

Fordítóprogramok

8. előadás

Aszalós László

2015. november 4.

SLR nyelvtan

Készítse el az alábbi kiterjesztett nyelvtan LR(0) kanonikus halmazait!

- $S' \rightarrow E,$
- $E \rightarrow T \mid E + T,$
- $T \rightarrow i \mid (E)$

Megoldás

- $\mathcal{I}_0 = V(\lambda) = \{[S' \rightarrow \bullet E], [E \rightarrow \bullet T], [E \rightarrow \bullet E + T], [T \rightarrow \bullet i], [T \rightarrow \bullet (E)]\}$
- $\mathcal{I}_1 = V(E) = \{[S' \rightarrow E \bullet], [E \rightarrow E \bullet + T]\}$
- $\mathcal{I}_2 = V(T) = \{[E \rightarrow T \bullet]\}$
- $\mathcal{I}_3 = V(i) = \{[T \rightarrow i \bullet]\}$
- $\mathcal{I}_4 = V(()) = \{[T \rightarrow (\bullet E)], [E \rightarrow \bullet T], [E \rightarrow \bullet E + T], [T \rightarrow \bullet i], [T \rightarrow \bullet (E)]\}$
- $\mathcal{I}_5 = V(E+) = \{[E \rightarrow E + \bullet T], [T \rightarrow \bullet i], [T \rightarrow \bullet (E)]\}$
- $\mathcal{I}_6 = V((E) = \{[T \rightarrow (E \bullet)], [E \rightarrow E \bullet + T]\}$
- $\mathcal{I}_7 = V(E + T) = \{[E \rightarrow E + T \bullet]\}$
- $\mathcal{I}_8 = V((E)) = \{[E \rightarrow (E) \bullet]\}$

Ebben az esetben a \mathcal{I}_1 LR(0) állapotban nem egyértelmű, hogy redukálni kell, vagy léptetni. Ezért felhasználjuk a *Follow* halmazokból kinyerhető információkat is.

SLR táblázat

Konstrukció

Az *action* és *goto* táblázatot a következőképpen töltjük fel:

- ha $[A \rightarrow \alpha \bullet a\beta] \in \mathcal{I}_i$ és $read(\mathcal{I}_i, a) = \mathcal{I}_j$, akkor $action[i, a] = s_j$
- ha $[A \rightarrow \alpha \bullet] \in \mathcal{I}_i$, $A \neq S'$ és $a \in Follow_1(A)$, akkor $action[i, a] = r_l$, ahol $A \rightarrow \alpha$ a nyelvtan l -dik szabálya
- ha $[S' \rightarrow S \bullet] \in \mathcal{I}_i$, akkor $action[i, \#] = acc$
- ha $read[\mathcal{I}_i, A] = \mathcal{I}_j$, akkor $goto[i, A] = j$

SLR nyelvtan

Ha az *action* és *goto* táblázat minden elemére a fenti eljárással maximum egy értéket adunk meg, akkor a táblázatok kitöltését konfliktusmentesnek, a megfelelő nyelvtant *SLR nyelvtannak* nevezzük.

Példa

Készítse el az előbbi feladatban szereplő kiegészített nyelvtan SLR elemzőjét!

Megoldás

	i	$+$	$($	$)$	$\#$	E	T
0	s_3		s_4			1	2
1		s_5			acc		
2		r_1		r_1	r_1		
3		r_3		r_3	r_3		
4	s_3		s_4			6	2
5	s_3		s_4				7
6		s_5		s_8			
7		r_2		r_2	r_2		
8		r_4		r_4	r_4		

Példa

Az előbbi táblázat segítségével elemezze az $i + (i + i)$ inputot!

Megoldás

Az elemzés menete igen hasonlít az LR(0) elemzésre: $(\#0, i + (i + i)\#) \xrightarrow{s_3}$
 $(\#0i3, +(i + i)\#) \xrightarrow{r_3} (\#0T2, +(i + i)\#) \xrightarrow{r_1} (\#0E1, +(i + i)\#) \xrightarrow{s_5}$
 $(\#0E1 + 5, (i + i)\#) \xrightarrow{s_4} (\#0E1 + 5(4, i + i)\#) \xrightarrow{s_3}$
 $(\#0E1 + 5(4i3, +i)\#) \xrightarrow{r_3} (\#0E1 + 5(4T2, +i)\#) \xrightarrow{r_1}$
 $(\#0E1 + 5(4E6, +i)\#) \xrightarrow{s_5} (\#0E1 + 5(4E6 + 5, i)\#) \xrightarrow{s_3}$
 $(\#0E1 + 5(4E6 + 5i3,)\#) \xrightarrow{r_3} (\#0E1 + 5(4E6 + 5T7,)\#) \xrightarrow{r_2}$
 $(\#0E1 + 5(4E6,)\#) \xrightarrow{s_8} (\#0E1 + 5(4E6)8, \#) \xrightarrow{r_4} (\#0E1 + 5T7, \#) \xrightarrow{r_2}$
 $(\#0E1, \#) \xrightarrow{acc}$

LR(1) elemzés

Példa

Készítse el az alábbi kiterjesztett nyelvtan SLR elemzőjét!

- $S' \rightarrow S$
- $S \rightarrow A := B \mid B$
- $A \rightarrow fB \mid i$
- $B \rightarrow A$

Megoldás

- $[I_0] = V(\lambda) = \{[S' \rightarrow \bullet S], [S \rightarrow \bullet A := B], [S \rightarrow \bullet B], [A \rightarrow \bullet fB], [A \rightarrow \bullet i], [B \rightarrow \bullet A]\}$
- $[I_1] = V(S) = \{[S' \rightarrow S \bullet]\}$
- $[I_2] = V(A) = \{[S \rightarrow A \bullet := B], [B \rightarrow A \bullet]\}$
- $[I_3] = V(B) = \{[S \rightarrow B \bullet]\}$
- $[I_4] = V(f) = \{[A \rightarrow f \bullet B], [B \rightarrow \bullet A], [A \rightarrow \bullet fB], [A \rightarrow \bullet i]\}$
- $[I_5] = V(i) = \{[A \rightarrow \bullet i]\}$
- $[I_6] = V(A :=) = \{[S \rightarrow A := \bullet B], [B \rightarrow \bullet A], [A \rightarrow \bullet fB], [A \rightarrow \bullet i]\}$
- $[I_7] = V(fB) = \{[A \rightarrow fB \bullet]\}$
- $[I_8] = V(fA) = \{[B \rightarrow A \bullet]\}$
- $[I_9] = V(A := B) = \{[S \rightarrow A := B \bullet]\}$

Az SLR táblázatok kitöltése nem konfliktusmentes (*action*[2, :=]), ezért nem SLR nyelvtan.

LR(1) elem

Ha a G' nyelvtan helyettesítési szabálya $A \rightarrow \alpha\beta$, akkor a nyelvtan $\{LR(1) \text{ eleme} \mid [A \rightarrow \alpha \bullet \beta, a], \text{ ahol } A \rightarrow \alpha \bullet \beta \text{ az LR(1)-elem magja, és az } a \in T \cup \{\#\} \text{ az LR(1)-elem előreolvasási szimbóluma.}\}$

Egy G' nyelvtan $[A \rightarrow \alpha \bullet \beta, a]$ LR(1)-eleme *érvényes* a $\gamma\alpha$ járható prefixre nézve, ha $S' \xRightarrow{*} \gamma Ax \Rightarrow \gamma\alpha\beta x$, és az a az x első szimbóluma, vagy ha $x = \lambda$, akkor $a = \#$.

closure

Legyen \mathcal{H} halmaz egy nyelvtan egy LR(1)-elemhalmaza. Ekkor a $\text{closure}(\mathcal{H})$ halmaz a következő elemeket tartalmazza:

- a \mathcal{H} halmaz minden eleme legyen eleme a $\text{closure}(\mathcal{H})$ halmaznak is,
- ha a $[A \rightarrow \alpha \bullet B\beta, a] \in \text{closure}(\mathcal{H})$ és $B \rightarrow \gamma$ a nyelvtan egy helyettesítési szabálya, akkor legyen $[B \rightarrow \bullet\gamma, b] \in \text{closure}(\mathcal{H})$ minden $b \in \text{First}_1(\beta a)$ -ra.

LR(1) kanonikus halmaz

read

Legyen \mathcal{H} halmaz egy nyelvtan egy LR(1)-elemhalmaza. Ekkor a $read(\mathcal{H}, X)$ halmaz a következő LR(1)-elemeket tartalmazza: * ha $[A \rightarrow \alpha \bullet X\beta, a] \in \mathcal{H}$, akkor a $closure([A \rightarrow \alpha X \bullet \beta, a])$ minden eleme legyen a $read(\mathcal{H}, X)$ halmaz eleme.

kanonikus halmaz

$\mathcal{H}_0, \dots, \mathcal{H}_m$ elemek *kanonikus halmazai* a következők:

- Legyen $\mathcal{H}_0 = closure([S' \rightarrow \bullet S, \#])$
- a többi halmazt az LR(0) kanonikus halmazoknál ismertetett módon a *read* függvénnyel lehet előállítani.

LR(1) elemzés

Definíció

Egy γ járható prefixre érvényes LR(1)-elemek halmaza az a \mathcal{H}_k ka-no-ni-kus elemhalmaz, amelyik az elemző determinisztikus véges automatájának ahhoz a k állapotához tartozik, amelyikbe az automata a γ hatására kerül.

Példa

Készítse el az alábbi kiterjesztett nyelvtan LR(1) elemzőjét!

- $S \rightarrow AA,$
- $A \rightarrow aA|b$

Megoldás

- $\mathcal{H}_0 = \{[S' \rightarrow \bullet S, \#], [S \rightarrow \bullet AA, \#], [A \rightarrow \bullet aA, a/b], [A \rightarrow \bullet b, a/b]\}$
- $\mathcal{H}_1 = \{[S' \rightarrow S\bullet, \#]\}$
- $\mathcal{H}_2 = \{[S \rightarrow A\bullet A, \#], [A \rightarrow \bullet aA, \#], [A \rightarrow \bullet b, \#]\}$
- $\mathcal{H}_3 = \{[A \rightarrow a\bullet A, a/b], [A \rightarrow \bullet aA, a/b], [A \rightarrow \bullet b, a/b]\}$
- $\mathcal{H}_4 = \{[A \rightarrow b\bullet, a/b]\}$
- $\mathcal{H}_5 = \{[S \rightarrow AA\bullet, \#]\}$
- $\mathcal{H}_6 = \{[A \rightarrow a\bullet A, \#], [A \rightarrow \bullet aA, \#], [A \rightarrow \bullet b, \#]\}$
- $\mathcal{H}_7 = \{[A \rightarrow b\bullet, \#]\}$
- $\mathcal{H}_8 = \{[A \rightarrow aA\bullet, a/b]\}$
- $\mathcal{H}_9 = \{[A \rightarrow aA\bullet, \#]\}$

LR(1) elemző táblázat

Az LR(1) elemző táblázat szerkezete az SLR elemző táblázatokéhoz hasonló.

A táblázat i -dik sorát a \mathcal{H}_i LR(1)-elemeiből töltjük ki.

- ha $[A \rightarrow \alpha \bullet a\beta, b] \in \mathcal{H}_i$ és $read(\mathcal{H}_i, a) = \mathcal{H}_j$, akkor legyen $action[i, a] = s_j$,
- ha $[A \rightarrow \alpha \bullet, a] \in \mathcal{H}_i$ és $A \neq S'$, akkor legyen $action[i, a] = r_l$, ahol az $A \rightarrow \alpha$ a nyelvtan l -dik szabálya.
- ha $[S' \rightarrow S \bullet, \#] \in \mathcal{H}_i$, akkor legyen $action[i, \#] = acc$
- ha $read(\mathcal{H}_i, A) = \mathcal{H}_j$, akkor legyen $goto[i, A] = j$.

LALR(1) elemző

Az LR(1)-elemek kanonikus halmazai között vannak olyan halmazpárok, hogy az egyik halmazban levő minden elemnek van megfelelője a másik halmazban (a mag azonos, csak az előreolvasási szimbólumokban különböznek).

Az ilyen kanonikus halmazokat egyesítsük, és az átsorszámozás után kapott halmazokat nevezzük *egyesített LR(1) kanonikus halmazoknak* vagy *LALR(1) kanonikus halmazoknak*. Ezekből a halmazokból létrejövő elemzőt *LALR(1) elemzőnek* nevezzük.

Előző példában a 3/6, 4/7 és 8/9 halmazok egyesíthetők.

Definíció

Ha a G' kiegészített nyelvtanra az LALR(1) elemzőtáblázat kitöltése konfliktusmentes, akkor a nyelvtant *LALR(1) nyelvtannak* nevezzük.