

编译原理

预测分析程序

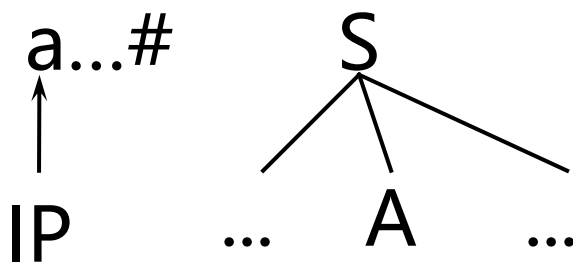
编译原理

回顾LL(1)分析法

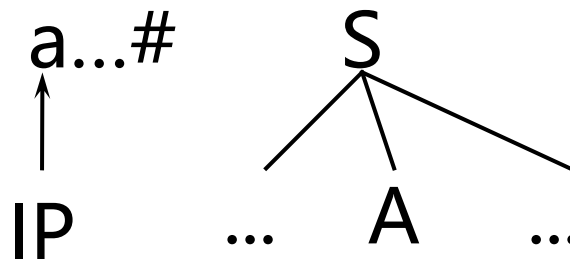
自上而下分析

► 基本思想

- 从文法的开始符号出发，向下推导，推出句子
- 针对输入串，试图用一切可能的办法，从文法开始符号(根结点)出发，自上而下地为输入串建立一棵语法树



LL(1)分析法



- ▶ 假设要用非终结符 A 进行匹配，面临的输入符号为 a ， A 的所有产生式为

$$A \rightarrow \alpha_1 \mid \alpha_2 \mid \dots \mid \alpha_n$$

1. 若 $a \in \text{FIRST}(\alpha_i)$ ，则指派 α_i 执行匹配任务；
2. 若 a 不属于任何一个候选首符集，则：
 - (1) 若 ϵ 属于某个 $\text{FIRST}(\alpha_i)$ 且 $a \in \text{FOLLOW}(A)$ ，则让 A 与 ϵ 自动匹配。
 - (2) 否则， a 的出现是一种语法错误。

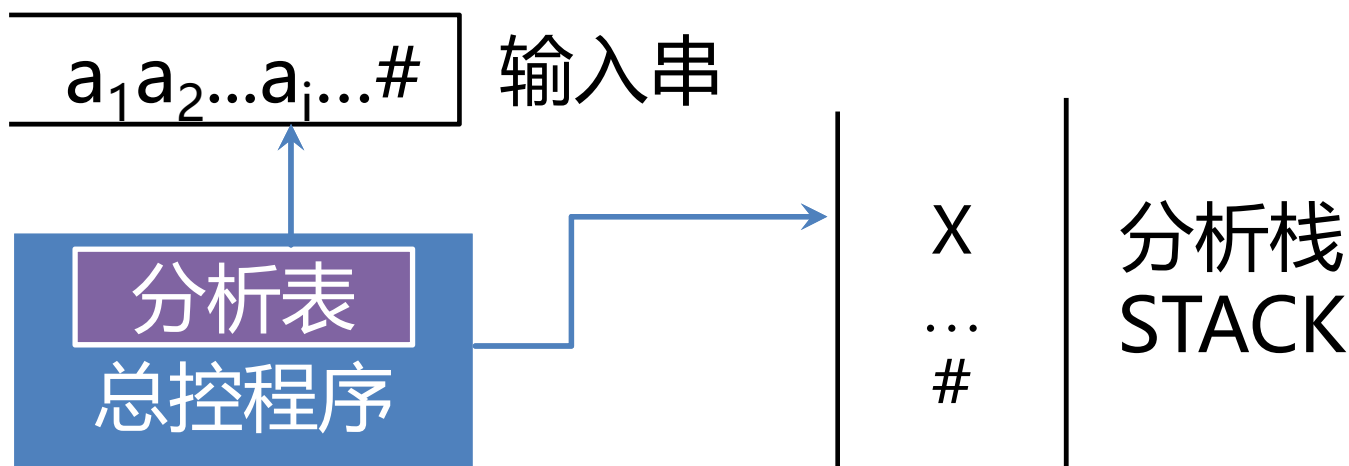
编译原理

预测分析程序的工作原理

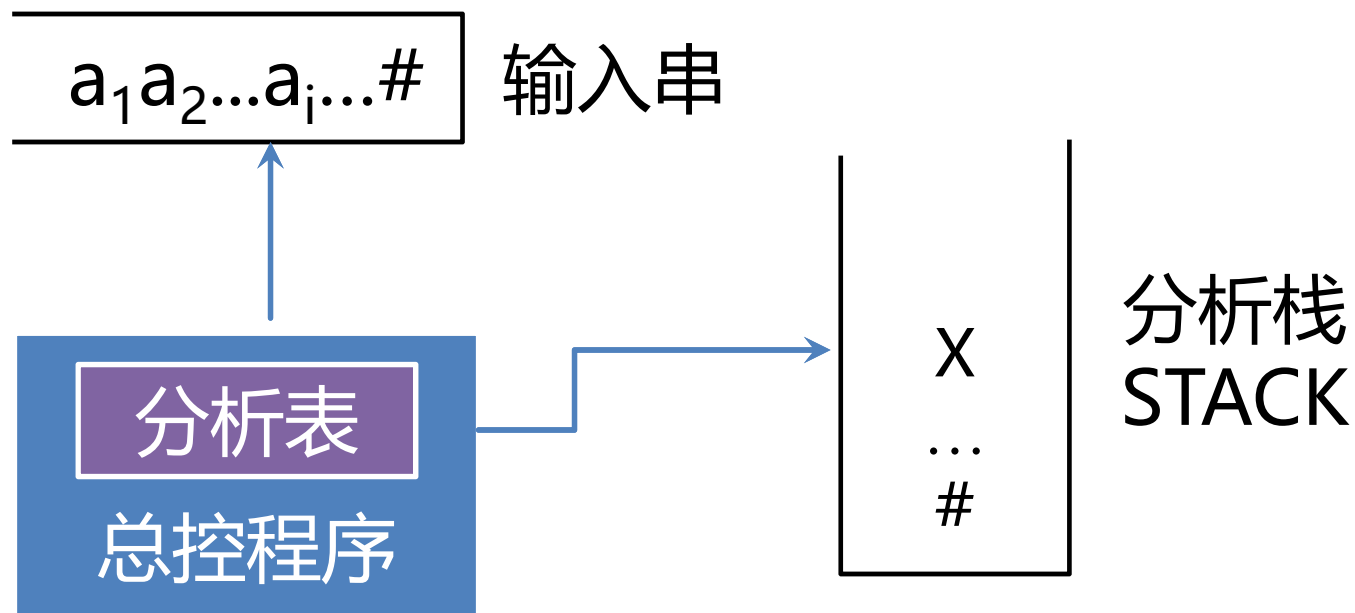
预测分析程序构成

- 计算思维的典型方法
 - 知识与控制的分离
 - 自动化

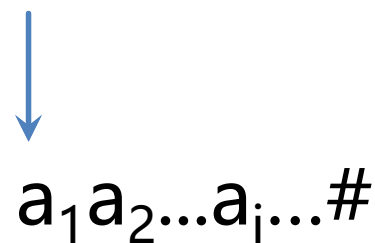
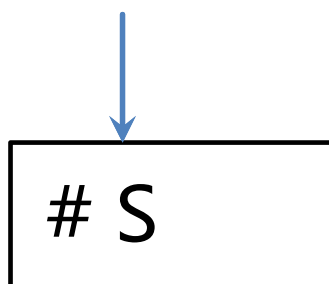
- ▶ **总控程序**，根据现行栈顶符号和当前输入符号，执行动作
- ▶ **分析表** $M[A, a]$ 矩阵， $A \in V_N$ ， $a \in V_T$ 是终结符或 '#'
- ▶ **分析栈** **STACK** 用于存放文法符号



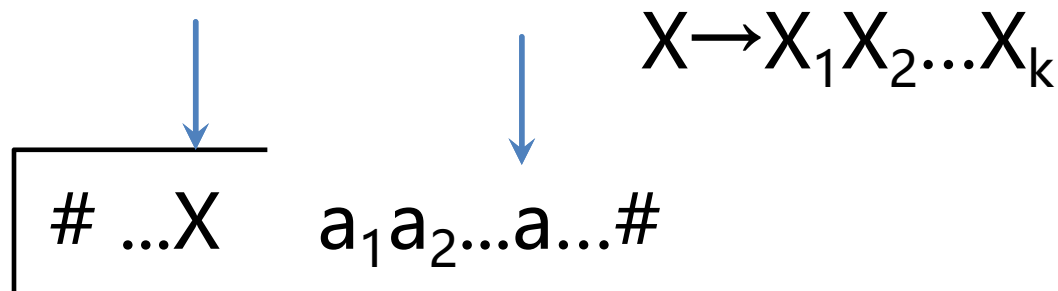
预测分析过程



分析开始时:



预测分析过程



- ▶ 总控程序根据当前栈顶符号X和输入符号a，执行下列三动作之一：
 1. 若 $X = a = \text{'\#'}'$ ，则宣布分析成功，停止分析。
 2. 若 $X = a \neq \text{'\#'}'$ ，则把X从STACK栈顶逐出，让a指向下一个输入符号。
 3. 若X是一个非终结符，则查看分析表M。
- ▶ 若 $M[X, a]$ 中存放着关于X的一个产生式，把X逐出STACK栈顶，把产生式的右部符号串按反序——推进STACK栈(若右部符号为 ϵ ，则意味不推什么东西进栈)。
- ▶ 若 $M[X, a]$ 中存放着“出错标志”，则调用出错诊察程序ERROR。

总控程序实现

BEGIN

首先把 ' #' 然后把文法开始符号推进STACK栈;
把第一个输入符号读进a;

FLAG:=TRUE;

WHILE FLAG DO

BEGIN

把STACK栈顶符号上托出去并放在X中;

IF $X \in V_T$ THEN

IF $X = a$ THEN 把下一输入符号读进a

ELSE ERROR

总控程序实现

ELSE IF $X = \#$ THEN

IF $X = a$ THEN FLAG:=FALSE

ELSE ERROR

ELSE IF $M[X, a] = \{X \rightarrow X_1 X_2 \dots X_k\}$ THEN

把 X_k, X_{k-1}, \dots, X_1 ——推进STACK栈

/* 若 $X_1 X_2 \dots X_k = \varepsilon$, 不推什么进栈 */

ELSE ERROR

END OF WHILE;

STOP /*分析成功, 过程完毕*/

END

编译原理

预测分析示例

预测分析示例

► 对于文法G(E):

$$E \rightarrow TE'$$

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

$$T \rightarrow FT'$$

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

$$F \rightarrow (E) \mid i$$

输入串为 $i_1*i_2+i_3$ ，利用分析表进行预测分析

	i	+	*	()	#
E	$E \rightarrow TE'$			$E \rightarrow TE'$		
E'		$E' \rightarrow +TE'$			$E' \rightarrow \varepsilon$	$E' \rightarrow \varepsilon$
T	$T \rightarrow FT'$			$T \rightarrow FT'$		
T'		$T' \rightarrow \varepsilon$	$T' \rightarrow *FT'$		$T' \rightarrow \varepsilon$	$T' \rightarrow \varepsilon$
F	$F \rightarrow i$			$F \rightarrow (E)$		

预测分析示例

► 对于文法G(E):

$$E \rightarrow TE'$$

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

$$T \rightarrow FT'$$

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

$$F \rightarrow (E) \mid i$$

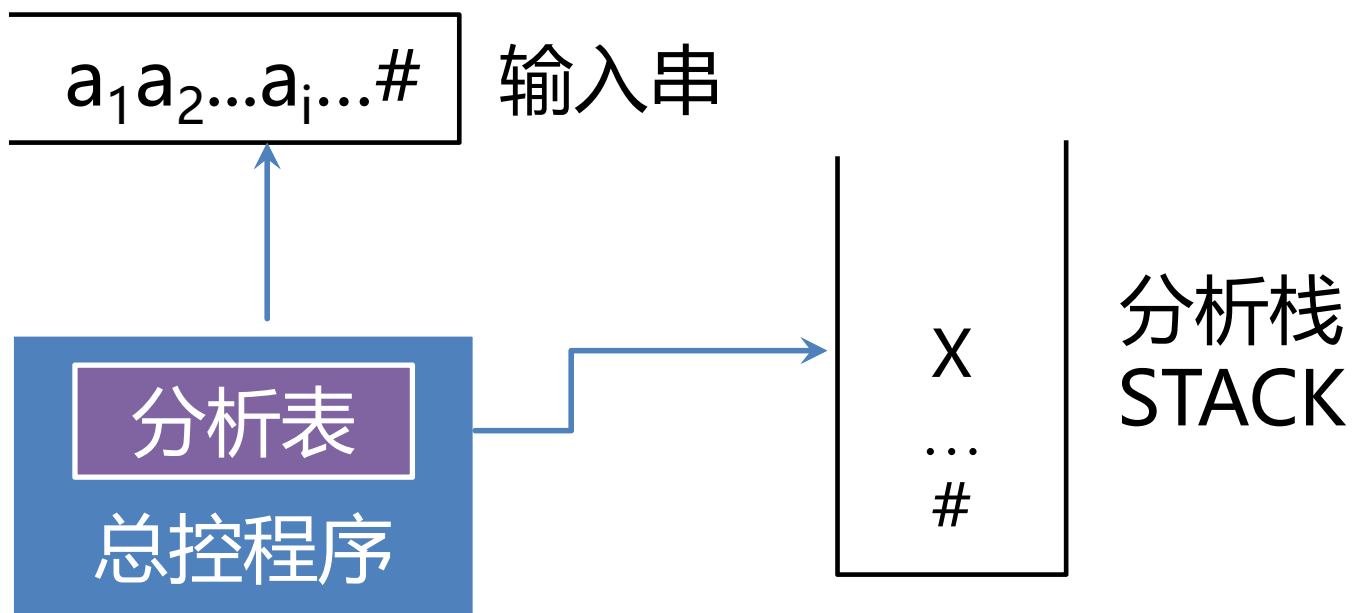
输入串为 $i_1*i_2+i_3$ ，利用分析表进行预测分析

	i	+	*	()	#
E	$E \rightarrow TE'$			$E \rightarrow TE'$		
E'		$E' \rightarrow +TE'$			$E' \rightarrow \varepsilon$	$E' \rightarrow \varepsilon$
T	$T \rightarrow FT'$			$T \rightarrow FT'$		
T'		$T' \rightarrow \varepsilon$	$T' \rightarrow *FT'$		$T' \rightarrow \varepsilon$	$T' \rightarrow \varepsilon$
F	$F \rightarrow i$			$F \rightarrow (E)$		

编译原理

构造预测分析表

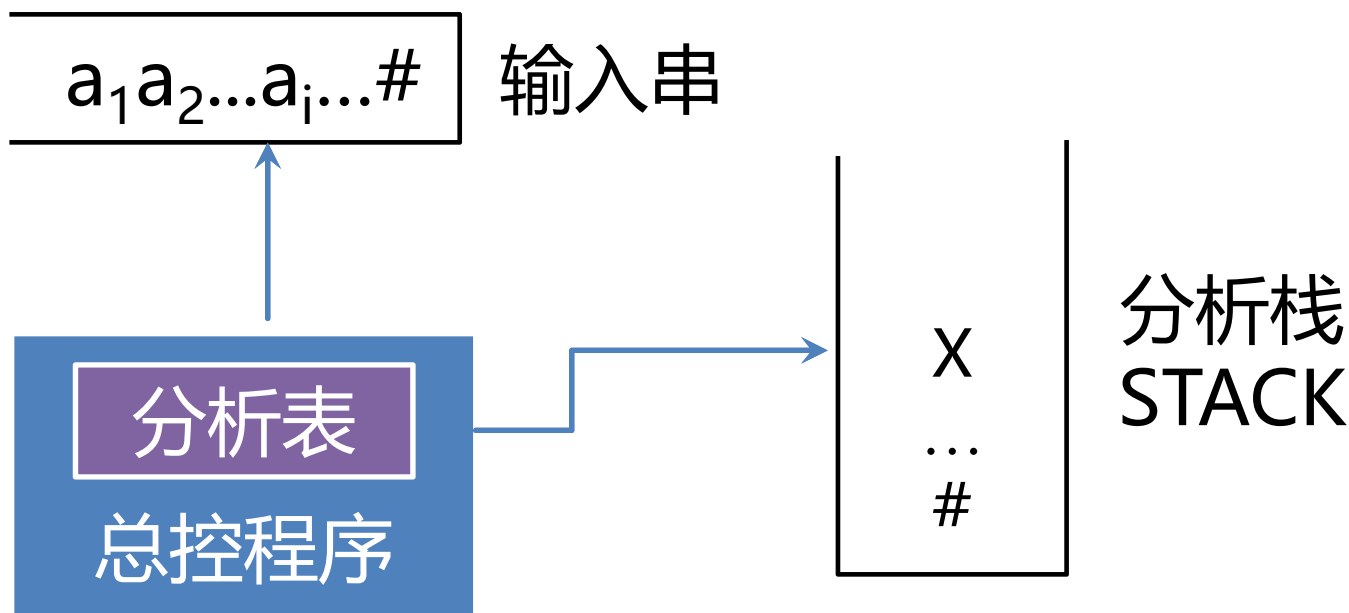
预测分析程序构成



- 计算思维的典型方法
 - 知识与控制的分离
 - 自动化

分析表 $M[A, a]$ 的构造

- ▶ 构造 $FIRST(\alpha)$ 和 $FOLLOW(A)$
- ▶ 构造分析表 $M[A, a]$



分析表的构造思想

► 对于文法G(E):

$E \rightarrow TE'$

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

$T \rightarrow FT'$

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

$F \rightarrow (E) \mid i$

输入串为 $i_1*i_2+i_3$, 利用分析表进行预

	i	+	*	()	#
E	$E \rightarrow TE'$			$E \rightarrow TE'$		
E'		$E' \rightarrow +TE'$			$E' \rightarrow \varepsilon$	$E' \rightarrow \varepsilon$
T	$T \rightarrow FT'$			$T \rightarrow FT'$		
T'		$T' \rightarrow \varepsilon$	$T' \rightarrow *FT'$		$T' \rightarrow \varepsilon$	$T' \rightarrow \varepsilon$
F	$F \rightarrow i$			$F \rightarrow (E)$		

步骤	符号栈	输入串	所用产生式
0	#E	$i_1*i_2+i_3\#$	
1	#E'T	$i_1*i_2+i_3\#$	$E \rightarrow TE'$
2	#E'T'F	$i_1*i_2+i_3\#$	$T \rightarrow FT'$
3	#E'T'i	$i_1*i_2+i_3\#$	$F \rightarrow i$
4	#E'T'	$*i_2+i_3\#$	
5	#E'T'F*	$*i_2+i_3\#$	$T' \rightarrow *FT'$
6	#E'T'F	$i_2+i_3\#$	
7	#E'T'i	$i_2+i_3\#$	$F \rightarrow i$
8	#E'T'	$+i_3\#$	
9	#E'	$+i_3\#$	$T' \rightarrow \varepsilon$
10	#E'T+	$+i_3\#$	$E' \rightarrow +TE'$
11	#E'T	$i_3\#$	
11	#E'T	$i_3\#$	
12	#E'T'F	$i_3\#$	$T \rightarrow FT'$
13	#E'T'i	$i_3\#$	$F \rightarrow i$
14	#E'T'	#	
15	#E'	#	$T' \rightarrow \varepsilon$
16	#	#	$E' \rightarrow \varepsilon$

分析表 $M[A, a]$ 的构造算法

► 构造 G 的分析表 $M[A, a]$ ，确定每个产生式 $A \rightarrow \alpha$ 在表中的位置

1. 对文法 G 的每个产生式 $A \rightarrow \alpha$ 执行第2步和第3步；
2. 对每个终结符 $a \in \text{FIRST}(\alpha)$ ，把 $A \rightarrow \alpha$ 加至 $M[A, a]$ 中；
3. 若 $\epsilon \in \text{FIRST}(\alpha)$ ，则对任何 $b \in \text{FOLLOW}(A)$ 把 $A \rightarrow \alpha$ 加至 $M[A, b]$ 中。
4. 把所有无定义的 $M[A, a]$ 标上“出错标志”。

练习：分析表的构造

► 对于文法G(E):

$$E \rightarrow TE'$$
$$E' \rightarrow +TE' \mid \varepsilon$$
$$T \rightarrow FT'$$
$$T' \rightarrow *FT' \mid \varepsilon$$
$$F \rightarrow (E) \mid i$$

构造每个非终结符的FIRST和FOLLOW集合：

$$\text{FIRST}(E) = \{ (, i \}$$
$$\text{FIRST}(E') = \{ +, \varepsilon \}$$
$$\text{FIRST}(T) = \{ (, i \}$$
$$\text{FIRST}(T') = \{ *, \varepsilon \}$$
$$\text{FIRST}(F) = \{ (, i \}$$
$$\text{FOLLOW}(E) = \{), \# \}$$
$$\text{FOLLOW}(E') = \{), \# \}$$
$$\text{FOLLOW}(T) = \{ +,), \# \}$$
$$\text{FOLLOW}(T') = \{ +,), \# \}$$
$$\text{FOLLOW}(F) = \{ *, +,), \# \}$$

分析表的构造

$\text{FIRST}(E) = \{ (, i \}$ $\text{FOLLOW}(E) = \{), \# \}$
 $\text{FIRST}(E') = \{ +, \epsilon \}$ $\text{FOLLOW}(E') = \{), \# \}$
 $\text{FIRST}(T) = \{ (, i \}$ $\text{FOLLOW}(T) = \{ +,), \# \}$
 $\text{FIRST}(T') = \{ *, \epsilon \}$ $\text{FOLLOW}(T') = \{ +,), \# \}$
 $\text{FIRST}(F) = \{ (, i \}$ $\text{FOLLOW}(F) = \{ *, +,), \# \}$

► 对于文法G(E):

$E \rightarrow TE'$

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

$T \rightarrow FT'$

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

$F \rightarrow (E) \mid i$

	i	+	*	()	#
E						
E'						
T						
T'						
F						

构造该文法的预测分析表。

1. 对文法G的每个产生式 $A \rightarrow \alpha$ 执行第2步和第3步；
2. 对每个终结符 $a \in \text{FIRST}(\alpha)$ ，把 $A \rightarrow \alpha$ 加至 $M[A, a]$ 中；
3. 若 $\epsilon \in \text{FIRST}(\alpha)$ ，则对任何 $b \in \text{FOLLOW}(A)$ 把 $A \rightarrow \alpha$ 加至 $M[A, b]$ 中。
4. 把所有无定义的 $M[A, a]$ 标上“出错标志”。

分析表的构造

$\text{FIRST}(E) = \{ (, i \}$ $\text{FOLLOW}(E) = \{), \# \}$
 $\text{FIRST}(E') = \{ +, \varepsilon \}$ $\text{FOLLOW}(E') = \{), \# \}$
 $\text{FIRST}(T) = \{ (, i \}$ $\text{FOLLOW}(T) = \{ +,), \# \}$
 $\text{FIRST}(T') = \{ *, \varepsilon \}$ $\text{FOLLOW}(T') = \{ +,), \# \}$
 $\text{FIRST}(F) = \{ (, i \}$ $\text{FOLLOW}(F) = \{ *, +,), \# \}$

► 对于文法G(E):

$E \rightarrow TE'$
 $E' \rightarrow +TE' \mid \varepsilon$
 $T \rightarrow FT'$
 $T' \rightarrow *FT' \mid \varepsilon$
 $F \rightarrow (E) \mid i$

	i	+	*	()	#
E	$E \rightarrow TE'$			$E \rightarrow TE'$		
E'		$E' \rightarrow +TE'$			$E' \rightarrow \varepsilon$	$E' \rightarrow \varepsilon$
T	$T \rightarrow FT'$			$T \rightarrow FT'$		
T'		$T' \rightarrow \varepsilon$	$T' \rightarrow *FT'$		$T' \rightarrow \varepsilon$	$T' \rightarrow \varepsilon$
F	$F \rightarrow i$			$F \rightarrow (E)$		

构造该文法的预测分析表。

1. 对文法G的每个产生式 $A \rightarrow \alpha$ 执行第2步和第3步；
2. 对每个终结符 $a \in \text{FIRST}(\alpha)$ ，把 $A \rightarrow \alpha$ 加至 $M[A, a]$ 中；
3. 若 $\varepsilon \in \text{FIRST}(\alpha)$ ，则对任何 $b \in \text{FOLLOW}(A)$ 把 $A \rightarrow \alpha$ 加至 $M[A, b]$ 中。
4. 把所有无定义的 $M[A, a]$ 标上“出错标志”。

预测分析示例

► 对于文法G(E):

$E \rightarrow TE'$

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

$T \rightarrow FT'$

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

$F \rightarrow (E) \mid i$

输入串为 $i_1*i_2+i_3$, 利用分析表进行预

	i	+	*	()	#
E	$E \rightarrow TE'$			$E \rightarrow TE'$		
E'		$E' \rightarrow +TE'$			$E' \rightarrow \varepsilon$	$E' \rightarrow \varepsilon$
T	$T \rightarrow FT'$			$T \rightarrow FT'$		
T'		$T' \rightarrow \varepsilon$	$T' \rightarrow *FT'$		$T' \rightarrow \varepsilon$	$T' \rightarrow \varepsilon$
F	$F \rightarrow i$			$F \rightarrow (E)$		

步骤	符号栈	输入串	所用产生式
0	#E	$i_1*i_2+i_3\#$	
1	#E'T	$i_1*i_2+i_3\#$	$E \rightarrow TE'$
2	#E'T'F	$i_1*i_2+i_3\#$	$T \rightarrow FT'$
3	#E'T'i	$i_1*i_2+i_3\#$	$F \rightarrow i$
4	#E'T'	$*i_2+i_3\#$	
5	#E'T'F*	$*i_2+i_3\#$	$T' \rightarrow *FT'$
6	#E'T'F	$i_2+i_3\#$	
7	#E'T'i	$i_2+i_3\#$	$F \rightarrow i$
8	#E'T'	$+i_3\#$	
9	#E'	$+i_3\#$	$T' \rightarrow \varepsilon$
10	#E'T+	$+i_3\#$	$E' \rightarrow +TE'$
11	#E'T	$i_3\#$	
11	#E'T	$i_3\#$	
12	#E'T'F	$i_3\#$	$T \rightarrow FT'$
13	#E'T'i	$i_3\#$	$F \rightarrow i$
14	#E'T'	#	
15	#E'	#	$T' \rightarrow \varepsilon$
16	#	#	$E' \rightarrow \varepsilon$

LL(1)文法与二义性

- ▶ 如果G是左递归或二义的，那么，M至少含有一个多重定义入口。因此，消除左递归和提取左因子将有助于获得无多重定义的分析表M。
- ▶ 可以证明，一个文法G的预测分析表M不含多重定义入口，当且仅当该文法为LL(1)的。
- ▶ LL(1)文法不是二义的。

LL(1)文法与二义性

► G(S):

$$S \rightarrow iCtS \mid iCtSeS \mid a$$

$$C \rightarrow b$$

提取左因子之后，改写成：

► G(S):

$$S \rightarrow iCtSS' \mid a$$

$$S' \rightarrow eS \mid \varepsilon$$

$$C \rightarrow b$$

LL(1)文法与二义性

if ... then if ... then ... else ...

if ... then if ... then ...
else ...

if ... then if ... then ...
else ...

► G(S):

$$S \rightarrow iCtS \mid iCtSeS \mid a$$
$$C \rightarrow b$$

提取左因子之后，改写成：

► G(S):

$$S \rightarrow iCtSS' \mid a$$
$$S' \rightarrow eS \mid \varepsilon$$
$$C \rightarrow b$$

	a	b	e	i	t	#
S	$S \rightarrow a$			$S \rightarrow iCtSS'$		
S'			$S' \rightarrow eS$ $S' \rightarrow \varepsilon$			$S' \rightarrow \varepsilon$
C		$C \rightarrow b$				

小结

- ▶ 预测分析程序的结构
- ▶ 预测分析程序的原理
- ▶ 预测分析表的构造
 - ▶ 消除左递归，消除回溯
 - ▶ 计算FIRST、FOLLOW集合
 - ▶ 构造预测分析表