Kapitola 1

Nerozhodnutelnost a neúplnost

V této, závěrečné kapitole se budeme zabývat tím, jak lze s teoriemi pracovat algoritmicky. Zlatým hřebem budou *Gödelovy věty o neúplnosti* z roku 1931, které ukazují limity formálního přístupu, a které zastavily desetiletí trvající program formalizace matematiky. Nemáme zde dostatek prostoru k uvedení formálních definic a úplných důkazů, proto se místy budeme pohybovat na poněkud intuitivní úrovni. Zaměříme se na pochopení smyslu tvrzení a myšlenek důkazů.

Pojem *algoritmu* budeme chápat také jen intuitivně. Pokud bychom ho chtěli formalizovat, potom nejběžnější (ale zdaleka ne jedinou) volbou je koncept *Turingova stroje*.[1]

1.1 Rekurzivní axiomatizace a rozhodnutelnost

V důkazových systémech, kterými jsme se zabývali (tablo metoda, rezoluce, hilbertův kalkulus) jsme povolili, aby teorie T, ve které dokazujeme, byla nekonečná. Vůbec jsem se ale zatím nezabývali tím, jak je zadaná. Pokud chceme ověřit, že je daný objekt (tablo, rezoluční strom, posloupnost formulí) korektním důkazem, potřebujeme nějaký algoritmický přístup ke všem axiomům T.

Jednou z možností by bylo požadovat *enumerátor* T, tj. algoritmus, který vypisuje na výstup axiomy z T, a každý axiom někdy vypíše.[2] Potom by bylo snadné potvrdit, že je daný důkaz korektní. Pokud bychom ale dostali důkaz, který použil chybný axiom, který v T není, nikdy bychom se to nedozvěděli: nekonečně dlouho bychom čekali, zda jej enumerátor přeci jen nevypíše. Požadujeme proto silnější vlastnost, která umožňuje rozpoznat i chybné důkazy *rekurzivní axiomatizaci*:[3]

Definice 1.1.1 (Rekurzivní axiomatizace)**.** Teorie T je *rekurzivně axiomatizovaná*, pokud existuje algoritmus, který pro každou vstupní formuli φ doběhne a odpoví, zda $\varphi \in T$.

Poznámka 1.1.2. Ve skutečnosti by nám stačil enumerátor pro T, pokud by bylo garantováno, že vypisuje axiomy v lexikografickém uspořádání. To už je ekvivalentní rekurzivní axiomatizaci. (Rozmyslete si proč.)

[1] Viz přednáška NTIN090 Základy složitosti a vyčíslitelnosti.

[2] Nutným předpokladem je, aby T byla spočetná. K tomu stačí předpokládat, že jazyk je spočetný.

[3] Slovo *rekurzivní* zde neznamená běžně známou rekurzi, ale odkazuje na formalizaci algoritmu pomocí 'rekurzivních funkcí' od Gödela. Rekurzivní funkce zde znamená totéž, co vyčíslitelná nějakým Turingovým strojem, a teorii vyčíslitelnosti (*computability theory*) se někdy také říká *recursion theory*.

Zaměříme se na otázku, zda můžeme v dané teorii T 'algoritmicky rozhodovat pravdu' (tj. platnost vstupní formule). Pokud ano, říkáme, že je teorie *rozhodnutelná*. To je ale poměrně silná vlastnost, definujeme proto také *částečnou rozhodnutelnost*, která znamená, že pokud formule platí, algoritmus nám to řekne, ale pokud neplatí, nikdy se nemusíme dočkat odpovědi.

Definice 1.1.3 (Rozhodnutelnost)**.** O teorii T říkáme, že je

- *rozhodnutelná*, pokud existuje algoritmus, který pro každou vstupní formuli φ doběhne a odpoví, zda $T \models \varphi$,
- *částečně rozhodnutelná*, pokud existuje algoritmus, který pro každou vstupní formuli:
  - pokud $T \models \varphi$, doběhne a odpoví "ano",
  - pokud $T \not\models \varphi$, buď nedoběhne, nebo doběhne a odpoví "ne".

Můžeme jako obvykle předpokládat, že φ v definici je sentence. Ukážeme si jednoduché tvrzení:

Tvrzení 1.1.4. *Nechť T je rekurzivně axiomatizovaná. Potom:*

(i) T je částečně rozhodnutelná,

(ii) je-li T navíc kompletní, potom je rozhodnutelná.

Důkaz. Algoritmem ukazujícím částečnou rozhodnutelnost je konstrukce systematického tabla pro Fφ.[4] Pokud φ v T platí, konstrukce skončí v konečně mnoha krocích a snadno ověříme, že je tablo sporné, jinak ale skončit nemusí.

Je-li T kompletní, víme, že $T \vdash \varphi$ právě když $T \nvdash \varphi$. Budeme tedy paralelně konstruovat tablo pro Fφ a tablo pro Tφ (důkaz a zamítnutí φ z T): jedna z konstrukcí po konečně mnoha krocích skončí. □

1.1.1 Rekurzivně spočetná kompletace

Požadavek kompletnosti je příliš silný, ukážeme, že stačí pokud jsme schopni efektivně popsat všechny kompletní jednoduché extenze.[5]

Definice 1.1.5 (Rekurzivně spočetná kompletace)**.** Řekneme, že teorie T má *rekurzivně spočetnou kompletaci*, pokud (nějaká) množina až na ekvivalenci všech jednoduchých kompletních extenzí teorie T je *rekurzivně spočetná*, tj. existuje algoritmus, který pro danou vstupní dvojici přirozených čísel (i, j) vypíše na výstup i-tý axiom j-té extenze (v nějakém pevně daném uspořádání[6]), nebo odpoví, že takový axiom už neexistuje.[7]

Tvrzení 1.1.6. *Pokud je teorie T rekurzivně axiomatizovaná a má rekurzivně spočetnou kompletaci, potom je T rozhodnutelná.*

[4] Zde nám stačí enumerátor axiomů T, nebo postupně generujeme všechny sentence (např. v lexikografickém pořadí) a pro každou testujeme, zda je axiomem.

[5] Tj. 'všechny modely až na elementární ekvivalenci'.

[6] Zde potřebujeme, aby byl jazyk spočetný.

[7] Jeli extenzí méně než j, nebo má-li j-tá extenze méně než i axiomů.

Důkaz. Pro danou sentenci φ buď $T \vdash \varphi$, nebo existuje protipříklad $\mathcal{A} \models \varphi$, tedy kompletní jednoduchá extenze T_i teorie T taková, že $T_i \nvdash \varphi$. Z kompletnosti ale plyne, že $T_i \vdash \neg\varphi$. Náš algoritmus bude paralelně konstruovat tablo důkaz φ z T a (postupně) tablo důkazy $\neg\varphi$ ze všech kompletních jednoduchých extenzí $T_1, T_2, \ldots$ teorie T.[8] Víme, že alespoň jedno z paralelně konstruovaných tabel je sporné, a můžeme předpokládat, že konečné (neprodlužujeme-li sporné větve tabla), tedy algoritmus ho po konečně mnoha krocích zkonstruuje. □

Cvičení 1.1. Ukažte, že následující teorie mají rekurzivně spočetnou kompletaci:

- Teorie čisté rovnosti (prázdná teorie v jazyce $L = \langle\rangle$ s rovností),
- Teorie unárního predikátu (prázdná teorie v jazyce $L = \langle U \rangle$ s rovností, kde U je unární relační symbol),
- Teorie hustých lineárních uspořádání DeLO* (kompletní jednoduché extenze jsou popsané v Důsledku **??**),

Jde o rekurzivně axiomatizované teorie (neboť jsou konečné), jsou tedy rozhodnutelné.

Příklad 1.1.7. Na závěr uveďme bez důkazu několik dalších příkladů rozhodnutelných teorií:

- Teorie Booleových algeber (Alfred Tarski 1940),
- Teorie algebraicky uzavřených těles (Tarski 1949),
- Teorie komutativních grup (Wanda Szmielew 1955).

Tyto teorie jsou také nekompletní, ale rekurzivně axiomatizované a mají rekurzivně spočetnou kompletaci.

1.1.2 Rekurzivní axiomatizovatelnost

V předchozí kapitole, konkrétně v Sekci **??**, jsme se zabývali otázkou, kdy lze popsat nějakou třídu struktur [resp. teorii] pomocí axiomů [určitého tvaru]. Nyní se zaměřme na otázku, kdy to lze udělat *algoritmicky*.

Definice 1.1.8 (Rekurzivní axiomatizovatelnost)**.** Třída L-struktur $K \subseteq \mathrm{M}_L$ je *rekurzivně axiomatizovatelná*, pokud existuje rekurzivně axiomatizovaná L-teorie T taková, že $K = M_L(T)$. Teorie T' je *rekurzivně axiomatizovatelná*, pokud je rekurzivně axiomatizovatelná třída jejích modelů, neboli pokud je T' ekvivalentní nějaké rekurzivně axiomatizované teorii.

Poznámka 1.1.9. Podobně bychom mohli definovat *rekurzivně spočetnou axiomatizovatelnost*.

Ukažme si následující jednoduché tvrzení:

Tvrzení 1.1.10. *Je-li $\mathcal{A}$ konečná struktura v konečném jazyce s rovností, potom je teorie* $\mathrm{Th}(\mathcal{A})$ *rekurzivně axiomatizovatelná.*

Poznámka 1.1.11. Z toho plyne i že $\mathrm{Th}(\mathcal{A})$ je rozhodnutelná, což ale není překvapivé: platnost sentence φ v konečné struktuře $\mathcal{A}$ můžeme snadno ověřit.

[8] Nevadí, že je jich nekonečně mnoho, můžeme využít tzv. *dovetailing*: Provedeme 1. krok konstrukce 1. tabla, potom 2. krok 1. tabla a 1. krok 2. tabla, 3. krok 1. tabla, 2. krok 2. tabla, 1. krok 3. tabla, atd.

Důkaz. Očíslujme prvky domény jako $A = \{a_1, \ldots, a_n\}$. Teorii $\mathrm{Th}(\mathcal{A})$ lze axiomatizovat jedinou sentencí, která je tvaru "existuje právě n prvků $a_1, \ldots, a_n$ splňujících právě ty *základní vztahy* o funkčních hodnotách a relacích, které platí ve struktuře $\mathcal{A}$".[9] □

Uveďme několik standardních příkladů struktur, které lze 'algoritmicky popsat':

Příklad 1.1.12. Pro následující struktury je $\mathrm{Th}(\mathcal{A})$ rekurzivně axiomatizovatelná, a tedy i rozhodnutelná:

- $\langle \mathbb{Z}, \leq \rangle$, jde o tzv.teorii *diskrétních lineárních uspořádání*,
- $\langle \mathbb{Q}, \leq \rangle$, jde o teorii DeLO,
- $\langle \mathbb{N}, S, 0 \rangle$, teorie *následníka s nulou*,
- $\langle \mathbb{N}, S, +, 0 \rangle$, *Presburgerova aritmetika*,
- $\langle \mathbb{R}, +, -, \cdot, 0, 1 \rangle$, teorie *reálně uzavřených těles*,[10]
- $\langle \mathbb{C}, +, -, \cdot, 0, 1 \rangle$, teorie *algebraicky uzavřených těles charakteristiky 0*.

Důsledek 1.1.13. *Pro struktury uvedené v Příkladu 1.1.12 platí, že* $\mathrm{Th}(\mathcal{A})$ *je rozhodnutelná.*

Poznámka 1.1.14. Jak ale vyplývá z První Gödelovy věty o neúplnosti (viz níže), teorie *standardního modelu aritmetiky*, tj. struktury $\underline{\mathbb{N}} = \langle \mathbb{N}, S, +, \cdot, 0, \leq \rangle$ *nemá* rekurzivně axiomatizovatelnou teorii.

1.2 Aritmetika

Vlastnosti přirozených čísel hrají důležitou roli nejen v matematice, ale například také v kryptografii. Připomeňme, že jazyk aritmetiky je jazyk $L = \langle S, +, \cdot, 0, \leq \rangle$ s rovností. Jak jsme zmínili v Poznámce 1.1.14, tzv. *standardní model aritmetiky* $\underline{\mathbb{N}} = \langle \mathbb{N}, S, +, \cdot, 0, \leq \rangle$ nemá rekurzivně axiomatizovatelnou teorii. Proto používáme rekurzivně axiomatizované teorie, které se snaží vlastnosti $\underline{\mathbb{N}}$ popsat částečně; těmto teoriím říkáme *aritmetiky*.

1.2.1 Robinsonova a Peanova aritmetika

Uvedeme jen dva nejdůležitější příklady aritmetik: *Robinsonovu* a *Peanovu*.

Definice 1.2.1 (Robinsonova aritmetika)**.** *Robinsonova aritmetika* je teorie Q v jazyce aritmetiky sestávající z následujících (konečně mnoha) axiomů:

$$\begin{array}{ll} \neg S(x) = 0 & x \cdot 0 = 0 \\ S(x) = S(y) \to x = y & x \cdot S(y) = x \cdot y + x \\ x + 0 = x & \neg x = 0 \to (\exists y)(x = S(y)) \\ x + S(y) = S(x + y) & x \leq y \leftrightarrow (\exists z)(z + x = y) \end{array}$$

[9] Například, pokud $f^{\mathcal{A}}(a_4, a_2) = a_{17}$, přidáme do konjunkce atomickou formuli $f(x_{a_4}, x_{a_2}) = x_{a_{17}}$ (kde x_{a_i} jsou proměnné odpovídající jednotlivým prvkům). A pokud $(a_3, a_3, a_1) \in R^{\mathcal{A}}$, přidáme $R(x_{a_3}, x_{a_3}, x_{a_1})$.

[10] Tento významný výsledek A. Tarského (1949) také znamená, že lze algoritmicky rozhodovat, které vlastnosti platí v Euklidovské geometrii.

Robinsonova aritmetika je velmi slabá, nelze v ní dokázat například komutativitu ani asociativitu sčítání či násobení, nebo tranzitivitu uspořádání.

Na druhou stranu v ní lze dokázat všechna *existenční tvrzení o numerálech*, která jsou pravdivá v $\underline{\mathbb{N}}$. Tím myslíme formule, které v prenexním tvaru mají pouze existenční kvantifikátory, a do kterých jsme za volné proměnné substituovali *numerály* $\underline{n} = S(\dots S(0)\dots)$.

Příklad 1.2.2. Například, pro formuli $\varphi(x, y)$ tvaru $(\exists z)(x+z = y)$ je $Q \vdash \varphi(\underline{1}, \underline{2})$, kde $\underline{1} = S(0)$ a $\underline{2} = S(S(0))$.

Platí tedy následující tvrzení, které ponecháme bez důkazu.

Tvrzení 1.2.3. *Je-li $\varphi(x_1, \dots, x_n)$ existenční formule a $a_1, \dots, a_n \in \mathbb{N}$, potom platí:*

$$Q \vdash \varphi(x_1/\underline{a_1}, \dots, x_n/\underline{a_n}) \text{ právě když } \underline{\mathbb{N}} \models \varphi[e(x_1/a_1, \dots, x_n/a_n)]$$

Užitečným rozšířením Robinsonovy aritmetiky je tzv. Peanova aritmetika, ve které lze *dokazovat indukcí*:

Definice 1.2.4 (Peanova aritmetika)**.** *Peanova aritmetika PA* je extenze Robinsonovy aritmetiky Q o *schéma indukce*, tj. pro každou L-formuli $\varphi(x, \overline{y})$ přidáme následující axiom:

$$(\varphi(0, \overline{y}) \wedge (\forall x)(\varphi(x, \overline{y}) \to \varphi(S(x), \overline{y}))) \to (\forall x)\varphi(x, \overline{y})$$

Peanova aritmetika je mnohem lepší aproximací teorie $\mathrm{Th}(\underline{\mathbb{N}})$, lze v ní dokázat všechny 'základní' vlastnosti platné v $\underline{\mathbb{N}}$ (například komutativitu a asociativitu sčítání). Stále ale existují sentence v jazyce aritmetiky, které platí v $\underline{\mathbb{N}}$, ale v Peanově aritmetice jsou nezávislé.[11]

Poznámka 1.2.5. Pokud bychom se přesunuli do logiky *2. řádu*, potom bychom už mohli strukturu $\underline{\mathbb{N}}$ axiomatizovat (až na izomorfismus), a to extenzí Peanovy aritmetiky o následující formuli 2. řádu, tzv. *axiom indukce*:

$$(\forall X)((X(0) \wedge (\forall x)(X(x) \to X(S(x)))) \to (\forall x)X(x))$$

Připomeňme, že X reprezentuje (libovolnou) unární relaci, neboli podmnožinu univerza. Použitím axiomu indukce na množinu následníků 0 získáme, že každý prvek (daného modelu) je následníkem nuly. Tak můžeme sestrojit izomorfismus s $\underline{\mathbb{N}}$.

1.3 Nerozhodnutelnost predikátové logiky

V této sekci si ukážeme, že nelze (algoritmicky) rozhodovat logickou platnost formulí prvního řádu.[12]

Věta 1.3.1 (O nerozhodnutelnosti predikátové logiky)**.** *Neexistuje algoritmus, který by pro danou vstupní formuli φ rozhodl, zda je logicky platná.*[13]

Protože zatím neznáme potřebný formalismus týkající se algoritmů, např. pojem Turingova stroje, zvolíme jako výchozí bod jiný *nerozhodnutelný problém*. Nejznámějším je tzv. *Halting problem*, tj. otázka, zda se daný program zastaví na daném vstupu.[14] My si ale usnadníme práci tím, že zvolíme jiný nerozhodnutelný problém, tzv. *Hilbertův desátý problém*.[15]

[11] Jak si ukážeme v Gödelově První větě o neúplnosti.

[12] Jinými slovy, že prázdná teorie $T = \emptyset$ není rozhodnutelná.

[13] Tj. zda je formule φ tautologie, neboli zda $\models \varphi$. Zde mluvíme o formulích 1. řádu, v libovolném jazyce.

[14] Jeho nerozhodnutelnost si dokážete v předmětech NTIN071 Automaty a gramatiky a poté znovu v NTIN090 Základy složitosti a vyčíslitelnosti.

[15] Hilbert jej vyslovil v roce 1900, a publikoval v roce 1902 spolu s 22 dalšími problémy, které významně

1.3.1 Hilbertův desátý problém

Mějme polynom $p(x_1, \ldots, x_n)$ s celočíselnými koeficienty. Hilbertův desátý problém se ptá po algoritmu, který rozhodne, zda má takový vstupní polynom celočíselný kořen, neboli zda má *Diofantická rovnice* $p(x_1, \ldots, x_n) = 0$ (celočíselné) řešení:

> "Nalezněte algoritmus, který po konečně mnoha krocích určí, zda daná Diofantická rovnice s libovolným počtem proměnných a celočíselnými koeficienty má celočíselné řešení."

Kdyby se Hilbert dožil vyřešení svého desátého problému v roce 1970, byl by překvapen, že žádný takový algoritmus neexistuje.

Věta 1.3.2 (Matiyasevich, Davis, Putnam, Robinson)**.** *Problém existence celočíselného řešení dané Diofantické rovnice s celočíselnými koeficienty je (algoritmicky) nerozhodnutelný.*

Důkaz zde pro nedostatek místa neuvedeme. K důkazu nerozhodnutelnosti ve skutečnosti použijeme následující důsledek, který mluví o polynomech s přirozenými koeficienty, a o řešení v přirozených číslech.

Důsledek 1.3.3. *Neexistuje algoritmus, který by pro danou dvojici polynomů $p(x_1, \ldots, x_n)$, $q(x_1, \ldots, x_n)$ s* přirozenými *koeficienty rozhodl, zda mají přirozené řešení, tj. zda platí:*

$$\underline{\mathbb{N}} \models (\exists x_1) \ldots (\exists x_n)\ p(x_1, \ldots, x_n) = q(x_1, \ldots, x_n)$$

Důkaz důsledku. Důkaz je snadný, využívá faktu, že každé celé číslo lze vyjádřit jako rozdíl dvojice přirozených čísel, a naopak, každé přirozené číslo lze vyjádřit jako součet čtyř čtverců (celých čísel).[16] Každou Diofantickou rovnici lze tedy transformovat na rovnici z důsledku, a naopak. □

1.3.2 Důkaz nerozhodnutelnosti

Připomeňme, že Robinsonova aritmetika Q má jen konečně mnoho axiomů, $\underline{\mathbb{N}}$ je jejím modelem, a lze v ní dokázat všechna *existenční tvrzení o numerálech* platná v $\underline{\mathbb{N}}$. Nyní jsme připraveni dokázat Větu o nerozhodnutelnosti predikátové logiky.

Důkaz věty o nerozhodnutelnosti predikátové logiky. Uvažme formuli φ tvaru

$$(\exists x_1) \ldots (\exists x_n)\ p(x_1, \ldots, x_n) = q(x_1, \ldots, x_n)$$

kde p a q jsou polynomy s přirozenými koeficienty. Dle Tvrzení 1.2.3 platí:

$$\underline{\mathbb{N}} \models \varphi \text{ právě když } Q \vdash \varphi$$

Označme jako ψ_Q konjunkci (generálních uzávěrů) všech axiomů Q. Zřejmě $Q \vdash \varphi$, právě když $\psi_Q \vdash \varphi$, což platí právě když $\vdash \psi_Q \to \varphi$. Dle Věty o úplnosti je to ale ekvivalentní $\models \psi_Q \to \varphi$. Dostáváme tedy následující ekvivalenci:

$$\underline{\mathbb{N}} \models \varphi \text{ právě když } \vdash \psi_Q \to \varphi$$

ovlivnily matematiku 20., i 21. století. Některé zůstávají nevyřešeny, např. Riemannova hypotéza, viz Wikipedia.

[16] Tzv. Lagrangeova věta o čtyřech čtvercích.

To znamená, že pokud existoval algoritmus rozhodující logickou platnost, mohli bychom rozhodovat i existenci přirozeného řešení rovnice $p(x_1, \ldots, x_n) = q(x_1, \ldots, x_n)$, neboli Hilbertův desátý problém by byl rozhodnutelný.[17] Což by byl spor. □

1.4 Gödelovy věty

Na závěr přednášky představíme slavné Gödelovy věty o neúplnosti, jejichž pochopení by mělo být samozřejmou součástí vzdělání každého informatika. Pokusíme se vysvětlit i princip důkazů, ale vynecháme veškeré technické detaily.

1.4.1 První věta o neúplnosti

Nejprve vyslovíme Gödelovu *První větu o neúplnosti*, a vysvětlíme smysl jejích předpokladů.

Věta 1.4.1 (První věta o neúplnosti)**.** *Pro každou bezespornou rekurzivně axiomatizovanou extenzi T Robinsonovy aritmetiky existuje sentence, která je pravdivá v $\underline{\mathbb{N}}$, ale není dokazatelná v T.*

Takové sentenci se říká *Gödelova sentence*. Velmi neformálně řečeno, Gödelova První věta o neúplnosti říká, že vlastnosti aritmetiky přirozených čísel nelze 'rozumně', efektivně popsat (v logice 1. řádu), každý takový popis je nutně 'neúplný'. Je důležité si uvědomit, že mluvíme o *pravdivosti* ve standardním modelu aritmetiky, tj. ve struktuře $\underline{\mathbb{N}}$, zatímco *dokazatelnost* je v teorii T. (Z Věty o úplnosti samozřejmě plyne, že každá sentence *pravdivá v T* je v T i dokazatelná.)

Bezespornost je nutným předpokladem, neboť ve sporné teorii je dokazatelná každá sentence. Připomeňme, že rekurzivní axiomatizovanost můžeme chápat jako 'efektivní zadání' axiomů (pomocí algoritmu), bez této vlastnosti by taková teorie nebyla užitečná. Požadavek aby teorie byla extenzí Robinsonovy aritmetiky chápejte jako předpoklad, že má alespoň 'základní aritmetickou sílu', že v ní lze určitým způsobem 'mluvit' o přirozených číslech. Existují různé varianty tohoto předpokladu, s jinými teoriemi než je Robinsonova aritmetika, a není například nutné, aby šlo přímo o extenzi, stačí, když je v teorii Robinsonova aritmetika v jistém smyslu 'definovatelná'. Ale teorie, ve které 'nelze zakódovat přirozená čísla' (a zde je důležité, že můžeme mluvit nejen o sčítání, ale i o násobení), je 'příliš slabá'.

Je dobré si uvědomit, že speciálně platí i následující tvrzení 'o nekompletnosti':

Důsledek 1.4.2. *Splňuje-li teorie T předpoklady První věty o neúplnosti a je-li navíc $\underline{\mathbb{N}}$ modelem teorie T, potom T není kompletní.*

Důkaz. Předpokládejme pro spor, že T je kompletní. Vezměme sentenci φ, která je pravdivá v $\underline{\mathbb{N}}$ ale není dokazatelná v T. Díky kompletnosti víme, že $T \vdash \neg\varphi$, potom ale Věta o korektnosti říká, že $T \models \neg\varphi$, tedy φ je lživá v $\underline{\mathbb{N}}$, což je spor. □

Zajímavé je nejen samotné tvrzení První věty o neúplnosti, ale také její důkaz: Gödel v něm přišel se zcela novou, na svou dobu převratnou důkazovou technikou. Sentence sestrojená v důkazu formalizuje tvrzení *"Nejsem dokazatelná v T"*, důkaz je založen na následujících dvou principech, které níže poněkud neformálně popíšeme:

[17] Ukazujeme, že existuje *redukce* 'těžkého' problému (Hilbertova desátého) na náš problém, tedy i náš problém je 'těžký'.

- *aritmetizace syntaxe*, tedy zakódování sentencí a jejich dokazatelnosti do přirozených čísel,
- *self-reference*, tedy schopnost sentence 'mluvit sama o sobě' (o svém kódu).

Aritmetizace dokazatelnosti

Konečné syntaktické objekty, jako jsou termy, formule, konečná tabla, a tedy i tablo důkazy, lze 'rozumně' zakódovat do přirozených čísel.[18] Konkrétní způsob jak to lze provést, tzv. *Gödelovo číslování*, jako technický detail přeskočíme. Stačí nám, že jsme schopni objekty 'algoritmicky' kódovat a dekódovat (případně 'simulovat manipulaci s objekty' na jejich kódech).

Označme kód formule φ jako $\ulcorner\varphi\urcorner$, podobně pro jiné syntaktické objekty. Numerál odpovídající kódu φ, tedy $\ulcorner\varphi\urcorner$-tý numerál, budeme značit $\underline{\varphi}$. Pro danou teorii T definujme následující binární relaci na množině všech přirozených čísel:

$$(n, m) \in \mathrm{Proof}_T \quad \text{právě když} \quad n = \ulcorner\varphi\urcorner \text{ a } m = \ulcorner\tau\urcorner, \text{ kde } \tau \text{ je tablo důkaz sentence } \varphi \text{ z } T$$

Máme-li efektivní přístup k axiomům, umíme také efektivně zkontrolovat zda τ je opravdu důkazem φ (kde τ a φ získáme dekódováním m a n), tedy platí:

Pozorování 1.4.3. *Je-li T rekurzivně axiomatizovaná, je relace* $\mathrm{Proof}_T \subseteq \mathbb{N}^2$ rekurzivní.

Klíčovou, ale velmi technickou částí důkazu První věty je následující tvrzení, které ponecháme bez důkazu.

Tvrzení 1.4.4. *Je-li T navíc extenzí Robinsonovy aritmetiky Q, potom existuje formule $Prf_T(x, y)$ v jazyce aritmetiky, která* reprezentuje *relaci* Proof_T*, tj. pro každá $n, m \in \mathbb{N}$ platí:*

- *Je-li $(n, m) \in \mathrm{Proof}_T$, potom $Q \vdash Prf_T(\underline{n}, \underline{m})$,*
- *jinak $Q \vdash \neg Prf_T(\underline{n}, \underline{m})$.*

Formule $Prf_T(x, y)$ tedy vyjadřuje *"y je důkaz x v T"*.[19] Potom můžeme vyjádřit, že *"x je dokazatelná v T"*, a to formulí $(\exists y)Prf_T(x, y)$. Budeme potřebovat následující důsledek:

Důsledek 1.4.5 (O predikátu dokazatelnosti)**.** *Je-li $T \vdash \varphi$, potom $\underline{\mathbb{N}} \models (\exists y)Prf_T(\underline{\varphi}, y)$ a také $T \vdash (\exists y)Prf_T(\underline{\varphi}, y)$.*

Umíme tedy vyjádřit, že daná sentence je, nebo není, dokazatelná. Jak ale může sentence říci 'sama o sobě', že není dokazatelná? K tomu použijeme *princip self-reference*.

Self-reference

Abychom ilustrovali princip self-reference, pro názornost si místo logické sentence představme větu v češtině, a místo vlastnosti "být dokazatelný" tvrzení o počtu písmen. Podívejme se na následující větu:

`Tato věta má 22 znaků.`

[18] Představte si jakýkoliv rozumný způsob, jak daný objekt zapsat do souboru. Soubor v binárním kódu je posloupnost 0 a 1. Připíšeme na začátek jedničku, abychom nezačínali nulou, a máme binární zápis přirozeného čísla.

[19] Přesněji, tablo jehož kódem je y je důkazem sentence jejíž kódem je y.

V přirozeném jazyce snadno vyjádříme self-referenci zájmenem "Tato", z kontextu víme, že myslíme větu samou. Ve formálních systémech ale typicky nemáme self-referenci přímo k dispozici. *Přímou referenci* obvykle máme k dispozici, stačí umět 'mluvit' o posloupnostech symbolů, jako v následujícím příkladě:

```
Následující věta má 29 znaků. "Následující věta má 29 znaků."
```

Zde se ale není žádná self-reference. Pomůžeme si trikem, kterému budeme říkat 'zdvojení':

```
Následující věta zapsaná jednou a ještě jednou v uvozovkách má 149
znaků. "Následující věta zapsaná jednou a ještě jednou v uvozovkách
má 149 znaků."
```

Pomocí přímé reference a zdvojení tedy můžeme získat self-referenci.

Poznámka 1.4.6. Stejný princip lze použít k sestrojení programu v C, jehož výstupem je jeho vlastní kód (34 je ASCII kód uvozovek):

```
main(){char *c="main(){char *c=%c%s%c; printf(c,34,c,34);}"; printf(c,34,c,34);}
```

1.4.2 Důkaz a důsledky

V této podsekci dokážeme První Gödelovu větu o neúplnosti a řekneme si i něco o jejích důsledcích. Budeme potřebovat následující větu, která popisuje, jak technicky využijeme princip self-reference. Lze na ní nahlížet jako na formu 'diagonalizačního argumentu',[20] proto se tomuto tvrzení také někdy říká *diagonální lemma*.

Věta 1.4.7 (Věta o pevném bodě)**.** *Je-li T extenzí Robinsonovy aritmetiky, potom pro každou formuli $\varphi(x)$ (v jazyce teorie T) existuje sentence ψ taková, že platí:*

$$T \vdash \psi \leftrightarrow \varphi(\underline{\psi})$$

Sentence ψ je tedy *self-referenční*, říká o sobě: "splňuji vlastnost φ".[21] Vysvětlíme si jen myšlenku důkazu. Všimněte si, jak se v důkazu použije přímá reference a zdvojení.

Důkaz. Uvažme *zdvojující funkci*, funkci $d\colon \mathbb{N} \to \mathbb{N}$ takovou, že pro každou formuli $\chi(x)$ platí:

$$d(\ulcorner \chi(x) \urcorner) = \ulcorner \chi(\underline{\chi(x)}) \urcorner$$

Funkce d tedy dostane na vstupu přirozené číslo n, které dekóduje jako formuli v jedné proměnné, dosadí do této formule numerál $\underline{n}$,[22] a výslednou sentenci znovu zakóduje.

S využitím předpokladu, že T je extenzí Q, lze dokázat, že tato funkce je v T *reprezentovatelná*. Pro jednoduchost předpokládejme, že je reprezentovatelná termem,[23] a označme ho také d. To znamená, že pro každou formuli $\chi(x)$ platí:

$$T \vdash d(\underline{\chi(x)}) = \underline{\chi(\underline{\chi(x)})}$$

[20] Diagonalizací se myslí argument připomínající *Cantorův diagonální argument*, známý z důkazu nespočetnosti $\mathbb{R}$. Podobný argument, používající self-referenci, potkáme třeba v *Holičově paradoxu*, nebo v důkazu nerozhodnutelnosti *Halting problému*.

[21] Přesněji, říká to o numerálu odpovídajícímu jejímu kódu.

[22] Zde *numerál* odpovídá 'uvozovkám' z předchozího neformálního popisu self-reference, a $d(\ulcorner \chi \urcorner)$ znamená "χ napsaná jednou a ještě jednou v uvozovkách."

[23] Byť ve skutečnosti je reprezentovaná (složitou) formulí.

Tedy Robinsonova aritmetika dokazuje *o numerálech*, že d opravdu 'zdvojuje'.

Hledaná self-referenční sentence ψ je sentence:[24]

$$\varphi(d(\underline{\varphi(d(x))}))$$

Chceme dokázat, že platí $T \vdash \psi \leftrightarrow \varphi(\underline{\psi})$, neboli $T \vdash \varphi(d(\underline{\varphi(d(x))})) \leftrightarrow \varphi(\underline{\varphi(d(\underline{\varphi(d(x))}))})$. K tomu stačí ověřit, že:

$$T \vdash d(\underline{\varphi(d(x))}) = \underline{\varphi(d(\underline{\varphi(d(x))}))}$$

To ale víme z reprezentovatelnosti d, kde za formuli $\chi(x)$ dosadíme $\varphi(d(x))$. □

Než přistoupíme k samotnému důkazu Gödelovy věty, ukážeme si jako rozcvičku jeden důsledek Věty o pevném bodě: *Definicí pravdy* v aritmetické teorii T myslíme formuli $\tau(x)$ takovou, že pro každou sentenci ψ platí:

$$T \vdash \psi \leftrightarrow \tau(\underline{\psi})$$

Pokud by definice pravdy existovala, znamenalo by to, že místo dokazování sentence stačí spočíst její kód, substituovat příslušný numerál do τ, a vyhodnotit.

Věta 1.4.8 (Nedefinovatelnost pravdy)**.** *V žádném bezesporném rozšíření Robinsonovy aritmetiky neexistuje definice pravdy.*

Důkaz využívá *Paradox lháře*, vyjádříme větu "Nejsem pravdivá v T".

Důkaz. Předpokládejme pro spor, že existuje definice pravdy $\tau(x)$. Použijeme Větu o pevném bodě, kde za formuli $\varphi(x)$ vezmeme $\neg\tau(x)$. Dostáváme existenci sentence ψ takové, že:

$$T \vdash \psi \leftrightarrow \neg\tau(\underline{\psi})$$

Protože $\tau(x)$ je definice pravdy, platí ale i $T \vdash \psi \leftrightarrow \tau(\underline{\psi})$, tedy i $T \vdash \tau(\underline{\psi}) \leftrightarrow \neg\tau(\underline{\psi})$. To by ale znamenalo, že T je sporná. □

Důkaz Gödelovy věty používá tentýž trik, ale pro větu "Nejsem dokazatelá v T".

Důkaz První věty o neúplnosti. Mějme bezespornou rekurzivně axiomatizovanou extenzi T Robinsonovy aritmetiky. Chceme najít Gödelovu sentenci ψ_T, která je pravdivá v $\underline{\mathbb{N}}$, ale není dokazatelná v T.

Takovou sentenci získáme z Věty o pevném bodě jako sentenci vyjadřující "Nejsem dokazatelná v T". Nechť $\varphi(x)$ je formule $\neg(\exists y)Prf_T(x, y)$ ("x není dokazatelná v T"). Podle Věty o pevném bodě existuje sentence ψ_T splňující:

$$T \vdash \psi_T \leftrightarrow \neg(\exists y)Prf_T(\underline{\psi_T}, y)$$

a stejná ekvivalence platí i v $\underline{\mathbb{N}}$.

Nejprve ukážeme, že ψ_T není dokazatelná v T. Předpokládejme pro spor, že $T \vdash \psi_T$. Ze self-reference tedy víme, že platí $T \vdash \neg(\exists y)Prf_T(\underline{\psi_T}, y)$. Z Důsledku 1.4.5 o predikátu dokazatelnosti ale dostáváme $T \vdash (\exists y)Prf_T(\underline{\psi_T}, y)$, což by znamenalo, že T je sporná.

Zbývá ukázat, že ψ_T je pravdivá v $\underline{\mathbb{N}}$. Přepokládejme opět pro spor, že $\underline{\mathbb{N}} \not\models \psi_T$, a protože je to sentence, dostáváme $\underline{\mathbb{N}} \models \neg\psi_T$. Potom ale $\underline{\mathbb{N}} \models (\exists y)Prf_T(\underline{\psi_T}, y)$, a z Důsledku 1.4.5 plyne, že $T \vdash \psi_T$. Což neplatí, jak jsme ukázali výše. □

[24] Následující věta zapsaná jednou a ještě jednou v uvozovkách má vlastnost φ. "Následující věta zapsaná jednou a ještě jednou v uvozovkách má vlastnost φ."

[TODO]

Na závěr si ukážeme dva důsledky a jedno zesílení. Následující okamžitý důsledek už jsme zmínili dříve:

Důsledek 1.4.9. *Je-li T rekurzivně axiomatizovaná extenze Robinsonovy aritmetiky a je-li navíc $\underline{\mathbb{N}}$ modelem teorie T, potom T není kompletní.*

Důkaz. Sporná teorie není kompletní. Pokud by T byla bezesporná, splňovala by předpoklady První věty o neúplnosti, tedy by v ní nebyla dokazatelná Gödelova sentence ψ_T. Pokud by byla kompletní, musela by dokazovat $\neg\psi_T$. To by ale znamenalo, že platí i $\underline{\mathbb{N}} \models \neg\psi_T$, přičemž víme, že ψ_T je v $\underline{\mathbb{N}}$ pravdivá. □

Z toho plyne, že nelze rekurzivně axiomatizovat standardní model přirozených čísel:

Důsledek 1.4.10. *Teorie* $\mathrm{Th}(\underline{\mathbb{N}})$ *není rekurzivně axiomatizovatelná.*

Důkaz. Teorie $\mathrm{Th}(\underline{\mathbb{N}})$ je extenzí Robinsonovy aritmetiky a platí v modelu $\underline{\mathbb{N}}$. Pokud by byla rekurzivně axiomatizovaná, podle předchozího důsledku by nemohla být kompletní, což ale je. □

Jedním ze zesílení Gödelovy První věty je následující tvrzení, které uvedeme bez důkazu. Ukazuje, že předpoklad $\underline{\mathbb{N}} \models T$ v prvním důsledku výše je ve skutečnosti nadbytečný.

Věta 1.4.11 (Rosserův trik, 1936)**.** *V každé bezesporné rekurzivně axiomatizované extenzi Robinsonovy aritmetiky existuje nezávislá sentence. Tedy taková teorie není kompletní.*

1.4.3 Druhá věta o neúplnosti

Druhá Gödelova věta o neúplnosti říká, neformálně řečeno, že efektivně daná, dostatečně bohatá teorie nemůže sama dokázat svou bezespornost. Bezespornost ("konzistenci") vyjádříme následující sentencí, kterou označíme jako Con_T:

$$\neg(\exists y)Prf_T(\underline{0 = S(0)}, y)$$

Všimněte si, že platí $\underline{\mathbb{N}} \models Con_T$, právě když $T \nvdash 0 = S(0)$. Neboli sentence Con_T opravdu vyjadřuje, že *"Teorie T je bezesporná"*.

Věta 1.4.12 (Druhá věta o neúplnosti)**.** *Pro každou bezespornou rekurzivně axiomatizovanou extenzi T Peanovy aritmetiky platí, že Con_T není dokazatelná v T.*

Všimněte si, že sentence Con_T je přitom pravdivá v $\underline{\mathbb{N}}$ (neboť T je opravdu bezesporná). Zmiňme také, že není třeba plná síla Peanovy aritmetiky, stačí slabší předpoklad. Nyní si ukážeme hlavní myšlenku důkazu Druhé věty:

Důkaz Druhé věty o neúplnosti. Vezměme Gödelovu sentenci ψ_T vyjadřující "nejsem dokazatelná v T". V důkazu První věty o neúplnosti (konkrétně v první části) jsme ukázali, že:

> "Pokud je T bezesporná, potom ψ_T není dokazatelná v T."

Z toho jednak plyne, že $T \nvdash \psi_T$, neboť T bezesporná je. Na druhou stranu to lze formulovat jako "platí $Con_T \to \psi_T$" a je-li T extenze Peanovy aritmetiky, lze důkaz tohoto tvrzení zformalizovat v rámci teorie T, tedy ukázat, že:

$$T \vdash Con_T \to \psi_T$$

Kdyby platilo $T \vdash Con_T$, dostali bychom i $T \vdash \psi_T$, což by byl spor. □

Na závěr si ukážeme tři důsledky Druhé věty.

Důsledek 1.4.13. *Existuje model PA, ve kterém platí sentence* $(\exists y) Prf_{PA}(\underline{0 = S(0)}, y)$.

Důkaz. Sentence Con_{PA} není dokazatelná, tedy ani pravdivá v PA. Platí ale v $\underline{\mathbb{N}}$ (neboť PA je bezesporná), což znamená, že je Con_{PA} nezávislá v PA. V nějakém modelu tedy musí platit její negace, která je ekvivalentní $(\exists y) Prf_{PA}(\underline{0 = S(0)}, y)$.

□

Uvědomme si, že musí jít o nestandardní model PA, svědkem musí být *nestandardní* prvek (tj. takový, který není hodnotou žádného numerálu).

Důsledek 1.4.14. *Existuje bezesporná rekurzivně axiomatizovaná extenze T Peanovy aritmetiky, která 'dokazuje svou spornost', tj. taková, že $T \vdash \neg Con_T$.*

Důkaz. Uvažme teorii $T = PA \cup \{\neg Con_{PA}\}$. Tato teorie je bezesporná, neboť $PA \nvdash Con_{PA}$. Také triviálně platí $T \vdash \neg Con_{PA}$ (tj. T 'dokazuje spornost' teorie PA). Protože je $PA \subseteq T$, platí i $T \vdash \neg Con_T$. □

Zde si uvědomme, že $\underline{\mathbb{N}}$ nemůže být modelem teorie T.

Nakonec se podívejme na teorii ZFC, tj. Zermelovu–Fraenkelovu teorii množin s axiomem výběru, na které je založena formalizace matematiky. Tato teorie není formálně vzato extenzí PA, ale není problém v ní Peanovu aritmetiku (v jistém smyslu) 'interpretovat'. To znamená, že ani tato teorie neumí dokázat svou vlastní bezespornost.

Důsledek 1.4.15. *Je-li teorie množin ZFC bezesporná, nemůže být sentence Con_{ZFC} v teorii ZFC dokazatelná.*

Pokud by tedy někdo v rámci teorie ZFC dokázal, že je ZFC bezesporná, znamenalo by to, že je ZFC sporná. Což bude taková pěkná tečka za naší přednáškou.