Riešenia prvej sady bodovaných domácich úloh

Peter Kostolányi

21. marca 2017

Úloha 1. *Poriadne* dokážte uzavretosť triedy $\mathscr{L}_{CF}$ na substitúciu. (Vyberte si ľubovoľnú konštrukciu a *poriadne* dokážte jej správnosť.)

Riešenie. Nech L je bezkontextový jazyk generovaný bezkontextovou gramatikou $G = (N, T, P, \sigma)$. Nech τ je bezkontextová substitúcia na T^*. Zostrojíme bezkontextovú gramatiku G' takú, že $L(G') = \tau(L)$.

Nech pre každé $c \in T$ je $G_c = (N_c, T_c, P_c, \sigma_c)$ bezkontextová gramatika taká, že $L(G_c) = \tau(c)$. Predpokladajme navyše, že platí

$$\begin{aligned} N_c \cap N_d &= \emptyset \qquad && \forall c, d \in T, c \neq d, \\ N_c \cap N &= \emptyset && \forall c \in T, \\ \left(N \cup \bigcup_{c \in T} N_c\right) \cap \left(\bigcup_{c\in T} T_c\right) &= \emptyset. \end{aligned}$$

Uvažujme homomorfizmus h na $(N \cup T)^*$ taký, že pre všetky $\xi \in N$ je $h(\xi) = \xi$ a pre všetky $c \in T$ je $h(c) = \sigma_c$. Gramatiku G' potom môžeme zostrojiť nasledovne: $G' = (N', T', P', \sigma')$, kde

$$\begin{aligned} N' &= N \cup \bigcup_{c\in T} N_c, \\ T' &= \bigcup_{c \in T} T_c, \\ P' &= \{\xi \to h(w) \mid (\xi \to w) \in P\} \cup \bigcup_{c \in T} P_c, \\ \sigma' &= \sigma. \end{aligned}$$

Dokážeme, že $L(G') = \tau(L)$.

$\subseteq$: Nech $c \in T$ a $\xi \in N_c$. Dokážeme najprv, že ak pre nejaké $w \in (N' \cup T')^*$ a $n \in \mathbb{N}$ platí $\xi \Rightarrow^n_{G'} w$, tak $w \in (N_c \cup T_c)^*$. Indukciou vzhľadom na n.

1. Nech $n = 0$. Potom nutne $w = \xi \in N_c$ a tvrdenie platí.
2. Nech tvrdenie platí pre $n = k$. Ukážeme, že platí aj pre $n = k + 1$.
   Nech $\xi \Rightarrow^{k+1}_{G'} w$. Nech w' je slovo také, že $\xi \Rightarrow^k_{G'} w' \Rightarrow_{G'} w$. Z indukčného predpokladu vyplýva, že $w' \in (N_c \cup T_c)^*$. Keďže $w' \Rightarrow_{G'} w$, musí existovať $\eta \in N_c$, slová $w_1', w_2' \in (N_c \cup T_c)^*$ a slovo $x \in (N' \cup T')^*$ tak, že $w' = w_1' \eta w_2'$, $w = w_1' x w_2'$ a $\eta \to x \in P'$. Zrejme ale $(P' \cap N_c \times (N' \cup T')^*) \subseteq (N_c \times (N_c \cup T_c)^*)$. Preto $x \in (N_c \cup T_c)^*$, a teda aj $w = w_1' x w_2' \in (N_c \cup T_c)^*$.

Z dokázaného tvrdenia okrem iného vyplýva, že zo žiadneho výskytu neterminálu $\xi \in N' - N$ nemožno odvodiť slovo obsahujúce neterminál z N. To znamená, že pre každé $w \in L(G')$ existuje odvodenie slova w v gramatike G' také, že ľubovoľné použitie pravidla z

$$P_0 := \{\xi \to h(w) \mid (\xi \to w) \in P\}$$

predchádza ľubovoľnému použitiu pravidla z P_c pre niektoré $c \in T$.

Nech $w \in L(G')$. Uvažujme odvodenie slova w v G' spĺňajúce podmienku opísanú vyššie a označme symbolom x vetnú formu, ktorá sa v tomto odvodení vyskytuje po poslednom použití pravidla z P_0 (a pred prvým použitím niektorého zo zvyšných pravidiel). Uvažované odvodenie teda možno zapísať ako

$$\sigma \Rightarrow^n_{G'} x \Rightarrow^m_{G'} w,$$

kde $n, m \in \mathbb{N}$. V odvodení $\sigma \Rightarrow_{G'}^{n} x$ sa pritom používajú iba pravidlá z P_0 a v odvodení $x \Rightarrow_{G'}^{m} w$ sa používajú iba pravidlá z množín P_c, kde $c \in T$.

Odovodenie $\sigma \Rightarrow_{G'}^{n} x$ možno ďalej rozpísať ako $\sigma \Rightarrow_{G'} x_1 \Rightarrow_{G'} x_2 \Rightarrow_{G'} \ldots \Rightarrow_{G'} x_n = x$. Z definície množiny P_0 potom vyplýva, že v gramatike G existuje odvodenie $\sigma \Rightarrow_G y_1 \Rightarrow_G y_2 \Rightarrow_G \ldots \Rightarrow_G y_n$ také, že pre $i = 1, \ldots, n$ platí $h(y_1) = x_i$ (toto tvrdenie možno ľahko dokázať indukciou). Keby navyše slovo y_n obsahovalo neterminál z N, obsahovalo by tento neterminál aj slovo $x = h(y_n)$; z x by sa potom nedalo odvodiť terminálne slovo w iba s použitím pravidiel z $P' - P_0$. Preto $x = h(y) = \sigma_{a_1} \ldots \sigma_{a_r}$ pre nejaké slovo $y = a_1 \ldots a_r \in L$, kde $r \in \mathbb{N}$ a $a_1, \ldots, a_r \in T$.

Bez ujmy na všeobecnosti teraz predpokladajme, že odvodenie $x = \sigma_{a_1} \ldots \sigma_{a_r} \Rightarrow_{G'}^{m} w$ je ľavé krajné. To potom možno rozpísať ako

$$\sigma_{a_1}\sigma_{a_2} \ldots \sigma_{a_r} \Rightarrow_{G'}^{*} w_1\sigma_{a_2} \ldots \sigma_{a_r} \Rightarrow_{G'}^{*} \ldots \Rightarrow_{G'}^{*} w_1 w_2 \ldots w_r = w,$$

kde $w_1, w_2, \ldots, w_r \in (T')^*$. Indukciou pritom možno ľahko dokázať, že pre $i = 1, \ldots, r$ možno slovo w_i odvodiť z neterminálu σ_{a_i} výlučne použitím pravidiel z množiny P_{a_i} – z toho následne vyplýva, že $w_i \in \tau(a_i)$. V dôsledku týchto pozorovaní dostávame

$$w = w_1 \ldots w_r \in \tau(a_1) \ldots \tau(a_r) = \tau(y) \subseteq \tau(L),$$

čo bolo treba dokázať.

$\supseteq$: Nech $w \in \tau(L)$. Potom existuje $u \in L$ také, že $w \in \tau(u)$. Ak navyše $u = a_1 \ldots a_r$ pre nejaké $r \in \mathbb{N}$ a $a_1, \ldots, a_r \in T$, tak $w = w_1 \ldots w_r$, kde $w_i \in \tau(a_i)$ pre $i = 1, \ldots, r$.

Uvažujme teraz odvodenie

$$\sigma \Rightarrow_G u_1 \Rightarrow_G u_2 \Rightarrow_G \ldots \Rightarrow_G u_n = u$$

slova u v gramatike G. Indukciou možno ľahko dokázať, že v gramatike G' existuje odvodenie

$$\sigma \Rightarrow_{G'} h(u_1) \Rightarrow_{G'} h(u_2) \Rightarrow_{G'} \ldots \Rightarrow_{G'} h(u_n) = h(u) = \sigma_{a_1} \ldots \sigma_{a_r}.$$

Keďže pre $i = 1, \ldots, r$ platí $L(G_{a_i}) = \tau(a_i)$, nutne musí platiť aj $\sigma_{a_i} \Rightarrow_{G_{a_i}}^{*} w_i$ – a keďže $P_{a_i} \subseteq P'$, platí aj $\sigma_{a_i} \Rightarrow_{G'}^{*} w_i$. Preto

$$\sigma \Rightarrow_{G'}^{*} \sigma_{a_1}\sigma_{a_2} \ldots \sigma_{a_r} \Rightarrow_{G'}^{*} w_1\sigma_{a_2} \ldots \sigma_{a_r} \Rightarrow_{G'}^{*} \ldots \Rightarrow_{G'}^{*} w_1 w_2 \ldots w_r = w.$$

Preto $w \in L(G')$, čo bolo treba dokázať. □

Úloha 2. Formálne definujte *zásobníkový prekladač* ako variant a-prekladača, ktorý má navyše k dispozícii zásobník a môže s ním pracovať obdobne ako zásobníkový automat. Definujte zobrazenie zásobníkovým prekladačom ako operáciu na jazykoch a nájdite všetky jazyky, na ktoré možno zobraziť jazyk $\{\varepsilon\}$. Nájdite dvojicu jazykov L_1 a L_2 takých, že jazyk L_1 nemožno preložiť na jazyk L_2 žiadnym zásobníkovým prekladačom.

Riešenie. Zásobníkový prekladač je usporiadaná osmica $M = (K, \Sigma_1, \Sigma_2, \Gamma, H, q_0, Z_0, F)$, kde K je neprázdna konečná množina stavov, Σ_1 je vstupná abeceda, Σ_2 je výstupná abeceda, Γ je zásobníková abeceda, $H \subseteq_{kon} K \times \Sigma_1^* \times \Gamma \times K \times \Sigma_2^* \times \Gamma^*$ je prechodová relácia, $q_0 \in K$ je počiatočný stav, $Z_0 \in \Gamma$ je počiatočný zásobníkový symbol a $F \subseteq K$ je množina akceptačných stavov.

Konfigurácia zásobníkového prekladača $M = (K, \Sigma_1, \Sigma_2, \Gamma, H, q_0, Z_0, F)$ je štvorica $(q, u, v, \gamma) \in K \times \Sigma_1^* \times \Sigma_2^* \times \Gamma^*$, kde jednotlivé zložky reprezentujú stav, nedočítanú časť vstupu, doposiaľ zapísanú časť výstupu a obsah zásobníka (s dnom naľavo).

Krok výpočtu zásobníkového prekladača $M = (K, \Sigma_1, \Sigma_2, \Gamma, H, q_0, Z_0, F)$ je binárna relácia $\vdash$ na konfiguráciách prekladača M taká, že pre všetky $p, q \in K$, $x, u \in \Sigma_1^*$, $v, y \in \Sigma_2^*$, $\gamma, \beta \in \Gamma^*$

a $Z \in \Gamma$ platí $(p, xu, v, \gamma Z) \vdash (q, u, vy, \gamma\beta)$ práve vtedy, keď $(p, x, Z, q, y, \beta) \in H$ (pričom v relácii $\vdash$ nie sú žiadne ďalšie dvojice konfigurácií).

Nech $L \subseteq \Sigma_1^*$ je jazyk. *Obraz jazyka* L pri zobrazení zásobníkovým prekladačom M definujeme ako jazyk

$$M(L) = \{v \in \Sigma_2^* \mid \exists u \in L\ \exists q \in F\ \exists \gamma \in \Gamma^* : (q_0, u, \varepsilon, Z_0) \vdash^* (q, \varepsilon, v, \gamma)\}.$$

Nech $\mathscr{L}$ je trieda všetkých jazykov L, pre ktoré existuje zásobníkový prekladač M taký, že $M(\{\varepsilon\}) = L$. Dokážeme, že $\mathscr{L} = \mathscr{L}_{CF}$.

$\subseteq$: Nech $L \in \mathscr{L}$ – existuje teda zásobníkový prekladač M taký, že $M(\{\varepsilon\}) = L$. Dokážeme, že $L \in \mathscr{L}_{CF}$.

Nech $M = (K, \Sigma_1, \Sigma_2, \Gamma, H, q_0, Z_0, F)$. Bez ujmy na všeobecnosti môžeme predpokladať, že $H \subseteq_{kon} K \times \Sigma_1^* \times \Gamma \times K \times (\Sigma_2 \cup \{\varepsilon\}) \times \Gamma^*$ – ľahko možno dokázať, že skutočne ide o normálny tvar zásobníkových prekladačov. Zostrojíme zásobníkový automat A taký, že $L(A) = L$.

Položme $A = (K, \Sigma_2, \Gamma, \delta, q_0, Z_0, F)$ a pre všetky $p \in K$, $z \in \Sigma_2 \cup \{\varepsilon\}$ a $Z \in \Gamma$ definujme

$$\delta(p, z, Z) = \{(q, \gamma) \in K \times \Gamma^* \mid (p, \varepsilon, Z, q, z, \gamma) \in H\}.$$

Ľahko vidieť, že pre automat A skutočne platí $L(A) = L$.

$\supseteq$: Nech $L \in \mathscr{L}_{CF}$ a nech $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je zásobníkový automat taký, že $L(A) = L$. Zostrojíme zásobníkový prekladač M taký, že $M(\{\varepsilon\}) = L$ – priamym dôsledkom potom bude príslušnosť L do $\mathscr{L}$.

Položme $M = (K, \Sigma, \Sigma, \Gamma, H, q_0, Z_0, F)$, kde

$$H = \{(p, \varepsilon, Z, q, z, \gamma) \in K \times \Sigma_1^* \times \Gamma \times K \times \Sigma_2^* \times \Gamma^* \mid (q, \gamma) \in \delta(p, z, Z)\}.$$

Dôkaz rovnosti $M(\{\varepsilon\}) = L$ vynechávame.

Z dokázaného tvrdenia už priamo vyplýva, že napríklad pre dvojicu jazykov $L_1 = \{\varepsilon\}$ a $L_2 = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ neexistuje žiaden zásobníkový prekladač M taký, že $M(L_1) = L_2$. $\square$