Riešenia druhej sady domácich úloh

Peter Kostolányi

25. októbra 2023

Úloha 1. Zostrojte bezkontextovú gramatiku generujúcu jazyk

$$L = \{a^i b a^j b a^k b a^{i+k} \mid i, j, k \in \mathbb{N}\}$$

(nad abecedou $\Sigma = \{a, b\}$). Správnosť svojej konštrukcie dokážte matematickou indukciou.

Riešenie. Dokážeme, že jazyk L je generovaný bezkontextovou gramatikou $G = (N, T, P, \sigma)$, kde $N = \{\sigma, \alpha, \beta\}$, $T = \{a, b\}$ a

$$\begin{aligned} P = \{\sigma &\to a\sigma a \mid b\alpha \\ \alpha &\to a\alpha \mid b\beta \\ \beta &\to a\beta a \mid b\}. \end{aligned}$$

Namiesto rovnosti $L(G) = L$ dokážeme silnejšie tvrdenie $F(G) = F$, kde F je jazyk obsahujúci práve

(*i*) všetky slová $a^i \sigma a^i$, kde $i \in \mathbb{N}$;

(*ii*) všetky slová $a^i b a^j \alpha a^i$, kde $i, j \in \mathbb{N}$;

(*iii*) a všetky slová $a^i b a^j b a^k s a^{i+k}$, kde $i, j, k \in \mathbb{N}$ a $s \in \{\beta, b\}$.

Zrejme pritom $F \cap T^* = L$, takže z rovnosti $F(G) = F$ naozaj vyplynie aj $L(G) = L$.

$\subseteq$: Nech $x \in F(G)$, čiže $\sigma \Rightarrow^n x$ pre nejaké $n \in \mathbb{N}$. Indukciou vzhľadom na n dokážeme, že $x \in F$.

Pre $n = 0$ je $x = \sigma$, čo je slovo tvaru (*i*). Nech teraz tvrdenie platí pre $n = \ell$; uvažujme $n = \ell + 1$. Odvodenie $\sigma \Rightarrow^{\ell+1} x$ vieme v takom prípade prepísať ako $\sigma \Rightarrow^{\ell} y \Rightarrow x$, kde na odvodenie $\sigma \Rightarrow^{\ell} y$ sa vzťahuje indukčný predpoklad. To znamená, že slovo y je jedného z tvarov (*i*) až (*iii*).

Ak $y = a^i \sigma a^i$ pre nejaké $i \in \mathbb{N}$, musí slovo x vzniknúť zo slova y použitím niektorého z pravidiel $\sigma \to a\sigma a$ alebo $\sigma \to b\alpha$ na jediný výskyt neterminálu σ v slove y. V prvom prípade dostávame $x = a^{i+1}\sigma a^{i+1}$, čo je slovo tvaru (*i*). V zostávajúcom prípade zas $x = a^i b\alpha a^i$, čo je slovo tvaru (*ii*). V oboch prípadoch je teda $x \in F$.

Ak $y = a^i b a^j \alpha a^i$ pre nejaké $i, j \in \mathbb{N}$, musí slovo x vzniknúť zo slova y použitím niektorého z pravidiel $\alpha \to a\alpha$ alebo $\alpha \to b\beta$ na jediný výskyt neterminálu α v slove y. V prvom prípade je $x = a^i b a^{j+1} \alpha a^i$, čo je slovo tvaru (*ii*); v druhom prípade je $x = a^i b a^j b \beta a^i$, čo je slovo tvaru (*iii*). V oboch prípadoch teda opäť $x \in F$.

Ak napokon $y = a^i b a^j b a^k s a^{i+k}$ pre nejaké $i, j, k \in \mathbb{N}$ a $s \in \{\beta, b\}$, musí slovo y vďaka vzťahu $y \Rightarrow x$ obsahovať aspoň jeden neterminál – z čoho $s = \beta$ a $y = a^i b a^j b a^k \beta a^{i+k}$. Slovo x teraz musí vzniknúť zo slova y použitím niektorého z pravidiel $\beta \to a\beta a$ alebo $\beta \to b$ na neterminál β. V prvom prípade $x = a^i b a^j b a^{k+1} \beta a^{i+k+1}$ a v druhom $x = a^i b a^j b a^k b a^{i+k}$. V oboch prípadoch ide o slovo tvaru (*iii*), a teda $x \in F$.

$\supseteq$: Dokážeme, že všetky slová v jazyku F – čiže všetky slová tvaru (*i*), (*ii*), alebo (*iii*) – sú vetnými formami v gramatike G.

Urobme tak najprv pre slová tvaru (*i*) a indukciou vzhľadom na i dokážme, že pre všetky $i \in \mathbb{N}$ je $\sigma \Rightarrow^* a^i \sigma a^i$. Pre $i = 0$ je skutočne $\sigma \Rightarrow^0 \sigma$. Nech teraz tvrdenie platí pre $i = k$ a uvažujme $i = k+1$. Z indukčného predpokladu potom $\sigma \Rightarrow^* a^k \sigma a^k \Rightarrow a^{k+1} \sigma a^{k+1}$, kde v poslednom kroku bolo použité pravidlo $\sigma \to a\sigma a$.

Podobne pre slová tvaru (*ii*) teraz indukciou vzhľadom na j dokážme, že pre všetky $i, j \in \mathbb{N}$ je $\sigma \Rightarrow^* a^i b a^j \alpha a^i$. Pre $j = 0$ z tvrdenia pre (*i*) dostávame $\sigma \Rightarrow^* a^i \sigma a^i$ pre všetky $i \in \mathbb{N}$, z čoho $\sigma \Rightarrow^* a^i \sigma a^i \Rightarrow a^i b \alpha a^i$, kde v poslednom kroku bolo použité pravidlo $\sigma \to b\alpha$. Nech ďalej tvrdenie platí pre $j = \ell$ a uvažujme $j = \ell + 1$. Z indukčného predpokladu potom pre všetky $i \in \mathbb{N}$ vyplýva $\sigma \Rightarrow^* a^i b a^\ell \alpha a^i \Rightarrow a^i b a^{\ell+1} \alpha a^i$, kde v poslednom kroku odvodenia bolo použité pravidlo $\alpha \to a\alpha$.

Pre slová tvaru (*iii*) najprv indukciou vzhľadom na k dokážme, že pre všetky $i, j, k \in \mathbb{N}$ je $\sigma \Rightarrow^* a^i b a^j b a^k \beta a^{i+k}$. Pre $k = 0$ z tvrdenia pre (*ii*) dostávame $\sigma \Rightarrow^* a^i b a^j \alpha a^i \Rightarrow a^i b a^j b \beta a^i$ pre všetky $i, j \in \mathbb{N}$, kde v poslednom kroku bolo použité pravidlo $\alpha \to b\beta$. Nech ďalej tvrdenie platí pre $k = m$ a uvažujme $k = m + 1$. Z indukčného predpokladu potom pre všetky $i, j \in \mathbb{N}$ vyplýva $\sigma \Rightarrow^* a^i b a^j b a^m \beta a^{i+m} \Rightarrow a^i b a^j b a^{m+1} \beta a^{i+m+1}$, kde v poslednom kroku odvodenia bolo použité pravidlo $\beta \to a\beta a$.

Zostáva napokon dokázať, že pre všetky $i, j, k \in \mathbb{N}$ je $\sigma \Rightarrow^* a^i b a^j b a^k b a^{i+k}$. Z dokázaného ale $\sigma \Rightarrow^* a^i b a^j b a^k \beta a^{i+k}$ a použitím $\beta \to b$ dostávame $\sigma \Rightarrow^* a^i b a^j b a^k \beta a^{i+k} \Rightarrow a^i b a^j b a^k b a^{i+k}$. ☐

Úloha 2. *Superbezkontextovou gramatikou* pre účely tejto úlohy nazveme štvoricu $G = (N, T, P, S)$, kde N a T sú disjunktné abecedy neterminálnych resp. terminálnych symbolov, $P \subseteq N \times (N \cup T)^*$ je *konečná* množina pravidiel a $S \subseteq N$ je množina počiatočných neterminálov. *Krok odvodenia* gramatiky G je binárna relácia $\Rightarrow_G$ na $(N \cup T)^*$ taká, že pre $u, v \in (N \cup T)^*$ je $u \Rightarrow_G v$ práve vtedy, keď existujú $u_1, u_2, x \in (N \cup T)^*$ a $\xi \in N$ také, že $u = u_1 \xi u_2$, $v = u_1 x u_2$ a $\xi \to x \in P$. *Jazyk generovaný gramatikou* G napokon definujeme ako

$$L(G) = \{w \in T^* \mid \exists \sigma \in S : \sigma \Rightarrow_G^* w\}.$$

Poriadne dokážte, že ku každej superbezkontextovej gramatike existuje ekvivalentná bežná bezkontextová gramatika. Tvrdenia známe z prednášky dokazovať nemusíte.

Riešenie. Nech $G = (N, T, P, S)$ je superbezkontextová gramatika. Pre každý neterminál $\sigma \in S$ označme ako G_σ bezkontextovú gramatiku $G_\sigma = (N, T, P, \sigma)$. Potom

$$L(G) = \{w \in T^* \mid \exists \sigma \in S : \sigma \Rightarrow_G^* w\} = \bigcup_{\sigma \in S} \{w \in T^* \mid \sigma \Rightarrow_G^* w\} = \bigcup_{\sigma \in S} L(G_\sigma).$$

Jazyk $L(G)$ je teda *konečným* zjednotením bezkontextových jazykov $L(G_\sigma)$ cez $\sigma \in S$. Vďaka uzavretosti triedy $\mathscr{L}_{CF}$ na zjednotenie tak musí byť tiež bezkontextový – a teda existuje bezkontextová gramatika G' taká, že $L(G') = L(G)$. ☐