7. Преобразования КС-грамматик: устранение левой рекурсии, левая факторизация.

Устранение левой рекурсии

опр. Нетерминал A называют леворекурсивный, если допустим $A \to^* A\gamma$

опр. Непосредственной левой рекурсией назвыаются правила вида $A \to A\gamma$

Наличие левой рекурсии делает невозможным применение некоторых методов нисходящего анализа (например, метода рекурсивного спуска), так как они могут войти в бесконечный цикл.

Алгоритм устранения непосредственной левой рекурсии (nl)

Непосред леворекурсив, если $(A \to A\gamma) \in P$

Устранение непосредств левой рекурсии ($\overline{\text{nl}}$

$A \to A\alpha_1 | \dots | A\alpha_m | \beta_1 | \dots | \beta_k$

добавим новый нетерм A'

$A \to \beta_1 A' | \dots | \beta_k A'$

$A' \to \alpha_1 A' | \dots | \alpha_n A' | \lambda$

$A \Rightarrow \beta_j \alpha_i, \dots, \alpha_{i+1}$

$A' \to \alpha_1 A' | \dots | \alpha_n A' | \lambda$

$A \to A\alpha_1 | \dots | A_n \alpha_n | \beta_1 | \dots | \beta_n$

$\forall i = \overline{1,m} \quad \beta_i[1] \neq A$

$\Downarrow \quad \overline{nl}(P, A)$

$A \to \beta_1 A' | \dots | \beta_m A'$

$A' \to \alpha_1 A' | \dots | \alpha_n A' | \lambda$

Пример:

$E \to E + T | T$

$T \to T * F | F$

$F \to (E) | x$

$E \to TE'$

$E' \to +TE' | \lambda$

$T \to FT'$

$T' \to *FT' | \lambda$

Алгоритм

$G = (\Sigma, \Gamma, P, S)$ — λ-свободная, $\Gamma = \{A_1, \dots, A_n\}$

$\bar{\Gamma} = \Gamma$

$\bar{P} = P$

цикл по $i = \overline{1,n}$

{

  цикл по $j = \overline{1, i-1}$

  {

    $K_j = \{\beta \mid (A_j \to \beta) \in \bar{P}\}$

    если $\forall (A_i \to A_j \alpha) \in P$

      $\bar{P} = \bar{P} \setminus (A_i \to A_j \alpha) \cup \bigcup_{K_j} (A_i \to \beta \alpha)$

  }

  если $\exists (A_i \to A_i \alpha) \in P$

    $\bar{\Gamma} = \bar{\Gamma} \cup \{A_i'\}$

    пл $(\bar{P}, A_i)$

}

Док-во корректности:

После завершения алгоритма, $\forall (A_i \to A_j \alpha) \in \bar{P}$ $i<j$

индукция по i

Б.И. $i=1$ внутреннего цикла нет

если $(A_1 \to A_1\alpha) \in P$, то мы устраним непосредст рекурсию

Ш.И. Пусть после $k-1$ шагов выполняется П.И. после внутреннего цикла если $(A_i \to A_j \alpha) \in \overline{P}$, то $i \le j$ $\boxtimes$

Пример:

$S \to Aa \mid AB \mid B$ $\qquad$ $A_1 \to A_2 a \mid A_2 A_3 \mid A_3$

$A \to SB \mid ac$ $\qquad$ $A_2 \to A_1 A_3 \mid ac$

$B \to Ac \mid b$ $\qquad$ $A_3 \to A_2 c \mid b$

Для A_1 все Ok

Для A_2 : $k_1 = \{A_2 a, A_2 A_3, A_3\}$

$A_2 \to A_1 A_3$ заменили $\quad A_2 \to A_2 \underbrace{a A_3}_{\alpha_1}$

$A_2 \to A_2 \underbrace{A_3 A_3}_{\alpha_2}$

$A_2 \to \underbrace{A_3 A_3}_{\beta_1}$

$A_2 \to \underbrace{ac}_{\beta_2}$

Преобразовываем через пе: A_2

$A_2 \to A_3 A_3 A_2' \mid ac A_2'$

$A_2' \to a A_3 A_2' \mid A_3 A_3 A_2' \mid \lambda$

A_3: $\quad k_2 = \{A_3 A_3 A_2', ac A_2'\}$

$A_3 \to A_2 c \mid b$ заменили:

$$A_3 \to A_3 \overbrace{A_3 A_2' c}^{\alpha_1} \mid \overbrace{a c A_2' c}^{\beta_1} \mid \overbrace{b}^{\beta_2}$$

Преобразуем через пл:

$$A_3 \to a c A_2' c A_3' \mid b A_3'$$

$$A_3' \to A_3 A_2' c \mid \lambda$$

итого

$$A_1 \to A_2 a \mid A_2 A_3 \mid A_3$$

$$A_2 \to A_3 A_3 A_2' \mid a c A_2'$$

$$A_2' \to a A_3 A_2' \mid A_3 A_3 A_2' \mid \lambda$$

$$A_3 \to a c A_2' c A_3' \mid b A_3'$$

$$A_3' \to A_3 A_2' c \mid \lambda$$

Для устранения рекурсии используется метод динамического программирования, который последовательно переупорядочивает нетерминалы и заменяет правила таким образом, чтобы индекс левого нетерминала в правой части был строго больше индекса нетерминала в левой части.

Любая КС-грамматика имеет эквивалентную ей грамматику без левой рекурсии.

Левая факторизация

Левая факторизация — это преобразование грамматики, целью которого является устранение у нетерминалов альтернатив, имеющих общий непустой префикс.

Это необходимо для детерминированного выбора правила в процессе нисходящего анализа: когда несколько правил начинаются одинаково, анализатор не может решить, какое из них применить, не заглядывая далеко вперед во входную цепочку.

Левая факторизация

$A \to \alpha\beta_1 | \dots | \alpha\beta_m | \gamma_1 | \dots | \gamma_n \quad \alpha \neq \lambda$

Алгоритм: $\quad G = (\Sigma, \Gamma, P, S)$, $\bar{\Gamma} = \Gamma \quad \bar{P} = P$

цикл по всем $A \in \bar{\Gamma}$

1. $K = \{\alpha \mid (A \to \alpha\beta_1), (A \to \alpha\beta_2) \in \bar{P}, \alpha \neq \lambda\}$

пока $k \neq \emptyset$:

2. $\alpha' = \alpha \in k : |\alpha| = \max\limits_{k} \{|\alpha_i|\}$

3. $\bar{\Gamma} = \bar{\Gamma} \cup \{A'\}$

4. $\bar{P} = \bar{P} \setminus \{(A \to \alpha'\beta)\} \cup (A \to \alpha' A') \cup \{(A' \to \beta)\}$
(для всех таких β)

5 $k = k \setminus \{\alpha'\}$ перейти на 2

Пример:

$S \to Abc \mid AbB \mid Ac \mid ABB$

$A \to Bc \mid b$

$B \to ac$

$C \to cA$

$\{Ab, A\}$ общие префиксы

$S \to AbD \mid Ac \mid ABB$

$D \to c \mid B$ ← добавили

$S \to AE$

$E \to bD \mid c \mid BB$ ← добавили

$D \to c \mid B$