Neúplnost a nerozhodnutelnost

* Godel: aritmetika nemůže být úplná, Turing: zastavení se nedá rozhodnout.
* Jak to spolu souvisí?
* Jak to ... ?

Neúplnost a nerozhodnutelnost

Gödel:
Bezesporná teorie zahrnující Peanovu aritmetiku nemůže být syntakticky úplná.
(Existují pravdivé formálně nedokazatelné věty o artimetice přirozených čísel.)

Neúplnost a nerozhodnutelnost

Gödel:
Bezesporná teorie zahrnující Peanovu aritmetiku nemůže být syntakticky úplná.
(Existují pravdivé formálně nedokazatelné věty o artimetice přirozených čísel.)

Turing:
Problém zastavení je nerozhodnutelný.
(Existují algoritmicky neřešitelné problémy.)

Seberefence

* Athéňané nikdy nelžou, Kréťané vždy lžou. Kdo může říct „Právě lžu.“?
* „Tato věta není dokazatelná.“ Je to pravdivá věta? Je dokazatelná?
* M je množina množin, které nejsou prvkem sama sebe. Je M prvkem sama sebe?
* Stroj T zastaví na kódu stroje T' právě tehdy, když T' nezastaví na vlastním kódu. Zastaví T na vlastním kódu?

Část I

Gödelův důkaz první věty

Syntaxe a sémantika PL v rychlíku

* ***Jazyk*** L – množina proměnných, predikátových a funkčních symbolů s aritami.
* ***Term*** – $t ::= x \mid f(x_1, \ldots, x_n)$ pro n-ární f
* ***Formule*** – $\varphi ::= p(t_1, \ldots, t_n) \mid \neg(\varphi) \mid \varphi \wedge \varphi \mid \exists x(\varphi)$ pro n-ární p
* ***Interpretace/realizace*** I jazyka L – funkce α_I přiřadí prvky z domény D_I proměnným, relace pred. symbolům, funkce funkčním symbolům.
* $\alpha_I(f(x_1, \ldots, x_n)) = \alpha_I(f)(\alpha_I(x_1), \ldots, \alpha_I(x_n))$ pro n-ární f
* $I \models p(t_1, \ldots, t_n) \Leftrightarrow (\alpha_I(t_1), \ldots, \alpha_I(t_n)) \in \alpha_I(p)$ pro n-ární p
* $I \models \varphi_1 \wedge \varphi_2 \Leftrightarrow I \models \varphi_1$ a $I \models \varphi_2$
* $I \models \neg\varphi \Leftrightarrow I \not\models \varphi$
* $I \models \exists x\varphi \Leftrightarrow I[x/v] \models \varphi$ pro nějaké $v \in D_I$, kde $I[x/v]$ je jako I, jen $\alpha_{I[x/v]}(x) = v$. (předp. že x je volná v φ)

Důkazový systém predikátové logiky

* ***Schémata výrokových axiomů***

(A1) $\varphi \to (\psi \to \varphi)$
(A2) $(\varphi \to (\psi \to \eta)) \to ((\varphi \to \psi) \to (\varphi \to \eta))$
(A3) $(\neg\psi \to \neg\varphi) \to ((\neg\psi \to \varphi) \to \psi)$

kde φ, ψ, η jsou formule PL.

* ***Schéma axiomů kvantifikátoru***:
Není-li x volné ve φ, pak

$$(\forall x(\varphi \to \psi)) \to (\varphi \to (\forall x \psi))$$

* ***Schéma axiomů substituce***:

$$(\forall x\, \varphi) \to \varphi[x/t]$$

kde t je term substituovatelný za x do φ.

* ***Axiomy rovnosti***:
Pro libovolné funkční a predikátové symboly f/n a p/n a proměnné $x, x_1, \ldots, x_n, y_1, \ldots, y_n$

$$x = x$$
$$x_1 = y_1 \to (\ldots x_n = y_n \to f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n)\ldots)$$
$$x_1 = y_1 \to (\ldots x_n = y_n \to p(x_1, \ldots, x_n) \to p(y_1, \ldots, y_n)\ldots)$$

* ***Pravidlo Modus ponens***:

Z předpokladů φ a $(\varphi \to \psi)$ odvodíme závěr ψ.

* ***Pravidlo zobecnění (generalizace)***:

Z předpokladu φ odvodíme závěr $(\forall x \varphi)$.

Peanova aritmetika T_{PA}

1. $\forall x \neg(S(x) = 0)$ (nula je první)
2. $\forall xy(S(x) = S(y) \to x = y)$ (každý má jiného následníka)
3. pro formule φ jazyka T_{PA} s jednou volnou proměnnou:

$$\big[\varphi(0) \wedge (\forall x(\varphi(x) \to \varphi(S(x))))\big] \to \forall x\,(\varphi(x)) \qquad \text{(axiom indukce)}$$

4. $\forall x(x + 0 = x)$ (0 je neutrální k $+$)
5. $\forall xy(x + S(y) = S(x + y))$ (def. sčítání)
6. $\forall x(x \cdot 0 = 0)$ (0 je nulová k $\cdot$)
7. $\forall xy(x \cdot S(y) = x \cdot y + x)$ (def. násobení)

Důkaz formule φ z množiny předpokladů T,
je sekvencí formulí $\varphi_1, \ldots, \varphi_n$,
kde $\varphi_n = \varphi$ a pro každé $i : 1 \leq i \leq n$,
formule φ_i je axiomem nebo prvkem T
nebo vznikla z $\varphi_1, \ldots \varphi_{i-1}$ aplikací odvozovacích pravidel.
Píšeme $\boldsymbol{T \vdash \varphi}$.

* ***Model*** teorie: interpretace $\mathcal{M}$, kde $\mathcal{M} \models \varphi$ pro všechny spec. axiomy $\varphi \in T$.
* ***Logický důsledek*** teorie. $T \models \varphi$: φ je platná ve všech modelech T.
* ***Dokazatelnost*** v teorii. $T \vdash \varphi$: Existuje důkaz φ z předpokladů T.

* PL je ***korektní***: Pokud $T \vdash \varphi$, potom $T \models \varphi$.
* PL je ***sémanticky úplná***: Pokud $T \models \varphi$, potom $T \vdash \varphi$. (Gödel)
* PL je ***efektivní***: Množina důkazů je rozhodnutelná.

* ***Bezespornost***: není možné $T \vdash \varphi$ a $T \vdash \neg\varphi$.
* ***Efektivnost***: množina axiomů je rozhodnutelná (a tedy i množina důkazů v T).
* ***Syntaktická úplnost***: pro uzav. formuli, buď $T \vdash \varphi$ nebo $T \vdash \neg\varphi$.

* T_{PA} je efektivní a bezesporná.
* Není úplná (Gödel).

Princip důkazu na příkladu

Princip důkazu na příkladu

* Stroj tisknoucí řetězce z $\{\neg, N, P, (,), \}^*$.

Princip důkazu na příkladu

* Stroj tisknoucí řetězce z $\{\neg, N, P, (,), \}^*$.
* Řetězce formy $P(X)$, $\neg P(X)$, $PN(X)$ a $\neg PN(X)$ jsou výroky (kde X je řetězec).

Princip důkazu na příkladu

* Stroj tisknoucí řetězce z $\{\neg, N, P, (,), \}^*$.
* Řetězce formy $P(X)$, $\neg P(X)$, $PN(X)$ a $\neg PN(X)$ jsou výroky (kde X je řetězec).
* Výroky mají význam:
  - $P(X)$: X je tisknutelný.
  - $PN(X)$: norma X je tisknutelná. Norma řetězce X je řetězec $X(X)$.
  - $\neg V$: není pravda, že V.

Princip důkazu na příkladu

* Stroj tisknoucí řetězce z $\{\neg, N, P, (,), \}^*$.
* Řetězce formy $P(X)$, $\neg P(X)$, $PN(X)$ a $\neg PN(X)$ jsou výroky (kde X je řetězec).
* Výroky mají význam:
  - $P(X)$: X je tisknutelný.
  - $PN(X)$: norma X je tisknutelná. Norma řetězce X je řetězec $X(X)$.
  - $\neg V$: není pravda, že V.
* Systém je korektní. Tiskne pouze pravdivé výroky.

Princip důkazu na příkladu

* Stroj tisknoucí řetězce z $\{\neg, N, P, (,), \}^*$.
* Řetězce formy $P(X)$, $\neg P(X)$, $PN(X)$ a $\neg PN(X)$ jsou výroky (kde X je řetězec).
* Výroky mají význam:
  - $P(X)$: X je tisknutelný.
  - $PN(X)$: norma X je tisknutelná. Norma řetězce X je řetězec $X(X)$.
  - $\neg V$: není pravda, že V.
* Systém je korektní. Tiskne pouze pravdivé výroky.

Může systém být úplný? Být schopen vytisknout všechny pravdivé výroky?

Princip důkazu na příkladu

* Stroj tisknoucí řetězce z $\{\neg, N, P, (,), \}^*$.
* Řetězce formy $P(X)$, $\neg P(X)$, $PN(X)$ a $\neg PN(X)$ jsou výroky (kde X je řetězec).
* Výroky mají význam:
  - $P(X)$: X je tisknutelný.
  - $PN(X)$: norma X je tisknutelná. Norma řetězce X je řetězec $X(X)$.
  - $\neg V$: není pravda, že V.
* Systém je korektní. Tiskne pouze pravdivé výroky.

Může systém být úplný? Být schopen vytisknout všechny pravdivé výroky?

* Formule, která implikuje netisknutelnost sebe sama?

Princip důkazu na příkladu

* Stroj tisknoucí řetězce z $\{\neg, N, P, (,), \}^*$.
* Řetězce formy $P(X)$, $\neg P(X)$, $PN(X)$ a $\neg PN(X)$ jsou výroky (kde X je řetězec).
* Výroky mají význam:
  - $P(X)$: X je tisknutelný.
  - $PN(X)$: norma X je tisknutelná. Norma řetězce X je řetězec $X(X)$.
  - $\neg V$: není pravda, že V.
* Systém je korektní. Tiskne pouze pravdivé výroky.

Může systém být úplný? Být schopen vytisknout všechny pravdivé výroky?

* Formule, která implikuje netisknutelnost sebe sama?

$$\neg PN(\neg PN)$$

Princip důkazu na příkladu

* Stroj tisknoucí řetězce z $\{\neg, N, P, (,), \}^*$.
* Řetězce formy $P(X)$, $\neg P(X)$, $PN(X)$ a $\neg PN(X)$ jsou výroky (kde X je řetězec).
* Výroky mají význam:
  - $P(X)$: X je tisknutelný.
  - $PN(X)$: norma X je tisknutelná. Norma řetězce X je řetězec $X(X)$.
  - $\neg V$: není pravda, že V.
* Systém je korektní. Tiskne pouze pravdivé výroky.

Může systém být úplný? Být schopen vytisknout všechny pravdivé výroky?

* Formule, která implikuje netisknutelnost sebe sama?

$$\neg PN(\neg PN) \left\{ \begin{array}{l} \text{Nepravdivá, tedy tisknutelná. Spor s korektností.} \end{array} \right.$$

Princip důkazu na příkladu

* Stroj tisknoucí řetězce z $\{\neg, N, P, (,), \}^*$.
* Řetězce formy $P(X)$, $\neg P(X)$, $PN(X)$ a $\neg PN(X)$ jsou výroky (kde X je řetězec).
* Výroky mají význam:
  - $P(X)$: X je tisknutelný.
  - $PN(X)$: norma X je tisknutelná. Norma řetězce X je řetězec $X(X)$.
  - $\neg V$: není pravda, že V.
* Systém je korektní. Tiskne pouze pravdivé výroky.

Může systém být úplný? Být schopen vytisknout všechny pravdivé výroky?

* Formule, která implikuje netisknutelnost sebe sama?

$$\neg PN(\neg PN) \begin{cases} \text{Nepravdivá, tedy tisknutelná. Spor s korektností.} \\ \text{Pravdivá, tedy netisknutelná.} \end{cases}$$

Základní pozorování: Jména čísel a čísla slov

* Formální zápisy jako formule, důkazy, popisy Turingových strojů ... jsou slova nad konečnou abecedou.
* Slova se dají uspořádat, třeba abecedně, a číslovat. zápisu item Např. číslo slova je hodnota jeho ASCI zápisu interpretovaného jako binární číslo.
* Můžeme mluvit o čísle slova a číslovce (slově/jméně čísla).

Sebereference, abstraktní systém

Abstraktní formální systém

* E je množina ***výrazů***, slov nad nějakou konečnou abecedou. Expressions
* $S \subseteq E$ je množina ***výroků***. Sentences
* T je množina ***pravdivých výroků***. True sent.
* $P \subseteq V$ je množina ***dokazatelných výroků***. Proovable
* $R \subseteq V$ je množina ***vyvratitelných výroků***. Refutable
* $H \subseteq E$ je množina ***predikátů***, kde pro každé $h \in H$ a $n \in \mathbb{N}$, $h(n)$ je výrok.
* Korektnost: $P \subseteq T$, $R \subseteq V \setminus T$.
* Syntaktická úplnost: $V = P \cup R$.

Sebereference, notace

Sebereference, notace

* Očíslujeme výrazy: Nechť každý výraz e má ***Gödelovo číslo*** $G(e)$.

Sebereference, notace

* Očíslujeme výrazy: Nechť každý výraz e má ***Gödelovo číslo*** $G(e)$.
* E_n značí výraz s G. číslem n.

Sebereference, notace

* Očíslujeme výrazy: Nechť každý výraz e má ***Gödelovo číslo*** $G(e)$.
* E_n značí výraz s G. číslem n.
* Výraz $E_n(n)$ nazýváme ***diagonalizací*** n.

Sebereference, notace

* Očíslujeme výrazy: Nechť každý výraz e má ***Gödelovo číslo*** $G(e)$.
* E_n značí výraz s G. číslem n.
* Výraz $E_n(n)$ nazýváme ***diagonalizací*** n.
* Predikát p ***vyjadřuje*** množinu $A \subseteq \mathbb{N}$, pokud $n \in A \leftrightarrow p(n) \in T$.

Sebereference, notace

* Očíslujeme výrazy: Nechť každý výraz e má ***Gödelovo číslo*** $G(e)$.
* E_n značí výraz s G. číslem n.
* Výraz $E_n(n)$ nazýváme ***diagonalizací*** n.
* Predikát p ***vyjadřuje*** množinu $A \subseteq \mathbb{N}$, pokud $n \in A \leftrightarrow p(n) \in T$.
* $\tilde{A}$ značíme komplement $\mathbb{N} \setminus A$ mn. A.

Sebereference, notace

* Očíslujeme výrazy: Nechť každý výraz e má ***Gödelovo číslo*** $G(e)$.
* E_n značí výraz s G. číslem n.
* Výraz $E_n(n)$ nazýváme ***diagonalizací*** n.
* Predikát p ***vyjadřuje*** množinu $A \subseteq \mathbb{N}$, pokud $n \in A \leftrightarrow p(n) \in T$.
* $\tilde{A}$ značíme komplement $\mathbb{N} \setminus A$ mn. A.
* Pro $A \subseteq \mathbb{N}$, A^* je množina taková, že $n \in A^* \leftrightarrow E_n(n) \in A$.

Tarski: V systému se sebereferencí nevyjádříme pravdivost

Dostatečně expresivní systém nemůže vyjádřit pravdivost vlastních výroků. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

Tarski: V systému se sebereferencí nevyjádříme pravdivost

Dostatečně expresivní systém nemůže vyjádřit pravdivost vlastních výroků. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

Důkaz:

Tarski: V systému se sebereferencí nevyjádříme pravdivost

Dostatečně expresivní systém nemůže vyjádřit pravdivost vlastních výroků. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

Důkaz:

* Nechť je T vyjádřitelná.

Tarski: V systému se sebereferencí nevyjádříme pravdivost

Dostatečně expresivní systém nemůže vyjádřit pravdivost vlastních výroků. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

Důkaz:

* Nechť je T vyjádřitelná.
* Z G2 je $\tilde{T}$ vyjádřitelná.

Tarski: V systému se sebereferencí nevyjádříme pravdivost

Dostatečně expresivní systém nemůže vyjádřit pravdivost vlastních výroků. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

Důkaz:

* Nechť je T vyjádřitelná.
* Z G2 je $\tilde{T}$ vyjádřitelná.
* Z G1 je $\tilde{T}^*$ vyjádřitelná.

Tarski: V systému se sebereferencí nevyjádříme pravdivost

Dostatečně expresivní systém nemůže vyjádřit pravdivost vlastních výroků. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

Důkaz:

* Nechť je T vyjádřitelná.
* Z G2 je $\tilde{T}$ vyjádřitelná.
* Z G1 je $\tilde{T}^*$ vyjádřitelná.
* Nechť Z vyjadřuje $\tilde{T}^*$ a nechť $G(Z) = z$.
  $Z(n)$ vyjadřuje, že predikát aplikovaný na vlastní G. číslo n není pravdivý.
* $Z(z)$ říká „Jsem nepravdivá.“

Tarski: V systému se sebereferencí nevyjádříme pravdivost

Dostatečně expresivní systém nemůže vyjádřit pravdivost vlastních výroků. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

Důkaz:

* Nechť je T vyjádřitelná.
* Z G2 je $\tilde{T}$ vyjádřitelná.
* Z G1 je $\tilde{T}^*$ vyjádřitelná.
* Nechť Z vyjadřuje $\tilde{T}^*$ a nechť $G(Z) = z$.
  $Z(n)$ vyjadřuje, že predikát aplikovaný na vlastní G. číslo n není pravdivý.
* $Z(z)$ říká „Jsem nepravdivá.“
* $Z(z) \in T \Leftrightarrow z \in \tilde{T}^* \Leftrightarrow Z(z) \in \tilde{T} \Leftrightarrow Z(z) \notin T$ Spor.

Sebereference a neúplnost

Pokud je $\tilde{P}^*$ vyjádřitelná a systém je korektní, potom není úplný.

Sebereference a neúplnost

Pokud je $\tilde{P}^*$ vyjádřitelná a systém je korektní, potom není úplný.

Důkaz:

* Nechť H vyjadřuje $\tilde{P}^*$ a $h = G(H)$.
  $H(n)$ vyjadřuje, že E_n aplikovaná vlastní číslo, n, není dokazatelná.

Sebereference a neúplnost

Pokud je $\tilde{P}^*$ vyjádřitelná a systém je korektní, potom není úplný.

Důkaz:

* Nechť H vyjadřuje $\tilde{P}^*$ a $h = G(H)$.
  $H(n)$ vyjadřuje, že E_n aplikovaná vlastní číslo, n, není dokazatelná.
* $H(h)$ říká „Nejsem dokazatelná.“

Sebereference a neúplnost

Pokud je $\tilde{P}^*$ vyjádřitelná a systém je korektní, potom není úplný.

Důkaz:

* Nechť H vyjadřuje $\tilde{P}^*$ a $h = G(H)$.
  $H(n)$ vyjadřuje, že E_n aplikovaná vlastní číslo, n, není dokazatelná.
* $H(h)$ říká „Nejsem dokazatelná.“
* $H(h) \notin T$, potom $H(h) \in P$ z definice $H(h)$. To by byl spor s korektností.

Sebereference a neúplnost

Pokud je $\tilde{P}^*$ vyjádřitelná a systém je korektní, potom není úplný.

Důkaz:

* Nechť H vyjadřuje $\tilde{P}^*$ a $h = G(H)$.
  $H(n)$ vyjadřuje, že E_n aplikovaná vlastní číslo, n, není dokazatelná.
* $H(h)$ říká „Nejsem dokazatelná.“
* $H(h) \notin T$, potom $H(h) \in P$ z definice $H(h)$. To by byl spor s korektností.
  Tedy $H(h) \in T$, a z definice $H(h)$, $H(h) \notin P$. $H(h)$ je pravdivá nedokazatelná formule.

Sebereference a neúplnost

Pokud je $\tilde{P}^*$ vyjádřitelná a systém je korektní, potom není úplný.

Důkaz:

* Nechť H vyjadřuje $\tilde{P}^*$ a $h = G(H)$.
  $H(n)$ vyjadřuje, že E_n aplikovaná vlastní číslo, n, není dokazatelná.
* $H(h)$ říká „Nejsem dokazatelná.“
* $H(h) \notin T$, potom $H(h) \in P$ z definice $H(h)$. To by byl spor s korektností.
  Tedy $H(h) \in T$, a z definice $H(h)$, $H(h) \notin P$. $H(h)$ je pravdivá nedokazatelná formule.

Dostatečně expresivní systém nemůže být úplný. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

G3 P je vyjádřitelná.

Sebereference a neúplnost verze 2

Pokud je R^* vyjádřitelná a systém je korektní, potom není úplný.

Důkaz:

* Nechť K vyjadřuje R^* a $k = G(K)$.
  $K(n)$ vyjadřuje, že E_n aplikovaná vlastní číslo, n, je vyvratitelná.
* Potom $K(k)$ říká „Jsem vyvratitelná.“
* $K(k) \in T$, potom $K(k) \in R$ z definice $K(k)$. To by byl spor s korektností.
  Tedy $K(k) \notin T$, a z definice $K(k)$, $K(k) \notin R$. $K(k)$ je nepravdivá nevyvratitelná formule.

Dostatečně expresivní systém nemůže být úplný. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G3' R je vyjádřitelná.

Kostra Gödelova důkazu

Víme, že dostatečně expresivní systém nemůže být úplný. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

G3 P je vyjádřitelná.

Gödel dokázal, že systém T_{PA} je dostatečně expresivní.

G2 je super snadná, v PL máme negaci. G1 je poměrně snadná. G3 je velmi komplikovaná.

Kostra Gödelova důkazu

Víme, že dostatečně expresivní systém nemůže být úplný. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

G3 P je vyjádřitelná.

Gödel dokázal, že systém T_{PA} je dostatečně expresivní.
G2 je super snadná, v PL máme negaci. G1 je poměrně snadná. G3 je velmi komplikovaná.

Číslování formulí je třeba zvolit tak, aby se G1 a G3 dokazovaly dobře (Gödel použil jiné).

Kostra Gödelova důkazu

Víme, že dostatečně expresivní systém nemůže být úplný. Stačí, když

- G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.
- G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.
- G3 P je vyjádřitelná.

Gödel dokázal, že systém T_{PA} je dostatečně expresivní.
G2 je super snadná, v PL máme negaci. G1 je poměrně snadná. G3 je velmi komplikovaná.

Číslování formulí je třeba zvolit tak, aby se G1 a G3 dokazovaly dobře (Gödel použil jiné).

* Každý používaný symbol a má číslo $G(a)$:

$'$	0	(	)	f	,	v	$\neg$	$\rightarrow$	$\forall$	$=$	$\leq$	$\sharp$
0	1	2	3	4	5	6	7	8	9	10	11	12

Kostra Gödelova důkazu

Víme, že dostatečně expresivní systém nemůže být úplný. Stačí, když

- G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.
- G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.
- G3 P je vyjádřitelná.

Gödel dokázal, že systém T_{PA} je dostatečně expresivní.
G2 je super snadná, v PL máme negaci. G1 je poměrně snadná. G3 je velmi komplikovaná.

Číslování formulí je třeba zvolit tak, aby se G1 a G3 dokazovaly dobře (Gödel použil jiné).

* Každý používaný symbol a má číslo $G(a)$:

$'$	0	(	)	f	,	v	$\neg$	$\rightarrow$	$\forall$	$=$	$\leq$	$\sharp$
0	1	2	3	4	5	6	7	8	9	10	11	12

* Číslovky $0, 0', 0'', 0''', \ldots,$

Kostra Gödelova důkazu

Víme, že dostatečně expresivní systém nemůže být úplný. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

G3 P je vyjádřitelná.

Gödel dokázal, že systém T_{PA} je dostatečně expresivní.
G2 je super snadná, v PL máme negaci. G1 je poměrně snadná. G3 je velmi komplikovaná.

Číslování formulí je třeba zvolit tak, aby se G1 a G3 dokazovaly dobře (Gödel použil jiné).

* Každý používaný symbol a má číslo $G(a)$:

$'$	0	(	)	f	,	v	$\neg$	$\rightarrow$	$\forall$	$=$	$\leq$	$\sharp$
0	1	2	3	4	5	6	7	8	9	10	11	12

* Číslovky $0, 0', 0'', 0''', \ldots,$
* $+$, $*$ a exponent: $f', f'', f''',$

Kostra Gödelova důkazu

Víme, že dostatečně expresivní systém nemůže být úplný. Stačí, když

G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.

G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.

G3 P je vyjádřitelná.

Gödel dokázal, že systém T_{PA} je dostatečně expresivní.
G2 je super snadná, v PL máme negaci. G1 je poměrně snadná. G3 je velmi komplikovaná.

Číslování formulí je třeba zvolit tak, aby se G1 a G3 dokazovaly dobře (Gödel použil jiné).

* Každý používaný symbol a má číslo $G(a)$:

$'$	0	(	)	f	,	v	$\neg$	$\rightarrow$	$\forall$	$=$	$\leq$	$\sharp$
0	1	2	3	4	5	6	7	8	9	10	11	12

* Číslovky $0, 0', 0'', 0''', \ldots,$
* $+$, $*$ a exponent: $f', f'', f''',$
* jména proměnných $v, v', v'', v''', \ldots,$

Kostra Gödelova důkazu

Víme, že dostatečně expresivní systém nemůže být úplný. Stačí, když

- G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.
- G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.
- G3 P je vyjádřitelná.

Gödel dokázal, že systém T_{PA} je dostatečně expresivní.
G2 je super snadná, v PL máme negaci. G1 je poměrně snadná. G3 je velmi komplikovaná.

Číslování formulí je třeba zvolit tak, aby se G1 a G3 dokazovaly dobře (Gödel použil jiné).

* Každý používaný symbol a má číslo $G(a)$:

$'$	0	(	)	f	,	v	$\neg$	$\rightarrow$	$\forall$	$=$	$\leq$	$\sharp$
0	1	2	3	4	5	6	7	8	9	10	11	12

* Číslovky $0, 0', 0'', 0''', \ldots$,
* $+$, $*$ a exponent: f', f'', f''',
* jména proměnných $v, v', v'', v''', \ldots$,
* $\sharp$ je oddělovač formulí v důkazech.

Kostra Gödelova důkazu

Víme, že dostatečně expresivní systém nemůže být úplný. Stačí, když

- G1 A^* je vyjádřitelná pro každou vyjádřitelnou A.
- G2 $\tilde{A}$ je vyjádřitelná pro každou vyjádřitelnou A.
- G3 P je vyjádřitelná.

Gödel dokázal, že systém T_{PA} je dostatečně expresivní.
G2 je super snadná, v PL máme negaci. G1 je poměrně snadná. G3 je velmi komplikovaná.

Číslování formulí je třeba zvolit tak, aby se G1 a G3 dokazovaly dobře (Gödel použil jiné).

- Každý používaný symbol a má číslo $G(a)$:

$'$	0	(	)	f	,	v	$\neg$	$\rightarrow$	$\forall$	$=$	$\leq$	$\sharp$
0	1	2	3	4	5	6	7	8	9	10	11	12

- Číslovky $0, 0', 0'', 0''', \ldots$,
- $+$, $*$ a exponent: f', f'', f''',
- jména proměnných $v, v', v'', v''', \ldots$,
- $\sharp$ je oddělovač formulí v důkazech.

Slovo $w = a_0 \cdots a_n \in \mathbb{N}^*$, je zápisem ***Gödelova čísla*** ve třináctkové soustavě:

$$G(w) = (a_0 * 13^n) + (a_1 * 13^{n-1}) + \cdots + (a_n * 13^0)$$

Kostra Gödelova důkazu: G2

Pro vyjádřitelnou A je A^* vyjádřitelná:

Kostra Gödelova důkazu: G2

Pro vyjádřitelnou A je A^* vyjádřitelná:

* Konkatenace je vyjádřitelná. $G(u.v) = G(u) * 13^{|v|} + G(v)$. Značíme $G(u) \circ G(v)$.

Kostra Gödelova důkazu: G2

Pro vyjádřitelnou A je A^* vyjádřitelná:

- Konkatenace je vyjádřitelná. $G(u.v) = G(u) * 13^{|v|} + G(v)$. Značíme $G(u) \circ G(v)$. Musíme jen vyjádřit $|v|$ jako funkci $G(v)$ ...

Kostra Gödelova důkazu: G2

Pro vyjádřitelnou A je A^* vyjádřitelná:

* Konkatenace je vyjádřitelná. $G(u.v) = G(u) * 13^{|v|} + G(v)$. Značíme $G(u) \circ G(v)$. Musíme jen vyjádřit $|v|$ jako funkci $G(v)$ ...
* Nechť $\bar{n}$ je číslovka s hodnotou n. $\bar{n}$ je $0\overbrace{'''\ldots'''}^{n}$. Proto $G(\bar{n}) = 13^n$.

Kostra Gödelova důkazu: G2

Pro vyjádřitelnou A je A^* vyjádřitelná:

* Konkatenace je vyjádřitelná. $G(u.v) = G(u) * 13^{|v|} + G(v)$. Značíme $G(u) \circ G(v)$. Musíme jen vyjádřit $|v|$ jako funkci $G(v)$ ...
* Nechť $\bar{n}$ je číslovka s hodnotou n. $\bar{n}$ je $0\overbrace{'''\ldots'''}^{n}$. Proto $G(\bar{n}) = 13^n$.
* $G(E_e(n))$ je vyjádřitelné na základě e a n:
  Pokud E_e má volnou proměnnou v, $E_e(n)$ je ekvivalentní $\forall v(v = \bar{n} \rightarrow E_e)$.

Kostra Gödelova důkazu: G2

Pro vyjádřitelnou A je A^* vyjádřitelná:

* Konkatenace je vyjádřitelná. $G(u.v) = G(u) * 13^{|v|} + G(v)$. Značíme $G(u) \circ G(v)$. Musíme jen vyjádřit $|v|$ jako funkci $G(v)$ ...
* Nechť $\bar{n}$ je číslovka s hodnotou n. $\bar{n}$ je $0\overbrace{'''\ldots'''}^{n}$. Proto $G(\bar{n}) = 13^n$.
* $G(E_e(n))$ je vyjádřitelné na základě e a n:
  Pokud E_e má volnou proměnnou v, $E_e(n)$ je ekvivalentní $\forall v(v = \bar{n} \to E_e)$.
* $G(E_e(n)) = k \circ 13^n \circ 8 \circ e \circ 3$

$$\overbrace{\forall v(v =}^{k}\ \overbrace{\bar{n}}^{13^n}\ \overbrace{\to}^{8}\ \overbrace{E_e}^{e}\ \overbrace{)}^{3}$$

Kostra Gödelova důkazu: G2

Pro vyjádřitelnou A je A^* vyjádřitelná:

* Konkatenace je vyjádřitelná. $G(u.v) = G(u) * 13^{|v|} + G(v)$. Značíme $G(u) \circ G(v)$. Musíme jen vyjádřit $|v|$ jako funkci $G(v)$ ...
* Nechť $\bar{n}$ je číslovka s hodnotou n. $\bar{n}$ je $0\overbrace{'''\ldots'''}^{n}$. Proto $G(\bar{n}) = 13^n$.
* $G(E_e(n))$ je vyjádřitelné na základě e a n:
  Pokud E_e má volnou proměnnou v, $E_e(n)$ je ekvivalentní $\forall v(v = \bar{n} \to E_e)$.
* $G(E_e(n)) = k \circ 13^n \circ 8 \circ e \circ 3$

$$\overbrace{\forall v(v =}^{k}\ \overbrace{\bar{n}}^{13^n}\ \overbrace{\to}^{8}\ \overbrace{E_e}^{e}\ \overbrace{)}^{3}$$

* Pro predikát F vyjadřující A, A^* je vyjádřena predikátem F^* definovaným jako

$$F^*(x) \ =$$

Kostra Gödelova důkazu: G2

Pro vyjádřitelnou A je A^* vyjádřitelná:

* Konkatenace je vyjádřitelná. $G(u.v) = G(u) * 13^{|v|} + G(v)$. Značíme $G(u) \circ G(v)$. Musíme jen vyjádřit $|v|$ jako funkci $G(v)$ ...
* Nechť $\bar{n}$ je číslovka s hodnotou n. $\bar{n}$ je $0\overbrace{'''\ldots'''}^{n}$. Proto $G(\bar{n}) = 13^n$.
* $G(E_e(n))$ je vyjádřitelné na základě e a n:
  Pokud E_e má volnou proměnnou v, $E_e(n)$ je ekvivalentní $\forall v(v = \bar{n} \to E_e)$.
* $G(E_e(n)) = k \circ 13^n \circ 8 \circ e \circ 3$

$$\overbrace{\forall v(v =}^{k}\ \overbrace{\bar{n}}^{13^n}\ \overbrace{\to}^{8}\ \overbrace{E_e}^{e}\ \overbrace{)}^{3}$$

* Pro predikát F vyjadřující A, A^* je vyjádřena predikátem F^* definovaným jako

$$F^*(x) \ =$$
$$\exists n(x = G(E_n(n)) \wedge F(n)) \ =$$

Kostra Gödelova důkazu: G2

Pro vyjádřitelnou A je A^* vyjádřitelná:

* Konkatenace je vyjádřitelná. $G(u.v) = G(u) * 13^{|v|} + G(v)$. Značíme $G(u) \circ G(v)$. Musíme jen vyjádřit $|v|$ jako funkci $G(v)$ ...
* Nechť $\bar{n}$ je číslovka s hodnotou n. $\bar{n}$ je $0\overbrace{'''\ldots'''}^{n}$. Proto $G(\bar{n}) = 13^n$.
* $G(E_e(n))$ je vyjádřitelné na základě e a n:
  Pokud E_e má volnou proměnnou v, $E_e(n)$ je ekvivalentní $\forall v(v = \bar{n} \to E_e)$.
* $G(E_e(n)) = k \circ 13^n \circ 8 \circ e \circ 3$

$$\overbrace{\forall v(v =}^{k}\ \overbrace{\bar{n}}^{13^n}\ \overbrace{\to}^{8}\ \overbrace{E_e}^{e}\ \overbrace{)}^{3}$$

* Pro predikát F vyjadřující A, A^* je vyjádřena predikátem F^* definovaným jako

$$F^*(x) \ =$$
$$\exists n(x = G(E_n(n)) \wedge F(n)) \ =$$
$$\exists n(x = k \circ 13^n \circ 8 \circ e \circ 3 \wedge F(n))$$

Zbytek Gödelova důkazu

Zbývá „jen“ formulemi aritmetiky

- vyjádřit $|w|$ jako funkci $G(w)$,
- vyjádřit x^y v T_{PA} (pomocí násobení a sčítání),
- vyjádřit P.

Zbytek Gödelova důkazu

Zbývá „jen“ formulemi aritmetiky

- vyjádřit $|w|$ jako funkci $G(w)$,
- vyjádřit x^y v T_{PA} (pomocí násobení a sčítání),
- vyjádřit P.

Technicky velmi komplikované, zvlášť vyjádřit P.
Je třeba definovat predikát D takový,
že $D(m, n)$ právě když m je G. číslem důkazu formule s G. číslem n.
P je potom vyjádřena predikátem $\exists m\, D(m, n)$.

Část II

Dokazování a počítání

Efektivnost $\Rightarrow$ generování důkazů

Pro efektivní log. systém existuje program ***Generátor důkazů***, který vypisuje (nekonečný) seznam důkazů, a každý důkaz časem vypíše.

foreach *řetězec* **do**
  if *řetězec* je důkazem **then**
    vypiš řetězec

Efektivnost ⇒ generování důkazů

Pro efektivní log. systém existuje program ***Generátor důkazů***, který vypisuje (nekonečný) seznam důkazů, a každý důkaz časem vypíše.

foreach *řetězec* **do**
  if *řetězec* je důkazem **then**
    vypiš řetězec

* Důkaz je řetězec symbolů z konečné abecedy symbolů.
* Řetězce je možné generovat v abecedním pořadí, na každý jednou dojde.
* Efektivnost: existuje program, který rozhodne, zda je řetězec důkazem.

Efektivnost $\Rightarrow$ generování důkazů

Pro efektivní log. systém existuje program ***Generátor důkazů***, který vypisuje (nekonečný) seznam důkazů, a každý důkaz časem vypíše.

foreach *řetězec* **do**
  if *řetězec* je důkazem **then**
    vypiš řetězec

* Důkaz je řetězec symbolů z konečné abecedy symbolů.
* Řetězce je možné generovat v abecedním pořadí, na každý jednou dojde.
* Efektivnost: existuje program, který rozhodne, zda je řetězec důkazem.

V efektivním systému je $\{\varphi \mid \varphi \text{ je dokazatelná}\}$ částečně rozhodnutelná.

Efektivnost + korektnost + synt. úplnost ⇒ rozhodování platnosti

Tedy umíme řešit „Entscheidungsproblem“.

Pro PL tedy existuje program ***Rozhodovač platnosti formulí***. Pro danou φ:

Spusť Generátor důkazů.
if *Generátor vypsal důkaz* φ **then** φ je platná
if *Generátor vypsal důkaz* $\neg\varphi$ **then** φ není platná

- Úplnost zaručuje, že program skončí, protože:
  - Pro každou větu máme $\models \varphi$, nebo $\models \neg\varphi$ (z def. sémantiky).
  - Ze sémantické úplnosti proto $\vdash \varphi$, nebo $\vdash \neg\varphi$.
- Korektnost zaručuje, že odpověď je správná.

Rozhodnutelná teorie

Teorie je (částečně) rozhodnutelná, pokud $\{\varphi \mid T \models \varphi\}$ je (částečně) rozhodnutelná.

Rozhodnutelná teorie

Teorie je (částečně) rozhodnutelná, pokud $\{\varphi \mid T \models \varphi\}$ je (částečně) rozhodnutelná.

Pro efektivní, bezespornou a úplnou teorii je $\{\varphi \mid T \vdash \varphi\} = \{\varphi \mid T \models \varphi\}$ rozhodnutelná mn. (generátor časem vygeneruje důkaz φ nebo $\neg\varphi$)

Rozhodnutelná teorie

Teorie je (částečně) rozhodnutelná, pokud $\{\varphi \mid T \models \varphi\}$ je (částečně) rozhodnutelná.

Pro efektivní, bezespornou a úplnou teorii je $\{\varphi \mid T \vdash \varphi\} = \{\varphi \mid T \models \varphi\}$ rozhodnutelná mn. (generátor časem vygeneruje důkaz φ nebo $\neg\varphi$)

Efektivní, bezesporná a úplná teorie je rozhodnutelná.
Efektivní a bezesporná teorie je částečně rozhodnutelná.

Sporná teorie je rozhodnutelná triviálně, protože každá formule je jejím důsledkem.

Důkaz a výpočet

* ***Důkaz* / *výpočet*** jsou velmi podobné věci:
  - Je to sekvence ***formulí* / *konfigurací***,
  - které jsou buď ***axiomy* / *iniciální konfigurace***
  - nebo jsou odvozeny z předchozích pomocí jednoduchých mechanických ***odvozovacích pravidel* / *pravidel daných přechodovou funkcí***.

Nerozhodnutelnost HP jako instance abstraktní Tarského věty

Nerozhodnutelnost HP jako instance abstraktní Tarského věty

Abstraktní formální systém, instanciovaný pro Turingovy stroje

* E je množina výrazů – řetězců.

Nerozhodnutelnost HP jako instance abstraktní Tarského věty

Abstraktní formální systém, instanciovaný pro Turingovy stroje

* E je množina výrazů – řetězců.
* $S \subseteq E$ je množina výroků – formy $M(n)$ kde M je zápis TS.

Nerozhodnutelnost HP jako instance abstraktní Tarského věty

Abstraktní formální systém, instanciovaný pro Turingovy stroje

* E je množina výrazů – řetězců.
* $S \subseteq E$ je množina výroků – formy $M(n)$ kde M je zápis TS.
* $T \subseteq S$ je množina pravdivých výroků – výroků $M(n)$ kde M zastaví na n.

Nerozhodnutelnost HP jako instance abstraktní Tarského věty

Abstraktní formální systém, instanciovaný pro Turingovy stroje

* E je množina výrazů – řetězců.
* $S \subseteq E$ je množina výroků – formy $M(n)$ kde M je zápis TS.
* $T \subseteq S$ je množina pravdivých výroků – výroků $M(n)$ kde M zastaví na n.
* $H \subseteq E$ je množina predikátů – kódů TS. Pro $M \in H$ a $n \in \mathbb{N}$ je $M(n)$ výrok.

G1: pro vyjádřitelnou A je také A^* vyjádřitelná:

Nerozhodnutelnost HP jako instance abstraktní Tarského věty

Abstraktní formální systém, instanciovaný pro Turingovy stroje

* E je množina výrazů – řetězců.
* $S \subseteq E$ je množina výroků – formy $M(n)$ kde M je zápis TS.
* $T \subseteq S$ je množina pravdivých výroků – výroků $M(n)$ kde M zastaví na n.
* $H \subseteq E$ je množina predikátů – kódů TS. Pro $M \in H$ a $n \in \mathbb{N}$ je $M(n)$ výrok.

G1: pro vyjádřitelnou A je také A^* vyjádřitelná:

* Nechť stroj M vyjadřuje A (zastaví právě pro elementy A).

Nerozhodnutelnost HP jako instance abstraktní Tarského věty

Abstraktní formální systém, instanciovaný pro Turingovy stroje

* E je množina výrazů – řetězců.
* $S \subseteq E$ je množina výroků – formy $M(n)$ kde M je zápis TS.
* $T \subseteq S$ je množina pravdivých výroků – výroků $M(n)$ kde M zastaví na n.
* $H \subseteq E$ je množina predikátů – kódů TS. Pro $M \in H$ a $n \in \mathbb{N}$ je $M(n)$ výrok.

G1: pro vyjádřitelnou A je také A^* vyjádřitelná:

* Nechť stroj M vyjadřuje A (zastaví právě pro elementy A).
* Potom A^* je vyjádřena strojem M^*, který pro vstup n simuluje M na $M_n(n)$. (konstrukce M_n na základě n je možná, viz kódování TS).

Nerozhodnutelnost HP jako instance abstraktní Tarského věty

Abstraktní formální systém, instanciovaný pro Turingovy stroje

* E je množina výrazů – řetězců.
* $S \subseteq E$ je množina výroků – formy $M(n)$ kde M je zápis TS.
* $T \subseteq S$ je množina pravdivých výroků – výroků $M(n)$ kde M zastaví na n.
* $H \subseteq E$ je množina predikátů – kódů TS. Pro $M \in H$ a $n \in \mathbb{N}$ je $M(n)$ výrok.

G1: pro vyjádřitelnou A je také A^* vyjádřitelná:

* Nechť stroj M vyjadřuje A (zastaví právě pro elementy A).
* Potom A^* je vyjádřena strojem M^*, který pro vstup n simuluje M na $M_n(n)$. (konstrukce M_n na základě n je možná, viz kódování TS).

$\tilde{T}$ tedy není vyjádřitelná, neexistuje TS K, který zastaví na $M(n)$ právě tehdy, když M nezastaví na n. Problém zastavení nemůže být rozhodnutelný, protože jinak by K bylo možné sestrojit jako komplementaci úplného stroje, který jej rozhoduje.

Podobnost Gödelova a Turingova důkazu

Gödel: Nemůžeme dokázat nebo vyvrátit každou formuli.
Sporem. Formule $\psi(x) : \neg\exists y\ D(y, x)$ říká, že formule s G. číslem x není dokazatelná, pokud je x dosazeno za její volnou proměnnou.

* $\psi(G(\psi))$ je dokazatelná. Pak, podle definice ψ, neexistuje důkaz $\psi(G(\psi))$.
* $\neg\psi(G(\psi))$ je dokazatelná. Pak, podle definice ψ, existuje důkaz $\psi(G(\psi))$.

Turing: Nemůžeme rozhodovat problém zastavení.
Sporem. TS M, který zastaví, právě když jeho vstup je kódem TS, který nezastaví na vlastním kódu.

* $M(\langle M\rangle)$ zastaví. Pak, podle definice M, M nezastaví s vlastním kódem na vstupu.
* $M(\langle M\rangle)$ nezastaví. Pak, podle definice M, M zastaví s vlastním kódem na vstupu.

Redukce problému zastavení na problém platnosti aritm. formule

Mějme DTS M který, pro jednoduchost, nikdy neskončí abnormálně.

Redukce problému zastavení na problém platnosti aritm. formule

Mějme DTS M který, pro jednoduchost, nikdy neskončí abnormálně.

Sestrojíme aritm. formuli φ_w^M platnou právě tehdy, když M zastaví na slově $w = a_1 \cdots a_n$.

Redukce problému zastavení na problém platnosti aritm. formule

Mějme DTS M který, pro jednoduchost, nikdy neskončí abnormálně.

Sestrojíme aritm. formuli φ_w^M platnou právě tehdy, když M zastaví na slově $w = a_1 \cdots a_n$.

Formule φ_w^M definuje vztah mezi stavem $S(k)$ v kroce k výpočtu M, pozicí hlavy $H(k)$ v kroce k, a znakem $Z(k, p)$ v kroce k na poli pásky p:

Redukce problému zastavení na problém platnosti aritm. formule

Mějme DTS M který, pro jednoduchost, nikdy neskončí abnormálně.

Sestrojíme aritm. formuli φ_w^M platnou právě tehdy, když M zastaví na slově $w = a_1 \cdots a_n$.

Formule φ_w^M definuje vztah mezi stavem $S(k)$ v kroce k výpočtu M, pozicí hlavy $H(k)$ v kroce k, a znakem $Z(k, p)$ v kroce k na poli pásky p:

Konkrétněji, φ_w^M specifikuje, 1) co platí na začátku běhu, 2) jak následující konfigurace závisí na předchozí, a 3) že stroj někdy zastaví:

$$\varphi_w^M \equiv \varphi_{init} \wedge \varphi_\Delta \wedge \varphi_{stop}$$

Redukce problému zastavení na problém platnosti aritm. formule

Mějme DTS M který, pro jednoduchost, nikdy neskončí abnormálně.

Sestrojíme aritm. formuli φ_w^M platnou právě tehdy, když M zastaví na slově $w = a_1 \cdots a_n$.

Formule φ_w^M definuje vztah mezi stavem $S(k)$ v kroce k výpočtu M, pozicí hlavy $H(k)$ v kroce k, a znakem $Z(k, p)$ v kroce k na poli pásky p:

Konkrétněji, φ_w^M specifikuje, 1) co platí na začátku běhu, 2) jak následující konfigurace závisí na předchozí, a 3) že stroj někdy zastaví:

$$\varphi_w^M \equiv \varphi_{init} \wedge \varphi_\Delta \wedge \varphi_{stop}$$

Na začátku jsme ve stavu 0, hlava je na začátku, a páska je formy $\Delta w \Delta\Delta \ldots$.

$\varphi_{init} \equiv$

Redukce problému zastavení na problém platnosti aritm. formule

Mějme DTS M který, pro jednoduchost, nikdy neskončí abnormálně.

Sestrojíme aritm. formuli φ_w^M platnou právě tehdy, když M zastaví na slově $w = a_1 \cdots a_n$.

Formule φ_w^M definuje vztah mezi stavem $S(k)$ v kroce k výpočtu M, pozicí hlavy $H(k)$ v kroce k, a znakem $Z(k, p)$ v kroce k na poli pásky p:

Konkrétněji, φ_w^M specifikuje, 1) co platí na začátku běhu, 2) jak následující konfigurace závisí na předchozí, a 3) že stroj někdy zastaví:

$$\varphi_w^M \equiv \varphi_{init} \wedge \varphi_\Delta \wedge \varphi_{stop}$$

Na začátku jsme ve stavu 0, hlava je na začátku, a páska je formy $\Delta w \Delta\Delta \ldots$.

$\varphi_{init} \equiv S(0) = q_0$

Redukce problému zastavení na problém platnosti aritm. formule

Mějme DTS M který, pro jednoduchost, nikdy neskončí abnormálně.

Sestrojíme aritm. formuli φ_w^M platnou právě tehdy, když M zastaví na slově $w = a_1 \cdots a_n$.

Formule φ_w^M definuje vztah mezi stavem $S(k)$ v kroce k výpočtu M, pozicí hlavy $H(k)$ v kroce k, a znakem $Z(k, p)$ v kroce k na poli pásky p:

Konkrétněji, φ_w^M specifikuje, 1) co platí na začátku běhu, 2) jak následující konfigurace závisí na předchozí, a 3) že stroj někdy zastaví:

$$\varphi_w^M \equiv \varphi_{init} \wedge \varphi_\Delta \wedge \varphi_{stop}$$

Na začátku jsme ve stavu 0, hlava je na začátku, a páska je formy $\Delta w \Delta\Delta \ldots$.

$\varphi_{init} \equiv S(0) = q_0 \wedge H(0) = 1$

Redukce problému zastavení na problém platnosti aritm. formule

Mějme DTS M který, pro jednoduchost, nikdy neskončí abnormálně.

Sestrojíme aritm. formuli φ_w^M platnou právě tehdy, když M zastaví na slově $w = a_1 \cdots a_n$.

Formule φ_w^M definuje vztah mezi stavem $S(k)$ v kroce k výpočtu M, pozicí hlavy $H(k)$ v kroce k, a znakem $Z(k, p)$ v kroce k na poli pásky p:

Konkrétněji, φ_w^M specifikuje, 1) co platí na začátku běhu, 2) jak následující konfigurace závisí na předchozí, a 3) že stroj někdy zastaví:

$$\varphi_w^M \equiv \varphi_{init} \wedge \varphi_\Delta \wedge \varphi_{stop}$$

Na začátku jsme ve stavu 0, hlava je na začátku, a páska je formy $\Delta w \Delta\Delta \ldots$.

$$\varphi_{init} \equiv S(0) = q_0 \wedge H(0) = 1 \wedge Z(0, 1) = \Delta$$

Redukce problému zastavení na problém platnosti aritm. formule

Mějme DTS M který, pro jednoduchost, nikdy neskončí abnormálně.

Sestrojíme aritm. formuli φ_w^M platnou právě tehdy, když M zastaví na slově $w = a_1 \cdots a_n$.

Formule φ_w^M definuje vztah mezi stavem $S(k)$ v kroce k výpočtu M, pozicí hlavy $H(k)$ v kroce k, a znakem $Z(k, p)$ v kroce k na poli pásky p:

Konkrétněji, φ_w^M specifikuje, 1) co platí na začátku běhu, 2) jak následující konfigurace závisí na předchozí, a 3) že stroj někdy zastaví:

$$\varphi_w^M \equiv \varphi_{init} \wedge \varphi_\Delta \wedge \varphi_{stop}$$

Na začátku jsme ve stavu 0, hlava je na začátku, a páska je formy $\Delta w \Delta\Delta \ldots$.

$$\varphi_{init} \equiv S(0) = q_0 \wedge H(0) = 1 \wedge Z(0,1) = \Delta \wedge (\bigwedge_{p=2}^{n+1} Z(0,p) = a_p)$$

Redukce problému zastavení na problém platnosti aritm. formule

Mějme DTS M který, pro jednoduchost, nikdy neskončí abnormálně.

Sestrojíme aritm. formuli φ_w^M platnou právě tehdy, když M zastaví na slově $w = a_1 \cdots a_n$.

Formule φ_w^M definuje vztah mezi stavem $S(k)$ v kroce k výpočtu M, pozicí hlavy $H(k)$ v kroce k, a znakem $Z(k, p)$ v kroce k na poli pásky p:

Konkrétněji, φ_w^M specifikuje, 1) co platí na začátku běhu, 2) jak následující konfigurace závisí na předchozí, a 3) že stroj někdy zastaví:

$$\varphi_w^M \equiv \varphi_{init} \wedge \varphi_\Delta \wedge \varphi_{stop}$$

Na začátku jsme ve stavu 0, hlava je na začátku, a páska je formy $\Delta w \Delta\Delta \ldots$.

$$\varphi_{init} \equiv S(0) = q_0 \wedge H(0) = 1 \wedge Z(0,1) = \Delta \wedge (\bigwedge_{p=2}^{n+1} Z(0,p) = a_p) \wedge (\forall p > n+1 : Z(0,p) = \Delta)$$

Redukce problému zastavení na problém platnosti aritm. formule

Mějme DTS M který, pro jednoduchost, nikdy neskončí abnormálně.

Sestrojíme aritm. formuli φ_w^M platnou právě tehdy, když M zastaví na slově $w = a_1 \cdots a_n$.

Formule φ_w^M definuje vztah mezi stavem $S(k)$ v kroce k výpočtu M, pozicí hlavy $H(k)$ v kroce k, a znakem $Z(k, p)$ v kroce k na poli pásky p:

Konkrétněji, φ_w^M specifikuje, 1) co platí na začátku běhu, 2) jak následující konfigurace závisí na předchozí, a 3) že stroj někdy zastaví:

$$\varphi_w^M \equiv \varphi_{init} \wedge \varphi_\Delta \wedge \varphi_{stop}$$

Na začátku jsme ve stavu 0, hlava je na začátku, a páska je formy $\Delta w \Delta\Delta \ldots$.

$$\varphi_{init} \equiv S(0) = q_0 \wedge H(0) = 1 \wedge Z(0,1) = \Delta \wedge (\bigwedge_{p=2}^{n+1} Z(0,p) = a_p) \wedge (\forall p > n+1 : Z(0,p) = \Delta)$$

Stroj zastaví, když se někdy dostane do koncového stavu.

$$\varphi_{stop} \equiv \exists k : S(k) = q_f$$

Pro každou k-tou konfiguraci výpočtu, s jakoukoliv pozicí hlavy p, $k+1$-tá konfigurace bude výsledkem kroku výpočtu, který závisí na momentálním stavu q a symbolu pod hlavou a.

$\varphi_\Delta \equiv \forall k \forall p \bigwedge_{q \in Q, a \in \Gamma} \varphi_{(q,a)}$, kde, pokud $\delta(q,a) = (q', X), X \in \{L, R\} \cup \Sigma$, potom

Pro každou k-tou konfiguraci výpočtu, s jakoukoliv pozicí hlavy p, $k+1$-tá konfigurace bude výsledkem kroku výpočtu, který závisí na momentálním stavu q a symbolu pod hlavou a.

$$\varphi_\Delta \equiv \forall k \forall p \bigwedge_{q \in Q, a \in \Gamma} \varphi_{(q,a)}, \quad \text{kde, pokud } \delta(q,a) = (q', X), X \in \{L, R\} \cup \Sigma, \text{ potom}$$

$$\varphi_{(q,a)} \equiv (S(k) = q \wedge H(k) = p \wedge Z(k,p) = a) \rightarrow (S(k+1) = q') \wedge H(k+1) = p' \wedge Z(k+1,p) = a' \wedge \forall \bar{p} \neq p : Z(k+1,\bar{p}) = Z(k,\bar{p}))$$

$$\text{kde } p' = \begin{cases} p & \text{pokud } X \in \Sigma \\ p+1 & \text{pokud } X = R \\ p-1 & \text{pokud } X = L \end{cases}, \qquad a' = \begin{cases} X & \text{pokud } X \in \Sigma \\ a & \text{pokud } X \in \{L, R\} \end{cases}.$$

Pro každou k-tou konfiguraci výpočtu, s jakoukoliv pozicí hlavy p, $k+1$-tá konfigurace bude výsledkem kroku výpočtu, který závisí na momentálním stavu q a symbolu pod hlavou a.

$$\varphi_\Delta \equiv \forall k \forall p \bigwedge_{q \in Q, a \in \Gamma} \varphi_{(q,a)}, \quad \text{kde, pokud } \delta(q,a) = (q', X), X \in \{L, R\} \cup \Sigma, \text{ potom}$$

$$\varphi_{(q,a)} \equiv (S(k) = q \wedge H(k) = p \wedge Z(k,p) = a) \rightarrow (S(k+1) = q') \wedge H(k+1) = p' \wedge Z(k+1,p) = a' \wedge \forall \bar{p} \neq p : Z(k+1,\bar{p}) = Z(k,\bar{p}))$$

$$\text{kde } p' = \begin{cases} p & \text{pokud } X \in \Sigma \\ p+1 & \text{pokud } X = R \\ p-1 & \text{pokud } X = L \end{cases}, \qquad a' = \begin{cases} X & \text{pokud } X \in \Sigma \\ a & \text{pokud } X \in \{L, R\} \end{cases}.$$

Důkaz (skoro) první Gödely věty redukcí z nezorhodnutelnosti HP

φ_w^M je platná právě tehdy, když T zastaví na w.

Důkaz (skoro) první Gödely věty redukcí z nezorhodnutelnosti HP

φ_w^M je platná právě tehdy, když T zastaví na w.

Tedy, platnost aritmetických formulí nemůže být rozhodnutelná,
protože potom by byl rozhodnutelný problém zastavení.

!!Nepoužili jsme T_{PA}, ale podobnou logiku.
Použili jsme Presburgerovu aritmetiku (T_{PA} bez násobení) obohacenou o neinterpretované funkční symboly (S, H, Z), pro která platí axiomy rovnosti. Označme ji T_{Presb+}.
Neinterpretované funkce umí kódovat násobení: $\forall xy\, f(0, x) = 0 \wedge f(y + 1, x) = x + f(y, x)$.
Bez (S, H, Z), čistě s T_{PA} (jen pomocí $+$ a $*$), by to nebylo na slajd ...

T_{Presb+} nemůže být úplná, protože pak by platnost aritm. formulí byla rozhodnutelná.

Stejně ani žádné rozšíření T_{Presb+} (efektivní a bezesporné), ani jakýkoliv jiný efektivní a korektní systém charakterizující sčítání přirozených čísel a neinterpretované funkce přesněji, nemůže být úplný.

O čem přemýšlejí vrány na elektrickém vedení

V žádném jednom efektivním rozumném systému nemůžeme formálně dokázat všechno, co je pravda (např. o přirozených číslech).
Můžeme ale dokazovací systémy dál vyvíjet,
např. objevovat nové axiomy, které umožní dokázat více.
Můžeme tak časem dokázat cokoliv, co je pravda?
Dvě možnosti:

1. Proces vymýšlení nových axiomů a systémů je také výpočtem stroje s Turingovskou silou.
   Pak je to jen komplikovaný TS generující teorémy.
   Aplikují se tedy věty o neúplnosti a nerozhodnutelnosti:
   Nemůžeme dokázat každou platnou formuli. Existují „nedokazatelné pravdy“.
2. Pokud můžeme každou formuli někdy nějak logicky zdůvodnit, dokázat, pak je lidské přemýšlení procesem s větší (nebo jinou) než Turingovskou silou, nedá se formalizovat jako efektivní a korektní systém.

Poznámky

- ***Úplnost teorie implikuje rozhodnutelnost*** (generátorem teorémů a důkazů).
  Ne naopak. Teorie může být rozhodnutelné, tj., existuje algoritmus rozhodující platnost v teorii, a stále neúplná.
- ***Čistá prvořádová logika s rovností*** (axiomy rovnosti: reflexivita, funkční a predikátová kongruence) ***je neúplná*** (např. jednoprvkový vs. nekonečný model a formule $\forall x \forall y (x = y)$, je možné ji zúplnit přidáním axiomu specifikujícího velikost domény).
  Je ale rozhodnutelná (Leopold Löwenheim, 1915).
- Úplnost a „přesnost definice“ není přesně to stejné. Peanova aritmetika je neúplná, ale Presburgerovat aritmetika (Peanova aritmetika bez násobení) je úplná. Protože T_{PA} je neúplná, má více modelů, a protože presburgerova aritmetika je podmnožinou speciálních axiomů T_{PA}, musí mít také více modelů. Presburgerova aritmetika je ale uplná, protože ***rozdíl mezi různými modely se v ní nedá vyjádřit***.
- Úplná teorie je rozhodnutelná generátorem teorémů/důkazů, ale většinou existuje i mnohem efektivnější algoritmus. Například pro Presburgerovu aritmetiku.

Formální versus neformální dokazování/přemýšlení

Věta

"Toto je nedokazatelná věta."

musí být pravdivá, a tedy nedokazatelná, protože jinak by byla dokazatelná, a to by byl spor s korektností dokazovacího systému.

* Přirozený jazyk je také systém se syntaxí a sémantikou a pravidly odvozování.
* Právě jsme v něm formulovali a dokázali větu „Toto je nedokazatelná věta.“,
* která je nedokazatelná ...