ПРЕОБРАЗОВАНИЕ КС-ГРАММАТИКИ В LL(1)-ГРАММАТИКУ

Будем определять, по какой причине КС-грамматика не является LL(1)-грамматикой, и устранять эту причину.

1) Леворекурсивная грамматика не может быть LL(1)-грамматикой.

$A \to \beta$
$A \to A\alpha$

Если $\beta \Rightarrow^* t\gamma$, то
1) $t \in$ ВЫБОР$(A \to \beta)$;
2) $A\alpha \Rightarrow \beta\alpha \Rightarrow^* t\gamma\alpha$ и $t \in$ ВЫБОР$(A \to A\alpha)$.
Следовательно, ВЫБОР$(A \to \beta) \cap$ ВЫБОР$(A \to A\alpha) \neq \varnothing$.

Если $\beta \Rightarrow^* \varepsilon$, то
1) СЛЕД$(A) \subseteq$ ВЫБОР$(A \to \beta)$;
2) $A \Rightarrow^* \varepsilon$ и ПЕРВ$(\alpha) \subseteq$ ПЕРВ$(A\alpha) \subseteq$ ВЫБОР$(A \to A\alpha)$.
По правилу $A \to A\alpha$ ПЕРВ$(\alpha) \subseteq$ СЛЕД(A).
Получаем, что ПЕРВ$(\alpha) \subseteq$ СЛЕД$(A) \subseteq$ ВЫБОР$(A \to \beta)$
и ПЕРВ$(\alpha) \subseteq$ ПЕРВ$(A\alpha) \subseteq$ ВЫБОР$(A \to A\alpha)$.
Следовательно, ВЫБОР$(A \to \beta) \cap$ ВЫБОР$(A \to A\alpha) \neq \varnothing$.

Чтобы преобразовать КС-грамматику в LL(1)-грамматику,
нужно устранить левую рекурсию.

2) Если в грамматике хотя бы два правила с одинаковой левой частью имеют общий префикс, из которого выводится непустая цепочка, то такая грамматика не может быть LL(1)-грамматикой.

$A \to \alpha\beta$

$A \to \alpha\gamma$

ПЕРВ(α) $\subseteq$ ПЕРВ($\alpha\beta$) $\subseteq$ ВЫБОР($A \to \alpha\beta$).

ПЕРВ(α) $\subseteq$ ПЕРВ($\alpha\gamma$) $\subseteq$ ВЫБОР($A \to \alpha\gamma$).

Следовательно, ВЫБОР($A \to \alpha\beta$) $\cap$ ВЫБОР($A \to \alpha\gamma$) $\neq \varnothing$.

Чтобы преобразовать КС-грамматику в LL(1)-грамматику,

нужно выполнять левую факторизацию.

3) Если в грамматике есть хотя бы два правила с одинаковой левой частью, из правых частей которых выводятся цепочки, имеющие общий префикс, из которого выводится непустая цепочка, то такая грамматика не может быть LL(1)-грамматикой.

$A \to \alpha$

$A \to \beta$

Пусть $\alpha =>^* \delta\gamma$ и $\beta =>^* \delta\varphi$

Если $\delta =>^* t\mu$, то

1) $t \in$ ПЕРВ(α) и $t \in$ ВЫБОР($A \to \alpha$);

2) $t \in$ ПЕРВ(β) и $t \in$ ВЫБОР($A \to \beta$).

Следовательно, ВЫБОР($A \to \alpha$) $\cap$ ВЫБОР($A \to \beta$) $\neq \varnothing$.

Чтобы преобразовать КС-грамматику в LL(1)-грамматику,

нужно выполнять замену края и левую факторизацию.

4) Если в грамматике есть два правила

$A \to \alpha$

$A \to \varepsilon$

и $x \in$ ПЕРВ(α) $\cap$ СЛЕД(A), то ВЫБОР($A \to \alpha$) $\cap$ ВЫБОР($A \to \varepsilon$) $\neq \varnothing$ и такая грамматика не может быть LL(1)-грамматикой.

Чтобы преобразовать такую КС-грамматику в LL(1)-грамматику,
нужно;

1) преобрзовать грамматику так, чтобы в правилах грамматики после нетерминала A стоял терминал x (используем замену нетерминала, который стоит после нетерминала A);
2) каждое правило вида $B \to \alpha A x \beta$ заменить на правило $B \to \alpha N \beta$, где N — новый нетерминал, и добавить правило $N \to Ax$;
3) к правилу вида $N \to Ax$ применять замену края и левую факторизацию. Если в процессе преобразований получится правило, содержащее в правой части цепочку Ax, то заменить её на нетерминал N.

Таким образом, получаем последовательность действий, которая может преобразовать произвольную КС-грамматику в LL(1)-грамматику (а может и не преобразовать):

1) устранить левую рекурсию;

2) выполнить левую факторизацию;

3) выполнить замену края;

4) п. 2 и 3 выполнять, пока можно;

5) пока в грамматике есть правила

$A \to \alpha$

$A \to \varepsilon$

и ВЫБОР($A \to \alpha$) $\cap$ ВЫБОР($A \to \varepsilon$) $\neq \varnothing$, то

1) преобрзовать грамматику так, чтобы в правилах грамматики после нетерминала A стоял терминал

$x \in$ ВЫБОР($A \to \alpha$) $\cap$ ВЫБОР($A \to \varepsilon$)

(используем замену нетерминала, который стоит после нетерминала A);

2) каждое правило вида $B \to \alpha A x \beta$ заменить на правило $B \to \alpha N \beta$, где N — новый нетерминал, и добавить правило $N \to Ax$;

3) к правилу вида $N \to Ax$ применять замену края и левую факторизацию. Если в процессе преобразований получится правило, содержащее в правой части цепочку Ax, то заменить её на нетерминал N.

Пример 1.

	Шаг 1	Шаг 2	Шаг 3
$A \to BaC$ {b,a} $A \to bA$ $A \to aC$ $B \to CaB$ {b} $B \to \varepsilon$ {a} $C \to ba$ $C \to bC$	$A \to BaC$ $A \to bA$ $A \to aC$ $B \to CaB$ $B \to \varepsilon$ $C \to bN_1$ $N_1 \to a$ $N_1 \to C$	$A \to BabN_1$ $A \to bA$ $A \to abN_1$ $B \to bN_1aB$ $B \to \varepsilon$ $N_1 \to a$ $N_1 \to bN_1$	$A \to bN_1aBabN_1$ $A \to abN_1$ $A \to bA$ $B \to bN_1aB$ $B \to \varepsilon$ $N_1 \to a$ $N_1 \to bN_1$
Шаг 4	Шаг 5	Шаг 6	Шаг 7
$A \to bN_2$ $A \to abN_1$ $B \to bN_1aB$ $B \to \varepsilon$ $N_1 \to a$ $N_1 \to bN_1$ $N_2 \to N_1aBabN_1$ $N_2 \to A$ А берём с шага 3	$A \to bN_2$ $A \to abN_1$ $B \to bN_1aB$ $B \to \varepsilon$ $N_1 \to a$ $N_1 \to bN_1$ $N_2 \to aaBabN_1$ $N_2 \to bN_1aBabN_1$ $N_2 \to abN_1$ $N_2 \to bA$	$A \to bN_2$ $A \to abN_1$ $B \to bN_1aB$ $B \to \varepsilon$ $N_1 \to a$ $N_1 \to bN_1$ $N_2 \to aN_3$ $N_2 \to bN_4$ $N_3 \to aBabN_1$ $N_3 \to bN_1$ $N_4 \to N_1aBabN_1$ $N_4 \to A$ $L(N_4) = L(N_2)$ См шаг 4	$A \to bN_2$ $A \to abN_1$ $B \to bN_1aB$ $B \to \varepsilon$ $N_1 \to a$ $N_1 \to bN_1$ $N_2 \to aN_3$ $N_2 \to bN_2$ $N_3 \to aBabN_1$ $N_3 \to bN_1$

Пример 2.

	Шаг 1	Шаг 2
$S \rightarrow aBa$ $S \rightarrow cbAaSc$ $S \rightarrow \varepsilon$ $\{c, \dashv\}$ $A \rightarrow abA$ $A \rightarrow \varepsilon$ $\{a\}$ $B \rightarrow aA$ $B \rightarrow \varepsilon$ $\{a\}$	$S \rightarrow aN_1$ $S \rightarrow cb\,N_2N_3$ $S \rightarrow \varepsilon$ $A \rightarrow abA$ $A \rightarrow \varepsilon$ $B \rightarrow aA$ $B \rightarrow \varepsilon$ $N_1 \rightarrow Ba$ $N_2 \rightarrow Aa$ $N_3 \rightarrow Sc$	$S \rightarrow aN_1$ $S \rightarrow cbN_2N_3$ $S \rightarrow \varepsilon$ $A \rightarrow abA$ $A \rightarrow \varepsilon$ $B \rightarrow aA$ $B \rightarrow \varepsilon$ $N_1 \rightarrow aAa$ $N_1 \rightarrow a$ $N_2 \rightarrow abAa$ $N_2 \rightarrow a$ $N_3 \rightarrow aN_1c$ $N_3 \rightarrow cbN_2N_3c$ $N_3 \rightarrow c$
Шаг 3	Шаг 4	Шаг 5
$S \rightarrow aN_1$ $S \rightarrow cbN_2N_3$ $S \rightarrow \varepsilon$ $A \rightarrow abA$ $A \rightarrow \varepsilon$ $B \rightarrow aA$ $B \rightarrow \varepsilon$ $N_1 \rightarrow aN_2$ $N_1 \rightarrow a$ $N_2 \rightarrow abN_2$ $N_2 \rightarrow a$ $N_3 \rightarrow aN_1c$ $N_3 \rightarrow cbN_2N_3c$ $N_3 \rightarrow c$	$S \rightarrow aN_1$ $S \rightarrow cbN_2N_3$ $S \rightarrow \varepsilon$ $A \rightarrow abA$ $A \rightarrow \varepsilon$ $B \rightarrow aA$ $B \rightarrow \varepsilon$ $N_1 \rightarrow aN_4$ $N_2 \rightarrow aN_5$ $N_3 \rightarrow aN_1c$ $N_3 \rightarrow cN_6$ $N_4 \rightarrow N_2$ $N_4 \rightarrow \varepsilon$ $N_5 \rightarrow bN_2$ $N_5 \rightarrow \varepsilon$ $N_6 \rightarrow bN_2N_3c$ $N_6 \rightarrow \varepsilon$	$S \rightarrow aN_1$ $S \rightarrow cbN_2N_3$ $S \rightarrow \varepsilon$ $\{\dashv\}$ $N_1 \rightarrow aN_4$ $N_2 \rightarrow aN_5$ $N_3 \rightarrow aN_1c$ $N_3 \rightarrow cN_6$ $N_4 \rightarrow N_2$ $\{a\}$ $N_4 \rightarrow \varepsilon$ $\{c, \dashv\}$ $N_5 \rightarrow bN_2$ $N_5 \rightarrow \varepsilon$ $\{a,c, \dashv\}$ $N_6 \rightarrow bN_2N_3c$ $N_6 \rightarrow \varepsilon$ $\{c, \dashv\}$