Lecture 5

SLR文法

徐辉 xuh@fudan.edu.cn



主要内容

- 一、SLR文法
- 二、更多文法

一、SLR文法

自底向上解析

```
语法规则:
                              解析对象: 1+2*-3
[1] E \rightarrow E OP1 E1
                              标签流:
                                         <UNUM><ADD><UNUM><MUL><SUB><UNUM>
[2]
       | E1
    E1 \rightarrow E1 OP2 E2
                                                16:[1]E
[4]
           E2
[5] E2 \rightarrow E3 OP3 E2
                                                OP1<sub>6:[11]</sub>
                                                                  E1 <sub>15:[3]</sub>
                                   5:[2] E
[6]
           E3
[7] E3 \rightarrow NUM
                                               <ADD> 10:[4]E1
                                   4:[4] E1
                                                                 OP2
                                                                          E2 14:[6]
                                                                    11:[13]
     [9] NUM \rightarrow \langle UNUM \rangle
                                                      9:[6] E2 <MUL>
                                   3:[6] E2
                                                                          E3 13:[7]
      <SUB> <UNUM>
                                                                         NUM<sub>12:[10]</sub>
[11] OP1 \rightarrow <ADD>
                                   2:[7] E3
                                                      8:[7] E3
       <SUB>
                                                                     <SUB> <UNUM>
                                   1:[9] NUM
                                                      7:[9] NUM
[13] OP2 → <MUL>
[14] | <DIV>
                                      <UNUM>
                                                         <UNUM>
[15] OP3 → <POW>
```

如何自动规约生成语法解析树?

- 两种操作:
 - 移进: 读入下一个字符
 - 规约:应用语法规则规约已读入字符
- 根据句柄状态应用语法规则(从左至右)逐步规约
- 如何精准选择移进或规约? 避免盲目搜索
- SLR(1)文法: Simple Left-to-Right, Rightmost, 前瞻一个字符
- SLR文法的基本要求: 无回溯
 - 同一个状态既可以移进,又可以规约
 - 同一个状态存在两个规约选项

句柄状态(规范项)分析

- 句柄: 语法规则中已解析的字符, 一般保存在栈上
- 每条语法规则 $X \to \beta$ 可能存在 $|\beta| + 1$ 种句柄状态

```
[1] E \rightarrow E OP1 E1
    | E1
[3] E1 \rightarrow E1 OP2 E2
     | E2
[4]
[5] E2 \rightarrow E3 OP3 E2
[6] | E3
[7] E3 \rightarrow NUM
[8] | <LPAR> E <RPAR>
[9] NUM \rightarrow \langle UNUM \rangle
[10] | <SUB> <UNUM>
[11] OP1 \rightarrow <ADD>
[12] | <SUB>
[13] OP2 → <MUL>
[14] | <DIV>
[15] OP3 → <POW>
```

```
句柄分析
```

```
[1] E \rightarrow \circ E OP1 E1
      \mathsf{E} \to \mathsf{E} \circ \mathsf{OP1} \mathsf{E1}
      E \rightarrow E OP1 \circ E1
      E \rightarrow E OP1 E1 \circ
       E \rightarrow \circ E1
[2] E \rightarrow E1 \circ
[3] E1 \rightarrow \circ E1 OP2 E2
[3] E1 \rightarrow E1 \circ OP2 E2
[3] E1 \rightarrow E1 OP2 \circ E2
[3] E1 \rightarrow E1 OP2 E2 \circ
```

语法增强

- 加入一个初始符号和初始规则,初始符号仅在初始规则出现
- 句柄为初始符号So,并且后一个字符是eof则解析结束

```
[1] E → ∘ E OP1 E1

[1] E → E ∘ OP1 E1

[1] E → E OP1 ∘ E1

[1] E → E OP1 E1 ∘

[2] E → ∘ E1

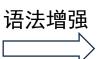
[2] E → E1 ∘

[3] E1 → ∘ E1 OP2 E2

[3] E1 → E1 ∘ OP2 E2

[3] E1 → E1 OP2 ∘ E2

[3] E1 → E1 OP2 ∘ E2
```



```
[0] S \rightarrow \circ E
       E \rightarrow \circ E OP1 E1
       E \rightarrow E \circ OP1 E1
       E \rightarrow E OP1 \circ E1
       E \rightarrow E OP1 E1 \circ
       E \rightarrow \circ E1
[2] E \rightarrow E1 \circ
[3] E1 \rightarrow \circ E1 OP2 E2
       E1 \rightarrow E1 \circ OP2 E2
       E1 → E1 OP2 ∘ E2
[3] E1 \rightarrow E1 OP2 E2 \circ
```

构建LR(0)自动机: 规范族

• 一条规范项可能由多种规范项直接或间接规约得到

```
[0] S → ∘ E

[1] E → ∘ E OP1 E1

[2] E → ∘ E1

···
```

```
分析规范族
```

While (Q has changed) //仅包含当前规范项 for each item $[A \to \beta \circ C\delta] \in Q$ for each production $[C \to \lambda] \in G$ if $[C \to \circ \lambda] \notin Q$ $Q \leftarrow Q \cup [C \to \circ \lambda]$

```
[0] S \rightarrow \circ E
[1] E \rightarrow E OP1 E1
[2] | E1
[3] E1 \rightarrow E1 OP2 E2
[4]
     | E2
[5] E2 \rightarrow E3 OP3 E2
[6]
    | E3
[7] E3 → NUM
[8] | <LPAR> E <RPAR>
[9] NUM \rightarrow \langle UNUM \rangle
[10] | <SUB> <UNUM>
[11] OP1 → <ADD>
[12] | <SUB>
[13] OP2 → <MUL>
      <DIV>
[14]
[15] OP3 → <POW>
```

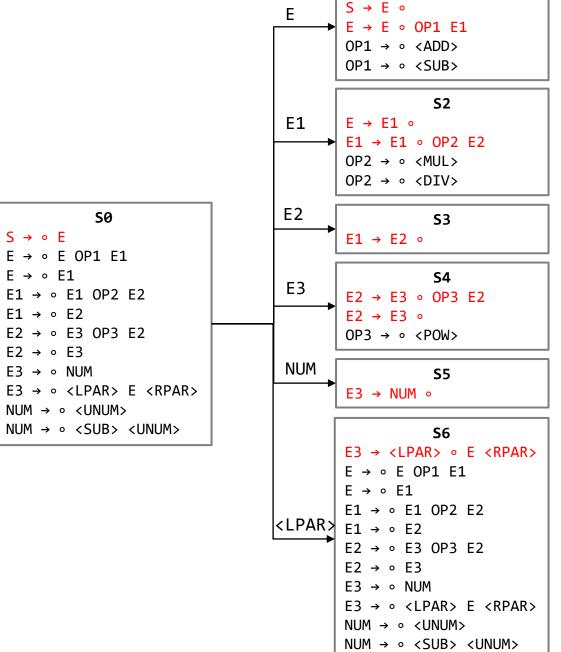
```
应用
                                S0
                S \rightarrow \circ E
               F \rightarrow \circ F \cap P1 = F1
               \mathsf{E} \to \mathsf{o} \mathsf{E} \mathsf{1}
分析规范族
                E1 \rightarrow \circ E1 OP2 E2
                E1 \rightarrow \circ E2
                E2 → ∘ E3 OP3 E2
               E2 → ∘ E3
                E3 → • NUM
                E3 → ∘ <LPAR> E <RPAR>
               NUM → ∘ <UNUM>
               NUM → ∘ <SUB> <UNUM>
```

核心项

非核心项

构建LR(0)自动机

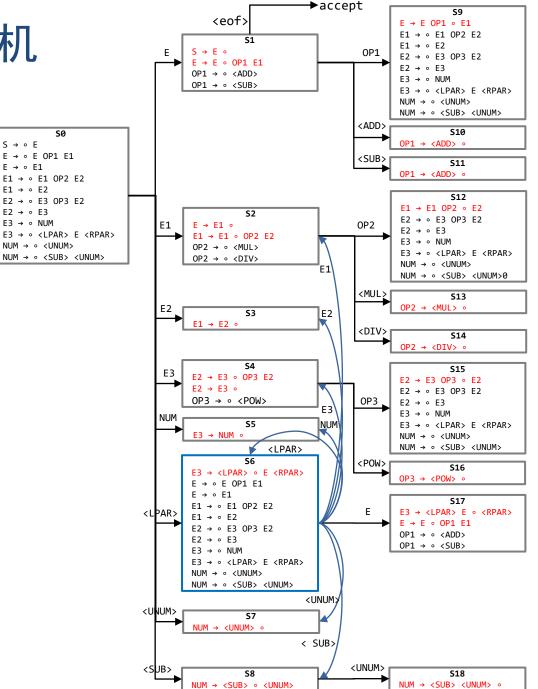
```
[0] S \rightarrow E
[1] E \rightarrow E OP1 E1
[2] | E1
[3] E1 \rightarrow E1 OP2 E2
      | E2
[4]
[5] E2 → E3 OP3 E2
[6]
       | E3
[7] E3 \rightarrow NUM
[8]
    [9] NUM \rightarrow \langle UNUM \rangle
[10]
      [11] OP1 → <ADD>
[12]
          <SUB>
[13] OP2 → <MUL>
[14] | <DIV>
[15] OP3 → <POW>
```

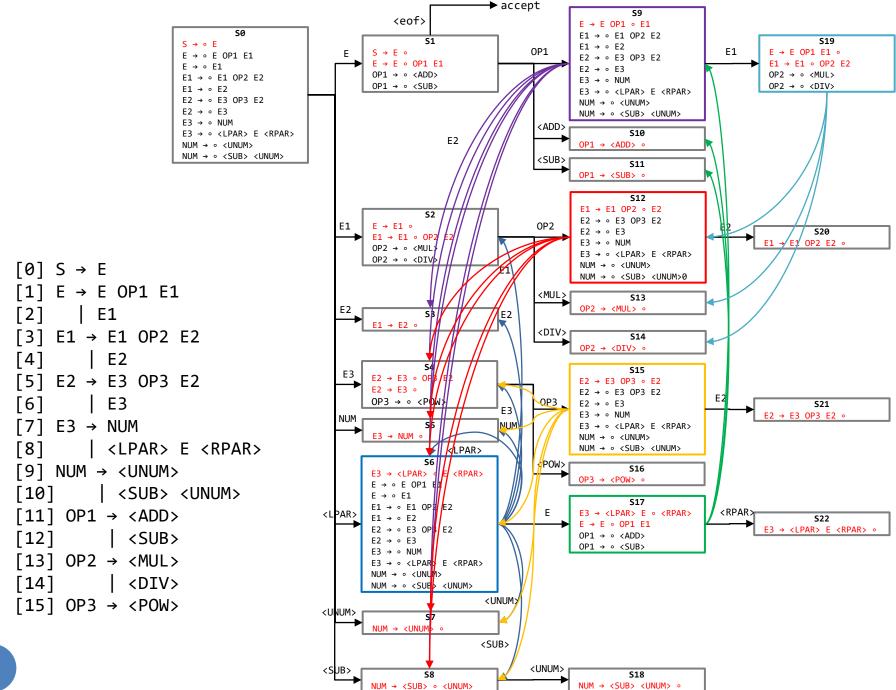


S1

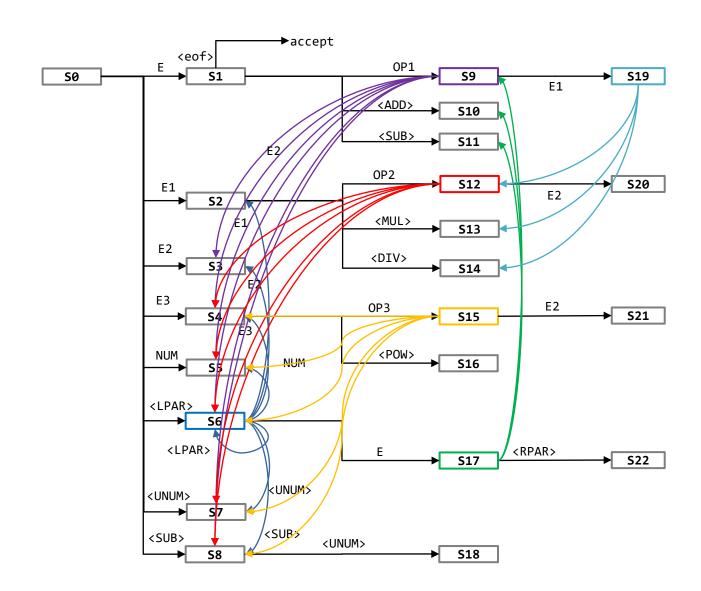
构建LR(0)自动机

```
[0] S \rightarrow E
[1] E \rightarrow E OP1 E1
[2]
    | E1
[3] E1 \rightarrow E1 OP2 E2
    | E2
[4]
[5] E2 → E3 OP3 E2
[6]
     l E3
[7] E3 → NUM
    | <LPAR> E <RPAR>
[9] NUM \rightarrow \langle UNUM \rangle
[10] | <SUB> <UNUM>
[11] OP1 → <ADD>
[12]
     <SUB>
[13] OP2 → <MUL>
[14] | <DIV>
[15] OP3 → <POW>
```



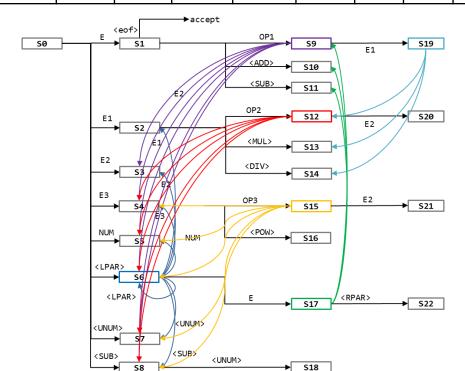


LR(0)自动机: 状态转移关系



LR(0)自动机的状态转移关系表

| 规范族 | E | E1 | E2 | E3 | OP1 | OP2 | 0P3 | NUM | <unum></unum> | <add></add> | | <mul></mul> | <div></div> | <pow></pow> | <lp></lp> | <rp></rp> | <eof></eof> |
|-------|-----------|-----------|----|-----------|-----|-----|-----|-----|---------------|-------------|-------------|-------------|-------------|-------------|-----------|-----------|-------------|
| SØ | S1 | S2 | S3 | S4 | | | | S5 | S7 | | S8 | | | | S6 | | |
| S1 | | | | | S9 | | | | | S10 | S11 | | | | | | accept |
| S2 | | | | | | S12 | | | | | | S13 | S14 | | | | |
| S3 | | | | | | | | | | | | | | | | | |
| S4 | | | | | | | S15 | | | | | | | S16 | | | |
| S5 | | | | | | | | | | | | | | | | | |
| S6 | S17 | S2 | S3 | S4 | | | | S5 | S7 | | S8 | | | | S6 | | |
| • • • | | | | | | | | | | | | | | | | | |
| S22 | | | | | | | | | | | | | | | | | |

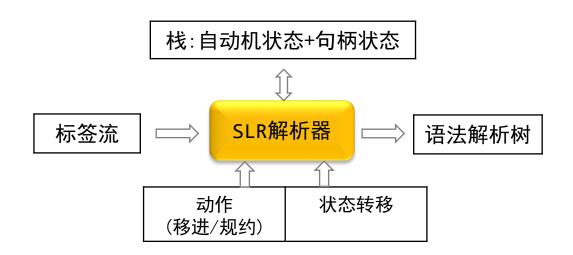


LR(0)自动机的状态转移关系表

| | , | , | | | | | | | | | | | | | | | |
|------------|-----------|------------|------------|-----------|------------|-----|-----|------------|---------------|-------------|-------------|-------------|-------------|-------------|------------|-----------|-------------|
| 规范族 | E | E1 | E2 | E3 | 0P1 | OP2 | 0P3 | NUM | <unum></unum> | <add></add> | | <mul></mul> | <div></div> | <pow></pow> | <lp></lp> | <rp></rp> | <eof></eof> |
| SØ | S1 | S2 | S 3 | S4 | | | | S 5 | S7 | | S8 | | | | S 6 | | |
| S1 | | | | | S9 | | | | | S10 | S11 | | | | | | accept |
| S2 | | | | | | S12 | | | | | | S13 | S14 | | | | |
| S 3 | | | | | | | | | | | | | | | | | |
| S4 | | | | | | | S15 | | | | | | | S16 | | | |
| S5 | | | | | | | | | | | | | | | | | |
| S6 | S17 | S2 | S 3 | S4 | | | | S5 | S7 | | S8 | | | | S6 | | |
| S7 | | | | | | | | | | | | | | | | | |
| S8 | | | | | | | | | S18 | | | | | | | | |
| S9 | | S19 | S 3 | S4 | | | | S5 | S7 | | S8 | | | | S6 | | |
| S10 | | | | | | | | | | | | | | | | | |
| S11 | | | | | | | | | | | | | | | | | |
| S12 | | | S20 | S4 | | | | S5 | S7 | | S8 | | | | S6 | | |
| S13 | | | | | | | | | | | | | | | | | |
| S14 | | | | | | | | | | | | | | | | | |
| S15 | | | S21 | S4 | | | | S5 | S7 | | S8 | | | | S6 | | |
| S16 | | | | | | | | | | | | | | | | | |
| S17 | | | | | S 9 | | | | | S10 | S11 | | | | | S22 | |
| S18 | | | | | | | | | | | | | | | | | |
| S19 | | | | | | S12 | | | | | | S13 | S14 | | | | |
| S20 | | | | | | | | | | | | | | | | | |
| S21 | | | | | | | | | | | | | | | | | |
| S22 | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |

构建SLR解析器

- 移进条件: 如果 $A \to \alpha \circ \alpha \beta \in S_i$,并且 $Goto(S_i, \alpha) = S_j$,设置 $Action(S_i, \alpha) = "Shift j"$
- 规约条件: 如果 $A \to \alpha \circ \in S_i$, $\forall a \in Follow(A)$, 设置 $Action(S_i, a) =$ " $Reduce\ A \to \alpha$ "



SLR解析表

| ₩₩₩ | | | | G | бото | | | | Action (Shift-Reduce) | | | | | | | | | |
|-----|-----|---------------|----------|---------|-------|--|--------|-----|-----------------------|-------------|-------------|-------------|-------------|-------------|-----------|-----------|-------------|--|
| 规范族 | Е | E1 | E2 | E3 | 0P1 | OP2 | 0P3 | NUM | <unum></unum> | <add></add> | | <mul></mul> | <div></div> | <pow></pow> | <lp></lp> | <rp></rp> | <eof></eof> | |
| Sø | S1 | S2 | S3 | S4 | | | | S5 | S7 | | 58 | | | | S6 | | | |
| S1 | | 52. | F | E1 • | ça | | | | | S10 | S11 | | | | | | acc | |
| S2 | | | | → E2 · | | S12 | | | | R[2] | R[2] | S13 | S14 | | | R[2] | R[2] | |
| S3 | | | | → E3 | | | | | | R[4] | R[4] | R[4] | R[4] | | | R[4] | R[4] | |
| S4 | | S5: | E3 | → NUM | 0 | | S15 | | | R[6] | R[6] | R[6] | R[6] | S16 | | R[6] | R[6] | |
| S5 | | [0] | S - | → E | | | | | | R[7] | R[7] | R[7] | R[7] | R[7] | | R[7] | R[7] | |
| S6 | S17 | [1] | Ε. | → E C |)P1 E | 1 | | S5 | S7 | | S8 | | | | S6 | | | |
| S7 | | [2] | | E1 | | | | | | R[9] | R[9] | R[9] | R[9] | R[9] | | R[9] | R[9] | |
| S8 | | [3] | E1 | → E1 | OP2 | . E2 | | | S18 | | | | | | | | | |
| S9 | | [4] | [4] E2 | | | | | | S7 | | S8 | | | | S6 | | | |
| S10 | | [5] | | → E3 | | E2 | | | R[11] | | R[11] | | | | R[11] | | | |
| S11 | | [6] | | E3 | | | | | R[12] | | R[12] | | | | R[12] | | | |
| S12 | | [7] | | → NU | | | | S5 | S7 | | S8 | | | | S6 | | | |
| S13 | | [8] | | | | · E <r< td=""><td>(PAK></td><td></td><td>R[13]</td><td></td><td>R[13]</td><td></td><td></td><td></td><td>R[13]</td><td></td><td></td></r<> | (PAK> | | R[13] | | R[13] | | | | R[13] | | | |
| S14 | | [9] - [10] | | M → < | | ı> · <unl< td=""><td>IMS</td><td></td><td>R[14]</td><td></td><td>R[14]</td><td></td><td></td><td></td><td>R[14]</td><td></td><td></td></unl<> | IMS | | R[14] | | R[14] | | | | R[14] | | | |
| S15 | | | - | | | |)IYI > | S5 | S7 | | 58 | | | | S6 | | | |
| S16 | | [12] | - | | | | | | R[15] | | R[15] | | | | R[15] | | | |
| S17 | | | _ | P2 → | | | | | | S10 | S11 | | | | | S22 | | |
| S18 | | [14] | _ | · - i | | | | | | R[10] | R[10] | | | R[10] | | R[10] | R[10] | |
| S19 | | | - | P3 → | | | | | | R[1] | R[1] | S13 | S14 | | | R[1] | R[1] | |
| S20 | | | | | | | | | | R[3] | R[3] | R[3] | R[3] | | | R[3] | R[3] | |
| S21 | | | | ! ! | | | | | | R[5] | R[5] | | | | | R[5] | R[5] | |
| S22 | | | | | | | | | | R[8] | R[8] | R[8] | R[8] | | | R[8] | R[8] | |

SLR查表解析应用示例

| 规范族 | | | | G | ото | | | | Action (Shift-Reduce) | | | | | | | | | |
|--------|--------------|---|---|-------------|-----|-----|--|--|--|--|---------------------------------|---------------|-------------|-------------|-----------|-----------|-----------------------|--|
| 规氾族 | Е | E1 | E2 | E3 | OP1 | OP2 | 0P3 | NUM | <unum></unum> | <add></add> | | <mul></mul> | <div></div> | <pow></pow> | <lp></lp> | <rp></rp> | <eof:< td=""></eof:<> | |
| S0 | S1 | S2 | S3 | S4 | | | | S5 | S7 | | S8 | | | | S6 | | | |
| S1 | | | | | S9 | | | | | S10 | S11 | | | | | | acc | |
| S2 | | | | | | S12 | | | | R[2] | R[2] | S13 | S14 | | | R[2] | R[2 | |
| S3 | | | | | | | | | | R[4] | R[4] | R[4] | R[4] | | | R[4] | R[4 | |
| S4 | | | | | | | S15 | | | R[6] | R[6] | R[6] | R[6] | S16 | | R[6] | R[6 | |
| S5 | | | | | | | | | | R[7] | R[7] | R[7] | R[7] | R[7] | | R[7] | R[7 | |
| S6 | S17 | S2 | S3 | S4 | | | | S5 | S7 | | S8 | | | | S6 | | | |
| S7 | | | | | | | | | | R[9] | R[9] | R[9] | R[9] | R[9] | | R[9] | R[9 | |
| S8 | | | | | | | | | S18 | | | | | | | | | |
| S9 | | S19 | S3 | S4 | | | | S5 | S7 | | S8 | | | | S6 | | | |
| S10 | | | | | | | | | R[11] | | R[11] | | | | R[11] | | | |
| Stack | | | Symb | ools | | | Input | | | | Action | 1 | | | | | | |
| SØ | | | | | | | <unui< td=""><td>M><mul< td=""><td>.><unum></unum></td><td><eof></eof></td><td>shift</td><td><unum></unum></td><td>goto s</td><td>S7</td><td></td><td></td><td></td></mul<></td></unui<> | M> <mul< td=""><td>.><unum></unum></td><td><eof></eof></td><td>shift</td><td><unum></unum></td><td>goto s</td><td>S7</td><td></td><td></td><td></td></mul<> | .> <unum></unum> | <eof></eof> | shift | <unum></unum> | goto s | S7 | | | | |
| S0,S7 | | | <uni< td=""><td>JM></td><td></td><td></td><td></td><td><mui< td=""><td>L><unum></unum></td><td colspan="8">:UNUM><eof> Reduce [9], back to S0, goto S5</eof></td></mui<></td></uni<> | JM> | | | | <mui< td=""><td>L><unum></unum></td><td colspan="8">:UNUM><eof> Reduce [9], back to S0, goto S5</eof></td></mui<> | L> <unum></unum> | :UNUM> <eof> Reduce [9], back to S0, goto S5</eof> | | | | | | | | |
| S0,S5 | | | NUM | | | | | <mui< td=""><td>L><unum></unum></td><td><eof></eof></td><td colspan="8">Reduce [7], back to S0, goto S4</td></mui<> | L> <unum></unum> | <eof></eof> | Reduce [7], back to S0, goto S4 | | | | | | | |
| S0,S4 | | | E3 | | | | | <mui< td=""><td>L><unum></unum></td><td colspan="8">1><eof> Reduce [6], back to S0, goto S3</eof></td></mui<> | L> <unum></unum> | 1> <eof> Reduce [6], back to S0, goto S3</eof> | | | | | | | | |
| S0,S3 | | E2 <mu< td=""><td><mui< td=""><td>L><unum></unum></td><td><eof></eof></td><td>Reduce</td><td>[4], b</td><td>ack to</td><td>S0, go</td><td>to S2</td><td></td><td></td></mui<></td></mu<> | | | | | | <mui< td=""><td>L><unum></unum></td><td><eof></eof></td><td>Reduce</td><td>[4], b</td><td>ack to</td><td>S0, go</td><td>to S2</td><td></td><td></td></mui<> | L> <unum></unum> | <eof></eof> | Reduce | [4], b | ack to | S0, go | to S2 | | | |
| 50,52 | | | E1 | | | | | <mu< td=""><td>L><unum< td=""><td>><eof></eof></td><td>Shift</td><td><mul>,</mul></td><td>goto S</td><td>13</td><td></td><td></td><td></td></unum<></td></mu<> | L> <unum< td=""><td>><eof></eof></td><td>Shift</td><td><mul>,</mul></td><td>goto S</td><td>13</td><td></td><td></td><td></td></unum<> | > <eof></eof> | Shift | <mul>,</mul> | goto S | 13 | | | | |
| S0,S2, | , S13 | | E1 · | <mul></mul> | | | | | <unum></unum> | <eof></eof> | Reduce | [13], | back to | o S2, g | oto Si | 12 | | |
| S0,S2, | ,S12 | | E1 (| 0P2 | | | S2,S12 E1 OP2 | | | | | | | | | | | |

SLR查表解析应用示例

| Stack | Symbols | Input | Action |
|---------------|----------------------|--|-----------------------------------|
| SØ | | <unum><mul><unum><eof></eof></unum></mul></unum> | shift <unum>, goto S7</unum> |
| S0,S7 | <unum></unum> | <mul><unum><eof></eof></unum></mul> | Reduce [9], back to S0, goto S5 |
| S0,S5 | NUM | <mul><unum><eof></eof></unum></mul> | Reduce [7], back to S0, goto S4 |
| S0,S4 | E3 | <mul><unum><eof></eof></unum></mul> | Reduce [6], back to S0, goto S3 |
| S0,S3 | E2 | <mul><unum><eof></eof></unum></mul> | Reduce [4], back to S0, goto S2 |
| 50,52 | E1 | <mul><unum><eof></eof></unum></mul> | Shift <mul>, goto S13</mul> |
| S0,S2,S13 | E1 <mul></mul> | <unum><eof></eof></unum> | Reduce [13], back to S2, goto S12 |
| S0,S2,S12 | E1 OP2 | <unum><eof></eof></unum> | Shift <unum>, goto S7</unum> |
| 50,52,512,57 | E1 OP2 <unum></unum> | <eof></eof> | Reduce [9], back to S12, goto S5 |
| S0,S2,S12,S5 | E1 OP2 NUM | <eof></eof> | Reduce [7], back to S12, goto S4 |
| 50,52,512,54 | E1 OP2 E3 | <eof></eof> | Reduce [6], back to S12, goto S20 |
| S0,S2,S12,S20 | E1 OP2 E2 | <eof></eof> | Reduce [4], back to S12, goto S2 |
| SØ | E1 | <eof></eof> | Reduce [3], back to s0, goto S2 |
| S0,S2 | E1 | <eof></eof> | Reduce [2], back to s0, goto S1 |
| S0,S1 | E | <eof></eof> | accept |

二、更多文法



思考: LL(1)和SLR哪个语法的表达能力更强?

- 如果一个语法是SLR, 是否一定是LL(1)?
- 如果一个语法是LL(1), 是否一定是SLR?

```
 \begin{bmatrix} 1 \end{bmatrix} S \rightarrow AaAb 
 \begin{bmatrix} 2 \end{bmatrix} \quad |BbBa 
 \begin{bmatrix} 3 \end{bmatrix} A \rightarrow \epsilon 
 \begin{bmatrix} 4 \end{bmatrix} B \rightarrow \epsilon
```

```
是LL(1)
```

- $First^+(S \rightarrow AaAb) = \{a\}$
- $First^+(S \rightarrow BbBa) = \{b\}$

不是SLR(1)

- $Follow(A) = Follow(B) = \{a, b\}$
- $Action(S_0, a) = reduce[3]$ 或 reduce[4]

LL(1)/SLR不够用怎么办?

- 表达能力太弱:解析表存在冲突
 - LR(1)>LALR>SLR: 规范族构造时考虑Follow信息
 - LL(K) > LL (1)
- 通用上下文无关文法解析算法:
 - 自底向上: CYK、GLR (Generalized LR)
 - 自顶向下: Earley算法