

Приведенные КС-грамматики

Символ $x \in (V_T \cup V_N)$ называется *недостижимым* в грамматике $G=(V_T, V_N, P, S)$, если он не появляется ни в одной сентенциальной форме этой грамматики.

Символ $A \in V_N$ называется *бесплодным* в грамматике $G=(V_T, V_N, P, S)$, если множество выводимых из этого символа терминальных цепочек пусто.

КС-грамматика называется *приведенной*, если в ней нет недостижимых и бесплодных символов.

Приведенные КС-грамматики

Алгоритм приведения грамматики:

1. Найти и удалить все бесплодные символы и правила, их содержащие.
2. Найти и удалить все недостижимые символы и правила, их содержащие.

Примечание. Если начальный символ грамматики окажется бесплодным, то следует удалить содержащие его правила, а сам символ оставить в алфавите нетерминалов V_N , так как по определению грамматики V_N обязан содержать начальный символ.

Для нахождения бесплодных и недостижимых символов полезен граф КС-грамматики:

- каждому символу из $V_T \cup V_N$ соответствует единственная вершина, помеченная этим символом; если в P есть правило с пустой правой частью ϵ , то граф имеет вершину, помеченную ϵ ;
- вершина X соединяется с вершиной Y стрелкой (дугой), если в грамматике есть правило $X \rightarrow \alpha Y \beta$, $\alpha, \beta \in (V_T \cup V_N)^*$;
- X соединяется с вершиной ϵ , если в грамматике есть правило $X \rightarrow \epsilon$.

Алгоритм удаления бесплодных символов:

1. Отметить терминальные вершины (вершины, помеченные терминальными символами), а также вершину ϵ , если таковая имеется.
2. Если в P есть правило $A \rightarrow \alpha$, где α состоит из уже отмеченных в графе символов, а вершина A не отмечена, то отметить эту вершину. Повторять шаг 2 пока возможно.
3. Из грамматики удалить неотмеченные символы и правила, их содержащие.

Алгоритм удаления бесплодных символов:

1. Отметить терминальные вершины (вершины, помеченные терминальными символами), а также вершину ε , если такая имеется.
2. Если в P есть правило $A \rightarrow \alpha$, где α состоит из уже отмеченных в графе символов, а вершина A не отмечена, то отметить эту вершину. Повторять шаг 2 пока возможно.
3. Из грамматики удалить неотмеченные символы и правила, их содержащие.

Алгоритм удаления недостижимых символов:

1. Отметить вершины, в которые есть путь из вершины S .
2. Удалить из грамматики неотмеченные символы и правила, их содержащие.

Пример. Дана грамматика

$G = (\{a, b, c, d, e\}, \{S, A, B, C, D\}, P, S)$

P: $S \rightarrow aAB \mid C$

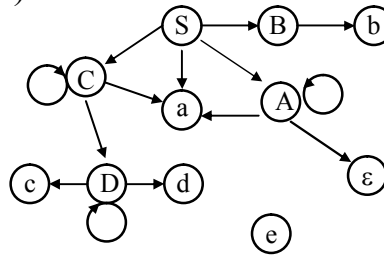
$D \rightarrow cDc \mid d$

$C \rightarrow aCD$

$A \rightarrow aA \mid a \mid \varepsilon$

$B \rightarrow b$

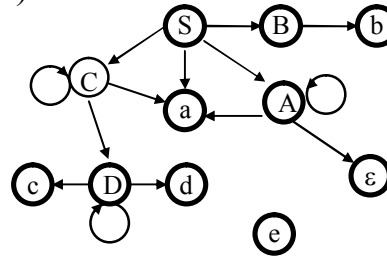
Граф грамматики G:



Пример. Дана грамматика
 $G = (\{a, b, c, d, e\}, \{S, A, B, C, D\}, P, S)$

P: $S \rightarrow aAB \mid C$
 $D \rightarrow cDc \mid d$
 $C \rightarrow aCD$
 $A \rightarrow aA \mid a \mid \varepsilon$
 $B \rightarrow b$

Граф грамматики G:

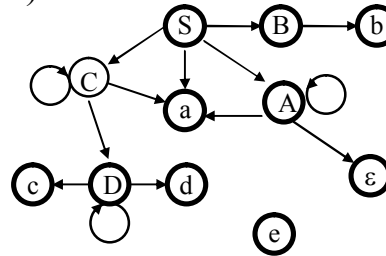


Не отмеченные жирным кружком символы бесплодны.

Пример. Дана грамматика
 $G = (\{a, b, c, d, e\}, \{S, A, B, C, D\}, P, S)$

$P:$
 $S \rightarrow aAB \mid C$
 $D \rightarrow cDc \mid d$
 $C \rightarrow aCD$
 $A \rightarrow aA \mid a \mid \varepsilon$
 $B \rightarrow b$

Граф грамматики $G:$



Не отмеченные жирным кружком символы бесплодны.

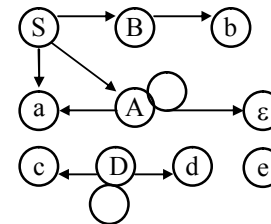
Удалив из G бесплодные символы, получим эквивалентную грамматику

$G_1 = (\{a, b, c, d, e\}, \{S, A, B, D\}, P_1, S)$

$P_1:$
 $S \rightarrow aAB$
 $D \rightarrow cDc \mid d$
 $A \rightarrow aA \mid a \mid \varepsilon$
 $B \rightarrow b$

G_1 не содержит бесплодных символов.

Граф грамматики $G_1:$



Находим недостижимые символы

$G_1 = (\{a, b, c, d, e\}, \{S, A, B, D\}, P_1, S)$

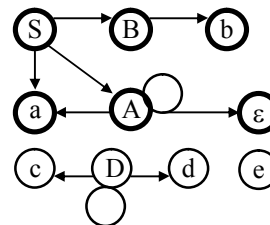
$P_1 : S \rightarrow aAB$

$D \rightarrow cDc \mid d$

$A \rightarrow aA \mid a \mid \varepsilon$

$B \rightarrow b$

Граф грамматики G_1 :



Здесь неотмеченные символы являются недостижимыми.

$G_1 = (\{a, b, \epsilon, \cancel{d}, \cancel{e}\}, \{S, A, B, \cancel{D}\}, P_1, S)$

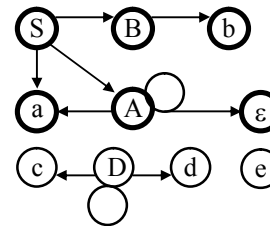
$P_1 : S \rightarrow aAB$

~~$D \rightarrow eDe \mid d$~~

$A \rightarrow aA \mid a \mid \epsilon$

$B \rightarrow b$

Граф грамматики G_1 :



Здесь неотмеченные символы являются недостижимыми.

Удалив из G_1 недостижимые символы, получим эквивалентную грамматику:

$G_2 = (\{a, b\}, \{A, B\}, P_2, S)$

$P_2 : S \rightarrow aAB$

$A \rightarrow aA \mid a \mid \epsilon$

$B \rightarrow b$

G_2 – приведенная грамматика

$L(G) = L(G_1) = L(G_2) = \{ a^n b \mid n \geq 1 \}$

Задача. Убедиться, что если в рассмотренном выше примере поменять местами шаги (1) и (2) алгоритма приведения грамматики, то результатом будет неприведенная грамматика.

Устранение правил с пустой правой частью из КС-грамматики

1. Построить множество $X = \{A \in N \mid A \Rightarrow \varepsilon\}$.
2. Удалить правила с пустой правой частью.
3. Если $S \in X$, то S' – новый начальный символ, $S' \rightarrow S \mid \varepsilon \in P$.
4. $\forall A \in X$ правило вида $B \rightarrow \alpha_1 A_1 \alpha_2 A_2 \dots \alpha_n A_n \alpha_{n+1}$,

где $\alpha_i \in ((N - X) \cup T)^*$

заменить 2^n правилами, соответствующими всем возможным комбинациям вхождений A между α_i :

$$B \rightarrow \alpha_1 \alpha_2 \dots \alpha_n \alpha_{n+1}$$

$$B \rightarrow \alpha_1 \alpha_2 \dots \alpha_n A_n \alpha_{n+1}$$

...

$$B \rightarrow \alpha_1 \alpha_2 A_2 \dots \alpha_n A_n \alpha_{n+1}$$

$$B \rightarrow \alpha_1 A_1 \alpha_2 A_2 \dots \alpha_n A_n \alpha_{n+1}$$

Замечание: если все $\alpha_i = \varepsilon \quad \forall i=1, \dots, n+1$, то правило $B \rightarrow \varepsilon$ не включать в новую грамматику.

5. Удалить бесполезные символы и правила, их содержащие.

Пример.
исходная
грамматика

$S \rightarrow BC \mid Ab$
 $B \rightarrow \varepsilon$
 $C \rightarrow c$
 $A \rightarrow Aa \mid \varepsilon$

эквивалентная
грамматика

$S \rightarrow C \mid b \mid Ab$
 $C \rightarrow c$
 $A \rightarrow Aa \mid a$