

Геделове теореме и примене у рачунарству

Ментор: проф. др Предраг Јаничић
Ученик: Филип Грујовић

Мај 2025

Садржај

1	Увод	2
2	Логичке основе	3
2.1	Синтакса предикатске логики	3
2.1.1	Језик предикатске логики	3
2.1.2	Слободне променљиве	5
2.2	Интерпретација у предикатској логици	5
2.2.1	Структуре	5
2.2.2	Интерпретација термова и формула	6
2.2.3	Логичке последице	7
2.2.4	Теорије и модели	8
2.3	Хилбертов систем	9
2.3.1	Аксиоме исказне логики	10
2.3.2	Докази из претпоставки	10
2.4	Пеанова аритметика	11
2.4.1	Сигнатура	11
2.4.2	Аксиоме	11
2.4.3	Битне скраћенице	12
3	Геделове теореме о непотпуности	13
3.1	Кратка историја	13
3.2	Геделово кодирање	15
3.3	Лема о фиксној тачки	16
3.4	Докази теорема	17
4	Значај Геделових теорема за рачунарство	20
4.1	Тјуринг и проблем заустављања	20
4.1.1	Тјурингове машине	20
4.1.2	Проблем заустављања	21
4.2	Черч-Тјурингова теза	22
4.3	Вештачка интелигенција	23
4.3.1	Геделове теореме, вештачка интелигенција, свест и лични ставови	24
	Додатак	25

Поглавље 1

Увод

Овај матурски рад је из области логике, па да бисмо могли да причамо о било чему унутар ње, потребно је дефинисати основне појмове математичке логике. Када будемо имали задовољавајућу основу, прећи ћемо на фокус овог рада, који су Геделове теореме о непотпуности. Након обраде њих, прећи ћемо на њихове даље импликације у области која је од средине прошлог века све релевантнија - рачунарству.

Поглавље 2

Логичке основе

У математици постоји више различитих формалних система који покушавају да нам омогуће записивање реченица „људског“ језика на математичком језику. Исказна логика нам омогућује да формализујемо реченице које је могуће исказати речима повезаним везницима и, или, не, следи и сличнима. На пример, неке од реченица које можемо превести на математички језик коришћењем исказне логике јесу: Ако данас буде облачно, падаће киша; Сутра ће или падати снег или киша; Ако оставим писмени задатак празним, добићу лошу оцену. Међутим, исказна логика се није показала довољно изражајном за неке потребе, те су математичари нашли решење у виду нешто јачег система. Предикатска логика (логика првог реда, предикатски рачун) представља одређени вид надоградње на исказну логику (која се зато понекад назива и логицом нултог реда) кроз употребу квантификатора (универзални - „за сваки“, екзистенцијални - „постоји“). Они нам омогућавају да проширимо домен реченица које можемо превести и на реченице типа: Сваки ученик који има бар једну негативну оцену неће ићи на екскурзију; Постоји кошаркаш који је нижи од просечног човека. Предикатска логика се показала довољном за аксиоматизацију многих формалних система у математици, укључујући и теорију бројева.

2.1 Синтакса предикатске логике

2.1.1 Језик предикатске логике

Да бисмо могли да правимо тврдње, прво се морамо сложити око језика који користимо. Језик предикатске логике се дели на *логички* и *нелогички* део.

Дефиниција 1. Логички део језика предикатске логике, скуп S дефинишемо као унију следећа четири скупа:

1. пребројивог скупа променљивих V ,
2. скупа логичких везника $\{\wedge, \vee, \Rightarrow, \Leftrightarrow, \neg\}$, где је $\neg$ унарни везник, а остали су бинарни,
3. скупа логичких константи $\{\top, \perp\}$,
4. скупа помоћних симбола $\{(\cdot), \cdot\}$,
- и 5. скупа квантификатора $\{\forall, \exists\}$ ¹.

Елементи скупа S су *логички симболи*.

Речник или *сигнатура* L састоји се од два највише пребројива дисјунктна скупа Σ функцијских симбола и Π предикатских симбола заједно са функцијом *арности* или функцијом која нам враћа „степен“ датог функцијског или предикатског симбола (укратко, може се схватити као мера сложености предикатског или функцијског симбола или број променљивих које он има). Функцијски симбол арности нула се назива и константом. Скупови Σ и Π чине нелогички део језика предикатске логики, па су и њихови елементи *нелогички симболи*. Сигнатуре дефинишемо у зависности од теорије коју покушавамо да опишемо, па ће се тако међу елементима сигнатуре теорије скупова могу на пример наћи симболи $\cap$ или $\cup$, а теорије бројева $+$ или $=$.

За дату сигнатуру $L = (\Sigma, \Pi, ar)$ било који низ симбола из Σ , Π или S називамо *речју* над L . Нека је скуп свих речи над L скуп A . Приметимо да, по овој дефиницији, реч не мора нужно имати „смисла“. Зато морамо увести појам *добре заснованости или формираности*.

Дефиниција 2. Скуп $T \subseteq A$ *добро формианих* термова над сигнатуром L и скупом променљивих V је најмањи подскуп A за који важи:

1. Ако је x променљива, онда $x \in T$,
2. Ако је c симбол константе, онда $c \in T$,
3. Ако је f n -арни функцијски симбол, $n \geq 1$, и свако t_i , $i \in 1, \dots, n$ је терм, тада је и $f(t_1, \dots, t_n) \in T$.

Дефиниција 3. Скуп $W \subseteq A$ *добро формианих* формула над сигнатуром L и скупом променљивих V је најмањи подскуп A за који важи:

1. Ако је Q нуларни предикатски симбол, онда $Q \in W$,
2. Ако је P n -арни предикатски симбол, $n \geq 1$, и свако t_i , $i \in 1, \dots, n$ је термин, тада је и $P(t_1, \dots, t_n) \in W$,
3. $\top, \perp \in W$,
4. Ако $\alpha, \beta \in W$, тада и $(\neg\alpha), (\alpha \wedge \beta), (\alpha \vee \beta), (\alpha \Rightarrow \beta), (\alpha \Leftrightarrow \beta) \in W$,
5. Ако $\alpha \in W$ и x је променљива, тада су и $(\forall x \alpha)$ и $(\exists x \alpha) \in W$.

¹Скуп квантификатора, као ни скуп логичких везника у овом случају није минималан, јер се сви чланови ових скупова могу дефинисати кроз нпр. $\neg$, $\vee$ и $\forall$, али се због прегледности и остали користе као подразумевани.

2.2. ИНТЕРПРЕТАЦИЈА У ПРЕДИКАТСКОЈ ЛОГИЦИ

Формуле изграђене од неког од прва 3 правила су *атомићке*, а остале су *сложене*.

2.1.2 Слободне променљиве

Дефиниција 4. *Слободно и везано* појављивање променљиве у формули дефинишемо овим правилима:

1. Свако појављивање променљиве у атомичкој формули је слободно.
2. Свако појављивање променљиве које је слободно у A је слободно и у $\neg A$, а свако које је везано у A је везано и у $\neg A$.
3. Свако појављивање променљиве које је слободно у A или у B је слободно и у било којој њиховој комбинацији направљеној бинарним логичким везницима, а свако које је везано у A или у B је везано и у било којој њиховој комбинацији направљеној бинарним логичким везницима.
4. Свака променљива која је квантификована у A је везана у сваком свом појављивању у A , као и у свом појављивању уз квантификатор.

Дефиниција 5. Променљива је слободна у формули ако има слободно појављивање у тој формули. Аналогно је и променљива је везана у формули ако има везано појављивање у тој формули.

2.2 Интерпретација у предикатској логици

Сада када смо јасно дефинисали слова и речи језика који користимо, време је да те речи добију неко значење. Увешћемо појам *структуре*, *валуације* и *интерпретације*, затим тиме дефинисати шта значи да је нека реченица *логичка последица* неке друге и на крају ћемо коначно стећи могућност да опишемо неку теорију језиком логики првог реда.

2.2.1 Структуре

Да бисмо интерпретирали формуле логики првог реда, потребно је да знамо три ствари:

1. У којем скупу се интерпретирају променљиве.
2. Како се интерпретирају основни константски и функцијски симболи.
3. Како се интерпретирају основни предикатски симболи.

Дефиниција 6. Фиксирајмо сигнатуру L . Структура U за дату сигнатуру састоји се од следећих података:

1. Непразан скуп $|U|$, који се назива *носач* структуре U . Елементи скупа U другачије називају и *објекти*.

2. За сваки n -арни функцијски симбол f , функција

$$f^U : |U|^n \rightarrow |U|,$$

која се назива *интерпретација* симбола f . Функција f^U мора бити дефинисана за све улазе из $|U|^n$.

3. За сваки n -арни предикатски симбол Q , релација

$$Q^U \subseteq |U|^n,$$

која се назива *интерпретација* симбола Q .

2.2.2 Интерпретација термова и формула

Сада желимо формално да дефинишемо шта значи да нека структура задовољава неку реченицу.

Да бисмо то учинили, морамо бити у могућности да интерпретирамо термове и формуле у датој структури. Прво уводимо појам валуације променљивих. Фиксирамо неки језик под сигнатуром L . Подсетимо се да ознаком V означавамо скуп променљивих, а ознаком T скуп термова тог језика.

Дефиниција 7. *Валуација* у структури U је пресликавање $s : V \rightarrow |U|$ које свакој променљивој језика додељује елемент из носача структуре.

Ако је s валуација, $x \in V$ променљива, а $a \in |U|$ неки објекат, тада ознаком $s(x \mid a)$ означавамо нову валуацију s' која се понаша као s , осим што x додељује вредност a :

$$s'(z) = \begin{cases} a & \text{ако } z = x, \\ s(z) & \text{ако } z \neq x. \end{cases}$$

Свака валуација $s : V \rightarrow |U|$ може се рекурзивно проширити на интерпретацију $\bar{s} : T \rightarrow |U|$ термова језика, и то на следећи начин:

$$\begin{aligned} \bar{s}(x) &= s(x), \\ \bar{s}(f(t_1, \dots, t_n)) &= f^U(\bar{s}(t_1), \dots, \bar{s}(t_n)) \quad \text{за } n > 0. \end{aligned}$$

Нека је U структура, $s : V \rightarrow |U|$ валуација и φ формула. Тада дефинишемо када структура U *задовољава* формулу φ под валуацијом s , што означавамо са $U \models \varphi[s]$, помоћу следећих рекурзивних правила ²:

²Строго формално гледано, горња дефиниција се не уклапа у стандардну шему рекурзије, јер дефинишемо предикат (а не функцију) рекурзивно. Формалније би било најпре дефинисати функцију model_U која сваком упарену формуле φ и валуације s додељује истинитосну вредност, тј. $\text{model}_U(\varphi, s) \in \{\top, \perp\}$, а затим се дефинише $U \models \varphi[s] \iff \text{model}_U(\varphi, s) = \top$.

$$\begin{aligned}
 U \models P(t_1, \dots, t_n)[s] &\iff (\bar{s}(t_1), \dots, \bar{s}(t_n)) \in P^U, \\
 U \models \top[s] &\text{унавек,} \\
 U \models \perp[s] &\text{никада,} \\
 U \models (\neg\varphi)[s] &\iff \text{није тачно да } U \models \varphi[s], \\
 U \models (\varphi_1 \wedge \varphi_2)[s] &\iff U \models \varphi_1[s] \text{ и } U \models \varphi_2[s], \\
 U \models (\varphi_1 \vee \varphi_2)[s] &\iff U \models \varphi_1[s] \text{ или } U \models \varphi_2[s], \\
 U \models (\varphi_1 \rightarrow \varphi_2)[s] &\iff \text{ако } U \models \varphi_1[s], \text{ онда } U \models \varphi_2[s], \\
 U \models (\forall x \varphi)[s] &\iff \text{за свако } a \in |U|, U \models \varphi[s(x \mid a)], \\
 U \models (\exists x \varphi)[s] &\iff \text{постоји } a \in |U| \text{ такво да } U \models \varphi[s(x \mid a)].
 \end{aligned}$$

2.2.3 Логичке последице

Дефиниција 8. Нека је Γ скуп добро формираних формула, и нека је φ добро формирана формула. Кажемо да Γ *логички имплицира (или повлачи)* φ , што означавамо са $\Gamma \models \varphi$, ако за сваку структуру U и сваку валуацију s у U важи: ако је $U \models \psi[s]$ за сваку $\psi \in \Gamma$, онда важи и $U \models \varphi[s]$.

Ради прегледности, углавном се користи скраћена нотација: $\psi_1, \dots, \psi_n \models \varphi$ уместо $\{\psi_1, \dots, \psi_n\} \models \varphi$, и $\models \varphi$ уместо $\emptyset \models \varphi$. Дакле, $\models \varphi$ значи да свака структура и свака валуација задовољавају формулу φ . У том случају кажемо да је формула φ *ваљана*.

Две формуле φ и ψ су *логички еквивалентне* ако важи и $\varphi \models \psi$ и $\psi \models \varphi$.

Пример 1. Покажимо да важи $\forall x Qx \models Qy$. Нека су U и s произвољна структура и произвољна валуација такви да важи $U \models \forall x Qx[s]$. Тада, на основу дефиниције интерпретације универзалног квантификатора, за свако $a \in |U|$ важи $U \models Qx[s(x \mid a)]$, односно $a \in Q^U$. Конкретно, $s(y) \in Q^U$, па важи $U \models Qy[s]$. Пошто су U и s били произвољни, закључујемо да $\forall x Qx \models Qy$.

Пример 2. Покажимо да не важи $Qy \models \forall x Qx$. Довољно је навести пример структуре и валуације за које лева страна важи, а десна не. Нека је $|U| = \{0, 1\}$ и $Q^U = \{0\}$. Нека је s валуација таква да је $s(y) = 0$. Тада важи $U \models Qy[s]$, али не важи $U \models \forall x Qx[s]$, јер на пример $s(x \mid 1) = 1 \notin Q^U$.

Напомена. Приметимо да су ови докази слични доказивању помоћу истинитосних табли у исказној логици: да бисмо доказали да логичка последица важи, морамо размотри све структуре и све валуације; док да бисмо показали да логичка последица *не* важи, довољно је дати један контра-пример. Међутим, за разлику од истинитосних табли које су коначне, структуре у предикатској логици не морају бити коначне, па не постоји очигледан метод за одлучивање да ли нека последица важи или не. (У ствари, може се доказати да такав метод не постоји ни не-очигледно, и баш то је на неки начин и фокус овог рада.)

2.2.4 Теорије и модели

Прво примећујемо да интерпретација реченице не зависи од валуације s .

Лема 1. 1. Ако су s и s' валуације које се слажу на слободним променљивима терма t , тада важи $\bar{s}(t) = \bar{s}'(t)$.

2. Ако су s и s' валуације које се слажу на слободним променљивима формуле φ , тада важи: $U \models \varphi[s]$ ако и само ако $U \models \varphi[s']$.

Доказ: Доказујемо индукцијом по грађи термова t .

База индукције: Ако је t променљива x , тада по дефиницији интерпретације термова имамо:

$$\bar{s}(x) = s(x), \quad \bar{s}'(x) = s'(x)$$

По претпоставци, $s(x) = s'(x)$ (јер се слажу на слободним променљивима термова), па следи:

$$\bar{s}(x) = \bar{s}'(x)$$

Индуктивни корак: Нека је $t = f(t_1, \dots, t_n)$, где је f функционални симбол арности n , а $t_1, \dots, t_n$ су подтермови.

По дефиницији интерпретације термова:

$$\bar{s}(t) = f^U(\bar{s}(t_1), \dots, \bar{s}(t_n)) \quad \text{и} \quad \bar{s}'(t) = f^U(\bar{s}'(t_1), \dots, \bar{s}'(t_n))$$

По индуктивној претпоставци, важи $\bar{s}(t_i) = \bar{s}'(t_i)$ за све $i = 1, \dots, n$, јер су сви t_i мањи термови од t , и s и s' се слажу на слободним променљивима свих t_i .

Дакле,

$$\bar{s}(t) = f^U(\bar{s}(t_1), \dots, \bar{s}(t_n)) = f^U(\bar{s}'(t_1), \dots, \bar{s}'(t_n)) = \bar{s}'(t).$$

■

Последица. Ако је σ реченица и U структура, тада истинитосна вредност израза $U \models \sigma[s]$ не зависи од валуације s (тј. или важи за све валуације, или ни за једну).

Дефиниција 9. Кажемо да је структура U модел реченице σ ако важи $U \models \sigma[s]$ за неку валуацију s . Пошто ова особина не зависи од s (према претходној последици), скраћено пишемо $U \models \sigma$.

Слично, ако је Σ скуп реченица, тада кажемо да је U модел скупа Σ , и то означавамо са $U \models \Sigma$.

Дефиниција 10. Нека је U структура. Скуп свих реченица које су задовољене у U назива се *теорија* структуре U , и означава се са $\text{Th}(U)$.

Општије, ако је C класа структура, скуп свих реченица које су задовољене у свакој структури $U \in C$ назива се *теорија класе* C , и пише се $\text{Th}(C)$.

Можемо посматрати и обрнути смер: ако имамо неки скуп реченица Σ , можемо посматрати класу структура које задовољавају Σ . Ова класа се означава као $\text{Mod}(\Sigma)$. У том контексту, реченице из Σ називају се *аксиомама*, и кажемо да је класа $\text{Mod}(\Sigma)$ *аксиоматизована* скупом Σ .

Дефиниција 11. Кажемо да је класа C *аксиоматибилна* ако постоји скуп реченица Σ такав да је $C = \text{Mod}(\Sigma)$. Кажемо да је класа C *коначно аксиоматибилна* ако постоји реченица σ таква да је $C = \text{Mod}(\sigma)$.

Класа која је коначно аксиоматибилна назива се такође елементарна класа. Општија аксиоматибилна класа назива се елементарном класом у ширем смислу.

Напомена. Пошто важи $\text{Mod}(\sigma_1, \dots, \sigma_n) = \text{Mod}(\sigma_1 \wedge \dots \wedge \sigma_n)$, свака класа која је аксиоматизабилна коначним скупом реченица аутоматски је аксиоматибилна једном реченицом. Дакле, дефинисањем елементарне класе као класе коју аксиоматизује једна реченица не губимо на општости.

Приметимо да за сваку структуру U и сваку реченицу σ важи:

$$\sigma \in \text{Th}(U) \iff U \models \sigma \iff U \in \text{Mod}(\sigma).$$

2.3 Хилбертов систем

Један од аспеката предикатске логике чине дедукциони системи. Иако их постоји мноштво, у овом раду бавићемо се Хилбертовим системом. Хилбертов систем дефинише мање логичких симбола него што их дефинише језик који смо раније описали, али се сви они могу додефинисати као „скраћенице” зарад прегледнијег записа, па ћемо и даље користити све симболе из скупа S каквог смо дефинисали у поглављу 2.2.1.

Хилбертов систем се састоји од језика и (I) скупа аксиома (добро-формираних формула) које називамо *аксиомама*, и (II) скупа *правила извођења*. *Правило извођења* се састоји од: (i) листе формула које називамо *претпоставке правила*, и (ii) формуле која се назива *закључак правила*.

Обично правила извођења немају више од две претпоставке.

Доказ у датом дедуктивном систему је коначни низ формула:

$$a_1, a_2, \dots, a_n$$

у којем је свака формула a_i или аксиом, или изведена из претходних формула a_j помоћу правила извођења.

Низ је *доказ формуле* a , ако је a последњи члан низа.

Дефиниција 12. Формула је *доказива* у датом систему ако постоји њен доказ. Свака формула која се појављује у доказу је доказива, јер сваки почетни сегмент доказа до те формуле (укључујући њу) такође представља валидан доказ.

2.3.1 Аксиоме исказне логике

(i)

$$a \Rightarrow (b \Rightarrow a)$$

(ii)

$$[a \Rightarrow (b \Rightarrow c)] \Rightarrow [(a \Rightarrow b) \Rightarrow (a \Rightarrow c)]$$

(iii)

$$(\neg a \Rightarrow \neg b) \Rightarrow (b \Rightarrow a)$$

Додају се додатне аксиоме ако се користе и други везници као примитивни. Уколико се сви други везници дефинишу помоћу $\neg$ и $\Rightarrow$, тада нису потребне додатне аксиоме.

Правило извођења: Модус поненс

Ако имамо:

$$a \Rightarrow b \quad \text{и} \quad a$$

тада можемо закључити:

$$b$$

Деф 2.12. Реченица a исказне логике је теорема ако је доказива у горе наведеном систему. Означава се као $\vdash a$.

2.3.2 Докази из претпоставки

Доказ из скупа Γ (добро формираних формула) је низ формула у којем је свака формула:

- логичка аксиома (у овом контексту: аксиома исказне логике),
- члан скупа Γ (претпоставка), или
- изведена из претходних формула применом правила извођења — у овом случају, **модус поненс**.

Такав низ представља доказ последње формуле у низу. Формула је **доказива из скупа Γ** ако постоји њен доказ из Γ . Означавамо то као:

$$\Gamma \vdash a$$

Ако скуп Γ садржи формуле $a_1, a_2, \dots, a_n$, тада пишемо:

$$a_1, a_2, \dots, a_n \vdash a$$

Користимо и запис Γ, a да означимо скуп претпоставки који укључује све чланове скупа Γ и формулу a .

Теорема о дедукцији гласи:

$$\Gamma, a \vdash b \iff \Gamma \vdash a \Rightarrow b$$

Доказ теореме се изводи преко индукције.

За потребе овог рада не морамо залазити дубље у овај систем.

2.4 Пеанова аритметика

Сада када коначно имамо довољно јаку логичку потпору, можемо почети да се бавимо аксиоматизацијом теорије бројева. Као и било коју другу област у математици, теорију бројева је могуће аксиоматизовати на више начина, међу којима су неки еквивалентни, а неки не. Иако је касније показано да их је могуће генерализовати и за слабије системе (као што је на пример Робинсонова аритметика), Геделове теореме су доказане за систем који се зове **Пеанова аритметика**, и из тог разлога ћемо и ми користити тај систем.

2.4.1 Сигнатура

Пеанова аритметика користи сигнатуру L :

1. константу 0
2. функцијски симбол S арности 1 (функција „следбеник“)
3. функцијске симболе $+$ и $*$ арности 2
4. предикатски симбол $=$ арности 2.

У суштини сви ови симболи значе управо оно што сви знамо да значе и не захтевају неко објашњење, осим евентуално функције s . Преко ње и нуле заправо уводимо различите бројеве, па је број 2 формално $s(s(0))$, а број 5 $s(s(s(s(s(0)))))$, а бројеви које иначе користимо ће бити само „скраћенице“.

2.4.2 Аксиоме

1. $(\forall x)(\forall y)(\forall z)(x = y \Rightarrow (x = z \Rightarrow y = z))$
2. $(\forall x)(\forall y)(x = y \Rightarrow s(x) = s(y))$
3. $(\forall x)(\neg(0 = s(x)))$
4. $(\forall x)(\forall y)(s(x) = s(y) \Rightarrow x = y)$
5. $(\forall x)(x + 0 = x)$
6. $(\forall x)(\forall y)(x + s(y) = s(x + y))$
7. $(\forall x)(x \cdot 0 = 0)$
8. $(\forall x)(\forall y)(x \cdot s(y) = (x \cdot y) + x)$
9. (**Аксиома индукције**) $(\Phi(0, \bar{y}) \wedge ((\forall x)(\Phi(x, \bar{y}) \Rightarrow \Phi(s(x), \bar{y})))) \Rightarrow (\forall x)\Phi(x, \bar{y})$,

за сваку формулу Пеанове аритметике $\Phi(x, \bar{y})$ где је $\bar{y}$ скраћеница за $y_1, \dots, y_k$.

У Пеановој оригиналној формулацији, аксиома индукције је дефинисана у језику логики другог реда (зато што квантификује предикате, што у логици првог реда није дозвољено):

Ако је K скуп такав да је $0 \in K$ и за сваки $n \in \mathbb{N}$, $n \in K \Rightarrow S(n) \in K$, тада су сви природни бројеви у K .

Међутим, ми у логици првог реда као нешто слабију алтернативу можемо узети управо индукцијску шему која је дата у аксиоми 9. Важно је имати у виду да је ова индукцијска шема слабија од оригиналне формулације, јер је број скупова природних бројева непребројив, док је број предикатских симбола у логици првог реда пребројив, па постоје скупови који не могу бити описани у језику логики првог реда.

Скуп аксиома генерисаних шемом 9 је *рекурзивно набројив*. Ригорозна дефиниција рекурзивне набројивости би превише оптеретила овај рад, али укратко, скуп је рекурзивно набројив ако постоји алгоритам³ који ће у неком тренутку излистати произвољан члан скупа; могуће са понављањима и могуће без краја. Ако се неки елемент налази у скупу, он ће сигурно у неком тренутку бити излистан, али ако није, никад нећемо бити сигурни да неће бити излистан. Скуп свих формула првог реда Φ је рекурзивно пребројив јер се могу систематски излистати (а не могу се систематски проверити), а за сваку формулу Φ може се механички генерисати одговарајућа аксиома. Поред тога што може бити рекурзивно набројив, скуп може бити и рекурзиван, што значи да за произвољан објекат можемо у коначној временској јединици са сигурношћу проверити да ли он датом скупу припада или не. Сви рекурзивни скупови су и рекурзивно набројиви, али обрнуто не важи.

Ако је скуп аксиома неке теорије рекурзивно набројив, за њу се може рећи да је *рекурзивно аксиоматизована*.

2.4.3 Битне скраћенице

У Пеановој аритметици ради прегледности уводимо скраћенице. Предикатски симболи $<$ и $\leq$ арности два дефинишу се на следећи начин:

$$u < v \Leftrightarrow \exists x(u + s(x) = v),$$

$$u \leq v \Leftrightarrow u < v \vee u = v.$$

Нумерали су скраћенице за термове који садрже више угнеждених „следбеник” функција и дефинишу се на следећи начин: $\bar{1} = S(0)$, $\bar{2} = S(S(0))$ итд. Нумерали представљају природне бројеве и њихове интеракције са операцијама ће бити управо онакве какве очекујемо.

³Формалније се може користити и Тјурингова машина; о њој ће бити речи у поглављу 4

Поглавље 3

Геделове теореме о непотпуности

Геделове теореме о непотпуности се често сматрају за најзначајније и најлепше резултате математичке логике XX века. У овом одељку упознаћемо се са околностима под којим су настале, објаснићемо идеју иза њих и на крају извести њихов формалан доказ у систему који смо описали у одељку 2 (који ћемо надаље називати теоријом T).

3.1 Кратка историја

Све је почело покушајима да се механизује процес закључивања. Чак и стари Грци су знали да је логичко резонување, бар делимично, процес којем владају одређени закони. Аристотел и Еуклид су поставили своје постулате још у IV веку пре нове ере (Еуклид геометрију, а Аристотел логичко закључивање, тзв. силогизме), међутим, након тога је било потребно да прођу стотине година пре него што је дошло до даљих помака у изучавању аксиоматизације математике.

Једно од већих достигнућа математике XIX века јесте откриће да постоји више „могућих” геометрија, где се на геометрију односи теорија о апстрактним појмовима права и тачака. Пре тога је владао консензус да је Еуклидова геометрија апсолутна истина, и да, упркос томе што може бити мањих пропуста у Еуклидовој репрезентацији, они су небитни и да ће сваки даљи напредак у геометрији бити управо у његовом систему. Откриће различитих геометрија пољуљало је идеју да математика истражује реалан свет: како је могуће да у једној реалности може постојати више врста „тачака” и „правих”? Касније у XIX веку, логичари **Џорџ Бул** и **Август де Морган** значајно надмашују Аристотела у постулирању дедуктивних метода закључивања. Истовремено, немачки математичар **Давид Хилберт** (1862–1943) ради на стриктнијој формализацији геометрије од Еуклидове, Ђузепе Пеано и

Готлоб Фреге раде на комбиновању формалног закључивања и теорије бројева, а **Георг Кантор** поставља теорију скупова, која упркос својој користности, отвара многе проблеме, међу којима је можда најистакнутији *Раселов парадокс*¹. **Бертранд Расел** и **Алфред Вајтхед** верују да је те проблеме мало комплекснијим системом могуће избећи и у раним 1910-им објављују чувену књигу „*Principia Mathematica*“, у којој наизглед решавају све досадашње несавршености. Њен циљ је био да у потпуности и без контрадикција садржи подлогу за целу математику, али нико заправо није могао да докаже ниједно од та два тврдње.

Хилбертов програм

На **Интернационалном конгресу математичара у Паризу 1900.** године Хилберт је поставио своја 23 питања (данас позната као „Хилбертова питања“). „Од ових, друго је било питање доказивања конзистентности „Пеанових аксиома“, на којима је, како је (Хилберт) показао, зависила ригорозност математике.“² Баш тај други проблем Хилберт поново представља 1928. године на конгресу у Болоњи. Поставио је три питања:

1. Да ли је математика *комплетна*?
2. Да ли је математика *конзистентна*?
3. Да ли је математика *одлучива*³?

Хилберт је био оптимистичан да су одговори на сва ова питања потврдни.

Међутим, на Хилбертово изненађење, само две године касније аустријски логичар **Курт Гедел** објављује доказ као одричан одговор на прва два питања и 1931. објављује рад „*О формално неодлучивим исказима Principia Mathematica и сродних система I*“ у коме се налазе докази за две Геделове теореме о непотпуности. Хилберт се у почетку оштро супротстављао, али је, како није могао да математички обори доказ, на крају ипак морао да прихвати истину. Геделове теореме нису биле крај негативним одговорима на његова питања. Амерички математичар, логичар и филозоф **Алонцо Черч** 1935. године објављује свој резултат да се интуитивни појам *ефективно израчуњливе* функције може формализовати *генералним рекурзивним функцијама* или еквивалентно такозваним *ламбда-рачуном* и доказује да проблем заустављања („Halting problem“) у ламбда-рачуна није ефективно израчуњлив. Годину дана касније, 1936, Черч и **Алан Тјуринг** независно један од другог објављују доказе да је одговор на треће Хилбертово питање негативан, при чему је Черч користио свој појам ламбда-рачуна, а Тјуринг, информатичар у срцу, уводи појам „израчунавања машином“. Овим резултатима Хилбертов програм, и заједно са њом идеја да „*Principia Mathematica*“ савршено описује целу математику, пада у воду.

¹ Нека је R скуп свих скупова који не садрже самог себе; тада ако R садржи себе он не може садржати себе и обрнуто

² Ендру Ходжс: „Алан Тјуринг: Енигма“

³ Овим појмом се нећемо бавити али је изузетно битан за логичаре.

3.2 Геделово кодирање

У самој сржи Геделових доказа налази се чувени *Епименидов парадокс*. У њему Епименид, који је и сам Крићанин, изјављује: „Сви Крићани су лажови.“, или поједностављено: „Ова реченица је лаж.“ Очигледно је где овде настаје парадокс, и када бисмо могли да преведемо ову реченицу на математички језик, имали бисмо парадокс и у самој математици. Ово, међутим, није толико очигледно. Прво нам треба начин да кажемо:

„реченица x није доказива“,

што бисмо језиком који смо раније дефинисали могли да изведемо. Проблем настаје када бисмо уместо x опет хтели да заменимо ту реченицу тако да она изгледа овако:

„реченица 'реченица x није доказива' није доказива.“

Математика једноставно не садржи наводнике или било шта налик њима. Оно чега се Гедел досетио јесте да дефинише функцију која свакој реченици додељује бројчану вредност, такозвану *Геделову функцију*. Пример такве функције је

$$g_e(x) = \prod_{i=1}^n pn(i)^{1+g_s(s_i)}$$

где је $pn(i)$ и-ти прост број а $g_s(x)$ бијективна функција која слика симболе $s_1, s_2, \dots, s_k$ из скупа $L \cup V$ у природне бројеве, и место симбола у реченици, а n број симбола у реченици. Очигледно је да је оваква функција 1-1, тј. сваке две различите функције ће имати различиту слику, али није бијекција јер не мора (и неће) сваки природан број имати своју реченицу. За реченицу A вредност $g_e(A)$ можемо писати и као $[A]$ и ту вредност називамо *Геделовим кодом*.

Дефинишемо предикатске симболе:

$Ded(x, y)$ ако је x код доказа формуле чији је код y ;

$Pr(x) = (\exists y)(Ded(y, x))$;

$Cons(T) = \neg Pr([0 = S(0)])$ у T . (користићемо $\perp$ уместо $0 = S(0)$ као и $Pr([0 = S(0)])$ ради прегледности)

Што се тиче конзистентности теорије, она подразумева да у њој нису истовремено доказиве тврдње A и $\neg A$. Гедел је у свом првобитном доказу користио нешто јаче тврдње које се назива ω -конзистентност⁴. Теорија је ω -конзистентна ако она, ни за једно P , **не доказује**

$$\neg \exists x P(x) \text{ и истовремено, за свако } n, P(n).$$

Једноставније, систем који је ω -конзистентан не каже „Ниједан број нема својство P^k “, а да такође каже „Сваки број га има, један по један“. Сваки

⁴Слово омега се у математичкој логици и теорији скупова користи као ознака за први ординал - број који долази после свих природних бројева. Овде је изабрано јер се конзистентност односи на грешку која укључује квантфиковање над свим природним бројевима.

ω -конзистентан систем је конзистентан, али обрнуто не важи, што значи да је ω -конзистентност јаче тврђење од конзистентности. Међутим, 1936. године **Џон Баркли Розер** нешто другачијом конструкцијом Геделове реченице заобилази појам ω -конзистентности и доказује Геделове теореме са слабијом претпоставком обичне конзистентности.

Претпоставићемо да у теорији T важе и следећи услови:

- H1: $\vdash A \Rightarrow \vdash Pr(\lceil A \rceil)$
- H2: $\vdash Pr(\lceil A \rceil) \Rightarrow \vdash Pr(\lceil Pr(\lceil A \rceil) \rceil)$
- H3: $\vdash Pr(\lceil A \Rightarrow B \rceil) \Rightarrow \vdash (Pr(\lceil A \rceil) \Rightarrow Pr(\lceil B \rceil))$

Свака теорија која садржи аксиоме Пеанове аритметике и Хилбертовог система довољно је изражајна за дефинисање ових појмова и самим тим Геделове теореме важе за њих.

3.3 Лема о фиксној тачки

Вратимо се на реченицу „реченица 'реченица x није доказива' није доказива." Уведимо функцију $\text{sub}(x, y)$ која у формули са кодом x замењује све слободне променљиве са вредношћу y и као бројчану вредност даје код те нове реченице. За сваку формулу A и број n (нумерал, али смо већ утврдили да је практично исто) важи

$$\text{sub}(\lceil A(x) \rceil, n) = \lceil A(n) \rceil.$$

Реченица са кодом $\text{sub}(\lceil A(x) \rceil, n)$ заправо тврди да реченица са кодом n има својство A . Нека је $\Psi(x)$ произвољна формула. Желимо да конструишемо реченицу Φ са кодом $\lceil \Phi \rceil$ која за реченицу са кодом $\lceil \Phi \rceil$ (за себе) тврди да има својство Ψ . Ако је

$$n = \lceil \Psi(\text{sub}(x, x)) \rceil \quad \text{и} \quad \Phi = \Psi(\text{sub}(n, n)),$$

тада имамо

$$\Phi = \Psi(\text{sub}(n, n)) = \Psi(\text{sub}(\lceil \Psi(\text{sub}(x, x)) \rceil, n)) = \Psi(\lceil \Psi(\text{sub}(n, n)) \rceil) = \Psi(\lceil \Phi \rceil),$$

па је

$$\vdash (\Phi \Leftrightarrow \Psi(\lceil \Phi \rceil)),$$

тј. Φ је таква реченица. Из произвољности функције $\Psi(x)$ следи да тврдња важи за сваку функцију коју је могуће дефинисати у теорији која садржи T . За потребе Геделових теорема није од значаја изглед ове реченице, већ само њено постојање које смо њеном конструкцијом доказали.

Лема о фиксној тачки се тако назива јер се у њој тражи фиксна тачка (вредност a за функцију $f(x)$ таква да је $f(a) = a$) за функцију $\Psi(\lceil x \rceil)$. Често се може чути и назив „Дијагоналациона лема“ због своје сличности са чувеним „Канторовим дијагоналним аргументом“, или пак самореференцијална лема јер реченица Φ прави тврдње које се односе на њу саму.

3.4 Докази теорема

Прва Геделова теорема о непотпуности односи се на постојање реченице која је семантички тачна али недоказива као теорема у теорији T .

Прва Геделова теорема о непотпуности. Ако теорија T садржи Пеанову аритметику, ако је рекурзивно аксиоматизована и конзистентна, онда је она и непотпуна; постоје искази језика T који се не могу доказати, као ни њихове негације.

Доказ. Из леме о фиксној тачки знамо да ће за функцију $\neg Pr(x)$ постојати нека реченица G која је њена фиксна тачка тј.

$$G \Leftrightarrow \neg Pr(G).$$

Прво претпоставимо $\vdash G$. Тада

1. $\vdash G \Rightarrow \vdash Pr(\lceil G \rceil)$,
2. $\vdash G \Rightarrow \vdash \neg Pr(\lceil G \rceil)$, из избора реченице G ,

а како је теорија T по тврдњи теореме конзистентна, претпоставка мора бити нетачна, па

$$\not\vdash G.$$

Даље, претпоставимо $\vdash \neg G$. Тада имамо

1. $\neg G \Rightarrow Pr(\lceil \neg G \rceil)$,
2. $Pr(\lceil \neg G \rceil) \Rightarrow Pr(\lceil G \Rightarrow (0 = S(0)) \rceil)$,
3. $Pr(\lceil G \Rightarrow \perp \rceil) \Rightarrow (Pr(\lceil G \rceil) \Rightarrow \perp)$.

па је из конзистентности теорије $T \vdash \neg G \Rightarrow \neg Pr(\lceil G \rceil)$ (*), а како из избора G имамо

$(\neg Pr(\lceil G \rceil) \Rightarrow G) \Rightarrow (\neg Pr(\lceil G \rceil) \Rightarrow G)$, па из тога и (*) имамо $\vdash \neg G \Rightarrow G$, што је из конзистентности теорије T контрадикција, па је претпоставка погрешна, што значи $\not\vdash \neg G$.

Показали смо да реченица G , као ни њена негација нису теореме теорије T , што је и био захтев. ■

Наизглед логичан начин да се „заобиђе“ прва Геделова теорема јесте да се G дода као аксиома теорије T ; она нам неће утицати на конзистентност теорије, а проблем неодлучивости реченице G је решен тиме што смо је „вештачки“ одлучен. Нажалост (или на срећу), Геделова теорема је отпорна на овај тип напада јер се она након додатка не односи на првобитну теорију T , већ на нову теорију $T' = T \cup G$ (која јер садржи теорију T садржи довољно аритметике), а онда ће у теорији T' постојати реченица G' која је неодлучива. Очигледно, чим у било коју теорију додамо њену неодлучиву реченицу G као аксиому, та нова теорија ће опет имати своју неодлучиву реченицу G' .

Остаје још једна нада за систем отпоран на прву Геделову теорему: ако смо у поглављу 2.4.2. при дефинисању аксиома Пеанове аритметике у дефиницији аксиоме 9 смели да користимо аксиоматску шему, то јест да Пеанову аритметику поставимо тако да има бесконачно много аксиома, чини се да можемо увести још једну аксиоматску шему: „За сваку теорију T , нека је њена Геделова реченица G_T аксиома.“ Ова шема ипак није валидна, и то из више разлога. Прво, квантфиковали смо над свим теоријама T , што у логици првог реда не можемо да радимо. Друго, скуп свих Геделових реченица није рекурзивно набројив, па ни теорија у којој смо дефинисали шему није рекурзивно аксиоматизована.

Интересантно је приметити да реченица G , упркос томе што је недоказива у теорији T , свакако исказује тачну тврдњу, јер она за себе управо тврди да је недоказива. То нам јасно показује могућу несагласност између онога што је *тачно* и онога што је *доказиво*. Може се (нетачно) наслућивати да је горе описана констукција једина таква да даје неодлучиву реченицу. Међутим, поред овако конструисане реченице G , крајем осамдесетих година прошлог века Џорџ Бул искористио је формализацију *Беријевог парадокса* уместо Епименидовога. Беријев парадокс је самореферентни парадокс који произилази из исказа сличног следећем: „Нека је x најмањи број ког није могуће дефинисати у осамдесет карактера.“ Тада се она никако не може односити на број на који се односи јер га управо она дефинише, и то за мање од 80 карактера.

Друга Геделова теорема о непотпуности. Ако теорија T садржи Пеанову аритметику, ако је рекурзивно аксиоматизована и конзистентна, она не може у својим оквирима доказати своју конзистентност.

Доказ. Изаберимо реченицу G на исти начин као у доказу прве теореме. Кад бисмо показали $\vdash \text{Cons}(T) \Rightarrow \vdash G$, како из прве теореме имамо $\nvdash G$ имали бисмо и $\nvdash \text{Cons}(T)$ (*).

1. $\text{Pr}(\lceil G \rceil) \Rightarrow \text{Pr}(\text{Pr}(\lceil G \rceil))$
2. $G \Rightarrow \neg \text{Pr}(\lceil G \rceil)$ из избора G
3. $\text{Pr}(\lceil G \rceil) \Rightarrow \neg G$ из 2 по својству које је доказиво у теорији T
4. $\text{Pr}(\lceil \text{Pr}(\lceil G \rceil) \Rightarrow \neg G \rceil)$

5. $Pr(\lceil Pr(\lceil G \rceil) \rceil) \Rightarrow Pr(\lceil \neg G \rceil)$
6. $Pr(\lceil G \rceil) \Rightarrow Pr(\lceil \neg G \rceil)$
7. $\neg G \Rightarrow (G \Rightarrow \perp)$, својство доказиво у теорији T
8. $Pr(\lceil \neg G \Rightarrow (G \Rightarrow \perp) \rceil)$
9. $Pr(\lceil \neg G \Rightarrow (G \Rightarrow \perp) \rceil) \Rightarrow (Pr(\lceil \neg G \rceil) \Rightarrow Pr(\lceil G \Rightarrow \perp \rceil))$
10. $Pr(\lceil \neg G \rceil) \Rightarrow Pr(\lceil G \Rightarrow \perp \rceil)$ из 8, 9 и правила modus ponens
11. $Pr(\lceil G \rceil) \Rightarrow Pr(\lceil G \Rightarrow \perp \rceil)$ из 6 и 10
12. $Pr(\lceil G \Rightarrow \perp \rceil) \Rightarrow (Pr(\lceil G \rceil) \Rightarrow \perp)$
13. $Pr(\lceil G \rceil) \Rightarrow (Pr(\lceil G \rceil) \Rightarrow \perp)$
14. $Pr(\lceil G \rceil) \Rightarrow \perp$
15. $\neg(\perp) \Rightarrow \neg Pr(\lceil G \rceil)$ из 14 следи користећи својство доказиво у теорији T
16. $Cons(T) \Rightarrow \neg Pr(\lceil G \rceil)$ из 15 по дефиницији Cons(T)
17. $\neg Pr(\lceil G \rceil) \Rightarrow G$ из избора G
18. $Cons(T) \Rightarrow G$

па из тога и (*) следи тврдња теореме. ■

Можемо повући паралелу између друге Геделове теореме и правног система: да ли суд заиста може доказати да су све његове процедуре фер користећи само те процедуре?

Поглавље 4

Значај Геделових теорема за рачунарство

Геделове теореме о непотпуности, упркос томе што су првобитно формулисане у контексту формалне математичке логики, имају последице и ван поља чисте математике. Како модерни рачунари прате неки математички модел (наравно то довољно јак да се Геделове теореме односе на њега) оне нам показују да ће механизација закључивања употребом рачунара, колико год моћна била, увек имати одређене границе. Интерпретиране у контексту *израчуњљивости*, теоретске тврдње у математици се осликавају на конкретна ограничења онога што програми и алгоритми могу да постигну. Геделове теореме су директно инспирисале резултате као што је Тјурингов чувени „Halting problem” - „Проблем заустављања” и дискусије о могућностима аутоматских доказивача теорема и вештачке интелигенције.

4.1 Тјуринг и проблем заустављања

Амерички логичар Алан Тјуринг 1936. у свом раду поставља питање „Да ли постоји програм који ће, за дати програм а и његов улаз х, моћи да одреди да ли ће програм а имати излаз (тј. да ли ће упасти у бесконачну петљу). По узору на Гедела, Тјуринг формулише формални систем који у математици треба да имитира неки реалан процес, само што уместо људског језика који је Гедел имитирао, Тјуринг моделира програме. Тјуринг формализује концепт програма или алгоритма појмом *аутоматске машине*, данас познате као *Тјурингова машина*.

4.1.1 Тјурингове машине

Тјурингове машине се састоје од:

1. Бесконачне меморијске траке подељене на ћелије, од којих свака садржи неки симбол из скупа V ;
2. Читача који може да прочита вредност са ћелије на којој се налази и да се помера лево или десно по траци;
3. Коначног скупа интерних стања S , укључујући једно почетно стање и скупа F могућих завршних стања;
4. Коначног скупа правила (програма) који управљају машином; формалније функција $f : (S \setminus F) \times V_i \rightarrow S \times V_i \times \{L, D\}$, где је V_i симбол на неком пољу i и L и D упутство да ли идемо улево или удесно (на поље V_{i-1} или V_{i+1}).

Тјурингове машине, као и друге формализације алгоритама (као што је и раније поменути Черчов лямбда-рачун) су еквивалентне, како међусобно, тако и многим програмским језицима који су данас у употреби; оне могу извршити све програме које модеран језик може.

4.1.2 Проблем заустављања

Доказ¹ проблема заустављања налачи доказу прве Геделове теореме: Нека је $h(p, x)$ *ефективно израчунљива функција*² која за дати програм p и (као што смо у доказима Геделових додељивали свакој реченици јединствен број, тако и овде сваки програм добија свој јединствен „Тјурингов код“, па је p у ствари код неког програма, а не сам тај програм) теоремама његов улаз (или n -торку његових улаза) x враћа вредност 1 ако програм p враћа излаз за $\text{input } x$, а 0 у супротном. Желимо да покажемо да таква функција h не постоји.

За сваку бинарну (има два улаза) ефективно израчунљиву функцију f , парцијална (не враћа излаз за све улазе) постоји функција g :

$$g(x) = \begin{cases} 0 & \text{ако } f(x, x) = 0, \\ \text{недефинисана} & \text{у супротном.} \end{cases}$$

Како су сви подпроцеси функције g израчунљиви (из претпоставке да је f израчунљива функција), постоји програм (алгоритам) са кодом e који рачуна функцију g .

Претпоставимо да је $f(e, e) = 0$. Тада је $g(e) = 0$. У овом случају програм g се зауставља па је по дефиницији функције h $h(e, e) = 1$. Даље, ако узмемо $f(e, e) \neq 0$, па је $g(e)$ недефинисано. Тада се програм e не зауставља за улаз e , па је $h(e, e) = 0$. Из произвољности функције f важи да је свака бинарна ефективно израчунљива функција h различита од f , па функција h дефинисана као у претпоставци не може постојати.

¹У тексту је дат само нацрт доказа.

²Израчунљиве функције су формализовани аналог интуитивног појма алгоритама, у смислу да је функција израчунљива ако постоји алгоритам који може да обави посао функције, тј. за дат улаз функција увек може да врати одговарајући излаз.

Овај резултат има далекосежне последице у пракси: он нам говори да не можемо направити савршен дебагер за произвољан програм: увек ће постојати нека бесконачна петља у неком програму коју произвољан код за откривање бесконачних петљи неће моћи да открије. Такође, не може се за произвољна два програма доказати да су еквивалентни: никад нећемо моћи да за уназад инжењерисану функцију докажемо да је иста оригиналу, или за оптимизовану верзију произвољног кода да је заправо еквивалентна оригиналу. Аналогно, али више узнемирујуће, програм који савршено детектује сваки рачунарски вирус би морао да реши проблем заустављања, што значи да савршен антивирус не може постојати.

4.2 Черч-Тјурингова теза

Средином тридесетих година двадесетог века логичари су се, након решавања питања о комплетности и конзистентности, окренули ка појму *израчунљивости*. Да би дефинисали шта тачно значи за функцију када за њу кажемо да је израчунљива, многи математичари су постављали своје системе: Тјуринг своје аутоматске машине, Черч ламбда-рачун, сам Гедел заједно са **Жаком Хербрандом** формализовао је дефиницију класе рекурзивних функција... Сваки од ових система је био дефинисан независно један од другог, међутим Черч, Тјуринг и Клин заједно доказују да су ове три класе функција једнаке: функција је Тјуринг-израчунљива ако и само ако је ламбда-израчунљива, и ако и само ако је рекурзивна. Овај резултат је навео математичаре на веровање да је сам концепт израчунљивости прецизно карактерисан са било којим од ова три различита процеса. Било би крајње чудно да су независно развијени системи „нетачни“ или „непотпуни“ на исти начин. Даљи покушаји формалног описивања израчунљивих функција само су даље охрабрили овај став.

Иако су Черч и Тјуринг оригинално мало другачије (и различито један од другог!) формулисали своју тезу, она у суштини гласи овако:

„Свака функција која се може израчунати механичком процедуром пратећи јасно дефинисана правила (тј. алгоритмом) је израчунљива и у интуитивном смислу.“ За разлику од ранијег доказа са Клином, ова тврдња Черча и Тјуринга тврди да су раније поменути системи не само међусобно еквивалентни, већ еквивалентни и са **неформалном идејом израчунљивости**. Важно је истаћи да Черч-Тјурингова теза није теорема, управо јер се бави неформалним појмом израчунљивости, већ више наличи дефиницији баш тог појма; она представља мост између формалне и неформалне израчунљивости. Гедел је за тезу прокоментарисао: „Тачну дефиницију механичке израчунљивости је, без сумње, поставио Тјуринг.“ Временом су се појавиле бројне варијације на Черч-Тјурингову тезу, па тако сада имамо **Физичку Черч-Тјурингову тезу**: „Све физички израчунљиве функције су и Тјуринг израчунљиве“; **Изоморфичку Черч-Тјурингову тезу**: „Не само што за сваку функцију израчунљиву

алгоритмом постоји Тјурингова машина која је рачуна, на неком нивоу комплексности постоји кореспонденција (изоморфизам) између алгоритма и корака Тјурингове машине ³"; **Јаку Черч-Тјурингову тезу** или **Черч-Тјуринг-Дојчову тезу**: „Универзум је еквивалентан Тјуринговој машини, дакле израчунавање нерекурзивних функција је физички немогуће“; **Черч-Тјурингову тезу теорије комплексности** и многе друге.

4.3 Вештачка интелигенција

Черч-Тјурингова теза нам говори да дигитални рачунари могу симулирати сваки могући вид јасно дефинисаног математичког резонувања људи.

У последњих неколико година присуствовали смо скоро невероватном напретку у области вештачке интелигенције. У кратком временском периоду, она је од нечега чиме су се само поједини бавили постала алат којим на дневној бази барата велики број људи за разне потребе: омогућила је разне пробоје у медицини (који се, на пример, тичу савијања протеина) и грађевини и другим наукама, олакшала (убрзала) је послове програмерима али и значајно унапредиле неке ненаучне дисциплине као што је шах, спорт (аутоматско препознавање прекршаја у неким спортовима, спортска статистика, чак и „тренери“ који препознају непожељне шаблоне у игри и сугеришу поправке)... Усред популаризације вештачке интелигенције намећу се многа питања: „Која су ограничења вештачке интелигенције?“ и „Да ли ће вештачка интелигенција икад моћи да заправо парира нашој, људској интелигенцији?“.

Многи системи вештачке интелигенције, посебно системи који се експлицитно баве логиком (на пример, аутоматски доказивачи теорема) су саграђени по моделу предикатске логике, па би то, због тога што су ти модели подложни Геделовим теоремама, значило да ће увек бити истина до којих они неће моћи да досегну; постојаће тачне тврдње које они неће моћи да дођу. Једно од могућих решења за овакве проблеме јесте такозвано *хијерархијско или мета-резонување*. Оно би подразумевало да у архитектуру система који резонује у некој логици додамо још један систем (или више система?) на нивоу изнад првобитног који ће резонovati о тој логици и имати одређен увид у њена ограничења. Такође, многи модели за вештачку интелигенцију нису детерминистички (вођени јасно дефинисаним правилима) већ пробабилистички (вођени вероватноћом), па тако Геделове теореме за њих не важе (они, међутим, имају друге проблеме што се тиче сигурности; ако се у општем случају на детерминистички систем не можемо ослонити због Геделових теорема, на пробабилистички не можемо по дефиницији).

³Ова верзија изложена је у књизи Дагласа Хофстадера: "Гедел, Ешер, Бах: бесконачна Златна Нит".

4.3.1 Геделове теореме, вештачка интелигенција, свест и лични ставови

Остаје, ипак, питање да ли ће „електронски мозак“ икад моћи да заиста мисли. Одговор на њега зависи од тога да ли наш биолошки мозак сматрамо своје врском „машином“ или не. Чувени енглески математичар-физичар **Роџер Пенроуз** даје следећи аргумент: Када би наш ум био формални систем, био би подложен ефектима Геделових теорема. Али, ми имамо моћ да „искочимо“ из система и интуитивно сагледамо истину. Дакле, људски ум није алгоритам и као такав се од стране алгоритма не може тачно симулирати. Међутим, многи теоретичари ВИ се не слажу са овим ставом. **Марвин Мински** и **Даглас Хофстадер** верују да, упркос томе што људски мозак изгледа као да надмашује формалност, ипак може бити само изузетно флексибилан, самомодификујући формални систем. Хофстадер људски мозак пореди са мрављом колонијом: мрави, на неки начин као и неурони нашег мозга, функционишу потпуно несвесно и детерминистички, али из саме сложености њихових међусобних интеракција израђа се оно што ми називамо „свест“⁴. Само зато што се нама чини да увиђамо „трансценденталну истину“, не значи да је она заправо трансцендентална, већ да је можда само посматрамо са вишег нивоа. Ниједна страна нема чврсте доказе за своје тврдње, па расправа за сада остаје отворено питање пре свега филозофског карактера.

Лично бих, иако Хофстадеров и Минскијев аргумент звучи потпуно разумно, ипак стао на страну Пенроузовог холистичког става који нам говори да је свест више од пуког збира својих делова. Желим да верујем да имамо неку срж, есенцију, суштину коју жице и струја не могу тек тако да емулирају.

⁴У претходном поделјку поменуто је мета-резонovanje; Хофстадер тврди да свест настаје онда када се нивои измешају и крену неконтролисано да утичу једни на друге па тиме и сами на себе

Додатак

Курт Гедел

Курт Гедел (Kurt Gödel) је рођен 1906. године у граду Brünn у Аустроугарској (данашње Брно у Чешкој) у немачкој породици. Са 18 година придружио се свом брату на Универзитету у Бечу, где је, упркос првобитној намери да студира теоријску физику, присуствовао курсевима који су се тicali примарно математике и филозофије. Након тога је изучавао теорију бројева, али када је присуствовао семинару о књизи Бертранда Расела „Увод у математичку филозофију“ заинтересовао се за математичку логику. Гедел је, након што је присуствовао Хилбертовом предавању у Болоњи где је овај изнео питања о структурном интегритету логице, изабрао управо та питања за свој докторски рад, који је 1929. године завршио, и у коме се, између осталог, налазе докази двеју теорема о непотпуности. Докторат му је додељен 1930. године и његов рад је тада објавила Бечка академија наука.

Од 1933. године Гедел је почео често да путује у Сједињене Америчке Државе, где је држао предавања и упознао Ајнштајна. Услед Хитлеровог доласка на власт у Немачкој 1933. године и растућег утицаја нацизма у Аустрији, Гедел је развио симптоме параноје и почео је да једе само јела која је његова жена Аделе скувала, из страха да не буде отрован. Након Аншлуса 1938. Геделова титула на универзитету је укинута и одбијена му је пријава за еквивалентну, пре свега због своје повезаности са јеврејским математичарима (првенствено Ханом који му је био ментор за докторски рад). На пролеће 1938. године у Принстону је објавио своје доказе за консистентност аксиоме избора и хипотезе континуитета, што су два резултата од изузетног значаја за теорију скупова. Након почетка Другог светског рата 1939. године Гедел бежи из Беча у Принстон да би избегао регрутацију. Током свог пута ка САД (који, узгред, није био преко Атлантика већ транс-сибирском железницом до Пацифика па бродом до Западне обале), Гедел је носио тајно писмо упућено Ајнштајну од једног беачког физичара да упозори Америчког председника Рузвелта на могућност да Хитлер развија нуклеарну бомбу, али га није уручио Ајнштајну (који је ипак већ био обавештен) јер је био убеђен да нацисти нису били способни за тако нешто. Гедел и Ајнштајн су истовремено живели

у Принстону и држали су веома високо мишљење један о другом. Постоји анегдота да је Ајнштајн једно време ишао на посао само да би у повратку шетао са Геделом.



Гедел је 1953. постао професор на Институту за Напредне студије у Принстону, где је до краја живота предавао. За време боравка на Институту, почео је да се бави и филозофијом и физиком. Ајнштајну је као поклон за његов 70. рођендан показао егзактна решења за његове једначине поља. Након што му је жена 1977. године хоспитализована Гедел је престао да једе. Био је тежак мање од 30 килограма када је у јануару 1978. године преминуо у Принстону.

Алан Тјуринг

Рођен 1912. године у Лондону, Алан Тјуринг је већ у основној школи од директорке препознат као изузетан таленат за математику. Са 13 година уписао је школу у Шерборну, где нису ценили његове способности за природне науке. Ипак, тамо је развио блиско пријатељство са вршњаком Кристофером Морком, који је често описиван и као Тјурингова прва љубав. Иако је њихов однос инспирисао Тјуринга, нагло је завршен Моркомљевом смрћу, која је на младог Алана оставила дубоку жалост. Након што је матурирао у Шерборну, добио је стипендију на Краљевом колеџу на Кембриџу где је 1935. године и дипломирао. За време свог мастер студија на истом месту, Тјуринг се бавио одлучивошћу проблема и 1936. године, отприлике у исто време као и Алонцо Черч, објављује свој одговор

на Хилбертово питање о одлучивости. У том раду је први пут дефинисана и Тјурингова машина. Неки називају овај Тјурингов рад и „најутицајнијим радом у историји математике". Од 1936. до 1938. Тјуринг је већину времена провео учећи од Черча на Принстону.



За време Другог светског рата Алан Тјуринг је премештен у **Блечли Парк**, криптографску истраживачку станицу у близини Лондона. Тамо је, заједно са тимом пажљиво одабраних научника, добио задатак да дешифрује немачку технологију за шифровање порука: чувену „Енигму". Решење тог задатка 1941. године значајно је допринело победи савезничких снага у Борби за Атлантик и тиме практично окончао немачко присуство у Атлантику, што је омогућило неометано снабдевање Британије и тиме драстично утицало на даљи ток рата. Исте године се верио за колегиницу из Блечли Парка Џоан Кларк, али њихова веза није била дуговечна; након што јој је признао своју хомосексуалност, одлучио је да откаже брак. Након рата Тјуринг је наставио рад у области рачунарства и значајно је допринео развоју модерних рачунара. Године 1951. се окренуо математичкој биологији и објавио рад „Хемијска база морфогенезе" у коме је предвидео осцилирајуће хемијске реакције годинама пре него што су први пут експериментално изведене.

Упркос својим заслугама у Британији је Алану Тјурингу 1952. суђено за „велику непристојност" због његовог сексуалног опредељења, када је као алтернативу затвору изабрао хормонску терапију. Умро је 1954. године од тровања цијанидом и и даље није разјашњено да ли је била реч о случајном тровању или самоубиству.

Литература

- [1] Predrag Janičić, *Matematička logika u računarstvu*, elektronsko izdanje, 2008.
- [2] Douglas Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*, Basic Books, 1979.
- [3] Peter Smith, *An introduction to Gödel's theorems*, Cambridge University Press, 2007.
- [4] Kurt Gödel, *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I (O формално неодлучивим исказима „Principia Mathematica“ u повезаних система I)*, Monatshefte für Mathematik, 1931.
- [5] Alan Turing, *On Computable Numbers, with an Application to the Entscheidungsproblem*, Proceedings of the London Mathematical Society, Series 2, Volume 42, 1936.
- [6] Roger Penrose, *The Emperor's New Mind*, Oxford University Press, 1989.
- [7] J.R. Lucas, *Minds, Machines and Gödel*, Cambridge University Press, 1961.
- [8] Peter Selinger, *Lecture Notes 3: The Language of First-Order Logic*, 2012.
- [9] Википедија, слободна енциклопедија.