

# 第5章 自底向上优先分析

# 自底向上优先分析

自底向上分析方法，也称移进-归约分析法

实现思想（是推导的逆过程）：

对输入符号串自左向右进行扫描，并将输入符逐个移入一个后进先出栈中，边移入边分析，一旦栈顶符号串形成某个句型的句柄或可归约串时（该句柄或可归约串对应某产生式的右部），就用该产生式的左部非终结符代替相应右部的文法符号串，称为一步归约。重复这一过程，直到归约到栈中只剩下文法的开始符号时，则分析成功。

- ❖ 自底向上的移进-归约过程是自顶向下的最右推导的逆过程，而最右推导称为规范推导，自左向右的归约过程也称规范归约。

例1: 文法

$$\left\{ \begin{array}{l} S \rightarrow aAcBe \\ A \rightarrow b \\ A \rightarrow Ab \\ B \rightarrow d \end{array} \right.$$

输入串 **abbcde#** 分析

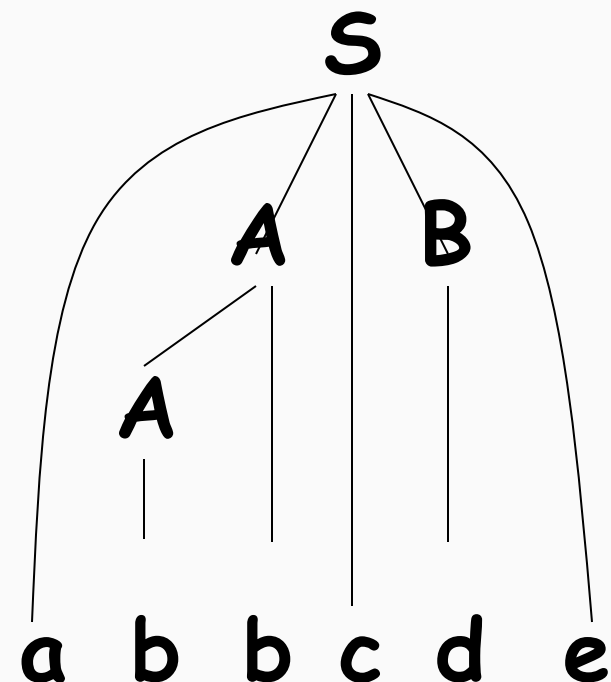
文法 $G[S]$ :

(1)  $S \rightarrow aAcBe$

(2)  $A \rightarrow b$

(3)  $A \rightarrow Ab$

(4)  $B \rightarrow d$



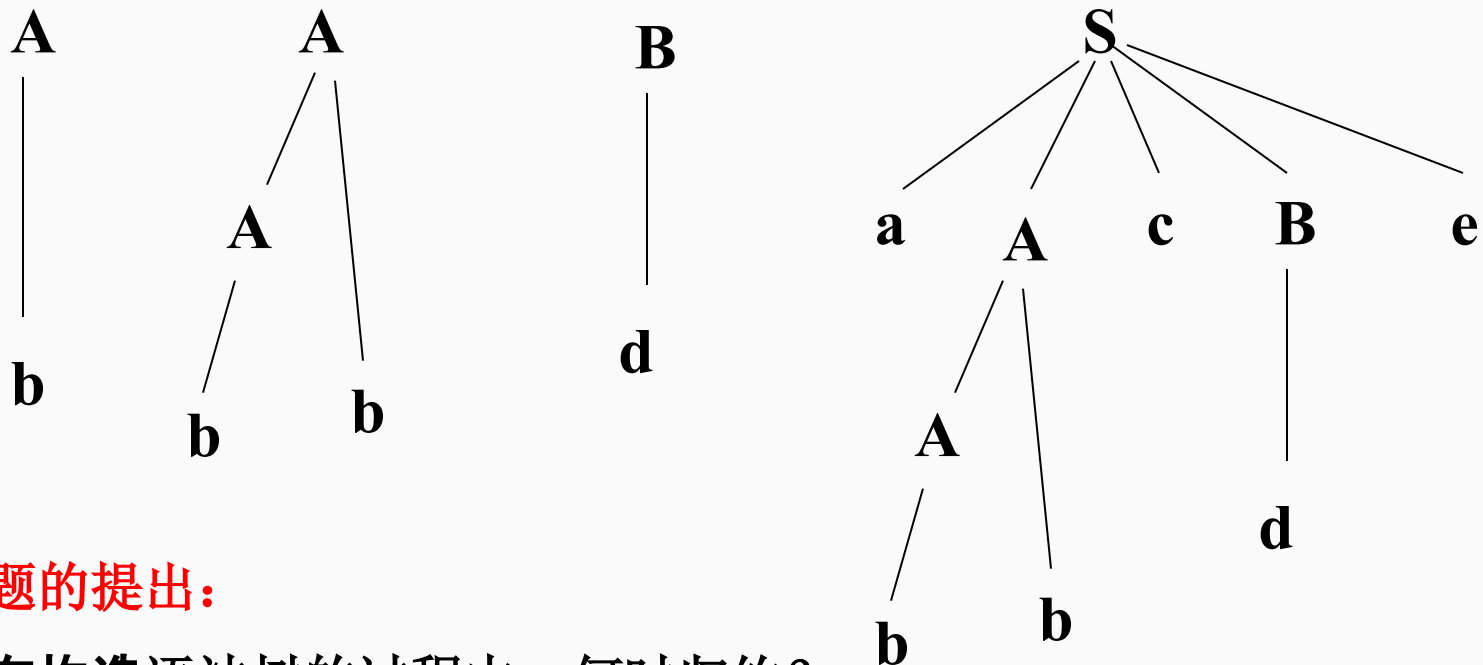
步骤	符号栈	输入符号串	动作
1)	#	abbcde#	移进
2)	#a	bbcde#	移进
3)	#ab	bcde#	归约( $A \rightarrow b$ )
4)	#aA	bcde#	移进
5)	#aAb	cde#	归约( $A \rightarrow Ab$ )
6)	#aA	cde#	移进
7)	#aAc	de#	移进
8)	#aAcd	e#	归约( $B \rightarrow d$ )
9)	#aAcB	e#	移进
10)	#aAcBe	#	归约
(1) $S \rightarrow aAcBe$	#	#	接受

对输入串abbcde#的移进-归约分析过程

分析符号串abbcde是否为 $G[S]$ 的句子?

$S \Rightarrow aAcBe \Rightarrow aAcde \Rightarrow aAbcde \Rightarrow abbcde$

上述的每一步归约可以构造一颗语法树：



问题的提出：

①在构造语法树的过程中，何时归约？

当可归约串或句柄出现在栈顶符号串中就可以归约。

②如何知道在栈顶符号串中已经形成可归约串或句柄？

通过自底向上分析算法中的优先关系来计算

## 自下而上分析的关键问题： 如何确定可归约串？

- ❖ 简单优先分析法：寻找句柄
- ❖ 算符优先分析法：寻找最左素短语

## 5.1 自底向上优先分析法概述

- ❖ 优先分析法又可分简单优先分析法和算符优先分析法
  - ✎ 简单优先分析法：按一定原则求出文法中所有符号（终结符和非终结符）的优先关系，按这种关系求出句柄。（规范归约——从左向右的归约）
  - ✎ 算符优先分析法：只考虑算符（终结符）之间的优先关系，不考虑非终结符之间的优先关系。按这种关系求出最左素短语。（不规范归约）
- ❖ 两种分析法比较
  - ✎ 简单优先分析法：准确规范，但分析效率低，实际使用价值不大；
  - ✎ 算符优先分析法：不规范，但分析速度快，适于实际使用。



句柄的定义：

令  $G$  是一文法， $S$  是文法的开始符号， $\alpha\beta\delta$  是文法  $G$  的一个句型。  
（为  $\alpha\beta\delta$  确定可归约串）如果有  $S \xRightarrow{*} \alpha A \delta$  且  $A \xRightarrow{+} \beta$ ， $\square$  称  $\beta$  是句型  $\alpha\beta\delta$  相对于非终结符  $A$  的短语。

若有  $A \Rightarrow \beta$ ， $\square$  称  $\beta$  是句型  $\alpha\beta\delta$  相对  $A \rightarrow \beta$  的直接短语。一个句型的最左直接短语称为该句型的句柄。

句柄是自底向上句法分析中当前时刻需要归约的符号串。如果能够自动计算出当前的句柄，则可执行自动句法分析。

## 5.2 简单优先分析法

### 1、优先关系的表示

$x = y$  表示X与Y的优先关系相等

$x \succ y$  表示X的优先性大于Y

$x \prec y$  表示X的优先性小于Y

对任意两个文法符号X、Y按其在句型中可能会出现相邻关系来确定优先关系：

## 确定优先关系的规则

- (1)  $X \equiv Y$  当且仅当  $G$  中存在产生式  $A \rightarrow \dots XY \dots$  (在语法树的同一层)
- (2)  $X \triangleleft Y$  当且仅当  $G$  中存在产生式  $A \rightarrow \dots XB \dots$ , 且  $B \Rightarrow Y \dots$  ( $Y$  在  $X$  的下一层)
- (3)  $X \triangleright Y$  当且仅当  $G$  中存在产生式  $A \rightarrow \dots BD \dots$ , 且  $B \Rightarrow \dots X$  和  $D \overset{*}{\Rightarrow} Y \dots$  ( $X$  在  $Y$  的下一层或  $X$  比  $Y$  先归约——规范归约/最左归约)

例：有文法G[S]:  $S \rightarrow bAb$

$$A \rightarrow ( B \mid a$$
$$B \rightarrow Aa )$$

解：文法符号优先关系推导如下：

(1) 求 $\stackrel{<}{\sim}$ 关系：

由 $S \rightarrow bAb$  ,  $A \rightarrow ( B$  ,  $B \rightarrow Aa$  )

$b \stackrel{<}{\sim} A$  ,  $A \stackrel{<}{\sim} b$  ,  $( \stackrel{<}{\sim} B$  ,  $A \stackrel{<}{\sim} a$  ,  $a \stackrel{<}{\sim} )$

(2) 求 $\prec$ 关系:

G[S]:  $S \rightarrow bAb$   
 $A \rightarrow ( B \mid a$   
 $B \rightarrow Aa )$

由 $S \rightarrow bAb$  且 $A \xRightarrow{+} ( B$  可得  $b \prec ($

$A \xRightarrow{+} a$  可得  $b \prec a$

由 $A \rightarrow ( B$  且 $B \xRightarrow{+} ( B \dots$  可得  $( \prec ($

$B \xRightarrow{+} aa \dots$  可得  $( \prec a$

$B \xRightarrow{+} Aa )$  可得  $( \prec A$

(3) 求  $\succ$  关系:

G[S]:  $S \rightarrow bAb$   
 $A \rightarrow ( B \mid a$   
 $B \rightarrow Aa )$

由  $S \rightarrow bAb$ , 且  $A \xRightarrow{+} \dots$  ) 可得  $) \succ b$

$A \xRightarrow{+} \dots B$  可得  $B \succ b$

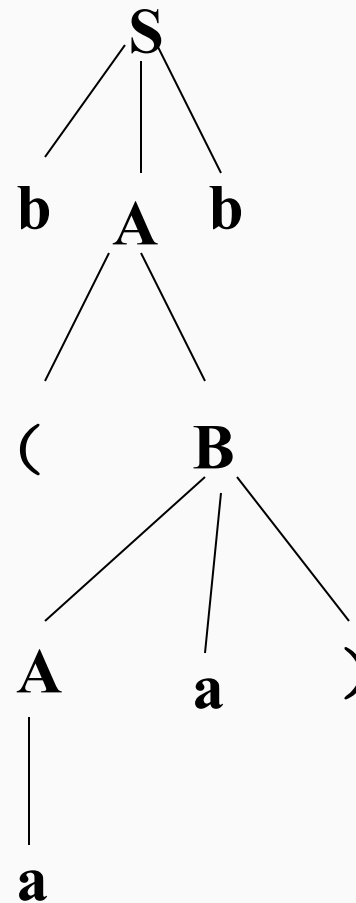
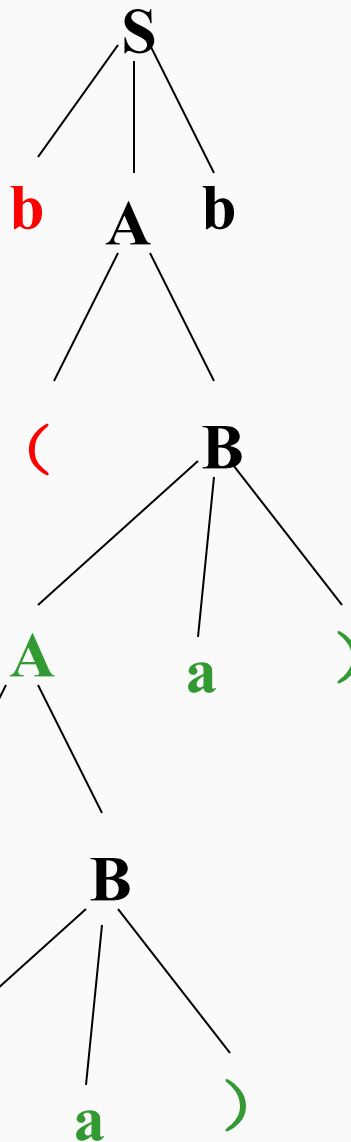
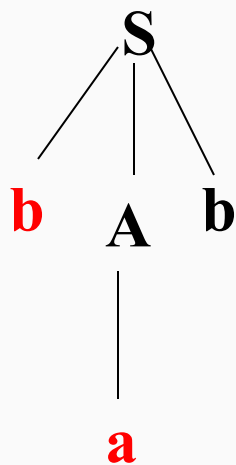
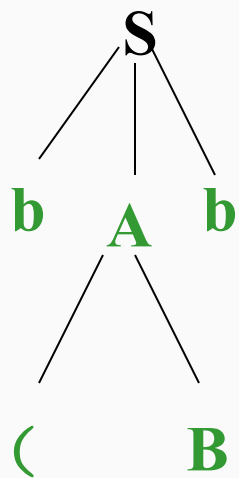
$A \xRightarrow{+} a$  可得  $a \succ b$

由  $B \rightarrow Aa$  ), 且  $A \xRightarrow{+} \dots$  ) 可得  $) \succ a$

$A \xRightarrow{+} a$  可得  $a \succ a$

$A \xRightarrow{+} \dots B$  可得  $B \succ a$

语法树结构如下：



G[S]:  $S \rightarrow bAb$   
 $A \rightarrow ( B \mid a$   
 $B \rightarrow Aa )$

从上图可以看出：

(1) 当 $b($ 、 $ba$ 出现在某一句型中，则 $($ 和 $a$ 在句柄中， $b$ 不在句柄中，则 $($ 和 $a$ 必须先归约。因此 $b \prec ($ ， $b \prec a$ 。

(2) 同样可以看出，当 $(($ ， $(a$ ， $(A$ 出现在某一句型中，右边的 $($ ， $a$ ， $A$ 出现在句柄中，左边的 $($ 不在句柄中。因此，左边 $( \prec$ 右边的 $($ ，左边 $( \prec a$ ，左边 $( \prec b$ 。

(3) 当 $ab$ 、 $aa$ 出现在某一句型中，左边 $a$ 在句柄中，则右边的 $a$ 、 $b$ 不可能在句柄中。因此有左边 $a \succ b$ ，左边 $a \succ a$ 。



- ❖ 因此，可以根据文法符号之间的优先关系确定句柄。
- ❖ 在规范归约（最左归约）过程中，出现在栈顶的优先级相同的连续符号串就是句柄。

## 优先关系矩阵表示法:

所有符号>#

	S	b	A	(	B	a	)	#
S								>
b			=	<		<		>
A		=				=		
(			<	<	=	<		
B		>				>		
a		>				>	=	
)		>				>		
#	<	<						=

## 如何确定优先关系?

文法 $G[S]$ :

(1)  $S \rightarrow bAb$

(2)  $A \rightarrow (B|a$

(3)  $B \rightarrow Aa)$

1.求=关系:

由(1):  $b=A \quad A=b$

由(2):  $(=B$

由(3):  $A=a \quad a=)$

2.求<关系:

由(1)(2):  $b<( \quad b<a$

由(2)(3):  $(<A \quad (<( \quad (<a$

3.求>关系:

由(1):  $B>b \quad a>b \quad )>b$

由(3):  $B>a \quad a>a \quad )>a$

4.#<所有符号, 所有符号>#

简单优先文法的定义：

- (1) 在文法符号集中，任意两个符号之间最多只有一种优先关系；
- (2) 在文法中任意两个产生式没有相同的右部。

# 简单优先分析法的操作步骤

首先根据已知优先文法构造优先关系矩阵，设置符号栈S，算法步骤如下：

- (1) 将输入符号 $a_1a_2\dots a_n\#$ 依次逐个存入符号栈S中，直到遇到栈顶符号 $a_i$ 的优先性大于待输入符号 $a_j$ 时为止。
- (2) 栈顶当前符号 $a_i$ 为句柄尾，由此向左在栈中找句柄的头符号 $a_k$ ，即找到 $a_{k-1} < a_k$ 为止。
- (3) 由句柄 $a_k\dots a_i$ 在文法中的产生式中查找右部为 $a_k\dots a_i$ 的产生式，若找到则用相应左部代替句柄，若找不到则为出错，这时可断定输入串不是该文法的句子。
- (4) 重复上述(1), (2), (3)步骤直到归约完输入符号串，栈中只剩文法的开始符号为止。

文法G[S]:

(1)  $S \rightarrow bAb$

(2)  $A \rightarrow (B|a$

(3)  $B \rightarrow Aa)$

步骤	符号栈	输入符号串	动作
1)	#	b(aa)b#	#<b,移进
2)	#b	(aa)b#	b<(,移进
3)	#b(	aa)b#	(<a,移进
4)	#b(a	a)b#	a>a,归约 $A \rightarrow a$
5)	#b(A	a)b#	$A=a$ ,移进
6)	#b(Aa	)b#	a=),移进
7)	#b(Aa)	b#	)>b,归约 $B \rightarrow Aa)$
8)	#b(B	b#	B>b,归约 $A \rightarrow (B$
9)	#bA	b#	$A=b$ ,移进
10)	#bAb	#	b>#,归约 $S \rightarrow bAb$
11)	#S	#	接受

简单优先关系矩阵

对输入串b(aa)#的简单优先分析过程

## 5.3 算符优先分析

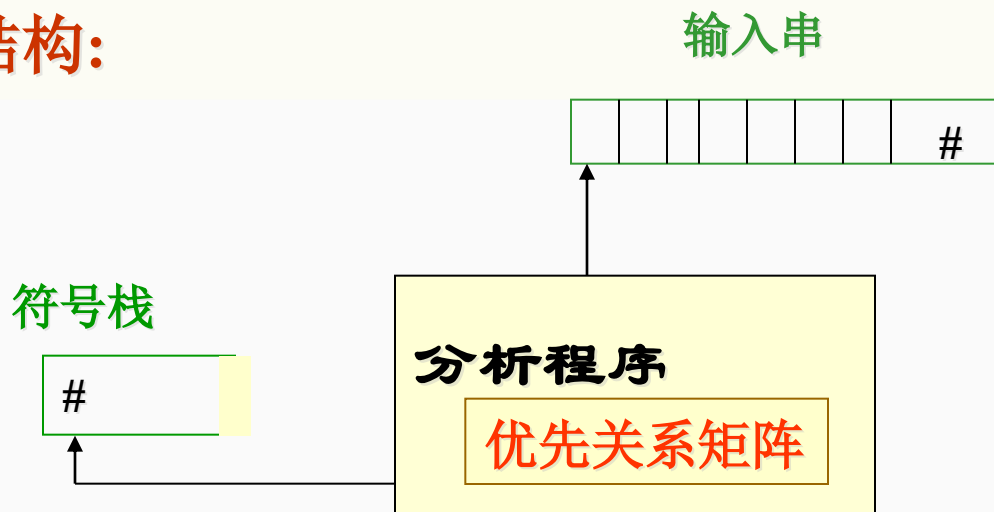
- 1) 这是一种经典的自底向上分析法，简单直观，并被广泛使用，开始主要是对表达式的分析，现在已不限于此。可以用于一大类上下无关的文法。
- 2) 称为算符优先分析是因为这种方法是仿效算术式的四则运算而建立起来的，作算术式的四则运算时，为了保证计算结果和过程的唯一性，规定了一个统一的四则运算法则，规定运算符之间的优先关系。
  1. 乘除的优先大于加减
  2. 同优先级的运算符左大于右
  3. 括号内的优先级大于括号外

于是：  $4+8-6/2*3$  运算过程和结果唯一。

### 3) 算符优先分析的特点:

仿效四则运算过程，预先规定相邻终结符之间的优先关系，然后利用这种优先关系来确定句型的可归约串，并进行归约。

### 4) 分析器结构:





# (1) 直观算符优先分析法

二义性文法 $G[E]$ :  $E \rightarrow E + E | E - E | E * E | E / E | E \uparrow E | (E) | i$

- 1)  $i$  的优先级最高
- 2)  $\uparrow$  先次于  $i$ ，右结合
- 3)  $*$  和  $/$  优先级次之，左结合
- 4)  $+$  和  $-$  优先级最低，左结合
- 5) 括号 ‘(’, ‘)’ 的优先级大于括号外的运算符，小于括号内的运算符，内括号的优先性大于外括号
- 6)  $\#$  的优先性低于与其相邻的算符

	+	-	*	/	$\uparrow$	(	)	$i$	#
+	>	>	<	<	<	<	>	<	>
-	>	>	<	<	<	<	>	<	>
*	>	>	>	>	<	<	>	<	>
/	>	>	>	>	<	<	>	<	>
$\uparrow$	>	>	>	>	<	<	>	<	>
(	<	<	<	<	<	<	=	<	
)	>	>	>	>	>		>		>
$i$	>	>	>	>	>		>		>
#	<	<	<	<	<	<		<	=

算符优先关系表

- ❖ 所给表达式文法虽然是二义性的，但我们人为直观地给出运算符之间的优先关系且这种优先关系是唯一的。
- ❖ 有了算符优先关系表，对一般表达式的输入串的计算结果就是唯一的。
- ❖ 类似地，通过定义终结符之间的优先关系，对一个输入串的归约过程就是唯一的。

## (2) 算符优先文法的定义

### 【算符文法定义】

设有一文法G，如果G中**没有**形如 $A \rightarrow \dots BC \dots$ 的产生式，其中B和C为非终结符，则称G为**算符文法** (或称**OG文法**)。

【要点】任何一个产生式中都不包含两个非终结符相邻的情况，就是是算符文法。或：两个非终结符之间一定通过1个或多个终结符相连。

**性质1：**在算符文法中任何句型都不包含两个相邻的非终结符。

**性质2：**如果 $Ab$ 或 $(bA)$ 出现在算符文法的句型 $\gamma$ 中，其中 $A \in V_N$   $b \in V_T$ ，则 $\gamma$ 中任何含 $b$ 的短语必含有 $A$ 。

（含 $b$ 的短语必含 $A$ ，含 $A$ 的短语不一定含 $b$ ）

## 【算符优先关系的定义】

设 $G$ 是一个不含 $\varepsilon$ 产生式的算符文法， $a$ 和 $b$ 是任意两个终结符， $A, B, C$ 是非终结符，算符优先关系如下：

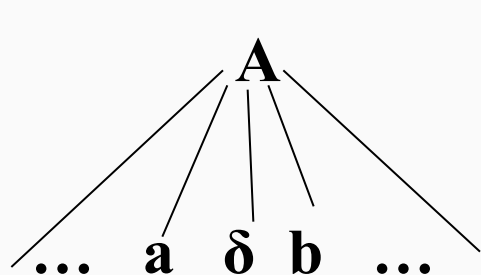
(1)  $a = b$  当且仅当 $G$ 中含有形如 $A \rightarrow \dots ab \dots$ 或 $A \rightarrow \dots aBb \dots$ 的产生式；

(2)  $a < b$  当且仅当 $G$ 中含有形如 $A \rightarrow \dots aB \dots$ 的产生式，且 $B \Rightarrow b \dots$ 或 $B \Rightarrow Cb \dots$ ；

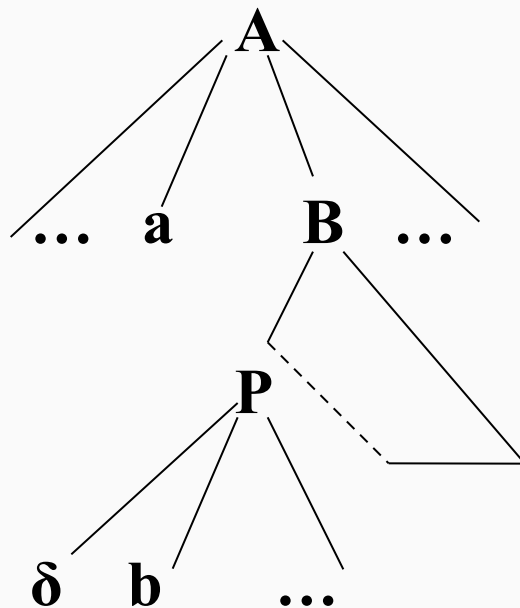
(3)  $a > b$  当且仅当 $G$ 中含有形如 $A \rightarrow \dots Bb \dots$ 的产生式，且 $B \Rightarrow \dots a$ 或 $B \Rightarrow \dots aC$ 。

与简单优先关系区别：区分终结符与非终结符，非终结符忽略不计

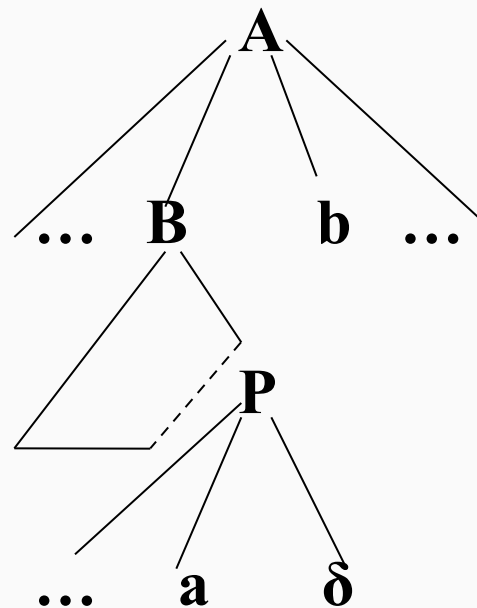
由语法树结构决定优先级:



$a=b$   
 $\delta$ 为 $\epsilon$ 或B



$a < b$   
 $\delta$ 为 $\epsilon$ 或C



$a > b$   
 $\delta$ 为 $\epsilon$ 或C

## 【算符优先文法的定义】

设有一不含 $\epsilon$ □生式的算符文法G，如果对任意两个终结符a，b之间至多只有=，<和>三种关系的一种成立，则称G是一个算符优先文法(也称OPG文法)。

即 $a = b$ ， $a < b$ ， $a > b$ 只有一种成立，但允许 $a < b$ ， $b < a$ 同时存在。

例：  $E \rightarrow E + E \mid E * E \mid (E) \mid i$       证明不是OPG文法。

因为：  $E \rightarrow E + E$  ,     $E \stackrel{+}{\Rightarrow} E * E$       则有  $+$   $<$   $*$

又因为：  $E \rightarrow E * E$ ,  $E \stackrel{+}{\Rightarrow} E + E$       则有  $+$   $>$   $*$

所以不是算符优先文法。



### (3) 算符优先关系表的构造

用表格形式来表示各终结符号的优先关系，这种表称为优先关系表。

构造优先关系表的方法：①按照定义来构造

②按关系图来构造

✓构造步骤：（根据算符优先关系的定义）

- 定义两个集合：firstVT集合lastVT集合。

$$\text{firstVT}(B) = \{b | B \xRightarrow{+} b... \text{ 或 } B \xRightarrow{+} Cb...\}$$

$$\text{lastVT}(B) = \{a | B \xRightarrow{+} ...a \text{ 或 } B \xRightarrow{+} ...aC\}$$

## 三种优先关系的计算:

a)  $=$  关系:

$A \rightarrow \dots ab \dots$

$A \rightarrow \dots aBb \dots$

} 则  $a = b$

b)  $<$  关系:

对于每个非终结符B的  
 $\text{firstVT}(B)$ 有形如 $A \rightarrow \dots aB \dots$ 中, 对  
每一个  $b \in \text{firstVT}(B)$ 。

} 则  $a < b$

c)  $>$ 关系:

每个非终结符 $B$ 的 $\text{lastVT}(B)$ 有  
形如 $A \rightarrow \dots Bb \dots$ 中, 对每一个  
 $a \in \text{lastVT}(B)$  } 则  $a > b$

例: 为文法

$E' \rightarrow \#E\#$

$T \rightarrow F$

$E \rightarrow E+T$

$F \rightarrow P \uparrow F | P$

$E \rightarrow T$

$P \rightarrow (E)$

$T \rightarrow T * F$

$P \rightarrow i$

构造算符优先关系表

## 根据定义构造FIRSTVT(A)的算法

1)若有产生式 $A \rightarrow a...$ 或 $A \rightarrow Ba...$ ，则 $a \in \text{FIRSTVT}(A)$ ，其中A为非终结符，a为终结符。

2)若有产生式 $A \rightarrow B...$ ，则 $\text{FIRSTVT}(B) \subset \text{FIRSTVT}(A)$ ，其中，A、B为非终结符，a为终结符。

## 根据定义构造LASTVT(A)的算法

1)若有产生式 $A \rightarrow \dots a$ 或 $A \rightarrow \dots aB$ ，则 $a \in \text{LASTVT}(A)$ ，其中A为非终结符，a为终结符。

2)若有产生式 $A \rightarrow \dots B$ ，则 $\text{LASTVT}(B) \subset \text{LASTVT}(A)$ ，其中，A、B为非终结符，a为终结符。

# 关系图法构造FIRSTVT(A)的方法

- 1) 图中结点为每个非终结符的FIRSTVT集或终结符。
- 2) 若有产生式 $A \rightarrow a...$ 或 $A \rightarrow Ba...$ ，其中A为非终结符，a为终结符。则构造由FIRSTVT(A)结点到终结符结点a用箭弧连接的图形。
- 3) 若有产生式 $A \rightarrow B...$ ，其中，A、B为非终结符，a为终结符。则构造由FIRSTVT(A)结点到FIRSTVT(B)结点用箭弧连接。
- 4) 若某一非终结符A的FIRSTVT(A)结点经箭弧有路径能够到达某终结符结点a，则 $a \in \text{FIRSTVT}(A)$ 。

## 关系图法构造LASTVT(A)的方法

- 1) 图中结点为每个非终结符的LASTVT集或终结符。
- 2) 若有产生式 $A \rightarrow \dots a$ 或 $A \rightarrow \dots aB$ ，其中A为非终结符，a为终结符。则构造由LASTVT(A)结点到终结符结点a用箭弧连接的图形。
- 3) 若有产生式 $A \rightarrow \dots B$ ，其中，A、B为非终结符，a为终结符。则构造由LASTVT(A)结点到LASTVT(B)结点用箭弧连接。
- 4) 若某一非终结符A的LASTVT(A)结点经箭弧有路径能够到达某终结符结点a，则 $a \in \text{LASTVT}(A)$ 。

## 构造优先关系矩阵的算法

```
FOR 每条规则  $U \rightarrow x_1 x_2 \dots x_n$  DO
  FOR  $i:=1$  TO  $n-1$  DO
    BEGIN
      IF  $x_i$  和  $x_{i+1}$  均为终结符, THEN 置  $x_i = x_{i+1}$ 
      IF  $i \leq n-2$ , 且  $x_i$  和  $x_{i+2}$  都为终结符号但
         $x_{i+1}$  为非终结符号 THEN 置  $x_i = x_{i+2}$ 
      IF  $x_i$  为终结符号  $x_{i+1}$  为非终结符号 THEN
        FOR FIRSTVT( $x_{i+1}$ ) 中的每个  $b$  DO
          置  $x_i < b$ 
      IF  $x_i$  为非终结符号  $x_{i+1}$  为终结符号 THEN
        FOR LASTVT( $x_i$ ) 中的每个  $a$  DO
          置  $a > x_{i+1}$ 
    END
  END
END
```



文法G[E]:

(0)  $E' \rightarrow \#E\#$

(1)  $E \rightarrow E+T$

(2)  $E \rightarrow T$

(3)  $T \rightarrow T * F$

(4)  $T \rightarrow F$

(5)  $F \rightarrow P \uparrow F | P$

(6)  $P \rightarrow (E)$

(7)  $P \rightarrow i$

$FIRSTVT(E') = \{\#\}$

$FIRSTVT(E) = \{+, *, \uparrow, (, i\}$

$FIRSTVT(T) = \{*, \uparrow, (, i\}$

$FIRSTVT(F) = \{\uparrow, (, i\}$

$FIRSTVT(P) = \{(, i\}$

$LASTVT(E') = \{\#\}$

$LASTVT(E) = \{+, *, \uparrow, ), i\}$

$LASTVT(T) = \{*, \uparrow, ), i\}$

$LASTVT(F) = \{\uparrow, ), i\}$

$LASTVT(P) = \{), i\}$

## 1) '='关系

由产生式(0)和(6),得

$\# = \#$ ,  $( = )$

## 2) '<'关系

找形如:  $A \rightarrow \dots aB \dots$  的产生式

$\#E$ : 则  $\# < FIRSTVT(E)$

$+T$ : 则  $+ < FIRSTVT(T)$

$*F$ : 则  $* < FIRSTVT(F)$

$\uparrow F$ : 则  $\uparrow < FIRSTVT(F)$

$(E$ : 则  $( < FIRSTVT(E)$

## 3) '>'关系

找形如:  $A \rightarrow \dots Bb \dots$  的产生式

$E\#$ , 则  $LASTVT(E) > \#$

$E+$ , 则  $LASTVT(E) > +$

$T*$ , 则  $LASTVT(T) > *$

$P\uparrow$ , 则  $LASTVT(P) > \uparrow$

$E)$ , 则  $LASTVT(E) > )$

	+	*	↑	(	)	i	#
+	>	<	<	<	>	<	>
*	>	>	<	<	>	<	>
↑	>	>	<	<	>	<	>
(	<	<	<	<	=	<	
)	>	>	>		>		>
i	>	>	>		>		>
#	<	<	<	<		<	=

## (4) 算符优先分析算法

有了算符优先分析表，就可以对任意给定的符号串进行归约分析，进而判断输入串是否为该文法的句子。

算符文法的任何一个句型为如下形式：

$$\#N_1a_1N_2a_2\ldots N_na_nN_{n+1}\#$$

其中 $N_i(1 \leq i \leq n+1)$ 为非终结符或空， $a_i(1 \leq i \leq n)$ 为终结符。

若有句型 $\dots N_i a_i \dots N_j a_j N_{j+1} \dots$ ，当 $a_i \dots N_j a_j$ 属于句柄，则 $N_i$ 和 $N_{j+1}$ 也在句柄中。该句柄中终结符之间的关系为：

$$a_{i-1} < a_i$$

$$a_i = a_{i+1} = \dots = a_{j-1} = a_j$$

$$a_j > a_{j+1}$$

算符优先文法在归约过程中只考虑终结符号之间的优先关系确定句柄，而非终结符无关，只需知道当前句柄归约为某一非终结符，不必知道该非终结符的名字是什么，这样也就去掉了单非终结符的归约。

文法 $G[E]$ :  $E \rightarrow E + E \mid E - E \mid E * E \mid E / E \mid E \uparrow E \mid (E) \mid i$

步骤	符号栈	输入符号串	动作
1)	#	i+i*i#	#<i,移进
2)	#i	+i*i#	#<i>+,规约
3)	#E	+i*i#	#<+,移进
4)	#E+	i*i#	+<i,移进
5)	#E+i	*i#	+<i>*,规约
6)	#E+E	*i#	+<*,移进
7)	#E+E*	i#	*<i,移进
8)	#E+E*i	#	*<i>#,规约
9)	#E+E*E	#	+<*>#,规约
10)	#E+E	#	#<+>#,规约
11)	#E	#	接受

算符优先关系表

对输入串i+i\*i的算符优先分析过程

## 实例比较

- ❖ 算符优先归约（表5.8）
- ❖ 规范归约（表5.7）
- ❖ 算符优先规约去掉单非终结符的归约，并且在归约时不考虑非终结符的名字，因此得到的不是真正的语法树，而是语法树的框架。（比较图5.7与图5.6）
- ❖ 算符优先分析法的可归约串不是句柄而是最左素短语。
- ❖ “算符优先分析的关键是如何找最左素短语”

## 最左素短语

定义：设有文法 $G[S]$ ，其句型的素短语是一个**短语**，它**至少**包含一个**终结符**，并除自身外不包含其它素短语，最左边的素短语称最左素短语。

与句柄的区别：至少包含一个终结符。（从而去掉了单非终结符的归约）

例：文法 $G[E]: E \rightarrow E+T | T$

$$T \rightarrow T * F | F$$

$$F \rightarrow P \uparrow F | P$$

$$P \rightarrow (E) | i$$

句型 $\#T+T*F+i\#$ 的语法树如下：

根据语法树可知：

句型#T+T\*F+i#的短语有：

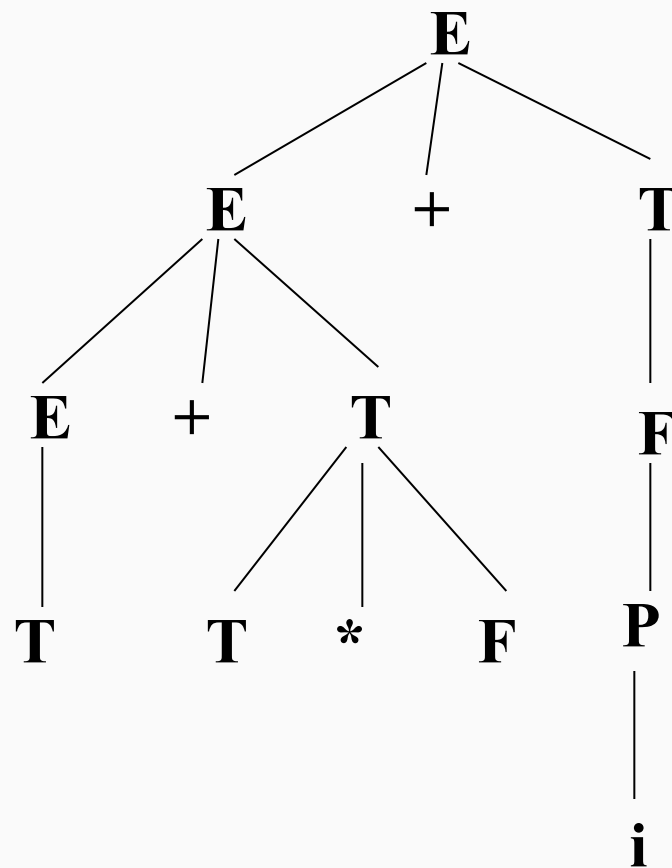
T — 相对非终结符E的短语

T\*F — 相对非终结符T的短语

T+T\*F — 相对非终结符E的短语

i — 相对非终结符P、F、T的短语

T+T\*F+i — 相对非终结符E的短语



根据素短语的定义可知：

$i$ 和 $T^*F$ 为素短语。

其中： $T+T^*F$  (含其他 $T^*F$ 素短语)和

$T+T^*F+i$  不是素短语。

$T^*F$ 为最左素短语。

$T$ 为句柄——最左直接短语



算符优先分析法的关键：

如何确定当前句型的最左素短语？

算符文法的任一句型形式为  $\#N_1a_1N_2a_2\ldots N_na_nN_{n+1}\#$

( $N_i \in V_N$  或  $N_i = \varepsilon$ ,  $a_i \in V_T$ )

**定理：** 一个OPG句型的最左素短语是满足下列条件的最左子串： $a_{j-1}N_ja_j\ldots N_ia_iN_{i+1}a_{i+1}$

其中  $a_{j-1} < a_j$

$a_j = a_{j+1}$ ,  $a_{j+1} = a_{j+2}, \ldots$ ,  $a_{i-2} = a_{i-1}$ ,  $a_{i-1} = a_i$

$a_i > a_{i+1}$

根据该定理,要找句型的最左素短语就是要找满足上述条件的最左子串.

## 算符优先分析法的实现:

详见P118 图5.8

基本部分是找句型的最左子串（最左素短语）并进行归约。

## (5) 优先函数:

算符优先关系表示法 { 矩阵表示法(空间大)  
优先函数法

### 1、优先函数的定义:

当  $a = b$ , 令  $f(a)=g(b)$

当  $a < b$ , 令  $f(a)<g(b)$

当  $a > b$ , 令  $f(a)>g(b)$

$f(a)$ 、 $g(b)$ 为优先函数,  
函数值用整数表示

## 2、由定义构造优先函数

### 构造规则

a) 对终结符  $a \in V_T$  (包括#号) 令  $f(a)=g(a)=1$  (初始化)

b) 如果  $a > b$ , 而  $f(a) \leq g(b)$ , 则令  $f(a)=g(b)+1$

c) 如果  $a < b$ , 而  $f(a) \geq g(b)$ , 则令  $g(b)=f(a)+1$

d) 如果  $a=b$ , 而  $f(a) \neq g(b)$  则令  
 $\min\{f(a), g(b)\} = \max\{f(a), g(b)\} + 1$

重复b)~d)过程, 直到**收敛**。若重复过程中有一个值  $> 2n$  ( $n$ 为终结符个数), 则该文法不存在算符优先函数。

例1：有优先表如下，构造优先函数。

	+	*	↑
+	>	<	<
*	>	>	<
↑	>	>	<

【解】 赋初值

	+	*	↑
f	1	1	1
g	1	1	1

最终结果为:

	+	*	↑
f	2	4	4
g	1	3	5

例2：已知优先关系表，构造优先函数。

	+	*	↑	i	(	)	#
+	>	<	<	<	<	>	>
*	>	>	<	<	<	>	>
↑	>	>	<	<	<	>	>
i	>	>	>			>	>
(	<	<	<	<	<	=	
)	>	>	>			>	>
#	<	<	<	<	<		=

# 构造过程

## (1) 初始化

	+	*	↑	i	(	)	#
f	1	1	1	1	1	1	1
g	1	1	1	1	1	1	1

## (2) 第一次迭代

	+	*	↑	i	(	)	#
f	2	4	4	6	1	6	1
g	2	3	5	5	5	1	1



### (3)第二次迭代

	+	*	↑	i	(	)	#
f	3	5	5	7	1	7	1
g	2	4	6	6	6	1	1

(4)第三次迭代同第二次，上图即为最终结果。

### 3、用关系图法构造优先函数

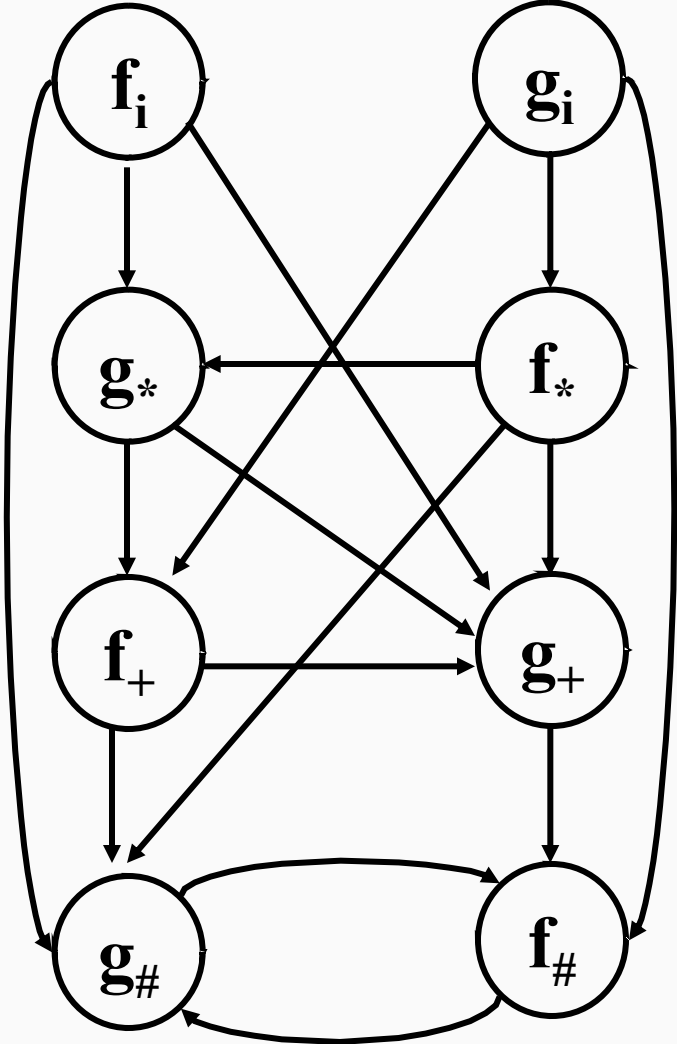
构造规则：

- a) 对所有终结符 $a \in V_T$  (包括#号)用有下脚标的 $f_a$ ,  $g_a$ 为结点名, 画出 $2n$ 个结点
- b) 如果 $a_i > a_j$ 或 $a_i = a_j$ , 则从 $f_{a_i}$ 到 $g_{a_j}$ 画一条箭弧  
如果 $a_i < a_j$ 或 $a_i = a_j$ , 则从 $g_{a_j}$ 到 $f_{a_i}$ 画一条箭弧
- c) 给每个结点赋一个数, 此数等于从该结点出发所能到达的结点 (包括该结点自身在内) 的个数。赋给结点 $f_{a_i}$ 的数, 就是 $f(a_i)$ 的值, 赋给 $g_{a_j}$ 的数, 就是 $g(a_j)$ 的值。
- d) 对构造出的优先函数, 按优先关系矩阵**检查一遍**是否满足优先关系的条件, 若不满足时, 则说明在关系图中存在3个或3个以上结点的回路, 不存在优先函数。

例3：已知优先关系表，构造优先函数。

	i	*	+	#
i		>	>	>
*	<	>	>	>
+	<	<	>	>
#	<	<	<	=

构造优先关系图如下：



	i	*	+	#
i		>	>	>
*	<	>	>	>
+	<	<	>	>
#	<	<	<	=

根据优先关系图求得的优先函数结果如下：

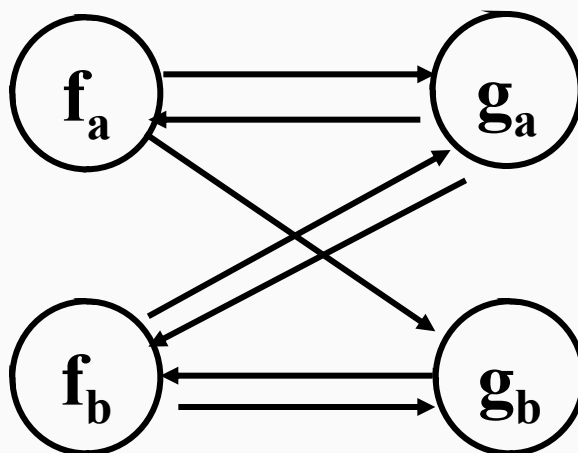
	i	*	+	#
f	6	6	4	2
g	7	5	3	2

其优先函数的优先关系与优先矩阵的优先关系是一致的。

例4： 已知优先关系表， 构造优先函数。

	a	b
a	=	>
b	=	=

构造优先关系图如下：



根据优先关系图求得的优先函数结果如下：

	a	b
a	4	4
b	4	4

其优先函数的优先关系与优先矩阵的优先关系相矛盾，所以不存在优先函数。



## 优先函数的缺点

- ❖ 在利用优先函数矩阵进行优先分析时，当两个终结符对无优先关系的情况时优先关系矩阵的相应元素为出错信息，而用优先函数进行优先分析时，对两个终结符对没有优先关系的情况不能区分，因而出错时不能准确地指出错误位置。

## (6) 算符优先分析法的局限性

由于算符优先分析法去掉了单非终结符之间的归约，尽管在分析过程中，当决定是否为句柄时采取一些检查措施，但仍**难完全避免把错误的句子得到正确的归约**。

通常一个使用语言的文法很难满足算符优先文法的条件，因而致使算符优先分析法**仅适用于表达式的语法分析**。

# 作业

❖ P121 练习1