Uzávěrové vlastnosti bezkontextových jazyků

Zkoumání bezkontextových jazyků a gramatik, případně zásobníkových automatů se blíží ke konci. V minulých kapitolách jsme si vztah všech tří pojmů přiblížili, nyní se ještě budeme zabývat množinou (resp. třídou) bezkontextových jazyků obecně. Uvedeme několik sdělení týkajících se jejich uzavřenosti na různé unární i binární množinové operace a naznačíme si důkazy. Nejdříve vše přehledně zapíšeme:

Věta 1 (uzavřenost CFL na operace).

Třída bezkontextových jazyků je uzavřena na operace:

a) sjednocení

b) zřetězení

c) iteraci

d) pozitivní iteraci

e) průnik s regulárním jazykem

Věta 2 (neuzavřenost CFL na operace).

Třída bezkontextových jazyků není uzavřena na operace:

a) průnik

b) doplněk do jazyka (komplement)

Důkaz (1a: sjednocení).

Mějme bezkontextové jazyky L_1, L_2. Sestrojíme pro ně bezkontextové gramatiky $G_1 = (N_1, \Sigma_1, P_1, S_1)$ a $G_2 = (N_2, \Sigma_2, P_2, S_2)$. Předpokládejme navíc, že $N_1 \cap N_2 = \emptyset$.

Vytvoříme novou gramatiku $G = (N_1 \cup N_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, P, S)$, kde $S \notin N_1 \cup N_2$ a pravidla P jsou následující:

$$P = P_1 \cup P_2 \cup \{S \to S_1 \mid S_2\}$$

Na počátku libovolné derivace v gramatice G se musíme rozhodnout, jakou cestou se vydáme, zda vybereme průchod gramatikou G_1 či G_2. Předpoklad $N_1 \cap N_2 = \emptyset$ je nezbytný, v derivaci nemůžeme mít možnost používat střídavě pravidla P_1 nebo P_2. Je patrné, že gramatika G generuje jazyk $L_1 \cup L_2$.

Důkaz (1b: zřetězení).

Buďte L_1, L_2 bezkontextové jazyky pro něž máme bezkontextové gramatiky $G_1 = (N_1, \Sigma_1, P_1, S_1)$ a $G_2 = (N_2, \Sigma_2, P_2, S_2)$. Předpokládejme opět, že $N_1 \cap N_2 = \emptyset$.

Sestrojíme bezkontextovou gramatiku $G = (N_1 \cup N_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, P, S)$, kde $S \notin N_1 \cup N_2$ je nový neterminál a pravidla P jsou následující:

$$P = P_1 \cup P_2 \cup \{S \to S_1 S_2\}$$

Řetězení slov z L_1, L_2 zajistíme úvodní derivací $S \Rightarrow S_1 S_2$. Protože $N_1 \cap N_2 = \emptyset$, nemůže se stát, že bychom v derivačním podstromu pro kořen S_1 použili pravidla gramatiky G_2 a naopak. Je patrné, že $L(G) = L_1.L_2$.

Důkaz (1c: iterace).

Buď L bezkontextový jazyk a $G = (N, \Sigma, P, S)$ bezkontextová gramatika, která jej generuje. Vytvoříme gramatiku $G^* = (N \cup \{S^*\}, \Sigma, P^*, S^*)$ s novým neterminálem S^*, kde pravidla z P obohatíme o dvě další:

$$P^* = P \cup \{S^* \to SS^* \mid \varepsilon\}$$

Na začátku každé derivace použijeme n-krát ($n \in \mathbb{N}_0$) pravidlo $S^* \to SS^*$ a poté $S^* \to \varepsilon$ – z S^* jsme odvodili řetězec n neterminálů S. Každé S poté derivujeme na slovo $w \in L$, tedy celkově z řetězce S^n odvodíme jazyk L^n. Jelikož n je libovolné, dostáváme $L(G^*) = L^* = \bigcup_{i=0}^{n} L^i$.

Důkaz (1d: pozitivní iterace).

Je analogií předchozího důkazu pro iteraci. Máme-li pro jazyk L bezkontextovou gramatiku $G = (N, \Sigma, P, S)$, pak novou gramatiku $G^+ = (N \cup \{S^+\}, \Sigma, P^+, S^+)$ vytvoříme tak, že k původním pravidlům P přidáme $S^+ \to S \mid S^+S$. Použitím stejné myšlenky jako v předchozím důkazu dostáváme $L(G^+) = L^+$.

Důkaz (1e: průnik s regulárním jazykem).

Buď L_R regulární jazyk a L_B bezkontextový jazyk.

1. K jazyku L_R sestrojíme konečný automat $M_R = (Q_R, \Sigma, \delta_R, q_0^R, F_R)$ a
2. pro jazyk L_B sestrojíme zásobníkový automat $M_B = (Q_B, \Sigma, \Gamma, \delta_B, q_0^B, Z_0, F_B)$.

PDA P pro jazyk $L_R \cap L_B$ zkonstruujeme pomocí synchronní paralelní kompozice:
$P = (Q', \Sigma, \Gamma, \delta', q_0', Z_0, F')$, kde

- $Q' = Q_R \times Q_B$,
- $q_0' = (q_0^R, q_0^B)$,
- $F' = F_R \times F_B$ – P akceptuje, pokud dojde do koncových stavů obou automatů současně,
- δ' sestrojíme takto: pro libovolné stavy $p \in Q_R, q \in Q_B$, slovo $w \in \Sigma$ a zásobníkový symbol $Z \in \Gamma$:

  $$\delta'((p, q), w, Z) = ((r, s), \alpha) \iff \delta_R(p, w) = r \wedge \delta_B(q, w, Z) = (s, \alpha),$$

  kde $r \in Q_R, s \in Q_B, \alpha \in \Gamma^*$.

Výsledný PDA P provádí synchronní paralelní kompozici automatů M_R, M_B čili akceptuje jazyk $L_R \cap L_B$. Důležité však je, že $L_R \cap L_B$ je tedy bezkontextový.

Příklad 1.

Ukažte pomocí uzávěrových vlastností, že jazyk $L = \{w \in \{a, b, c\}^* \mid \#_a(w) = \#_b(w) = \#_c(w)\}$ není bezkontextový.

Řešení. Stačí si uvědomit, že

$$L \cap a^*b^*c^* = \{a^n b^n c^n \mid n \in \mathbb{N}_0\}.$$

Jazyk $a^*b^*c^*$, jazyk $\{a^n b^n c^n \mid n \in \mathbb{N}_0\}$ není bezkontextový. Pokud bychom předpokládali, že L je bezkontextový, pak dle uzavřenosti třídy CFL na průnik s regulárním jazykem musí nutně být bezkontextový i jazyk $\{a^n b^n c^n \mid n \in \mathbb{N}_0\}$. To však určitě není. L tedy není bezkontextový.

Důkaz (2a: průnik).

Protipříkladem: Buďte

$$L_1 = \{a^n b^n c^m \mid m, n \geq 1\},$$
$$L_2 = \{a^m b^n c^n \mid m, n \geq 1\},$$

které jsou bezkontextové. Uvažujeme-li jejich průnik $L = L_1 \cap L_2$, zjistíme, že je to jazyk $\{a^n b^n c^n \mid n \geq 1\}$. Ten ale není podle Pumping lemmatu bezkontextový.

Poznámka.

Jistě budete namítat, jak je možné, že třída CFL je uzavřena na průnik s regulárním jazykem, ale už ne s bezkontextovým jazykem. Vysvětlení je nasnadě: provádíme-li synchronní paralelní kompozici dvou zásobníkových automatů, nevíme si rady s tím, jak ze dvou původních zásobníků „udělat“ jeden.

Důkaz (2b: doplněk).

Předpokládejme sporem, že třída bezkontextových jazyků je uzavřena na doplněk. Mějme dva bezkontextové jazyky L_1, L_2. Dle Morganových pravidel platí

$$L_1 \cap L_2 = \text{co-}(\text{co-}L_1 \cup \text{co-}L_2).$$

Předchozí rovnost je založena pouze na konstrukci sjednocení a doplňků do množin L_1 či L_2, díky níž získáme jazyk $L_1 \cap L_2$. Protože jsme předpokládali, že třída CFL je uzavřena na operaci doplněk, je i jazyk $L_1 \cap L_2$ bezkontextový. To je však spor: třída CFL není uzavřena na operaci průnik. Předpoklad uzavřenosti na doplněk je tedy chybný a třída bezkontextových jazyků není uzavřena na doplněk.

Příklad 2.

O každém z následujících tvrzení rozhodněte, zda jsou pravdivá či ne. Svá rozhodnutí vysvětlete.

1. Jsou-li L_1, L_2 bezkontextové jazyky, pak je $L_1 \cup L_2$ kontextový jazyk.
2. Je-li L_1 bezkontextový a $L_1 \cap L_2$ není bezkontextový, pak L_2 není také bezkontextový.
3. Je-li L_1 regulární a L_2 bezkontextový jazyk, pak jazyk co-$(L_1 \cap L_2)$ je bezkontextový.
4. Je-li L_1 konečný a L_2 bezkontextový jazyk, pak jazyk co-$(L_1 \cap L_2)$ je bezkontextový.