

编译原理

第六章 LR语法分析技术 (2)

方徽星

扬州大学 信息工程学院(505)

fanghuixing@yzu.edu.cn

2018年春季学期

本章主要内容

一. 自下向上语法分析

二. LR分析

- SLR
- LR(1)
- LALR

三. 使用二义性文法

四. Yacc

2.3 构造SLR语法分析表

SLR语法分析表构造算法

- 输入：一个文法 G 的增广文法 G'
- 输出： G' 的SLR语法分析表函数 $ACTION$ 和 $GOTO$
- 方法：
 1. 构造 G' 的规范LR(0)项集族 $C = \{I_0, I_1, \dots, I_n\}$
 2. 根据 I_i 构造得到状态 i
 3. 如果 $GOTO(I_i, A_N) = I_j$ ，则 $GOTO[i, A_N] = j$
 4. 规则2和3没有定义的条目均设置为“报错”
 5. 初始状态为由 $[S' \rightarrow \bullet S]$ 所在项集构造所得的状态

该算法得到的语法分析表称为文法 G 的SLR(1)分析表

2.3 构造SLR语法分析表

- 根据 I_i 构造得到状态 i ，状态 i 的语法分析动作按照如下规则设定：
 - 如果 $[A \rightarrow \alpha \bullet a_T \beta] \in I_i \wedge GOTO(I_i, a_T) = I_j$ ，则 $ACTION[i, a_T] := \text{“移进状态}j\text{”}$
 - 如果 $[A \rightarrow \alpha \bullet] \in I_i \wedge A \neq S'$ ，则(由Follow实现向前看1个符号) $\forall a_T \in Follow(A). ACTION[i, a_T] := \text{“归约}A \rightarrow \alpha\text{”}$
 - 如果 $[S' \rightarrow S \bullet] \in I_i$ ，则 $ACTION[i, \$] := \text{“接受”}$

如果由上述规则导致冲突发生，则文法不是SLR(1)文法，此时该算法无法生成一个语法分析器

2.3 构造SLR语法分析表

- 与SLR相关的两个概念：
 - 使用 G 的SLR(1)分析表的LR语法分析器称为 G 的**SLR(1)语法分析器**
 - 一个具有SLR(1)分析表的文法称为**SLR(1)的**

**SLR后的(1)常可省略，因为此处
不会处理多于1个向前看符号**

2.3 构造SLR语法分析表

- 例：为增广表达式文法构造SLR分析表

$$1. E \rightarrow E + T$$

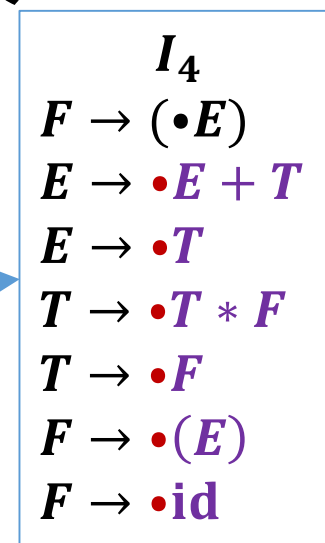
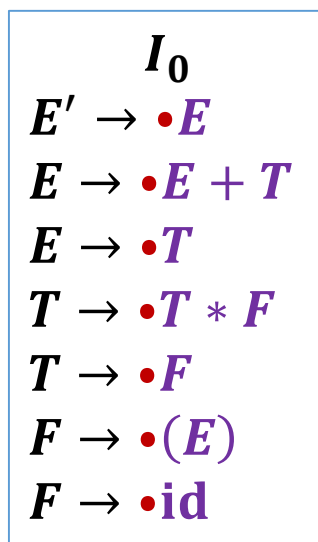
$$2. E \rightarrow T$$

$$3. T \rightarrow T * F$$

$$4. T \rightarrow F$$

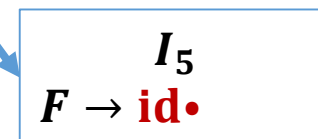
$$5. F \rightarrow (E)$$

$$6. F \rightarrow \mathbf{id}$$



(

id



- ✓ I_0 中的项 $F \rightarrow \bullet (E)$ 使得条目 $ACTION[0, (]$ = “移进状态4”
- ✓ I_0 中的项 $F \rightarrow \bullet \mathbf{id}$ 使得条目 $ACTION[0, \mathbf{id}]$ = “移进状态5”

2.3 构造SLR语法分析表

- 例：为增广表达式文法构造SLR分析表

1. $E \rightarrow E + T$

2. $E \rightarrow T$

3. $T \rightarrow T * F$

4. $T \rightarrow F$

5. $F \rightarrow (E)$

6. $F \rightarrow \text{id}$

$$\begin{array}{l} I_1 \\ E' \rightarrow E \bullet \\ E \rightarrow E \bullet + T \end{array}$$

+

\$

accept

$$\begin{array}{l} I_6 \\ E \rightarrow E + \bullet T \\ T \rightarrow \bullet T * F \\ T \rightarrow \bullet F \\ F \rightarrow \bullet (E) \\ F \rightarrow \bullet \text{id} \end{array}$$

- ✓ I_1 中的项 $E' \rightarrow E \bullet$ 使得 $ACTION[1, \$] = \text{“接受”}$
- ✓ I_1 中的项 $E \rightarrow E \bullet + T$ 使得 $ACTION[1, +] = \text{“移进状态6”}$

2.3 构造SLR语法分析表

- 例：为增广表达式文法构造SLR分析表

1. $E \rightarrow E + T$

2. $E \rightarrow T$

3. $T \rightarrow T * F$

4. $T \rightarrow F$

5. $F \rightarrow (E)$

6. $F \rightarrow \text{id}$

$$\begin{array}{c} I_2 \\ E \rightarrow T \bullet \\ T \rightarrow T \bullet * F \end{array}$$

*

$$\begin{array}{c} I_7 \\ T \rightarrow T * \bullet F \\ F \rightarrow \bullet (E) \\ F \rightarrow \bullet \text{id} \end{array}$$

✓ 因为 $\text{Follow}(E) = \{\$, +,)\}$, I_2 中的项 $E \rightarrow T \bullet$ 使得
 $\text{ACTION}[2, \$] = \text{ACTION}[2, +] = \text{ACTION}[2,)] =$
“归约 $E \rightarrow T$ ”

✓ I_2 中的项 $T \rightarrow T \bullet * F$ 使得 $\text{ACTION}[2, *] =$ “移进状态7”

2.3 构造SLR语法分析表

| 状态 | ACTION | | | | | | GOTO | | |
|----|--------|----|----|----|-----|-----|----------|----------|----------|
| | id | + | * | (|) | \$ | <i>E</i> | <i>T</i> | <i>F</i> |
| 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 | | | |

2.3 构造SLR语法分析表

- 例：每个SLR(1)文法都是无二义性的，但存在非SLR(1)的无二义性文法，如：

- $S \rightarrow L = R \mid R$
- $L \rightarrow * R \mid \text{id}$
- $R \rightarrow L$

I_2
 $S \rightarrow L \bullet = R$
 $R \rightarrow L \bullet$

=

I_6
 $S \rightarrow L = \bullet R$
 $R \rightarrow \bullet L$
 $L \rightarrow \bullet * R$
 $L \rightarrow \bullet \text{id}$

I_2 的第一个项 $S \rightarrow L \bullet = R$ 使得
 $\text{ACTION}[2, =] = \text{“移进状态6”}$

冲突

因为 $= \in \text{Follow}(R)$, I_2 的
第二个项 $R \rightarrow L \bullet$ 使得
 $\text{ACTION}[2, =] = \text{“归约 } R \rightarrow L \text{”}$

SLR分析器不够强大，具备的上下文信息不够

2.4 可行前缀(Viable Prefix)

- 给定文法 G ，如果存在如下的最右推导：

$$S \xRightarrow{*} \alpha N \omega \Rightarrow \alpha \beta_1 \beta_2 \omega, \quad \gamma = \alpha \beta_1 \in (V_N \cup V_T)^*$$

则称串 γ 是 G 的一个可行前缀

- 可以在可行前缀后面增加一些终结符号来得到一个最右句型

如从文法 G 的开始符号 S 进行推导，
可以得到串 α ：

$$S \xRightarrow{*} \alpha \in (V_N \cup V_T)^*$$

则称 α 为文法 G 的一个句型

2.4 可行前缀(Viable Prefix)

- 给定文法 G ，如果存在如下的最右推导：

$$S \xRightarrow{*} \alpha N \omega \Rightarrow \alpha \beta_1 \beta_2 \omega, \quad \gamma = \alpha \beta_1 \in (V_N \cup V_T)^*$$

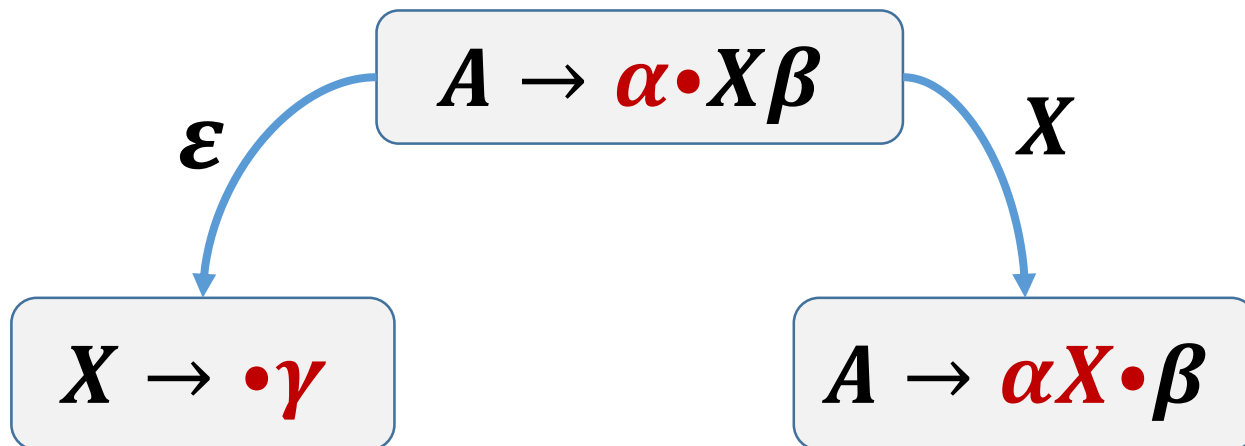
则LR(0)项 $[N \rightarrow \beta_1 \bullet \beta_2]$ 是串 γ 的一个有效项(Valid Item)

- 项 $[N \rightarrow \beta_1 \bullet \beta_2]$ 对可行前缀 $\alpha \beta_1$ 有效可以让我们知道：
 - 如果 $\beta_2 \neq \varepsilon$ ，则句柄没有完全入栈，因此选择移进
 - 如果 $\beta_2 = \varepsilon$ ，则 β_1 就是句柄，因此可以按照产生式 $N \rightarrow \beta_1$ 进行归约

可能存在多个不同的项要求对同一个可行前缀做不同的动作

2.4 可行前缀(Viable Prefix)

- 将项看作一个NFA N 的状态



$$CLOSURE(I) \equiv \epsilon\text{-closure}(T)$$

- I : 项集
- T : N 的状态集合
- $GOTO(I, X)$ 给出由NFA N 通过子集构造法得到的DFA中状态 I 在符号 X 上的转换

2.4 可行前缀(Viable Prefix)

- 例：考虑增广表达式文法，串 $\gamma = E + T *$ 是该文法的一个可行前缀，LR(0)自动机在读入 γ 后位于状态7上：

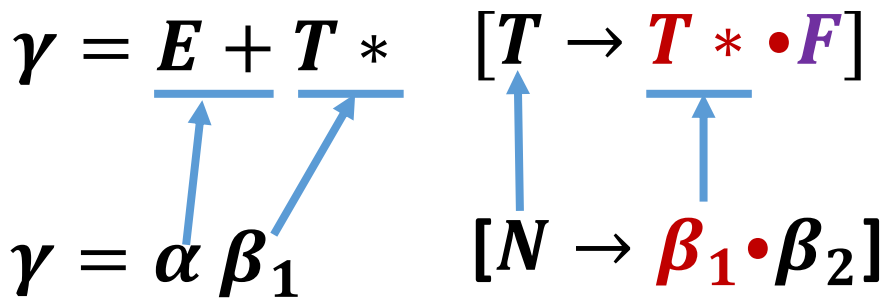
I_7
 $T \rightarrow T * \bullet F$
 $F \rightarrow \bullet (E)$
 $F \rightarrow \bullet id$

给定文法 G ，如果存在如下的最右推导：

$S \xRightarrow{*} \alpha N \omega \Rightarrow \alpha \beta_1 \beta_2 \omega$, $\gamma = \alpha \beta_1 \in (V_N \cup V_T)^*$
 则LR(0)项 $[N \rightarrow \beta_1 \bullet \beta_2]$ 是串 γ 的一个有效项

1.考虑最右推导

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

$\gamma = E + T *$ $[T \rightarrow T * \bullet F]$

 $\gamma = \alpha \beta_1$ $[N \rightarrow \beta_1 \bullet \beta_2]$
 $[T \rightarrow T * \bullet F]$ 是有效的！

2.4 可行前缀(Viable Prefix)

- 例：考虑增广表达式文法，串 $\gamma = E + T *$ 是该文法的一个可行前缀，LR(0)自动机在读入 γ 后位于状态7上：

I_7
 $T \rightarrow T * \bullet F$
 $F \rightarrow \bullet (E)$
 $F \rightarrow \bullet id$

给定文法 G ，如果存在如下的最右推导：

$S \xRightarrow{*} \alpha N \omega \Rightarrow \alpha \beta_1 \beta_2 \omega$, $\gamma = \alpha \beta_1 \in (V_N \cup V_T)^*$
则LR(0)项 $[N \rightarrow \beta_1 \bullet \beta_2]$ 是串 γ 的一个有效项

2.考虑最右推导

$E' \Rightarrow E$
 $\Rightarrow E + T$
 $\Rightarrow E + T * F$
 $\Rightarrow E + T * (E)$

$\gamma = E + T *$ $[F \rightarrow \bullet (E)]$

$\gamma = \alpha \beta_1$ $[N \rightarrow \beta_1 \bullet \beta_2]$

2.4 可行前缀(Viable Prefix)

- 例：考虑增广表达式文法，串 $\gamma = E + T *$ 是该文法的一个可行前缀，LR(0)自动机在读入 γ 后位于状态7上：

I_7
 $T \rightarrow T * \bullet F$
 $F \rightarrow \bullet (E)$
 $F \rightarrow \bullet id$

给定文法 G ，如果存在如下的最右推导：

$S \xRightarrow{*} \alpha N \omega \Rightarrow \alpha \beta_1 \beta_2 \omega$ ， $\gamma = \alpha \beta_1 \in (V_N \cup V_T)^*$
则LR(0)项 $[N \rightarrow \beta_1 \bullet \beta_2]$ 是串 γ 的一个有效项

2.考虑最右推导

$E' \Rightarrow E$
 $\Rightarrow E + T$
 $\Rightarrow E + T * F$
 $\Rightarrow E + T * (E)$

$\gamma = E + T * \epsilon$
 $\gamma = \alpha \beta_1$
 $[F \rightarrow \epsilon \bullet (E)]$
 $[N \rightarrow \beta_1 \bullet \beta_2]$
 $[F \rightarrow \bullet (E)]$ 是有效的！

2.4 可行前缀(Viable Prefix)

- 例：考虑增广表达式文法，串 $\gamma = E + T *$ 是该文法的一个可行前缀，LR(0)自动机在读入 γ 后位于状态7上：

I_7
 $T \rightarrow T * \bullet F$
 $F \rightarrow \bullet (E)$
 $F \rightarrow \bullet id$

给定文法 G ，如果存在如下的最右推导：

$S \xRightarrow{*} \alpha N \omega \Rightarrow \alpha \beta_1 \beta_2 \omega$ ， $\gamma = \alpha \beta_1 \in (V_N \cup V_T)^*$
则LR(0)项 $[N \rightarrow \beta_1 \bullet \beta_2]$ 是串 γ 的一个有效项

3.考虑最右推导

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

$\gamma = \underline{E + T} * \epsilon$
 $\gamma = \alpha \beta_1$
 $[F \rightarrow \epsilon \bullet id]$
 $[N \rightarrow \beta_1 \bullet \beta_2]$
 $[F \rightarrow \bullet id]$ 是有效的！

2.4 可行前缀(Viable Prefix)

- 例：考虑增广表达式文法，串 $\gamma = E + T *$ 是该文法的一个可行前缀，LR(0)自动机在读入 γ 后位于状态7上：

$$\begin{array}{c} I_7 \\ T \rightarrow T * \bullet F \\ F \rightarrow \bullet (E) \\ F \rightarrow \bullet id \end{array}$$

在LR(0)自动机中，从初始状态开始沿着标号为某个可行前缀 γ 的路径到达一个状态，则该状态对应的项集就是 γ 的有效项集

2.5 规范LR(1)项

- 例：重新考虑如下文法

- $S \rightarrow L = R \mid R$
- $L \rightarrow * R \mid \text{id}$
- $R \rightarrow L$

$$\begin{array}{c} I_2 \\ S \rightarrow L \bullet = R \\ R \rightarrow L \bullet \end{array}$$

当输入为符号 $=$ 时，有两种可能性：

1. 因为 $[R \rightarrow L \bullet] \in I_2$ ，所以按照 $R \rightarrow L$ 进行归约
2. 因为 $[S \rightarrow L \bullet = R] \in I_2$ ，所以执行移进动作

- 如果按照 $R \rightarrow L$ 进行归约则要求该文法能够产生以 $R = \dots$ 开头的最右句型
- 但该文法无法产生以 $R = \dots$ 开头的最右句型，因此状态2只和可行前缀 L 对应，不应该执行从 L 到 R 的归约

2.5 规范LR(1)项

- 如果在状态中包含更多信息，则可能可以排除不正确的归约
- 方法：对项可行精化，使它包含第二个分量，项的一般形式修改为

$$[A \rightarrow \alpha \bullet \beta, c]$$

第一个分量 第二个分量
(向前看符号)

其中 $A \rightarrow \alpha\beta$ 是一个产生式， c 是一个终结符号或 $\$$ ，上述项称为**LR(1)项**，1表示第二个分量的长度

2.5 规范LR(1)项

- 如果LR(1)项 $[A \rightarrow \alpha \bullet \beta, c]$ 中的 $\beta \neq \varepsilon$, 则向前看符号没有任何作用
- LR(1)项 $[A \rightarrow \alpha \bullet, c]$ 只在下一个输入符号等于 c 时才要求按照 $A \rightarrow \alpha$ 进行归约
 - 只有当栈顶状态中包含一个LR(1)项 $[A \rightarrow \alpha \bullet, c]$ 时, 才会在输入为 c 时按照 $A \rightarrow \alpha$ 进行归约
 - 这样的 c 的集合总是 $Follow(A)$ 的子集(可能是真子集)

2.5 规范LR(1)项

- LR(1)项 $[A \rightarrow \alpha \bullet \beta, c]$ 对于一个可行前缀 γ 有效的条件是存在一个最右推导

$$S \xRightarrow{*} \delta A \omega \Rightarrow \delta \alpha \beta \omega,$$

其中

- $\gamma = \delta \alpha$, 而且
- c 是 ω 的第一个符号, 或者 $\omega = \varepsilon \wedge c = \$$

由最右推导可知 ω 是一个仅由终结符号构成的串

2.5 规范LR(1)项

• 例：考虑如下文法

- $S \rightarrow BB$

- $B \rightarrow aB \mid b$

该文法有一个最右推导

$$\begin{array}{l}
 S \xRightarrow{*} \underline{aa} \underline{B} \underline{ab} \Rightarrow aaaBab \\
 S \xRightarrow{*} \delta \underline{A} \omega \Rightarrow \delta \underline{\alpha} \underline{\beta} \omega
 \end{array}$$

项 $[B \rightarrow a \bullet B, a]$ (a 是 $\omega = ab$ 的第一个符号) 对于可行前缀

$$\gamma = \delta \alpha = aa a$$

是有效的

2.5 规范LR(1)项

• 例：考虑如下文法

- $S \rightarrow BB$

- $B \rightarrow aB \mid b$

该文法还有一个最右推导

$$\begin{array}{c} S \xRightarrow{*} \underline{Ba}B \Rightarrow Ba\alpha B \\ \uparrow \quad \uparrow \\ S \xRightarrow{*} \delta A \omega \Rightarrow \delta \alpha \beta \omega \end{array}$$

项 $[B \rightarrow a \bullet B, \$]$ ($\omega = \varepsilon$) 对于可行前缀

$$\gamma = \delta \alpha = Ba\alpha$$

是有效的

2.6 构造LR(1)项集

- LR(1)项集族的构造算法
 - 输入：一个增广文法 G'
 - 输出：LR(1)项集族，其中的每个项集对文法 G' 的一个或多个可行前缀有效

① **SetOfItems CLOSURE(I) {**
 repeat
 for (I 中的每个项 $[A \rightarrow \alpha \bullet B \beta, a]$)
 for (G' 中的每个产生式 $B \rightarrow \gamma$)
 for ($First(\beta a)$ 中的每个终结符号 b)
 把 $[B \rightarrow \bullet \gamma, b]$ 放入 I 中；
 until 不能向 I 中加入更多的项；
 return I ;
}

2.6 构造LR(1)项集

- LR(1)项集族的构造算法
 - 输入：一个增广文法 G'
 - 输出：LR(1)项集族，其中的每个项集对文法 G' 的一个或多个可行前缀有效

②

```
SetOfItems GOTO( $I, X$ ) {  
     $J := \emptyset$  ;  
    for ( $I$ 中的每个项 $[A \rightarrow \alpha \bullet X \beta, a]$ )  
        将项 $[A \rightarrow \alpha X \bullet \beta, a]$ 加入到集合 $J$ 中 ;  
    return CLOSURE( $J$ )
```

2.6 构造LR(1)项集

- LR(1)项集族的构造算法
 - 输入：一个增广文法 G'
 - 输出：LR(1)项集族，其中的每个项集对文法 G' 的一个或多个可行前缀有效

③

```
void items( $G'$ ) {  
     $C := \{ \text{CLOSURE}(\{[S' \rightarrow \bullet S, \$])\} \}$  ;  
    repeat  
        for ( $C$ 中每个项集 $I$ )  
            for (每个文法符号)  
                if ( $\text{GOTO}(I, X) \neq \emptyset \wedge \text{GOTO}(I, X) \notin C$ )  
                    将 $\text{GOTO}(I, X)$ 加入 $C$ 中 ;  
    until 不在有新的项集加入到 $C$ 中 ;  
}
```

2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

首先计算 $\{[S' \rightarrow \bullet S, \$]\}$ 的闭包：

将 CC 归约为 S 之后
可遇见的
第一个终结符号

$$\begin{array}{c} [S' \rightarrow \epsilon \bullet S \epsilon, \$] \\ \uparrow \\ [A \rightarrow \alpha \bullet B \beta, a] \end{array}$$

对于每个产生式 $B \rightarrow \gamma$ 和
 $First(\beta a)$ 中的终结符号 b
将项 $[B \rightarrow \bullet \gamma, b]$ 加入闭包

$$\begin{array}{c} [S \rightarrow \bullet CC, \$] \\ \uparrow \\ [B \rightarrow \bullet \gamma, b] \end{array}$$

2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

首先计算 $\{[S' \rightarrow \bullet S, \$]\}$ 的闭包：

$$\begin{array}{c} I_0 \\ [S' \rightarrow \bullet S, \$] \\ [S \rightarrow \bullet CC, \$] \end{array}$$

再看由新加入的项 $[S \rightarrow \bullet CC, \$]$
能否引入更多的项？

2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

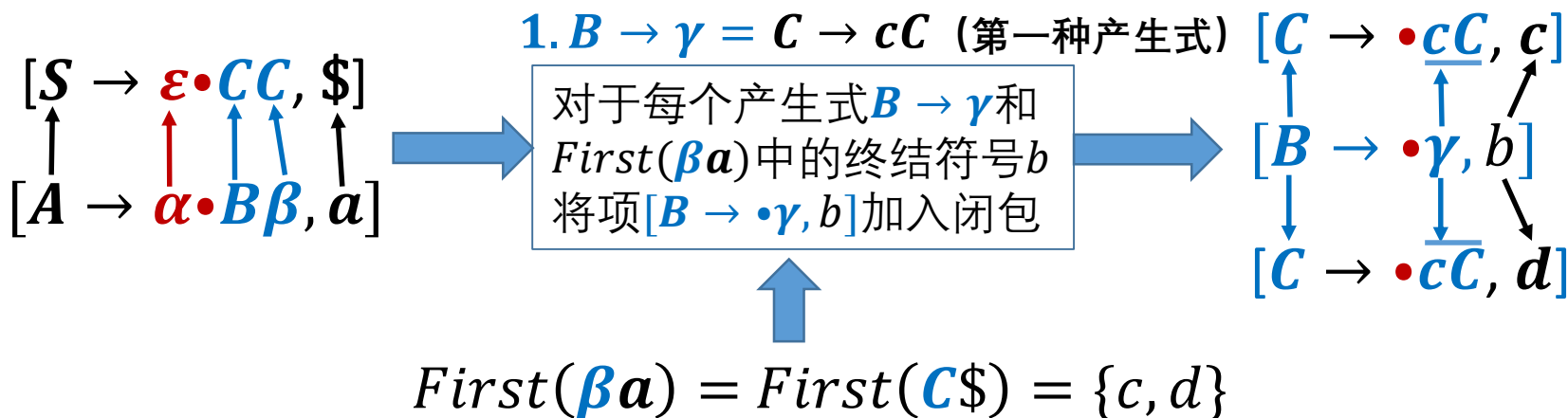
$$\bullet C \rightarrow cC \mid d$$

$$I_0$$

$$[S' \rightarrow \bullet S, \$]$$

$$[S \rightarrow \bullet CC, \$]$$

首先计算 $\{[S' \rightarrow \bullet S, \$]\}$ 的闭包：



2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

首先计算 $\{[S' \rightarrow \bullet S, \$]\}$ 的闭包：

$$\begin{array}{c} I_0 \\ [S' \rightarrow \bullet S, \$] \\ [S \rightarrow \bullet CC, \$] \\ [C \rightarrow \bullet cC, c] \\ [C \rightarrow \bullet cC, d] \end{array}$$

2.6 构造LR(1)项集

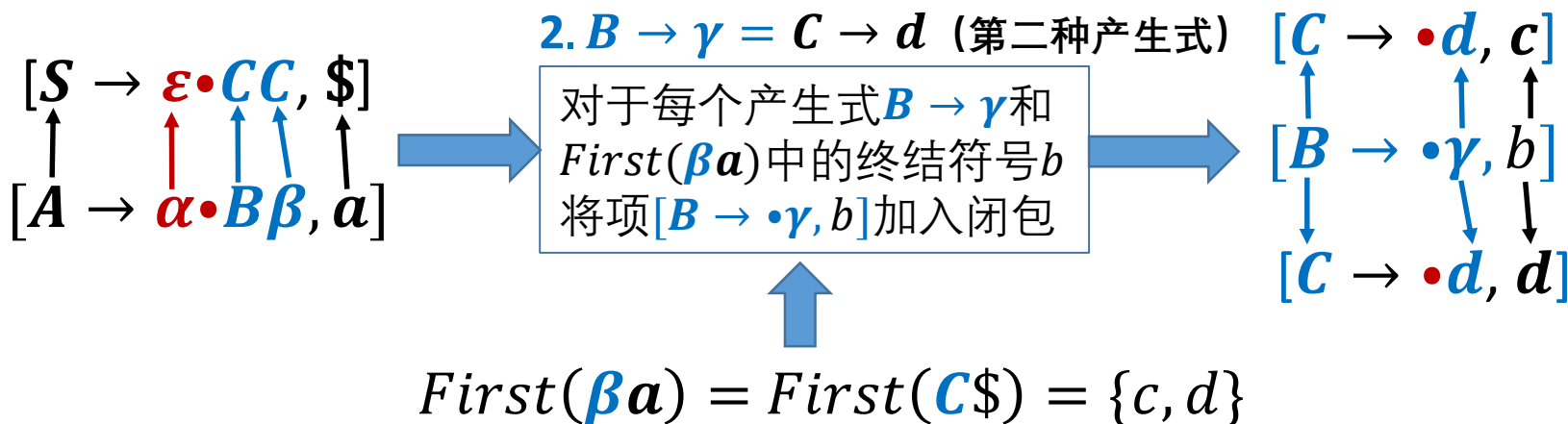
• 例：考虑增广文法：

- $S' \rightarrow S$
- $S \rightarrow CC$
- $C \rightarrow cC \mid d$

$$I_0$$

$$\begin{aligned} &[S' \rightarrow \bullet S, \$] \\ &[S \rightarrow \bullet CC, \$] \\ &[C \rightarrow \bullet cC, c] \\ &[C \rightarrow \bullet cC, d] \end{aligned}$$

首先计算 $\{[S' \rightarrow \bullet S, \$]\}$ 的闭包：



2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

首先计算 $\{[S' \rightarrow \bullet S, \$]\}$ 的闭包：

$$\begin{array}{l} I_0 \\ [S' \rightarrow \bullet S, \$] \\ [S \rightarrow \bullet CC, \$] \\ [C \rightarrow \bullet cC, c/d] \\ [C \rightarrow \bullet d, c/d] \end{array}$$

此时没有更多新项， I_0 已确定完成
下一步计算 $GOTO(I_0, X)$!

2.6 构造LR(1)项集

• 例：考虑增广文法：

• $S' \rightarrow S$

• $S \rightarrow CC$

• $C \rightarrow cC \mid d$

I_0
[$S' \rightarrow \bullet S, \$$]
[$S \rightarrow \bullet CC, \$$]
[$C \rightarrow \bullet cC, c/d$]
[$C \rightarrow \bullet d, c/d$]

对不同的文法符号 X 计算 $GOTO(I_0, X)$ ：

1. 对于 $X = S$ 求 $\{[S' \rightarrow S\bullet, \$]\}$ 的闭包，点 (\bullet) 后没有非终结符号，从而没有产生式可以处理，所以无法加入新的项，则

I_1
[$S' \rightarrow S\bullet, \$$]

2.6 构造LR(1)项集

• 例：考虑增广文法：

- $S' \rightarrow S$
- $S \rightarrow CC$
- $C \rightarrow cC \mid d$

$$I_0$$

$$[S' \rightarrow \bullet S, \$]$$

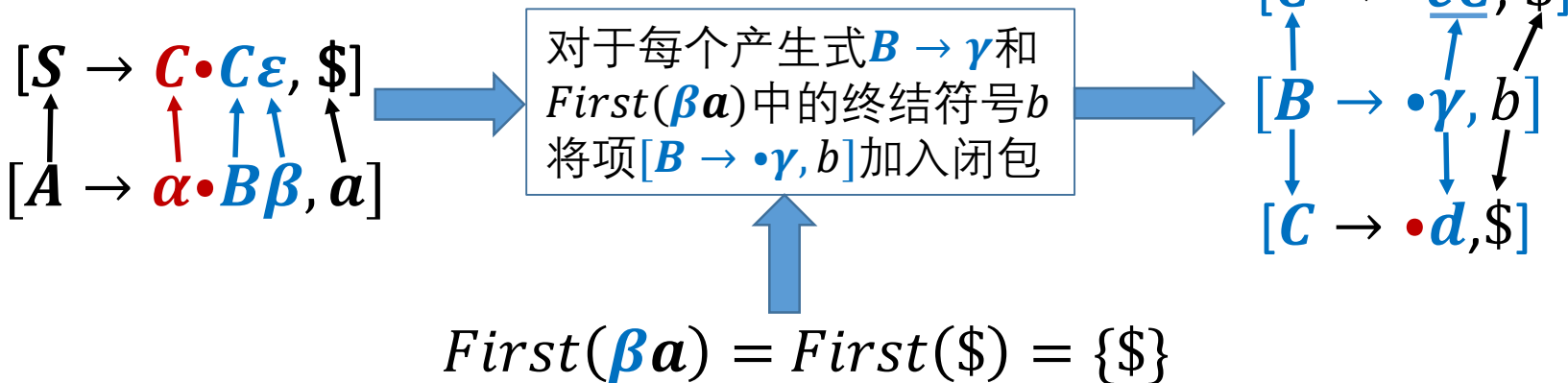
$$[S \rightarrow \bullet CC, \$]$$

$$[C \rightarrow \bullet cC, c/d]$$

$$[C \rightarrow \bullet d, c/d]$$

对不同的文法符号 X 计算 $GOTO(I_0, X)$ ：

2. 对于 $X = C$ 求 $\{[S \rightarrow \bullet C, \$]\}$ 的闭包



2.6 构造LR(1)项集

• 例：考虑增广文法：

• $S' \rightarrow S$

• $S \rightarrow CC$

• $C \rightarrow cC \mid d$

I_0
[$S' \rightarrow \bullet S, \$$]
[$S \rightarrow \bullet CC, \$$]
[$C \rightarrow \bullet cC, c/d$]
[$C \rightarrow \bullet d, c/d$]

对不同的文法符号 X 计算 $GOTO(I_0, X)$ ：

2. 对于 $X = C$ 求{[$S \rightarrow C \bullet C, \$$]}的闭包

I_2
[$S \rightarrow C \bullet C, \$$]
[$C \rightarrow \bullet cC, \$$]
[$C \rightarrow \bullet d, \$$]

2.6 构造LR(1)项集

• 例：考虑增广文法：

- $S' \rightarrow S$
- $S \rightarrow CC$
- $C \rightarrow cC \mid d$

$$I_0$$

$$[S' \rightarrow \bullet S, \$]$$

$$[S \rightarrow \bullet CC, \$]$$

$$[C \rightarrow \bullet cC, c/d]$$

$$[C \rightarrow \bullet d, c/d]$$

对不同的文法符号 X 计算 $GOTO(I_0, X)$ ：

3. 对于 $X = c$ 求 $\{[C \rightarrow \bullet cC, c/d]\}$ 的闭包

$$[C \rightarrow \bullet cC, c/d]$$

$$\uparrow$$

$$[A \rightarrow \alpha \bullet B \beta, a]$$

对于每个产生式 $B \rightarrow \gamma$ 和
 $First(\beta a)$ 中的终结符号 b
 将项 $[B \rightarrow \bullet \gamma, b]$ 加入闭包

$$[C \rightarrow \bullet cC, c/d]$$

$$\uparrow$$

$$[B \rightarrow \bullet \gamma, b]$$

$$\downarrow$$

$$[C \rightarrow \bullet d, c/d]$$

$$First(\beta a) = First(c/d) = \{c, d\}$$

2.6 构造LR(1)项集

• 例：考虑增广文法：

• $S' \rightarrow S$

• $S \rightarrow CC$

• $C \rightarrow cC \mid d$

对不同的文法符号 X 计算 $GOTO(I_0, X)$ ：

3. 对于 $X = c$ 求 $\{[C \rightarrow c \bullet C, c/d]\}$ 的闭包

I_0
[$S' \rightarrow \bullet S, \$$]
[$S \rightarrow \bullet CC, \$$]
[$C \rightarrow \bullet cC, c/d$]
[$C \rightarrow \bullet d, c/d$]

I_3
[$C \rightarrow c \bullet C, c/d$]
[$C \rightarrow \bullet cC, c/d$]
[$C \rightarrow \bullet d, c/d$]

2.6 构造LR(1)项集

• 例：考虑增广文法：

• $S' \rightarrow S$

• $S \rightarrow CC$

• $C \rightarrow cC \mid d$

I_0
[$S' \rightarrow \bullet S, \$$]
[$S \rightarrow \bullet CC, \$$]
[$C \rightarrow \bullet cC, c/d$]
[$C \rightarrow \bullet d, c/d$]

对不同的文法符号 X 计算 $GOTO(I_0, X)$ ：

4. 对于 $X = d$ 求 $\{[C \rightarrow d\bullet, c/d]\}$ 的闭包

[$C \rightarrow d\bullet\epsilon\epsilon, c/d$]

↑

[$A \rightarrow \alpha\bullet B\beta, a$]

对于每个产生式 $B \rightarrow \gamma$ 和
 $First(\beta a)$ 中的终结符号 b
将项 $[B \rightarrow \bullet\gamma, b]$ 加入闭包

不存在
以 ϵ 为头的产生式！

$First(\beta a) = First(c/d) = \{c, d\}$

2.6 构造LR(1)项集

• 例：考虑增广文法：

• $S' \rightarrow S$

• $S \rightarrow CC$

• $C \rightarrow cC \mid d$

对不同的文法符号 X 计算 $GOTO(I_0, X)$ ：

4. 对于 $X = d$ 求 $\{[C \rightarrow d\bullet, c/d]\}$ 的闭包

I_0
[$S' \rightarrow \bullet S, \$$]
[$S \rightarrow \bullet CC, \$$]
[$C \rightarrow \bullet cC, c/d$]
[$C \rightarrow \bullet d, c/d$]

I_4
[$C \rightarrow d\bullet, c/d$]

到此， I_0 上的 $GOTO$ 函数完成计算！

2.6 构造LR(1)项集

• 例：考虑增广文法：

- $S' \rightarrow S$

- $S \rightarrow CC$

- $C \rightarrow cC \mid d$

$$I_1$$
$$[S' \rightarrow S\bullet, \$]$$

对不同的文法符号 X 计算 $GOTO(I_1, X)$ ：

项 $[S' \rightarrow S\bullet, \$]$ 中红点(\bullet)在最后，
没有文法符号在点之后，所以
没有对应的 $GOTO$ 项集！

2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

$$\begin{array}{l} I_2 \\ [S \rightarrow \textcolor{red}{C} \bullet \textcolor{blue}{C}, \$] \\ [C \rightarrow \bullet \textcolor{blue}{c} \textcolor{blue}{C}, \$] \\ [C \rightarrow \bullet \textcolor{blue}{d}, \$] \end{array}$$

对不同的文法符号 X 计算 $GOTO(I_2, X)$ ：

1. 对于 $X = C$ 求 $\{[S \rightarrow \textcolor{red}{CC} \bullet, \$]\}$ 的闭包

$$\begin{array}{l} I_5 \\ [S \rightarrow \textcolor{red}{CC} \bullet, \$] \end{array}$$

点后没有非终结符号，所以闭包只包含 $[S \rightarrow \textcolor{red}{CC} \bullet, \$]$

2.6 构造LR(1)项集

• 例：考虑增广文法：

• $S' \rightarrow S$

• $S \rightarrow CC$

• $C \rightarrow cC \mid d$

I_2

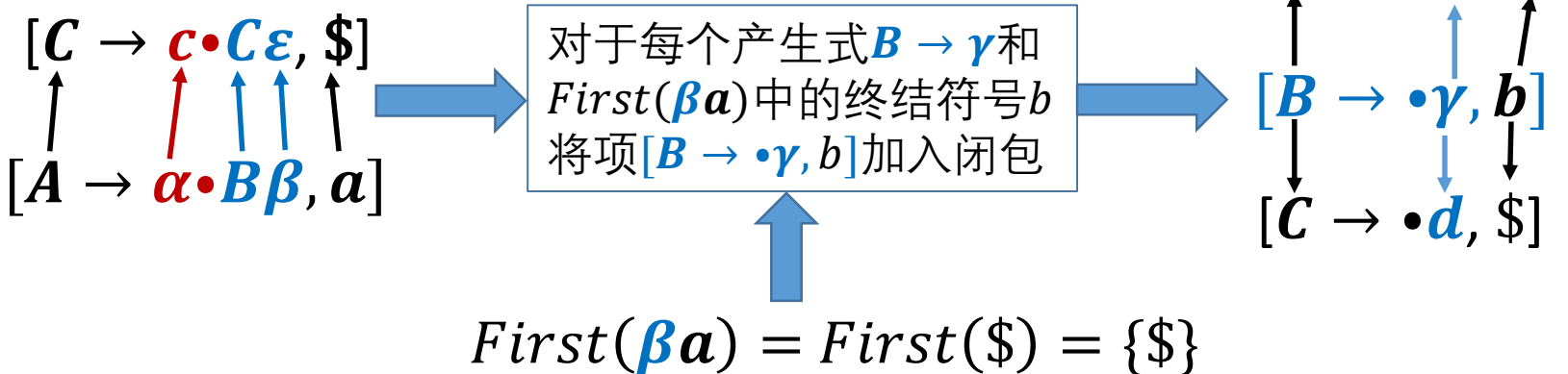
$[S \rightarrow \textcolor{red}{C} \bullet \textcolor{blue}{C}, \$]$

$[C \rightarrow \bullet \textcolor{blue}{c} \textcolor{blue}{C}, \$]$

$[C \rightarrow \bullet \textcolor{blue}{d}, \$]$

对不同的文法符号 X 计算 $GOTO(I_2, X)$ ：

2. 对于 $X = c$ 求 $\{[C \rightarrow \textcolor{red}{c} \bullet \textcolor{blue}{C}, \$]\}$ 的闭包



2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

$$\begin{array}{c} I_2 \\ [S \rightarrow \textcolor{red}{C} \bullet \textcolor{blue}{C}, \$] \\ [C \rightarrow \bullet \textcolor{blue}{c} \textcolor{blue}{C}, \$] \\ [C \rightarrow \bullet \textcolor{blue}{d}, \$] \end{array}$$

对不同的文法符号 X 计算 $GOTO(I_2, X)$ ：

2. 对于 $X = c$ 求 $\{[C \rightarrow \textcolor{red}{c} \bullet \textcolor{blue}{C}, \$]\}$ 的闭包

$$\begin{array}{c} I_6 \\ [C \rightarrow \textcolor{red}{c} \bullet \textcolor{blue}{C}, \$] \\ [C \rightarrow \bullet \textcolor{blue}{c} \textcolor{blue}{C}, \$] \\ [C \rightarrow \bullet \textcolor{blue}{d}, \$] \end{array}$$

2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

$$\begin{array}{l} I_2 \\ [S \rightarrow \textcolor{red}{C} \bullet \textcolor{blue}{C}, \$] \\ [C \rightarrow \bullet \textcolor{blue}{c} \textcolor{blue}{C}, \$] \\ [C \rightarrow \bullet \textcolor{blue}{d}, \$] \end{array}$$

对不同的文法符号 X 计算 $GOTO(I_2, X)$ ：

3. 对于 $X = d$ 求 $\{[C \rightarrow \textcolor{red}{d} \bullet, \$]\}$ 的闭包

$$\begin{array}{l} I_7 \\ [C \rightarrow \textcolor{red}{d} \bullet, \$] \end{array}$$

点后没有非终结符号，所以闭包只包含 $[C \rightarrow \textcolor{red}{d} \bullet, \$]$

2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

$$\begin{array}{c} I_3 \\ [C \rightarrow \textcolor{red}{c} \bullet \textcolor{blue}{C}, c/d] \\ [C \rightarrow \bullet \textcolor{blue}{c} C, c/d] \\ [C \rightarrow \bullet \textcolor{blue}{d}, c/d] \end{array}$$

对不同的文法符号 X 计算 $GOTO(I_3, X)$ ：

1. 对于 $X = C$ 求 $\{[C \rightarrow \textcolor{red}{c} C \bullet, c/d]\}$ 的闭包

$$\begin{array}{c} I_8 \\ [C \rightarrow \textcolor{red}{c} C \bullet, c/d] \end{array}$$

点后没有非终结符号，所以闭包只包含 $[C \rightarrow \textcolor{red}{c} C \bullet, c/d]$

2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

$$\begin{array}{c} I_3 \\ [C \rightarrow \textcolor{red}{c} \bullet \textcolor{blue}{C}, c/d] \\ [C \rightarrow \bullet \textcolor{blue}{c} C, c/d] \\ [C \rightarrow \bullet \textcolor{blue}{d}, c/d] \end{array}$$

对不同的文法符号 X 计算 $GOTO(I_3, X)$ ：

2. 对于 $X = c$ 求 $\{[C \rightarrow \textcolor{red}{c} \bullet \textcolor{blue}{C}, c/d]\}$ 的闭包

前面已经计算过：

$$\begin{array}{c} I_3 \\ [C \rightarrow \textcolor{red}{c} \bullet \textcolor{blue}{C}, c/d] \\ [C \rightarrow \bullet \textcolor{blue}{c} C, c/d] \\ [C \rightarrow \bullet \textcolor{blue}{d}, c/d] \end{array}$$

2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

$$\begin{array}{c} I_3 \\ [C \rightarrow \textcolor{red}{c}\bullet\textcolor{blue}{C}, c/d] \\ [C \rightarrow \bullet\textcolor{blue}{c}C, c/d] \\ [C \rightarrow \bullet\textcolor{blue}{d}, c/d] \end{array}$$

对不同的文法符号 X 计算 $GOTO(I_3, X)$ ：

3. 对于 $X = d$ 求 $\{[C \rightarrow \textcolor{red}{d}\bullet, c/d]\}$ 的闭包

前面已经计算过：

$$\begin{array}{c} I_4 \\ [C \rightarrow \textcolor{red}{d}\bullet, c/d] \end{array}$$

2.6 构造LR(1)项集

• 例：考虑增广文法：

- $S' \rightarrow S$

- $S \rightarrow CC$

- $C \rightarrow cC \mid d$

$$I_4$$
$$[C \rightarrow d\bullet, c/d]$$

对不同的文法符号 X 计算 $GOTO(I_4, X)$ ：

$[C \rightarrow d\bullet, c/d]$ 中点后没有非终结符号，
因此没有相应的 $GOTO$ 项集

2.6 构造LR(1)项集

• 例：考虑增广文法：

- $S' \rightarrow S$

- $S \rightarrow CC$

- $C \rightarrow cC \mid d$

$$I_5$$
$$[S \rightarrow CC\bullet, \$]$$

对不同的文法符号 X 计算 $GOTO(I_5, X)$ ：

$[S \rightarrow CC\bullet, \$]$ 中点后没有非终结符号，
因此没有相应的 $GOTO$ 项集

2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

$$\begin{array}{c} I_6 \\ [C \rightarrow \textcolor{red}{c} \bullet \textcolor{blue}{C}, \$] \\ [C \rightarrow \bullet \textcolor{blue}{c} C, \$] \\ [C \rightarrow \bullet \textcolor{blue}{d}, \$] \end{array}$$

对不同的文法符号 X 计算 $GOTO(I_6, X)$ ：

1. 对于 $X = C$ 求 $\{[C \rightarrow \textcolor{red}{c} C \bullet, \$]\}$ 的闭包

$$\begin{array}{c} I_9 \\ [C \rightarrow \textcolor{red}{c} C \bullet, \$] \end{array}$$

点后没有非终结符号，所以闭包只包含 $[C \rightarrow \textcolor{red}{c} C \bullet, \$]$

2.6 构造LR(1)项集

• 例：考虑增广文法：

• $S' \rightarrow S$

• $S \rightarrow CC$

• $C \rightarrow cC \mid d$

I_6
[$C \rightarrow c \bullet C, \$$]
[$C \rightarrow \bullet cC, \$$]
[$C \rightarrow \bullet d, \$$]

对不同的文法符号 X 计算 $GOTO(I_6, X)$ ：

2. 对于 $X = c$ 求{[$C \rightarrow c \bullet C, \$$]}的闭包

前面已经计算过：

I_6
[$C \rightarrow c \bullet C, \$$]
[$C \rightarrow \bullet cC, \$$]
[$C \rightarrow \bullet d, \$$]

2.6 构造LR(1)项集

• 例：考虑增广文法：

$$\bullet S' \rightarrow S$$

$$\bullet S \rightarrow CC$$

$$\bullet C \rightarrow cC \mid d$$

$$\begin{array}{c} I_6 \\ [C \rightarrow \textcolor{red}{c} \bullet \textcolor{blue}{C}, \$] \\ [C \rightarrow \bullet \textcolor{blue}{c} C, \$] \\ [C \rightarrow \bullet \textcolor{blue}{d}, \$] \end{array}$$

对不同的文法符号 X 计算 $GOTO(I_6, X)$ ：

3. 对于 $X = d$ 求 $\{[C \rightarrow \textcolor{red}{d} \bullet, \$]\}$ 的闭包

前面已经计算过：

$$I_7$$
$$[C \rightarrow \textcolor{red}{d} \bullet, \$]$$

2.6 构造LR(1)项集

• 例：考虑增广文法：

- $S' \rightarrow S$

- $S \rightarrow CC$

- $C \rightarrow cC \mid d$

$$I_7$$
$$[C \rightarrow d\bullet, \$]$$

对不同的文法符号 X 计算 $GOTO(I_7, X)$ ：

没有对应的 $GOTO$ 项集！

2.6 构造LR(1)项集

• 例：考虑增广文法：

- $S' \rightarrow S$

- $S \rightarrow CC$

- $C \rightarrow cC \mid d$

$$I_8$$
$$[C \rightarrow \textcolor{red}{cC}\bullet, c/d]$$

对不同的文法符号 X 计算 $GOTO(I_8, X)$ ：

没有对应的 $GOTO$ 项集！

2.6 构造LR(1)项集

• 例：考虑增广文法：

- $S' \rightarrow S$

- $S \rightarrow CC$

- $C \rightarrow cC \mid d$

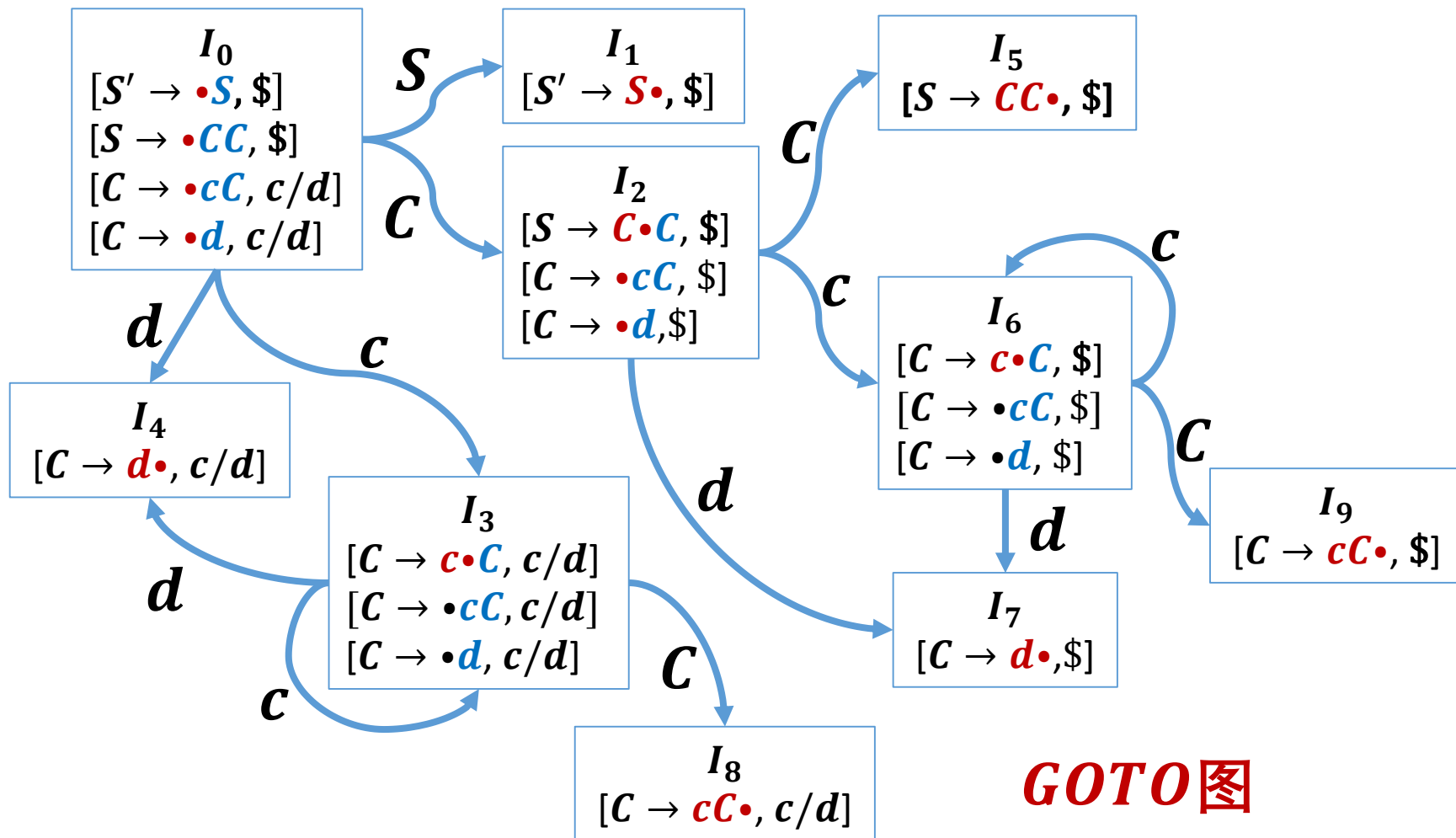
$$I_9$$
$$[C \rightarrow \textcolor{red}{c}C\bullet, \$]$$

对不同的文法符号 X 计算 $GOTO(I_9, X)$ ：

没有对应的 $GOTO$ 项集！

终于完成了所以项集计算！！！！

2.6 构造LR(1)项集



2.7 规范LR(1)语法分析表

- 规范LR(1)语法分析表构造算法
 - 输入：一个增广文法 G'
 - 输出： G' 的规范LR语法分析表的函数 $ACTION$ 和 $GOTO$
 - 方法：
 1. 构造 G' 的**LR(1)**项集族 $C' = \{I_0, I_1, \dots, I_n\}$
 2. 语法分析器的状态 i 根据 I_i 构造得到
 3. 如果 **$GOTO(I_i, A_N) = I_j$** ，则 **$GOTO[i, A_N] = j$**
 4. 规则2和3没有定义的条目均设置为“报错”
 5. 初始状态为由 $[S' \rightarrow \bullet S, \$]$ 所在项集构造所得的状态

2.7 规范LR(1)语法分析表

- 根据 I_i 构造得到状态 i ，状态 i 的语法分析动作按照如下规则设定：
 - 如果 $[A \rightarrow \alpha \bullet a_T \beta, b] \in I_i \wedge GOTO(I_i, a_T) = I_j$ ，则
 $ACTION[i, a_T] := \text{“移进状态}j\text{”}$
 - 如果 $[A \rightarrow \alpha \bullet, a] \in I_i \wedge A \neq S'$ ，则
 $ACTION[i, a] := \text{“归约}A \rightarrow \alpha\text{”}$
 - 如果 $[S' \rightarrow S \bullet, \$] \in I_i$ ，则
 $ACTION[i, \$] := \text{“接受”}$

如果由上述规则导致冲突发生，则文法不是LR(1)文法，此时该算法无法生成一个语法分析器

2.7 规范LR(1)语法分析表

- 例：考虑文法：

1. $S \rightarrow CC$

2. $C \rightarrow cC$

3. $C \rightarrow d$

| 状态 | ACTION | | | GOTO | |
|----|--------|-----|------|------|-----|
| | c | d | $\$$ | S | C |
| 0 | s3 | s4 | | 1 | 2 |
| 1 | | | acc | | |
| 2 | s6 | s7 | | | 5 |
| 3 | s3 | s4 | | | 8 |
| 4 | r3 | r4 | | | |
| 5 | | | r1 | | |
| 6 | s6 | s7 | | | 9 |
| 7 | | | r3 | | |
| 8 | r2 | r2 | | | |
| 9 | | | r2 | | |