o fiksnoj tački). Ono što omogućava taj dokaz jeste formulisanje aksioma koje karakterišu predikat pripadnosti skupu i njegovom komplementu isključivo pomoću pozitivnih formula. Gilmor je takođe pokazao da bi dodavanje bilo koje aksiome ili pravila koje implicira ekstenzionalnost parcijalnih skupova učinilo njegovu teoriju inkonzistentnom. Zahvaljujući tome, ti parcijalni skupovi mogu biti shvaćeni i kao intenzionalni objekti, odnosno, kao ekstenzije svojstava ili pojmova.

To sugeriše da bi teorija predložena od strane Gilmora i Fefermana, koju ćemo zvati *Gilmor-Fefermanova teorija*, mogla biti korišćena za formalizaciju relacije primene pojma. U toj interpretaciji ove teorije, svaka njena formula bila bi shvaćena kao da definiše pojam koji se primenjuje na objekte za koje važi njena pozitivna aproksimacija, a čiji se komplement primenjuje na objekte za koje važi pozitivna aproksimacija njene negacije. Time je ostavljena mogućnost da neki pojam ne bude svuda smisleno primenljiv, jer može postojati objekat za koji ne važi ni pozitivna aproksimacija formule kojom je taj pojam definisan, ni pozitivna aproksimacija njene negacije. Neka rečenica date teorije biće shvaćena kao da tvrdi smislenu primenu pojma samo ako je ekvivalentna svojoj pozitivnoj aproksimaciji. Da li je to slučaj ne mora uvek biti određeno. To znači da, u opštem slučaju, nije moguće odrediti da li neka rečenica te teorije tvrdi smislenu primenu pojma ili ne. To bi u ovom kontekstu odgovaralo Gedelovoj ideji da *pojam smislene primenljivosti* nije uvek smisleno primenljiv.

Gilmor-Fefermanova teorija takođe sugeriše i vrstu formula koje stoje za pojmove koji nisu svuda definisani, tj. čije instance ne opisuju nužno smislenu primenu pojma. To su formule koje sadrže određene negacije primene pojma, a koje se ne svode uvek na formule u kojima se te negacije ne javljaju. Upravo ova odlika te teorije može je činiti adekvatnom osnovom za analizu primene i smislene primenljivosti pojma koja bi vodila zasnivanju teorije pojmova. To tvrđenje pokušaćemo da opravdamo tako što ćemo nešto detaljnije razmotriti pitanja kojima teorija pojmova treba da se bavi, i koja mogu sugerisati poželjne osobine njene osnove. Time ćemo se baviti u ostatku ovog rada. Pre toga, daćemo kratak pregled rešenja paradoksa omogućena prethodno predstavljenim teorijama i razloge zbog kojih verujemo da takva vrsta rešenja odgovara intenzionalnim paradoksima.

3.6.4 *„Matematičko" rešenje intenzionalnih paradoksa*

Svaka od predstavljenih teorija omogućava rešenje intenzionalnih paradoksa koje je u matematičkom duhu, odnosno, na način ilustrovan primerima iz metamatematike i teorije izračunljivosti (kao i sa par primera iz teorije skupova). Zajednička ideja jeste da naizgled paradoksalni zaključci, koji se najčešće svode na dokaze postojanja fiksne tačke negacije, ne treba da budu odbačeni kao neprihvatljivi ili kao rezultati koji pokazuju neodrživost određenih, tipično samoreferencijalnih, metoda dokazivanja. Umesto toga, oni bi trebalo da budu usvojeni kao rezultati koji otkrivaju nešto o objektima o kojima je reč. Da bi ti rezultati mogli da budu uključeni u teoriju koja se tim objektima bavi, mora biti nađen način da se izbegne njihova najopasnija posledica - kontradikcija kojoj oni pod određenim uslovima vode. Način na koji je to postignuto u pomenutim oblastima je takođe uniforman. U teoriji izračunljivosti kontradikcije su izbegnute prepoznavanjem mogućnosti da izračunljiva funkcija koja ne potvrđuje pripadnost nekog objekta datom skupu, ne negira nužno tu pripadnost. Slično tome, rečenica koja nije dokaziva u nekoj teoriji ne mora nužno biti opovrgljiva u njoj. Ta mogućnost blokira izvođenje kontradikcije iz Gedelove rečenice koja bi bila slična onoj u osnovi PARADOKSA LAŽLJIVCA. Prilagođavanje takvog rešenja intenzionalnim paradoksima podrazumeva prihvatanje mogućnosti da pojam nije smisleno primenljiv na neki objekat u smislu da se ne može konzistentno tvrditi ni da se primenjuje ni da se ne primenjuje na taj objekat. Svaka od teorija koje su predstavljene u ovom poglavlju nudi takvu vrstu rešenja. One zadovoljavaju Gedelov zahtev da slučajevi besmislenih primena pojma budu shvaćeni kao granični slučajevi, i ne nameću nikakva opšta ograničenja kojima bi brojne primene pojmova bile proglašene besmislenim. Ono što je posebno bitno jeste da one dozvoljavaju smislenu samoprimenu pojma. Samo određene primene pojmova čiji opisi ne mogu biti konzistentno smatrani istinitim ili lažnim proglašene su besmislenim.

To je, međutim, dovoljno da onemogući uniformno povezivanje tvrđenja o primeni pojma sa primenom odgovarajućeg predikata unutar teorije zasnovane na klasičnoj logici, jer je svaka primena klasičnog predikata interpretirana kao istinita ili lažna. Iz tog razloga, nijedna od njih ne može biti pridružena tvrđenjima o besmislenoj primeni pojma. Jedan način prevazilaženja tog problema jeste reinterpretacija rečenica izgrađenih primenom predikata na način koji ostavlja mogućnost da te rečenice nisu ni istinite ni lažne. To vodi usvajanju neke trovrednosne logike. Druga opcija je da se izdvoje one primene predikata koje mogu biti konzistentno smatrane istinitim ili lažnim u kontekstu neograničene komprehenzije, i da se samo one smatraju relevantnim za pitanje primene odgovarajućeg pojma. To može biti postignuto na različite načine: proširenjem klasičnog jezika parcijalnim predikatom istinitosti ili intenzionalnom ekvivalencijom, kao i shvatanjem kao istinitih i lažnih samo takvih primena predikata koje su ekvivalentne onima formulisanim pomoću pozitivnih formula.

Svaka od teorija do koje se dolazi na taj način nudi traženu vrstu rešenja intenzionalnih paradoksa. Ali koja od njih i po svojim ostalim osobinama predstavlja adekvatnu osnovu buduće teorije pojmova? Pokušaćemo da odgovorimo na to pitanje u sledećem poglavlju gde ćemo razmotriti neke od zahteva za adekvatnu osnovu te teorije do kojih ćemo doći prateći Gedelove napomene.

3.7 POSTAVLJANJE OSNOVA TEORIJE POJMOVA

Kako bi obezbedila odgovarajuću osnovu za istraživanje formalnih osobina pojmova, teorija pre svega treba da omogući rešenje intenzionalnih paradoksa bez ograničavanja aksiome komprehenzije na određene formule ili određene domene u kojima neka formula treba da važi. Taj glavni zahtev za osnovu teorije pojmova, koji sve prethodno predstavljene teorije zadovoljavaju, nije i jedini. Razmatranjem nekih dodatnih zahteva pokušaćemo da utvrdimo koja od njih najlakše može biti prilagođena onoj vrsti istraživanja koje bi teorija pojmova trebalo da omogući.

U tom kontekstu, razlike između logičkih sistema na kojima su navedene teorije izgrađene mogu biti od suštinskog značaja. One su zasnovane ili na nekom sistemu *trovrednosne logike* ili na *klasičnoj logici*. Zato bismo, pri razmatranju toga koja od njih može da obezbedi najbolju osnovu za teoriju pojmova, morali da razmotrimo i *koja logička osnova bi najviše odgovarala toj teoriji?*

Najjači argument u prilog tome da to ne može biti bilo koji sistem trovrednosne logike jeste činjenica da ta logika onemogućava upotrebu nekih osnovnih principa rasuđivanja koji su uobičajeni u logici i matematici. Logička snaga teorije izgrađene na njoj, samim tim, u znatnoj je meri oslabljena. Iz tog razloga, trovrednosna logika ne bi mogla da predstavlja odgovarajuću logičku osnovu za teoriju koja bi, kako je Gedel očekivao, trebalo da zauzme centralno mesto u logici. Drugi razlog zbog kog pristup zasnovan na trovrednosnoj logici ne bi bio adekvatan u ovom kontekstu jeste i činjenica da rešenje intenzionalnih paradoksa koje on pruža ne doprinosi ni na koji način razumevanju pojmova i njihove primene, kao i pogrešnih pretpostavki o njima koje su dovele do paradoksa. Osim što koristi logiku koja je bolje od klasične logike prilagođena rasuđivanju koje se tiče rečenica bez istinosne vrednosti, među kojima su i one u osnovi paradoksa, taj pristup ne predlaže nikakve konkretne ideje o tome kojim tačno rečenicama treba poreći istinosnu vrednost i u kakvoj su one vezi sa pitanjem primene ili smislene primenljivosti pojma. Tako nam ovaj pristup ne donosi nikakva nova saznanja o osobinama pojmova i razlozima zbog kojih su neke njihove primene besmislene. To je, međutim, nešto što je Gedel očekivao od odgovarajućeg rešenja intenzionalnih paradoksa. Ono bi trebalo da unapredi naše razumevanje formalnih osobina pojmova u istoj meri u kojoj je rešenje skupovno-teorijskih paradoksa (zasnovano na iterativnom pojmu skupa) unapredilo naše razumevanje skupova.

Na osnovu navedenog možemo zaključiti da rešenje paradoksa zasnovano na trovrednosnoj logici ne predstavlja adekvatnu osnovu teorije pojmova. Ali da li bi neki pristup zasnovan na klasičnoj logici imao bolje šanse za to? Ono što bi u tom kontekstu bilo neophodno razmotriti jeste možemo li se u ekstenzionalnom logičkom okruženju kakvo pruža klasična logika, i dalje baviti intenzionalnim svojstvima objekata kao što su pojmovi? Kako bismo odgovorili na to pitanje, moramo razmotriti neke konkretne zahteve za buduću teoriju koji su u vezi sa intenzionalnom prirodom objekata kojima se ona bavi.

Prvi takav zahtev, koji proističe iz rasprave u prethodnom delu rada, jeste da ta teorija treba da prepozna parcijalnost pojmova, odnosno, činjenicu da oni nisu nužno svuda smisleno primenljivi, što bi trebalo istovremeno da omogući njenu konzistentnost. Videli smo da to može biti postignuto i u teorijama zasnovanim na klasičnoj logici. Još jedan zahtev, koji je Gedel posebno naglašavao, jeste da bi ta teorija trebalo da opiše principe građenja pojmova, kao i principe kojima je to građenje regulisano. Pošto bi logički veznici trebalo da imaju značajnu ulogu u građenju pojmova, značenje koje im logička osnova teorije pripisuje može biti posebno važno za pitanje njene adekvatnosti. Značenje logičkih veznika bi, osim njihovoj ulozi u građenju pojmova, takođe trebalo da odgovara i njihovoj ulozi u opisivanju relacija između pojmova. Videćemo da bi opis strukture koju formiraju pojmovi koji stoje u određenim relacijama, mogao da predstavlja i glavni cilj teorije pojmova. U nastavku ćemo razmotriti da li ovi zahtevi mogu biti zadovoljeni u teoriji zasnovanoj na klasičnoj logici ili bi ipak neki drugi logički sistem bio adekvatniji.

3.7.1 *Građenje pojmova*

> „Suština ovih objekata [pojmova] leži, čini se, u njihovoj sposobnosti da grade složene objekte, tako da ih nikada ne možemo razumeti (znati), a da i te složene objekte ne razumemo”[18] (Gödel, 2017b, p. 1).

Jedan od ciljeva teorije pojmova bilo bi ispravno razumevanja načina na koji su složeni pojmovi izgrađeni od drugih pojmova. Značaj koji je Gedel pripisivao tom cilju proizilazi iz njegovog stanovišta o pojmovima, koje je sugerisano citiranom napomenom, a po kojem je suštinska osobina pojmova to da se kombinuju i tako grade složene pojmove. To, čini se, implicira da bi neka vrsta *konstruktivističkog* pristupa u istraživanju formalnih osobina pojmova bila adekvatna. U tom pristupu ograničili bismo se na izučavanje onih osobina pojmova koje zavise od načina njihovog građenja i na opisu tog građenja zasnivali naša tvrđenja o pojmovima.

Složeni pojmovi izgrađeni su prema određenim principima koji bi trebalo da budu identifikovani i formalno predstavljeni unutar teorije pojmova. Gedel razmatra mogućnost da su ti principi takođe dati u obliku pojma ili neke pojmovne strukture:

> „Mogao bi čak postojati i pojam koji je jednako precizan kao matematički pojmovi (ili čak precizniji od njih), i koji sadrži princip svakog građenja pojmova” ... „Možda taj pojam nije jedinstveni objekat, već predstavlja pojmovnu strukturu sačinjenu od više međusobno povezanih pojmova”[19] (Gödel, 2020a, pp. 63-64).

18 „Das Wesen dieser Gegenstände [Begriffe] scheint aber in ihrer Fähigkeit, Zusammensetzungen zu bilden, zu liegen, sodass man sie nicht verstehen (kennen) kann, ohne auch die Zusammensetzungen zu verstehen.”

19 „Es gibt wahrscheinlich irgendeinen Begriff, der ebenso (oder noch mehr) präzise ist als die mat<hematischen> Begriffe, und welcher das Prinzip jeder Begriffsbildung enthält”... „Vielleicht ist dieser Begriff keine Einheit, sondern eine Begriffsstruktur, bestehend aus mehreren miteinander in Beziehung stehenden Begriffen.”

Prema Gedelovom mišljenju, standardni načini građenja pojmova, koji bi trebalo da odgovaraju načinu na koji su složene formule odgovarajuće teorije izgrađene, jesu nesumnjivi, ali neka ograničenja koja se tiču njihove primene ipak moraju da budu uvedena:

> „Sasvim je u redu formirati pojmove na poznati način: imamo evidenciju da su to smisleni i ispravni načini formiranja pojmova. Ono što je pogrešno nisu posebni načini njihovog formiranja, nego ideja da možemo proizvoljno formirati pojmove korišćenjem ispravnih principa" (Wang, 1996, 8.5.20).

Nepoštovanje principa građenja pojmova ili njihovih restrikcija moglo bi da bude upravo ono što vodi formulaciji paradoksa. Naime, posledica tog nepoštovanja bilo bi korišćenje nekih formula za izražavanje pojmova koje ne predstavljaju na odgovarajući način njihovu strukturu. Nekom instancom takve formule bi se onda mogla tvrditi primena pojma koju sama njegova struktura, prikrivena tom formulom, ne dozvoljava. Iz toga bi sledilo da je razlog zbog kog neke rečenice ne uspevaju da izraze smislenu primenu pojma taj što formule čije su te rečenice instance ne opisuju na odgovarajući način njegovu strukturu. Takve formule bi mogle da ukažu na principe građenja pojmova koji su njima prekršeni.

Među teorijama koje smo razmatrali u prethodnom delu rada, Gilmor-Fefermanova teorija izdvaja se po tome što se najeksplicitnije tiče forme rečenica koje, u našoj interpretaciji te teorije, izražavaju besmislenu primenu pojma. Te rečenice su instance onih formula koje nisu ekvivalentne nijednoj pozitivnoj formuli. Na osnovu toga bi se moglo zaključiti, u skladu sa prethodno rečenim, da su upravo pozitivne formule one koje na odgovarajući način predstavljaju strukturu nekog pojma određenu njegovom konstrukcijom. Sa druge strane, formule koje nisu ekvivalentne nijednoj pozitivnoj formuli mogu se shvatiti kao pokušaj definisanja pojma koji ne može biti izgrađen na način predstavljen tim formulama. Takve formule, kao što smo videli, sadrže negaciju primene nekog pojma koja se ne može svesti na primenu nekih drugih pojmova. Zato bi se principi građenja pojmova sugerisani tim formulama ticali *uloge negacije u građenju komplementa nekog pojma*. Iz njih bi sledilo da komplement pojma ne može uvek biti definisan pomoću negacije primene tog pojma, već samo tada kada se takva definicija može svesti na onu kojom se tvrdi primena nekog drugog pojma. Taj drugi pojam bio bi definisan pomoću pozitivne aproksimacije negacije formule kojom je prvi pojam definisan.

Ova teorija i rešenje paradoksa koje ona nudi mogu se lako predstaviti u svetlu gore pomenutog konstruktivističkog pristupa pojmovima koji je sugerisan Gedelovim napomenama. Ideja bi bila da pogrešna pretpostavka koja vodi paradoksima jeste da pojam možemo definisati samo tako što ćemo odrediti njegov domen primene kao sve ono na šta se neki drugi pojam ne primenjuje. Nasuprot tome, Gilmor-Fefermanova teorija bi u navedenoj interpretaciji zahtevala da, kako bismo mogli da tvrdimo bilo šta o tom pojmu, moramo dati njegov konstruktivniji opis pomoću odgovarajuće pozitivne formule. Ta formula bi trebalo da otkrije strukturu datog pojma, koja ostaje nevidljiva u njegovoj „negativnoj definiciji". Jedna od formula koja bi predstavljala takvu negativnu definiciju pojma je i formula $\neg T_1(x, x)$, kojom bi u našoj interpretaciji trebalo da bude definisan *pojam neprimenljivosti na sebe*. Iz gore navedenog sledi da ta formula daje definiciju pojma samo ako se može svesti na formulu $\bar{T}_1(x, x)$ koja predstavlja definiciju *pojma potpadanja pod sopstveni komplement*. Tako definisan pojam ne daje osnova za formulisanje bilo kakvog paradoksa.

To da regulatorni principi građenja pojmova treba da se odnose na način upotrebe logičkih veznika u tom građenju, posebno *negacije*, opravdano je ulogom koju bi ti veznici trebalo da imaju u građenju složenih pojmova. Naime, upravo bi oni trebalo da daju strukturu složenom pojmu:

> „U analizi pojma na njegove sastavne delove, logički pojmovi posebno grade „vezivno tkivo" koje ih drži zajedno. U tom smislu oni su ono „formalno", naime, struktura sastavnih delova"[20] (Gödel, 2017b, p. 1).

> „Pojam je celina sastavljena od primitivnih pojmova kao što je negacija, konjunkcija, postojanje, univerzalnost, objekat, pojam pojma, relacija potpadanja pod neki pojam (ili primene pojma na nešto), i tako dalje" (Wang, 1996, 8.6.17).

U svetlu uloge koju logički veznici imaju u građenju pojmova, pitanje značenja koje im pripisuje logička osnova teorije koja se tim građenjem bavi, dobija poseban značaj. Postoje određeni argumenti u prilog tome da bi opisu te uloge logičkih veznika najbolje odgovarala interpretacija koja im je pripisana u slaboj trovrednosnoj logici. Naime, čini se da ako se neki pojam ne primenjuje smisleno na određeni objekat, onda bi isto trebalo da važi i za sve one pojmove koji su izgrađeni od njega pomoću veznika. Isto je implicirano i Gedelovim zapažanjem u vezi sa Fregeovim predloženim rešenjem RASELOVOG PARADOKSA, prema kojem isključivanje određenog objekta iz domena smislene primenljivosti nekog predikata zahteva isključivanje istog objekta iz domena smislene primenljivosti svih složenih predikata izgrađenih od njega pomoću veznika. Takvo shvatanje veznika bi takođe bilo opravdano konstruktivističkim pristupom koji uzima za ozbiljno činjenicu da značenje nekog pojma zavisi od značenja pojmova od kojih je on izgrađen. Ako značenje primene nekog pojma na određeni objekat nije određeno, onda bi isto važilo i za svaki složeni pojam u čijem značenju on učestvuje.

Način na koji Gilmor-Fefermanova teorija tretira formule izgrađene od onih negacija primena pojma koje su prema toj teoriji besmislene, u skladu je sa slabom trovrednosnom interpretacijom veznika koji učestvuju u građenju tih formula. Naime, osim tih negacija, takođe će i sve druge formule koje ih sadrže biti izuzete iz skupa formula koje opisuju smislenu primenu odgovarajućeg pojma. Te formule neće biti ekvivalentne svojim pozitivnim aproksimacijama, pa aksiome komprehenzije neće njima biti instancirane. Tako bi Gilmor-Fefermanova teorija, iako nije zasnovana na trovrednosnoj logici, ipak omogućila adekvatno razumevanje veznika u kontekstu građenja pojmova.

Zauzimanje konstruktivističkog pristupa u istraživanju pojmova moglo bi u značajnoj meri da utiče na pitanje logičke osnove na kojoj treba zasnovati teoriju koja omogućava takvo istraživanje. Čini se da bi, umesto klasične logike čiji su zakoni prilagođeni rasuđivanju o nekim nezavisno postojećim objektima, tom pristupu pre mogla odgovarati *intuicionistička logika*. U narednom poglavlju daćemo grubi pregled konstruktivističkih ideja u filozofiji matematike, kao i intuicionističke logike koja bi tim idejama trebalo da obezbedi logičku osnovu. Izdvojićemo samo one odlike tog logičkog sistema za koje nam se čini da mogu biti relevantne za našu temu. Detaljniji prikaz intuicionističke logike i njene veze sa konstruktivizmom može se naći, na primer, u (van Dalen, 2013).

3.7.1.1 *Konstruktivizam u teoriji pojmova i intuicionistička logika*

Intuicionistička logika razvijena je u cilju prilagođavanja logičkog rasuđivanja konstruktivističkom stanovištu u filozofiji matematike. Glavni predstavnici tog stanovišta su Kroneker (Leopold Kronecker) - koji je predložio načelne ideje o njemu, i Brauer - koji je te ideje pretvorio u razrađeno stanovište u filozofiji matematike. Prema tom stanovištu, tvrđenje da neki matematički objekat postoji je opravdano samo ako je poznat način na koji se taj objekat može konstruisati. Tvrđenja o svojstvima matematičkih objekata opravdana su takođe samo ako su dokaziva na konstruktivan način (tj. bez korišćenja pravila *jakog svođenja na apsurd* ili njemu ekvivalentnog). Gedel nije

20 „Insbesondere bilden bei der Analyse eines Begriffes in seine Bestandteile die logischen Begriffe „den Kitt", welcher sie zusammenhält. In diesem Sinn sind sie das „Formale", nämlich die Struktur der Bestandteile."

zastupao konstruktivističko stanovište u vezi matematičkih objekata, ali njegove napomene o pojmovima govore u prilog tome da bi mogao smatrati konstruktivistički pristup u njihovom istraživanju adekvatnim. Iako, po njegovom mišljenju, jedino klasična logika pruža odgovarajuću osnovu za matematičku teoriju koja se bavi ekstenzionalnim objektima, intuicionistička logika bi mogla bolje da odgovara specifično logičkom predmetu koji čine intenzionalni objekti kao što su pojmovi.

Umesto o *istinitosti* iskaza, u intuicionističkoj logici se pre govori o njihovoj *dokazivosti.* Kada tvrdimo neki iskaz, tvrdimo da je on dokaziv i to na konstruktivan način. U skladu sa tim, intuicionistički veznici nisu shvaćeni kao istinosne funkcije, već je njihovo značenje određeno vrstom dokaza koji odgovara formuli izgrađenoj pomoću njih. Tako je, na primer, disjunktivna formula dokaziva ako i samo ako se može dati dokaz bar jednog od njenih disjunkata, dok je konjunktivna formula dokaziva ako i samo ako to važi za oba njena konjunkta. Intuicionistička implikacija interpretirana je kao da tvrdi da se pomoću dokaza njenog antecedensa može doći do dokaza njenog konsekvensa. Pomoću nje je definisana i intuicionistička negacija kojom se tvrdi da se pomoću dokaza neke formule može doći do dokaza kontradikcije (cf. van Dalen, 2013, pp. 155-157). Dokazivost neke formule u intuicionističkoj logici tako ključno zavisi od njene strukture određene veznicima koji učestvuju u njenom formiranju. Ta struktura formule bi u adekvatnoj teoriji trebalo da odslikava strukturu njom izraženog pojma. Tako bi pitanje dokazivosti formule u intuicionističkoj logici moglo biti usko povezano sa pitanjem strukture pojma izraženog tom formulom.

Gedelovo mišljenje o intuicionističkoj logici kao osnovi teorije pojmova, sugerisano je njegovom primedbom koju iznosi u kontekstu razmatranja nepotpunosti matematičkih teorija i unapređenja razumevanja matematičkih pojmova:

> „*Intuicionistički* put do rešenja sastoji se u upotpunjavanju pojmovog sistema kroz izostavljanje - ili on služi samo tome da usmeri naše traganje za stvarnom situacijom na osnovu primera u kojem „potpuniji" može biti posmatran zajedno sa okrnjenim sistemom? [Na sličan način na koji je četvrta dimenzija rasvetljena pomoću dvodimenzionalnog.]"[21] (Gödel, 2020a, p. 27).

Ova Gedelova napomena sugeriše da je on video konstruktivistički pristup pojmovima i, u skladu sa tim, upotrebu intuicionističke logike, kao mogući prelazni stadijum ka nekom objektivnijem, nekonstruktivnom razumevanju sveta pojmova. Korišćenje intuicionističke logike moglo bi da približi svet pojmova našem ograničenom razumevanju i omogući nam da izgradimo solidnu osnovu na kojoj ćemo to razumevanje dalje razvijati. Za početak bismo se ograničili na opisivanje načina na koji su pojmovi izgrađeni i na tome zasnivali tvrđenja o njihovim osobinama. Ti aspekti pojmova jesu ono što nam je u vezi sa njima i najbliže i što možemo znati sa izvesnom sigurnošću. Ograničavanje na poznate konstruktivne aspekte pojmova takođe bi trebalo da nas zaštiti od paradoksa koji su, po Gedelovom mišljenju, posledica ekstrapolacije našeg ograničenog razumevanja pojmova i njihovih svojstava na ceo svet pojmova.

Sa druge strane, sam prelazak na intuicionističku logiku ne sprečava izvođenje paradoksa koji su najčešće formulisani pomoću klasičnih principa rasuđivanja (kao što je *princip isključenja trećeg*). Unutar teorije zasnovane na intuicionističkoj logici, u kojoj taj princip ne važi, paradoksi će i dalje biti izvodivi pomoću rasuđivanja koje je u osnovi KARIJEVOG PARADOKSA. Pošto intuicionistička logika ne sprečava izvođenje paradoksa, njihovo rešenje bi i dalje bilo traženo u Gilmor-Fefermanovoj teoriji koja se jednako dobro može izgraditi na intuicionističkoj, kao i na klasičnoj logičkoj osnovi (Feferman, 1984, p. 78).

21 „Der *int*<*uitionistische*> Weg zur Auflösung besteht darin, das Begriffssystem durch Weglassung zu vervollständigen — oder dient das nur dazu, dass wir die wahre Situation an einem Beispiel verfolgen können, wo man sowohl das „vollständigere" als das verstümmel<te> System sieht? [Ähnlich wie die vierte Dimension durch das Zweidimensionale erläutert wird.]"

Osnovna razlika između intuicionističke i klasične logike leži u njihovoj interpretaciji implikacije. Kao i ostali veznici, implikacija je u klasičnoj logici shvaćena na potpuno ekstenzionalan način, tj. kao istinosna funkcija (ta implikacija naziva se *materijalnom implikacijom*). Kao takva, ona je u potpunosti određena istinosnim vrednostima njenog antecedensa i konsekvensa i nezavisna od postojanja bilo kakve dodatne veze između njih zasnovane na njihovom značenju. Intuicionistička implikacija je intenzionalnija. Veza između formula koju ona izražava zasniva se u većoj meri na njihovim intenzionalnim karakteristikama koje su određene strukturom tih formula. Kao što je već rečeno, ta implikacija tiče se mogućnosti da se od dokaza jedne dođe do dokaza druge formule. To „dolaženje” od dokaza jedne do dokaza druge formule možemo razumeti kao *dedukciju* jedne formule iz druge. Tako shvaćena intuicionistička implikacija bila bi usko vezana za dedukciju koju bi odslikavala. Ona bi tako odgovarala implikaciji na koju Gedel referira u sledećoj napomeni:

> „Da je objektivno R i $\check{R}$ isto, posebno se jasno vidi u slučaju *implikacije*, jer se u kontekstu logičkog ispitivanja suštine *implikacije*, relacija posledice i *implikacija* ne razmatraju odvojeno. (To je jedna ista stvar)”[22] (Gödel, 2020b, p. 19).

U beleškama za kurs na Notr Dejmu, Gedel opisuje intenzionalne veznike kao one koji ne mogu biti zadati istinosnom tablicom. Intenzionalna implikacija je pritom shvaćena kao veznik pomoću kojeg se mogu formulisati tvrđenja o tome da je jedna formula logička posledica druge u smislu da sledi iz nje nekim logičkim zaključivanjem (Gödel, 2017a, pp. 18-19). Tako opisanu implikaciju Gedel naziva ’striktnom implikacijom’ i povezuje je sa modalnom logikom. S obzirom na to da sve što kaže o intenzionalnoj implikaciji odgovara njenoj intuicionističkoj interpretaciji, nije u potpunosti jasno zbog čega Gedel ipak daje prednost striktnoj implikaciji (vidi: Adžić & Došen, 2016, pp. 473). U svakom slučaju, jasno je da je Gedel razumeo implikaciju koja bi bila usko povezana sa dedukcijom kao intenzionalnu. Po tome bi ona bila srodnija od klasične implikacije predmetu istraživanja teorije pojmova koji predstavljaju formalna svojstva samog intenzionalnog značenja.

Opisivanjem veze između dokazivosti neke dve formule, intuicionistička implikacija, kao i ekvivalencija definisana pomoću nje, takođe bi omogućile opisivanje relacija između pojmova izraženih tim formulama. Upravo bi ispravno razumevanje tih relacija između pojmova mogao biti jedan od glavnih ili čak glavni zadatak za buduću teoriju pojmova.

3.7.2 *Relacije između pojmova*

> „Pod analizom nekog pojma podrazumeva se i nalaženje fundamentalnih relacija u kojima on stoji prema drugim stvarima, a ne samo njegovo razdvajanje na elementarne sastavne delove. Ovo bi bila u nekom smislu „spoljašnja” analiza, zapravo takva spoljašnja analiza je i cilj”[23] (Gödel, 2020a, p. 79).

Krajnji cilj teorije pojmova bi tada bio potpuni opis mreže koju grade međusobno povezani pojmovi i unutar koje je njihovo značenje određeno. Šta tačno određuje značenje nekog pojma za sada nam nije u potpunosti poznato. Ali jasno je da njega ne čine objekti koji potpadaju pod pojam na način na koji objekti koji su elementi nekog skupa čine taj skup. Iz tog razloga, hijerarhijska slika sveta pojmova ponuđena u teoriji tipova ne bi adekvatno predstavljala način

22 „Dass objektiv R und $\check{R}$ dasselbe ist, sieht man besonders deutlich bei der *Impli<kation>*, denn, wenn man logische Untersuchungen über das Wesen der *Impli<kation>* anstellt, so hat man nicht die Relation des Folgens und der *Impli<kation>* gesondert zu betrachten. (Es ist ein Wesen).”

23 „Unter Analyse eines Begriffes versteht man auch die Auffindung der fundamentalen Beziehungen zu anderen Dingen, nicht bloss sein Aufspalten in elementare Bestandteile. Das ist also gewissermassen eine Analyse von „aussen“, insbesondere der Zweck ist eine solche Analyse von aussen”.

na koji je značenje nekog pojma određeno na osnovu značenja drugih pojmova. Gedel izgleda dolazi do sličnog zaključka (moj kurziv):

> „Iako nemamo razvijenu teoriju pojmova, ipak o pojmovima znamo dovoljno da bismo znali da možemo imati nešto kao hijerarhiju pojmova (ili takođe klasa) koja liči na hijerarhiju skupova i sadrži je kao segment. Ali takva hijerarhija je izvodiva i periferna je u odnosu na teoriju pojmova; ona takođe zauzima dosta drugačije mesto; na primer, ona ne može zadovoljiti uslov da sadrži pojam *pojma* koji se primenjuje na sebe ili univerzum svih klasa koje pripadaju sebi. *Uzimanje te hijerarhije za teoriju pojmova je primer pokušaja eliminisanja intenzionalnih paradoksa na proizvoljan način*" (Wang, 1996, 8.6.20).

Ova Gedelova napomena govori u prilog ideji, praćenoj u ovom radu, da mogućnost samoprimene nekog pojma otkriva nešto bitno o njegovoj prirodi i odnosu prema drugim pojmovima u kojem je njegovo značenje određeno. Usvajanje slike sveta pojmova opisane teorijom tipova onemogućilo bi praćenje traga u vidu mogućnosti njihove samoprimene koji bi trebalo da vodi ispravnom razumevanju prirode pojmova.

Osim hijerarhije pojmova zasnovane na domenu njihove smislene primenljivosti, jedna druga hijerarhija pojmova može se zamisliti - ona koja bi bila zasnovana na građenju složenih od jednostavnijih pojmova. U osnovi takve hijerarhije bili bi neki sasvim jednostavni pojmovi koji sami nisu izgrađeni od drugih pojmova i od kojih bi, na neki način, svi ostali pojmovi trebalo da budu izgrađeni. Čini se je Gedel očekivao da buduća teorija pojmova pruži baš takav opis sveta pojmova. Jednostavni pojmovi bili bi, između ostalih, i logički pojmovi, kao i pojmovi same teorije pojmova - pojam, relacija primene i smislene primenljivosti pojma, klasa, itd. Ostali pojmovi bi trebalo da budu izgrađeni od njih prema principima građenja koje postulira teorija pojmova.[24]

Takva predstava sveta pojmova ne bi, međutim, trebalo da sugeriše da je značenje nekog pojma određeno isključivo pojmovima koji učestvuju u njegovom građenju i načinom na koji su oni međusobno povezani (kako bi tada značenje jednostavnih pojmova bilo određeno?). Ono takođe može zavisiti od konteksta u kom se taj pojam javlja, tj. od njegovog učešća u građenju značenja drugih pojmova, kao i u građenju iskaza i dokaza. Značenje nekog pojma tada može biti objašnjeno i pomoću složenijih intenzionalnih entiteta u čijem građenju on učestvuje. Tako, na primer, pojam identiteta ne možemo objasniti pomoću pojmova koji ga čine, pošto je on jedan jednostavan pojam, već ga umesto toga definišemo pomoću pojmova koji su složeniji od njega i koji ga, po svoj prilici, i sami pretpostavljaju (cf. Gödel, 2017b, p. 64).

Objašnjenja jednostavnih pojmova pomoću složenih mogu se naći pre svega u jednoj matematičkoj disciplini - *teoriji kategorija* (vidi: Došen, 2017). Može se reći da se teorija kategorija bavi matematičkim strukturama. Kao takva, ona zauzima spoljašnju perspektivu u odnosu na matematičke (ili logičke) objekte koji su deo tih struktura, i omogućava njihovo objašnjenje na osnovu njihovog mesta u tim strukturama i uloge koju oni imaju u njihovom formiranju. Došen daje primer pojma uređenog para i proizvoda koji su u teoriji kategorija objašnjeni pomoću veoma složenog aparata koji uključuje pojam adjunkcije između dijagonalnog funktora i bifunktora proizvoda. Ovi složeni pojmovi koji treba da omoguće objašnjenje pojma uređenog para i proizvoda, i sami pretpostavljaju te pojmove. Po Došenovom mišljenju, objašnjenja slična pomenutom „mogu služiti kao prototip nečega što se može očekivati od sveobuhvatne samoreferencijalne intenzionalne matematike budućnosti kakvu je Gedel predvideo" (Došen, 2017).

Umesto slike prema kojoj značenje nekog pojma zavisi isključivo od značenja jednostavnijih pojmova od kojih je on izgrađen, a koje je unapred određeno i nezavisno od njega, jedna druga slika o međusobnom odnosu pojmova proizilazi iz prethodno rečenog. Prema toj novoj slici,

24 Takav opis građenja pojmova odgovarao bi opisu građenja rekurzivnih funkcija.

značenje nekog pojma može zavisiti od značenja složenih pojmova u čijem građenju i on sam učestvuje. Odnos zavisnosti značenja između pojmova u opisanoj hijerarhiji išao bi tada u oba smera. Tako ponovo dolazimo do toga da samoreferencija, ovog puta u obliku samosadržavanja i samokarakterizacije pojma, može imati značajnu ulogu u ispravnom razumevanju sveta pojmova. Značaj te vrste samoreferencije za rasvetljavanje formalnih osobina pojmova ukazuje na neodrživost predstave o njihovom međusobnom odnosu prema kojoj bi oni pojmovi koji učestvuju u građenju složenih pojmova dolazili pre njih i bili nezavisno određeni. Umesto toga, samoreferencija bi bila ugrađena u samu strukturu sveta pojmova i njihovih međusobnih relacija u kojima je njihovo značenje određeno.

Adekvatno predstavljanje načina na koji je značenje nekog pojma određeno bi, kako je nagovešteno Gedelovom napomenom sa početka ovog odeljka, zahtevalo zauzimanje „spoljašnje" perspektive u odnosu na taj pojam. Ona bi omogućila prepoznavanje uloge koju drugi pojmovi imaju u njegovom značenju, a koja nije nagoveštena njegovom strukturom. Taj pristup omogućio bi objašnjenja pojmova slična onima koje nudi teorija kategorija. Da se u rasvetljavanju značenja pojmova koje karakteriše samoreferencija treba ugledati na objašnjenja data u teoriji kategorija, sugeriše i sledeća Gedelova napomena:

> „Teorija kategorija razvijena je u cilju pokazivanja neadekvatnosti teorije skupova. Ona je u većoj meri zainteresovana za izvodljive formulacije određenih matematičkih argumenata koji naizgled koriste samoreferenciju" (Wang, 1996, 4.3.5).

Kao što smo spomenuli u prethodnom poglavlju, relacije između pojmova na kojima se zasniva njihovo značenje mogle bi se pokazati u konstruktivnim dokazima u kojima formule koje izražavaju te pojmove učestvuju, i u međusobnom odnosu tih dokaza koji je opisan pomoću intuicionističke implikacije. Zahvaljujući rezultatima teorije dokaza znamo da su intuicionistički dokazi i njihovi međusobni odnosi predstavljivi unutar teorije kategorija. Naime, struktura koju intuicionistički dokazi grade odgovara jednoj od struktura koje ta teorija izučava. To govori u prilog tome da bi, zahvaljujući njenoj vezi sa dokazima, intuicionistička logika mogla da obezbedi okruženje, slično onom koje pruža teorija kategorija, u kojoj bi „spoljašnja" analiza pojmova bila moguća.

Da sumiramo, razmatranje ciljeva teorije pojmova koji se sastoje u opisu građenja pojmova i njihovih relacija, ukazuje na to da bi najprikladnija logička osnova te teorije bila intuicionistička logika. To je u skladu sa konstruktivističkim pristupom za koji smo uzeli da je preporučen od strane Gedela kao pristup kojim bi adekvatno istraživanje pojmova moglo da započne. Istovremeno, zahvaljujući njenoj bliskoj vezi sa dokazima ili dokazivošću njenih formula, intuicionistička logika bi mogla da pruži dobru osnovu za razumevanje odnosa između pojmova odslikanih u dokazima u kojima učestvuju formule koje te pojmove izražavaju. Kako smo pokušali da pokažemo u ovom poglavlju, teorija koja može biti izgrađena na intuicionističkoj logičkoj osnovi i koja bi pružila adekvatno rešenje intenzionalnih paradoksa i istovremeno inicirala istraživanje o principima građenja pojmova, jeste Gilmor-Fefermanova teorija. Rešenje paradoksa koje ta teorija nudi zahteva minimalne izmene našeg intuitivnog razumevanja pojmova i njihove primene. Te izmene tiču se građenja pojmova pomoću logičkih veznika i mogu predstavljati korak ka njegovom boljem razumevanju. One ne impliciraju zabranu samoprimene i samokarakterizacije pojmova, što se u svetlu prethodnih razmatranja pokazuje kao posebno bitno. Na taj način, praćenjem Gedelovih napomena koje se tiču zadataka i poželjnih osobina buduće teorije pojmova, dolazimo do toga da Gilmor-Fefermanova teorija pruža adekvatno rešenje intenzionalnih paradoksa. Ta teorija mogla bi, samim tim, da predstavlja osnovu čijim bi proširenjem odgovarajućim aksiomama ili pravilima bila izgrađena teorija pojmova kakvu je Gedel zamislio.

4

ZAKLJUČAK

Glavni cilj ovog rada jeste rasvetljavanje Gedelove ideje o teoriji pojmova - njenog predmeta i glavnih zadataka, kao i načina na koji oni mogu biti ispunjeni. Pokušali smo da prateći te ideje dođemo do nekih konkretnih predloga o tome kako bi osnova takve teorije, koja bi trebalo da omogući odgovarajuće rešenje intenzionalnih paradoksa, trebalo da izgleda. Jedna od uloga teorije pojmova jeste da omogući razumevanje određenih osobina lingvističkih izraza, među kojima su i logički i matematički simboli, koje ostaju van domašaja skupovno-teorijske semantike. Među tim osobinama, samoreferencija ima posebno važno mesto, sudeći po njenoj ulozi u logičkim i matematičkim teorijama i po onome šta bi mogla da otkrije o intenzionalnom značenju. Iz tog razloga je i značajan deo ovog rada posvećen upravo razmatranju načina na koji samoreferencija može biti uključena u teoriju pojmova.

Osim što bi omogućila objašnjenje nekih važnih aspekata i osobina lingvističkih izraza, teorija pojmova bi, po Gedelovom mišljenju, imala i mnogo širi logički i matematički značaj. Kako bismo razumeli njenu ulogu u kontekstu dominantno ekstenzionalne logike i matematike, pokušaćemo da razjasnimo zbog čega joj je Gedel uopšte pridavao toliki značaj i insistirao na potrebi za njenim zasnivanjem. Odgovor na to pitanje može se naći u Gedelovom filozofskom stanovištu u pogledu ontološkog statusa pojmova, i sa njim povezanim idejama o prirodi matematičkog znanja i načinu na koji se do njega dolazi.

4.1 GEDELOV POJMOVNI REALIZAM

Gedel je dobro poznat po zastupanju platonističkog stanovišta u filozofiji matematike, prema kojem su matematički objekti i njihova svojstva nezavisni od našeg znanja i uopšte naše sposobnosti da ih saznamo. Ono što je specifično, ali i manje poznato u vezi Gedelovog platonizma, jeste da je on nezavisno postojanje pripisivao ne samo matematičkim objektima, kao što su brojevi ili skupovi, nego i pojmovima pod koje ti i drugi objekti potpadaju. Pojmove, prema tom stanovištu, ne možemo slobodno konstruisati niti menjati, već ih samo možemo u određenoj meri razumeti i na osnovu tog razumevanja koristiti. U tako shvaćenim pojmovima sadržano je intenzionalno značenje izraza kako običnog, tako i formalnih jezika, koji referiraju na objekte na koje se ti pojmovi primenjuju. Umesto svođenja intenzionalnog značenja matematičkih pojmova na to kako su nam oni dati, Gedel mu pripisuje objektivno postojanje nezavisno od jezika i naših mentalnih. sposobnosti[1] Ova specifična Gedelova pozicija ima značajne posledice po njegovo razumevanje prirode matematičkog znanja i načina na koji se do njega dolazi.

Uobičajena pretpostavka među Gedelovim interpretatorima i kritičarima njegovog stanovišta jeste da je Gedel postulirao neko mistično čulo kojim opažamo nezavisno postojeće matematičke objekte. Njihovim opažanjem treba da dođemo do saznanja o tim objektima koje zatim izražavamo pomoću matematičkih aksioma. Brojne kritike Gedelove pozicije u filozofiji matematike usmerene su baš na taj aspekat navodno njegovog stanovišta (Adžić, 2014, pp. 139-142). Sami Gedelovi tekstovi mogu u nekoj meri motivisati tu interpretaciju. Takvu ulogu pre svega ima njegovo poređenje čulnog opažanja koje nas uverava u postojanje fizičkog sveta, i takozvane *matematičke intuicije* ili opažanja istinitosti matematičkih aksioma koje bi trebalo da ima sličnu ulogu u odnosu na matematički svet opisan tim aksiomama.

1 To govori protiv Fefermanovog izjednačavanja platonističkog i ekstenzionalističkog stanovišta u filozofiji matematike (Feferman, 1985, p. 41).

Prema alternativnoj interpretaciji, Gedel ne pretpostavlja da postoji bilo kakav neposredan epistemički odnos između nas i matematičkih objekata, koji bi ličio na čulno opažanje. Ono što navodi na pogrešan zaključak mogu biti metafore kojima se on služi kako bi opisao iskustvo sticanja matematičkog znanja. Prema toj interpretaciji, koja je potkrepljena brojnim Gedelovim tvrđenjima, matematička intuicija o kojoj on govori nije usmerena na matematičke objekte, nego na pojmove pod koje ti objekti potpadaju. Matematička intuicija koja se tiče pojmova ne mora biti shvaćena kao neka posebna sposobnost kojom se stiče isključivo matematičko znanje, već se može shvatiti kao sposobnost *razumevanja* značenja matematičkih pojmova (Martin, 2005; Adžic, 2014, pp. 142-153). Razumevanje nekog pojma, kao što je *pojam skupa* ili *pojam broja*, trebalo bi da donese određeno znanje o objektima koji pod taj pojam potpadaju (cf. Gödel, 1964; 1951, pp. 320-321). Znanje koje stičemo na taj način je u nekom smislu neposredno jer nije izvedeno iz nekog drugog znanja. Moguće je da Gedel koristi termin opažanje baš kako bi naglasio tu činjenicu. Tako stečeno znanje možemo izraziti pomoću aksioma odgovarajuće formalne teorije i njenih definicija:

> „Istina je da matematički pojmovi mogu biti uvedeni pomoću pravila za baratanje simbolima, na osnovu empirijskih činjenica koje se tiču fizičkih simbola. Međutim, u stvarnosti je situacija baš suprotna: pravila korišćenja simbola su tako odabrana da izražavaju svojstva prethodno shvaćenih matematičkih pojmova ili objekata" (Gödel, 1959, p. 354, fusnota 45).

Aksiome do kojih dolazimo na taj način Gedel naziva analitičkim, pri čemu naglašava da to ne znači da su one „istinite na osnovu definicije", nego „istinite na osnovu prirode pojmova koji se u njima javljaju" (Gödel, 1951, p. 321). Takve aksiome karakteriše neka vrsta intrinsične nužnosti koja potiče od toga što su njima opisana svojstva objekata koja im pripadaju samo na osnovu značenja pojma pod koji ti objekti potpadaju. Takve aksiome imaju takozvano *unutrašnje opravdanje.*

Razumevanje matematičkih pojmova, koje nam otkriva svojstva objekata koji pod te pojmove potpadaju, često nije savršeno. Time se objašnjava nepotpunost matematičkih teorija kao što je *ZFC*, odnosno činjenica da ta teorija ne omogućava potpuni opis sveta skupova. To se pokazuje, na primer, u neodlučivosti *hipoteze kontinuuma* pomoću *ZFC* aksioma koje se mogu shvatiti kao da opisuju iterativni pojam skupa. Naše razumevanje pojmova može takođe biti neadekvatno, zbog čega možemo imati pogrešnu predstavu o osobinama objekata koji pod njih potpadaju. To se pre svega pokazuje u paradoksima. Paradoksi su, prema Gedelu, jedan od načina na koji se objektivna priroda matematičkih i logičkih objekata buni protiv njihovog proizvoljnog razumevanja (Wang, 1996, 7.4.5).

> „Naše saznanje sveta pojmova može biti jednako ograničeno i nepotpuno kao i naše saznanje sveta stvari. Ne može se poreći da je, u pojedinim slučajevima, ovo saznanje ne samo nepotpuno nego i nerazgovetno. Ovo je slučaj sa paradoksima teorije skupova, koji se često, po mom mišljenju potpuno neopravdano, smatraju pobijanjem platonizma. Naši vizuelni doživljaji ponekad protivreče našim taktilnim doživljajima, na primer u slučaju štapa uronjenog u vodu, ali niko ko je pri čistoj svesti ne bi iz ovoga zaključio da spoljašnji svet ne postoji" (Gödel, 1951, p. 321, prevod: Adžić, 2014, p. 148).

> „Nemoć prirodnog razumevanja u kontekstu matematike se naizgled zasniva na tome da su određeni pojmovi za nas nejasni, dok preostali pojmovni sistem [dat kroz „čistu intuiciju"] nije „potpun", tj. brojni problemi mogu biti formulisani koje on ne može da reši"[2] (Gödel, 2020a, p. 27).

2 „Die Hilflosigkeit des natürlichen Verstandes der Math<ematik> gegenüber beruht offenbar darauf, dass gewisse Begriffe für uns verdunkelt sind und dass das übrigbleibende [durch „reine Anschauung" gegebene] Begriffssystem nicht „vollständig" ist, d. h., viele Probleme zu formulieren gestattet, die es nicht lösen kann."

Do razumevanja matematičkih pojmova, poput *pojma skupa*, često se ne dolazi odjednom, već se ono postepeno razvija i menja kroz upotrebu. Razvoj matematičkih teorija čije aksiome i definicije izražavaju nesavršeno razumevanje nekog pojma može imati važnu ulogu u njegovom unapređenju. Na primer, izvođenjem posledica iz tih aksioma i definicija može biti rasvetljeno značenje pojma sadržanog u njima, i određene njegove dimenzije mogu doći do izražaja.

Razvoj teorije koja sadrži neko nepotpuno matematičko znanje moguć je i na drugi način. Naime, osim aksioma koje izražavaju razumevanje pojmova kojima se ta teorija bavi, i zbog toga imaju unutrašnje opravdanje, mogle bi postojati i aksiome koje, iako intuitivno nisu zasnovane na značenju tih pojmova, ipak imaju neke važne posledice kojima je njihovo korišćenje opravdano. Relevantne posledice aksioma bile bi omogućavanje nekih dokaza ili metoda dokazivanja pomoću kojih do tada neodlučive rečenice teorije postaju odlučive, ili unapređivanje prethodno postojećih dokaza, njihovo pojednostavljivanje i povezivanje (cf. Gödel, 1964, p. 261). Vrsta opravdanja koje aksiome stiču na osnovu takvih posledica naziva se *spoljašnjim opravdanjem*. Razvoj teorije i unapređenje razumevanja pojma koje te aksiome omogućavaju, može voditi tome da njihova zasnovanost na značenju tog pojma na kraju ipak bude prepoznata.

Gedelovo stanovište u filozofiji matematike, ili njegova navedena interpretacija, ukazuju na to da je on matematičko znanje video kao znanje pojmovne prirode. Ono bi tako suštinski zavisilo od naše sposobnosti da razumemo matematičke pojmove, da to razumevanje na adekvatan način izrazimo i izvedemo sve njegove posledice. Ako je to slučaj, onda nepotpunost ili neadekvatnost neke teorije potiču od nepotpunog ili neadekvatnog razumevanja pojma izraženog njenim aksiomama. Napredovanje u razumevanju matematičkih pojmova, kao što je *pojam skupa*, trebalo bi tada da omogući razvoj teorije skupova i da je učini jačom, što bi za posledicu imalo odlučivost određenih tvrđenja o skupovima koja su na osnovu njenih sadašnjih aksioma neodlučiva.

Time što bi unapredila naše razumevanje univerzalnih formalnih osobina pojmova i njihovih relacija, teorija pojmova bi mogla da doprinese i ispravnom razumevanju matematičkih pojmova, pre svega *pojma skupa*, i prepoznavanju posledica koje iz njih slede. Krajnji cilj bilo bi dostizanje ispravnog i potpunog razumevanja tog pojma koje bi trebalo da omogući i dalji razvoj teorije zasnovane na njemu. Sadašnje razumevanje pojma skupa ne omogućava potpuni opis univerzuma skupova. Čini se da je glavni razlog to što aksiome zasnovane na iterativnom pojmu skupa opisuju induktivno građenje skupova koje se ne završava, jer svaki korak u tom građenju daje osnovu za primenu sledećeg koraka (Gödel, 1951, pp. 306-307). Na taj način gradi se hijerarhija skupova koju naizgled nije moguće u celini opisati aksiomama poput onih koje pripadaju teoriji *ZFC*. Svaki skup aksioma koji opisuje neki deo hijerarhije skupova omogućava formulisanje novih aksioma, kao što su sve jače aksiome beskonačnosti, koje opisuju njene još više nivoe. Proširenje teorije skupova novim aksiomama koje opisuju sve više nivoe hijerarhije moglo bi takođe da se shvati kao napredovanje u razumevanju pojma skupa ili izvođenje novih posledica iz tog pojma. Bolje razumevanje *pojma skupa* omogućeno teorijom pojmova, moglo bi da dovede do formulisanja novih aksioma koje bi dale potpuniji opis sveta skupova. Ono bi takođe moglo da vodi formulisanju zajedničke forme svih tih aksioma kojima teorija skupova može biti proširena, čime bi i svi koraci u građenju hijerarhije skupova bili opisani na jedan nekonstruktivan način (cf. Gödel, 1946, p. 151).

Iz toga sledi da bi teorija pojmova mogla da ima značajnu ulogu u unapređenju i korigovanju matematičkog znanja. Ona bi mogla da utiče na razvoj teorije skupova i drugih matematičkih teorija koje formalizuju neko nesavršeno razumevanje matematičkih pojmova. To bi mogao da bude deo razloga zbog kog je Gedel insistirao na zasnivanju te teorije.

Ideja o teoriji pojmova može se dovesti u vezu i sa Lajbnicovom idejom o univerzalnom računu koji bi sadržao osnovne pojmove ne samo matematike, nego i nauke i filozofije, i koji bi trebalo da obezbedi formalni metod pomoću kojeg bi celokupno znanje iz tih oblasti bilo izvodivo iz aksioma koje opisuju te pojmove. Teoriju pojmova možemo razumeti kao teoriju koja bi mogla da obezbedi logičku osnovu takvog računa (za Gedelove napomene o toj Lajbnicovoj ideji vidi: Gödel, 1944, pp. 140-141).

4.2 AKSIOME TEORIJE POJMOVA

Potraga za teorijom pojmova, pre nego njenim mogućim primenama, motivisana je interesom za njen predmet koji čine objekti veoma bliski nekim fundamentalnim matematičkim objektima - skupovima, a čija formalna svojstva nisu ni blizu tako dobro shvaćena kao svojstva skupova u današnjoj matematici. Aksiome te teorije bi, poput aksioma koje se tiču skupova, trebalo da budu opravdane na osnovu razumevanja odgovarajućeg pojma - *pojma pojma* u ovom slučaju, i njegovih posledica koje se tiču formalnih osobina pojmova. Međutim, naše razumevanje tog pojma je nedovoljno razvijeno da bi nam sugerisalo aksiome pomoću kojih bi teorija pojmova bila zasnovana i koje bi, na prvom mestu, omogućile odgovarajuće rešenje intenzionalnih paradoksa. Iz tog razloga ne možemo se nadati da ćemo do zasnivanja te teorije i rešavanja intenzionalnih paradoksa doći pomoću aksioma sa unutrašnjim opravdanjem.

Kako smo nastojali da pokažemo u ovom radu, umesto na osnovu *pojma pojma*, tj. na način inherentan predmetu teorije pojmova, do odgovarajućeg rešenja intenzionalnih paradoksa možemo doći i tako što ćemo slediti način na koji su kontradikcije u osnovama paradoksa izbegnute u nekim drugim teorijama. U tom pogledu, posebno je relevantna teorija izračunljivosti, koja se bavi intenzionalnim objektima srodnim pojmovima, u odnosu na koje se takođe javlja neka vrsta samoreferencije, ali u kojoj je protivrečnost kojoj ta samoreferencija vodi uspešno izbegnuta. Način na koji je to postignuto - prepoznavanjem parcijalnosti izračunljivih funkcija, sugeriše način na koji isto može biti postignuto i u pogledu intenzionalnih paradoksa - prepoznavanjem parcijalnosti pojmova. Tako je ideja koja vodi izbegavanju paradoksa u oba slučaja ta da objekti u odnosu na koje se ti paradoksi javljaju nisu svuda definisani (i u tom smislu su parcijalni), i da to jesu li definisani za neki objekat nije uvek određeno. Ono što govori u prilog takvom rešenju intenzionalnih paradoksa jeste to što ono omogućava korišćenje samoreferencije koja, osim toga što ukazuje na važne formalne osobine pojmova, i u metodološkom smislu može doprineti razvoju teorije. Može se ispostaviti da prihvatanje parcijalnosti pojmova na kojem je to rešenje zasnovano ima i unutrašnje opravdanje sadržano u značenju *pojma pojma*. Da bismo eventualno došli do tog uvida, moramo razviti naše razumevanje pojmova i to gradeći teoriju na aksiomama sa spoljašnjim opravdanjem koje garantuju njenu neprotivrečnost. Aksiome predložene u Gilmor-Fefermanovoj teoriji, osim što formalizuju parcijalnost pojmova koja vodi izbegavanju intenzionalnih paradoksa, sugerišu i način na koji bi mogli biti opisani principi građenja pojmova i njihova ograničenja (naime, pomoću pozitivnih formula i strukture koja ih odlikuje). Time oni stiču spoljašnje opravdanje koje preporučuje Gilmor-Fefermanovu teoriju kao osnovu za izgradnju teorije pojmova.

Dalji razvoj Gilmor-Fefermanove teorije kao teorije pojmova zahtevao bi njeno proširenje aksiomama koje bi opisivale građenje pojmova pomoću logičkih veznika, i njegova ograničenja. Među aksiomama te teorije posebno važno mesto imale bi one koje bi određivale uslove identiteta dva pojma. Ti uslovi zahtevali bi, na primer, da za jednake pojmove važi da je domen njihove primene, kao i domen primene njihovih komplemenata, jednak. Tako, za razliku od skupova, pojmovi ne bi bili jednaki samo na osnovu toga što pod njih potpadaju isti objekti. Dodatni uslov da i pod njihove komplemente potpadaju isti objekti, bio bi opravdan na osnovu mogućnosti da ti pojmovi nisu svuda smisleno primenljivi. Možda ključno pitanje koje bi odredilo dalji razvoj ove teorije jeste da li su ti uslovi dovoljni za identitet neka dva pojma? Moglo bi se tvrditi da je osim toga potrebno i da ti pojmovi imaju istu strukturu, kao i da budu izgrađeni od pojmova sa istim značenjem. Uslov bi u tom slučaju bio i da su pozitivne aproksimacije formula koje izražavaju ta dva pojma, a koje pokazuju strukturu tih pojmova, jednake. U svetlu prethodnih razmatranja, još jedan uslov bi mogao doći u obzir, a to je da dati pojmovi učestvuju u izgradnji istih složenih pojmova, iskaza i dokaza. Različite opcije za uslove identiteta pojmova vodile bi zasnivanju teorija koje u manjoj ili većoj meri odstupaju od ekstenzionalnog stanovišta i u tom smislu su u manjoj ili većoj meri intenzionalne.

Ono što bi na najbolji način svedočilo o vrednosti ovog rada i ispravnosti razmatranja iznetih u njemu bio bi razvoj teorije koja koristi Gilmor-Fefermanove aksiome, a koja se na osnovu onoga što Gedel o njoj kaže može nazvati teorijom pojmova. Takva teorija dozvoljavala bi primenu pojma na samog sebe i koristila impredikativne definicije, bez opasnosti da to vodi njenoj protivrečnosti. Zahvaljujući tome, ona bi pružila filozofsko opravdanje drugih vrsta samoreferencije na koje nailazimo širom logike i matematike, ali i otvorila vrata istraživanju drugih aspekata i osobina intenzionalnog značenja.

5

LITERATURA

Aczel, P., (1988) *Non-Well-Founded Sets*, CSLI Lecture Notes, No. 14, CSLI Publications, Stanford.

Adžić, M., (2014) „Gedel o aksiomatizaciji teorije skupova", Doktorska disertacija, Univerzitet u Beogradu.

Adžić, M. & Došen, K., (2016) "Gödel's Notre Dame Course", *The Bulletin of Symbolic Logic*, Vol. 22, No. 4, pp. 469-481.

Beklemishev, L. D., (2010) "Gödel Incompleteness Theorems and the Limits of their Applicability I", *Russian Mathematical Surveys*, Vol. 65, No. 5, pp. 61-106.

Bochvar, D. A., (1981) "On a Three-Valued Logical Calculus and Its Application to the Analysis of the Paradoxes of the Classical Extended Functional Calculus", *History and Philosophy of Logic*, Vol. 2, pp. 87-112.

Boolos, G., (1993) *The Logic of Provability*, Cambridge University Press, Cambridge.

Bolzano, B., (2014) *Paradoxes of the Infinite*, Steele, D. A. (editor), Routledge and Kegan Paul, London.

Burgess, J. P., (2013) *Saul Kripke: Puzzles and Mysteries*, Polity Press, Cambridge.

Cantini, A., (1996) *Logical Frameworks for Truth and Abstraction: An Axiomatic Study*, Abramsky, S. et al. (editors), *Studies in Logic and the Foundations of Mathematics*, Vol. 135, North-Holland, Amsterdam.

Cantini, A., (2009), "Paradoxes, Self-Reference and Truth in the 20th Century", in Gabbay, D. & Woods, J. (editors), *Handbook of the History of Logic*, North-Holland, Amsterdam, pp. 875-1013.

Casti, J. L., (1996) *Five Golden Rules: Great Theories of 20th-Century Mathematics and Why They Matter*, John Wiley & Sons, New York.

Devlin, K., (2003) "The Forgotten Revolution", internet link: www.maa.org/external_archive/devlin/devlin_03_03.html.

Došen, K., (2017) "Should the *explicans* be simpler than the *explicandum*?, an abstract for the talk at the conference: Proof Theory as Mathesis Universalis, Loveno di Menaggio (Como).

Enderton, H. B., (1977) "Elements of Recursion Theory", in Barwise, J. (editor), *Handbook of Mathematical Logic*, North-Holland, Amsterdam.

Feferman, S., (1984) "Toward Useful Type-Free Theories I", *The Journal of Symbolic Logic*, Vol. 49, No. 1, pp. 75-111.

Feferman, S., (1985) "Intensionality in Mathematics", *Journal of Philosophical Logic* Vol. 14, pp. 41-55.

Feferman, S., (2011) "Axiomatizing Truth: How and Why?", Pillars of Truth Conference, math.stanford.edu/~feferman/papers/AxTruthSchwiFest.pdf.

Fraenkel, A. A., Bar-Hillel, Y. & Levy, A., (1973) *Foundations of Set Theory*, Heyting, A. et al. (editors) *Studies of Logic and the Foundations of Mathematics*, Vol. 67, North-Holland, Amsterdam.

Gilmore, P. C., (1967) "The Consistency of Partial Set Theory without Extensionality", in Jech, T. & Scott, D., (editors) *Proceedings of Symposia in Pure Mathematics*, Vol. 13, Part 2, American Mathematical Society, Providence (1974), pp. 147-153.

Gödel, K., (1931) "On Formally Undecidable Propositions", in Gödel (1986), pp. 145-195.

Gödel, K., (1934) "On Undecidable Propositions of Formal Mathematical Systems", in Gödel (1986), pp. 346-371.

Gödel, K., (1944) "Russell's Mathematical Logic", in Gödel (1990), pp. 119-141.

Gödel, K., (1946) "Remarks before the Princeton Bicentennial Conference on Problems in Mathematics", in Gödel (1990), pp. 150-153.

Gödel, K., (1951) "Some Basic Theorems on the Foundations of Mathematics and Their Implications", in Gödel (1995), pp. 304-323.

Gödel, K., (1959?) "Is mathematics a syntax of language?", in Gödel (1995), pp. 334– 356.

Gödel, K., (1964) "What is Cantor's Continuum Problem", in Gödel (1990), pp. 254-270.

Gödel, K., (1986) *Kurt Gödel Collected works I: Publications 1929 - 1936,* Feferman, S., et al. (editors), Oxford University Press, Oxford.

Gödel, K., (1990) *Kurt Gödel Collected works II: Publications 1938 - 1974*, Feferman, S., et al. (editors), Oxfrod University Press, Oxford.

Gödel, K., (1995) *Kurt Gödel Collected works III: Unpublished Essays and Lectures*, Feferman, S., et al. (editors), Oxford University Press, Oxford.

Gödel, K., (2017a) *Logic Lectures: Gödel's Basic Logic Course at Notre Dame*, Adžić, M. & Došen, K. (editors), Logical Society, Belgrade, arxiv.org/abs/1705.02601v3.

Gödel, K., (2017b) *Kurt Gödel Maxims and Philosophical Remarks. Vol. X*, Crocco, G., Van Atten, M., Cantu, P. & Engelen, E-M. (editors), hal.archives-ouvertes.fr/hal-01459188.

Gödel, K., (2020a) *Kurt Gödel Maxims and Philosophical Remarks. Vol. IX*, Crocco, G., Van Atten, M., Cantu, P. & Rollinger, R. (editors), hal.archives-ouvertes.fr/hal-02892852.

Gödel, K., (2020b) *Kurt Gödel Maxims and Philosophical Remarks. Vol. XI*, Crocco, G., Van Atten, M., Cantu, P. & Rollinger, R. (editors), hal.archives-ouvertes.fr/hal-02968031.

Gödel, K., (2021) *Kurt Gödel Maxims and Philosophical Remarks. Vol. XII*, Crocco, G., Van Atten, M., Cantu, P. & Rollinger, R. (editors), hal.archives-ouvertes.fr/hal-03142656.

Halbach, V. & Visser, A., (2014) "Self-Reference in Arithmetic I", *Review of Symbolic Logic*, Vol. 7, No. 4, pp. 671-691.

Halbach, V. & Visser, A., (2014) "Self-Reference in Arithmetic II", *Review of Symbolic Logic*, Vol. 7, No. 4, pp. 692-712.

Hilbert, D. & Bernays, P., (1968) *Grundlagen der Mathematik*, Springer, Heidelberg.

Hintikka, J., (1956) "Identity, Variables, and Impredicative Definitions", *The Journal of Symbolic Logic*, Vol. 21, No. 3, pp. 225-245.

Hintikka, J., (1957) "Vicious Circle Principle and the Paradoxes", *The Journal of Symbolic Logic*, Vol. 22, No. 3, pp. 245-249.

Hrbacek, K. & Jech, T., (1999) *Introduction to Set Theory*, Marcel Dekker, New York.

Kleene, S. C., (1936) "General Recursive Functions of Natural Numbers", *Mathematische Annalen*, Vol. 112, pp. 727-742.

Kleene S. C., (1943) "Recursive Predicates and Quantifiers", *Transactions of the American Mathematical Society*, Vol. 53, pp. 41-73.

Kleene S. C., (1950) "A Symmetric Form of Gödel's Theorem", *Indagationes Mathematicae*, Vol. 12, pp. 244-246.

Kleene, S. C., (1971) *Introduction to Metamathematics*, North-Holland, Amsterdam.

Kripke, S., (1975) "Outline of a Theory of Truth", *The Journal of Philosophy*, Vol. 72, No. 19, pp. 690-716.

Löb, M. H., (1955) "Solution of a Problem of Leon Henkin", *The Journal of Symbolic Logic*, Vol. 20, No. 2, pp. 115-118.

Łukasiewicz, J., (1970) *Jan Łukasiewicz Selected Works*, Borkowski, L. (editor), North-Holland, Amsterdam.

Martin, D. A., (2005) "Gödel's Conceptual Realism", *The Bulletin of Symbolic Logic*, Vol. 11, No. 2, pp. 207-224.

Moore, G. H. & Garciadiego, A., (1981) "Burali-Forti's Paradox: A Reappraisal of its Origins", *Historia Mathematica*, Vol. 8, pp. 319-350.

Müller-Stach, S., (2017) "Richard Dedekind: Style and Influence", *Mathematische Semesterberichte*, Vol. 64, pp. 179–186.

Peckhaus, V., (2004) "Paradoxes in Göttingen", in Link, G. (editor), *One Hundred Years of Russell's Paradox: Mathematics, Logic, Philosophy*, De Gruyter, Berlin, pp. 501-515.

Peckhaus, V. & Kahle, R., (2002) "Hilbert's Paradox", *Historia Mathematica*, Vol. 29, pp. 157-175.

Post, E. L., (1944) "Recursively Enumerable Sets of Positive Integers and their Decision Problems", *Bulletin of the American Mathematical Society*, Vol. 50, No. 5, pp. 284-316.

Quine, W. V., (1937) "New Foundations for Mathematical Logic", *The American Mathematical Monthly*, Vol. 44, No. 2, pp. 70-80.

Quine, W. V., (1955) "On Frege's Way Out", *Mind*, Vol. 64, No. 254, pp. 145-159.

Rautenberg, W., (2010) *A Concise Introduction to Mathematical Logic*, Springer, Heidelberg.

Reinhardt, W. N., (1986) “Some Remarks on Extending and Interpreting Theories with a Partial Predicate for Truth”, *Journal of Philosophical Logic*, Vol. 15, pp. 219-251.

Resnik, M. D., (1974) “The Frege-Hilbert Controversy”, *Philosophy and Phenomenological Research*, Vol. 34, No. 3, pp. 386-403.

Rogers, H. Jr., (1987) *Theory of Recursive Functions and Effective Computability*, The MIT Press, Cambridge (MA).

Rosser, B., (1936) “Extensions of Some Theorems of Gödel and Church”, *The Journal of Symbolic Logic*, Vol. 1, No. 3, pp. 87-91.

Russell, B., (1907) “On Some Difficulties in the Theory of Transfinite Numbers and Order Types”, *Proceedings of the London Mathematical Society*, S. 2, Vol. 3, pp. 29-53.

Scott, D., (1975) “Combinators and Classes”, in Böhm C. (editor), *λ-Calculus and Computer Science Theory*, Springer, Heidelberg.

Shapiro, J. H., (2016) *A Fixed-Point Farrago*, Springer, Heidelberg.

Smoryński, C., (1991) “The Development of Self-Reference: Löb’s Theorem”, in Drucker, T. (editor), *Perspectives on the History of Mathematical Logic*, pp. 110-133.

Smoryński, C., (2002) “Modal Logic and Self-Reference”, in Gabbay, D. & Guenther, F. (editors), *Handbook of Philosophical Logic* Vol. 2, D. Reidel, Dordrecht, pp. 1-53.

Smullyan, R. M., (1957) “Languages in Which Self Reference is Possible”, *The Journal of Symbolic Logic*, Vol. 22, No. 1, pp. 55-67.

Smullyan, R. M., (1978) *What is the Name of This Book: The Riddle of Dracula and Other Logical Puzzles*, Prentice-Hall, Englewood Cliffs.

Smullyan, R. M., (1980) *This Book Needs No Title: A Budget of Living Paradoxes*, Prentice-Hall, Englewood Cliffs.

Smullyan, R. M., (1984) “Fixed Points and Self-reference”, *International Journal of Mathematics and Mathematical Sciences*, Vol 7. No. 2, pp. 283-289.

Smullyan, R. M., (1994) *Diagonalization and Self-Reference*, Oxford University Press, Oxford.

Tait, W. W., (2000) "Cantor's *Grundlagen* and the Paradoxes of Set Theory", in Sher, G. & Tieszen, R. (editors), *Between Logic and Intuition: Essays in Honor of Charles Parsons*, Cambridge University Press, Cambridge, pp. 269–290.

Thomson, J. F., (1962) "On Some Paradoxes", in Butler, R. J. (editor) *Analytical Philosophy*, Barnes & Noble, New York, pp. 104-119.

van Dalen, D., (2013) *Logic and Structure*, Springer, Heidelberg.

Wang, H., (1961) "The Calculus of Partial Predicates and Its Extension to Set Theory I", *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, Vol. 7, pp. 283-288.

Wang, H., (1987) *Reflections on Kurt Gödel*, The MIT Press, Cambridge (MA).

Wang, H., (1996) *A Logical Journey: From Gödel to Philosophy*, The MIT Press, Cambridge (MA).

Yanofsky, N. S., (2003) "A Universal Approach to Self-Referential Paradoxes, Incompleteness and Fixed Points", *Bulletin of Symbolic Logic*, Vol. 9, No. 3, pp. 362-386.

BIOGRAFIJA

Jovana Kostić se rodila 1991. godine. Završila je gimnaziju u Beogradu. Studije filozofije na Filozofskom fakultetu u Beogradu upisala je 2010. godine, a završila ih 2014. godine sa prosečnom ocenom 9,94 i ocenom 10 na diplomskom ispitu. Njen diplomski rad nosi naslov *Da li kondicionali običnog jezika imaju istinosnu vrednost?* i pripada oblasti filozofije jezika.

Master studije filozofije na Filozofskom fakultetu u Beogradu upisala je 2014. godine, a završila 2015. godine sa prosečnom ocenom 10. Naslov njenog završnog master rada iz oblasti filozofije matematike je *Gedel i matematički jezik*.

Na doktorske studije filozofije upisala se 2015. godine. Predmeti koje je pohađala na doktorskim studijama su većinom logički predmeti: Teorija dedukcije, Filozofija logike I, Filozofija logike II, Filozofija nauke, Filozofija naučnog istraživanja i Supstrukturalne logike. Sve ispite položila je sa ocenom 10.

Od 2018. godine zaposlena je kao istraživač-saradnik na projektu „Dinamički sistemi u prirodi i društvu: filozofski i empirijski aspekti“ čiji je nosilac Institut za filozofiju Filozofskog fakulteta u Beogradu, a koji finansira Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije. Učestvuje u nastavi na kursevima iz logike i srodnih oblasti koji se održavaju na Odeljenju za filozofiju.

Tokom doktorskih studija provela je dva semestra na Eberhard Karls Univerzitetu u Tibingenu, Nemačka, gde je pohađala kurseve iz logike na Odeljenju za teorijsko računarstvo Fakulteta za nauku. Do sada je objavila više naučnih radova i učestvovala na domaćim i stranim konferencijama i seminarima.

Образац 5.

Изјава о ауторству

Име и презиме аутора _____________Јована Костић_____________________

Број индекса _______________________ОФ15-08___________________________

Изјављујем

да је докторска дисертација под насловом

Самореференција и теорија појмова

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

Потпис аутора

У Београду, __________________

____

Образац 6.

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора Јована Костић

Број индекса ОФ 15-08

Студијски програм Филозофија

Наслов рада Самореференција и теорија појмова

Ментор др Милош Аџић, доцент на Филозофском факултету Универзитета у Београду

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањена у **Дигиталном репозиторијуму Универзитета у Београду.**

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис аутора

У Београду, ________________________

____

Образац 7.

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Самореференција и теорија појмова

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)

2. Ауторство – некомерцијално (CC BY-NC)

3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)

4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)

5. Ауторство – без прерада (CC BY-ND)

6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.
Кратак опис лиценци је саставни део ове изјаве).

Потпис аутора

У Београду, ________________________

____

1. **Ауторство**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. **Ауторство – некомерцијално**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. **Ауторство – некомерцијално – без прерада**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. **Ауторство – некомерцијално – делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. **Ауторство – без прерада**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. **Ауторство – делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.