Stručný úvod do problematiky Gödelových vět o neúplnosti

Miloš Jakubíček

Úvod

Následující text si klade za cíl zasvětit čtenáře do problematiky, jež je zpravidla zahrnována mezi nejtěžší oblasti disciplíny zvané *logika*, a jeho účelem tedy zdaleka není vyčerpat dané téma, nýbrž pokusit se stručně vyložit základní pojmy, ukázat s přihlédnutím k historickému kontextu, čeho se tato problematika týká, nastínit metody použité v důkazech obou Gödelových vět a zmínit další důsledky těchto objevů. U čtenáře se předpokládají znalosti na úrovni středoškolské matematiky, ochota k abstraktnímu myšlení a pochopení faktu, že se jedná pouze o výchozí text pro případné další studium, ke kterému lze vřele doporučit literaturu uvedenou na konci kapitoly.

Kurt Gödel a historické souvislosti

Dříve než se vrhneme do jádra problému, není od věci připomenout osobu Kurta Gödela a souvislosti, v jakých jeho dílo vznikalo. Kurt Gödel, narozený v Brně roku 1906, publikoval svůj článek „Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme“ („O formálně nerozhodnutelných větách v Principia Mathematica a příbuzných systémech“) [1] roku 1931, tj. ve věku pouhých 25 let. Principia Mathematica [2] je obsáhlé dílo A. N. Whiteheada a B. Russela, jehož cílem bylo zachytit celou matematiku (resp. „vyšší aritmetiku“, tj. teorii množin a čísel) jako systém určitých fixních *axiomů* a *inferenčních pravidel*, ze kterých jsou dále vyvozovány další platná tvrzení, tzv. *teorémy*. Tento systém měl být tzv. *úplný*, tj. všechna pravdivá aritmetická tvrzení měla být z axiomů odvoditelná pomocí inferenčních pravidel, tedy měla být teorémem tohoto systému. Zároveň měl ovšem splňovat podmínku *bezespornosti*, tedy vlastnosti takové, že o tomtéž výroku nemělo být možné korektně vyvodit, že je platný a zároveň není platný (je i není teorémem). Jak se dále pokusím vyložit, právě Kurt Gödel dokázal, že cíl, který si Principia Mathematica kladla, je neuskutečnitelný.

Gödelovo očíslování

Klíčem k pochopení Gödelových důkazů a zároveň podstatou jeho objevu je způsob, jakým svoje důkazy konstruoval. Na tomto místě je třeba osvětlit, v čem byla Gödelova myšlenka nová a proč byla vůbec ke konstrukci jeho důkazu zapotřebí. Mluvíme-li o nějakém formálním systému, např. matematice, je třeba zásadní způsobem rozlišit tvrzení *matematická* ($1+1=2$) od tvrzení *o* matematice, tedy tvrzení *metamatematických* („$1+1=2$“ je jednoduchá rovnice). Pak je ovšem třeba zodpovědět otázku, jakým způsobem metamatematická tvrzení (a tedy naše úvahy a důkazy o tomto systému) formulovat – není totiž zase tak obtížné zdánlivě vyvodit spor v nějakém formálním systému, jestliže způsob, jakým *o* tomto systému budeme mluvit (například přirozený jazyk v celé své kráse), bude mít mnohem silnější vyjadřovací schopnosti než systém samotný (což zřejmě přirozený jazyk splňuje).

Příkladem takového zdánlivého sporu je tzv. *Richardův paradox*, formulovaný francouzským matematikem Julesem Richardem v roce 1905 a zadaný následujícím způsobem: uvažujme libovolný přirozený jazyk a množinu přirozených čísel $\mathbb{N}$. Dále předpokládejme, že v daném přirozeném jazyce lze o přirozených číslech mluvit a přiřazovat jim tak různé vlastnosti, jako např. *„být sudé“, „být prvočíslo“, „mít jako poslední cifru jedničku“* apod. Vycházíme-li z toho, že každý přirozený jazyk obsahuje pouze konečně mnoho slov, pak každá taková formulace vlastnosti čísel je také konečná a můžeme všechna tato tvrzení *o číslech* lexikograficky uspořádat a tím každému tvrzení přiřadit nějaké pořadí, tj. přirozené číslo.

Shodou okolností se však může (ale nemusí) stát, že číslo (pořadí) určitého tvrzení bude mít vlastnost, kterou toto tvrzení formuluje – například že na 8. místě bude tvrzení *„být sudé“*. O takovém číslu prohlásíme, že *není Richardovské* a obráceně ostatní čísla tvrzení (která *nemají* vlastnost, kterou tato tvrzení popisují) označíme jako *Richardovská*. Tím jsme ovšem opět korektně formulovali vlastnost přirozených čísel (*„být Richardovské“*). Takové tvrzení tedy nutně je zahrnuto v našem uspořádání a má nějaké pořadí x. Ptejme se nyní, zda pro x platí, že je Richardovské. Jestliže ano, pak nemá vlastnost, kterou popisuje x-té tvrzení, tedy nemá vlastnost *„být Richardovské“* a tedy není Richardovské. Naopak, pokud by x nebylo Richardovské, tak má vlastnost, kterou popisuje x-té tvrzení, tedy má vlastnost *„být Richardovské“* a tedy je Richardovské. Zkrátka a dobře x je Richardovské právě tehdy, když není Richardovské a spor je na světě.

Je to však opravdu spor? Pokud se blíže podíváme na způsob, jakým byl zkonstruován, je zřejmé, že jeho základ spočívá ve smíšení tvrzení *matematických* a tvrzení *o* matematice (tedy *metamatematických*) a je možný jen tehdy, pokud toto připustíme.

Právě takovému postupu se ovšem Gödel geniálním způsobem vyhnul tím, že ukázal, jak lze metamatematická tvrzení formulovat prostřednictvím matematiky samotné. Výsledkem takového postupu bylo přiřazení *jedinečného Gödelova čísla* libovolné formuli z Principia Mathematica, a to takovým způsobem, že z něj lze zpětně dekódovat přesný tvar původní formule. Aniž bych zcela popisoval Gödelův postup, pokusím se alespoň nastínit jeho princip:

Formule se skládají z konečně mnoha symbolů, kterým Gödel prostě přiřadil jednoznačný číselný identifikátor – Gödelovo číslo. Určení Gödelova čísla pro celou formuli (tj. posloupnost takových symbolů) pak probíhá tak, že každému symbolu přiřadil po řadě prvočíslo umocněné na Gödelovo číslo daného symbolu a číslo celé formule získal jako součin těchto mocnin prvočísel. Tento postup byl zřejmě jednoznačný a navíc, protože rozklad čísla na součin prvočísel (tzv. *faktorizace*) je také jednoznačně určen, umožňoval i zpětné vyvozování formule z daného Gödelova čísla.

Jeden příklad za všechny: mějme formuli $(x \wedge y) \Rightarrow z$ a předpokládejme, že všechny 7 symbolů, které se v této formuli vyskytují, mají po řadě přidělena Gödelova čísla od 1 do 7 (tj. „(“ je přiřazena 1, „x“ je přiřazena 2 atd.). Pak Gödelovo číslo celé formule f je následující součin: $2^1 \cdot 3^2 \cdot 5^3 \cdot 7^4 \cdot 11^5 \cdot 13^6 \cdot 17^7 = 172321976576031262654749075 0$. Na příkladu této jednoduché formule je vidět, že Gödelova čísla nabývají obrovských hodnot, nic to ovšem nemění na skutečnosti, že je to teoreticky možný a především korektní způsob jednoznačného přiřazení.

Užitím Gödelova očíslování (někdy nazývaného nepříliš pěkným způsobem jako *Gödelizace*) lze mluvit o aritmetických formulích (resp. o formulích kteréhokoli dostatečně silného formálního systému) opět prostřednictvím aritmetiky samotné jako o relacích čísel. Takto lze například vyjádřit, že formule z našeho příkladu s Gödelovým číslem f zahrnuje symbol implikace „$\Rightarrow$“ tak, že prvočíselný rozklad čísla f obsahuje právě šestou mocninu prvočísla 13, což lze aritmeticky vyjádřit. Gödel založil svůj důkaz právě na tom, že ukázal, jak lze aritmeticky konstruovat složitější formule vyjadřující se o (ne)dokazatelnosti nějakého tvrzení.

1. Gödelova věta

Každý formální systém zahrnující alespoň aritmetiku přirozených čísel buď není bezesporný, nebo není úplný.

Gödel ukázal, jak lze s pomocí jím definovaného očíslování sestrojit formuli G, která (sama o sobě) tvrdí, že „G není dokazatelná“. Nástin jeho postupu je přibližně následující: Nechť Y je nějaká pravdivá formule a X posloupnost odvození formule Y v daném formálním systému, tedy důkaz Y. Protože X i Y jsou korektně utvořené formule, lze najít jejich Gödelovo číslo, které označíme malými písmeny x a y. Pak existuje (složitá) aritmetická relace mezi těmito dvěma čísly x a y, která tento vztah vyjadřuje a kterou Gödel označoval $dem(x, y)$[1], a formule $(\exists x)dem(x, y)$ reprezentuje *uvnitř* daného formálního systému tvrzení, že existuje formule (s Gödelovým číslem x), která je důkazem formule s Gödelovým číslem y, jinými slovy, že Y je dokazatelná formule. Obdobně formule $(\nexists x)dem(x, y)$ tvrdí, že Y není dokazatelná formule.

Jádro jeho postupu spočívá v tom, že přesně ukázal, jak lze najít číslo g takové, že formule G mající Gödelovo číslo g tvrdí „Formule s Gödelovým číslem g není dokazatelná“, a jak tuto formuli způsobem uvedeným v předchozím odstavci vyjádřit – $(\nexists x)dem(x, g)$. Co ovšem plyne z G? Je-li G dokazatelná, pak platí G, tj. že G není dokazatelná. Naopak, není-li G dokazatelná, pak platí to, co G tvrdí, a G je tedy dokazatelná. Výsledkem je, že G je dokazatelná právě tehdy, když G není dokazatelná, tj. $G \iff \neg G$, tedy spor.

Protože platí, že jestliže systém obsahuje G nebo $\neg G$, pak není bezesporný, je zřejmé, že jestliže budeme bezespornost požadovat, nesmí daný formální systém obsahovat ani G ani $\neg G$. Proto formule G je v daném systému *nerozhodnutelná*, a tedy nutně neexistuje žádný důkaz formule G. Pak je ovšem formule G pravdivá, neboť právě sama o sobě tvrdí, že není dokazatelná. Našli jsme tedy pravdivé tvrzení, které v daném systému není dokazatelné, a proto je tento systém evidentně *neúplný*[2]. Z druhé strany platí, že bude-li systém *úplný*, pak nutně obsahuje G i $\neg G$, a není tedy *bezesporný*, čímž je 1. Gödelova věta dokázána.

Kromě toho Gödel dokázal, že tento systém je tzv. *podstatně neúplný*, tj. že i kdybychom rozšířili množinu jeho axiomů a inferenčních pravidel, stále bude možné najít takovou formuli G, pro kterou bude platit výše uvedený spor.

[1] „dem“ je zde zkratka z německého „Demonstration“, tj. důkaz.

[2] I kdybychom nevěděli, že pravdivá je právě formule G, systém by byl stejně neúplný – zřejmě totiž buď G nebo $\neg G$ je pravdivá a systém neobsahuje žádnou z nich.

2. Gödelova věta

Žádný formální systém zahrnující alespoň aritmetiku přirozených čísel nemůže dokázat vlastní bezespornost.

Na výsledcích 1. Gödelovy věty staví druhý zásadní Gödelův objev: Gödel ukázal, že tvrzení „Formální systém je bezesporný" lze formulovat prostřednictvím ekvivalentního tvrzení, že je tento systém neúplný, tj. že existuje alespoň jedna pravdivá formule, která v tomto systému není dokazatelná. Totéž lze vyjádřit tak, že existuje Gödelovo číslo b takové, že pro žádné Gödelovo číslo a neplatí, že a je Gödelovým číslem odvození (důkazu) formule s Gödelovým číslem b, tj. existuje formule $F = (\exists b)(\nexists a)dem(a, b)$.

Je-li formule F pravdivá, pak takový systém není úplný, ale je bezesporný, jak plyne z 1. Gödelovy věty. Jak jsem uvedl v první odstavci, neúplnost lze ekvivalentně zapsat tak, že pro nějakou pravdivou formuli platí, že tato formule není dokazatelná – to je ovšem případ například formule G použité v důkazu 1. Gödelovy věty. Zřejmě tedy platí implikace $F \Rightarrow G$, neboli – přepsáno do řeči Gödelových čísel v aritmetice – platí, že $(\exists b)(\nexists a)dem(a, b) \Rightarrow (\nexists x)dem(x, g)$. Co by to ovšem znamenalo, kdyby F byla dokazatelná? Je-li F dokazatelná (tj. systém je bezesporný), pak je dokazatelná i G a tedy i $\neg G$ a systém bezesporný není! Nelze tedy dokázat bezespornost daného formálního systému *uvnitř* systému samotného.

Stojí za připomenutí, že naopak platí, že pokud nějaký logický systém obsahuje spor, tak v něm lze dokázat libovolné tvrzení. Tato skutečnost plyne z faktu, že formule $p \Rightarrow (\neg p \Rightarrow q)$ je *tautologií*, tj. vždy platným tvrzením[3]. Protože systém obsahuje spor (*kontradikci*), pak v něm lze nalézt platné formule A a $\neg A$. Dosazením do uvedené tautologie vzniká formule $A \Rightarrow (\neg A \Rightarrow B)$, která platí pro libovolné B, a tedy (podle pravidla *modus ponens*) platí i B.

Důsledky Gödelových vět

První Gödelova věta definuje konkrétní teoretické hranice každého formálního systému zahrnujícího aritmetiku, a tak zasahuje mnoho oblastí zahrnujících logiku, matematiku, informatiku i filosofii. Její dopady logické a matematické by měly být z předchozí textu alespoň částečně zřejmé. Z informatického pohledu jsou Gödelovy výsledky úzce spjaty s teorií vyčíslitelnosti a kladou zásadní teoretická omezení na jakoukoli výpočetní sílu užitou pro algoritmizaci různých úloh (nejznámější – a mezi všemi velmi význačnou – takovou nerozhodnutelnou úlohou je tzv. *problém zastavení*). V případě mnoha obecně

[3] Přesněji: jedná se o tvrzení platné pro libovolnou evaluaci hodnot proměnných p, q.

nerozhodnutelných problémů lze ovšem zkonstruovat algoritmus, který rozhoduje daný problém pro určitou (nezřídka velmi podstatnou) podmnožinu možných zadání.

Další důsledky první Gödelovy věty plynou pro problematiku umělé inteligence a zpracování přirozeného jazyka (týkají se například logické analýzy přirozeného jazyka) – už pro predikátovou logiku 1. řádu platí, že množina jejích teorémů není *rekurzivní*, ale pouze *rekurzivně spočetná*, tedy neexistuje algoritmus, který v konečném čase rozhodne, zda zadané tvrzení je či není teorémem.

Ve všech těchto oblastech se však dopad první Gödelovy věty dá shrnout v poněkud nepřesné, přesto velmi výmluvné tvrzení, že jednou pro vždy stanovil hranice algoritmizace a automatizace a ukázal, že člověk nemůže být zcela nahrazen strojem. Jeho výsledky však nemají pouze negativní charakter, naopak, lze je vykládat i jako obrovskou motivaci k využívání výpočetního stroje té největší vyjadřovací síly, jakou známe – lidského mozku. Myslím, že to velmi dobře vyjádřili Ernest Nagel a James R. Newman ve své knize pojednávající o Gödelově důkazu [3, s. 89], kde praví: „Je to příležitost nikoli pro sklíčenost, nýbrž pro obnovené ocenění sil tvůrčího rozumu.“ Obdobně se dá důsledek druhé Gödelovy věty formulovat tak, že u formálních systémů, kterých se týká, nám (kromě jejich pečlivého návrhu samozřejmě) nezbývá, než *věřit* tomu, že jsou bezesporné.

Literatura

[1] GÖDEL, Kurt. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik.* 1931, Heft 38, s. 173–198.

[2] WHITEHEAD, Alfred North, BERTRAND, Russell. *Principia Mathematica.* Cambridge : Cambridge University Press, 1910, 1912, 1913. 3 sv. (počet stran neznámý).

[3] NAGEL, Ernest, NEWMAN, James R. *Gödelův důkaz.* Redakce a předmluva Douglas R. Hofstadter; překlad Rostislav Niederle. 1. vyd. Brno : VUTIUM, 2006. 126 s. edice Quantum. ISBN 80-214-3174-1.

[4] SMULLYAN, Raymond. *Navěky nerozhodnuto : Úvod do logiky a zábavný průvodce ke Gödelovým objevům.* Překlad PhDr. Petr Hromek; recenzent RNDr. Vítězslav Švejdar, CSc.. 1. vyd. Praha : Academia, 2003. 308 s. ISBN 80-200-1068-8.