# 编译原理

第一章 编译程序概述

第二章 PL/0编译程序的实现

第三章 文法和语言

第四章 词法分析

第五章 自顶向下语法分析方法

第六章 自底向上优先分析方法

第七章 LR分析方法

第八章 语法制导翻译和中间代码生成

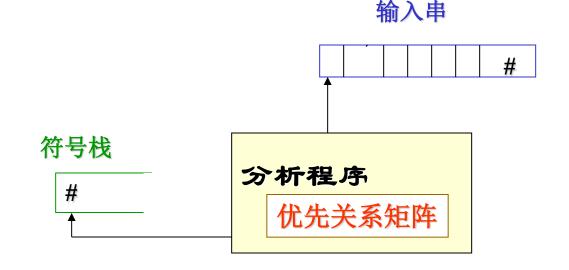
第九章 符号表

第一〇章 代码优化

第一一章 代码生成

## 自底向上优先分析

自底向上分析也称移进归约分析(是推导的逆过程)



简单优先文法:确定所有符号优先关系的规则

算符优先文法:确定算符优先关系的规则

firstVT(), lastVT()

## 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(最左直接短语)

• 算符优先分析法: 寻找最左素短语

## 二、简单优先分析法

1、优先关系的表示

x = y 表示x = 5y的优先关系相等

x > y 表示x的优先性大于y

x < · 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的相邻关系来确定优先关系:

# 确定优先关系的规则

- (1) X = Y 当且仅当G中存在产生式A  $\rightarrow ... XY ... (在语法树的同一层)$
- (2) X < Y 当且仅当G中存在产生式 $A \rightarrow ... XB..., 且<math>B \Rightarrow Y... (Y \in X \text{ 的下一层})$
- (3) X > . Y 当且仅当G中存在产生式A → ... BD...,且B ⇒ ... X和D ⇒ Y ... (X在 Y的 下一层或X比 Y先归约——规范归约/最左 归约)

## 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→…BC…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)

P.108 证明:反证法,如果b不和A一起规约。

## 【算符优先关系的定义】

设G是一个算符文法, 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→...ab... 或 A→...aBb...的产生式;
- (2) a < .b 当且仅当G中含有形如A  $\rightarrow$  ... aB ... 的产生式,且B  $\stackrel{+}{\Rightarrow}$  b ... 或B  $\stackrel{+}{\Rightarrow}$  Cb ...;
- (3) a >. b 当且仅当G中含有形如A→...Bb...的产生式,且B → ...a或B → ...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

## (3) 算符优先关系表

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。

firstVT(B)={
$$b|B \stackrel{+}{\Rightarrow} b...$$
 或  $B \stackrel{+}{\Rightarrow} Cb...$ }

lastVT(B)={a|B
$$\stackrel{+}{\Rightarrow}$$
...a 或 B $\stackrel{+}{\Rightarrow}$ ...aC}

## 三种优先关系的计算:

a) =. 关系:

$$A \rightarrow \dots ab \dots$$
 则  $a = b$ 

b) <.关系:

对于每个非终结符B的firstVT(B)有 形 如 A→...aB... 中 , 对 每 一 个 b∈firstVT(B)。

### c) >.关系:

每个非终结符B的lastVT(B)有 形如A→...Bb...中,对每一个 a∈lastVT(B)

> FOR 每条规则U→x<sub>1</sub>x<sub>2</sub>...x<sub>n</sub> DO FOR i:=1 TO n-1 DO BEGIN

> > IF xi和xi+1均为终结符,THEN 置 xi=.xi+1

IF i≤n-2, 且x<sub>i</sub>和x<sub>i+2</sub>都为终结符号但 x<sub>i+1</sub>为非终结符号 THEN 置 x<sub>i</sub>=.x<sub>i+2</sub>

IF x<sub>i</sub>为终结符号x<sub>i+1</sub>为非终结符号 THEN FOR FIRSTVT(x<sub>i+1</sub>)中的每个b DO 置x<sub>i</sub><.b

IF x<sub>i</sub>为非终结符号x<sub>i+1</sub>为终结符号 THEN FOR LASTVT(x<sub>i</sub>)中的每个a DO 置a>.x<sub>i+1</sub>

**END** 

## 构造FIRSTVT(U)的算法

1)若有规则U → b...或U → Vb...(存在U → b...或U → Vb...) 则b∈FIRSTVT(U)

2)若有规则 $U \rightarrow V...$ 且 $b \in FIRSTVT(V)$ ,则 $b \in FIRSTVT(U)$ 

说明:因为V⇒b...或V⇒Wb...,所以有U⇒V...⇒b...或 U⇒V... ⇒Wb...

#### 具体方法如下:

```
设一个栈S和一个二维布尔数组F
    F[U,b]=TRUE iff b \in FIRSTVT(U)
PROCEDURE INSERT(U,b)
       IF NOT F[U,b] THEN
         BEGIN
          F[U,b]:=TRUE
          把(U,b)推进S栈 /* b∈FIRSTVT(U) */
          END
BEGIN {main}
   FOR 每个非终结符号U和终结符b DO
         F[U,b]:=FALSE /*赋初值*/
   FOR 每个形如U→b...或U→Vb... 的规则 DO
        INSERT(U,b)
```

WHILE S栈非空 DO BEIGN

把S栈的顶项弹出,记为 (V,b)/\* b∈FIRSTVT(V)\*/FOR 每条形如U→V...的规则 DO INSTER(U,b); /\* b∈FIRSTVT(U)\*/

**END OF WHILE** 

**END** 

上述算法的工作结果是得到一个二维的布尔数组F,从F可以得到任何非终结符号U的FIRSTVT FIRSTVT(U)={b|F[U,b]=TRUE}

## 构造LASTVT(U)的算法

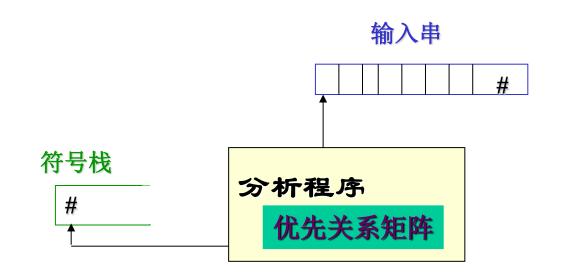
- 1.若有规则U→...a或U→=...aV,则a∈LASTVT(U)
- 2.若有规则U→...V,且a∈LASTVT(V)则a∈LASTVT(U)

```
设一个栈ST,和一个布尔数组B
PROCEDURE INSERT(U,a)
IF NOT B[U,a] THEN
BEGIN
B[U,a]→TRUE;把(U,a)推进ST栈;
END;
```

```
BEGIN
  FOR 每个非终结符号U和终结符号a
                            DO
          B[U,a]:=FALSE;
  FOR 每个形如U→...a或U→...aV的规则 DO
         INSERT (U,a);
  WHILE ST栈非空 DO
     BEGIN
       把ST栈的栈顶弹出,记为(V,a);
       FOR 每条形如U→…V的规则
                           DO
          INSERT(U,a);
     END OF WHILE;
END;
```

## 算符优先分析法的实现:

详见P117图6.8



当栈内终结符的优先级<=栈外的终结符的优先级时,移进;栈内终结符的优先级>栈外的终结符的优先级时,归约。表明找到了素短语的尾,再往前找其头,并进行规约。

## 算符优先分析句型的性质

- $\#N_1a_1N_2a_2....N_na_nN_{n+1}\#为句型$ 
  - $-N_i$ 为非终结符或空, $a_i$ 为终结符
  - 若a<sub>i</sub>...N<sub>i</sub>a<sub>i</sub>属于句柄,则N<sub>i</sub>和N<sub>i+1</sub>也在句柄中
    - $a_{i-1} < \cdot a_i$
    - $a_i = \cdot a_{i+1} \dots = \cdot a_i$
    - $a_j > a_{j+1}$

## · 如果aNb(或ab)出现在句型r中

- -a<·b,则在r中必含有b而不含a的短语存在
- -a >b,则在r中必含有a而不含b的短语存在
- $-a=\cdot b$ ,则在r中含有a的短语必含有b

# 实例比较

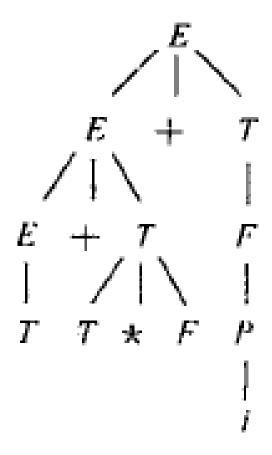
- 算符优先归约 (P115表6.8)
- 规范归约(P115表6.7)
- 前者(1)去掉单非终结符的归约, (2)并且 在归约时不考虑非终结符的名字, 因此得到的 不是真正的语法树, 而是语法树的框架。(比 较P116图6.7与图6.6)
- 算符优先分析法的可归约串不是句柄而是最左 素短语。
- P116"算符优先分析的关键是如何找最左素短语……"

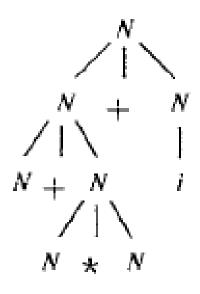
表 6.7 对输入串 i +i #的规范归约过程

步骤	栈	剩余输入串	句柄	归约用产生式
(1)	#	i+i #		
(2)	# i	+i #	i	$P \rightarrow i$
(3)	# P	+ <i>i</i> #	P	$F \rightarrow P$
(4)	# F	+i #	F	T→F
(5)	# T	+i #	T	$E \rightarrow T$
(6)	# E	+ i #		
(7)	# E+	i #		
(8)	# E+i	#	i	$P \rightarrow i$
(9)	$\sharp E+P$ .	#	P	$F \rightarrow P$
(10)	# E+F	#	F	$T \rightarrow F$
(11)	# E+T	#	E+T	$E \rightarrow E + T$
(12)	# E	#		接受

表 6.8 对输入串 i +i #的算符优先归约过程

步骤	栈	优先关系	当前符号	剩余输入串	移进或归约
(1)	#	<	i	+ i #	移进
(2)	# i	⊳	+	i #	归约
(3)	# F	< <	+	į #	移进
(4)	# F+	< <	i	#	移进
(5)	# F+i	⊳	#		归约
(6)	# F+F	⇒	#		归约
(7)	# F	=	#		接受





## 五、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→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:

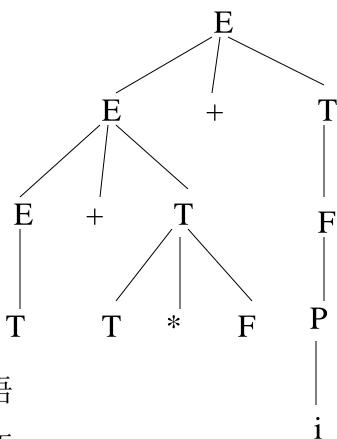
T— 相对非终结符E的短语

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 T

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



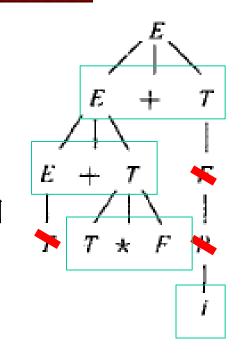
根据素短语的定义可知:

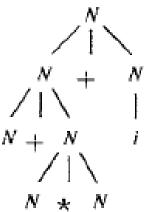
i和T\*F为素短语。

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  $\#N_1a_1N_2a_2...N_na_nN_{n+1}\#$   $(N_i \in V_N 或 N_i = \epsilon, a_i \in V_T)$ 

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: aj-1Njaj...NiaiNi+1ai+1

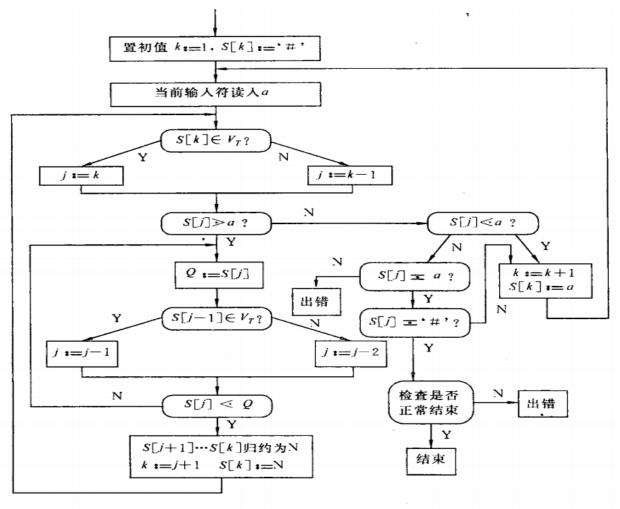
其中 a<sub>j-1</sub><.a<sub>j</sub>

$$a_{j}=.a_{j+1}, a_{j+1}=.a_{j+2},..., a_{i-2}=.a_{i-1}, a_{i-1}=.a_{i}$$

 $a_{i} > . a_{i+1}$ 

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:



26

## 六、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:

f(a)、g(b)为优先函数, 函数值用整数表示 2、优先函数的构造

构造规则

- a) 对终结符a∈ $V_T$ (包括#号)令f(a)=g(a)=1(初始化)
- b)如果a >. b,而 $f(a) \le g(b)$ ,则令f(a) = g(b) + 1
- c)如果a <.b,而 $f(a) \ge g(b)$ ,则令g(b) = f(a) + 1
- d) 如果 a=.b,而  $f(a)\neq g(b)$ 则令  $min\{f(a),g(b)\}=max\{f(a),g(b)\}$

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。

	+	*	<b>↑</b>
+	>.	<.	<.
*	>.	>.	<.
$\uparrow$	>.	>.	<.

【解】(1)赋初值

	+	*	<b></b>
f	1	1	1
g	1	1	1

### (2) 对a>.b关系有:

$$+>$$
. +  $f(+)=g(+)+1$ 

	+	*	<b></b>
f	2	1	1
g	1	1	1

	+	*	<b>↑</b>
+	>.	<.	<.
*	>.	>.	<.
<b>↑</b>	>.	>.	<.

	+	*	<b>↑</b>
f	2	2	1
g	1	1	1

$$\uparrow >$$
.+  $\uparrow >$ .\*

	+	*	<b></b>
f	2	2	2
g	1	1	1

## (3) 对a <.b关系有

$$+ <. * g(*)= f(+)+1$$
  
+ <.  $\uparrow g(\uparrow)=f(+)+1$ 

	+	*	<b>↑</b>
f	2	2	2
g	1	3	3

	+	*	<b>↑</b>
+	>.	<.	<.
*	>.	>.	<.
<b></b>	>.	>.	<.

	+	*	<b>↑</b>
f	2	2	2
g	1	3	3

(4) 对a =. b没有

重复过程(2)、(3)

## (4) 对a>.b关系

	+	*	$\uparrow$
f	2	2	2
g	1	3	3

$$* >.+ * >.* f(*)=g(*)+1$$

	+	*	<b></b>
f	2	4	2
g	1	3	3

	+	*	<b></b>
+	>.	<b>√</b>	<b>&lt;</b>
*	>.	>.	<.
$\uparrow$	>.	>.	<.

$$\uparrow$$
 >. +  $\uparrow$  >. \*  $f(\uparrow)=g(*)+1$ 

	+	*	<b>↑</b>
f	2	4	4
g	1	3	3

## (5) 对a <.b关系

	+	*	$\uparrow$
f	2	4	4
g	1	3	3

	+	*	<b>↑</b>
+	<b>^•</b>	<b>√</b>	<.
*	>.	>.	<.
<b>←</b>	>.	>.	<.

$$* <. \uparrow \qquad \uparrow <. \uparrow g(\uparrow) = f(\uparrow) + 1$$

	+	*	<b>↑</b>
f	2	4	4
g	1	3	5

## 重复以上过程得:

	+	*	<b></b>
f	2	4	4
g	1	3	5

### 所以存在优先函数为:

	+	*	$\uparrow$
f	2	4	4
g	1	3	5

	+	*	<b>↑</b>
+	>.	<.	<.
*	>.	>.	<.
$\uparrow$	>.	>.	<.

$$\stackrel{\text{\tiny $\omega$}}{=} a = .b, \Leftrightarrow f(a) = g(b)$$

$$\stackrel{\text{\tiny $\omega$}}{=} a \stackrel{\text{\tiny $c$}}{=} b$$
,  $\diamondsuit f(a) < g(b)$ 

$$\stackrel{\text{def}}{=} a > .b$$
,  $\diamondsuit f(a) > g(b)$ 

### 结果同上步, 因此收敛

## 例2: 己知优先关系表,构造优先函数。

	+	*	<b>↑</b>	i	(	)	#
+	>.	<.	<.	<.	<.	>.	>.
*	>.	>.	<.	<.	<.	>.	>.
$\uparrow$	>.	>.	<.	<.	<.	>.	>.
i	>.	>.	>.			>.	>.
(	<.	<.	<.	<.	<.	<u> </u>	
)	>.	>.	>.			>.	>.
#	<.	<.	<.	<.	<.		<u></u>

## 构造过程

## (1) 初始化

	+	*	<b>↑</b>	i	(	)	#
f	1	1	1	1	1	1	1
g	1	1	1	1	1	1	1

### (2) 对a>.b关系

	+	*	<b>↑</b>	i	(	)	#
f	2	1	1	1	1	1	1
g	1	1	1	1	1	1	1

	+	*	<b>↑</b>	i	(	)	#
f	2	2	1	1	1	1	1
g	1	1	1	1	1	1	1

	+	*	<b>↑</b>	i	(	)	#
f	2	2	2	1	1	1	1
g	1	1	1	1	1	1	1

$$i > + i > * i > . \uparrow i > . \downarrow$$

	+	*	<b>↑</b>	i	(	)	#
f	2	2	2	2	1	1	1
g	1	1	1	1	1	1	1

<del>&gt;.</del> †i		*	$\uparrow$	i	(	)	#
f	2	2	2	2	1	2	1
g	1	1	1	1	1	1	1

#### (3) 对a<.b关系

	+	*	<b>↑</b>	i	(	)	#
f	2	2	2	2	1	2	1
g	1	3	3	3	3	1	1

	+	*	<b>↑</b>	i	(	)	#
f	2	2	2	2	1	2	1
g	1	3	3	3	3	1	1

	+	*	<b>↑</b>	i	(	)	#
f	2	2	2	2	1	2	1
g	1	3	3	3	3	1	1

	+	*	<b>↑</b>	i	(	)	#
f	2	2	2	2	1	2	1
g	2	3	3	3	3	1	1

	+	*	<b></b>	i	(	)	#
f	2	2	2	2	1	2	1
g	2	3	3	3	3	1	1

$$(\leq)$$
 # $\leq$ #

	+	*	<b>↑</b>	i	(	)	#
f	2	2	2	2	1	2	1
g	2	3	3	3	3	1	1

#### 第二次重复以上过程(2,3,4步)

对于a>.b关系

	+	*	<b>↑</b>	i	(	)	#
f	3	4	4	4	1	4	1
g	2	3	3	3	3	1	1

#### 对于a <.b关系

	+	*	<b>↑</b>	i	(	)	#
f	3	4	4	4	1	4	1
g	2	4	5	5	5	1	1

#### 对于a≤b关系

	+	*	<b>↑</b>	i	(	)	#
f	3	4	4	4	1	4	1
g	2	4	5	5	5	1	1

第三次重复以上过程

对于a>.b关系

	+	*	<b>↑</b>	i	(	)	#
f	3	5	5	6	1	4	1
g	2	4	5	5	5	1	1

#### 对于a <.b关系

	+	*	<b>↑</b>	i	(	)	#
f	3	5	5	6	1	4	1
g	2	4	6	6	6	1	1

#### 对于a≤b关系

	+	*	<b>↑</b>	i	(	)	#
f	3	5	5	6	1	4	1
g	2	4	6	6	6	1	1

#### 第四次重复过程

对于a>.b关系

	+	*	<b>←</b>	i	(	)	#
f	3	5	5	7	1	7	1
g	2	4	6	6	6	1	1

对于a <.b关系

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

#### 对于a≤b关系

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

### 第五次重复过程

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

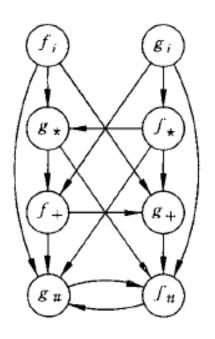
第五次结果同第四次,表示收敛了。 因此优先函数为:

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

## 关系图法构建优先函数

- (1) 任一终结符a, 建立两个节点f<sub>a</sub>和g<sub>a</sub>
- (2) a>.a|a=.a,  $f_a \rightarrow g_a$
- (3)  $a < .a \mid a = .a$ ,  $f_a \leftarrow g_a$
- (4) 优先函数按优先关系矩阵检查一遍,是 否一致,否则存在3个或以上节点的回路。

# 例题



	i	*	+	#				
i		>	⇒	⇒				
*	< .	>	>	⊳				
+	< <	<:	⊳	⊳				
#	<	< .	. <	☲				

表 6.11 优先函数关系表

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

表 6.12 优先关系矩阵

	и	ь
и	Ξ	⇒
b	≖	≖

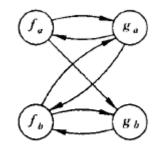


图 6.10 优先关系图

表 6.13 优先函数表

	а	b
f	4	4
g	4	4

# 局限性

$$S \rightarrow S; D \mid D$$
  
 $D \rightarrow D(T) \mid H$   
 $H \rightarrow a \mid (S)$   
 $T \rightarrow T + S \mid S$ 

	;	(	)	а	+	#
;	⊳	<	>	<	⊳	>
(	<:	⋖	=	< −	< -	
)	⇒	⇒	>		>	>
а	⇒	⊳	>		>	>
+	⋖	<	>	<	>	
#	< -	∢		<		=

- (a+a)#
- 被成功规约
- 却不是合法的句子

## 例题分析

**例题 1** 已知布尔表达式文法 G[B]为: B→BoT|T

T→TaF|F

F→nF|(B)|t|f

- G[B]是算符优先文法吗?
- 2. 若 G[B]是算符优先文法,请给出输入串 nto fat # 的分析过程。

# 作业

- 计算 例6.1 (P102) 文法的简单优先关系矩阵
- P122 练习 1、2