

1-消除左递归

原理

3.2.3 消除左递归

假设我们要为文法 3-4 构造一个预测分析器。下面两个产生式

$$\begin{aligned} E &\rightarrow E + T \\ E &\rightarrow T \end{aligned}$$

肯定会导致在 LL(1) 分析表中有双重定义的登记项，因为任何属于 $\text{FIRST}(T)$ 的单词同时也属于 $\text{FIRST}(E + T)$ 。问题发生的原因是 E 作为 E 的产生式的第一个右部符号出现，这种情况称为左递归 (left recursion)。具有左递归的文法不是 LL(1) 文法。

为了消除左递归，我们将利用右递归来重写产生式。为此需要引入一个新的非终结符 E' ，并将产生式重写为：

$$\begin{aligned} E &\rightarrow T E' \\ E' &\rightarrow + T E' \\ E' &\rightarrow \end{aligned}$$

这 3 个产生式推导出的（关于 T 和 $+$ 的）字符串集合与原来那两个产生式推导出的字符串集合相同，但是它没有左递归。

一般地，对于产生式 $X \rightarrow X\gamma$, $X \rightarrow \alpha$ ，其中 α 不以 X 开始，我们知道由它们可推导出形如 $\alpha\gamma^*$ 的字符串，即一个 α 其后跟随零至多个 γ 的字符串。因此我们可以利用右递归来重写正则表达式：

$$\left(\begin{array}{l} X \rightarrow X\gamma_1 \\ X \rightarrow X\gamma_2 \\ X \rightarrow \alpha_1 \\ X \rightarrow \alpha_2 \end{array} \right) \Rightarrow \left(\begin{array}{l} X \rightarrow \alpha_1 X' \\ X \rightarrow \alpha_2 X' \\ X' \rightarrow \gamma_1 X' \\ X' \rightarrow \gamma_2 X' \\ X' \rightarrow \end{array} \right)$$

对文法 3-4 应用这种转换，可以得到文法 3-7。

为构造预测分析器，我们首先要计算 nullable、FIRST 和 FOLLOW 集合（见表 3-1）。文法 3-7 的预测分析器如表 3-2 所示。

文法 3-7

$S \rightarrow E \$$	$T \rightarrow F T'$	
$E \rightarrow T E'$		$F \rightarrow \text{id}$
	$T' \rightarrow * F T'$	$F \rightarrow \text{num}$
$E' \rightarrow + T E'$	$T' \rightarrow / F T'$	$F \rightarrow (E)$
$E' \rightarrow - T E'$	$T' \rightarrow$	
$E' \rightarrow$		

表 3-1 文法 3-3 的 nullable、FIRST 和 FOLLOW

	nullable	FIRST	FOLLOW
S	no	(id num	
E	no	(id num) \$
E'	yes	+ -) \$
T	no	(id num) + - \$
T'	yes	* /) + - \$
F	no	(id num) * / + - \$

表 3-2 文法 3-7 的预测分析表。其中省略了 num、/和 - 对应的列，因为它们和表中的其他项是类似的

	+	*	id	()	\$
S			$S \rightarrow E \$$	$S \rightarrow E \$$		
E			$E \rightarrow T E'$	$E \rightarrow T E'$		
E'	$E' \rightarrow + T E'$				$E' \rightarrow$	$E' \rightarrow$
T			$T \rightarrow F T'$	$T \rightarrow F T'$		
T'	$T' \rightarrow$	$T' \rightarrow * F T'$			$T' \rightarrow$	$T' \rightarrow$
F			$F \rightarrow \text{id}$	$F \rightarrow (E)$		

从遍历+回溯的方法去匹配改为顺序迭代不回溯的方法去匹配后带来了速度提升

算法

直接消除左递归法：

$$P \rightarrow P\alpha_1 \mid P\alpha_2 \mid \dots \mid P\alpha_m \mid \beta_1 \mid \beta_2 \mid \dots \mid \beta_n$$

变换成：

$$P \rightarrow \beta_1 P' \mid \beta_2 P' \mid \dots \mid \beta_n P'$$
$$P' \rightarrow \alpha_1 P' \mid \alpha_2 P' \mid \dots \mid \alpha_m P' \mid \varepsilon$$

消除间接左递归法：

消除左递归的算法

1. 把文法G的所有非终结符按任一种顺序排列 P_1, P_2, \dots, P_n ；按此顺序执行：

2. FOR $i:=1$ TO n DO 把 P_i 的规则改造成 $P_i \rightarrow a... | P_{i+1}... | P_{i+2}... | \dots | P_{i+k}...$

BEGIN

FOR $j:=1$ TO $i-1$ DO

把形如 $P_i \rightarrow P_j \gamma$ 的规则改写成

$$P_i \rightarrow a... | P_{j+1}... | P_{j+2}... | \dots | P_{j+k}...$$

$P_i \rightarrow \delta_1 \gamma | \delta_2 \gamma | \dots | \delta_k \gamma$ ；(其中 $P_j \rightarrow \delta_1 | \delta_2 | \dots | \delta_k$ 是关于 P_j 的所有规则)

消除关于 P_i 规则的直接左递归性

$$P_i \rightarrow a... | P_i... | P_{i+1}... | \dots | P_{i+k}...$$

END

$$P_i \rightarrow a... | P_{i+1}... | P_{i+2}... | \dots | P_{i+k}...$$

3. 化简由2所得的文法，去除从开始符号出发永远无法到达的非终结符的产生规则。

练习题

1.

▶ 消除文法G(S)的左递归

$$S \rightarrow Qc | c$$

$$Q \rightarrow Rb | b$$

$$R \rightarrow Sa | a$$

答案

1.

- 1 $S \rightarrow c|bc|abc|Sabc$
- 2 其他的文法消除也类似可以算出来
- 3 $Q \rightarrow b|ab|cab|Qcab$
- 4 $R \rightarrow a|ca|bca|Rbca$