

LL(k) gramatiky

1. Pro danou gramatiku zjistěte, zda je LL(1), a sestavte rozkladovou tabulku.

S \rightarrow aAB (1)
S \rightarrow bA (2)
S \rightarrow e (3)
A \rightarrow aAb (4)
A \rightarrow e (5)
B \rightarrow bB (6)
B \rightarrow e (7)

Řešení

follow(S) = {e}
follow(A) = {b, e}
follow(B) = {e}

Pro každá dvě pravidla $x \rightarrow \alpha_1$ a $x \rightarrow \alpha_2$ platí, že $\text{first}(\alpha_1 \text{ follow}(x))$ a $\text{first}(\alpha_2 \text{ follow}(x))$ jsou disjunktní množiny. Proto je gramatika LL(1) a můžeme sestavit rozkladovou tabulku:

M	a	b	e
S	1	2	3
A	4	5	5
B		6	7

2. Sestavte rozkladovou tabulku pro G:

S \rightarrow ABa (1)
A \rightarrow e (2)
A \rightarrow aA (3)
B \rightarrow bSAb (4)
B \rightarrow c (5)

Řešení

M	a	b	c	e
S	1	1	1	
A	3	2	2	
B		4	5	

3. Převedte na LL(1) a sestavte rozkladovou tabulku

S \rightarrow aS (1)
S \rightarrow a (2)

Řešení

Pravidla (1) a (2) začínají stejným terminálním symbolem; gramatika tedy není LL(1). Tuto tzv. *first-first kolizi* ale můžeme snadno odstranit levou faktorizací:

S \rightarrow aS' (1)
S' \rightarrow S (2)
S' \rightarrow e (3)

Protože $\text{follow}(S') = \{e\}$, bude rozkladová tabulka:

M	a	e
S	1	
S'	2	3

4. Převedte G(N, T, P, A) na LL(1):

A \rightarrow Bc | Dd
B \rightarrow bx | y
D \rightarrow Bz

Řešení

Dosadíme za D:

A \rightarrow Bc | Bzd
B \rightarrow bx | y
(symbol D je nedostupný)

Dosadíme za B:

A \rightarrow bxc | yc | bxzd | yzd
(symbol B je nedostupný)

Odstraníme first-first kolizi:

A \rightarrow bxA' | yA'
A' \rightarrow c | zd
A'' \rightarrow c | zd

Gramatika by šla samozřejmě dále zjednodušit. Jelikož je ale jazyk L(G) konečný, nemá gramatika příliš velké opodstatnění a nebudeme se jí dále zabývat.

5. Převedte G(N, T, P, A) na LL(1):

A \rightarrow aB | CB
B \rightarrow cB | d
C \rightarrow aC | bB

Řešení

Budeme odstraňovat first-first kolizi mezi ab a cb; nejprve dosadíme za C:

A \rightarrow aB | aCB | bBB
B \rightarrow cB | d
C \rightarrow aC | bB

U symbolu A odstraníme first-first kolizi:

A \rightarrow aA' | bBB
A' \rightarrow B | CB
B \rightarrow cB | d
C \rightarrow aC | bB

6. Převedte G(N, T, P, A) na LL(1):

A \rightarrow BaC
B \rightarrow e | abC
C \rightarrow c | cBC

Řešení

Hned u druhého pravidla si můžeme všimnout *first-follow kolize*: nevíme, zda symbol B přepsat na e nebo na abc, protože pak obě větné formy začínají na a. Pro odstranění použijeme pohlcení follow symbolu:

A \rightarrow [Ba]C
[Ba] \rightarrow a | abCa
B \rightarrow e | abC
C \rightarrow c | cBC

Vytvořili jsme nový neterminální symbol [Ba], který vznikl sloučením nepohodlného follow symbolu a pravidel pro B. Tím ovšem vznikla first-first kolize, kterou je třeba odstranit. Krom toho odstraníme first-first kolizi pro symbol C.

A \rightarrow [Ba]C
[Ba] \rightarrow a[Ba]'
[Ba]' \rightarrow e | bCa
B \rightarrow e | abC
C \rightarrow cC'
C' \rightarrow e | BC

Snadno se přesvědčíme, že nová gramatika ještě není LL(1). Další postup by byl analogický.

7. Pro danou gramatiku zjistěte, zda je LL(2).

S \rightarrow aAaa (1)
S \rightarrow bAba (2)
A \rightarrow ba (3)
A \rightarrow e (4)

Gramatika není LL(1), protože $\text{first}(bA) = \{b\}$ a $\text{follow}(A) = \{a, b\}$, dochází tedy k first-follow kolizi. K rozhodnutí, zda je gramatika LL(2), budeme potřebovat:

$\text{follow}_2(A) = \{aa, ba\}$

Gramatika je LL(k), jsou-li pro větné formy

S $\rightarrow^* wAx \rightarrow^* wax$

a

S $\rightarrow^* wAx \rightarrow^* wbx$

(w, x, a, b jsou řetězce symbolů, a \neq b) množiny $\text{first}_k(ax)$ a $\text{first}_k(bx)$ disjunktní. V naší gramatice:

pro S \rightarrow aAaa
 $\text{first}_2(bAaa) = \{bb, ba\}$
 $\text{first}_2(aa) = \{aa\}$
pro S \rightarrow bAba
 $\text{first}_2(bAba) = \{bb\}$
 $\text{first}_2(ba) = \{ba\}$

Gramatika je tedy LL(2). Rozkladovou tabulku ovšem vytvořit nemůžeme, protože gramatika není silná LL(2). Co to znamená? Řekněme, že rozpoznáváme slovo "bbba":

(bbba, S, e) :- podle symbolu b použijeme pravidlo 2

(bba, Aba, 2) :- podle symbolu b se nedokážeme rozhodnout pro pravidlo 3 nebo 4;

musíme k rozhodnutí použít ještě jeden symbol (bb)
 (ba, Aba, 23) :- symbol b opět nestačí, musíme k rozhodnutí použít
 ještě jeden symbol (ba)
 (ba, ba, 234) :- teď už je to jasné

Zkusíme ještě "abaa":

(abaa, S, e) :- podle prvního symbolu (a) použijeme pravidlo 1
 (baa, Aaa, 1) :- teď se dokážeme rozhodnout podle symbolů ba
 (aa, Aaa, 13) :- symbol a pro rozhodnutí stačí
 (aa, aa, 134) :- teď už je to jasné

V předchozích příkladech jsme viděli, že jednou jsme se podle symbolu "b" rozhodnout dokázali, jindy jsme potřebovali dva symboly. Navíc, byly-li na vstupu symboly "ba", použili jsme jednou pravidlo 4, podruhé 3. K rozhodnutí docházelo na základě znalosti předchozích derivací.

Silné LL(k) gramatiky znalost předchozích derivací nevyžadují. Jsou-li tedy $x \rightarrow \alpha$ a $x \rightarrow \beta$ dvě různá pravidla, musí pro ně platit, že

$\text{first}_k(\alpha \text{ follow}_k(X))$

a

$\text{first}_k(\beta \text{ follow}_k(X))$

jsou disjunktí množiny.

V naší gramatice platí, že $\text{first}_2(\text{ba follow}_2(A)) = \{\text{ba}, \text{bb}\}$ a $\text{first}_2(\text{e follow}_2(A)) = \{\text{aa}, \text{ba}\}$. Proto není G silná LL(2) gramatika.

Pro úplnost dodejme, že každá LL(1) gramatika je zároveň silná LL(1) gramatika.

8. Sestavte rozkladovou tabulku pro G:

S \rightarrow e (1)
 S \rightarrow abA (2)
 A \rightarrow Saa (3)
 A \rightarrow b (4)

Řešení

Platí $\text{follow}(S) = \{a, e\}$, gramatika tedy není LL(1) - viz pravidlo 3. Zkusíme, zda je gramatika LL(2):

$\text{first}_2(\text{e follow}_2(S)) = \{\text{aa}, \text{e}\}$
 $\text{first}_2(\text{abA follow}_2(S)) = \{\text{ab}\}$
 $\text{first}_2(\text{Saa follow}_2(A)) = \{\text{ab}, \text{aa}\}$
 $\text{first}_2(\text{b follow}_2(A)) = \{\text{ba}, \text{b}\}$

Gramatika je dokonce silná LL(2) gramatika, můžeme sestavit rozkladovou tabulku:

M	aa	ab	a	ba	bb	b	e
S	1	2					1
A	3	3		4		4	

Poslední změna: 04.12.2012