第三章：语法分析

LR方法

LR(0)自动机

1.1 LR方法基本思想

符号栈 #α 有终极符也有非终极符

输入流 β# 终极符

- 假如要分析的串没有语法错误，则αβ一定是文法的一个句型
- LR方法的主要思想是，从输入流依此把符号移入符号栈，直至栈顶出现一个句柄；之后对句柄进行归约，直至栈顶不出现句柄；重复上述过程，直至最终归约为一个开始符，且输入流为空。

1.2 LR类分析方法相关定义

- 规范句型：用最右推导导出的句型(也称右句型）。
- 规范前缀：若存在规范句型αβ，且β是终极符串或空串，则称α为规范前缀。
- 归约规范活前缀：若α是含句柄的规范前缀，且句柄在α的最右端，即有α=α'π，且π是句柄，则称活前缀α为归约规范活前缀（简称可归前缀）。

1.3 LR类分析方法的关键问题

- 如何判断分析栈内容是否为可归前缀
- 能唯一的确定可归前缀中的句柄

1.4 可归前缀的判断

派生定理:

- 开始符产生式的右部是可归前缀。
- 如果$\alpha_1 A \alpha_2$是可归前缀，且$A \rightarrow \beta$是产生式，则$\alpha_1 \beta$也是可归前缀。

设有文法G[S]:

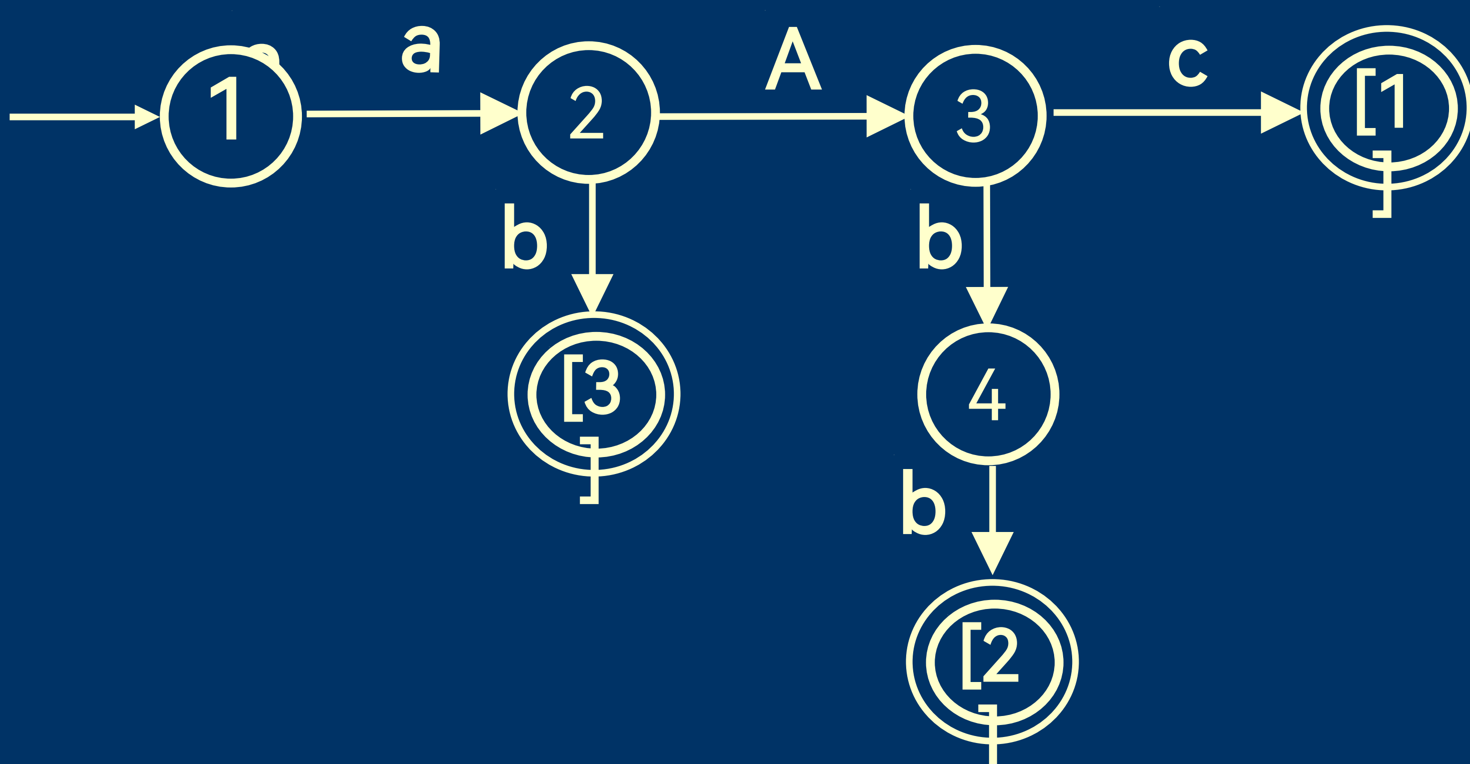
$S \rightarrow aAc$ [1]

$A \rightarrow Abb$ [2]

$A \rightarrow b$ [3] 可归前缀为：aAc，aAbb，ab

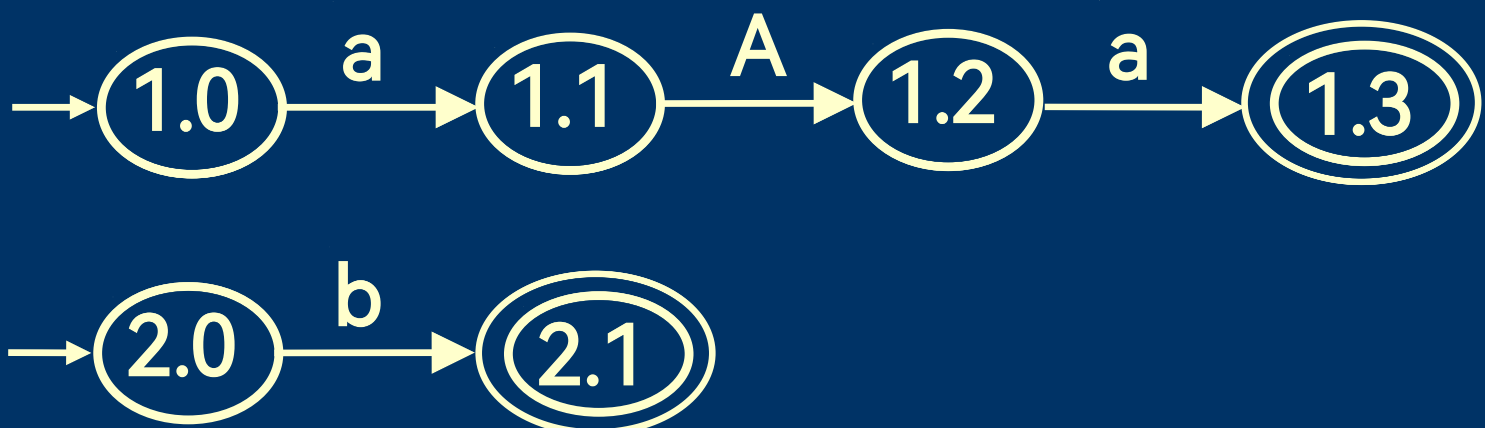
1.5 归约活前缀自动机

- 可归前缀的自动机表示为：aAc，aAbb，ab



2.1 文法自动机（GA）

- PA自动机：假设有产生式A→aAa[1] A→b[2]，则可构建如下多入口自动机，称做A的产生式自动机，记为PA(A),注意状态编码
- PA(A):



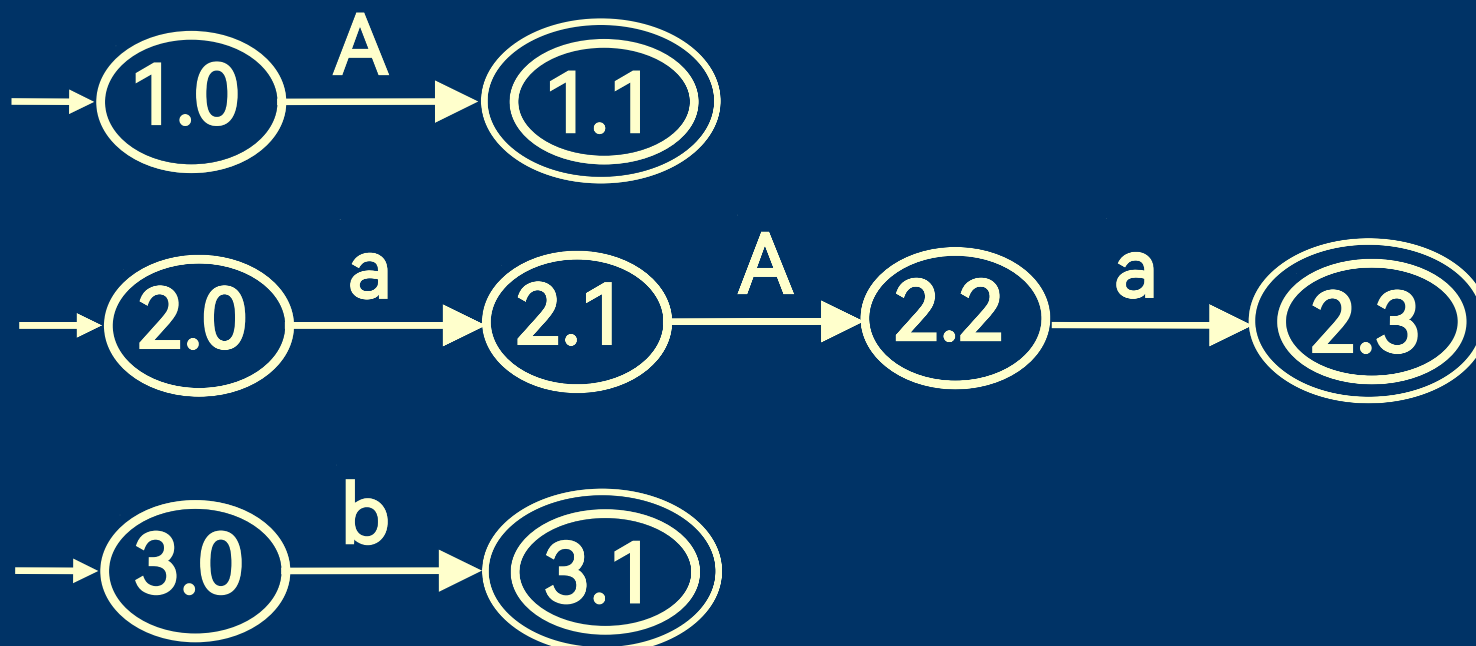
2.1 文法自动机（GA）

- GA自动机：假设给定文法G的一组PA自动机$PA(A_1),PA(A_2),\ldots,PA(A_n)$，则称按如下方法构造出来的自动机为G文法自动机，记为GA(G)，恰好接受G的所有可归前缀。

1. 将PA(S)的初始状态作为GA(G)的初始状态，其中S是文法的开始符
2. 连接各PA自动机，设$PA(A_i)$中的某一状态有A_j输出边，则从该状态到$PA(A_j)$的初始状态画ε边
3. 把原所有$PA(A_i)$的终止状态作为GA(G)的终止状态

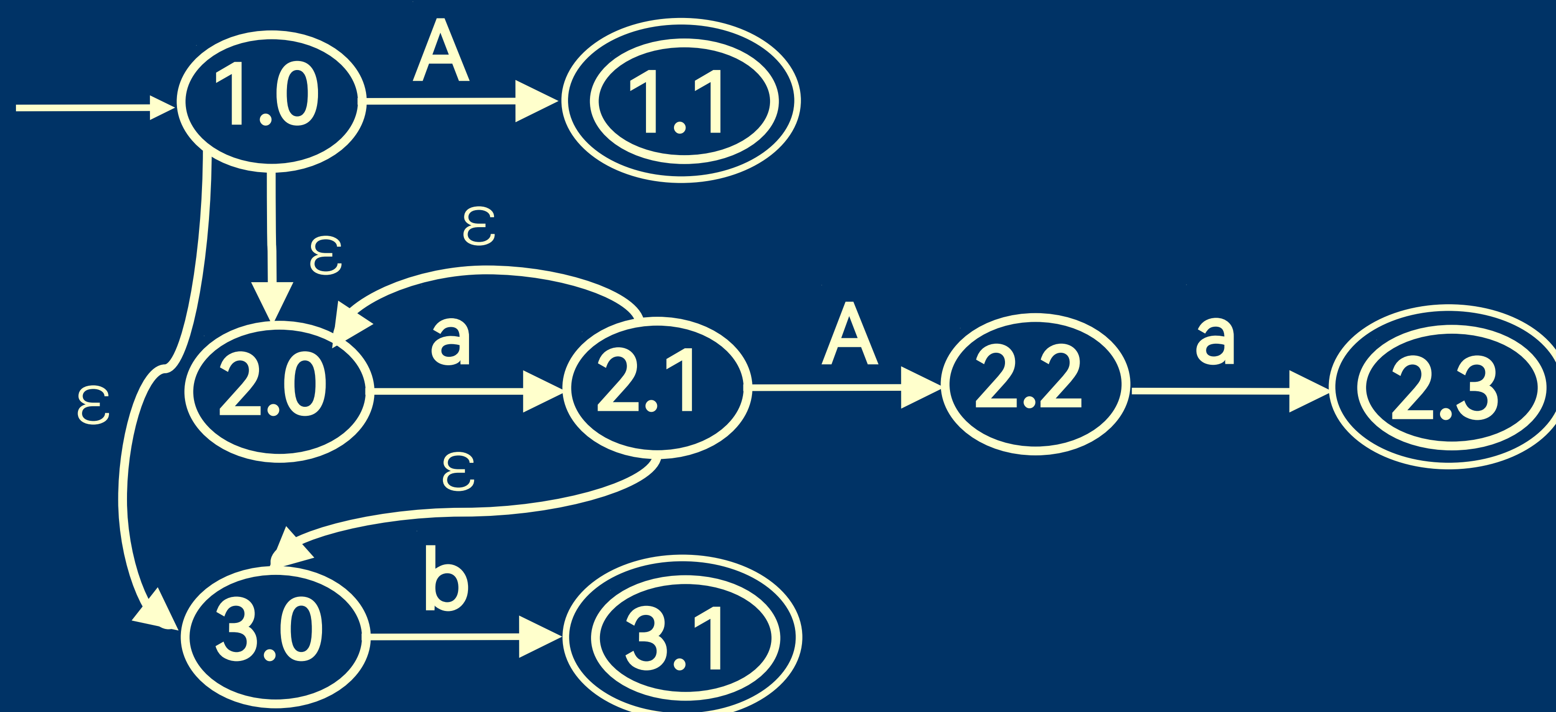
例子:文法的PA自动机

- G: Z→A[1] A→aAa[2] A→b[3]
- PA(Z)
- PA(A)



例子：非确定GA(G)

- GA(G):



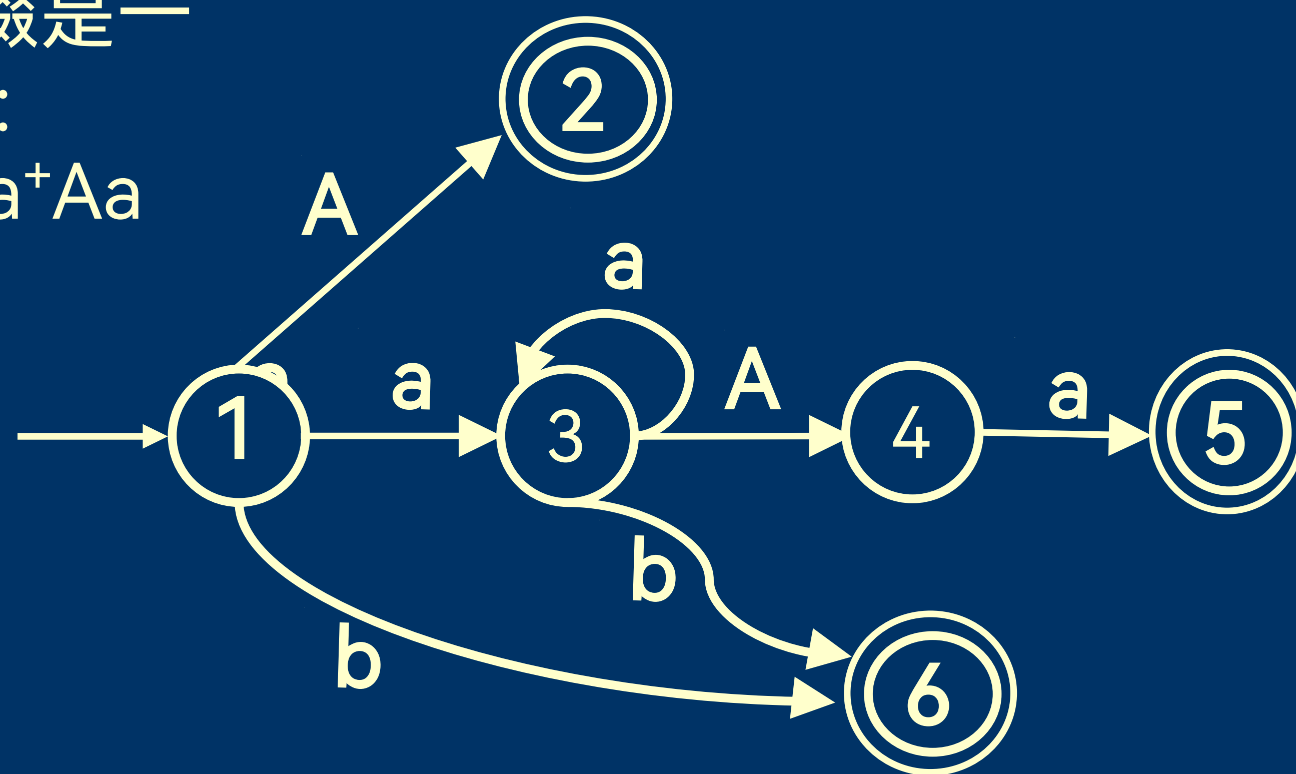
例子:转换过程、结点关系

	A	a	b
[10,20,30]	[11]	[21,20,30]	[31]
[11]			
[21,20,30]	[22]	[21,20,30]	[31]
[31]			
[22]		[23]	
[23]			

结点对应关系
1:[1.0,2.0,3.0]
2:[1.1]
3:[2.1,2.0,3.0]
4:[2.2]
5:[2.3]
6:[3.1]

例子:确定GA(G)_DFA

文法G的可归前缀是一个无穷集合，为:
A，aAa，a^+b，a^+Aa



2.2 LR(0)项目

- LR(0)项目是带一个圆点'•'的产生式。每个项目都标志分析时的某一状态。
- 对于文法Z→A[1] A→aAa[2] A→b[3]

来说，LR(0)项目有:

Z→•A <1,0> Z→A• <1,1>

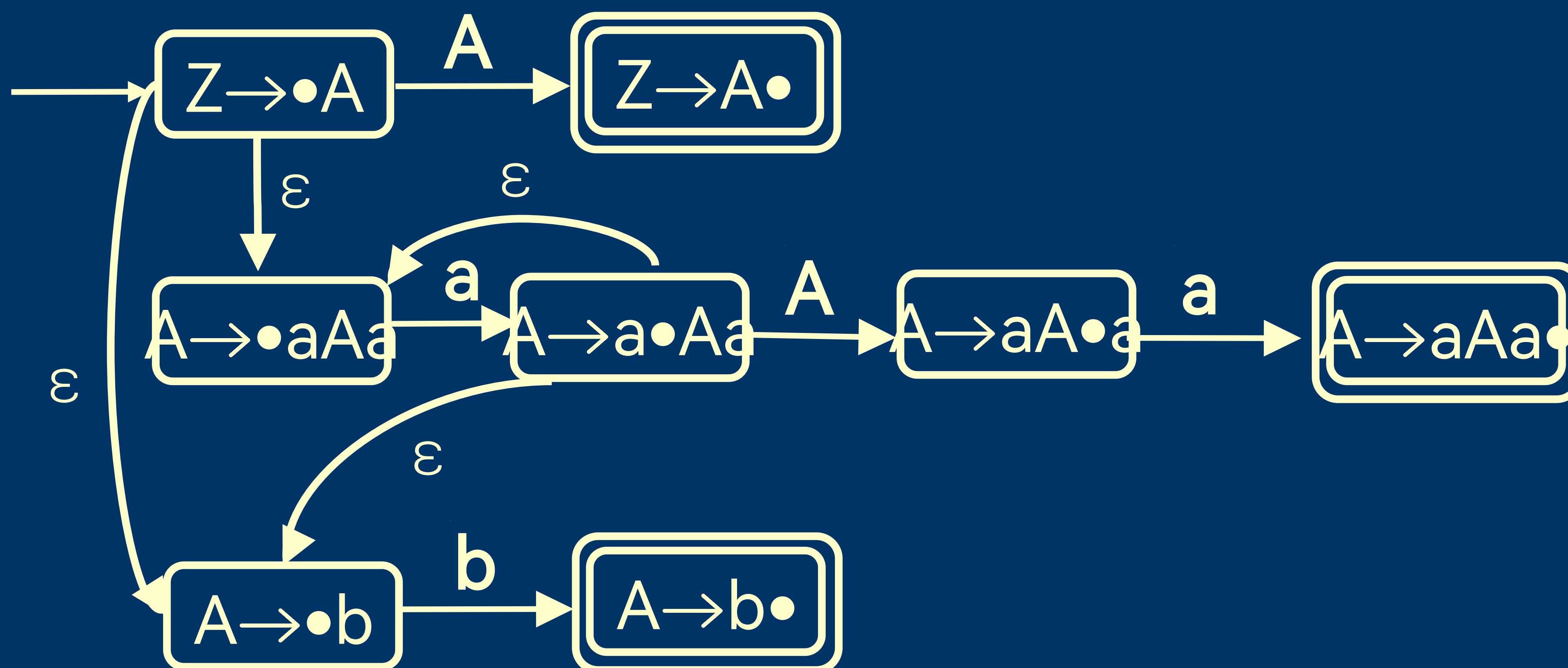
A→•aAa <2,0> A→a•Aa <2,1>

A→aA•a <2,2> A→aAa• <2,3>

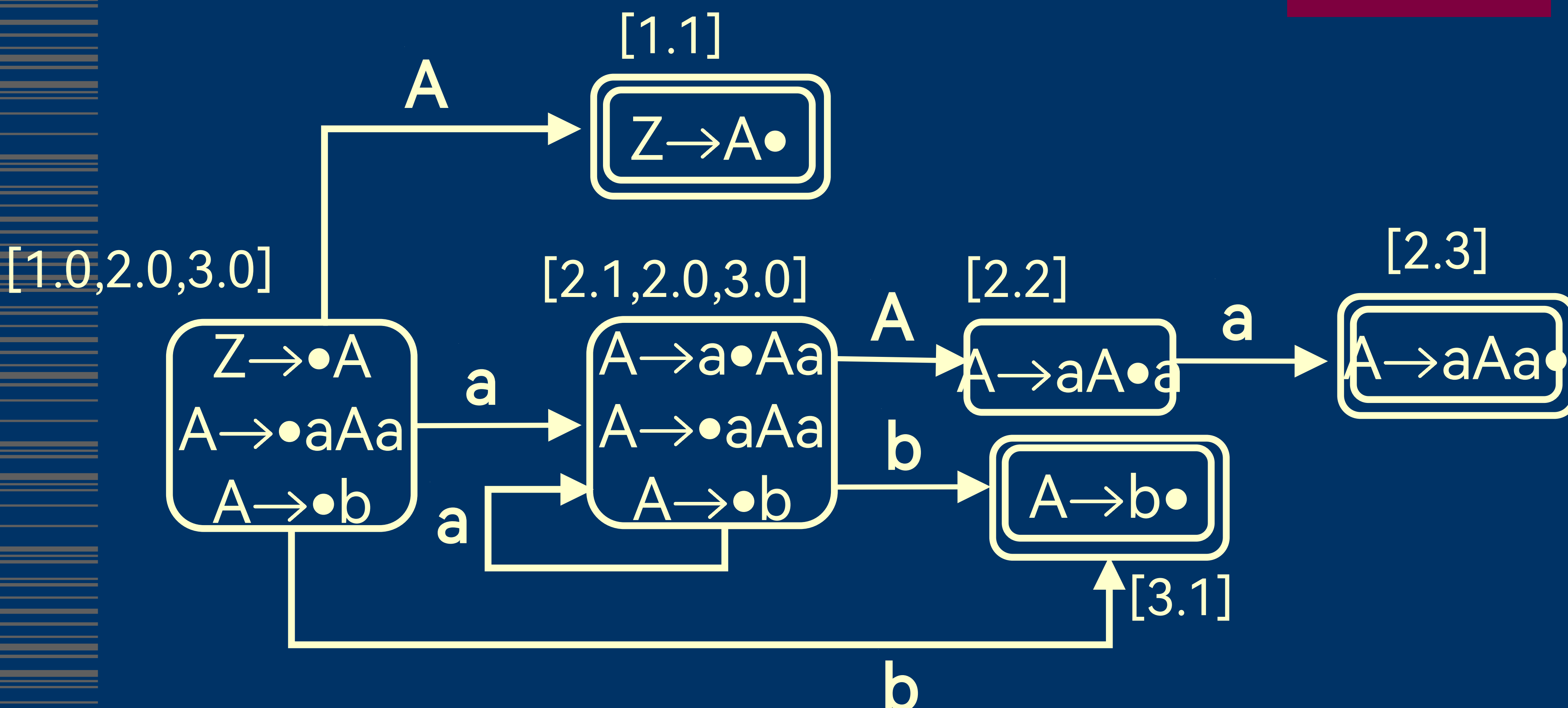
A→•b <3,0> A→b• <3,1>

2.3 文法的LR(0)自动机

LR(0)_GA(G):



2.4 LR(0)_GA(G)的转换



2.4 LR(0)_GA(G)的转换

- 项目集的投影：假设IS是LR(0)项目集，则称下面$IS_{(X)}$ 为IS关于X的投影集：

$$IS_{(X)} = \{A \to \alpha X \bullet \beta \mid A \to \alpha \bullet X\beta \in IS,\ X \in (V_T \cup V_N)\}$$

2.4 LR(0)_GA(G)的转换

NFA中状态集ε的闭包:

- 状态集I的ε闭包：设I是NFA M状态集的子集，定义I的ε闭包ε-CLOSURE(I)为:

1. 若q ∈I ,则q ∈ε_CLOSURE(I).
2. 若q ∈I,那么从q出发经任意条ε弧而能到达的任何状态q’都属于

ε-CLOSURE(I).

2.4 LR(0)_GA(G)的转换

项目集的闭包和状态集ε的闭包对应:

- 项目集的闭包：假设IS是LR(0)项目集，则称下面CLOSURE(IS)为IS的闭包集:

$CLOSURE(IS) = IS \cup \{A \to \bullet\pi \mid$
$Y \to \beta\bullet A\eta \in CLOSURE(IS)$ $A \to \pi$是产生式 $\}$

2.4 LR(0)_GA(G)的转换

NFA到DFA的转换

- 状态集I的a转换：若$I=\{S_1, \ldots, S_m\}$是NFA的状态集的一个子集，$a \in \Sigma$, 则定义:

  $Ia = \varepsilon\text{-CLOSURE}(J)$

其中:

  $J = f(S_1,a) \cup f(S_2,a) \ldots \cup f(S_m,a)$

2.4 LR(0)_GA(G)的转换

GO函数与NFA状态集I的a转换对应：

- 状态转换函数
- 设IS为LR(0)的项目集，X是语法符号，则定义GO(IS,X)=CLOSURE($IS_{(X)}$)

2.5 可归前缀的状态机(图)构造

1. 产生初始项目集IS_0，且$IS_0 \in ISS$
IS_0=CLOSURE(Z→•α),其中Z为开始符。
2. 若$IS_i \in ISS$，$X \in V_T \cup V_N$，则定义IS_j=GO(IS_i,X)，若IS_j不空且不属于ISS则将IS_j加入ISS，建立IS_i到IS_j的X映射,重复该过程，直到ISS不产生新状态。
3. 含有原PA中的终止状态的状态为可归前缀状态机的终止状态

例子

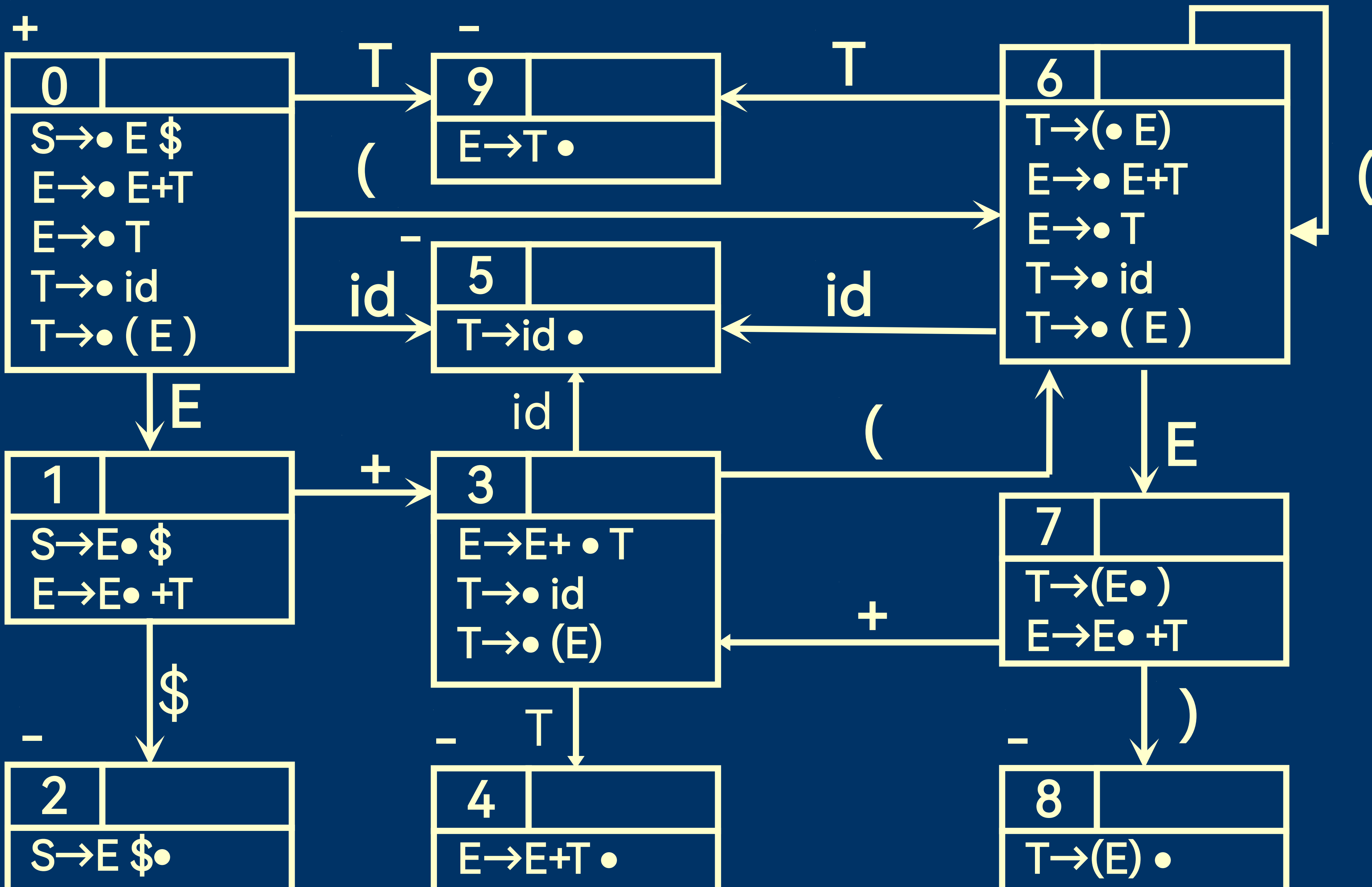
设有文法G(S):

$$S \rightarrow E\ \$$$

$$E \rightarrow E + T$$

$$E \rightarrow T$$

$$T \rightarrow id | (E)$$



0
S→• E $
E→• E+T
E→• T
T→• id
T→• (E)
1
S→E• $
E→E• +T
2
S→E $•
3
E→E+ • T
T→• id
T→• (E)
4
E→E+T •
5
T→id •
6
T→(• E)
E→• E+T
E→• T
T→• id
T→• (E)
7
T→(E•)
E→E• +T
8
T→(E) •
9
E→T •
T
(
id
E
+
$
T
id
(
E
+
)
+
–