

CYK, Zásobníkové automaty a jejich konstrukce

BI-AAG (9.cvičení)

Bc. Eliška Šestáková
sestaeli@fit.cvut.cz

20.11.2014

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 9.1

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 9.2

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 9.3

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

Zásobníkový automat

- $R = (Q, T, G, \delta, q_0, Z_0, F)$
- Jazyk definovaný (přijímaný) ZA $R = (Q, T, G, \delta, q_0, Z_0, F)$:
 - přechodem do koncového stavu
$$L(R) = \{w : (q_0, w, Z_0) \vdash^* (q, \varepsilon, \gamma), \gamma \in G^*, q \in F, w \in T^*\},$$
 - s prázdným zásobníkem
$$L_\varepsilon(R) = \{w : (q_0, w, Z_0) \vdash^* (q, \varepsilon, \varepsilon), q \in Q, w \in T^*\}.$$

Příklad 9.4

Převeďte zásobníkový automat

$R = (\{q, r\}, \{+, *, (,), a\}, \{E, T, F, +, *, (,), a, \#\}, \delta, q, \#, \{r\})$,
přijímající koncovým stavem, na zásobníkový automat přijímající
prázdným zásobníkem. Zobrazení δ je definováno takto:

$$\begin{aligned}\delta(q, b, \varepsilon) &= \{(q, b)\} \text{ pro všechna } b \in \{+, *, (,), a\}, \\ \delta(q, \varepsilon, E + T) &= \{(q, E)\}, \\ \delta(q, \varepsilon, T) &= \{(q, E)\}, \\ \delta(q, \varepsilon, T * F) &= \{(q, T)\}, \\ \delta(q, \varepsilon, F) &= \{(q, T)\}, \\ \delta(q, \varepsilon, (E)) &= \{(q, F)\}, \\ \delta(q, \varepsilon, a) &= \{(q, F)\}, \\ \delta(q, \varepsilon, \#E) &= \{(r, \#E)\},\end{aligned}$$

Příklad 9.5

Sestrojte zásobníkový automat pro jazyk $L = \{a^n b^n : n \geq 0\}$.
Bezkontextová gramatika, která tuto strukturu generuje má tvar:

$$G = (\{S\}, \{a, b\}, \{S \rightarrow aSb \mid \epsilon\}, S)$$

Příklad 9.6

Sestrojte zásobníkový automat pro analýzu závorkové struktury.
Bezkontextová gramatika, která tuto strukturu generuje má tvar:

$$G = (\{S\}, \{ (,) \}, \{ S \rightarrow (S)S \mid \epsilon \}, S)$$

Zásobníkový automat jako model syntaktického analyzátoru

- Syntaktická analýza – proces analýzy řetězců w s cílem určit:
 - zdali $w \in L(G)$
 - strukturu řetězce (posloupnost derivací, derivační strom):
 - levý rozklad věty
 - pravý rozklad věty
- Metody syntaktické analýzy
 - shora dolů (top down) – nalezení levého rozkladu věty v gramatice
 - zdola nahoru (bottom up) – nalezení pravého rozkladu věty v gramatice

Konstrukce zásobníkového automatu

1 Metodou shora dolů

$R = (\{q\}, T, N \cup T, \delta, q, S, \emptyset)$, kde δ :

- $\delta(q, \varepsilon, A) = \{(q, \alpha) : A \rightarrow \alpha \in P\}, \forall A, A \in N$ (**expanze**),
- $\delta(q, a, a) = \{(q, \varepsilon)\}, \forall a, a \in T$ (**srovnání**).

Vrchol zásobníku u tohoto typu automatu bude vždy vlevo.

2 Metodou zdola nahoru

$R = (\{q, r\}, T, N \cup T \cup \{\#\}, \delta, q, \#, \{r\})$, kde δ :

- $\delta(q, a, \varepsilon) = \{(q, a)\}, \forall a, a \in T$, (**přesun**),
- $\delta(q, \varepsilon, \alpha) = \{(q, A) : A \rightarrow \alpha \in P\}$ (**redukce**),
- $\delta(q, \varepsilon, \#S) = \{(r, \varepsilon)\}$ (**přijetí**).

Vrchol zásobníku u tohoto typu automatu bude vždy vpravo.

Příklad 9.7

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

- (1) $S \rightarrow AB$
- (2) $A \rightarrow aAb$
- (3) $A \rightarrow ab$
- (4) $B \rightarrow cBd$
- (5) $B \rightarrow cd$

- Vytvořte ZA jako model syntaktického analyzátoru pro metodu:
 - shora dolů
 - zdola nahoru
- Proved'te analýzu řetězce *aabbcd* pro obě metody.