

Cvičenie 8 - Syntaktická analýza bottom-up, LR(0), SLR(1)

Ing. Viliam Hromada, PhD.

C-510
Ústav informatiky a matematiky
FEI STU

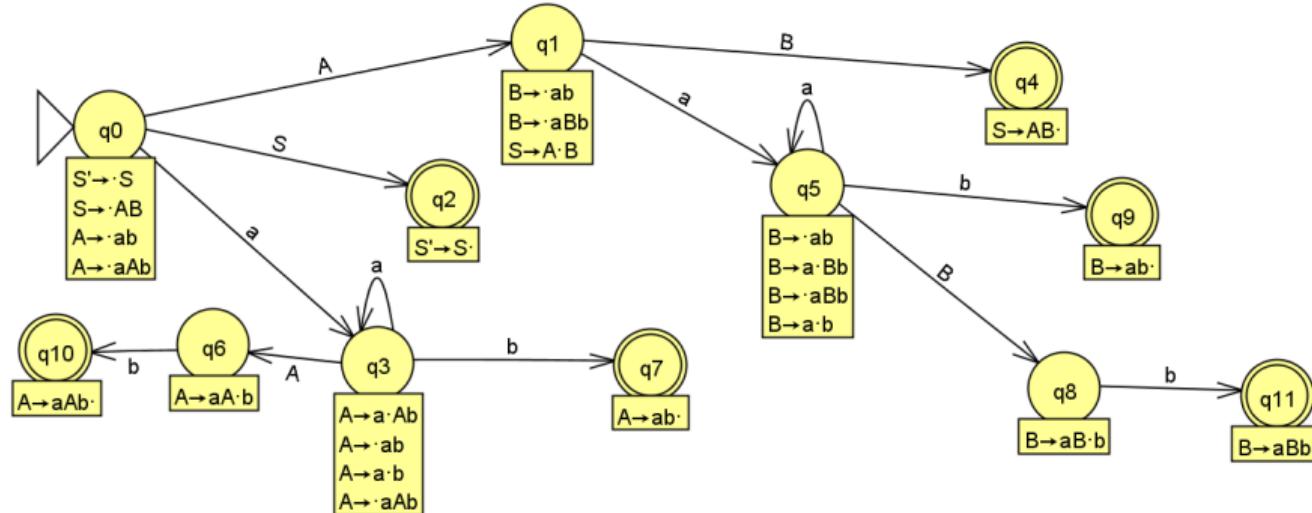
viliam.hromada@stuba.sk

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. $B \rightarrow ab$

Príklad - $LR(0)$ automat



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}
	P	P	A	P	$R1$	P	P	$R3$	P	$R5$	$R2$	$R4$

Čísla pri redukcii sú podľa číslования pravidiel na slajde č. 44

Príklad - GOTO

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

Príklad - činnosť analyzátoru

r.	Zásobník ⊑	Zvyšok vstupu:	Akcia
1	q_0	$aabbab$	P
2	$q_0 q_3$	$abbab$	P
2	$q_0 q_3 q_3$	$bbab$	P
3	$q_0 q_3 q_3 q_7$	bab	R3
4	$q_0 q_3 q_6$	bab	P
5	$q_0 q_3 q_6 q_{10}$	ab	R2
6	$q_0 q_1$	ab	P
7	$q_0 q_1 q_5$	b	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

Príklad č. 2

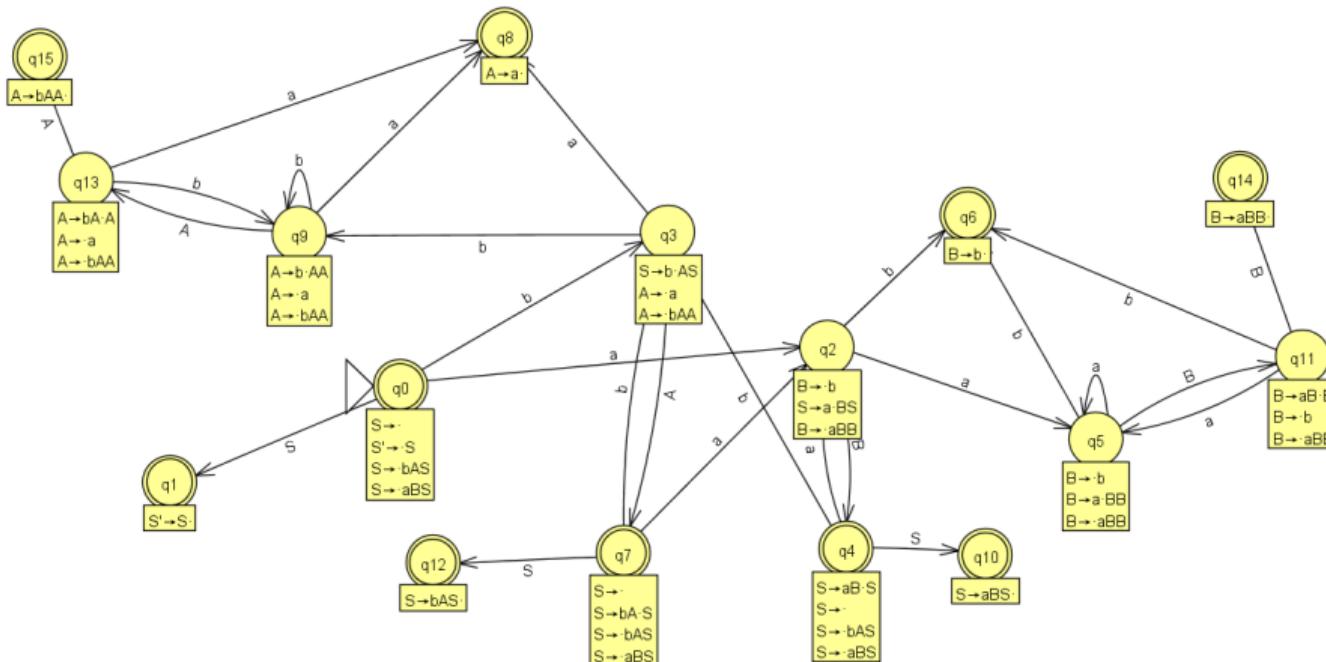
Nech je daná gramatika $G = (N, T, P, S)$, kde S je počiatočný neterminál,
 $N = \{S, A, B\}$, $T = \{a, b\}$, pravidlá:

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.

$LR(0)$ -automat



ACTION - konflikty

<i>ACTION</i>	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
	$P/R3$	A	P	P	$P/R3$	P	$R6$	$P/R3$

<i>ACTION</i>	q_8	q_9	q_{10}	q_{11}	q_{12}	q_{13}	q_{14}	q_{15}
	$R4$	P	$R1$	P	$R2$	P	$R7$	$R5$

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

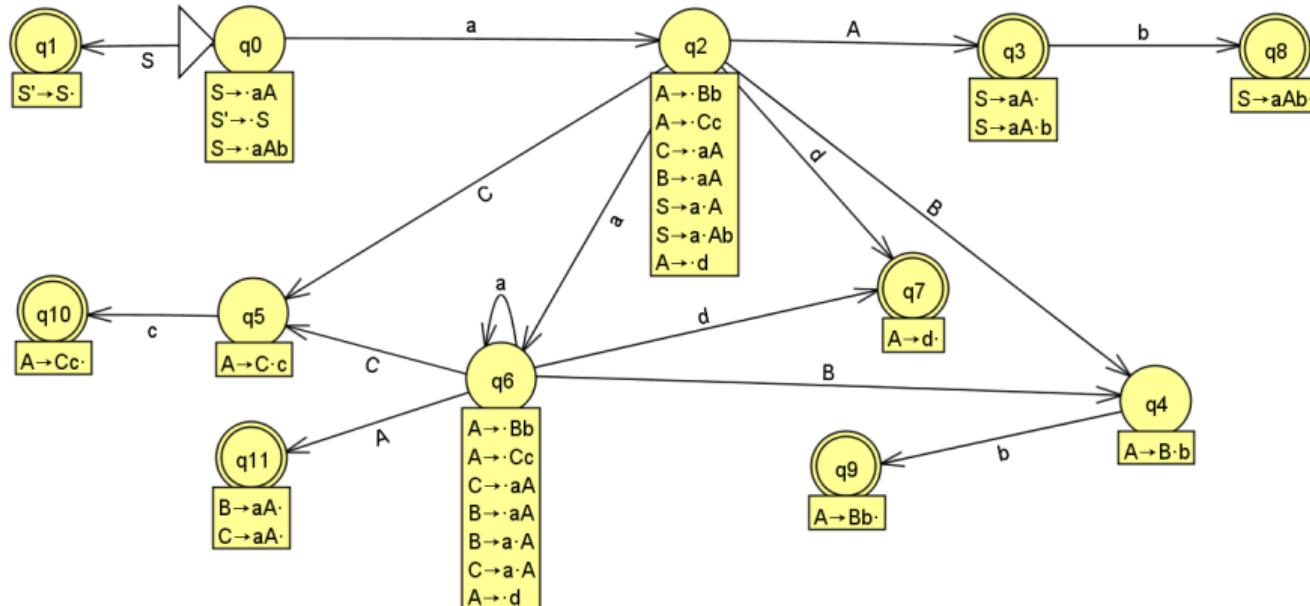
Príklad č. 3

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 íšť o $LR(0)$ gramatiku.

Tabuľka GOTO

$GOTO$	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				

Príklad č. 4

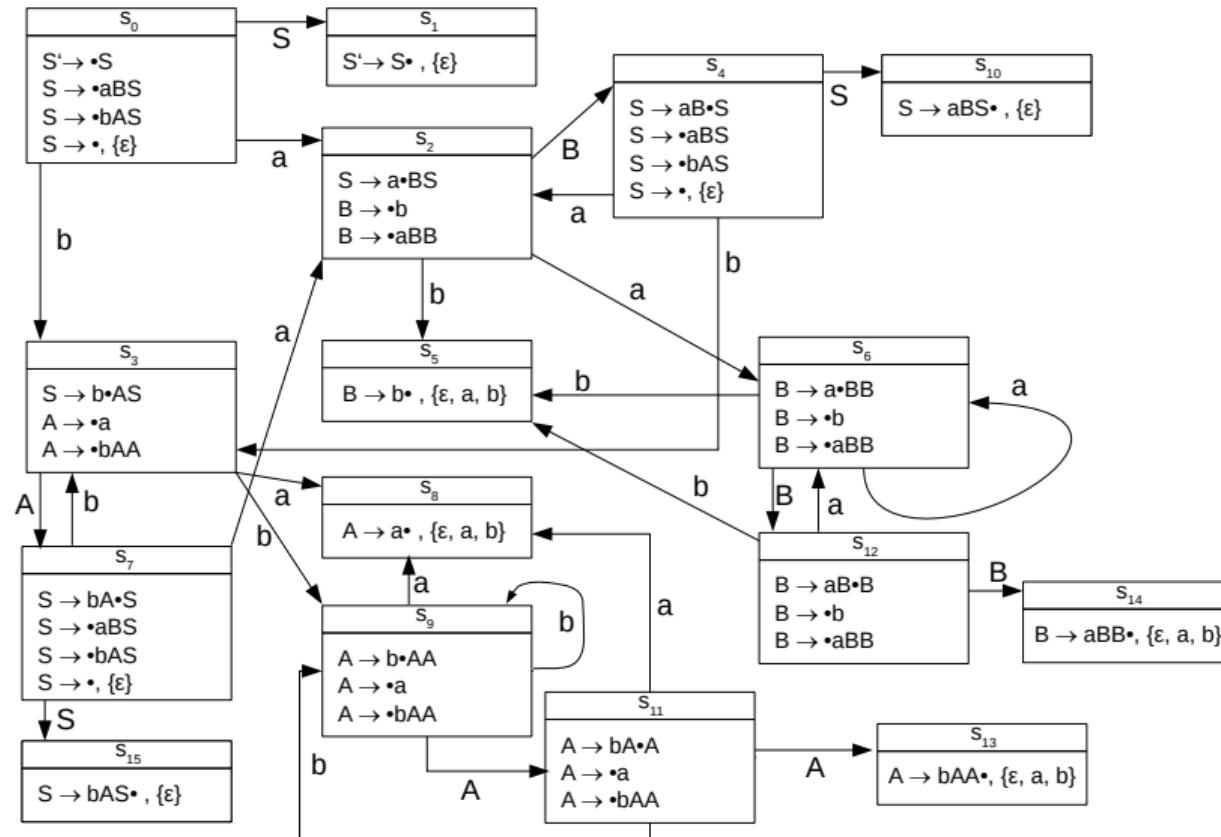
Nasledovná gramatika sa nedala spracovať $LR(0)$ -analyzátorm (vid'. príklad č. 2)- pretože obsahuje ε -pravidlo a to v $LR(0)$ -analyzátoroch predstavuje vždy problém. V $SLR(1)$ -analýze to už problém byť nemusí...

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. $B \rightarrow aBB$

Príklad č. 4 - FOLLOW

Najprv zostrojme množinu *FOLLOW* pre neterminály S, A, B :

<i>FIRST</i>	S	A	B
	ε, a, b	a, b	a, b
<i>FOLLOW</i>	ε	ε, a, b	ε, a, b



ACTION&GOTO

ACTION	s0	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s15
a	P		P	P	P	R6	P	P	R4	P		P	P	R5	R7	
b	P		P	P	P	R6	P	P	R4	P		P	P	R5	R7	
ϵ	R3	A			R3	R6		R3	R4		R1			R5	R7	R2

GOTO	s0	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s15
S		s1														
A					s7						s11		s13			
B				s4												s14
C																
a	s2			s6	s8	s2		s6	s2		s8		s8	s6		
b	s3			s5	s9	s3		s5	s3		s9		s9	s5		

Poznámka k príkladu

- Pre uvedenú gramatiku boli v $LR(0)$ -analýzátore (viď príklad č. 2) konflikty, pretože $LR(0)$ -analýzátor sa nerozhoduje na základe vstupných symbolov.
- Redukcia podľa pravidla $S \rightarrow \varepsilon$ sa totižto mohla robiť vždy, keď bola signalizovaná v stavoch $LR(0)$ -automatu, čo bolo v konflikte s presunmi symbolov a, b .
- V $SLR(1)$ -analýzátore sa táto redukcia bude vykonávať už **len vtedy**, keď je signalizovaná v stave $LR(0)$ -automatu **a zároveň, keď je vstup prázdny** - pretože $FOLLOW(S) = \{\varepsilon\}$
- Tým sa odstráni konflikt s presunmi symbolov a, b a gramatika sa dá deterministicky syntakticky spracovať napr. $SLR(1)$ -analýzátorom.

Príklad č. 5

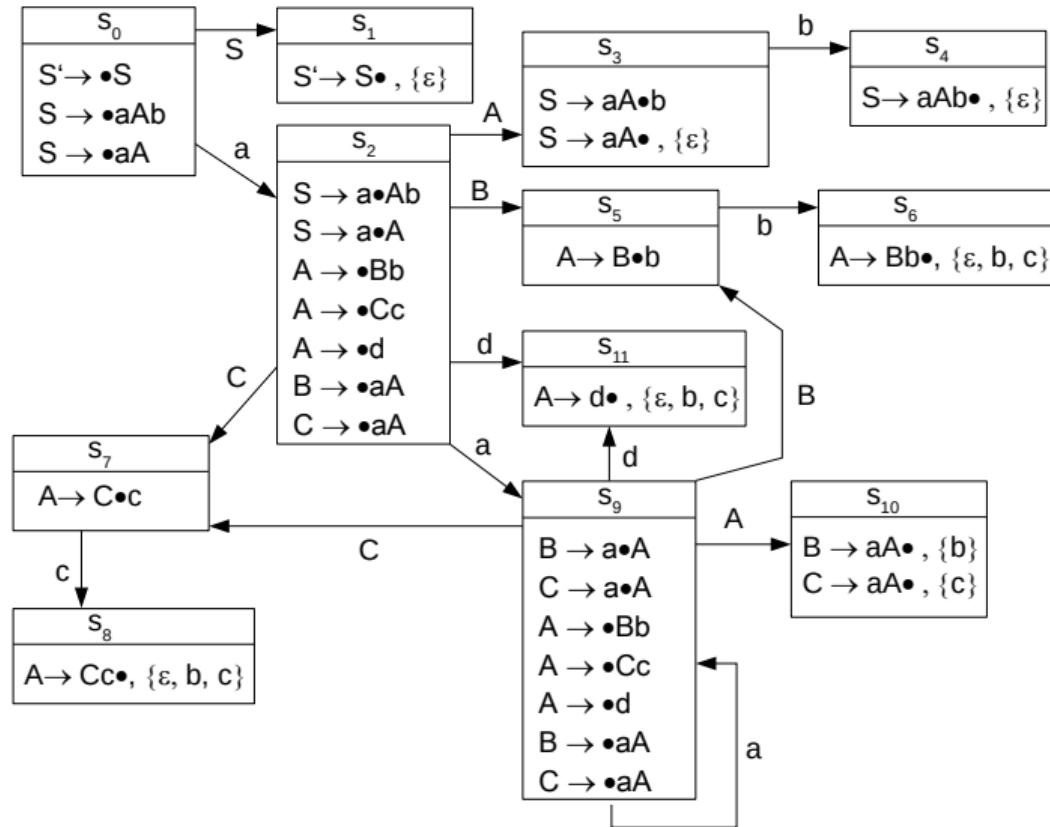
Nasledovná gramatika v $LR(0)$ -analyzátore obsahovala P-R, R-R konflikty (vid' príklad č. 3). Zistime, či sa dá spracovať pomocou $SLR(1)$ -analyzátora, t.j. či sa konflikty odstránia, keď začneme uvažovať pri rozhodovaní aj vstupné symboly. Gramatika $G = (\{S, A, B, C\}, \{a, b, c, d\}, P, S)$, kde 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. $C \rightarrow aA$

Príklad č. 5 - FOLLOW

Najprv zostrojme množinu *FOLLOW* pre neterminály S, A, B, C :

<i>FIRST</i>	S	A	B	C
	a	a, d	a	a
<i>FOLLOW</i>	ϵ	ϵ, b, c	b	c



ACTION&GOTO

ACTION	s0	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11
a	P		P						P			
b				P		P	R3		R4		R6	R5
c							R3	P	R4		R7	R5
d			P						P			
ϵ		A		R2	R1		R3		R4			R5

GOTO	s0	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11
S	s1											
A			s3						s10			
B			s5						s5			
C			s7						s7			
a	s2		s9						s9			
b				s4	s6							
c							s8					
d			s11						s11			

Poznámka k príkladu

- V $LR(0)$ -analyzátore pre danú gramatiku boli konflikty (vid' príklad č. 3) - pre stav $LR(0)$ -automatu tu označený ako s_3 tam bol konflikt presun-redukcia, pre stav tu označený ako s_{10} konflikt redukcia-redukcia.
- V stave s_3 už konflikt P-R nie je, pretože redukcia sa vykoná už len v prípade, že je vstup prázdny - pretože v danej gramatike sa nemôže v pravej vetnej forme vyskytnúť Sb , lebo $b \notin FOLLOW(S)$.
- V stave s_{10} už konflikt R-R nie je, pretože množiny $FOLLOW(B)$ a $FOLLOW(C)$ sú disjunktné, a teda sa vieme rozhodnúť, či chceme redukciou získať neterminál B alebo C - podľa toho, aký terminál je na vstupe, a teda aká časť pravej vetnej formy vznikne v zásobníku.

Analýza reťazca $aadb$

Vykonajte syntaktickú analýzu reťazca $aadb$

r.	Zásobník \sqsubset	Zvyšok vstupu	Akcia
1	s_0	$aadb$	P
2	$s_0 s_2$	adb	P
3	$s_0 s_2 s_9$	bb	P
4	$s_0 s_2 s_9 s_{11}$	bb	R5
5	$s_0 s_2 s_9 s_{10}$	bb	R6
6	$s_0 s_2 s_5$	bb	P
7	$s_0 s_2 s_5 s_6$	b	R3
8	$s_0 s_2 s_3$	b	P
9	$s_0 s_2 s_3 s_4$	ε	R1
10	$s_0 s_1$	ε	A

Slovo **patrí** do jazyka. Jeho pravú deriváciu dostaneme postupnou aplikáciou pravidiel č. 1, 3, 6, 5.

$$S \Rightarrow aAb \Rightarrow aBbb \Rightarrow aaAbb \Rightarrow aadb.$$

Analýza reťazca *baadbb*

Vykonajte syntaktickú analýzu reťazca *baadbb*

r.	Zásobník	Zvyšok vstupu	Akcia
1	s_0	<i>baadbb</i>	ERROR

- Zásek analyzátoru už v prvom kroku, pretože v *ACTION* nie je definovaná akcia pre symbol s_0 na vrchu zásobníka a *b* na vstupe.
- To je len logické, lebo pohľad na pravidlá gramatiky naznačuje, že všetky v nej generované slová musia začínať znakom *a*.
- Slovo *baadbb* teda nepatrí do jazyka generovaného danou gramatikou a nemá v nej odvodenie.

Analýza reťazca *adbb*

Vykonaljte syntaktickú analýzu reťazca *adbb*

r.	Zásobník \sqsubset	Zvyšok vstupu	Akcia
1	s_0	<i>adbb</i>	P
2	s_0s_2	<i>dbb</i>	P
3	$s_0s_2s_{11}$	<i>bb</i>	R5
4	$s_0s_2s_3$	<i>bb</i>	P
5	$s_0s_2s_3s_4$	<i>b</i>	ERROR

- Zásek analyzátora v piatom kroku, pretože v *ACTION* nie je definovaná akcia pre symbol s_4 na vrchu zásobníka a b na vstupe.
- Je tam síce definovaná redukcia podľa pravidla č. 1, ale len v prípade, že **na vstupe nič nie je** (t.j. na vstupe je už len ε).
- Slovo *adbb* teda nepatrí do jazyka generovaného danou gramatikou a nemá v nej odvodenie.

Analýza reťazca *aad*

Vykonaljte syntaktickú analýzu reťazca *aad*

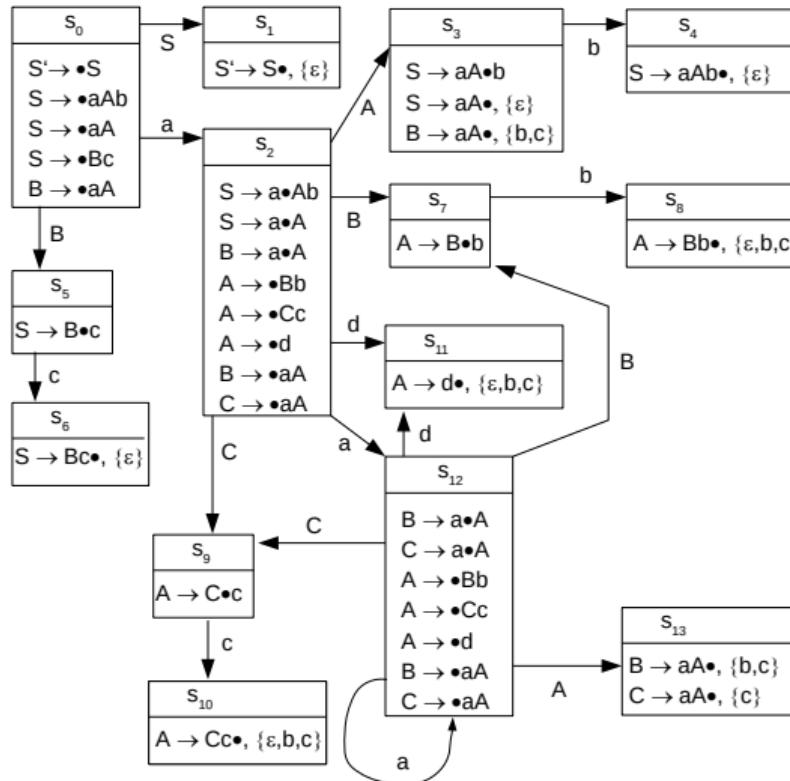
r.	Zásobník \sqsubset	Zvyšok vstupu	Akcia
1	s_0	<i>aad</i>	P
2	s_0s_2	<i>ad</i>	P
3	$s_0s_2s_9$	<i>d</i>	P
4	$s_0s_2s_9s_{11}$	ϵ	R5
5	$s_0s_2s_9s_{10}$	ϵ	ERROR

- Zásek analyzátora v piatom kroku, pretože v *ACTION* nie je definovaná akcia pre symbol s_{10} na vrchu zásobníka a prázdný vstup.
- Pre s_{10} sú definované 2 redukcie, avšak na vstupe by malo byť alebo *b*, alebo *c*.
- Slovo *aad* teda nepatrí do jazyka generovaného danou gramatikou a nemá v nej odvodenie.

Príklad č. 6

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

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



ACTION&GOTO

ACTION	s0	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13
a	P		P									P		
b				P/R7				P	R4		R5	R6		R7
c				R7		P			R4	P	R5	R6		R7/R8
d			P									P		
ϵ		A		R2	R1		R3		R4		R5	R6		

GOTO	s0	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13
S		s1												
A				s3									s13	
B		s5		s7									s7	
C				s9									s9	
a		s2		s12									s12	
b					s4				s8					
c						s6				s10				
d													s11	

- Vidíme, že tabuľka *ACTION* obsahuje konflikt Presun-redukcia pre stav s_3 a vstup b a konflikt redukcia-redukcia pre stav s_{13} a vstup c .
- Príslušná gramatika teda **nie je $SLR(1)$ gramatika.**