

Приведенные КС-грамматики

Дана грамматика G :

$$S \rightarrow aSB \mid C$$

$$D \rightarrow cDc \mid d$$

$$C \rightarrow a$$

Приведенные КС-грамматики

Дана грамматика G :

$$S \rightarrow aSB \mid C$$

$$D \rightarrow cDc \mid d$$

$$C \rightarrow a$$

Какой язык она порождает?

$$L(G)=?$$

Приведенные КС-грамматики

G :

$$\begin{aligned} S &\rightarrow aSB \mid C \\ D &\rightarrow cDc \mid d \\ C &\rightarrow a \end{aligned}$$

$$L(G) = \{a\}$$

G' :

$$S \rightarrow a$$

$$L(G') = \{a\}$$

Приведенные КС-грамматики

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

Символ $A \in N$ называется *непорождающим* (или *бесплодным*) в грамматике $G = \langle T, N, P, S \rangle$, 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 N , так как по определению грамматики N обязан содержать начальный символ. В алфавите терминалов T также следует оставить хотя бы один символ, так как алфавит не может быть пустым по определению.

Для нахождения бесплодных и недостижимых символов полезен граф «достижимости» для КС-грамматики:

- каждому символу из $T \cup N$ 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 $X \rightarrow \alpha Y \beta$, $\alpha, \beta \in (T \cup N)^*$;
- X соединяется с вершиной ε , если в грамматике есть правило $X \rightarrow \varepsilon$.

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 (достижимы из вершины S).
2. Удалить из грамматики неотмеченные символы и правила, их содержащие.

Пример

Дана грамматика

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

$P:$ $S \rightarrow aAB \mid C$

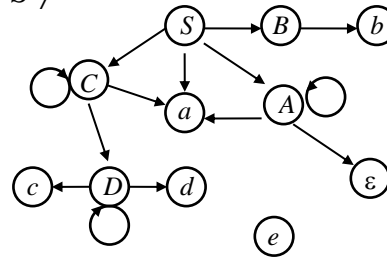
$D \rightarrow cDc \mid d$

$C \rightarrow aCD$

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

$B \rightarrow b$

Граф грамматики G :



Дана грамматика

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

$$P: \quad S \rightarrow aAB \mid C$$

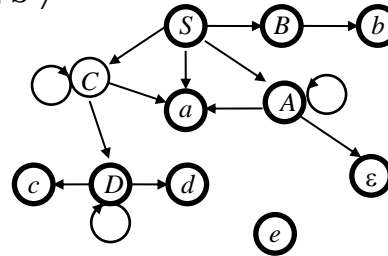
$$D \rightarrow cDc \mid d$$

$$C \rightarrow aCD$$

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

$$B \rightarrow b$$

Граф грамматики G :



Не отмеченные жирным кружком символы бесплодны.

Дана грамматика

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

$$P: \quad S \rightarrow aAB \mid C$$

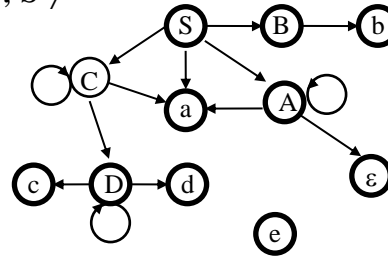
$$D \rightarrow cDc \mid d$$

$$C \rightarrow aCD$$

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

$$B \rightarrow b$$

Граф грамматики G :



Не отмеченные жирным кружком символы бесплодны.

Удалив из G бесплодные символы, получим эквивалентную грамматику

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

$$P_1: \quad S \rightarrow aAB$$

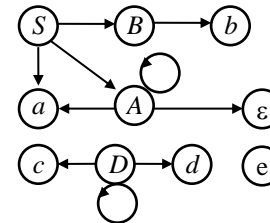
$$D \rightarrow cDc \mid d$$

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

$$B \rightarrow b$$

G_1 не содержит бесплодных символов.

Граф грамматики G_1 :



Находим недостижимые символы

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

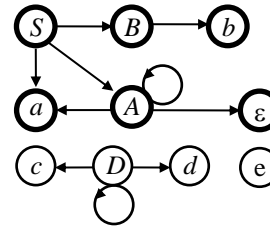
$$P_1: S \rightarrow aAB$$

$$D \rightarrow cDc \mid d$$

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

$$B \rightarrow b$$

Граф грамматики G_1 :



Здесь неотмеченные символы являются недостижимыми.

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

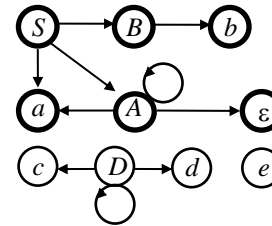
$$P_1: S \rightarrow aAB$$

$$\cancel{D} \rightarrow \cancel{cD\epsilon} \mid \cancel{d}$$

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

$$B \rightarrow b$$

Граф грамматики G_1 :



Здесь неотмеченные символы являются недостижимыми.

Удалив из G_1 недостижимые символы, получим эквивалентную грамматику:

$$G_2 = \langle \{a, b\}, \{S, A, B\}, P_2, S \rangle$$

$$P_2: S \rightarrow aAB$$

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

$$B \rightarrow b$$

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. $\forall A_i \in X$ 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_i 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 получившееся правило

$B \rightarrow \varepsilon$ не включать в новую грамматику.

5. Удалить бесполезные символы и правила, их содержащие.

Пример

исходная	$S \rightarrow BC \mid Ab$	эквивалентная	$S \rightarrow C \mid b \mid Ab$
грамматика	$B \rightarrow \varepsilon$	грамматика	$C \rightarrow c$
	$C \rightarrow c$		$A \rightarrow Aa \mid a$
	$A \rightarrow Aa \mid \varepsilon$		