Riešenia úloh z malej písomky

Peter Kostolányi

6. novembra 2023

Úloha 1. Pre abecedu Σ a jazyk $L \subseteq \Sigma^*$ označme ako $\mathrm{MAX}(L)$ jazyk

$$\mathrm{MAX}(L) = \{w \in L \mid \forall x \in \Sigma^+ : wx \notin L\}$$

pozostávajúci zo všetkých slov z jazyka L, ktoré nie sú *vlastným* prefixom[1] žiadneho iného slova z jazyka L.

Nech teraz Σ je abeceda a $L_1, L_2 \subseteq \Sigma^*$ jazyky. Porovnajte jazyky

$$\mathrm{MAX}(L_1 \cup L_2) \quad \text{a} \quad \mathrm{MAX}(L_1) \cup \mathrm{MAX}(L_2).$$

Svoje tvrdenia dokážte.

Riešenie. Dokážeme, že $\mathrm{MAX}(L_1 \cup L_2) \subseteq \mathrm{MAX}(L_1) \cup \mathrm{MAX}(L_2)$, kým opačná inklúzia vo všeobecnosti neplatí.

$\subseteq$: Nech $w \in \mathrm{MAX}(L_1 \cup L_2)$. Potom $w \in L_1 \cup L_2$ a pre všetky $x \in \Sigma^+$ je $wx \notin L_1 \cup L_2$. Keďže $w \in L_1 \cup L_2$, je $w \in L_1$ alebo $w \in L_2$. Ak $w \in L_1$, z $wx \notin L_1 \cup L_2$ pre všetky $x \in \Sigma^+$ vyplýva aj $wx \notin L_1$, z čoho $w \in \mathrm{MAX}(L_1)$, a teda aj $w \in \mathrm{MAX}(L_1) \cup \mathrm{MAX}(L_2)$. Podobne ak $w \in L_2$, z $wx \notin L_1 \cup L_2$ pre všetky $x \in \Sigma^+$ máme aj $wx \notin L_2$, z čoho $w \in \mathrm{MAX}(L_2)$, a teda aj $w \in \mathrm{MAX}(L_1) \cup \mathrm{MAX}(L_2)$.

$\not\supseteq$: Nech $\Sigma = \{a\}$, $L_1 = \{a\}$ a $L_2 = \{aa\}$. Potom $a \in \mathrm{MAX}(L_1)$, z čoho $a \in \mathrm{MAX}(L_1) \cup \mathrm{MAX}(L_2)$. Na druhej strane ale $a \notin \mathrm{MAX}(L_1 \cup L_2) = \mathrm{MAX}(\{a, aa\}) = \{aa\}$. □

Úloha 2. Zostrojte bezkontextovú gramatiku generujúcu jazyk $L = \{a^i b^j a^i \mid i, j \in \mathbb{N}\}$. Správnosť svojej konštrukcie dokážte poriadnou matematickou indukciou.

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

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

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*) a všetky slová $a^i b^j s a^i$, kde $i, j \in \mathbb{N}$ a $s \in \{\alpha, \varepsilon\}$.

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 $n = k$; uvažujme $n = k + 1$. Odvodenie $\sigma \Rightarrow^{k+1} x$ vieme v takom prípade prepísať ako $\sigma \Rightarrow^k y \Rightarrow x$, kde na odvodenie $\sigma \Rightarrow^k y$ sa vzťahuje indukčný predpoklad. To znamená, že slovo y je tvaru (*i*) alebo (*ii*).

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 $\sigma \to \alpha$ 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 $x = a^i \alpha a^i$, čo je slovo tvaru (*ii*). V oboch prípadoch je teda $x \in F$.

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

[1] Pod vlastným prefixom slova u rozumieme prefix slova u rôzny od u samotného.

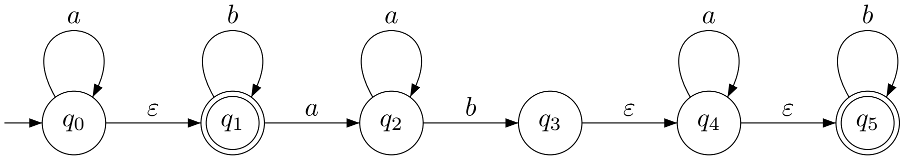
$\supseteq$: Dokážeme, že všetky slová v jazyku F – čiže všetky slová tvaru (i) alebo (ii) – sú vetnými formami v gramatike G.

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

Zostáva napokon dokázať, že pre všetky $i, j \in \mathbb{N}$ je $\sigma \Rightarrow^* a^i b^j a^i$. Z dokázaného ale $\sigma \Rightarrow^* a^i b^j \alpha a^i$ a použitím pravidla $\alpha \to \varepsilon$ dostávame $\sigma \Rightarrow^* a^i b^j \alpha a^i \Rightarrow a^i b^j a^i$. □

Úloha 3. Uvažujme nedeterministický konečný automat A nad abecedou $\Sigma = \{a, b\}$ daný nasledujúcim diagramom.



Nájdite nedeterministický konečný automat A' neobsahujúci prechody na prázdne slovo taký, že $L(A') = L(A)$. Túto rovnosť nie je potrebné dokazovať v prípade použitia štandardnej konštrukcie; v opačnom prípade je dôkaz nutný.

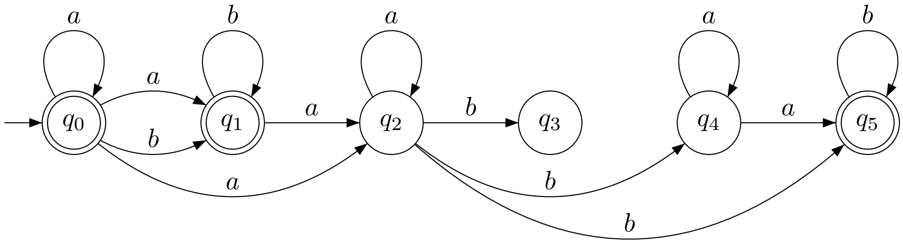
Riešenie. Nech $A = (K, \Sigma, \delta, q_0, F)$. Pre každý stav $q \in K$ najprv nájdeme „epsilonový chvost“ $[q]_\varepsilon$:

$$[q_0]_\varepsilon = \{q_0, q_1\}, \qquad [q_1]_\varepsilon = \{q_1\}, \qquad [q_2]_\varepsilon = \{q_2\},$$
$$[q_3]_\varepsilon = \{q_3, q_4, q_5\}, \qquad [q_4]_\varepsilon = \{q_4, q_5\}, \qquad [q_5]_\varepsilon = \{q_5\}.$$

Následne odstránime prechody na ε vedúce z q_0 do q_1, z q_3 do q_4 a z q_4 do q_5.

Pre všetky stavy $q \in K \setminus \{q_0\}$ a $c \in \Sigma$ ďalej pre všetky $q' \in \delta(q, c)$ povedieme nové prechody na c zo stavu q aj do všetkých stavov $p \in [q']_\varepsilon$. Takto pribudnú prechody na b z q_2 do q_4 a q_5 a prechod na a z q_4 do q_5.

V rámci ďalšieho kroku pre všetky $q \in [q_0]_\varepsilon$, $c \in \Sigma$ a $q' \in \delta(q, c)$ povedieme nové prechody na c zo stavu q_0 do všetkých stavov $p \in [q']_\varepsilon$. Takto dostaneme prechody na a aj na b z q_0 do q_1, ako aj prechod na a z q_0 do q_2.



Obr. 1: Prechodový diagram nedeterministického konečného automatu A'.

V poslednom kroku sa stane akceptačným aj stav q_0, pretože $q_1 \in [q_0]_\varepsilon \cap F$. Výsledný automat $A' = (K', \Sigma', \delta', q_0', F')$ je potom znázornený diagramom na obrázku 1. □

Úloha 4 (Bonus)**.** Pre jazyk $L \subseteq \Sigma^*$ nad abecedou Σ označme ako $\mathrm{MIN}(L)$ jazyk všetkých slov w z jazyka L takých, že žiaden *vlastný* prefix slova w nie je v jazyku L:

$$\mathrm{MIN}(L) = \{w \in L \mid \text{neexistujú žiadne } u \in L \text{ a } v \in \Sigma^+ \text{ také, že } w = uv\}.$$

Jazyk $\mathrm{MAX}(L)$ definujeme rovnako ako v prvej úlohe.

Nech je teraz Σ abeceda a $L \subseteq \Sigma^*$ jazyk. Nájdite maximálnu možnú kardinalitu množiny jazykov

$$\mathscr{M}(L) = \{M_k(M_{k-1}(\ldots(M_2(M_1(L)))\ldots)) \mid k \in \mathbb{N}; M_1, \ldots, M_k \in \{\mathrm{MIN}, \mathrm{MAX}\}\}.$$

Svoje tvrdenia dokážte.

Riešenie. Jazyk $L \subseteq \Sigma^*$ je *bezprefixový*, ak neexistujú žiadne $u \in \Sigma^*$ a $v \in \Sigma^+$ také, že $u \in L$ a zároveň $uv \in L$.

Ľahko dokážeme, že pre všetky $L \subseteq \Sigma^*$ sú jazyky $\mathrm{MIN}(L)$ a $\mathrm{MAX}(L)$ bezprefixové. Keby totiž pre $u \in \Sigma^*$ a $v \in \Sigma^+$ bolo $u \in \mathrm{MIN}(L)$ aj $uv \in \mathrm{MIN}(L)$, mali by sme tiež $u \in L$ a $uv \in L$, z čoho $uv \notin \mathrm{MIN}(L)$: spor. Podobne pre $u \in \Sigma^*$ a $v \in \Sigma^+$ také, že $u \in \mathrm{MAX}(L)$ a $uv \in \mathrm{MAX}(L)$ nutne $u \in L$ aj $uv \in L$, z čoho $u \notin \mathrm{MAX}(L)$: spor.

Pre všetky bezprefixové jazyky $L \subseteq \Sigma^*$ ďalej $\mathrm{MIN}(L) = L$ a $\mathrm{MAX}(L) = L$. Inklúzie $\mathrm{MIN}(L) \subseteq L$ a $\mathrm{MAX}(L) \subseteq L$ vyplývajú priamo z definície tej-ktorej operácie. Ak naopak $w \in L$, musí byť aj $w \in \mathrm{MIN}(L)$, pretože v opačnom prípade by pre nejaké $u \in L$ a $v \in \Sigma^+$ bolo $w = uv$ a príslušnosť obidvoch slov u a $uv = w$ do L by odporovala bezprefixovosti jazyka L. Podobne pre všetky $w \in L$ musí byť aj $w \in \mathrm{MAX}(L)$, pretože v opačnom prípade by existovalo $x \in \Sigma^+$ také, že $wx \in L$ a príslušnosť obidvoch slov w a wx do L by opäť znamenala spor s bezprefixovosťou jazyka L.

Z dokázaného vyplýva, že pre všetky $L \subseteq \Sigma^*$ je $\mathrm{MIN}(\mathrm{MIN}(L)) = \mathrm{MAX}(\mathrm{MIN}(L)) = \mathrm{MIN}(L)$ a $\mathrm{MIN}(\mathrm{MAX}(L)) = \mathrm{MAX}(\mathrm{MAX}(L)) = \mathrm{MAX}(L)$.

Indukciou vzhľadom na $k \in \mathbb{N}$ už teraz ľahko dokážeme, že pre všetky $M_1, \ldots, M_k \in \{\mathrm{MIN}, \mathrm{MAX}\}$ je

$$M_k(M_{k-1}(\ldots(M_2(M_1(L)))\ldots)) \in \{L, \mathrm{MIN}(L), \mathrm{MAX}(L)\}.$$

Pre $k = 0$ je totiž $M_k(M_{k-1}(\ldots(M_2(M_1(L)))\ldots)) = L$; ak teraz tvrdenie platí pre $k = s$, je $M_s(M_{s-1}(\ldots(M_2(M_1(L)))\ldots)) \in \{L, \mathrm{MIN}(L), \mathrm{MAX}(L)\}$ a z dokázaného vyplýva, že aj pre ľubovoľné $M_{s+1} \in \{\mathrm{MIN}, \mathrm{MAX}\}$ je

$$M_{s+1}(M_s(\ldots(M_2(M_1(L)))\ldots)) \in \{\mathrm{MIN}(L), \mathrm{MAX}(L)\} \subseteq \{L, \mathrm{MIN}(L), \mathrm{MAX}(L)\}.$$

Preto nutne $|\mathscr{M}(L)| \leq 3$. Napríklad pre $L = \{a, aa\}$ je navyše $\mathrm{MIN}(L) = \{a\}$ a $\mathrm{MAX}(L) = \{aa\}$, pričom všetky tri jazyky patria do $\mathscr{M}(L)$. *Najväčší možný počet prvkov množiny $\mathscr{M}(L)$ je teda* 3. $\square$