



编译原理

构造SLR分析表

计算机科学与技术学院 王中卿

目录

- 构造SLR分析表
- 从文法构造识别可行前缀的DFA
 - 拓广文法
 - 构造LR(0)项目集规范族
- 从DFA构造分析表
 - *action*函数
 - *goto*函数



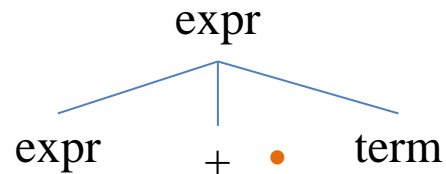
构造SLR分析表

- 构造SLR分析表本质上就是构造一个基于文法LR(0)项目的LR(0)自动机。
 - LR(0)项目
 - LR(0)自动机



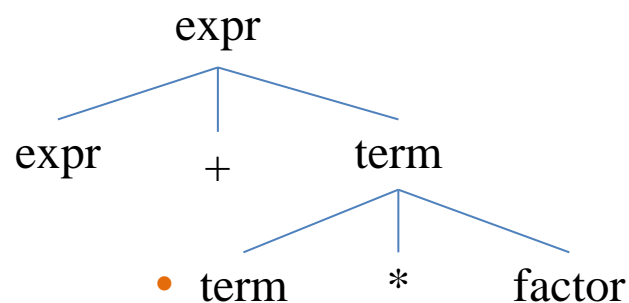
构造SLR分析表

- LR(0)项目
 - 在右部的某个地方加点的产生式
 - 加点的目的是用来表示分析过程中的状态



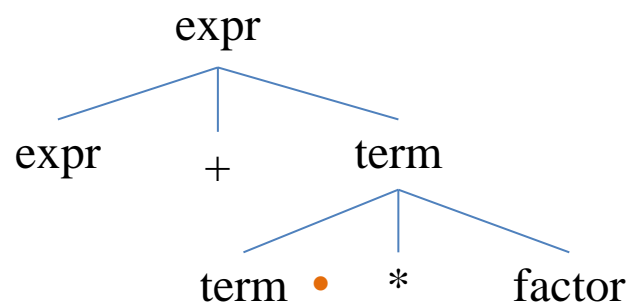
构造SLR分析表

- LR(0)项目
 - 在右部的某个地方加点的产生式
 - 加点的目的是用来表示分析过程中的状态



构造SLR分析表

- LR(0)项目
 - 在右部的某个地方加点的产生式
 - 加点的目的是用来表示分析过程中的状态



构造SLR分析表

- 例，从 $A \rightarrow XYZ$ 文法可得到如下四个项目；
 - $A \rightarrow \cdot XYZ$
 - $A \rightarrow X \cdot YZ$
 - $A \rightarrow XY \cdot Z$
 - $A \rightarrow XYZ \cdot$
- 对于产生式 $A \rightarrow \varepsilon$ 只有一个项目和它对应
 - $A \rightarrow \cdot$



构造SLR分析表

- LR(0)自动机
 - 根据文法LR(0)项目构造识别可行前缀的DFA
- LR(0)项 \Rightarrow LR(0)自动机
 - 文法的所有LR(0)项构成一组规范LR(0)项集，这些规范项集对应LR(0)自动机的状态。



构造SLR分析表

- 构造SLR分析表的两大步骤
 - 从文法构造识别可行前缀的DFA
 - 从上述DFA构造分析表



构造SLR分析表

- 从文法构造识别可行前缀的DFA
 - 拓广文法
 - $E' \rightarrow E$
 - $E \rightarrow E + T \mid T$
 - $T \rightarrow T * F \mid F$
 - $F \rightarrow (E) \mid \mathbf{id}$



构造SLR分析表

- 从文法构造识别可行前缀的DFA

- 构造LR(0)项目集规范族

I_0 :

$E' \rightarrow \cdot E$

(核心项目)

$E \rightarrow \cdot E + T$

$E \rightarrow \cdot T$

(非核心项目,
通过对核心项目求闭包而获得)

$T \rightarrow \cdot T * F$

$T \rightarrow \cdot F$

$F \rightarrow \cdot (E)$

$F \rightarrow \cdot \text{id}$



构造SLR分析表

- 从文法构造识别可行前缀的DFA

- 构造LR(0)项目集规范族

I_0 :

$E' \rightarrow \cdot E$

$E \rightarrow \cdot E + T$

$E \rightarrow \cdot T$

$T \rightarrow \cdot T * F$

$T \rightarrow \cdot F$

$F \rightarrow \cdot (E)$

$F \rightarrow \cdot \text{id}$

\xrightarrow{E}

I_1 :

$E' \rightarrow E \cdot$

$E \rightarrow E \cdot + T$

$I_1 := \text{goto} (I_0, E)$

\xrightarrow{T}

I_2 :

$E \rightarrow T \cdot$

$T \rightarrow T \cdot * F$

\xrightarrow{F}

I_3 :

$T \rightarrow F \cdot$



构造SLR分析表

- 从文法构造识别可行前缀的DFA
 - 构造LR(0)项目集规范族

I_0 :

$E' \rightarrow \cdot E$

$E \rightarrow \cdot E + T$

$E \rightarrow \cdot T$

$T \rightarrow \cdot T * F$

$T \rightarrow \cdot F$

$F \rightarrow \cdot (E)$

$F \rightarrow \cdot \text{id}$

(

I_4 :

$F \rightarrow (\cdot E)$

$E \rightarrow \cdot E + T$

$E \rightarrow \cdot T$

$T \rightarrow \cdot T * F$

$T \rightarrow \cdot F$

$F \rightarrow \cdot (E)$

$F \rightarrow \cdot \text{id}$

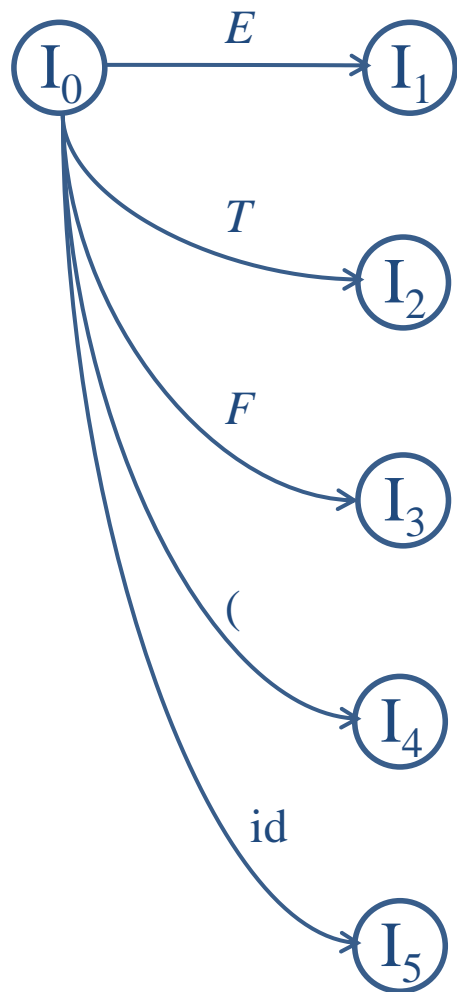
id

I_5 :

$F \rightarrow \text{id} \cdot$



构造SLR分析表



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

$\downarrow +$

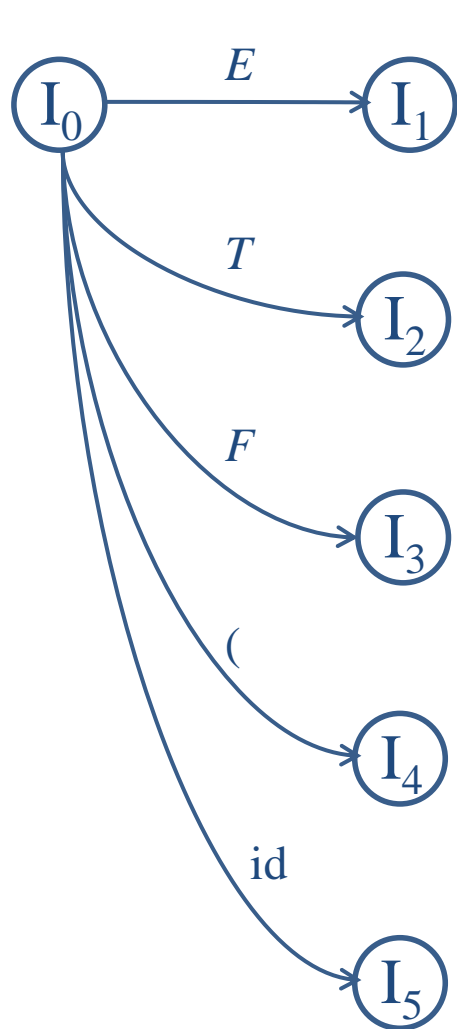
$I_6:$
 $E \rightarrow E + \cdot T$
 $T \rightarrow \cdot T * F$
 $T \rightarrow \cdot F$
 $F \rightarrow \cdot (E)$
 $F \rightarrow \cdot id$

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

$\downarrow *$

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

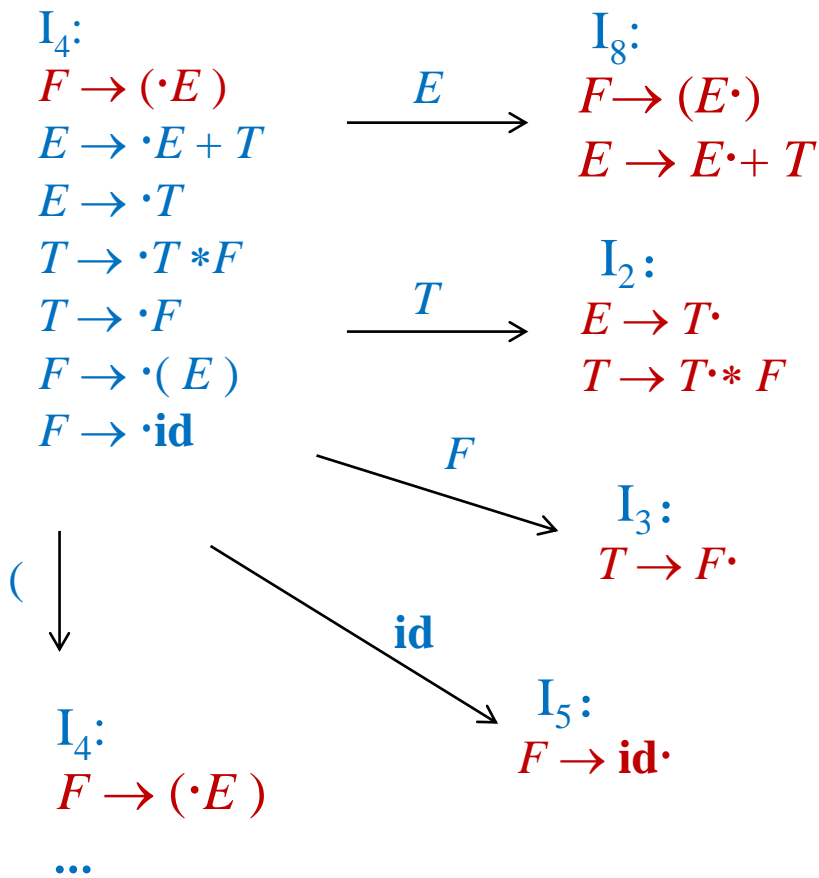
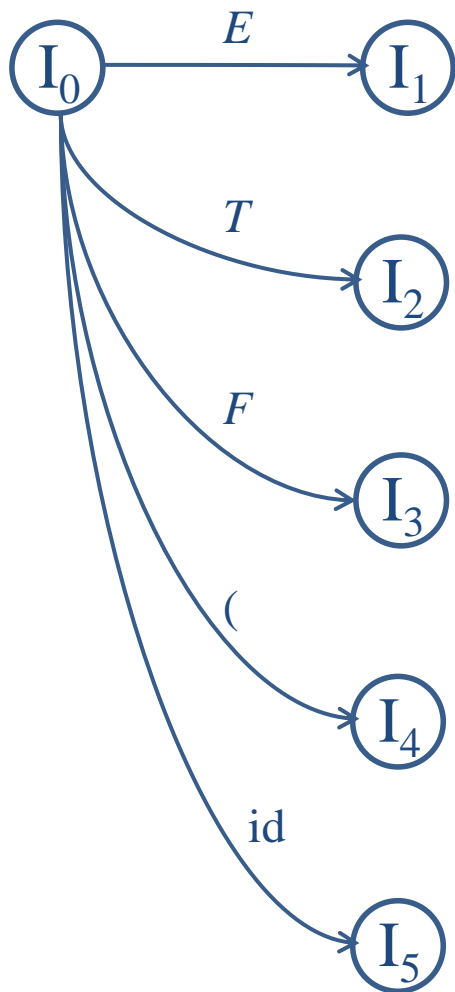
构造SLR分析表



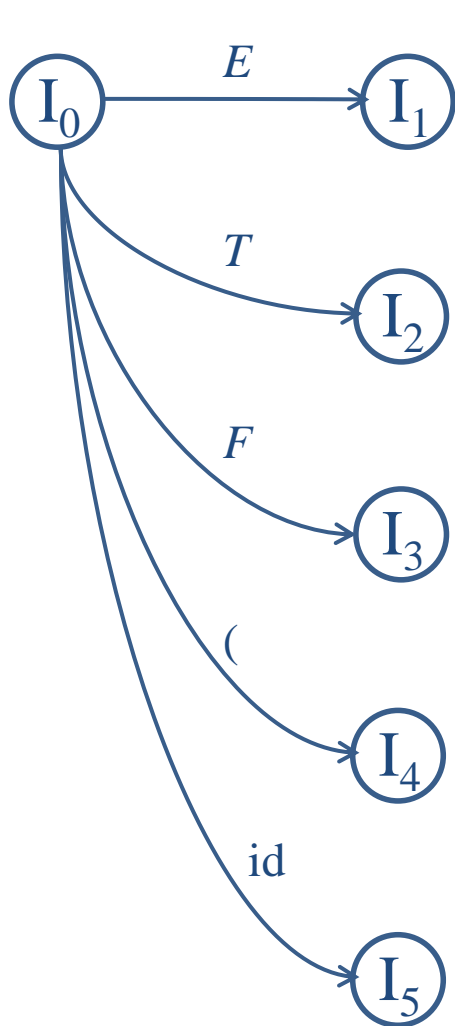
I_3 :
 $T \rightarrow F \cdot$

无状态转换

构造SLR分析表



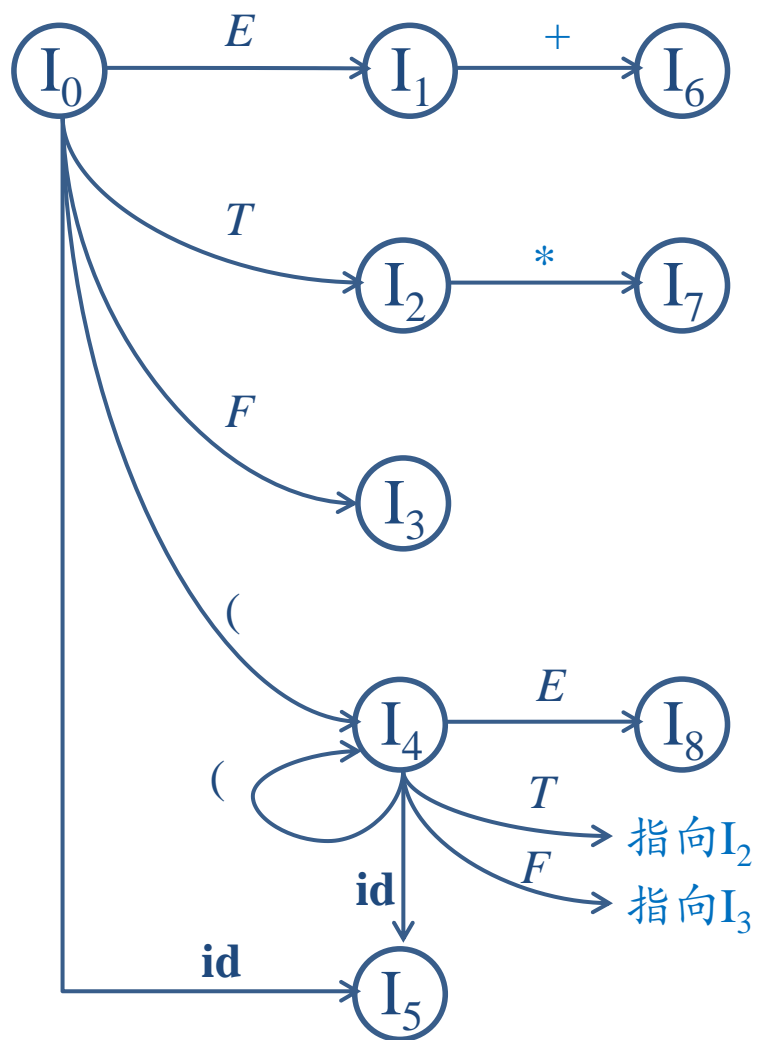
构造SLR分析表



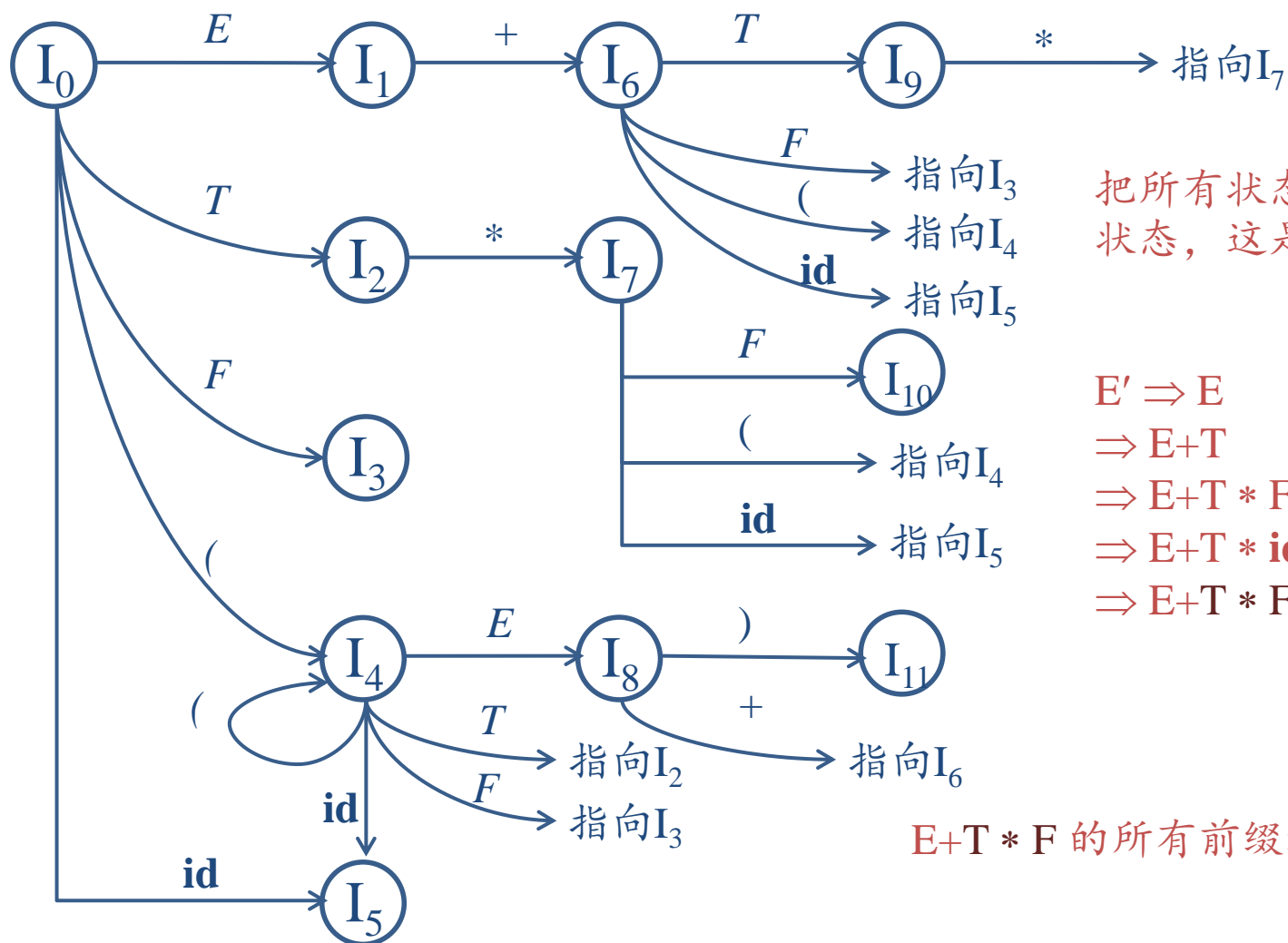
$I_5:$
 $F \rightarrow id \cdot$

无状态转换

构造SLR分析表



构造SLR分析表



把所有状态都作为接受状态，这是一个DFA

$E' \Rightarrow E$
 $\Rightarrow E+T$
 $\Rightarrow E+T * F$
 $\Rightarrow E+T * id$
 $\Rightarrow E+T * F * id$

$E+T * F$ 的所有前缀都可接受



构造SLR分析表

- 也可以构造一个识别可行前缀的NFA

I_0 :

每个项目一个状态

$E' \rightarrow \cdot E$

$E \rightarrow \cdot E + T$

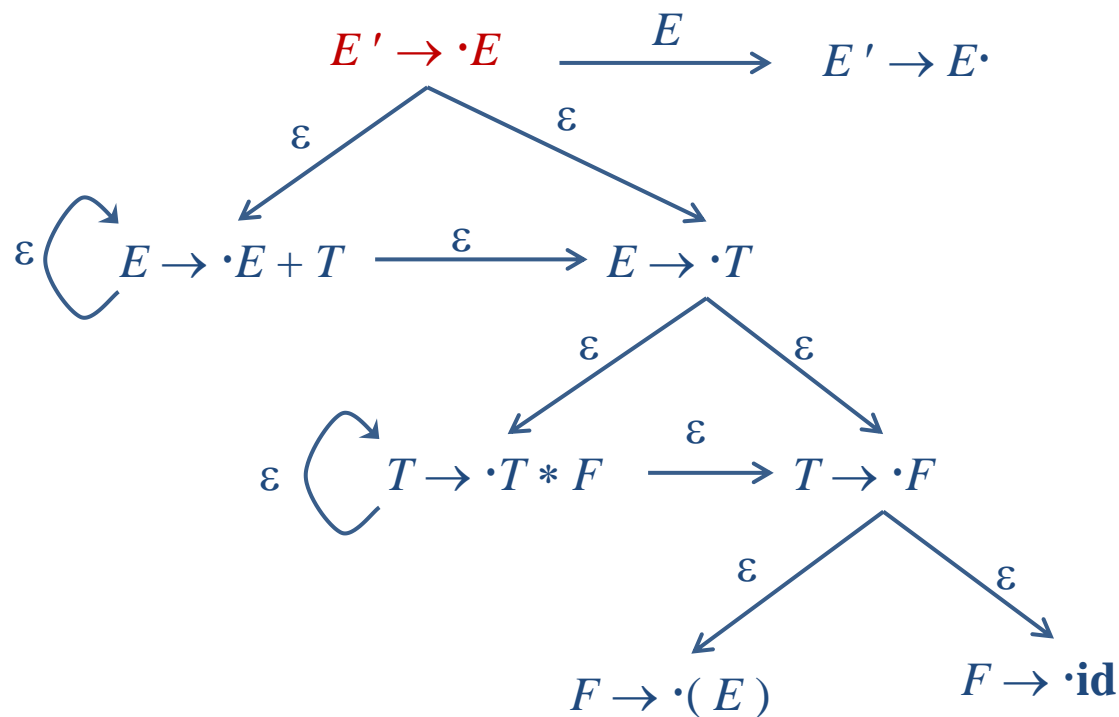
$E \rightarrow \cdot T$

$T \rightarrow \cdot T * F$

$T \rightarrow \cdot F$

$F \rightarrow \cdot (E)$

$F \rightarrow \cdot \text{id}$



构造SLR分析表

- 从文法构造的识别可行前缀的DFA的一些特点

- 概念：有效项目

如果 $S' \Rightarrow_{rm}^* \alpha A w \Rightarrow_{rm} \alpha \beta_1 \beta_2 w$ ，那么就说项目 $A \rightarrow \beta_1 \cdot \beta_2$ 对可行前缀 $\alpha \beta_1$ 是有效的

- 一个项目可能对好几个可行前缀都是有效的

$E \rightarrow \cdot E + T$	对 ϵ 和 (这两个活前缀都有效
$E' \Rightarrow E \Rightarrow E + T$	$(\alpha, \beta_1 \text{ 都为空})$
$E' \Rightarrow E \Rightarrow (E) \Rightarrow (E + T)$	$(\alpha = \text{"("}, \beta_1 \text{ 为空})$

该DFA读过 ϵ 和(后到达不同的状态，那么项目 $E \rightarrow \cdot E + T$ 就出现在对应的不同项目集中



构造SLR分析表

- 从文法构造的识别可行前缀的DFA的一些特点

- 概念：有效项目

如果 $S' \Rightarrow_{rm}^* \alpha A w \Rightarrow_{rm} \alpha \beta_1 \beta_2 w$ ，那么就说项目 $A \rightarrow \beta_1 \cdot \beta_2$ 对可行前缀 $\alpha \beta_1$ 是有效的

- 一个项目可能对好几个可行前缀都是有效的

- 从项目 $A \rightarrow \beta_1 \cdot \beta_2$ 对活前缀 $\alpha \beta_1$ 有效这个事实可以知道
 - 如果 $\beta_2 \neq \varepsilon$ ，应该移进
 - 如果 $\beta_2 = \varepsilon$ ，应该用产生式 $A \rightarrow \beta_1$ 归约



构造SLR分析表

- 从文法构造的识别可行前缀的DFA的一些特点

- 概念：有效项目

如果 $S' \Rightarrow_{rm}^* \alpha A w \Rightarrow_{rm} \alpha \beta_1 \beta_2 w$, 那么就说项目 $A \rightarrow \beta_1 \cdot \beta_2$ 对可行前缀 $\alpha \beta_1$ 是有效的

- 一个项目可能对好几个可行前缀都是有效的
 - 一个可行前缀可能有多个有效项目
 - 一个可行前缀 γ 的有效项目集就是从这个DFA的初态出发, 沿着标记为 γ 的路径到达的那个项目集 (状态)



构造SLR分析表

- 例，串 $E + T *$ 是可行前缀，读完它后，DFA处于状态 I_7

$$I_7: \quad T \rightarrow T * \cdot F, \quad F \rightarrow \cdot (E), \quad F \rightarrow \cdot \text{id}$$

$$\begin{aligned} E' &\Rightarrow E \\ &\Rightarrow E + T \\ &\Rightarrow E + T * F \\ &\Rightarrow E + T * \text{id} \\ &\Rightarrow E + T * F * \text{id} \end{aligned}$$

$$\begin{aligned} E' &\Rightarrow E \\ &\Rightarrow E + T \\ &\Rightarrow E + T * F \\ &\Rightarrow E + T * (E) \end{aligned}$$

$$\begin{aligned} E' &\Rightarrow E \\ &\Rightarrow E + T \\ &\Rightarrow E + T * F \\ &\Rightarrow E + T * \text{id} \end{aligned}$$



构造SLR分析表

- 从DFA构造SLR分析表

- 状态 i 从 I_i 构造，它的 action 函数如下确定：
 - 如果 $[A \rightarrow \alpha \cdot a \beta]$ 在 I_i 中，并且 $\text{goto}(I_i, a) = I_j$ ，那么置 $\text{action}[i, a]$ 为 sj
 - 如果 $[A \rightarrow \alpha \cdot]$ 在 I_i 中，那么对 $\text{FOLLOW}(A)$ 中的所有 a ，置 $\text{action}[i, a]$ 为 rj ， j 是产生式 $A \rightarrow \alpha$ 的编号
 - 如果 $[S' \rightarrow S \cdot]$ 在 I_i 中，那么置 $\text{action}[i, \$]$ 为接受 acc

如果出现动作冲突，那么该文法就不是SLR(1)的



构造SLR分析表

- 从DFA构造SLR分析表
 - 状态 i 从 I_i 构造，它的 action 函数如下确定：
 - 如果 $[A \rightarrow \alpha \cdot a \beta]$ 在 I_i 中，并且 $\text{goto}(I_i, a) = I_j$ ，那么置 $\text{action}[i, a]$ 为 sj
 - 如果 $[A \rightarrow \alpha \cdot]$ 在 I_i 中，那么对 $\text{FOLLOW}(A)$ 中的所有 a ，置 $\text{action}[i, a]$ 为 rj ， j 是产生式 $A \rightarrow \alpha$ 的编号
 - 如果 $[S' \rightarrow S \cdot]$ 在 I_i 中，那么置 $\text{action}[i, \$]$ 为接受 acc
 - 使用下面规则构造状态 i 的 goto 函数：
 - 对所有的非终结符 A ，如果 $\text{goto}(I_i, A) = I_j$ ，那么 $\text{goto}[i, A] = j$
 - 不能由上面两步定义的条目都置为 error
 - 分析器的初始状态是包含 $[S' \rightarrow \cdot S]$ 的项目集对应的状态



构造SLR分析表

- 从DFA构造SLR分析表

- 例, I_2 :

$$E \rightarrow T \cdot$$

$$T \rightarrow T \cdot * F$$

因为 $\text{FOLLOW}(E) = \{ \$, +,) \}$, 所以
 $\text{action}[2, \$] = \text{action}[2, +] = \text{action}[2,)] = r2$
 $\text{action}[2, *] = s7$

状态	动 作						转 移		
	id	+	*	()	\$	E	T	F
0	s5			s4			1	2	3
1		s6				acc			
<u>2</u>		r2	s7		r2	r2			
3		r4	r4		r4	r4			



构造SLR分析表

- 例，SLR(1)文法的描述能力有限
 - 每个SLR(1)文法都是无二义性的，但是存在很多不是SLR(1)的无二义性文法。

$S \rightarrow V = E$
 $S \rightarrow E$
 $V \rightarrow * E$
 $V \rightarrow \text{id}$
 $E \rightarrow V$

I_0 :
 $S' \rightarrow \cdot S$
 $S \rightarrow \cdot V = E$
 $S \rightarrow \cdot E$
 $V \rightarrow \cdot * E$
 $V \rightarrow \cdot \text{id}$
 $E \rightarrow \cdot V$

V

I_2 :
 $S \rightarrow V \cdot = E$
 $E \rightarrow V \cdot$

第一项目使得 $\text{action}[2, =] = s6$
第二项目使得 $\text{action}[2, =]$ 为按 $E \rightarrow V$
归约，因为 $=$ 是 E 的一个后继符

$=$ 是 E 的一个后继符：

$S \$ \Rightarrow V = E \$ \Rightarrow * E = E \$$



构造SLR分析表

- 例，SLR(1)文法的描述能力有限
 - 每个SLR(1)文法都是无二义性的，但是存在很多不是SLR(1)的无二义性文法。

$S \rightarrow V = E$
 $S \rightarrow E$
 $V \rightarrow * E$
 $V \rightarrow \text{id}$
 $E \rightarrow V$

$I_0:$
 $S' \rightarrow \cdot S$
 $S \rightarrow \cdot V = E$
 $S \rightarrow \cdot E$
 $V \rightarrow \cdot * E$
 $V \rightarrow \cdot \text{id}$
 $E \rightarrow \cdot V$

V

$I_2:$
 $S \rightarrow V \cdot = E$
 $E \rightarrow V \cdot$

第一项目使得 $\text{action}[2, =] = s6$
 第二项目使得 $\text{action}[2, =]$ 为按 $E \rightarrow V$
 归约，因为 $=$ 是 E 的一个后继符

在所关注场合 E 的后继是 $\$$:

$S \$ \Rightarrow V = E \$ \Rightarrow * E = E \$$
 $S \$ \Rightarrow E \$ \Rightarrow V \$$





谢谢！

Thanks!