

Bezkontextové jazyky a gramatiky, CYK

BI-AAG (8.cvičení)

Bc. Eliška Šestáková
sestaeli@fit.cvut.cz

13.11.2014

- Klasifikace jazyků a gramatik, množina bezkontextových jazyků
- Nejednoznačná (víceznačná) / jednoznačná BG
- Typy BG (bez cyklů, vlastní, redukováná, \dots)
- Transformace BG (vyloučení ϵ -pravidel, jednoduchých pravidel, nedosažitelných a zbytečných symbolů, rekurzivity zleva)
- CNF – Chomského normální tvar
- CYK (Cocke-Younger-Kasami) algoritmus
- * Normální tvar podle Greibachové

Nejednoznačná (víceznačná) / jednoznačná BG

- $G_1 = (\{S\}, \{0, 1\}, P, S)$, kde P :

- $S \rightarrow 0S1 \mid SS \mid \epsilon$

$$w_1 = 001101$$

- $G_2 = (\{S\}, \{+, *, 0, 1, \dots, 9\}, P, S)$, kde P :

- $S \rightarrow S + S \mid S * S \mid 0 \mid 1 \mid \dots \mid 9$

$$w_2 = 3 + 4 * 5$$

pozn. Zkuste upravit gramatiky tak, aby byly jednoznačné.

Názvosloví spojené s bezkontextovými gramatikami

- bez ϵ -pravidel (tj. pravidel $A \rightarrow \epsilon$)
- bez jednoduchých pravidel (tj. pravidel $A \rightarrow B$, $A, B \in N$)
- redukovaná – neobsahuje zbytečné symboly
- bez cyklů – neexistuje derivace $A \Rightarrow^+ A$
(bez jednoduchých a ϵ -pravidel \Rightarrow bez cyklů)
- vlastní (= bez cyklů, bez ϵ -pravidel, bez zbytečných symbolů)

Příklad 8.1

Je dána gramatika $G = (\{S, A, B, C\}, \{b, c\}, P, S)$, kde P obsahuje pravidla:

- $S \rightarrow ABC$
- $A \rightarrow aA \mid \epsilon$
- $B \rightarrow bB \mid \epsilon$
- $C \rightarrow \epsilon$

Odstraňte z této gramatiky ϵ -pravidla.

Příklad 8.2

Je dána gramatika $G = (\{S, A, B, C, D\}, \{a, b, c, d\}, P, S)$, kde P obsahuje pravidla:

- $S \rightarrow A \mid B$
- $A \rightarrow C \mid aA \mid bS$
- $B \rightarrow D \mid cB \mid dS$
- $C \rightarrow bC \mid a$
- $D \rightarrow dD \mid c$

Odstraňte z této gramatiky jednoduchá pravidla.

Příklad 8.3

Je dána gramatika $G = (\{S, A, B, C\}, \{b, c\}, P, S)$, kde P obsahuje pravidla:

- $S \rightarrow AB$
- $A \rightarrow C$
- $B \rightarrow bB$
- $C \rightarrow c$
- $D \rightarrow bc$

Odstraňte z této gramatiky zbytečné symboly.

Příklad 8.4

Je dána gramatika $G = (\{E, T, F\}, \{+, *, a, (,)\}, P, E)$, kde P obsahuje pravidla:

- $E \rightarrow E + T \mid T$
- $T \rightarrow T * F \mid F$
- $F \rightarrow (E) \mid a$

Odstraňte z této gramatiky rekurzivitu zleva.

Chomského normální tvar (CNF)

- $G = (N, T, P, S)$, kde P :
 - $N \rightarrow NN \mid T$
 - případně $S \rightarrow \epsilon$, pokud se S nevyskytuje na pravé straně žádného pravidla
- Každý bezkontextový jazyk je generovaný nějakou gramatikou v Chomského normálním tvaru.
- Algoritmus převodu gramatiky do CNF – vstup vlastní BG, bez jednoduchých pravidel
- Jak transformovat BG na vlastní a bez jednoduchých pravidel?
 - 1 vyloučení zbytečných symbolů
 - 2 vyloučení ϵ -pravidel
 - 3 vyloučení jednoduchých pravidel
 - 4 * vyloučení zbytečných symbolů

Příklad 8.5

Je dána vlastní gramatika bez ϵ -pravidel $G = (\{S, A, B\}, \{a, b\}, P, S)$, kde P obsahuje pravidla:

- $S \rightarrow aB \mid bA$
- $A \rightarrow aS \mid bAA \mid a$
- $B \rightarrow bS \mid aBB \mid b$

Převeďte tuto gramatiku do normálního tvaru podle Chomského.

Příklad 8.6

Sestrojte gramatiku v Chomského normálním tvaru, která generuje jazyk: $L = \{a^m b^n : n \geq m \geq 1\}$

CYK (Cocke-Younger-Kasami) algoritmus

- vstup:
 - BG v Chomského normálním tvaru,
 - $w \in T^*$
- výstup:
 - ano, pokud $w \in L(G)$,
 - ne, jinak.
- $O(n^3)$, kde n je délka vstupního řetězce

Příklad 8.7

Nechť $G = (\{A, B, C\}, \{0, 1\}, P, A)$, kde P :

- $A \rightarrow BC \mid AB \mid 1$
- $B \rightarrow AA \mid 0$
- $C \rightarrow CB \mid 1 \mid 0$

je gramatika v Chomského normálním tvaru.

Ukažte průběh algoritmu CYK pro větu

$$w = 110100$$

Normální tvar podle Greibachové

- $G = (N, T, P, S)$, kde P :
 - $N \rightarrow a\alpha$, kde $a \in T$ a $\alpha \in N^*$
 - případně $S \rightarrow \epsilon$, pokud se S nevyskytuje na pravé straně žádného pravidla