Teorie jazyků a automatů— 12. týden

Marie Demlová
http://math.feld.cvut.cz/en/people/demlova

9. 12. 2024

Uzávěrové vlastnosti bezkontextových jazyků

(a) **Tvrzení.**

Třída bezkontextových jazyků je uzavřena na průniky s regulárními jazyky.

(b) **Tvrzení.**

Bezkontextové jazyky jsou uzavřeny na reverzi.

(c) **Tvrzení.**

Třída bezkontextových jazyků je uzavřena na inversní homomorfní obrazy.

Tvz (a) $\forall L \in CF$, $R \in Reg$, tak $L \cap R \in CF$ $\quad L(A) = L$

D.: $L \in CF$, tj. ex. ZA $A = (Q_A, \Sigma, \Gamma_A, \delta_A, q_A, Z_A, F_A)$

$R \in Reg$, tj. ex. DFA $M = (Q_M, \Sigma, \delta_M, F_M)$ $L(M) = R$

→ ZA $B = (Q_B, \Sigma, \Gamma_A, \delta_B, q_B, Z_A, F_B)$

$Q_B = \{(q,p) \mid q \in Q_A \mid p \in Q_M\}$ $q_B = (q_A, q_M)$

$F_B = F_A \times F_M$

$a \in \Sigma \cup \{\varepsilon\}$ $\delta_B((q,p), a, X) \ni ((r,s), \gamma)$ iff $(r,\gamma) \in \delta_A(q, a, X)$

$s = \delta_M^*(p,a) < \begin{matrix} \delta_M(p,a) & \text{if } a \in \Sigma \\ p & \text{if } a = \varepsilon \end{matrix}$

$w \in L(B)$ iff $((q_A, q_M), w, Z_A) \vdash_B^* ((r,s), \varepsilon, \gamma)$

a $(r,s) \in F_B = F_A \times F_M$, tj. $r \in F_A$, $s \in F_M$

Indukcí podle počtu kroků $\vdash_B^*$

$((q,p), u, X\alpha) \vdash_B^* ((z,t), v, \beta)$ iff $(q, u, X\alpha) \vdash_A^*$

$u = xv$

$(z, x, \beta) \cdot \delta_M^*(p, x) = t$

(*) $(q_A, w, Z_A) \vdash_A^* (x, \varepsilon, \gamma)$, $\delta_M^*(q_M, w) = s$

(A) Tvr. Ak $L \in CF$, tak $L^R \in CF$

D $L \in CF$ tj ex. CF $g = (N, \Sigma, S, P)$ $\overset{2}{\Rightarrow} g^R =$

$L(g) = L$

$= (N, \Sigma, S, P^R)$

$P^R = \{X \to \alpha \mid X \to \alpha^R \in P\}$

$L(g^R) = L^R$

$w \in L(g)$ iff $S \Rightarrow_g^* w$, odpovídá levá derivace

$X_1 \to m_2 X_2 \alpha_2$

$S \Rightarrow_g m_1 X_1 \alpha_1 \Rightarrow_g m_1 m_2 X_2 \alpha_2 \alpha_1 \Rightarrow_g \dots \Rightarrow_g m_1 m_2 \dots m_k X_k \nu_k \Rightarrow_g$

$X \to m_{k+1}$

$\Rightarrow_g m_1 m_2 \dots m_k m_{k+1} \nu_k = w$

pravá derivace v g^R

$X_1 \to \alpha_1^R X_2 m_2^R$

$S \Rightarrow_{g^R}^{12} \alpha_1 X_1 m_1^R \Rightarrow \alpha_1^R \alpha_2^R X_2 m_2^R m_1^R \Rightarrow \dots$

$\Rightarrow_{g^R} (\alpha_3 \alpha_2 \alpha_1)^R X_3 (m_1 m_2 m_3)^R \dots \Rightarrow \nu_k^R X_k (m_1 \dots m_k)^R \Rightarrow$

$\Rightarrow \nu_k^R m_{k+1}^R (m_1 \dots m_k)^R = \overbrace{(m_1 \dots m_{k+1} \nu_k)^R}^{w^R}$

(c) substituce: Σ, Π

$\sigma : \Sigma^* \to \mathcal{P}(\Pi^*)$ $\quad \sigma(a) \subseteq \Pi^*$ $\quad a \in \Sigma$ $\quad \sigma(\varepsilon) = \{\varepsilon\}$

$\sigma(wa) = \sigma(w)\sigma(a)$

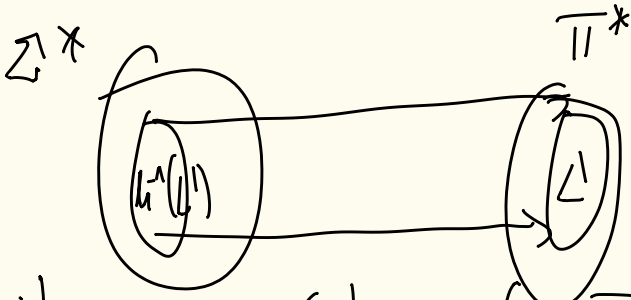
$\sigma(a_1 \dots a_k) = \sigma(a_1) \cdots \sigma(a_k)$

homomorfismus $h : \Sigma^* \to \Pi^*$ $\quad h(a) \in \Pi^*$ $\quad a \in \Sigma$ $\quad h(\varepsilon) = \varepsilon$

$h(wa) = h(w)h(a)$

inverzní homomorfismus:

$L \subseteq \Pi^*$ $\quad h^{-1}(L') = \{w \in \Sigma^* \mid h(w) \in L'\}$



Tv. Je-li σ substituce a $\sigma(a) \in CF\ (\text{nad } \Pi)$ pro každé $a \in \Sigma$ a L je $CF\ (\text{nad } \Sigma)$, pak $\sigma(L)$ je $CF\ (\text{nad } \Pi)$

D. $\forall a \in \Sigma$ a $g_a = (N_a, \Pi, S_a, P_a)$, $L(g_a) = \sigma(a)$

$N_a \cap N_b = \emptyset$ $\quad \forall a \neq b$

ex. $g = (N, \Sigma, S, P)$

$\hookrightarrow g_1$

$w \in L$ iff $S \Rightarrow^* w$

$S \Rightarrow_{g_1}^* S_{a_1} \dots S_{a_k}$

$S \to cSh \mid c$

$S \Rightarrow^* a^k S h^k \Rightarrow a^k c h^k$

$S \to S_a S S h \mid S_c$

$S \Rightarrow S_a^k S_c S h^k$

$g_1 = (N_1, \Pi, S, P_1)$ $\quad P_1 = \bigcup_{a \in \Sigma} P_a$ a pravidla $X \to \gamma$ která vznikla z pravidla $X \to \alpha \in P$, kde každý terminální $a \in \Sigma$ byl nahrazen S_a.

$h : \Sigma^* \to \Pi^*$

$h(a) \in \Pi^*$ $\quad h'(a) = \{h(a)\}$

if $L \in CF$ then $h(L)$ je CF

$h : \Sigma^* \to \Pi^*$ $L_1 \subseteq \Pi^*$ $L_1 CF$, pak $h^{-1}(L_1) \in CF$

$L' \in Reg \to h^{-1}(L') \in Reg$ $\quad f(q,a) = (f')^*(q, h(a))$

Bezkontextové gramatiky

Tvrzení.

Třída bezkontextových jazyků je uzavřena na substituce a tedy i homomorfismy.

Tvrzení.

Třída bezkontextových jazyků je uzavřena na inversní homomorfní obrazy.

Bezkontextové gramatiky

Bezkontextový jazyk L je Dyckův jazyk, jestliže je generován gramatikou $\mathcal{G} = (\{S\}, \Sigma, S, P)$, kde $\Sigma = \{a_1, a'_1, a_2, a'_2, \ldots, a_n, a'_n\}$ a P:

$$S \to \varepsilon \mid SS \mid a_1 S a'_1 \mid \ldots \mid a_n S a'_n$$

Věta.

Pro každý bezkontextový jazyk L existuje regulární jazyk R, Dyckův jazyk D a homomorfismus h takový, že $L = h(R \cap D)$.

Bezkontextové gramatiky

Spojení pravidel. Máme pravidlo $A \to \alpha B \beta$. Jestliže

$$B \to \gamma_1 \mid \ldots \mid \gamma_k$$

všechna pravidla s pravou stranou B

jsou všechna pravidla s levou stranou B, pak nahradíme $A \to \alpha B \beta$ pravidly

$$A \to \alpha\gamma_1\beta \mid \ldots, \mid \alpha\gamma_k\beta$$

Bezkontextové gramatiky

Odstranění levé rekurze. Máme pravidla

$$A \to A\alpha_1,\ A \to A\alpha_2,\ \ldots,\ A \to A\alpha_k$$

a pravidla

$$A \to \beta_1,\ A \to \beta_2,\ \ldots,\ A \to \beta_m$$

jsou všechna ostatní pravidla.

Pak tato pravidla nahradíme pravidly

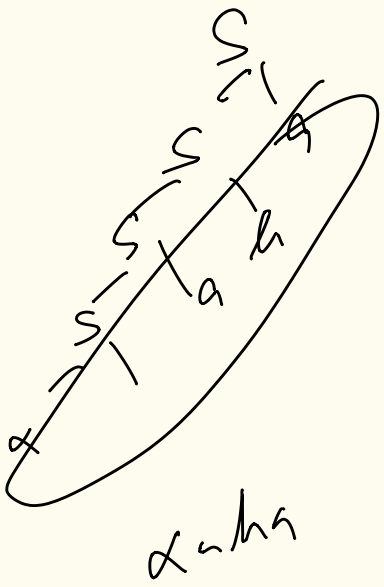
$$A \to \beta_i,\ A \to \beta_i Z,\ Z \to \alpha_j,\ Z \to \alpha_j Z, i = 1, \ldots m, j = 1, \ldots, k;$$

kde Z je nový neterminál.

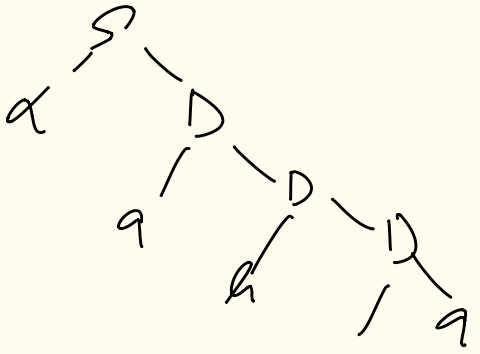
Pr. $S \to Sa \mid Sb \mid aA \mid ab$ $\quad S \Rightarrow Sa \overset{S \to Sb}{\Rightarrow} Sba \ldots$

$S \to aA \mid ab \mid aAD \mid abD$
$D \to aD \mid bD \mid a \mid b$

$S \to aAD \mid abD$
$S \to aD \mid bD \mid \varepsilon$



$\alpha = aba$



Greibachové normální forma (tvar)

CF gramatika $\mathcal{G}$ je v Greibachové normální formě, jestliže všechna pravidla mají tvar:

$$A \to a\alpha, \quad \text{kde } a \in \Sigma \text{ a } \alpha \in N^\star.$$

Věta. Ke každému CF jazyku L existuje CF gramatika v Greibachové normálním tvaru $\mathcal{G}$ taková, že

$$L(\mathcal{G}) = L - \{\varepsilon\}.$$

Greibachové normální forma (tvar)

Postup nalezení gramatiky v Greibachové normálním tvaru.

Máme nevypouštěcí gramatiku $\mathcal{G}$.

(1) Očíslujeme její neterminály, tj. $N = \{A_1, A_2, \ldots, A_n\}$.

(2) Postupným spojováním pravidel a odstraňováním levé rekurze získáme pouze pravidla, kde pravá strana začíná buď terminálem nebo jsou tvaru $A_i \to A_j\alpha$ pro $i < j$.

(3) Zajistíme, aby pravidla obsahující nové neterminály (vzniklé při odstraňování levé rekurze) také měla požadovaný tvar.

Greibachové normální forma (tvar)

Postup nalezení gramatiky v Greibachové normálním tvaru – 2.

(4) Postupným spojováním pravidel (od n do 1) dosáhneme toho, aby pravidla s levou stranou A_i měla tvar $A_i \to a\alpha$, kde $a \in \Sigma$.

(5) Jestliže nyní máme terminální symbol uvnitř pravé strany některého pravidla, zavedeme za něj nový neterminál.

Greibachové normální forma (tvar)

Příklad. Je dána bezkontextová gramatika $\mathcal{G} = (N, \Sigma, S, P)$, kde $N = \{E, T, F\}$, $\Sigma = \{+, \star, (,), a\}$, $S = E$ a

$$\begin{aligned} P: \quad & E \to E + T \mid T \\ & T \to T \star F \mid F \\ & F \to (E) \mid a \end{aligned}$$

Gramatiku převeďte do Greibachové normálního tvaru.