

Chapter 5 语法分析-自下而上分析



回顾

- ▶ 算符优先文法
 - ▶ 算符优先级顺序
- ▶ 算符优先方法归约方法
 - ▶ $a_{i-1} < a_i \doteq \cdots a_j > a_{j+1}$
- ▶ LR(0)分析法
 - ▶ LR(0)项目的计算
 - ▶ LR(0)分析表的生成

SLR分析表的构造

- ▶ LR(0)文法太简单，没有实用价值。
 - ▶ 假定一个LR(0)规范族中含有如下的一个项目集
 - ▶ (状态) $I = \{X \rightarrow \alpha \cdot b\beta, A \rightarrow \alpha \cdot, B \rightarrow \alpha \cdot\}$ 。
 $FOLLOW(A)$ 和 $FOLLOW(B)$ 的交集为 \emptyset ，且不包含 b ，那么，当状态 I 面临任何输入符号 a 时，可以：
 1. 若 $a=b$ ，则移进；
 2. 若 $a \in FOLLOW(A)$ ，用产生式 $A \rightarrow \alpha$ 进行归约；
 3. 若 $a \in FOLLOW(B)$ ，用产生式 $B \rightarrow \alpha$ 进行归约；
 4. 此外，报错。
-

-
- ▶ 假定LR(0)规范族的一个项目集 $I = \{A_1 \rightarrow \alpha \cdot a_1 \beta_1, A_2 \rightarrow \alpha \cdot a_2 \beta_2, \dots, A_m \rightarrow \alpha \cdot a_m \beta_m, B_1 \rightarrow \alpha ; B_2 \rightarrow \alpha ; \dots, B_n \rightarrow \alpha \}$ 如果集合 $\{a_1, \dots, a_m\}$, $\text{FOLLOW}(B_1), \dots, \text{FOLLOW}(B_n)$ 两两不相交(包括不得有两个FOLLOW集合有#), 则
1. 若a是某个 a_i , $i=1,2,\dots,m$, 则移进;
 2. 若 $a \in \text{FOLLOW}(B_i)$, $i=1,2,\dots,n$, 则用产生式 $B_i \rightarrow \alpha$ 进行归约;
 3. 此外, 报错。
- ▶ 冲突性动作的这种解决办法叫做SLR(I)解决办法
-

构造SLR(1)分析表方法:

- ▶ 首先把G拓广为G', 对G'构造LR(0)项目集规范族C和活前缀识别自动机的状态转换函数GO.
- ▶ 拓广文法: 增加 $S' \rightarrow S$, 新起始符号
- ▶ 然后使用C和GO, 按下面的算法构造SLR分析表:
 - ▶ 令每个项目集 I_k 的下标k作为分析器的状态, 包含项目 $S' \rightarrow \cdot S$ 的集合 I_k 的下标k为分析器的初态。



分析表的ACTION和GOTO子表构造方法：

1. 若项目 $A \rightarrow \alpha \cdot a\beta$ 属于 I_k 且 $GO(I_k, a) = I_j$, a 为终结符, 则置 $ACTION[k, a]$ 为 “sj”;
2. 若项目 $A \rightarrow \alpha \cdot$ 属于 I_k , 那么, 对任何终结符 a , $a \in FOLLOW(A)$, 置 $ACTION[k, a]$ 为 “rj”; 其中, 假定 $A \rightarrow \alpha$ 为文法 G' 的第 j 个产生式;
3. 若项目 $S' \rightarrow S \cdot$ 属于 I_k , 则置 $ACTION[k, \#]$ 为 “acc”;
4. 若 $GO(I_k, A) = I_j$, A 为非终结符, 则置 $GOTO[k, A] = j$;
5. 分析表中凡不能用规则1至4填入信息的空白格均置上“出错标志”。



-
- ▶ 按上述方法构造出的**ACTION**与**GOTO**表如果不含多重入口，则称该文法为**SLR(I)文法**。
 - ▶ 使用**SLR**表的分析器叫做一个**SLR分析器**。
 - ▶ 每个**SLR(I)**文法都是无二义的。但也存在许多无二义文法不是**SLR(I)**的。

► 例 考察下面的拓广文法：

(0) $S' \rightarrow E$

(1) $E \rightarrow E + T$

(2) $E \rightarrow T$

(3) $T \rightarrow T * F$

(4) $T \rightarrow F$

(5) $F \rightarrow (E)$

(6) $F \rightarrow i$



► 这个文法的LR(0)项目集规范族为：

I_0 : $S' \rightarrow \cdot E$
 $E \rightarrow \cdot E + T$
 $E \rightarrow \cdot T$
 $T \rightarrow \cdot T * F$
 $T \rightarrow \cdot F$
 $F \rightarrow \cdot (E)$
 $F \rightarrow \cdot i$

I_1 : $S' \rightarrow E \cdot$
 $E \rightarrow E \cdot + T$

I_2 : $E \rightarrow T \cdot$
 $T \rightarrow T \cdot * F$

► I_3 : $T \rightarrow F \cdot$

I_4 : $F \rightarrow (\cdot E)$
 $E \rightarrow \cdot E + T$
 $E \rightarrow \cdot T$
 $T \rightarrow \cdot T * F$
 $T \rightarrow \cdot F$
 $F \rightarrow \cdot (E)$
 $F \rightarrow \cdot i$

I_5 : $F \rightarrow i \cdot$

I_6 : $E \rightarrow E + \cdot T$
 $T \rightarrow \cdot T * F$
 $T \rightarrow \cdot F$
 $F \rightarrow \cdot (E)$
 $F \rightarrow \cdot i$

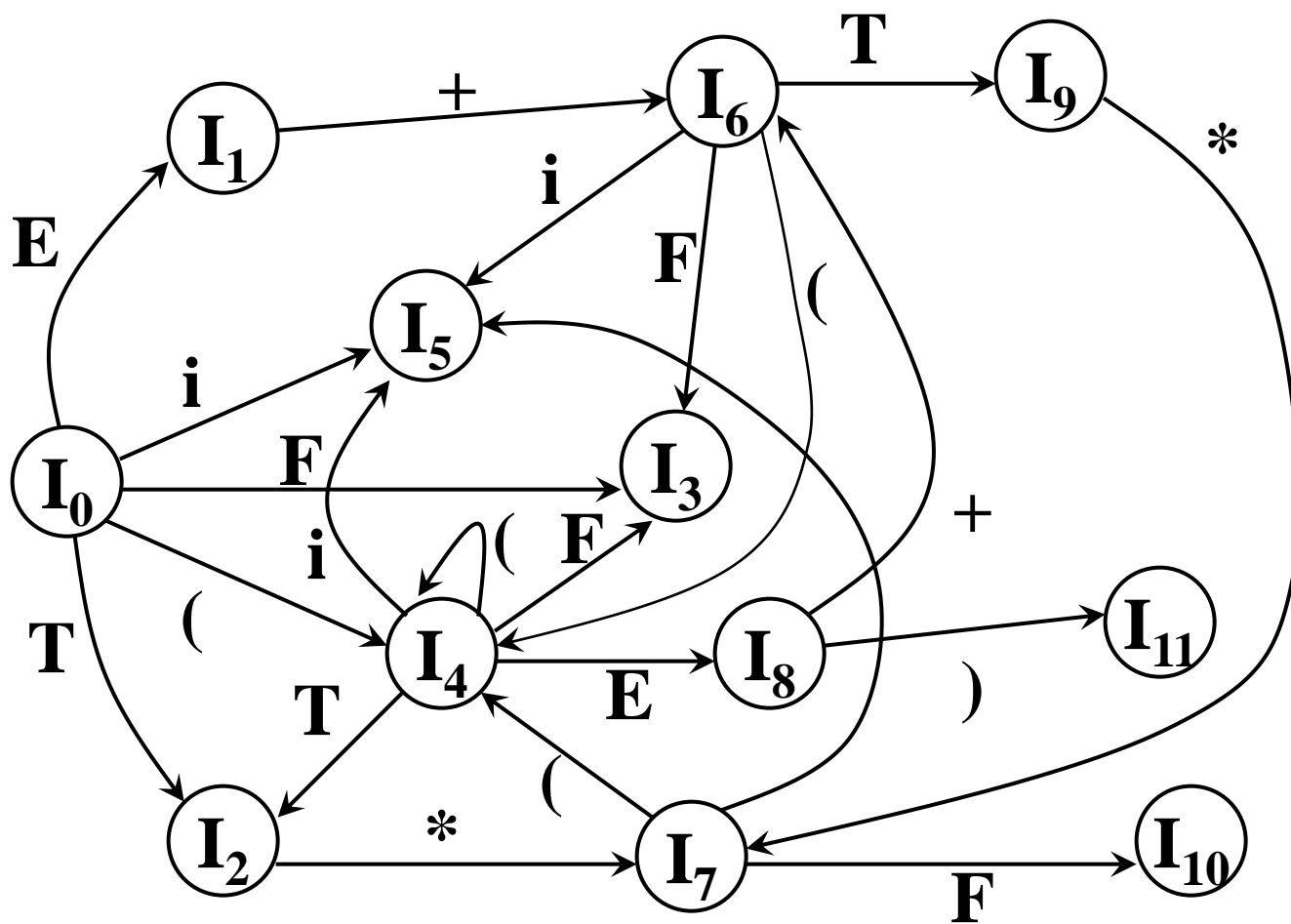
I_7 : $T \rightarrow T * \cdot F$
 $F \rightarrow \cdot (E)$
 $F \rightarrow \cdot i$

I_8 : $F \rightarrow (E \cdot)$
 $E \rightarrow E \cdot + T$

I_9 : $E \rightarrow E + T \cdot$
 $T \rightarrow T \cdot * F$

I_{10} : $T \rightarrow T * F \cdot$

I_{11} : $F \rightarrow (E) \cdot$



-
- ▶ I_1 、 I_2 和 I_9 都含有“移进-归约”冲突。
 - ▶ $\text{FOLLOW}(E) = \{\#,), +\}$,

$I_1:$ $S' \rightarrow E \cdot$
 $E \rightarrow E \cdot + T$

$I_2:$ $E \rightarrow T \cdot$
 $T \rightarrow T \cdot * F$

$I_9:$ $E \rightarrow E + T \cdot$
 $T \rightarrow T \cdot * F$

	Action						Goto		
状态	i	+	*	()	#	E	T	F
0	s5			s4			1	2	3
1		s6				acc			
2		r2	s7		r2	r2			
3		r4	r4		r4	r4			
4	s5			s4			8	2	3
5		r6	r6		r6	r6			
6	s5			s4				9	3
7	s5			s4					10
8		s6			s11				
9		r1	s7		r1	r1			
10		r3	r3		r3	r3			
11		r5	r5		r5	r5			



问题

- ▶ **SLR在方法中,如果项目集 I_i 含项目 $A \rightarrow \alpha$.而且下一输入符号 $a \in \text{FOLLOW}(A)$,则状态 i 面临 a 时,可选用“用 $A \rightarrow \alpha$ 归约”动作。但在有些情况下,当状态 i 显现于栈顶时,栈里的活前缀未必允许把 α 归约为 A ,因为可能根本就不存在一个形如“ $\beta A a$ ”的规范句型。因此,在这种情况下,用“ $A \rightarrow \alpha$ ”归约不一定合适。**
- ▶ **FOLLOW集合提供的信息太泛!**

SLR分析的问题

▶ 如下文法:

▶ $S' \rightarrow S$ $S \rightarrow aAa|aBb|bAb|bBa$

▶ $A \rightarrow a|b$ $B \rightarrow a|b$

I0 $S' \rightarrow \cdot S$ $S \rightarrow \cdot aAa | \cdot aBb | \cdot bAb | \cdot bBa$

\Downarrow a

I1 $S \rightarrow a \cdot Aa$ $A \rightarrow \cdot a$ $A \rightarrow \cdot b$
 $S \rightarrow a \cdot Bb$ $B \rightarrow \cdot a$ $B \rightarrow \cdot b$

\Downarrow a

I2 $A \rightarrow a \cdot$ $B \rightarrow a \cdot$

由于Follow(A)和Follow(B)都是
{a,b}, 无法判断应该如何归约

LR(1)分析

- ▶ 将后续的一个字符加入项目之中
- ▶ $(A \rightarrow \alpha \cdot B\beta, a)$

LR(0)项目

$S \xRightarrow{*} \delta A \omega \xRightarrow{+} \delta \alpha \cdot \beta \omega$ 得到 $(A \rightarrow \alpha \cdot \beta)$

LR(1)项目

$S \xRightarrow{*} \delta A \textcolor{red}{a} \omega \xRightarrow{+} \delta \alpha \cdot \beta \textcolor{red}{a} \omega$ 得到 $(A \rightarrow \alpha \cdot \beta, \textcolor{red}{a})$

LR(0)的两个函数应该如何修改?

LR(1)分析

$$\begin{array}{lll} (A \rightarrow \alpha \cdot b\beta, a) & \xrightarrow{b} & (A \rightarrow \alpha b \cdot \beta, a) \\ (A \rightarrow \alpha \cdot B\beta, a) & \xrightarrow{B} & (A \rightarrow \alpha B \cdot \beta, a) \end{array} \quad \begin{array}{l} \text{Go函数} \\ \text{没有变化} \end{array}$$

$$\begin{array}{ccc} (A \rightarrow \alpha \cdot B\beta, a) & \xrightarrow{\varepsilon - \text{closure}} & (B \rightarrow \cdot c, d) \\ B \rightarrow c & & d = \text{First}(\beta a) \end{array}$$

$\varepsilon - \text{closure}$ 需要修改

LR(1)分析

► 拓广文法

► $S' \rightarrow S, S \rightarrow BB, B \rightarrow aB, B \rightarrow b$

$$\begin{array}{l} (S' \rightarrow \cdot S, \#) \\ S \rightarrow BB \end{array} \Longrightarrow (S \rightarrow \cdot BB, First(\epsilon \#))$$

$$\begin{array}{l} (A \rightarrow \alpha \cdot B\beta, a) \\ B \rightarrow c \end{array} \xrightarrow{\epsilon - closure} \begin{array}{l} (B \rightarrow \cdot c, d) \\ d = First(\beta a) \end{array}$$

LR(1)分析

► 拓广文法

► $S' \rightarrow S, S \rightarrow BB, B \rightarrow aB, B \rightarrow b$

$$(S' \rightarrow \cdot S, \#) \Longrightarrow (S \rightarrow \cdot BB, \#)$$

$$\begin{array}{ccc} (S \rightarrow \cdot B\beta, \#) & \Longrightarrow & (B \rightarrow \cdot aB, First(\beta\#)) \\ B \rightarrow aB & & (B \rightarrow \cdot aB, a/b) \end{array}$$

$$\begin{array}{ccc} (A \rightarrow \alpha \cdot B\beta, a) & \xrightarrow{\varepsilon - closure} & (B \rightarrow \cdot c, d) \\ B \rightarrow c & & d = First(\beta a) \end{array}$$

Exercise

- ▶ 例12 已知文法G[S]:

$$S \rightarrow *A$$

$$A \rightarrow 0A1 \mid *$$

- ▶ (1) 求文法G的各非终结符号的FIRSTVT集和LASTVT集;
- ▶ (2) 构造文法G的优先关系矩阵, 并判断该文法是否是算符优先文法;

Exercise

- ▶ (1) 求文法G的各非终结符号的FIRSTVT集和LASTVT集。
- ▶ 根据非终结符号的FIRSTVT集定义得到
 - ▶ $\text{FIRSTVT}(S) = \{*\}$
 - ▶ $\text{FIRSTVT}(A) = \{0, *\}$
- ▶ 根据非终结符号的LASTVT集定义得到
 - ▶ $\text{LASTVT}(S) = \{*, 1\}$
 - ▶ $\text{LASTVT}(A) = \{1, *\}$

Exercise

对 $S \rightarrow *A$, $(* \not\leq 0)$, $(* \not\leq *)$

对 $A \rightarrow 0A1$,

$(0 \dot{=} 1)$

$(0 \not\leq 0)$, $(0 \not\leq *)$

$(1 \not\geq 1)$, $(* \not\geq 1)$

	0	1	*
0	$\not\leq$	$\dot{=}$	$\not\leq$
1		$\not\geq$	
*	$\not\leq$	$\not\geq$	$\not\leq$

练习：

- ▶ 设文法 $G(S)$:
- ▶ $S \rightarrow (A) \mid a$
- ▶ $A \rightarrow A+S \mid S$

- ▶ (1) 构造非终结符的FIRSTVT和LASTVT集合
- ▶ (2) 构造优先关系表

练习

- ▶ 已知文法 $G(S)$:
- ▶ $S \rightarrow aS \mid bS \mid a$
- ▶ (1) 构造该文法的LR (0) 项目集规范族
- ▶ (2) 构造识别该文法所产生的活前缀的DFA
- ▶ (3) 构造其SLR分析表, 并判断该文法是否是SLR(I)文法

练习

- ▶ 已知文法 $G = (\{a, b\}, \{S, B\}, S, P)$, 其中 P :
- ▶ $S \rightarrow BB$
- ▶ $B \rightarrow aB | b$

- ▶ (1) 构造该文法的LR (0) 项目集规范族
- ▶ (2) 构造识别该文法所产生的活前缀的DFA
- ▶ (3) 构造其LR分析表, 并写出aabab的分析过程

练习

- ▶ 已知文法 $G(P)$:
- ▶ $P \rightarrow aPb \mid Q$
- ▶ $Q \rightarrow bQc \mid bSc$
- ▶ $S \rightarrow Sa \mid a$

- ▶ 构造SLR分析表，以说明它是不是SLR文法

练习

- ▶ 已知文法 $G(S)$:
- ▶ $S \rightarrow iSeS \mid iS \mid a$
- ▶ 构造SLR分析表，以说明它是不是SLR文法

第5题

► 文法G(S): $S \rightarrow AS \mid b$ $A \rightarrow SA \mid a$

► 该文法的LR(0)项目有:

1. $S \rightarrow \cdot AS$

2. $S \rightarrow A \cdot S$

3. $S \rightarrow AS \cdot$

4. $S \rightarrow \cdot b$

5. $S \rightarrow b \cdot$

6. $A \rightarrow \cdot SA$

7. $A \rightarrow S \cdot A$

8. $A \rightarrow SA \cdot$

9. $A \rightarrow \cdot a$

10. $A \rightarrow a \cdot$

第5题

- ▶ 该文法的拓广文法为：

$$S' \rightarrow S \quad S \rightarrow AS \mid b \quad A \rightarrow SA \mid a$$

- ▶ LR(0)项目有：

$$S' \rightarrow \cdot S;$$

$$S' \rightarrow S \cdot ;$$

$$S \rightarrow \cdot AS;$$

$$S \rightarrow A \cdot S;$$

$$S \rightarrow AS \cdot ;$$

$$S \rightarrow \cdot b;$$

$$S \rightarrow b \cdot ;$$

$$A \rightarrow \cdot SA;$$

$$A \rightarrow S \cdot A;$$

$$A \rightarrow SA \cdot ;$$

$$A \rightarrow \cdot a;$$

$$A \rightarrow a \cdot ;$$

