

## § 3.2 正则文法和状态转换图

### § 3.2.1 由正则文法构造状态转换图

#### 1、状态图：

是一个由正则文法确定的有限的方向图, 结点对应状态, 含有一个初始状态和若干个终止状态.

#### 2、右线性文法的状态图

右线性文法的规则呈：

$$A \rightarrow aB \text{ 或 } A \rightarrow a$$

$$A, B \in V_n, a \in V_t$$

## (1) 状态图

$A \rightarrow aB$  或  $A \rightarrow a$

$A, B \in V_n, a \in V_t$

①  $G[S]$  的每一个非终结符号代表一结点(状态)

A

B

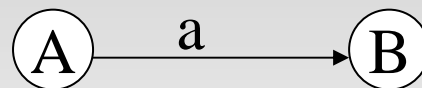
S

② 开始符号S作为初始状态

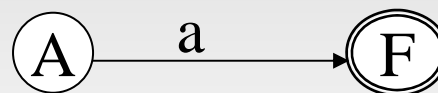
设一符号F不属于V作为终止状态

F

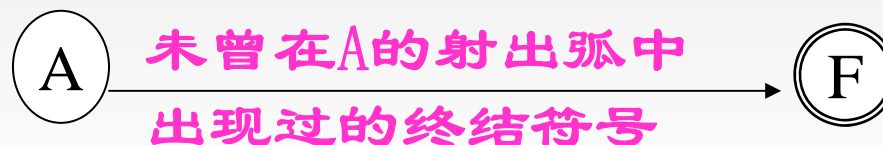
③ 形如  $A \rightarrow aB$  的规则



④ 形如  $A \rightarrow a$  的规则



特别:  $A \rightarrow \varepsilon$



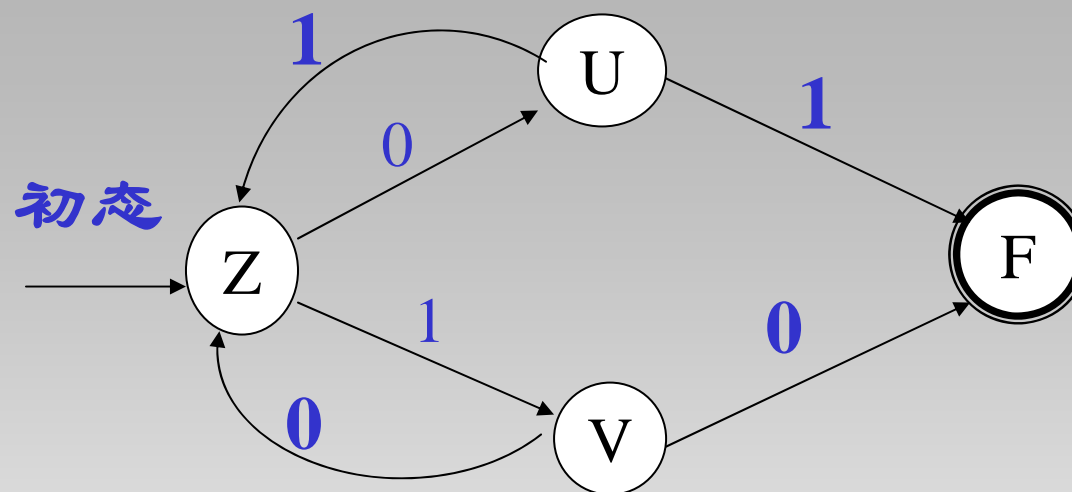
例:  $G[Z]$ :

$Z \rightarrow 0U \mid 1V$

$U \rightarrow 1Z \mid 1$

$V \rightarrow 0Z \mid 0$

状态图:



0101

1010

~~0011~~

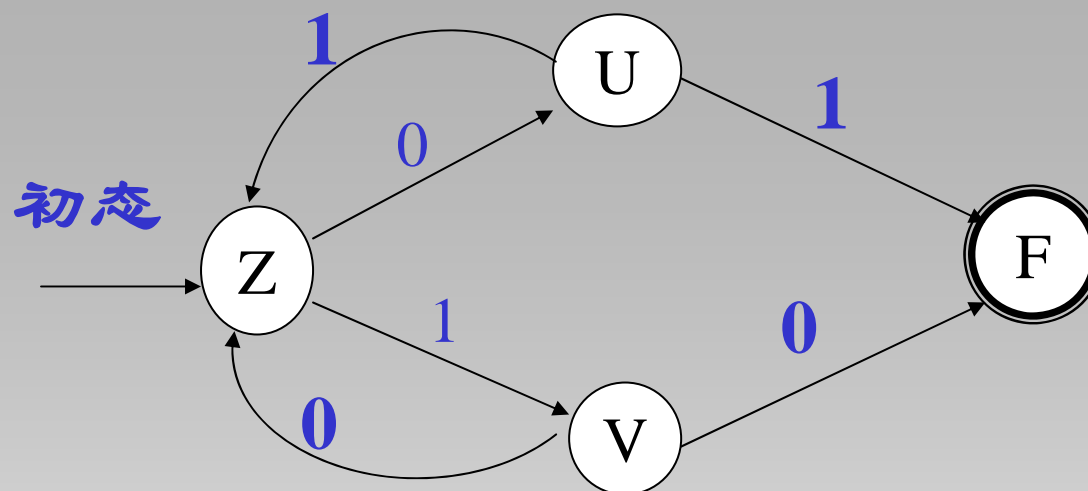
~~1100~~

例:  $G[Z]$ :

$Z \rightarrow 0U \mid 1V$

$U \rightarrow 1Z \mid 1$

$V \rightarrow 0Z \mid 0$



例:  $\omega = 011001$

通过状态图可以确定  $\omega$  是文法的句子.

分析过程是自上而下 (推导) ?

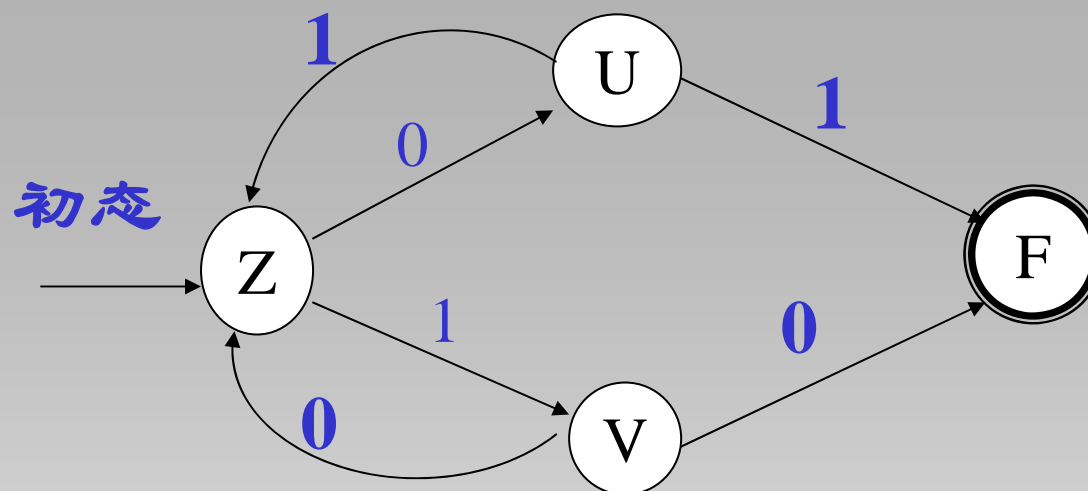
自下而上 (归约) ?

例:  $G[Z]$ :

$Z \rightarrow 0U \mid 1V$

$U \rightarrow 1Z \mid 1$

$V \rightarrow 0Z \mid 0$



例:  $\omega = 011001$

② 此过程是一种推导过程. (最右 (左) 推导)

相当于:  $Z \Rightarrow 0U$   
 $\Rightarrow 01Z$   
 $\Rightarrow 011V$   
 $\Rightarrow 0110Z$   
 $\Rightarrow 01100U$   
 $\Rightarrow 011001$

## 小 结

1. 右线性文法的状态转换图对符号串 $W$ 的识别的方法是一个**自顶向下的分析算法**
2. 推导过程中产生的句型都是**规范句型**
3. 右线性文法的状态转换图能识别的恰好是**对应文法的全部句子**

### 3、左线性文法的状态图

左线性文法的规则呈：  
 $A \rightarrow Ba$  或  $A \rightarrow a$   
 $A, B \in V_n, a \in V_t$

#### (1) 状态图

①  $G[S]$  的每一个非终结符号代表一结点(状态)  $\textcircled{A}$   $\textcircled{B}$

② 开始符号  $S$  作为终止状态  $\textcircled{\textcircled{S}}$

设一符号  $R$  不属于  $V$  作为初始状态  $\textcircled{R}$

③ 形如  $A \rightarrow Ba$  的规则  $\textcircled{B} \xrightarrow{a} \textcircled{A}$

④ 形如  $A \rightarrow a$  的规则  $\textcircled{R} \xrightarrow{a} \textcircled{\textcircled{S}}$

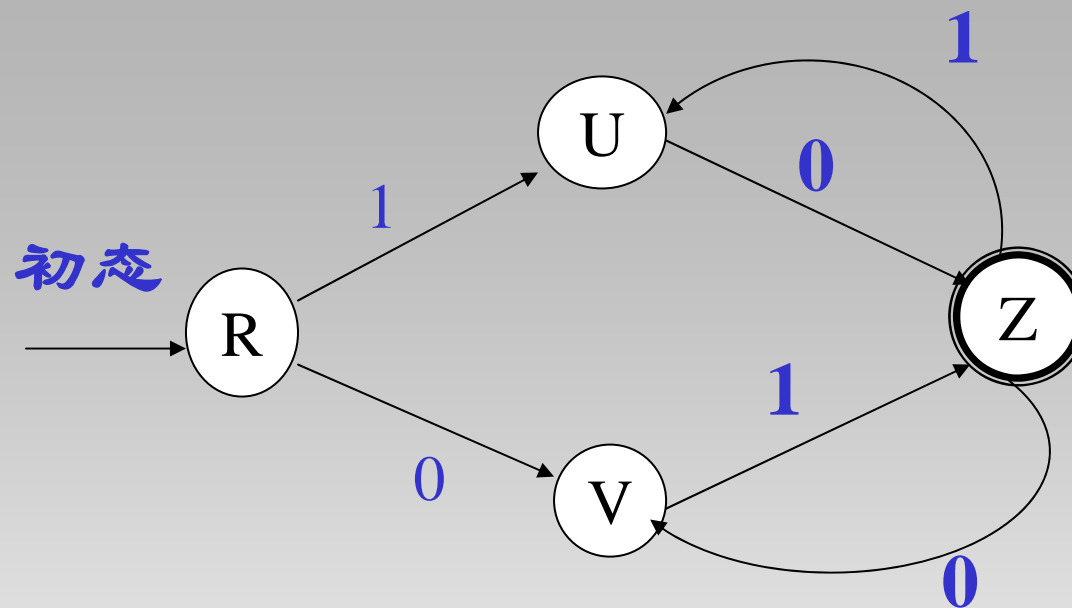
例:  $G[Z]$ :

$Z \rightarrow U0 \mid V1$

$U \rightarrow Z1 \mid 1$

$V \rightarrow Z0 \mid 0$

状态图:



0101

1010

~~0011~~

~~1100~~



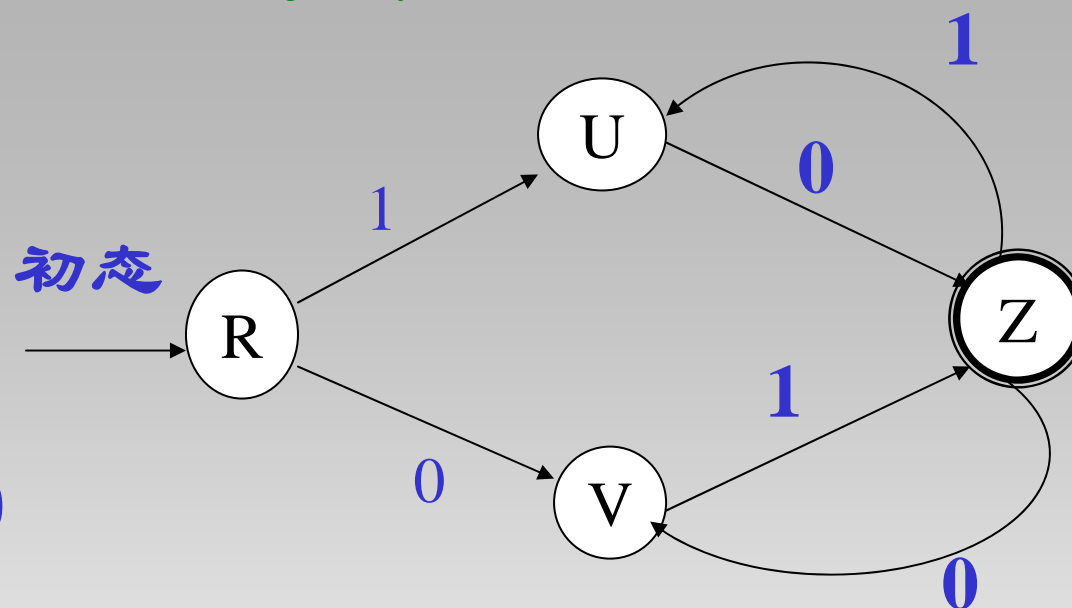
例:  $G[Z]:$

$Z \rightarrow U0 \mid V1$

$U \rightarrow Z1 \mid 1$

$V \rightarrow Z0 \mid 0$

状态图:



例:  $\omega = 100110$

通过状态图可以确定  $\omega$  是文法的句子.

分析过程是自上而下 (推导) ?

自下而上 (归约) ?

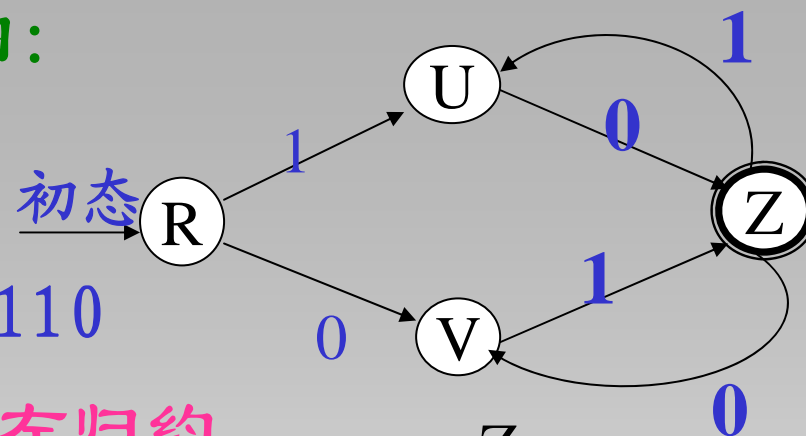
例:  $G[Z]$ :

$Z \rightarrow U0 \mid V1$

$U \rightarrow Z1 \mid 1$

$V \rightarrow Z0 \mid 0$

状态图:



例:  $\omega = 100110$

此过程是一种归约过程 最左归约

100110

U00110

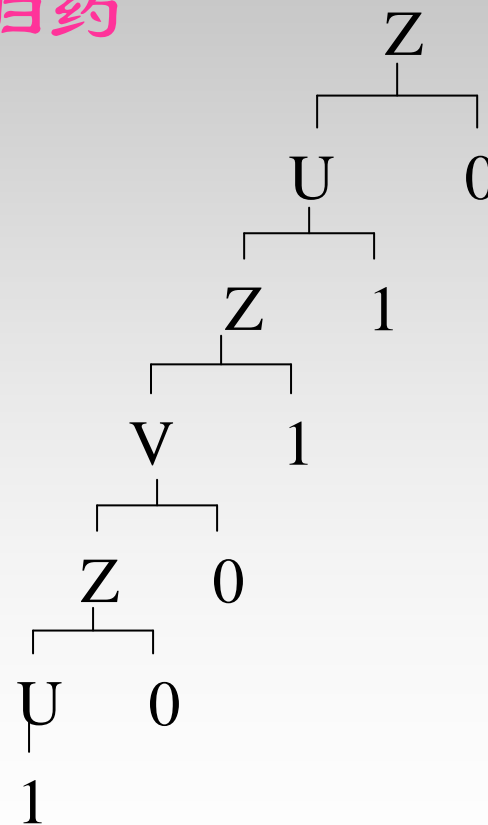
Z0110

V110

Z10

U0

Z



## 4、状态转换图的一种实现

### -----状态转换矩阵

状态转换矩阵是状态图在计算机实现的一种方法

	$a_1$	$a_2$	$\dots$	$a_j$	$a_m$
$S_1$	$B_{11}$	$B_{12}$	$\dots$		$B_{1m}$
$S_2$	$B_{21}$	$B_{22}$	$\dots$		$B_{2m}$
$\dots$					
$S_i$				$B_{ij}$	
$S_n$	$B_{n1}$	$B_{n2}$	$\dots$		$B_{nm}$

$$B_{ij}=B[I,j] \begin{cases} S_k & \text{当前状态 } S_i, \text{ 读输入符号 } a_j \text{ 的下一状态} \\ \text{空白 (error)} & \text{当前状态 } S_i, \text{ 不能读输入符号 } a_j \end{cases}$$

例：G[<无符号数>]： **右线性正则文法**

0<无符号数>  $\rightarrow$  d <余留无符号数> | . <小数部分> | d

1<余留无符号数>  $\rightarrow$  d <余留无符号数> | . <十进小数>  
| E <指数部分> | . | d

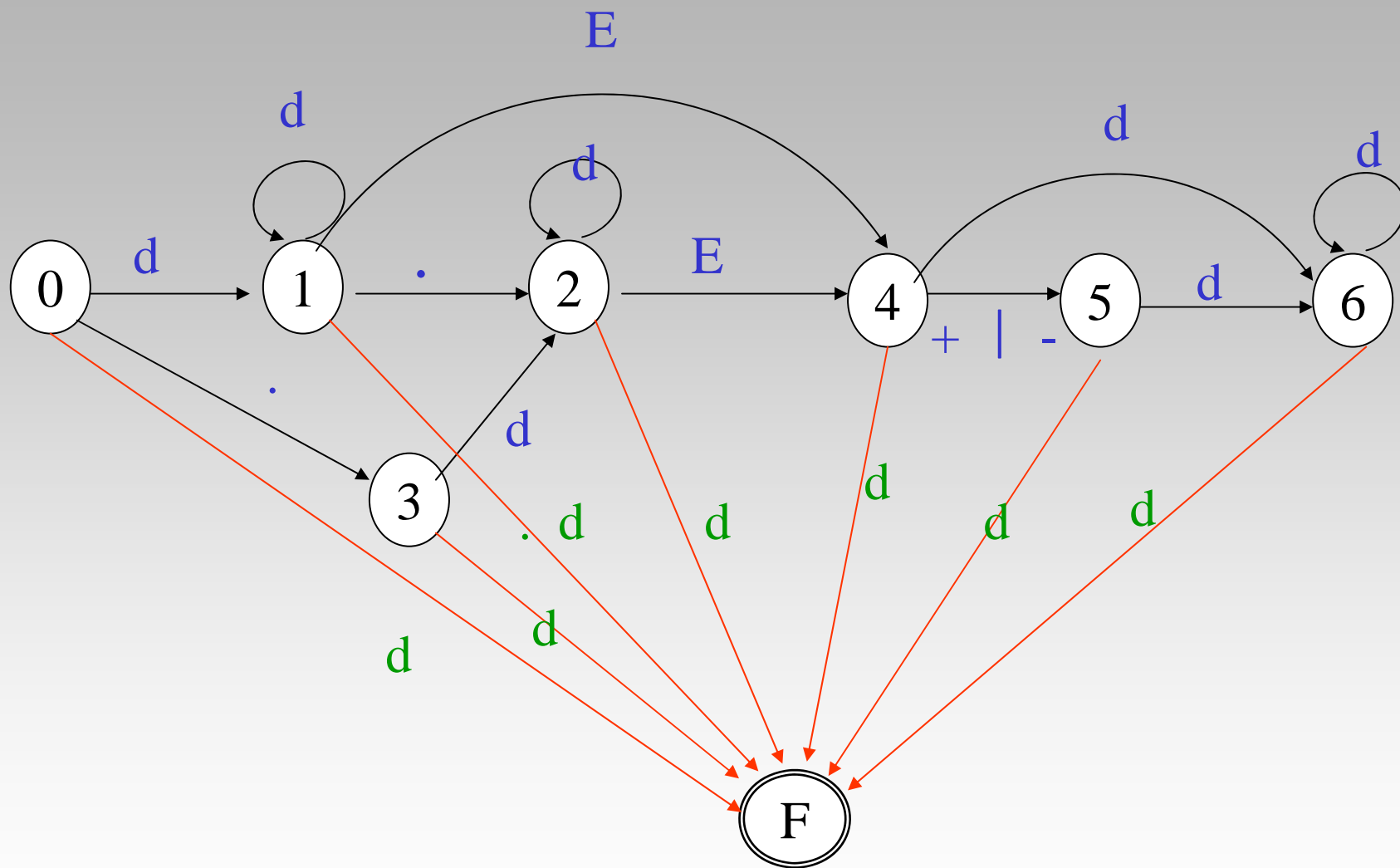
2<十进小数>  $\rightarrow$  E<指数部分> | d <十进小数> | d

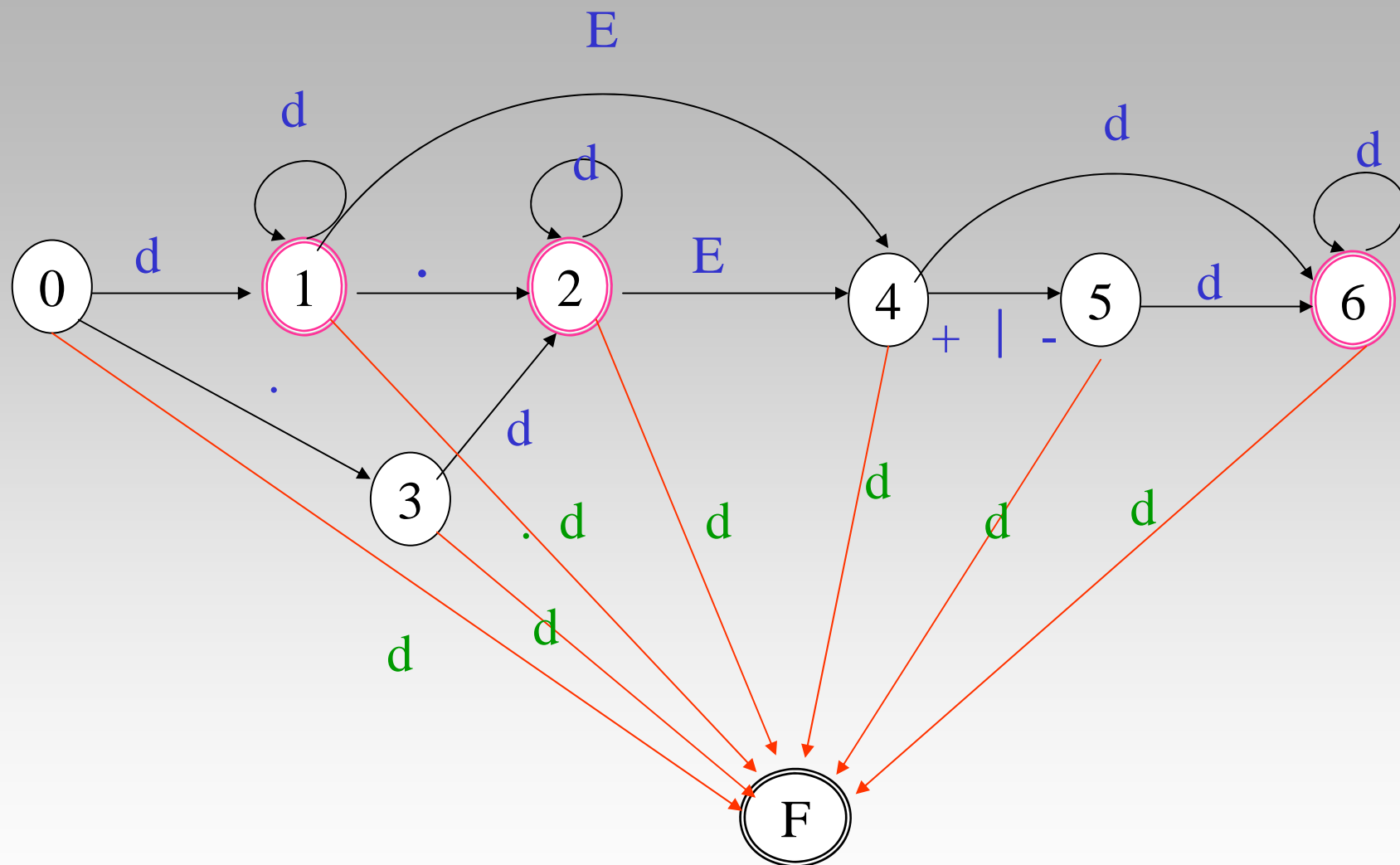
3<小数部分>  $\rightarrow$  d <十进小数> | d

4<指数部分>  $\rightarrow$  d <余留指数部分> | + <整指数> |  
- <整指数> | d

5<整指数>  $\rightarrow$  d <余留整指数> | d

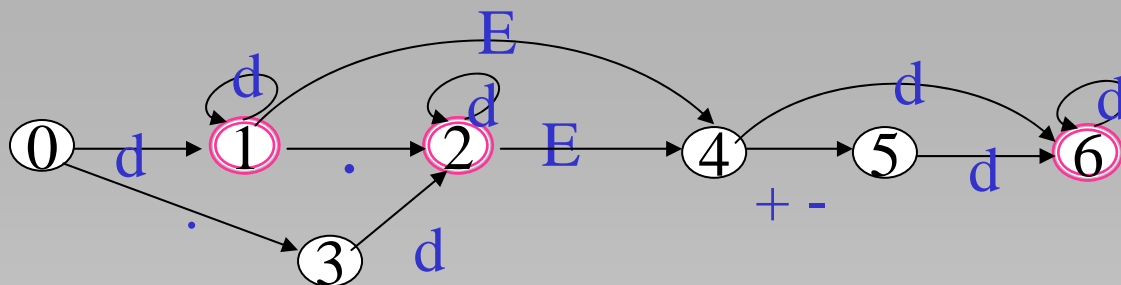
6<余留整指数>  $\rightarrow$  d <余留整指数> | d





## 无符号数写成

$$N = w * 10^{e * p - n}$$



无符号数一般形式： $d_m d_{m-1} \dots d_1 d_0 \cdot d_{-1} d_{-2} \dots d_{-n} E \pm d d \dots d$

改写为： $d_m d_{m-1} \dots d_1 d_0 d_{-1} d_{-2} \dots d_{-n} \times 10^{\pm dd \dots d-n}$

**w**: 尾数累加器 (初值为0)

**p**: 指数累加器 (初值为0)

**n**: 十进小数累加器 (初值为0)

**e**: 十进指数的符号 (初值为1,  
遇负号为-1)

遇到整数部分中d:  $w = w * 10 + d$

遇到小数点后d:  $w = w * 10 + d$   
 $n = n + 1$

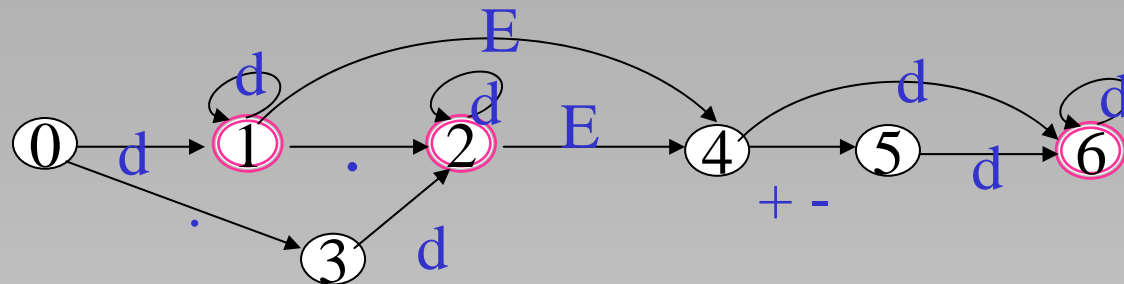
遇到E后负号:  $e = -1$

遇到指数部分中d:  $p = p * 10 + d$

遇到终态: 计算  $N = w * 10^{e * p - n}$

无符号数写成

$$N = w * 10^{e * p - n}$$



当前状态	扫描符号	语义程序	下一状态
0	d	$n=0;p=0;e=1;w=d;$	1
	.	$n=0;p=0;e=1; w=0;$	2
	other	error	NULL
1	d	$w=w*10+d$	1
	.		2
	E		4
	other	计算 $N=w*10^{e*p-n}$	结束
.....			