

Cvičenie - LR(0)-analyzátory

Ing. Viliam Hromada, PhD.

C-510
Ústav informatiky a matematiky
FEI STU

`viliam.hromada@stuba.sk`



Príklad č. 1

Zostrojte $LR(0)$ analyzátor pre gramatiku generujúcu jazyk
 $a^n b^n a^m b^m, n \geq 1, m \geq 1$:

1. $S \rightarrow AB$
2. $A \rightarrow aAb$
3. $A \rightarrow ab$
4. $B \rightarrow aBb$
5. $B \rightarrow ab$



- Ako prvé skontrolujeme, či je gramatika redukovaná.
- Keďže v tejto gramatike $N_T = \{S, A, B\}$ a následne $V_D = \{S, A, B, a, b\}$, táto gramatika je redukovaná a môžeme pristúpiť ku konštrukcii $LR(0)$ -automatu.
- Najprv do gramatiky pridáme nový počiatočný neterminál S' a pravidlo $S' \rightarrow S$.



- Počiatočný stav $LR(0)$ -automatu je tvorený $CLOSURE_0$ položkou $S' \rightarrow \bullet S$.
- Teda k položke $S' \rightarrow \bullet S$ pridáme všetky položky, ktoré vznikli z pravidiel, ktoré majú na ľavej strane neterminál S , pričom symbol \bullet pridáme na začiatok pravej strany, t.j. pridáme položky:
 - $S \rightarrow \bullet AB$
- Ak nám pribudnú nové položky, ktoré majú za symbolom \bullet neterminál, v našom prípade $S \rightarrow \bullet AB$, pridáme teraz položky pre všetky pravidlá, ktoré majú na ľavej strane neterminál A , pričom symbol \bullet pridáme na začiatok pravej strany, t.j. pridáme položky:
 - $A \rightarrow \bullet aAb$
 - $A \rightarrow \bullet ab$

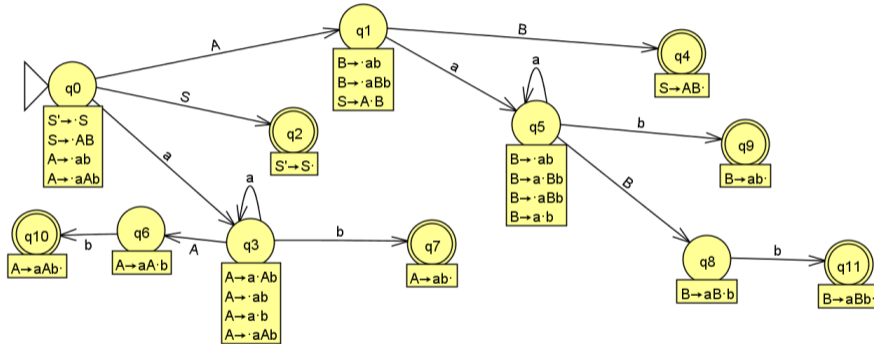
- Položky, ktoré pribudli, nemajú za symbolom \bullet neterminály, teda nové položky už nepridáme.
- Počiatočný stav $LR(0)$ -automatu, tu označený q_0 , je teda stav tvorený položkami
 - $S' \rightarrow \bullet S$
 - $S \rightarrow \bullet AB$
 - $A \rightarrow \bullet aAb$
 - $A \rightarrow \bullet ab$



- Zo stavu q_0 následne vedieme prechody na tie symboly, ktoré stoja za symbolom \bullet v jednotlivých položkách. Prechody vedieme do stavov, ktoré vzniknú ako $CLOSURE_0$ položiek, v ktorých posunieme \bullet za ten symbol, na ktorý vedieme prechod, t.j.:
 - V stave q_0 je položka $S \rightarrow \bullet AB$, v ktorej je za \bullet symbol A . Vedeime teda prechod na symbol A do stavu, označme ho napr. q_1 , ktorý je daný ako $q_1 = CLOSURE_0(\{S \rightarrow A \bullet B\})$.
 - V stave q_0 je položka $S' \rightarrow \bullet S$, v ktorej je za \bullet symbol S . Vedeime teda prechod na symbol S do stavu, označme ho napr. q_2 , ktorý je daný ako $q_2 = CLOSURE_0(\{S' \rightarrow S \bullet\})$.
 - V stave q_0 sú položky $A \rightarrow \bullet ab$, $A \rightarrow \bullet aAb$, v ktorých je za \bullet symbol a . Vedeime teda prechod na symbol a do stavu, označme ho napr. q_3 , ktorý je daný ako $q_3 = CLOSURE_0(\{A \rightarrow a \bullet b, A \rightarrow a \bullet Ab\})$.

- Tento postup opakujeme pre novo-vznikajúce stavy, kým nedostávame nové stavy.
- T.j. pre každý stav vyšetříme najprv, aké všetky položky doň patria pomocou *CLOSURE* a následne riešime prechody na jednotlivé symboly, ktoré stoja za • v položkách daného stavu.
- V tomto prípade dostávame výsledný *LR(0)*-automat uvedený na ďalšom slajde.

Príklad - LR(0) automat



Príklad - ACTION

Tabuľku *ACTION LR(0)*-analyzátoru pre jednotlivé stavy q_i zostrojíme nasledovným spôsobom podľa toho, aké položky sa v stave nachádzajú:

- Ak sa v stave q_i nachádza položka, v ktorej za \bullet stojí **terminál**, je v danom stave signalizovaný **presun**.
- Ak sa v stave q_i nachádza položka, v ktorej \bullet stojí **na konci nejakého pravidla**, je v danom stave signalizovaná **redukcia podľa tohto pravidla**.
- Ak sa v stave q_i nachádza položka $S' \rightarrow S\bullet$, je v danom stave signalizovaná **akceptácia vstupného slova**. Akceptáciu je však možné robiť **len v prípade, že bol vstup celý prečítaný!**



Príklad - ACTION

<i>ACTION</i>	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	q_9	q_{10}	q_{11}
	<i>P</i>	<i>P</i>	<i>A</i>	<i>P</i>	<i>R1</i>	<i>P</i>	<i>P</i>	<i>R3</i>	<i>P</i>	<i>R5</i>	<i>R2</i>	<i>R4</i>

Čísla pri redukcii sú podľa číslovania pravidiel na slajde č. 2.



Príklad - GOTO

<i>GOTO</i>	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	q_9	q_{10}	q_{11}
<i>a</i>	q_3	q_5		q_3		q_5						
<i>b</i>				q_7		q_9	q_{10}		q_{11}			
<i>S'</i>												
<i>S</i>	q_2											
<i>A</i>	q_1			q_6								
<i>B</i>		q_4				q_8						

Príklad - činnosť analyzátora

r.	Zásobník \square	Zvyšok vstupu:	Akcia
1	q_0	<i>aabbab</i>	P
2	$q_0 q_3$	<i>abbab</i>	P
2	$q_0 q_3 q_3$	<i>bbab</i>	P
3	$q_0 q_3 q_3 q_7$	<i>bab</i>	R3
4	$q_0 q_3 q_6$	<i>bab</i>	P
5	$q_0 q_3 q_6 q_{10}$	<i>ab</i>	R2
6	$q_0 q_1$	<i>ab</i>	P
7	$q_0 q_1 q_5$	<i>b</i>	P
8	$q_0 q_1 q_5 q_9$	ε	R5
9	$q_0 q_1 q_4$	ε	R1
10	$q_0 q_2$	ε	A



Činnosť analyzátor - reťazec ab

Zásobník \square	Zvyšok vstupu:	Akcia
q_0	ab	P
q_0q_3	b	P
$q_0q_3q_7$	ϵ	P
q_0q_1	ϵ	SYNTAX ERROR

Pri spracovaní reťazca ab nastala situácia, že vstup je celý prečítaný a na vrchu zásobníka je stavový symbol q_1 , ktorý signalizuje **Presun**. Keďže vstup je celý prečítaný, akciu Presun **nie je možné vykonať**, teda dochádza k **syntaktickej chybe**, čo znamená, že ab nemá v gramatike deriváciu.



Činnosť analyzátora - reťazec b

Zásobník \square	Zvyšok vstupu:	Akcia
q_0	b	SYNTAX ERROR

Pri spracovaní reťazca b nastala situácia, že na vstupe je symbol b , na vrchu zásobníka je stavový symbol q_0 a je signalizovaný **presun**. Teoreticky by sa teda mal presunúť do zásobníka **stavový symbol**, ktorý je v tabuľke *GOTO* na pozícii $GOTO[q_0, b]$. Avšak v tabuľke sa na pozícii $GOTO[q_0, b]$ **žiaden stavový symbol nenachádza** a dochádza teda k **syntaktickej chybe** a reťazec b nemá v gramatike deriváciu.



Činnosť analyzátoru - reťazec *ababa*

Zásobník \square	Zvyšok vstupu:	Akcia
q_0	<i>ababa</i>	P
$q_0 q_3$	<i>baba</i>	P
$q_0 q_3 q_7$	<i>aba</i>	R3
$q_0 q_1$	<i>aba</i>	P
$q_0 q_1 q_5$	<i>ba</i>	P
$q_0 q_1 q_5 q_9$	<i>a</i>	R5
$q_0 q_1 q_4$	<i>a</i>	R1
$q_0 q_2$	<i>a</i>	SYNTAX ERROR

V stave q_2 je signalizovaná akceptácia. Avšak, keďže zostala nespracovaná časť vstupu *a*, k akceptácii nemôže prísť a nastáva **syntaktická chyba**. Reťazec *ababa* teda nemá v gramatike deriváciu.



Príklad č. 2

Nech je daná gramatika $G = (N, T, P, S)$, kde S je počiatkový neterminál,
 $N = \{S, A, B\}$, $T = \{a, b\}$, pravidlá:

1. $S \rightarrow aBS$
2. $S \rightarrow bAS$
3. $S \rightarrow \varepsilon$
4. $A \rightarrow a$
5. $A \rightarrow bAA$
6. $B \rightarrow b$
7. $B \rightarrow aBB$

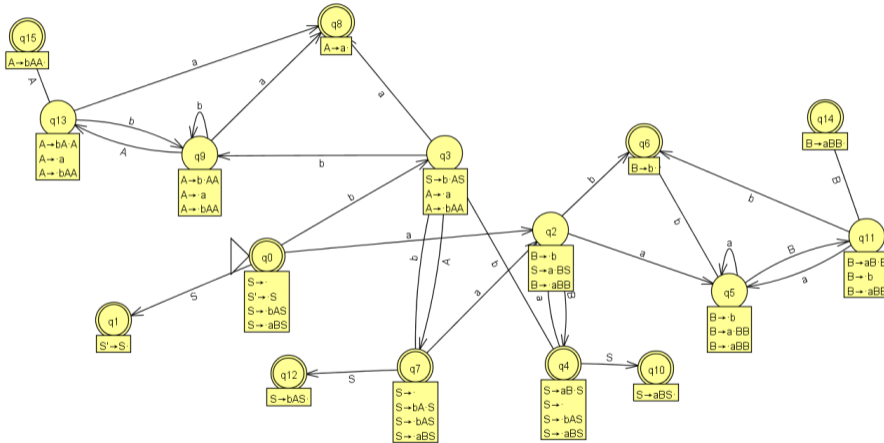
Ide o gramatiku, ktorej jazyk $L = \{w \in \{a, b\}^* \mid \#_a(w) = \#_b(w)\}$, t.j. ide o reťazce z písmen a, b ktoré majú rovnaký počet a -čok a b -čok.
Zistite, či ide o $LR(0)$ -gramatiku.



- Aby sme zistili, či ide o $LR(0)$ -gramatiku, zostrojíme tabuľku *ACTION* $LR(0)$ -analyzátoru.
- Ak tabuľka neobsahuje konflikty, gramatika **je** $LR(0)$ -gramatikou.
- Ak tabuľka obsahuje aspoň 1 konflikt, gramatika **nie je** $LR(0)$ -gramatikou.



LR(0)-automat



ACTION - konflikty

<i>ACTION</i>	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
	<i>P/R3</i>	<i>A</i>	<i>P</i>	<i>P</i>	<i>P/R3</i>	<i>P</i>	<i>R6</i>	<i>P/R3</i>
<i>ACTION</i>	q_8	q_9	q_{10}	q_{11}	q_{12}	q_{13}	q_{14}	q_{15}
	<i>R4</i>	<i>P</i>	<i>R1</i>	<i>P</i>	<i>R2</i>	<i>P</i>	<i>R7</i>	<i>R5</i>

V akciách pre stavy q_0 , q_4 , q_7 sú konflikty, preto aj príslušná gramatika **nie je LR(0)** gramatika.



Príklad č. 3

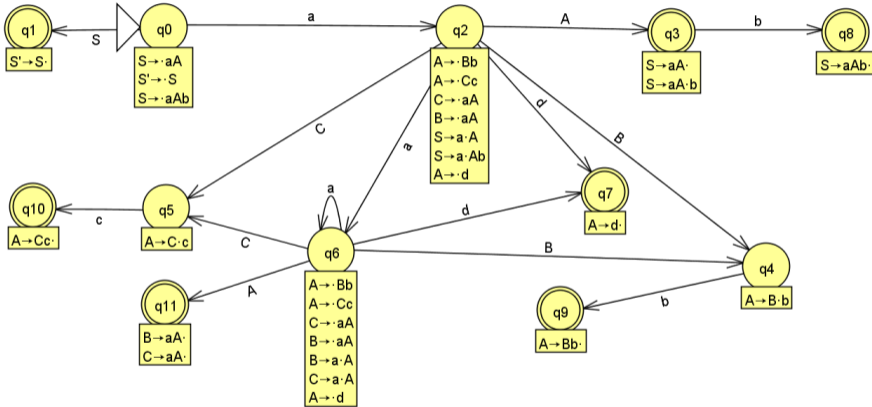
Nech gramatika $G = (\{S, A, B, C\}, \{a, b, c, d\}, P, S)$, kde pravidlá:

1. $S \rightarrow aAb$
2. $S \rightarrow aA$
3. $A \rightarrow Bb$
4. $A \rightarrow Cc$
5. $A \rightarrow d$
6. $B \rightarrow aA$
7. $C \rightarrow aA$

Zistite, či ide o $LR(0)$ gramatiku.



LR(0)-automat



Tabuľka ACTION

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	q_9	q_{10}	q_{11}
	P	A	P	P / R2	P	P	P	R5	R1	R3	R4	R6 / R7

V stave q_3 je konflikt **presun - redukcia** a v stave q_{11} je konflikt **redukcia - redukcia**. Nemôže teda ísť o $LR(0)$ gramatiku.



Tabuľka *GOTO*

<i>GOTO</i>	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	q_9	q_{10}
S'											
S	q_1										
A			q_3				q_{11}				
B			q_4				q_4				
C			q_5				q_5				
a	q_2		q_6				q_6				
b				q_8	q_9						
c						q_{10}					
d			q_7				q_7				