

§ 3. LR(1)-автомат, LR(1)-анализатор, LR(1)-грамматика

Опр. LR(1)-**пунктом** расширенной грамматики G называется набор (A, β_1, β_2, a) , где $A \rightarrow \beta_1 \beta_2$ – правило грамматики, $a \in \Sigma \cup \{\mid\}$.

Обозначение: $[A \rightarrow \beta_1 \bullet \beta_2, a]$.

Опр. **Ядром** LR(1)-пункта называется пункт $[A \rightarrow \beta_1 \bullet \beta_2]$.

Опр. LR(1)-пункт $[A \rightarrow \beta_1 \bullet \beta_2, a]$ называется **допустимым** для активного префикса $\gamma = \gamma' \beta_1$ некоторой r -формы, если существует вывод $S' \Rightarrow^* \gamma' A v \Rightarrow \gamma' \beta_1 \beta_2 v \Rightarrow^* uv$, и цепочка $v \mid$ начинается символом a .

Комментарий: в процессе анализа после переноса в стек β_2 во входном потоке будет a , поскольку v – необработанная часть цепочки.

Опр. Автомат LR(1)-пунктов содержит переходы двух видов:

$$1) \delta([A \rightarrow \beta_1 \bullet X \beta_3, a], X) = [A \rightarrow \beta_1 X \bullet \beta_3, a];$$

$$2) \delta([A \rightarrow \beta_1 \bullet B \beta_3, a], \varepsilon) = [B \rightarrow \bullet \beta, b], \text{ где } b \in \text{FIRST}(\beta_3 a).$$

Начальное состояние автомата: LR(1)-пункт $[S' \rightarrow \bullet S, \text{—} |]$.

Заключительные – все состояния.

Комментарий: в процессе построения переходов такого автомата появляются только допустимые LR(1)-пункты (покажем позже).

Опр. LR(1)-автоматом грамматики G называется ϵ -замыкание автомата LR(1)-пунктов.

Комментарий: в задачах на практических занятиях будем строить ϵ -замыкание без поиска ϵ -НКА LR(1)-пунктов.

Пример. $G = \{S' \rightarrow S, S \rightarrow ac \mid bDc \mid Da, D \rightarrow a\}$.

(стр. 171 учебного пособия А.Ф. Шура и А.П. Замятина содержит ошибку)

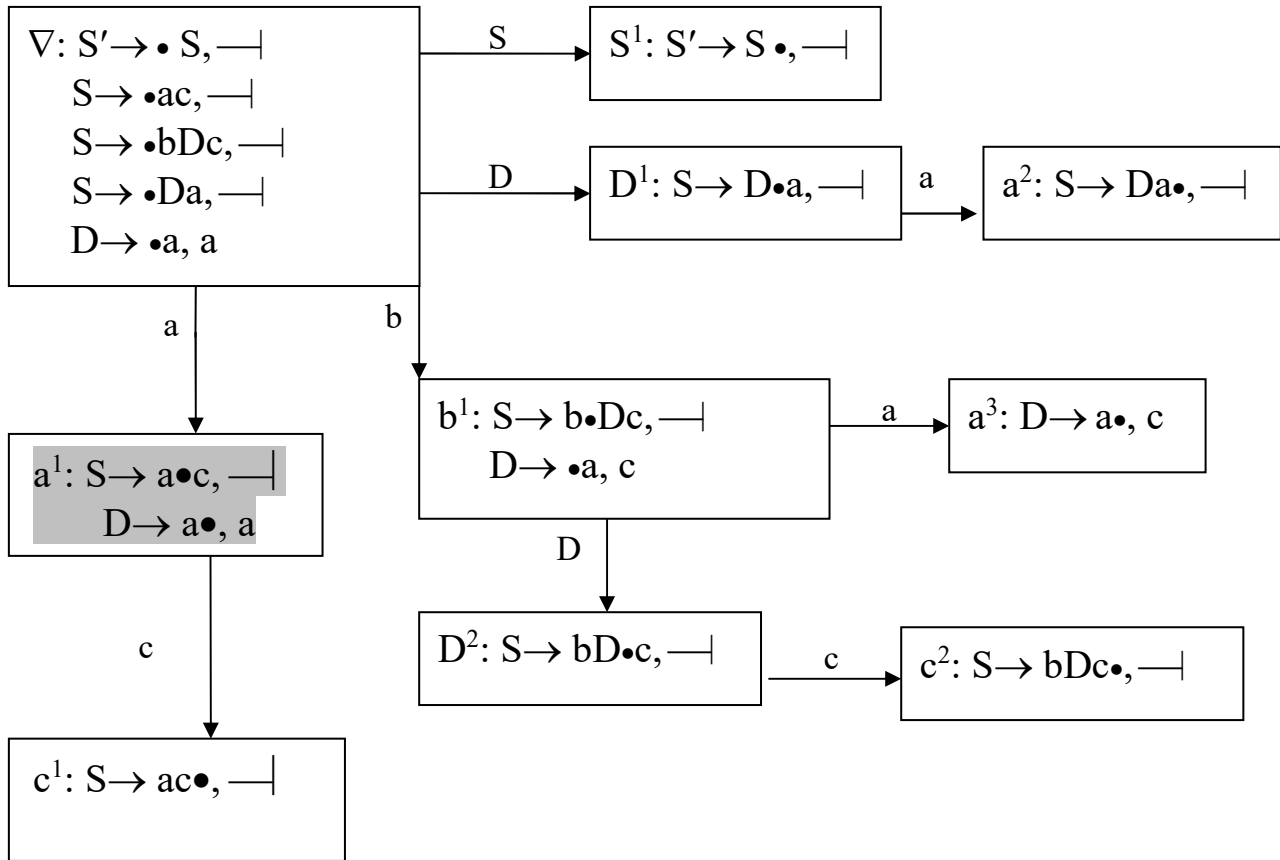
1) Построим массив FIRST для нетерминалов грамматики:

нетерм.	FIRST(..)
S'	
S	
D	

$G = \{S' \rightarrow S, S \rightarrow ac \mid bDc \mid Da, D \rightarrow a\}.$

нетерм.	FIRST(..)
S'	a, b
S	a, b
D	a

2) LR(1)-автомат:



Алгоритм построения LR(1)-анализатора

Заполнение ACTION:

- 1) В строке $q = [S' \rightarrow S\bullet, \text{---}]$ в столбце --- заносим $\checkmark$ (допуск).
- 2) В каждой строке q , не заполненной в 1):
для каждого пункта $[A \rightarrow \alpha\bullet, a]$, в столбце a заносим $\otimes_n$ (свертка);
для каждого пункта $[A \rightarrow \beta_1 \bullet a\beta_3, b]$ в столбце a заносим $a^i = \leftarrow \delta(q, a)$ (перенос).

Заполнение GOTO:

В каждой строке q для каждого пункта $[A \rightarrow \beta_1 \bullet B\beta_3, b]$ в столбце B заносим $B^j = \delta(q, B)$.

Напоминание: обозначение a для элементов $\Sigma \cup \{\text{---}\}$,
обозначение B для элементов Γ .

Опр. Грамматика называется **LR(1)-грамматикой**, если построенная таблица анализатора не содержит конфликтов.

Комментарий: $\{ \text{SLR}(1)\text{-грамм.} \} \subset \{ \text{LR}(1)\text{-грамм.} \}$

Точное доказательство будет следовать из основной теоремы LR-анализа, основная идея будет следовать из факта, что при выводе

$S' \Rightarrow^* \gamma' A v \Rightarrow \gamma' \beta_1 B \beta_3 v \Rightarrow^* uv$, соответствующем переходу

$\delta([A \rightarrow \beta_1 \bullet B \beta_3, a], \varepsilon) = [B \rightarrow \bullet \beta, b]$, выполняется

$\text{FIRST}(\beta_3 a) \subseteq \text{FOLLOW}(B)$.

Следствие: $\text{LR}(0) \subset \text{SLR}(1) \subset \text{LR}(1)$.

Пример, рассмотренный ранее $G = \{S' \rightarrow S, S \rightarrow ac \mid bDc \mid Da, D \rightarrow a\}$.

	ACTION				GOTO	
	<i>a</i>	<i>b</i>	<i>c</i>	$\mid$	S	D
∇						
a^1						
a^2						
a^3						
b^1						
c^1						
c^2						
S^1						
D^1						
D^2						

	ACTION				GOTO	
	a	b	c	$\neg$	S	D
∇	$\leftarrow a^1$	$\leftarrow b^1$			S^1	D^1
a^1	$\otimes 5$		$\leftarrow c^1$			
a^2				$\otimes 4$		
a^3			$\otimes 5$			
b^1	$\leftarrow a^3$					D^2
c^1				$\otimes 2$		
c^2				$\otimes 3$		
S^1				$\vee$		
D^1	$\leftarrow a^2$					
D^2			$\leftarrow c^2$			

Вопрос: является ли G LR(1)-грамматикой?

Упр. Является ли G LR(0)-грамматикой, SLR(1)-грамматикой?

Указание: рассмотреть $FOLLOW(D)$.