

第三章：语法分析

LR(0) 语法分析

例子

构造下列文法的LR(0)状态机。

G[S]:

$S \rightarrow aAd^{[1]}$

$A \rightarrow Bc^{[2]}$

$B \rightarrow b^{[3]}$

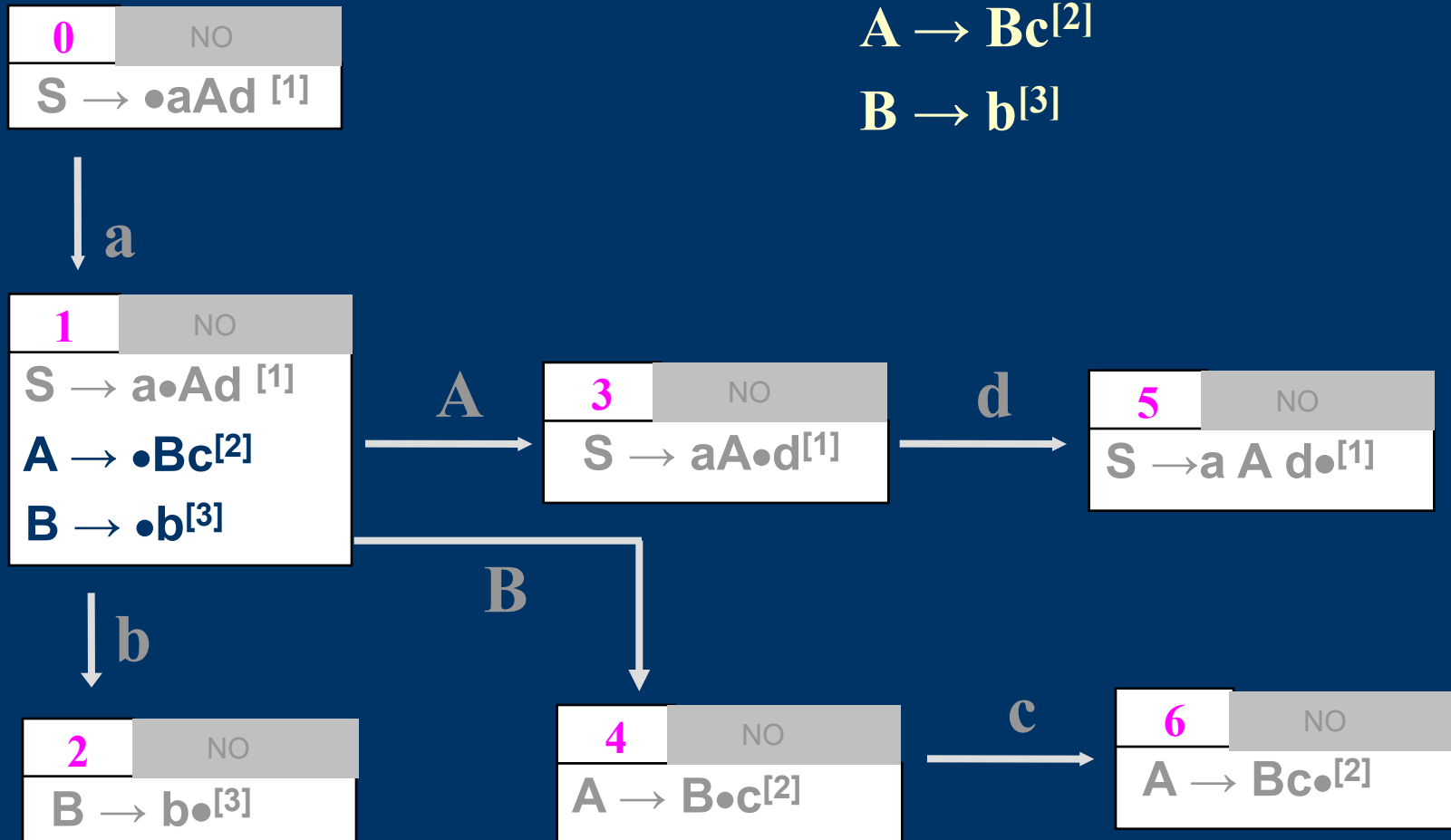
构造过程:

$G[S]$:

$S \rightarrow aAd^{[1]}$

$A \rightarrow Bc^{[2]}$

$B \rightarrow b^{[3]}$



$G_{[S]}$ 的 LRSM₀

1. 活前缀状态机提供的语法分析信息

1、移入信息：如果活前缀状态机中状态 IS_i 包含形如 $A \rightarrow \alpha \bullet a \beta$ 的项目，即 IS_i 有 a 的输出边，其中 a 是终极符，则表示 IS_i 状态遇当前输入符为 a 时应将其移入符号栈，状态机沿着其 a 的输出边转向其后继状态。

2、归约信息：如果状态 IS_i 包含形如 $X \rightarrow Y_1 Y_2 \dots Y_n \bullet$ 的项目，则表示 IS_i 状态可按该产生式做归约动作，约后状态机回退 n 个($Y_1 Y_2 \dots Y_n$ 的长度)状态至状态 IS_j ，随后沿着 IS_j 的 X 输出边转向其后继状态。

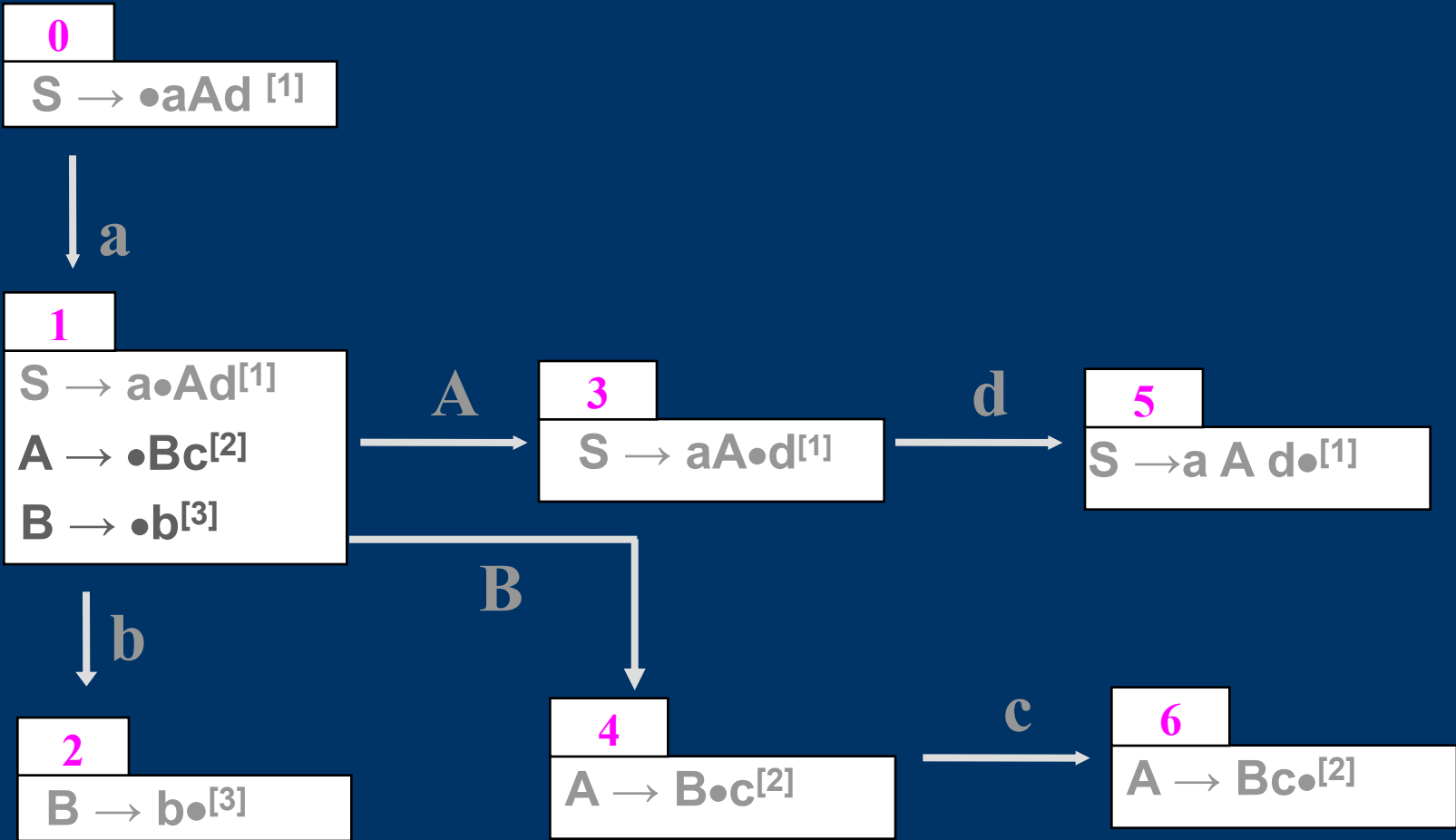
语法分析过程的直观描述:

a b c d #

符号栈

状态栈

#	a	B	d			
0	1	3	5			



活前缀状态机提供的语法分析信息（续）

设当前格局是：

#	X_1	X_2	...	X_k	...	X_t
S_{i0}	S_{i1}	S_{i2}	...	S_{ik}	...	S_{it}

$$a_i a_{i+1} \dots a_n \#$$

移入动作：设 S_{it} 的 a_i 输入边所指向的状态为 S^*

#	X_1	X_2	...	X_k	...	X_t	a_i
S_{i0}	S_{i1}	S_{i2}	...	S_{ik}	...	S_{it}	S^*

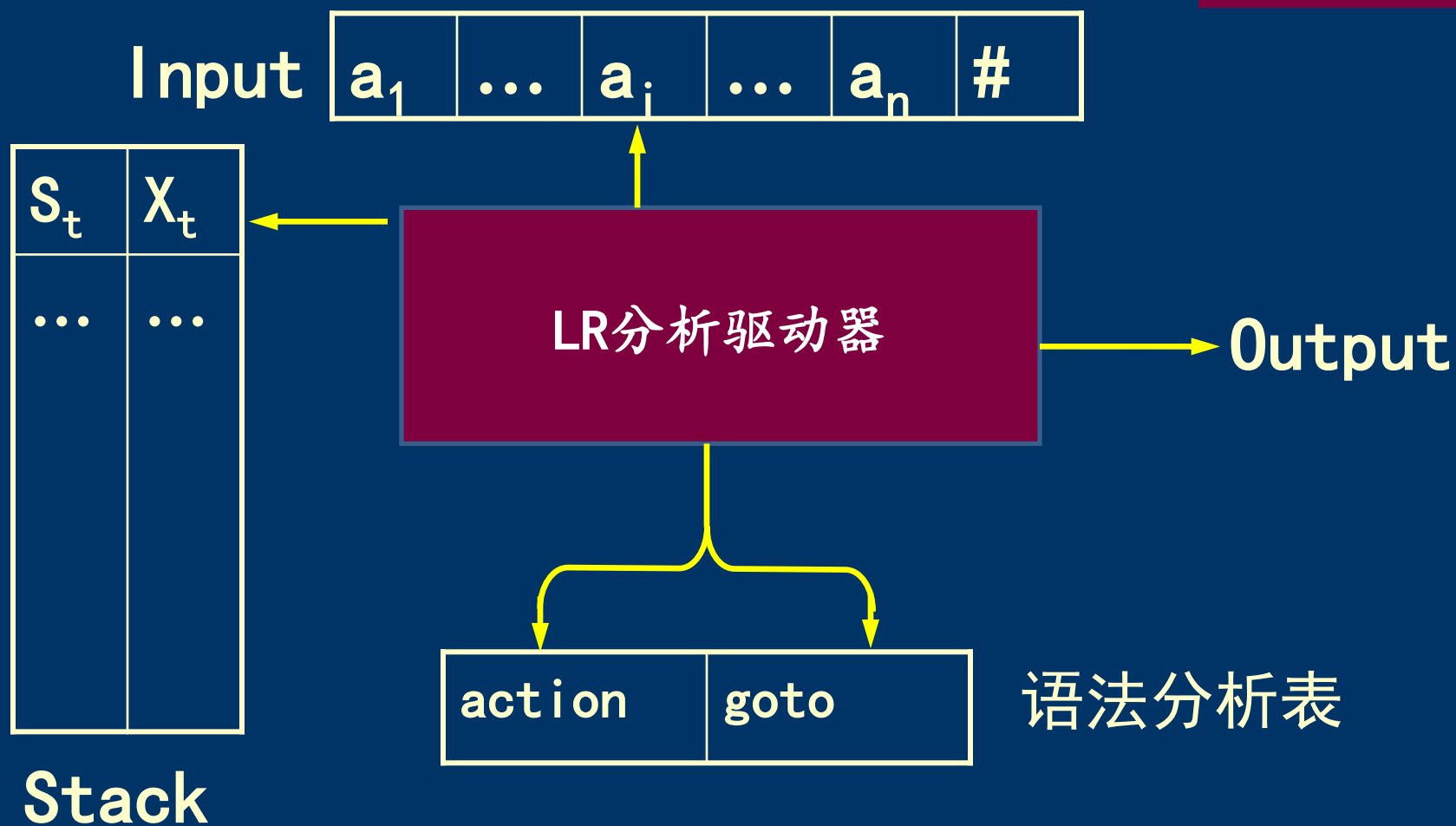
归约动作：设按 $A \rightarrow X_{k+1} X_{k+2} \dots X_t$ 进行归约，则首先归约为A

#	X_1	X_2	...	X_k	A
S_{i0}	S_{i1}	S_{i2}	...	S_{ik}	

S_{ik} 的A输出边所指向的状态设为 S^* ，则格局变为：

#	X_1	X_2	...	X_k	A
S_{i0}	S_{i1}	S_{i2}	...	S_{ik}	S^*

2. LR分析模型



2.1 LR分析表

- ◆ Action表:

状态 $\times (V_T \cup \{\#\}) \rightarrow$ 动作

动作包括: Shift / Reduce / Accept / Error (移入、归约、成功、失败)

- ◆ Goto表:

状态 $\times V_N \rightarrow$ 状态 | error

2.2 LR(0) 分析表的构造

假设 IS_k 为 LR(0) 项目集,

action 矩阵:

1. 若 $A \rightarrow \alpha \bullet a \beta \in IS_k$, 且 $G0(IS_k, a) = IS_i$, $a \in V_T$, 则 $action(IS_k, a) = S_i$, 表示移入动作。
2. 若 $A \rightarrow \alpha \bullet \in IS_k$, 则对任意 $a \in V_T \cup \{\#\}$, 令 $action(IS_k, a) = R_j$, 其中产生式 $A \rightarrow \alpha$ 的编号为 j , 表示用编号为 j 的产生式进行归约。
3. 若 $Z \rightarrow \alpha \bullet \in IS_k$, 且 Z 为拓广产生式的左部非终极符(文法的开始符), 则 $action(IS_k, \#) = Accept$ 。
4. 其它情形, 则 $Error(n)$, 表示出错标志, 也可不填。

goto 矩阵:

若 $G0(IS_k, A) = IS_i$, $A \in V_N$, 则 $goto(IS_k, A) = i$ 。

例：构造G[S]的基于LR(0) 方法的Action矩阵。

G[S]:
 $S \rightarrow aAd^{[1]}$
 $A \rightarrow Bc^{[2]}$
 $B \rightarrow b^{[3]}$

0

$S \rightarrow \bullet aAd^{[1]}$

a

1

$S \rightarrow a\bullet Ad^{[1]}$
 $A \rightarrow \bullet Bc^{[2]}$
 $B \rightarrow \bullet b^{[3]}$

b

2

$B \rightarrow b\bullet^{[3]}$

A

3

$S \rightarrow aA\bullet d^{[1]}$

d

5

$S \rightarrow aA d\bullet^{[1]}$

B

4

$A \rightarrow B\bullet c^{[2]}$

c

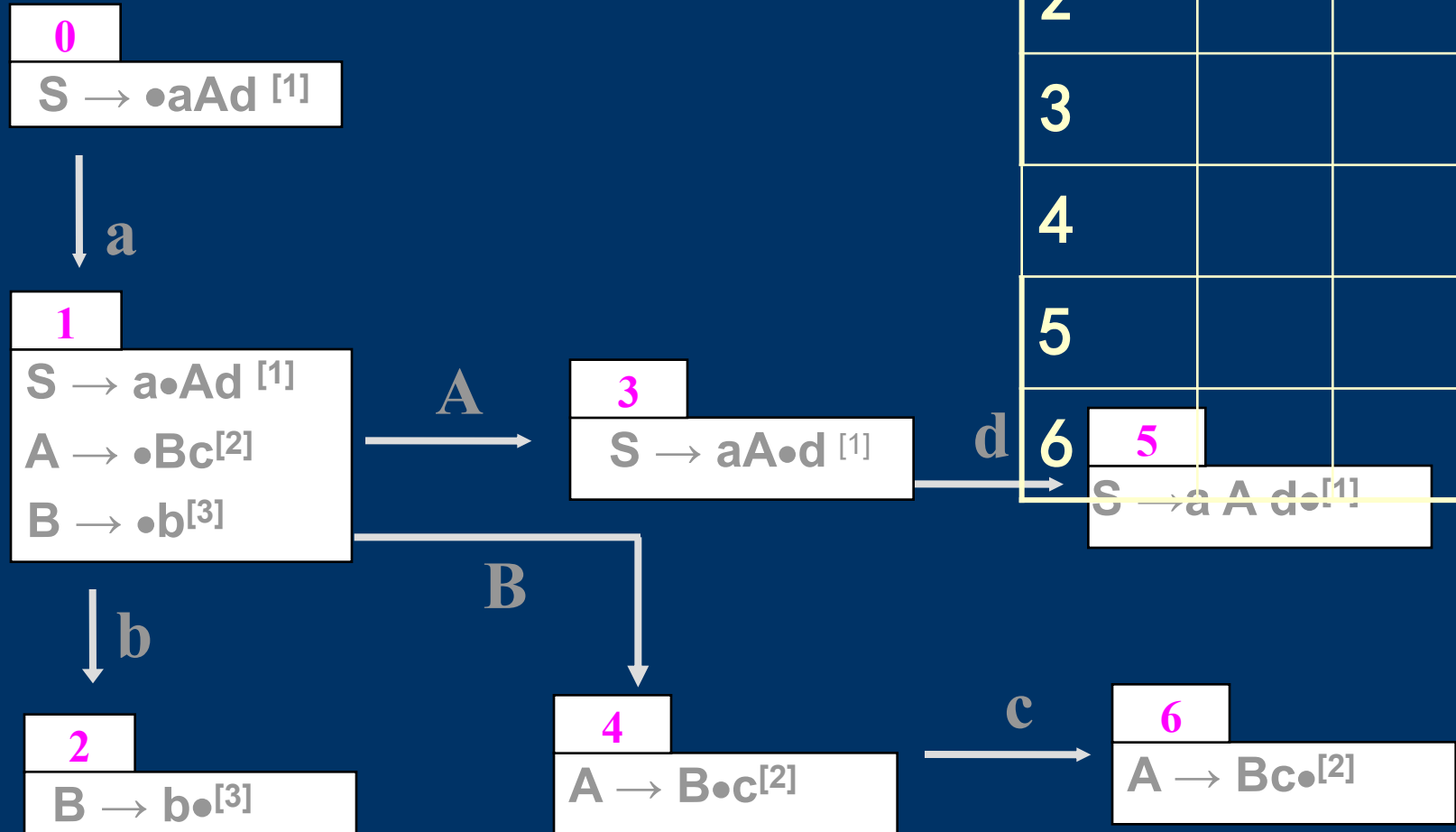
6

$A \rightarrow Bc\bullet^{[2]}$

	Action				
	a	b	c	d	#
0	S_1				
1		S_2			
2	R_3	R_3	R_3	R_3	R_3
3				S_5	
4			S_6		
5					Acc
6	R_2	R_2	R_2	R_2	R_2

例：构造G[S]的基于LR(0) 方法的GoTo矩阵。

G[S]:
 $S \rightarrow aAd^{[1]}$
 $A \rightarrow Bc^{[2]}$
 $B \rightarrow b^{[3]}$



	GoTo表		
	S	A	B
0			
1		3	4
2			
3			
4			
5			
6			

例：构造G[S]的分析表。

G[S]:

$S \rightarrow aAd^{[1]}$

$A \rightarrow Bc^{[2]}$

$B \rightarrow b^{[3]}$

	Act i on表					GoTo表		
	a	b	c	d	#	S	A	B
0	S ₁							
1		S ₂					3	4
2	R ₃	R ₃	R ₃	R ₃	R ₃			
3				S ₅				
4			S ₆					
5					Acc			
6	R ₂	R ₂	R ₂	R ₂	R ₂			

2.3 LR分析表提供的信息

- ◆ 合法性检查信息 $[A \rightarrow \alpha \bullet a \beta]$
- ◆ 移入/归约信息 $[A \rightarrow \alpha \bullet a \beta] [A \rightarrow \pi \bullet]$
- ◆ 移入/归约后的转向状态信息

2.4 LR驱动程序

- ◆ 状态栈、符号栈和输入流的开始格局为：
 $(\#S_1, \#, a_1a_2\cdots a_n\#)$
- ◆ 移入：若当前格局为
 $(\#S_1S_2\cdots S_n, \#X_1X_2\cdots X_n, a_ia_{i+1}\cdots a_n\#)$ ，且
 $\text{Action}(S_n, a_i) = S_j$ ， $a_i \in V_T$ ，则 a_i 入符号栈，第 j 个状态 S_j 入状态栈。即移入后的格局变为：
 $(\#S_1S_2\cdots S_n S_j, \#X_1X_2\cdots X_n a_i, a_{i+1}\cdots a_n\#)$

2.4 LR驱动程序

- ◆ 归约：若当前格局为

$(\#S_1S_2\cdots S_n, \#X_1X_2\cdots X_n, a_ia_{i+1}\cdots a_n\#)$ ，且 $Action(S_n, a) = R_j$ ， $a \in V_T \cup \{\#\}$ ，则按照第 j 个产生式进行归约，符号栈和状态栈相应元素退栈，归约后的文法符号入栈。假设第 j 个产生式为 $A \rightarrow \alpha$ ， $k = |\alpha|$ ($\alpha = X_{n-k+1}\cdots X_n$)，则归约后的格局变为：

$$(\#S_1S_2\cdots S_{n-k}S, \#X_1X_2\cdots X_{n-k}A, a_ia_{i+1}\cdots a_n\#)$$

其中 $S = Goto(S_{n-k}, A)$ 。

2.4 LR驱动程序

- ◆ 成功：若状态栈的栈顶状态为 S_i ，输入流当前值为 $\#$ ，且 $\text{action}(S_i, \#) = \text{Accept}$ ，则分析成功。
- ◆ 失败：若状态栈的栈顶状态为 S_i ，输入流当前值为 a ，且 $\text{action}(S_i, a) = \text{Error}$ 或空，则转向出错处理程序。

2.5 LR(0) 分析实例

构造LR(0)状态机

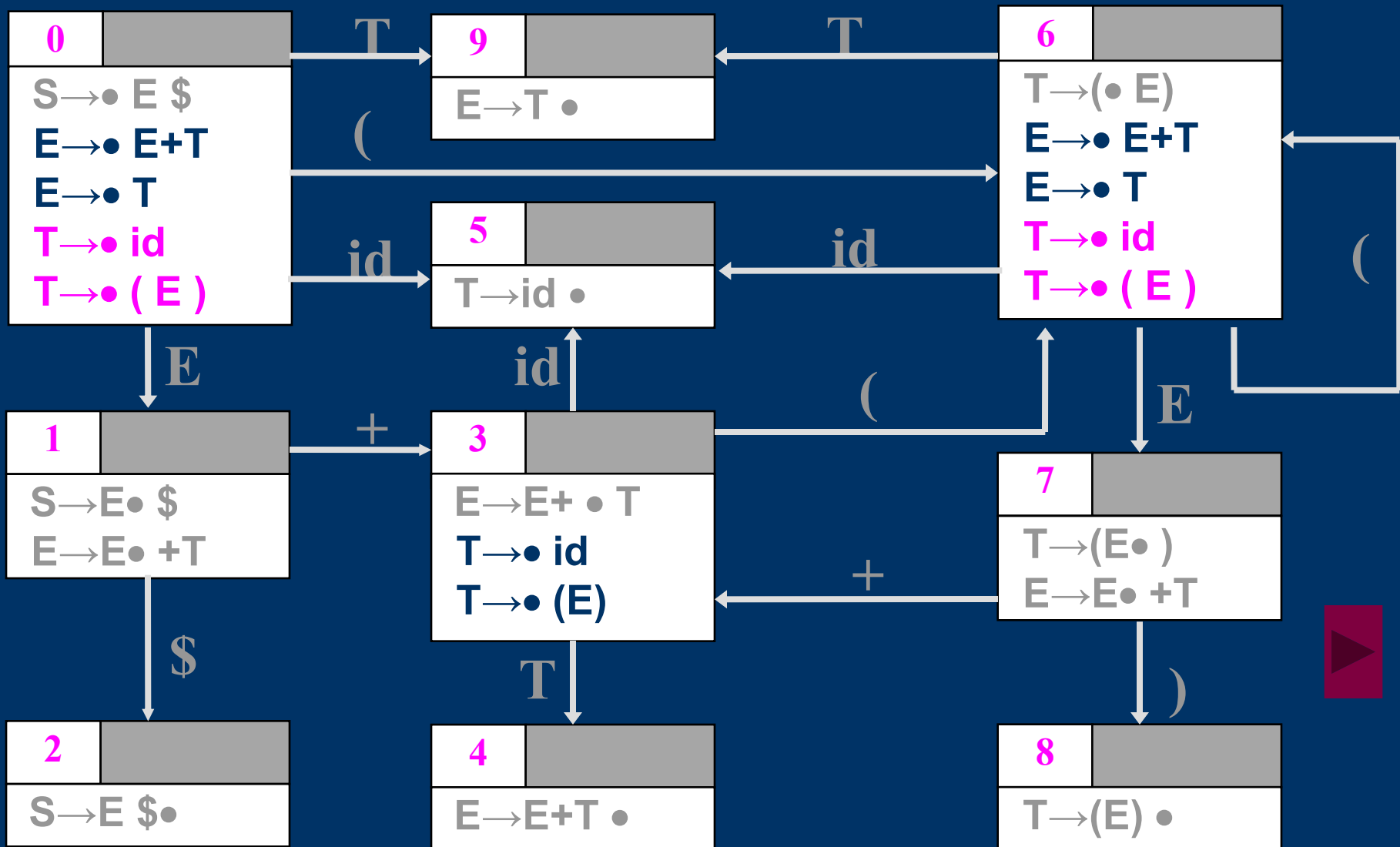
$$S \rightarrow E \$^{[1]}$$

$$E \rightarrow E + T^{[2]}$$

$$E \rightarrow T^{[3]}$$

$$T \rightarrow \text{id}^{[4]}$$

$$T \rightarrow (E)^{[5]}$$



	Action表						GoTo表	
	\$	+	id	()	#	E	T
0			S ₅	S ₆			1	9
1	S ₂	S ₃						
2						Acc		
3			S ₅	S ₆				4
4	R ₂	R ₂	R ₂	R ₂	R ₂	R ₂		
5	R ₄	R ₄	R ₄	R ₄	R ₄	R ₄		
6			S ₅	S ₆			7	9
7		S ₃			S ₈			
8	R ₅	R ₅	R ₅	R ₅	R ₅	R ₅		
9	R ₃	R ₃	R ₃	R ₃	R ₃	R ₃		

$S \rightarrow E \$^{[1]}$

$E \rightarrow E + T^{[2]} \mid T^{[3]}$

$T \rightarrow id^{[4]} \mid (E)^{[5]}$



$G[S]: \quad S \rightarrow E \$^{[1]}$
 $E \rightarrow E + T^{[2]} \mid T^{[3]}$
 $T \rightarrow id^{[4]} \mid (E)^{[5]}$

LR(0)分析实例

分析: $i+i\$$

状态栈	符号栈	输入串	Action	Goto
0	#	$i+i\$$	$A[0,i]=S_5$	
05	# i	$+i\$$	reduce4	$G[0,T]=9$
09	# T	$+i\$$	reduce3	$G[0,E]=1$
01	# E	$+i\$$	$A[1,+]=S_3$	
013	# E+	$i\$$	$A[3,i]=S_5$	
0135	# E+i	$\$$	reduce4	$G[3,T]=4$
0134	# E+T	$\$$	reduce2	$G[0,E]=1$
01	# E	$\$$	$A[1,\$]=S_2$	
012	# E\$	#	accept	



3.1 LR(0) 文法的限定条件

- ◆ 定义：在可归前缀图中，如果一个状态含有两个或两个以上的以上的归约项目，则称有**归约/归约冲突**。如果同时含有移入项目和归约项目，称为**移入/归约冲突**，这类状态称为二义性状态。

3.2 LRSM₀中含有二义性状态的例子

例如有文法：

[1] $M \rightarrow T$

[2] $T \rightarrow F$

[3] $T \rightarrow F * T$

[4] $F \rightarrow a$

3.2 LRSM₀中含有二义性状态的例子

