ε-свободные грамматики

Михайлова Инна Анатольевна

Институт математики и естественных наук.
Кафедра алгебры и фундаментальной информатики.

ε-свободные грамматики

- Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, нетерминал A называется аннулирующим, если $A \Rightarrow^* \varepsilon$. Множество аннулирующих символов грамматики G обозначим через $Ann(G)$.

ε-свободные грамматики

- Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, нетерминал A называется аннулирующим, если $A \Rightarrow^* \varepsilon$. Множество аннулирующих символов грамматики G обозначим через $Ann(G)$.
- КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется ε-свободной, если она либо не содержит аннулирующих символов, либо ее единственный аннулирующий символ - это аксиома и тогда $S \to \varepsilon \in P$ и аксиома не содержится в правых частях правил вывода грамматики.

- Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, нетерминал A называется аннулирующим, если $A \Rightarrow^* \varepsilon$. Множество аннулирующих символов грамматики G обозначим через $Ann(G)$.
- КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется ε-свободной, если она либо не содержит аннулирующих символов, либо ее единственный аннулирующий символ - это аксиома и тогда $S \to \varepsilon \in P$ и аксиома не содержится в правых частях правил вывода грамматики.

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

- Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, нетерминал A называется аннулирующим, если $A \Rightarrow^* \varepsilon$. Множество аннулирующих символов грамматики G обозначим через $Ann(G)$.
- КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется ε-свободной, если она либо не содержит аннулирующих символов, либо ее единственный аннулирующий символ - это аксиома и тогда $S \to \varepsilon \in P$ и аксиома не содержится в правых частях правил вывода грамматики.

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

- Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, найдем ее аннулирующие символы. Сначала $Ann(G) = \{A \in \Gamma \mid \exists A \to \varepsilon \in P\}$.

ε-свободные грамматики

- Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, нетерминал A называется аннулирующим, если $A \Rightarrow^* \varepsilon$. Множество аннулирующих символов грамматики G обозначим через $Ann(G)$.
- КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется ε-свободной, если она либо не содержит аннулирующих символов, либо ее единственный аннулирующий символ - это аксиома и тогда $S \to \varepsilon \in P$ и аксиома не содержится в правых частях правил вывода грамматики.

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

- Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, найдем ее аннулирующие символы. Сначала $Ann(G) = \{A \in \Gamma \mid \exists A \to \varepsilon \in P\}$.
- Затем добавим в $Ann(G)$ те нетерминалы B для которых существует такое правило $B \to \alpha$, что цепочка α состоит только из нетерминалов, которые уже принадлежат $Ann(G)$. Процесс заканчивается, когда новых символов добавить не получааается.

Теорема (продолжение)

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

Теорема (продолжение)

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

- На множестве цепочек над алфавитом $\Sigma \cup \Gamma$ зададим отношение $\alpha \trianglelefteq_G \beta$ тогда и только тогда, когда $\alpha \neq \varepsilon$ получается из β стиранием некоторого множества (возможно пустого) аннулирующих нетерминалов.

Теорема (продолжение)

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

- На множестве цепочек над алфавитом $\Sigma \cup \Gamma$ зададим отношение $\alpha \trianglelefteq_G \beta$ тогда и только тогда, когда $\alpha \neq \varepsilon$ получается из β стиранием некоторого множества (возможно пустого) аннулирующих нетерминалов.
- Построим грамматику $G' = (\Gamma, \Sigma, P', S)$, где

$$P' = P \setminus \{A \to \varepsilon \in P\} \cup \{B \to \alpha \mid \exists B \to \beta \in P \quad \alpha \trianglelefteq_G \beta\}$$

Теорема (продолжение)

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

- На множестве цепочек над алфавитом $\Sigma \cup \Gamma$ зададим отношение $\alpha \trianglelefteq_G \beta$ тогда и только тогда, когда $\alpha \neq \varepsilon$ получается из β стиранием некоторого множества (возможно пустого) аннулирующих нетерминалов.
- Построим грамматику $G' = (\Gamma, \Sigma, P', S)$, где

$$P' = P \setminus \{A \to \varepsilon \in P\} \cup \{B \to \alpha \mid \exists B \to \beta \in P \quad \alpha \trianglelefteq_G \beta\}$$

- Если в грамматике G аксиома S была аннулирующим символом, то добавим в грамматику G' новую аксиому S' и правила вывода $S' \to S$ и $S' \to \varepsilon$. Без ограничения общности будем считать, что S' — аксиома грамматики G'.

Теорема (продолжение)

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

- На множестве цепочек над алфавитом $\Sigma \cup \Gamma$ зададим отношение $\alpha \trianglelefteq_G \beta$ тогда и только тогда, когда $\alpha \neq \varepsilon$ получается из β стиранием некоторого множества (возможно пустого) аннулирующих нетерминалов.
- Построим грамматику $G' = (\Gamma, \Sigma, P', S)$, где

$$P' = P \setminus \{A \to \varepsilon \in P\} \cup \{B \to \alpha \mid \exists B \to \beta \in P \quad \alpha \trianglelefteq_G \beta\}$$

- Если в грамматике G аксиома S была аннулирующим символом, то добавим в грамматику G' новую аксиому S' и правила вывода $S' \to S$ и $S' \to \varepsilon$. Без ограничения общности будем считать, что S' — аксиома грамматики G'.
- Покажем, что $L(G) = L(G')$. Пусть $w \in L(G)$, тогда существует вывод $S \Rightarrow_G \eta_1 \Rightarrow_G \ldots \Rightarrow_G w$.

Теорема (продолжение)

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

- На множестве цепочек над алфавитом $\Sigma \cup \Gamma$ зададим отношение $\alpha \trianglelefteq_G \beta$ тогда и только тогда, когда $\alpha \neq \varepsilon$ получается из β стиранием некоторого множества (возможно пустого) аннулирующих нетерминалов.
- Построим грамматику $G' = (\Gamma, \Sigma, P', S)$, где

$$P' = P \setminus \{A \to \varepsilon \in P\} \cup \{B \to \alpha \mid \exists B \to \beta \in P \quad \alpha \trianglelefteq_G \beta\}$$

- Если в грамматике G аксиома S была аннулирующим символом, то добавим в грамматику G' новую аксиому S' и правила вывода $S' \to S$ и $S' \to \varepsilon$. Без ограничения общности будем считать, что S' — аксиома грамматики G'.
- Покажем, что $L(G) = L(G')$. Пусть $w \in L(G)$, тогда существует вывод $S \Rightarrow_G \eta_1 \Rightarrow_G \ldots \Rightarrow_G w$.
- Рассмотрим непосредственный вывод $\eta_i = \gamma_1 A \gamma_2$, $\eta_{i+1} = \gamma_1 \beta \gamma_2$, где $A \to \beta \in P$.

Теорема (продолжение)

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

- На множестве цепочек над алфавитом $\Sigma \cup \Gamma$ зададим отношение $\alpha \trianglelefteq_G \beta$ тогда и только тогда, когда $\alpha \neq \varepsilon$ получается из β стиранием некоторого множества (возможно пустого) аннулирующих нетерминалов.
- Построим грамматику $G' = (\Gamma, \Sigma, P', S)$, где

$$P' = P \setminus \{A \to \varepsilon \in P\} \cup \{B \to \alpha \mid \exists B \to \beta \in P \quad \alpha \trianglelefteq_G \beta\}$$

- Если в грамматике G аксиома S была аннулирующим символом, то добавим в грамматику G' новую аксиому S' и правила вывода $S' \to S$ и $S' \to \varepsilon$. Без ограничения общности будем считать, что S' — аксиома грамматики G'.
- Покажем, что $L(G) = L(G')$. Пусть $w \in L(G)$, тогда существует вывод $S \Rightarrow_G \eta_1 \Rightarrow_G \ldots \Rightarrow_G w$.
- Рассмотрим непосредственный вывод $\eta_i = \gamma_1 A \gamma_2$, $\eta_{i+1} = \gamma_1 \beta \gamma_2$, где $A \to \beta \in P$.
- Пусть $\beta = \beta_0 C_{i_1} \beta_1 C_{i_2} \ldots C_m \beta_m$ такова, что $C_{i_j} \Rightarrow_G^* \varepsilon$ в выводе w. Тогда $\beta_0 \beta_1 \ldots \beta_m \trianglelefteq_G \beta$ и существует правило $A \to \beta_0 \beta_1 \ldots \beta_m \in P'$.

Теорема (продолжение)

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

- На множестве цепочек над алфавитом $\Sigma \cup \Gamma$ зададим отношение $\alpha \trianglelefteq_G \beta$ тогда и только тогда, когда $\alpha \neq \varepsilon$ получается из β стиранием некоторого множества (возможно пустого) аннулирующих нетерминалов.
- Построим грамматику $G' = (\Gamma, \Sigma, P', S)$, где

$$P' = P \setminus \{A \to \varepsilon \in P\} \cup \{B \to \alpha \mid \exists B \to \beta \in P \quad \alpha \trianglelefteq_G \beta\}$$

- Если в грамматике G аксиома S была аннулирующим символом, то добавим в грамматику G' новую аксиому S' и правила вывода $S' \to S$ и $S' \to \varepsilon$. Без ограничения общности будем считать, что S' — аксиома грамматики G'.
- Покажем, что $L(G) = L(G')$. Пусть $w \in L(G)$, тогда существует вывод $S \Rightarrow_G \eta_1 \Rightarrow_G \ldots \Rightarrow_G w$.
- Рассмотрим непосредственный вывод $\eta_i = \gamma_1 A \gamma_2$, $\eta_{i+1} = \gamma_1 \beta \gamma_2$, где $A \to \beta \in P$.
- Пусть $\beta = \beta_0 C_{i_1} \beta_1 C_{i_2} \ldots C_m \beta_m$ такова, что $C_{i_j} \Rightarrow_G^* \varepsilon$ в выводе w. Тогда $\beta_0 \beta_1 \ldots \beta_m \trianglelefteq_G \beta$ и существует правило $A \to \beta_0 \beta_1 \ldots \beta_m \in P'$.
- Заменив в выводе в грамматике G каждое правило вывода на правило вывода в грамматике G', полученное так, как указано выше, мы получим вывод w в грамматике G'.

Теорема (продолжение)

Теорема

Для любой КС грамматики существует эквивалентная ей ε-свободная.

- На множестве цепочек над алфавитом $\Sigma \cup \Gamma$ зададим отношение $\alpha \trianglelefteq_G \beta$ тогда и только тогда, когда $\alpha \neq \varepsilon$ получается из β стиранием некоторого множества (возможно пустого) аннулирующих нетерминалов.
- Построим грамматику $G' = (\Gamma, \Sigma, P', S)$, где

$$P' = P \setminus \{A \to \varepsilon \in P\} \cup \{B \to \alpha \mid \exists B \to \beta \in P \quad \alpha \trianglelefteq_G \beta\}$$

- Если в грамматике G аксиома S была аннулирующим символом, то добавим в грамматику G' новую аксиому S' и правила вывода $S' \to S$ и $S' \to \varepsilon$. Без ограничения общности будем считать, что S' — аксиома грамматики G'.
- Покажем, что $L(G) = L(G')$. Пусть $w \in L(G)$, тогда существует вывод $S \Rightarrow_G \eta_1 \Rightarrow_G \ldots \Rightarrow_G w$.
- Рассмотрим непосредственный вывод $\eta_i = \gamma_1 A \gamma_2$, $\eta_{i+1} = \gamma_1 \beta \gamma_2$, где $A \to \beta \in P$.
- Пусть $\beta = \beta_0 C_{i_1} \beta_1 C_{i_2} \ldots C_m \beta_m$ такова, что $C_{i_j} \Rightarrow_G^* \varepsilon$ в выводе w. Тогда $\beta_0 \beta_1 \ldots \beta_m \trianglelefteq_G \beta$ и существует правило $A \to \beta_0 \beta_1 \ldots \beta_m \in P'$.
- Заменив в выводе в грамматике G каждое правило вывода на правило вывода в грамматике G', полученное так, как указано выше, мы получим вывод w в грамматике G'.
- Обратное включение доказывается аналогично.

Пример

Построить ε-свободную грамматику, эквивалентную данной

$$S \to AB \mid cC$$
$$A \to \varepsilon \mid aA$$
$$B \to AB \mid \varepsilon$$
$$C \to ABC \mid c$$

Пример

Построить ε-свободную грамматику, эквивалентную данной

$$\begin{aligned} S &\to AB \mid cC \\ A &\to \varepsilon \mid aA \\ B &\to AB \mid \varepsilon \\ C &\to ABC \mid c \end{aligned}$$

- Найдем аннулирующие символы: $Ann(G) = \{A, B, S\}$.

Пример

Построить ε-свободную грамматику, эквивалентную данной

$$
\begin{aligned}
S &\to AB \mid cC\\
A &\to \varepsilon \mid aA\\
B &\to AB \mid \varepsilon\\
C &\to ABC \mid c
\end{aligned}
$$

- Найдем аннулирующие символы: $Ann(G) = \{A, B, S\}$.
- Запишем измененные правила. Так как аксиома аннулирующий символ, но не встречается в правых частях, то не обязательно создавать новую аксиому, достаточно добавить правило $S \to \varepsilon$.

Грамматика G	Грамматика G'
$S \to AB$	$S \to AB \mid A \mid B$
$S \to cC$	$S \to cC$
$A \to aA$	$A \to aA \mid a$
$B \to AB$	$B \to AB \mid A$
$C \to ABC$	$C \to ABC \mid AC \mid BC \mid C$
$C \to c$	$C \to c$
	$S \to \varepsilon$