

Preklad

Peter Kostolányi

15. apríla 2025

1 Základné definície

Intuitívne si pod realizáciou prekladu možno predstaviť ľubovoľnú transformáciu slov u nad nejakou abecedou Σ_1 na slová v nad nejakou vo všeobecnosti inou abecedou Σ_2 . Niektoré slová u môžu mať aj viacero prípustných prekladov v ; iné zas nemusia byť preložiteľné vôbec. Preklad teda možno definovať ako reláciu $\Pi \subseteq \Sigma_1^* \times \Sigma_2^*$ takú, že $(u, v) \in \Pi$ práve vtedy, keď je slovo v prípustným prekladom slova u .

Definícia 1. Nech Σ_1 a Σ_2 sú abecedy. *Preklad* zo Σ_1^* do Σ_2^* je relácia $\Pi \subseteq \Sigma_1^* \times \Sigma_2^*$.

Keďže je preklad definovaný ako binárna relácia – čiže množina dvojíc – je užitočné mať k dispozícii mechanizmus umožňujúci pristupovať k jednotlivým komponentom. Na tento účel možno využiť štandardný pojem *i-tej projekcie*.

Definícia 2. Nech $M_1, M_2, \dots, M_n$ sú množiny a $i \in [n]$. Pod *i-tou projekciou* na karteziánskom súčine $M_1 \times M_2 \times \dots \times M_n$ rozumieme zobrazenie $\text{pr}_i: M_1 \times M_2 \times \dots \times M_n \rightarrow M_i$ dané pre všetky $(x_1, x_2, \dots, x_n) \in M_1 \times M_2 \times \dots \times M_n$ ako $\text{pr}_i(x_1, x_2, \dots, x_n) = x_i$. Túto definíciu možno prirodzeným spôsobom rozšíriť aj na relácie $R \subseteq M_1 \times M_2 \times \dots \times M_n$ – *i-ta projekcia* n -árnej relácie R je potom daná ako $\text{pr}_i(R) = \{\text{pr}_i(x_1, x_2, \dots, x_n) \mid (x_1, x_2, \dots, x_n) \in R\}$.

Nech teda Σ_1 a Σ_2 sú abecedy. Potom pre $(u, v) \in \Sigma_1^* \times \Sigma_2^*$ máme $\text{pr}_1(u, v) = u$ a $\text{pr}_2(u, v) = v$. Uvažujme teraz preklad Π daný ako $\Pi = \{a^n b^n, a^n b^n c^n \mid n \in \mathbb{N}\}$. Potom $\text{pr}_1(\Pi) = \{a^n b^n \mid n \in \mathbb{N}\}$ a $\text{pr}_2(\Pi) = \{a^n b^n c^n \mid n \in \mathbb{N}\}$. Prvá projekcia prekladu zo Σ_1^* do Σ_2^* je teda jazyk pozostávajúci zo všetkých slov nad abecedou Σ_1 , ktoré majú aspoň jeden prípustný preklad. Druhá projekcia je zas jazyk obsahujúci všetky slová nad abecedou Σ_2 , ktoré sú prípustným prekladom aspoň jedného slova nad abecedou Σ_1 .

2 Preklad realizovaný a-prekladačom

A-prekladače sme až doposiaľ chápali najmä ako zariadenia na transformáciu jazykov; zaujímali nás teda hlavne vlastnosti zobrazenia $L \mapsto M(L)$ pre daný a-prekladač M . A-prekladače sú ale v prvom rade zariadeniami realizujúcimi preklad.

Definícia 3. Nech $M = (K, \Sigma_1, \Sigma_2, H, q_0, F)$ je a-prekladač. *Preklad realizovaný a-prekladačom* M je relácia $\Pi(M) \subseteq \Sigma_1^* \times \Sigma_2^*$ daná ako

$$\Pi(M) = \{(u, v) \in \Sigma_1^* \times \Sigma_2^* \mid \exists q \in F : (q_0, u, \varepsilon) \vdash_M^* (q, \varepsilon, v)\}.$$

Preklady realizované a-prekladačmi sa niekedy zvyknú nazývať aj *racionálne relácie*. Pri takomto ponímaní možno a-prekladač považovať jednoducho za konečný automat nad monoidom $\Sigma_1^* \times \Sigma_2^*$ – teória a-prekladačov je tak jedným z hlavných východísk teórie automatov nad monoidmi.

Nasledujúce viac-menej očividné pozorovanie je užitočným nástrojom na dokazovanie, že daný preklad Π nie je realizovaný žiadnym a-prekladačom.

Tvrdenie 1. Nech $M = (K, \Sigma_1, \Sigma_2, H, q_0, F)$ je a-prekladač a $\Pi = \Pi(M)$. Potom sú obidva jazyky $\text{pr}_1(\Pi)$ a $\text{pr}_2(\Pi)$ regulárne.

Dôkaz. Podľa tvrdenia dokázaného na treťom cvičení možno bez ujmy na všeobecnosti predpokladať, že a-prekladač M je 1-ohraničený – čiže $H \subseteq K \times (\Sigma_1 \cup \{\varepsilon\}) \times (\Sigma_2 \cup \{\varepsilon\}) \times K$. Ľahko potom vidieť, že $\text{pr}_1(\Pi) = L(A_1)$ pre nedeterministický konečný automat $A_1 = (K, \Sigma_1, \delta_1, q_0, F)$ taký, že pre všetky $p \in K$ a $x \in \Sigma_1 \cup \{\varepsilon\}$ je $\delta_1(p, x) = \{q \in K \mid \exists y \in \Sigma_2 \cup \{\varepsilon\} : (p, x, y, q) \in H\}$. Podobne $\text{pr}_2(\Pi) = L(A_2)$ pre nedeterministický konečný automat $A_2 = (K, \Sigma_2, \delta_2, q_0, F)$ taký, že pre všetky $p \in K$ a $y \in \Sigma_2 \cup \{\varepsilon\}$ je $\delta_2(p, y) = \{q \in K \mid \exists x \in \Sigma_1 \cup \{\varepsilon\} : (p, x, y, q) \in H\}$. $\square$

3 Jednoduché prekladové schémy

Pod *jednoduchou prekladovou schémou* rozumieme bezkontextovú gramatiku, ktorá „pracuje na dvoch vetných formách naraz“. Postupnosť neterminálov v oboch vetných formách musí byť počas celého odvodenia rovnaká – vetné formy sa teda môžu líšiť iba v termináloch.

V každom kroku odvodenia sa vyberie jeden výskyt neterminálu, ktorý sa prepíše v oboch vetných formách podľa niektorého pravidla prekladovej schémy. Takéto pravidlo má dve pravé strany, ktoré sú vo všeobecnosti rôzne – prvá pravá strana sa použije na prvú generovanú vetnú formu a druhá na druhú. Aby ale postupnosti neterminálov ostali v oboch vetných formách rovnaké aj po použití pravidla, budú za prípustné považované iba tie pravidlá, ktorých pravé strany majú rovnaké postupnosti neterminálov.

Definícia 4. *Jednoduchá syntaxou riadená prekladová schéma* alebo skrátené *jednoduchá prekladová schéma* je päťica $\mathcal{S} = (N, T_1, T_2, P, \sigma)$, kde N je abeceda neterminálov, T_1 a T_2 sú abecedy terminálov, $N \cap (T_1 \cup T_2) = \emptyset$, $\sigma \in N$ je počiatočný neterminál a $P \subseteq N \times (N \cup T_1)^* \times (N \cup T_2)^*$ je konečná množina prepisovacích pravidiel taká, že pre všetky $(\xi, x, y) \in P$ je $h_N(x) = h_N(y)$ pre homomorfizmus $h_N: (N \cup T_1 \cup T_2)^* \rightarrow N^*$ daný ako $h(\xi) = \xi$ pre $\xi \in N$ a $h(c) = \varepsilon$ pre $c \in T_1 \cup T_2$.

Poznámka 1. Podobne ako pri bezkontextových gramatikách zvyčajne zapisujeme prepisovacie pravidlá $(\xi, x, y) \in P$ ako $\xi \rightarrow (x, y)$. Zápis $\xi \rightarrow (x_1, y_1), \xi \rightarrow (x_2, y_2), \dots, \xi \rightarrow (x_k, y_k)$ skrácujeme ako $\xi \rightarrow (x_1, y_1) \mid (x_2, y_2) \mid \dots \mid (x_k, y_k)$.

Definícia 5. Nech $\mathcal{S} = (N, T_1, T_2, P, \sigma)$ je jednoduchá prekladová schéma. *Krok odvodenia* v prekladovej schéme $\mathcal{S}$ je binárna relácia $\Rightarrow_{\mathcal{S}}$ na $(N \cup T_1)^* \times (N \cup T_2)^*$ taká, že pre všetky $u, u' \in (N \cup T_1)^*$ a $v, v' \in (N \cup T_2)^*$ je

$$(u, v) \Rightarrow_{\mathcal{S}} (u', v')$$

práve vtedy, keď existujú slová $u_1, u_2, x \in (N \cup T_1)^*$, $v_1, v_2, y \in (N \cup T_2)^*$ a neterminál $\xi \in N$ tak, že $(u, v) = (u_1 \xi u_2, v_1 \xi v_2)$, $(u', v') = (u_1 x u_2, v_1 y v_2)$, $\xi \rightarrow (x, y) \in P$ a $h_N(u_1) = h_N(v_1)$ pre homomorfizmus $h_N: (N \cup T_1 \cup T_2)^* \rightarrow N^*$ daný ako $h(\xi) = \xi$ pre $\xi \in N$ a $h(c) = \varepsilon$ pre $c \in T_1 \cup T_2$. Ak je prekladová schéma $\mathcal{S}$ zrejme z kontextu, píšeme namiesto $\Rightarrow_{\mathcal{S}}$ často iba $\Rightarrow$.

Poznámka 2. Podmienka $h_N(u_1) = h_N(v_1)$ v predchádzajúcej definícii zaručuje, že sa v oboch generovaných vetných formách prepíše rovnaký výskyt neterminálu ξ .

Definícia 6. Nech $\mathcal{S} = (N, T_1, T_2, P, \sigma)$ je jednoduchá prekladová schéma. *Preklad realizovaný prekladovou schémou* $\mathcal{S}$ je relácia $\Pi(\mathcal{S}) \subseteq T_1^* \times T_2^*$ daná ako

$$\Pi(\mathcal{S}) = \{(u, v) \in T_1^* \times T_2^* \mid (\sigma, \sigma) \Rightarrow_{\mathcal{S}}^* (u, v)\}.$$

Analógiou tvrdenia 1 je pre jednoduché prekladové schémy nasledujúce pozorovanie.

Tvrdenie 2. Nech $\mathcal{S} = (N, T_1, T_2, P, \sigma)$ je jednoduchá prekladová schéma a $\Pi = \Pi(\mathcal{S})$. Potom sú obidva jazyky $\text{pr}_1(\Pi)$ a $\text{pr}_2(\Pi)$ bezkontextové.

Dôkaz. Nech $G_1 = (N, T_1, P_1, \sigma)$ je bezkontextová gramatika s množinou prepisovacích pravidiel P_1 danou ako $P_1 = \{\xi \rightarrow x \mid \exists y \in (N \cup T_2)^*: \xi \rightarrow (x, y) \in P\}$. Zrejme $L(G_1) = \text{pr}_1(\Pi)$. Podobne pre bezkontextovú gramatiku $G_2 = (N, T_2, P_2, \sigma)$ s $P_2 = \{\xi \rightarrow y \mid \exists x \in (N \cup T_1)^*: \xi \rightarrow (x, y) \in P\}$ je $L(G_2) = \text{pr}_2(\Pi)$. $\square$

Podobne ako pri a-prekladačoch sa aj na jednoduché prekladové schémy môžeme dívať ako na operácie na jazykoch; má teda zmysel definovať obraz jazyka pri zobrazení jednoduchou prekladovou schémou.

Definícia 7. Nech $\mathcal{S} = (N, T_1, T_2, P, \sigma)$ je jednoduchá prekladová schéma a $L \subseteq T_1^*$ je jazyk. *Obraz jazyka* L pri zobrazení prekladovou schémou $\mathcal{S}$ je jazyk

$$\mathcal{S}(L) = \{v \in T_2^* \mid \exists u \in L : (\sigma, \sigma) \Rightarrow_{\mathcal{S}}^* (u, v)\}.$$

4 Preklady a transformácie jazykov

Máme teda k dispozícii dve zariadenia na preklad jazykov: a-prekladače a jednoduché prekladové schémy. Pre obidve tieto zariadenia máme definovaný nimi realizovaný preklad, ako aj obraz jazyka pri zobrazení daným zariadením. Nasledujúce tvrdenia vyplývajú priamo z definícií uvedených vyššie; ich dôkazy preto prenechávame čitateľovi.

Tvrdenie 3. *Nech $M = (K, \Sigma_1, \Sigma_2, H, q_0, F)$ je a-prekladač a $L \subseteq \Sigma_1^*$. Potom*

$$M(L) = \{v \in \Sigma_2^* \mid \exists u \in L: (u, v) \in \Pi(M)\}.$$

Tvrdenie 4. *Nech $\mathcal{S} = (N, T_1, T_2, P, \sigma)$ je jednoduchá prekladová schéma a $L \subseteq T_1^*$. Potom*

$$\mathcal{S}(L) = \{v \in T_2^* \mid \exists u \in L: (u, v) \in \Pi(\mathcal{S})\}.$$

Úloha 1. Nech $L_1 = \{a^n b^n \mid n \in \mathbb{N}\}$ a $L_2 = \{a^n b^n c^n \mid n \in \mathbb{N}\}$. Dokážte alebo vyvráťte nasledujúce tvrdenia:

- a) Existuje jednoduchá prekladová schéma $\mathcal{S}$ taká, že $\text{pr}_1(\Pi(\mathcal{S})) = L_1$ a $\text{pr}_2(\Pi(\mathcal{S})) = L_2$.
- b) Existuje jednoduchá prekladová schéma $\mathcal{S}$ taká, že $\mathcal{S}(L_1) = L_2$.

Riešenie.

- a) Takáto prekladová schéma neexistuje, pretože v opačnom prípade by podľa tvrdenia 2 musel byť jazyk L_2 bezkontextový.
- b) Zjavne stačí zostrojiť prekladovú schému $\mathcal{S} = (N, T_1, T_2, P, \sigma)$ takú, že

$$\Pi(\mathcal{S}) = \{(a^i b^j, a^i b^j c^j) \mid i, j \in \mathbb{N}\}.$$

V takom prípade totiž priamo z tvrdenia 4 dostávame $\mathcal{S}(L_1) = L_2$.

Schému $\mathcal{S}$ zostrojíme nasledovne: $N = \{\sigma, \alpha, \beta\}$, $T_1 = \{a, b\}$, $T_2 = \{a, b, c\}$ a

$$\begin{aligned} P = \{ & \sigma \rightarrow (\alpha\beta, \alpha\beta) \\ & \alpha \rightarrow (a\alpha, a\alpha) \mid (\varepsilon, \varepsilon) \\ & \beta \rightarrow (b\beta, b\beta c) \mid (\varepsilon, \varepsilon) \}. \end{aligned}$$

Zrejme ide o korektnú prekladovú schému a $\Pi(\mathcal{S}) = \{(a^i b^j, a^i b^j c^j) \mid i, j \in \mathbb{N}\}$. □