第三章：语法分析

文法等价变换

增加拓广产生式

- 定理一：对任一文法G_1都可以构造文法G_2，使得$L(G_1)=L(G_2)$，G_2的开始符不出现于任何产生式的右部。
- 证明：假设S是G_1的开始符，若S出现在某规则的右部，则引入新的语法符号S'，增加一条新产生式$S' \to S$即可，其中S'是新的开始符。

删除无用规则

- **定理二：** 对任一文法G_1都可以构造文法G_2，使得$L(G_1)=L(G_2)$，且G_2中的每个非终极符必出现在它的某个句型中。
- **证明：** 根据G_1，构造文法G_2的方法如下：

1. 令$\beta=\{S\}$，S是文法G_1的开始符
2. 递归扩充β
   $\beta=\beta \cup \{B|A\rightarrow xBy, B\in V_N, A\in\beta\}$。
3. 若$A\notin\beta$，则删除以A为左部的所有产生式。

删除无用规则

- 例子:

Z→aB|bc

B→bC|a

A→aB|a

C→c

D→AB|a

1.β={Z}

2.β={Z,B}

2.β={Z,B,C}

A,D无用

3.Z→aB|bc

B→bC|a

C→c

消除特型产生式

- 定义特型产生式：$A \to B$，且$B \in V_N$
- 定理三：对任一文法G_1，可以构造文法G_2，使得$L(G_1)=L(G_2)$，且在G_2中没有特型产生式。
- 证明：G_2的构造如下：

1. 对文法G_1中任意非终极符A，求集合 $\beta A=\{B \mid A \Rightarrow^+ B, B \in V_N\}$。
2. 若$B \in \beta A$，且$B \to \alpha$是文法G中的一个非特型产生式，则补充如下规则$A \to \alpha$。
3. 去掉文法G_1中的所有的特型产生式。
4. 去掉新的文法中的无用产生式。

消除特型产生式

- 例子

A→B|D|aB

B→C|b

C→c

D→B|d

βA={B,C,D}

βB={C}

βC={}

βD={B,C}

补充规则:

A→b|c|d

B→c

D→b|c

删去特型产生式:

A→b|c|d|aB

B→c|b

C→c

D→b|c|d

删去无用产生式:

A→b|c|d|aB

B→c|b

消除特型产生式

- 例子

$E \to E+T$

$E \to T$

$T \to T*F$

$T \to F$

$F \to (E)|i$

$\beta E=\{T,F\}$

$\beta T=\{F\}$

$\beta F=\{\}$

补充规则：

$E \to T*F|i|(E)$

$T \to (E)|i$

删去特型产生式：

$E \to E+T|T*F|i|(E)$

$T \to T*F|(E)|i$

$F \to (E)|i$

删去无用产生式：

$E \to E+T|T*F|i|(E)$

$T \to T*F|(E)|i$

$F \to (E)|i$

消除空产生式

- 空产生式: A→ε
- PASCAL语言例子:

<声明部分>→var <标号声明> <常量声明> <类型声明> <变量声明> <过程函数声明>

<变量声明>→<类型><标识符><变量声明>|ε

- 定理四: 对任一文法G_1，可构造文法G_2，使得$L(G_1)=L(G_2)$，且G_2中无空产生式。

消除空产生式

1. $\beta=\{A|A\to\varepsilon\}$，直接推出空的非终极符
2. $\beta=\beta\cup\{A|A \Rightarrow^{+}\alpha, \alpha\in\beta^{+}\}$，递归扩展隐式的 可以推出空的非终极符
3. 对于文法中任意产生式

   $A\to X_1\ldots X_{i-1}X_iX_{i+1}\ldots X_n(n\geq 2)$,若$X_i\in\beta$,

   补充规则 $A\to X_1X_2\ldots X_{i-1}X_{i+1}\ldots X_n$

   重复3，直到不产生新的产生式。
4. 删除所有形如$A\to\varepsilon$的产生式。

消除空产生式

- 例子

A→aBcD

B→b|ε

D→BB|d

β={B} 直接

β={B,D} 隐式

扩充:

由A→aBcD

有A→aBc

A→acD

由A→aBc

有A→ac

由D→BB 有D→B

删去空产生式

B→ε

A→aBcD|acD|aBc|ac

B→b

D→BB|B|d

提取公共前缀

- 公共前缀:A的不同产生式的右部具有相同的前缀
- $A \rightarrow \delta\beta_1 \mid \delta\beta_2 \mid \ldots \mid \delta\beta_n \mid \gamma_1 \mid \gamma_2 \mid \ldots \mid \gamma_m$ (其中每个γ不以δ开头)

则将这些规则写成:

$$A \rightarrow \delta A' \mid \gamma_1 \mid \gamma_2 \mid \ldots \mid \gamma_m$$

$$A' \rightarrow \beta_1 \mid \beta_2 \mid \ldots \mid \beta_n$$

提取公共前缀

- 例子:

A→aB|aC|cD

B→bB|bD|c

D→d

- 提取公共前缀

A→aA'

A'→B|C|D

B→bB'|c

B'→B|D

D→d

消除直接左递归

- 对于简单情形$A \rightarrow A\alpha|\beta$　即有$A \Rightarrow \beta\alpha^*$
  则转化　$A \rightarrow \beta A'$　$A' \rightarrow \alpha A'|\varepsilon$
- 对于一般情形
  $A \rightarrow A\alpha_1 \mid A\alpha_2 \mid \ldots \mid A\alpha_m \mid \beta_1 \mid \beta_2 \mid \ldots \mid \beta_n$
  则转化
  $A \rightarrow (\beta_1 \mid \beta_2 \mid \ldots \mid \beta_n)\ A'$
  $A' \rightarrow (\alpha_1 \mid \alpha_2 \mid \ldots \mid \alpha_m)\ A' \mid \varepsilon$

消除直接左递归

- 例子

$E \rightarrow E+T \mid T$

$T \rightarrow T*F \mid F$

$F \rightarrow (E) \mid i$

$A \rightarrow A\alpha \mid \beta$

$E \rightarrow E+T \mid T$

$\alpha = +T$

$\beta = T$

$E \rightarrow TE'$

$E' \rightarrow +TE' \mid \varepsilon$

$E \rightarrow TE'$

$E' \rightarrow +TE' \mid \varepsilon$

$T \rightarrow FT'$

$T' \rightarrow *FT' \mid \varepsilon$

$F \rightarrow (E) \mid i$

消除左递归

- 基本思想：

  类似于方程中的代入，把所有的形如 $A \rightarrow B\alpha$ 中的B作为左部的规则代入，然后消除文法中的直接左递归，最后进行化简，消去不可达产生式。

文法的一些表示方法

- 重复0次或者多次：{}

  A→a|aA 表示成A→a{a}

- 重复1次或者多次：[]

  A→a|aA 表示成A→[a]