KMI/FJ – cvičení 10

1 Gramatiky typu 0 a 1

- jazyky, které generují gramatiky typu 0 a 1, jsou na rámec tohoto kurzu, více se o nich dozvíte v předmětu *Vyčíslitelnost a složitost*

- gramatika $\mathcal{G}$ je *kontextově závislá* (typ 1), pokud jsou všechna pravidla ve tvaru:
  1. $\alpha A \beta \to \alpha\gamma\beta$, kde $A \in N$, $\gamma \neq \varepsilon$ a $\alpha, \gamma, \beta \in (\Sigma \cup N)^*$ (levá strana pravidla může obsahovat společně s přepisovaným neterminálem A i další symboly, které podmiňují aplikování daného pravidla, ale samy nemohou být přepsány)
  2. $S \to \varepsilon$, pokud se S nevyskytuje na pravé straně jakéhokoliv pravidla (pouze počáteční neterminál může generovat generovat prázdný řetězec)

Nezkracující gramatika

$\alpha \to \beta$, $|\alpha| \leq |\beta|$ $(S \to \varepsilon)$

- mějme mějme gramatiku $\mathcal{G}_1 = (\{S, A, B, C, W, Z\}, \{a, b, c\}, P, S\})$:

$$\begin{aligned} S &\to aBC \mid aSBC \\ CB &\to CZ \\ CZ &\to WZ \\ WZ &\to WC \\ WC &\to BC \end{aligned} \qquad \begin{aligned} aB &\to ab \\ bB &\to bb \\ bC &\to bc \\ cC &\to cc \end{aligned}$$

$aA \to Aa$

$X_a \to a$

$X_a A \to X_1 A$

$X_1 A \to X_1 X_2$

$X_1 X_2 \to A X_2$

$A X_2 \to A X_a$

  1. jaký jazyk generuje gramatika $\mathcal{G}_1$? $\{a^n b^n c^n \mid n \geq 1\}$
  2. rozepište celou derivaci řetězce *aabbcc*

S -> aSBC -> aaBCBC -> aaBCZC -> aaBWZC -> aaBWCC -> aaBBCC -> aabBCC -> aabbCC -> aabbcC -> aabbcc

$aabcBC \to aabcBC$ X

- pokud na pravidla neklademe žádné omezení, a tedy povolujeme libovolné pravidlo $\alpha \to \beta$ pro $\alpha, \beta \in (\Sigma \cup N)^*$, pak taková gramatika $\mathcal{G}$ je *bez omezení* (typ 0)

- mějme mějme gramatiku $\mathcal{G}_2 = (\{S, A, B, C, D, E\}, \{a, b\}, P, S\})$:

$$\begin{aligned} S &\to ABC \\ AB &\to aAD \mid bAE \mid \varepsilon \\ DC &\to BaC \\ EC &\to BbC \\ Da &\to aD \end{aligned} \qquad \begin{aligned} Db &\to bD \\ Ea &\to aE \\ Eb &\to bE \\ C &\to \varepsilon \\ aB &\to Ba \\ bB &\to Bb \end{aligned}$$

  1. jaký jazyk generuje gramatika $\mathcal{G}_2$? $\{ww \mid w \in \{a, b\}^*\}$
  2. rozepište celou derivaci řetězce *baabaa*

$S \to ABC \to bAEC \to bABbC \to baADbC \to baAbDC \to baAbBaC \to$

$\to baABbaC \to baaADbaC \to baaAbDaC \to baaAbaDC \to baaAbaBaC \to$

$\to baaAbBaaC \to baaABbaaC \to baabaaC \to baabaa$

2 Pumping lemma pro bezkontextové jazyky

- analogicky k pumping lemma pro regulární jazyky lze vyslovit pumping lemma i pro bezkontextové jazyky

- **Pumping lemma pro CFL:** nechť L je bezkontextový jazyk, pak existuje číslo $n \in \mathbb{N}$ tak, že každý řetězec $s \in L$ délky aspoň n lze rozdělit na tři části $s = uvxyz$, které splňují: pětl

  1. $|vy| > 0$
  2. $|vxy| \leq n$
  3. $uv^ixy^iz \in L$ pro všechna $i \geq 0$

xy^iz

- intuice: L je bezkontextový, proto existuje PDA P, který jej rozpoznává; pokud máme dostatečně dlouhý řetězec $s \in L$, pak se v něm musí nějaký podřetězec v opakovat, přičemž P si na zásobníku uloží výskyty tohoto podřetězce, a později je může srovnat s výskyty jiného podřetězce y (proto mezi počtem opakování v a y může být závislost)

- zásobník umožňuje pouze LIFO přístup, a proto P nemůže zároveň srovnávat výskyty více jak jedné dvojice; navíc řetězce ze zásobníku čteme v opačném pořadí, než byly uloženy (proto L_1 není CFL, přitom $\{ww^R \mid w \in \{0, 1\}^*\}$ je CFL)

- dokažte, že následující jazyky nejsou bezkontextové:

$L_1 = \{ww \mid w \in \{0, 1\}^*\}$

X $0^n 1 0^n 1$

$u = 0^{n-1}$
$v = 0$
$x = 1$
$y = 0$
$z = 0^{n-1}$

l $u = 0^i$
$v = 0$
$x = 0^{n-i-1}10$
$y = 0$
$z = 0^{n-i-1}1$

✓ $0^n 1^n 0^n 1^n \in L_1$

a) v a y jen jsou ve sobě 0, nebo 1

b) v je jen blok, y je druhý blok

c) x obsahuje 1^n

$u = 0^{n-1}$
$v = 0$
$x = 1^n$
$y = 0$
$z = 0^{n-1}1^n$

$|vxy| = n+2$

$L_2 = \{0^i \# 0^j \# 0^k \mid 0 < i < j < k\}$

$\# v \# y$ $\quad s = 0^n \# 0^{n+1} \# 0^{n+2}$

$u = 0^n \# 0^n$
$v = 0$
$x = \#$
$y = 0$
$z = 0^{n+1}$

$u \# v$

$vy\#\#$	$u\#y\#$	$\#vy\#$	$v\#\#y$	$\#v\#y$	$\#\#vy$
$i > j$	$i > k$	$j > k$	$i > j$	$i = j$	$j = k$

$0^n \# 0^n \# 0^{n+1} = u v^i x y^i z \notin L_2$

$\#\#vy$

$u = 0^n \# 0^n \#$
$v = 0$
$x = \varepsilon$
$y = \varepsilon$
$z = 0^{n+1}$

$u v^0 x y^0 z = 0^n \# 0^{n-1} \# 0^{n+1}$

3 Proč neplatí opačná implikace v pumping lemma?

- pumping lemma je tvrzení ve tvaru implikace: „pokud je jazyk regulární, případně bezkontextový, pak splňuje nějaké podmínky“

- lze ukázat, že opačná implikace neplatí (tedy neplatí tvrzení: „pokud platí podmínky v lemmatu, pak je jazyk regulární/bezkontextový“)

- pro jednoduchost uvažujme pumping lemma pro regulární jazyky, ale protipříklad můžeme najít i pro bezkontextové jazyky

- chceme najít jazyk, který splňuje pumping lemma, ale přitom není regulární

- mějme abecedu $\Sigma = \{a, b\}$, nějaký jazyk $L \subseteq \{b\}^*$ a jazyk $L_p = (\{a\}^+ \cdot L) \cup \{b\}^*$; dokažte následující tvrzení:
  1. L_p je regulární právě tehdy, když $(\{a\}^+ \cdot L)$ je regulární

     $\Leftarrow$: 'v' je uzávěrová operace RL

     $\Rightarrow$: $(\{a\}^+ \cdot L) = L_p \setminus \{b\}^*$

     "/" je uzávěrová operace RL

  2. $(\{a\}^+ \cdot L)$ je regulární právě tehdy, když L je regulární

     '$\Leftarrow$' průnik je uzávěrová operace RL

     '$\Rightarrow$'

  3. pokud L není regulární, pak L_p také není regulární a zároveň splňuje podmínky pumping lemma

$L = \{b^{2^n} \mid n \geq 0\}$

$n = 1$

a) $w = b^l$, $l \geq 1$ $\quad |xy| \leq n$ $\quad y \neq \varepsilon$

$x = \varepsilon$
$y = b$
$z = b^{l-1}$ $\quad xy^iz \in L_p$ $\quad (\{a\}^+ \subseteq L_p)$

b) $w = a^k b^l$

$x = \varepsilon$
$y = a$
$z = a^{k-1} b^l$

$xy^iz \in L_p$ $\quad (\{a\}^+ \cdot L \subseteq L_p)$

3 Pumping lemma

- užitečný nástroj, který umožňuje dokázat, že daný jazyk **není** regulární (neumožňuje dokázat, že jazyk je regulární)

- **Definice:** nechť L je regulární jazyk, pak existuje číslo $n \in \mathbb{N}$, tak že každý řetězec $w \in L$ délky aspoň n lze rozdělit na tři části $w = xyz$, které splňují:
  1. $|y| > 0$
  2. $|xy| \leq n$
  3. $xy^iz \in L$ pro všechna $i \geq 0$

- pokud je jazyk L konečný (a tedy i regulární), pak lemma platí triviálně – n zvolíme větší než je délka nejdelšího řetězce v L

- toto lemma se používá při důkazu sporem – nejprve předpokládáme, že daný jazyk je regulární a snažíme se dojít ke sporu s některou částí lemmatu

- dokažte, že jazyky nejsou regulární:
  1. $L_1 = \{ww \mid w \in \{0, 1\}^*\}$
  2. $L_2 = \{a^p \mid p \text{ je prvočíslo}\}$