※ 正规语言的三种表示法综合示例:

【例3.5】 L ={ ab^nc, b^n | n≥0 }, Σ= {a, b, c};

1. 正规文法描述:

G(S): S-> aA|bB|ε , A -> bA|c , B -> bB| ε

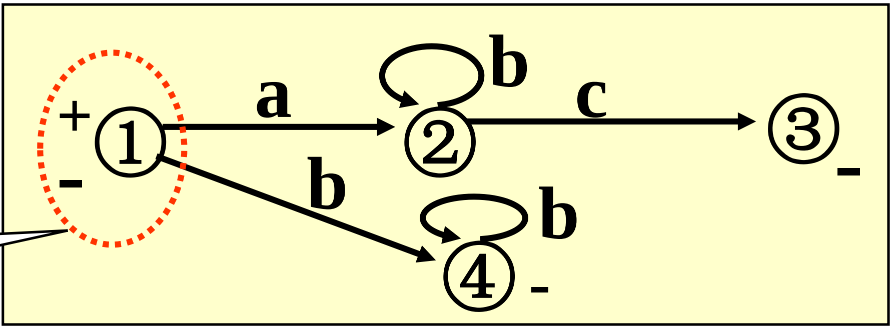
2. 正规式描述:

e = ab*c|b*

3. 有限自动机描述:

表示可接受**空串!**

FA:



【注】 凡是能由上述三种方法表示的语言，一定是正规语言；反之，凡是不能由上述三种方法表示的语言，一定不是正规语言。

3.2 有限自动机的定义和分类

3.2.1　有限自动机的定义

【定义】 有限自动机是一种数学模型，用于描述正规语言,可定义为五元组: FA=(Q, ∑, S, F, δ)

其中:

Q(有限状态集);

∑(字母表);

S(开始状态集, S ⊆ Q);

F(结束状态集, F ⊆ Q);

δ: 变换（二元函数）:

i,j∈Q,
a∈∑+{ε};

δ(i,a) =j **或** $i \xrightarrow{a} j$

【含义】

状态 i 遇符号 a, 变换为状态 j

3.2.2　有限自动机怎样描述语言

令　L(FA)为自动机FA所描述的正规语言；则

$$L(FA)=\{\ x \mid i \overset{x}{\Rightarrow} j,\ x\in\Sigma^*,\ i\in S,\ j\in F\ \}$$

含义是：

设　$x=a_1 a_2 \ldots a_n$，　$a_i\in\Sigma+\{\varepsilon\}$

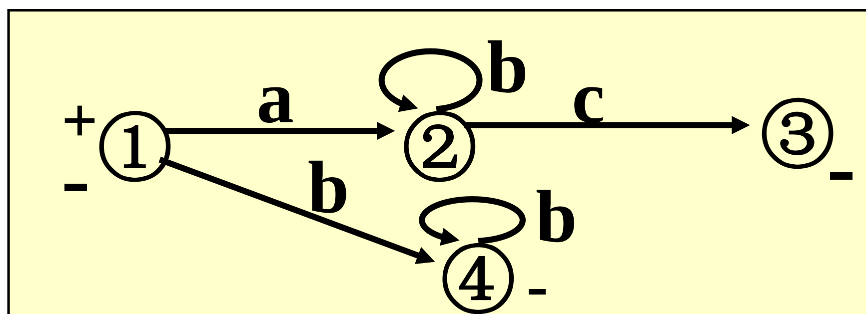
即：

则有　$i \overset{a_1}{\Rightarrow} i_1 \overset{a_2}{\Rightarrow} i_2 \ \ldots \ \overset{a_n}{\Rightarrow} j$

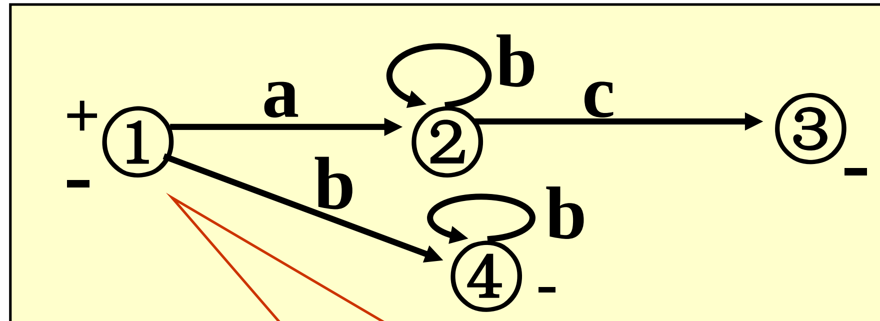
FA 存在一条从某初始状态 i 到某结束状态 j 的连续变换，且每次变换(=>)的边标记连成的符号串恰好为 x ；称x为FA接受，否则拒绝。

如 右图有限自动机：

则 L(FA)的识别过程如下所示：



※ L(FA)的生成(或识别)过程示例:



Ⅰ. 第一条通路：FA1

$+① \overset{a}{\Rightarrow} ② \overset{c}{\Rightarrow} ③\_$

$+① \overset{a}{\Rightarrow} ② \overset{b}{\Rightarrow} ② \overset{c}{\Rightarrow} ③\_$

$+① \overset{a}{\Rightarrow} ② \overset{b}{\Rightarrow} ② \overset{b}{\Rightarrow} ② \overset{c}{\Rightarrow} ③\_$

…

Ⅱ. 第二条通路：FA2

$+① \overset{\varepsilon}{\Rightarrow} ①\_$

$+① \overset{b}{\Rightarrow} ④\_$

$+① \overset{b}{\Rightarrow} ④ \overset{b}{\Rightarrow} ④\_$

…

接受空串的FA的典型特征！

$\therefore$ L(FA1)={ ab^nc | $n \geqslant 0$ }

$\therefore$ L(FA2)={ b^n | $n \geqslant 0$ }

因而

$\therefore$ L(FA)={ab^nc, b^n | $n \geqslant 0$}

3.2.3 有限自动机的两种表现形式

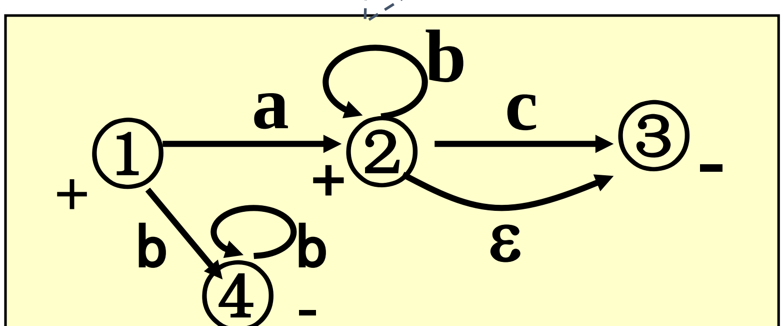
【例3.6】有限自动机 ：FA=(Q, Σ, S, F, δ)

其中：Q={1, 2, 3, 4}, Σ={a, b, c}, S={1, 2}, F={3, 4}

δ ： δ(1, a)=2; δ(1, b)=4; δ(2, b)=2;
δ(2, c)=3; δ(2, ε)=3; δ(4, b)=4;

FA 的两种表现形式：

(1) 状态图：



(2) 变换表：

开始状态
结束状态

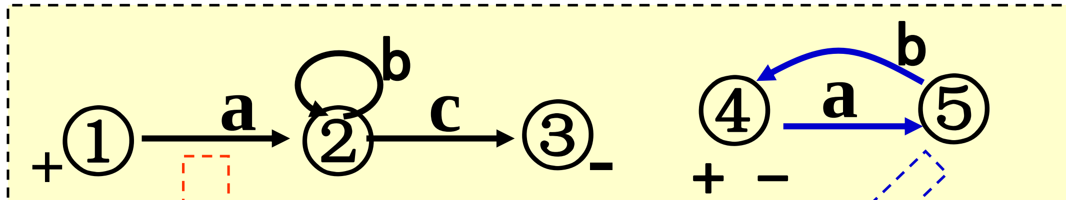
	a	b	c	ε
+1	2	4		
+2		2	3	3
−3				
−4		4		

※ 变换表结构：行(状态), 列(符号), 表项(变换后状态)

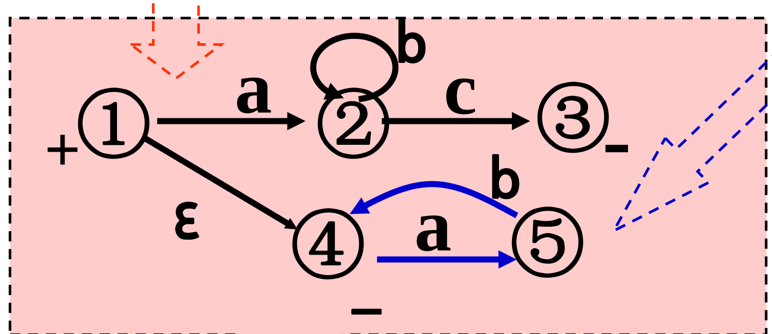
※ 有限自动机的构造示例:

【例3.7】 $A=\{\ ab^nc, (ab)^n | n\geq 0\ \}$ FA的构造:

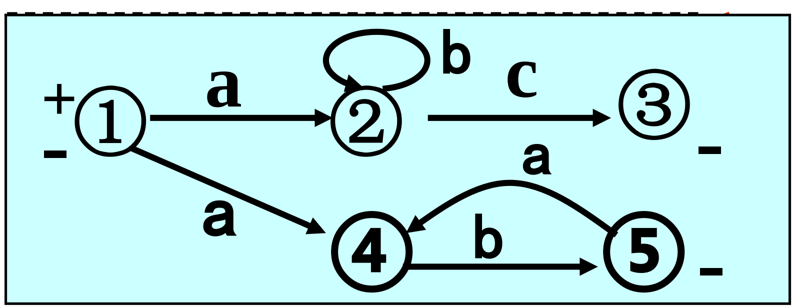
▪方法一：联合式



▪方法二：组合式1



或



有好用的吗?!

比较:

一：开始状态不唯一，不好用！

二：带有ε 边，也不好用！

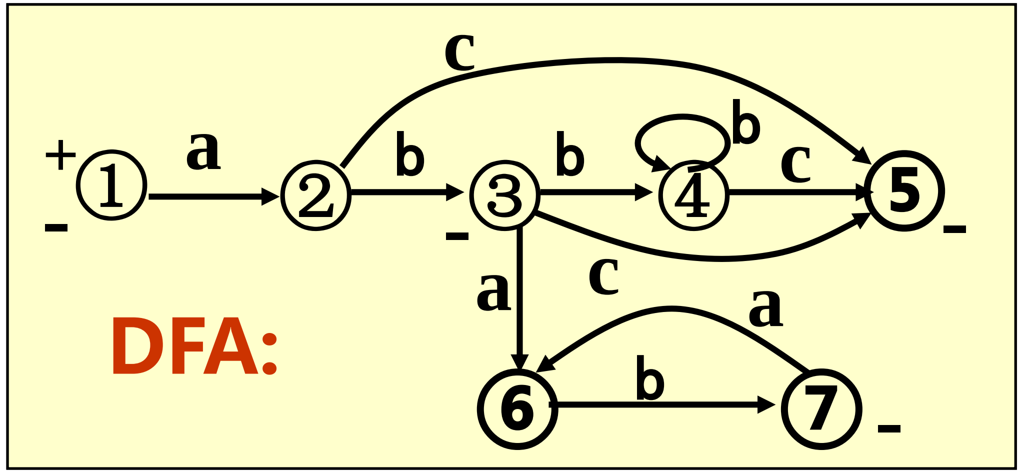
或：变换函数不单值，如

（ $\delta(1,a)=(2|4)$ ），仍不好用！

▪方法三：确定式

$A=\{ab^nc,(ab)^n|n\geq 0\}$

※ 确定的有限自动机如右图所示：



要领：构造的同时，要考虑连接！

3.2.4 有限自动机的分类

1. 确定的有限自动机(DFA)

特征：①开始状态唯一；②变换函数单值；③不带ε边。

2. 非确定的有限自动机(NFA)

-- 不能全部具备上述特征者！

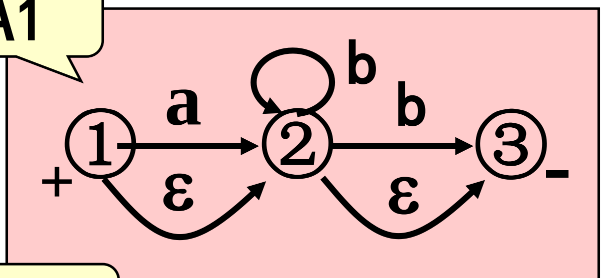
(1) 带有ε边的非确定的有限自动机(εNFA)

(2) 不带有ε边的非确定的有限自动机($\bar{\varepsilon}$NFA)

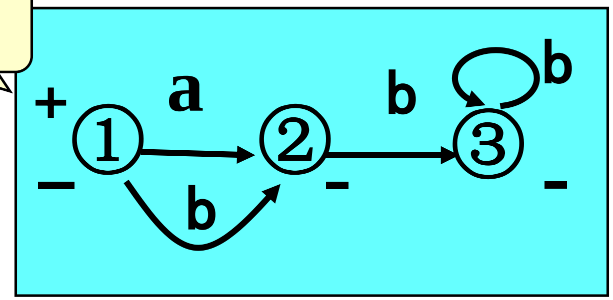
※ 有限自动机的分类示例

【例3.8 】试分别指出下述有限自动机的分类情况：

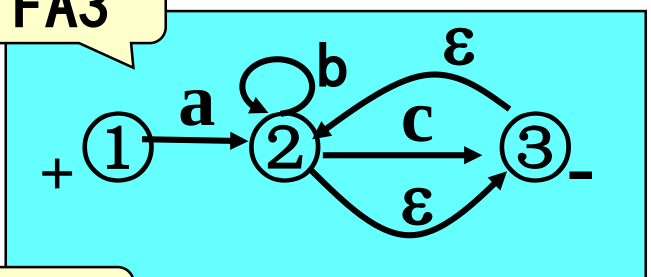
FA1



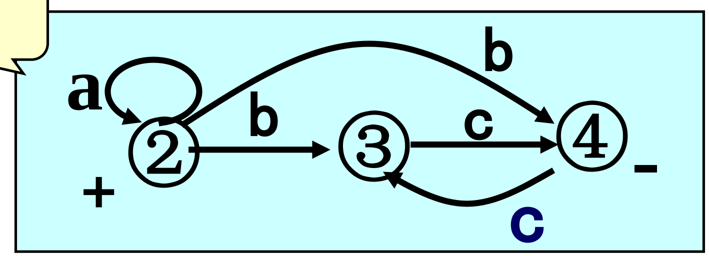
FA2



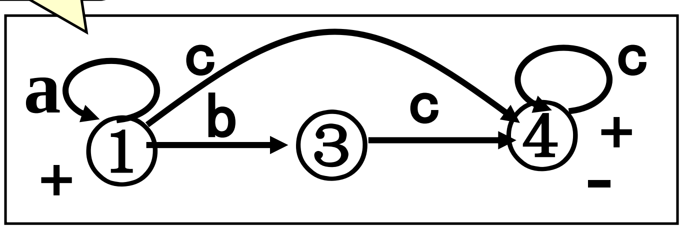
FA3



FA4



FA5



结论：

DFA： FA2

NFA： FA4, FA5（$\bar{\varepsilon}$ NFA）

FA1, FA3 （εNFA）

3.3 有限自动机的等价变换

有限自动机的等价变换，是指把FA1变换为FA2，且有 L(FA1)=L(FA2)，主要包含以下两个内容：

(1) 有限自动机的确定化(NFA=>DFA)；

非确定机（NFA）较易构造，但不好用！

确定机（DFA）较难构造，但好用！

※ 任何一非确定机NFA，皆可通过有效算法把其化为等价的确定机DFA。

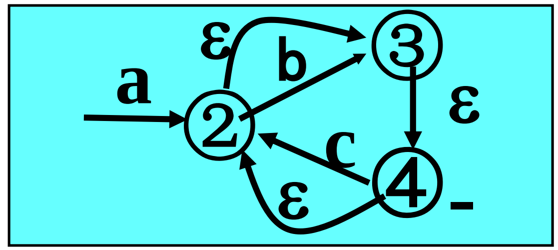
(2) 有限自动机的化简(最小化)；

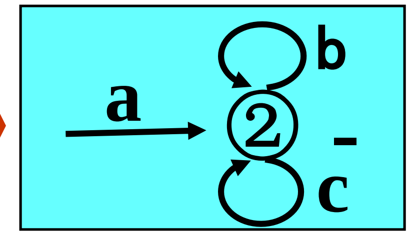
※ 对任何一确定机DFA1，皆可通过有效算法把其化为等价的另一个确定机DFA2，且DFA2的状态最少。

3.3.1 有限自动机的确定化

Ⅰ.消除 ε 边算法(εNFA => $\bar{\varepsilon}$NFA 或 DFA):

(1) 若存在ε闭路,则把ε闭路上各节点合而为一:

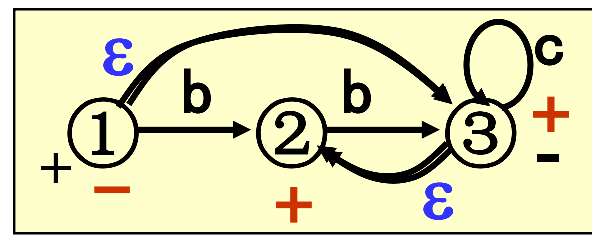




(2) 标记隐含的开始态和结束态:

•开始态的ε通路上的节点:+

•结束态逆向ε通路上节点:-

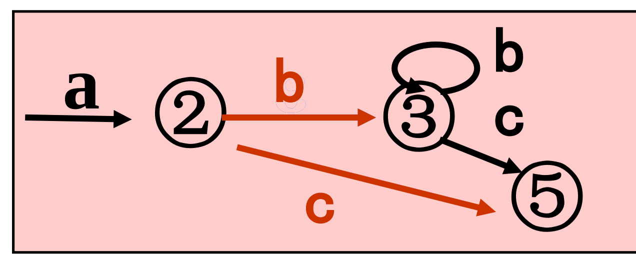


(3) 对剩余的ε边,逆向逐一进行:

① 删除一个ε边;同时

② 引进新边 :

凡由原ε边终点发出的边,也要由其始点发出。

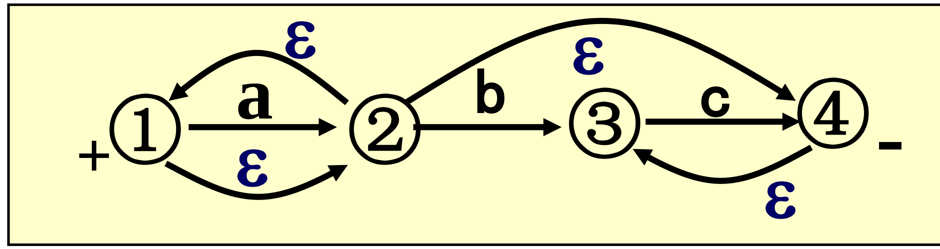


(4) 重复步骤(3),直到再无ε边为止。

※ 消除ε边算法示例:

【例3.9】考查

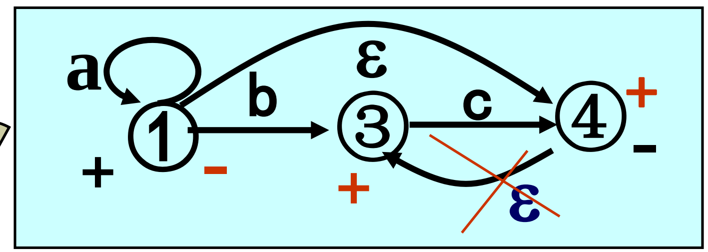
εNFA ：



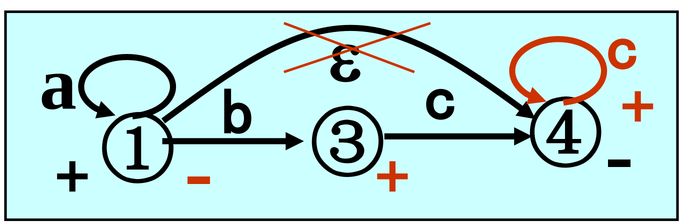
L($\overline{\varepsilon}$NFA)= ?

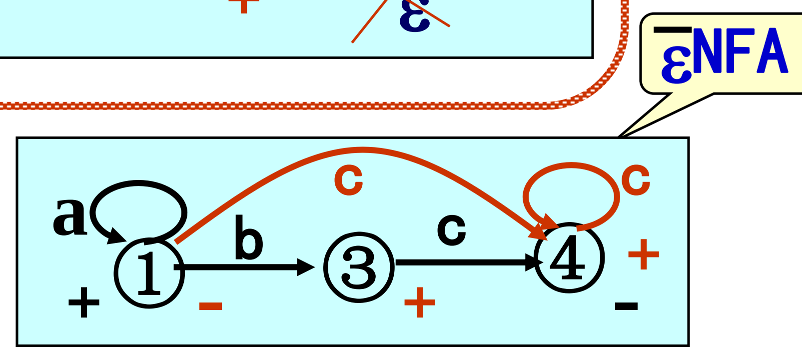
(1) ε闭路上的节点等价（① ≡ ②），可合而为一；

(2) 标记隐含的开始态和结束态： +④，+③，_①



(3) 逆序逐一删除ε边，同时引进新边：

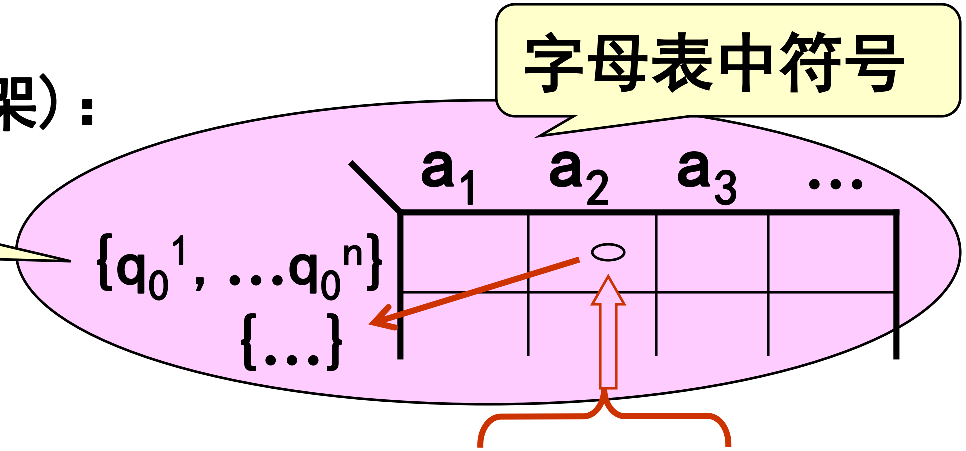




∴ L($\overline{\varepsilon}$NFA)={ $a^m, a^mbc^n, a^mc^n \mid m\geq 0, n\geq 1$ }

Ⅱ. $\overline{\varepsilon}$NFA的确定化算法（ $\overline{\varepsilon}$NFA =>DFA ）：

（1）构造DFA的变换表(框架)：



（2）按 $\overline{\varepsilon}$NFA 的变换函数

实施变换： $\delta(\{q_i^1, \ldots q_i^n\},\ a_k) = \{q_j^1, \ldots q_j^n\}$

（3）若 $\{q_j^1, \ldots q_j^n\}$ 未作为状态行标记，则作新行标记；

（4）重复步骤(2)(3)，直到在再不出现新状态集为止；

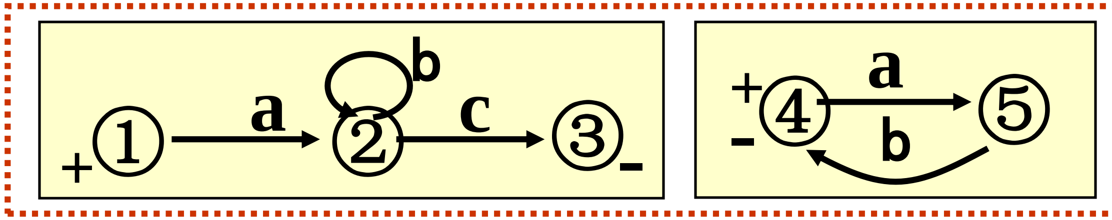
（5）标记DFA的开始态和结束态：

- 第一行 $\{q_0^1, \ldots q_0^n\}$，（右侧）标记 + ；
- 凡是状态行中含有 $\overline{\varepsilon}$NFA 的结束状态者，（右侧）标记 – 。

【注】 必要时，新产生的DFA可用状态图表式。

※ εNFA确定化示例:

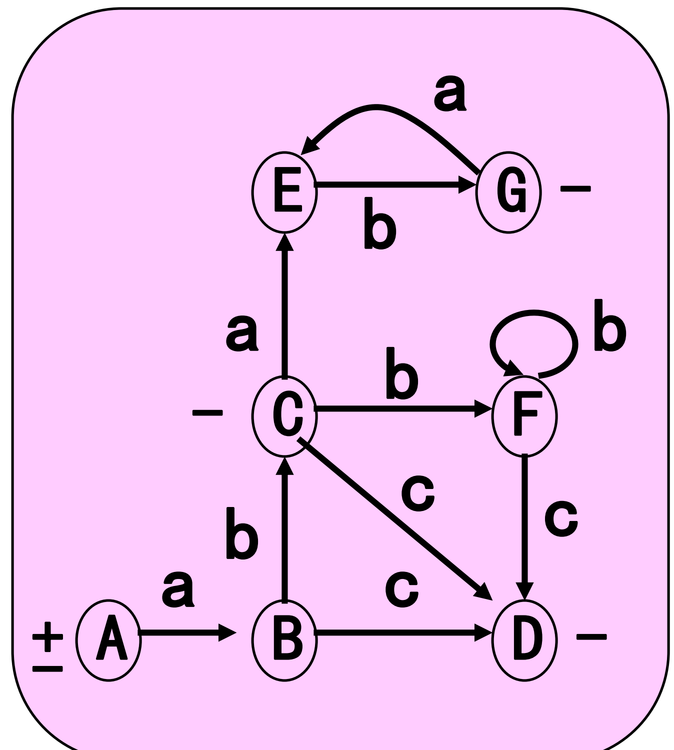
【例3.10】联合式自动机NFA:



※ 确定化过程:

DFA:

		a	b	c
±	A{1, 4}	B{2, 5}		
	B{2, 5}		C{2, 4}	D{3}
−	C{2, 4}	E{5}	F{2}	D{3}
−	D{3}			
	E{5}		G{4}	
	F{2}		F{2}	D{3}
−	G{4}	E{5}		



【注】A, B, C, ... 状态集的代码

3.3.2 有限自动机的最小化1

有限自动机的**最小化**,又称有限自动机的**化简**；是指：对给定的确定机DFA1,构造另一个确定机DFA2,使得 L(DFA1)=L(DFA2),且DFA2的状态最少。

有限自动机最小化算法，是指构造满足下述条件的确定有限自动机(称为**最小机**)：

(1) 删除无用状态；

(2) 合并等价状态。

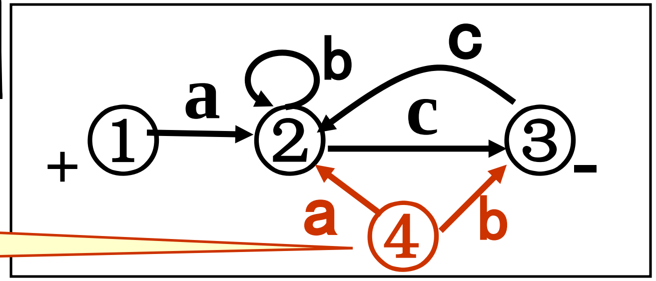
其中

【无用状态】是指由开始态达不到的状态(不可达)或者由其出发不能到达结束态的状态(不终结)。

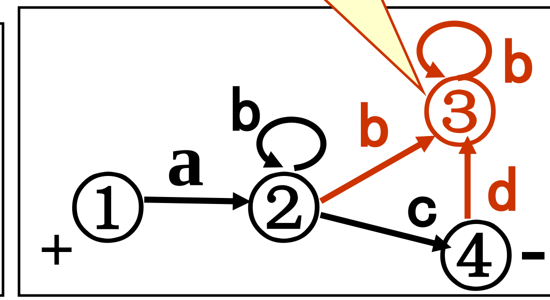
【等价状态】是指这样的两个状态，若分别把其看作开始态，二者接受的符号串集合相同。

※ 无用状态和等价状态示例:

【例3.11】

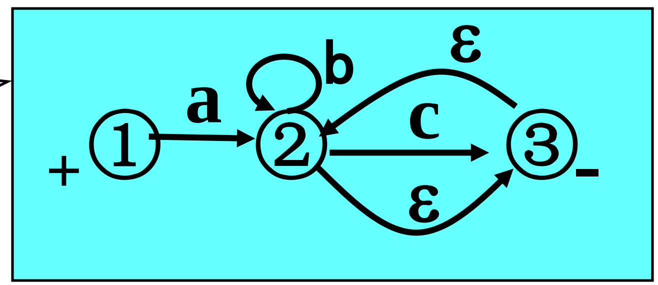


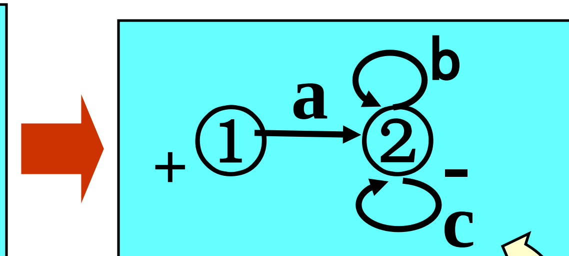
不可达



不终结

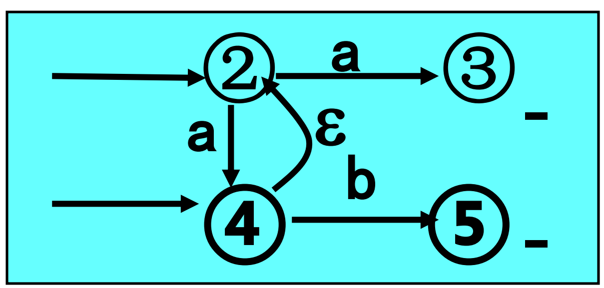
【例3.12】





∵ 2, 3节点构成ε闭路, ∴2, 3等价; 可合而为一

【例3.13】



判断等价性:

4≡5 ? ✕ 3≡5 ? √

2≡3 ? ✕ 2≡4 ? ✕

【注】如何判断两个状态的等价性? 稍后再讨论。

Ⅰ. 删除无用状态算法

※【删除不可达状态】 构造可达状态集 Q_{AR}

(1) 设 q_0 为开始态，则 令 $q_0 \in Q_{AR}$ ；

(2) 若 $q_i \in Q_{AR}$ 且有 $\delta(q_i, a) = q_j$ 则令 $q_j \in Q_{AR}$ ；

(3) 重复执行(2)，直到Q_{AR}不再增大为止。

(4) 从状态集Q中，删除不在Q_{AR}中的所有状态。

※【删除不终结状态】 构造可终结状态集 Q_{FN}

(1) 设 q_i 为结束态，则 令 $q_i \in Q_{FN}$ ；

(2) 若 $q_j \in Q_{FN}$ 且有 $\delta(q_i, a) = q_j$ 则令 $q_i \in Q_{FN}$ ；

(3) 重复执行(2)，直到Q_{FN}不再增大为止。

(4) 从状态集Q中，删除不在Q_{FN}中的所有状态。

Ⅱ. 合并等价状态原理与算法1

※【合并等价状态原理】

两个状态i，j等价，当且仅当满足下面两个条件：

① 必须同是结束态，或同不是结束态；

② 对所有字母表上符号，状态i，j必变换到等价状态。

【例3.14】把下述自动机最小化：

（1）初分成两个不等价子集：

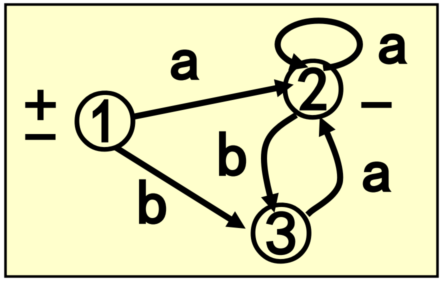
$Q_1=\{1,2\}$，$Q_2=\{3\}$

（2）还能分成不等价子集吗？

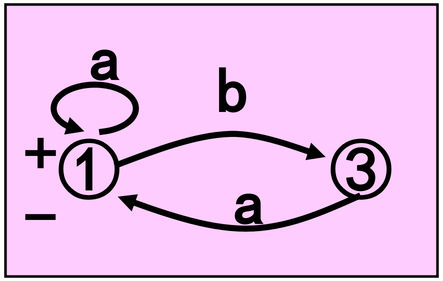
∵ $\delta(\{1,2\},a)=2$

又 $\delta(\{1,2\},b)=3$

∴ $1\equiv2$（等价）



合并后的最小自动机



II．合并等价状态原理与算法2

※【合并等价状态算法】— 划分不等价状态集

（1）初始，把状态集Q化分成两个不等价子集：

Q_1(结束状态集)，　Q_2(非结束状态集)；

（2）把每个Q_i再划分成不同的子集，条件是：

同一Q_i中两个状态i, j, 若对字母表中的某个符号，变换到已划分的不同的状态集中，则i, j应分离：

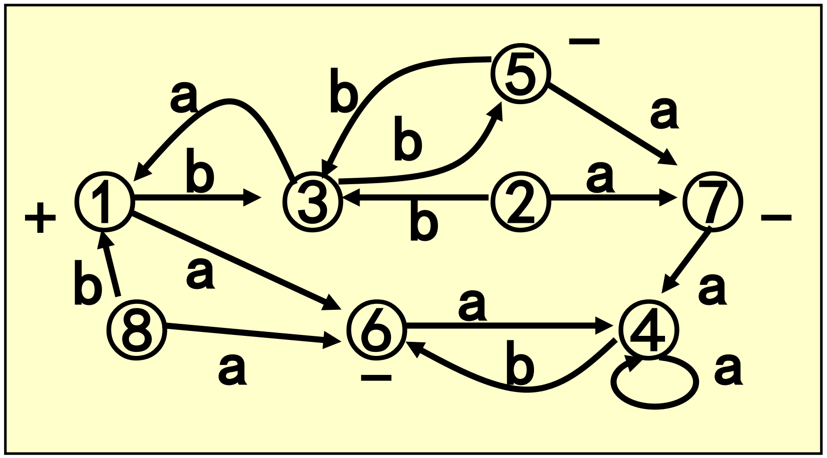
如 $\delta(i, a) \in Q_m$ ，$\delta(j, a) \in Q_n$ 且 $m \neq n$

（3）重复步骤(2)，直到再不能划分为止；

（4）合并最终划分的每个子集中的各状态(合而为一)。

※ 有限自动机化简示例:

【例3.15】 化简下述 DFA:



DFA的变换表:

		a	b
+	1	6	3
	3	1	5
−	6	4	
−	5	7	3
	4	4	6
−	7	4	

Ⅰ. 删除无用状态(不终结):

动态构造DFA变换表，即从开始状态 1 出发，把变换后的状态填入表项，并同时作为新行标记；如此下去，直到再不出现新状态为止。未出现的状态，就是无用的状态。

【注】 DFA 中的状态 2，8 被删除!

II． 合并等价状态：

不等价集

① 令 Q_{NE}= { {1, 3, 4}， {5, 6, 7} }

② 取 {1, 3, 4} ：

∵ $\delta(1, a)=6$ ，$\delta(\{3, 4\}, a)=\{1, 4\}$

∴ 划分成 $Q_1=\{1\}$， $Q_2=\{3, 4\}$；

即 Q_{NE}= { {1}，{3, 4}， {5, 6, 7} }

③ 取 {3, 4}：

∵ $\delta(3, a)=1$ ，$\delta(4, a)=4$

∴ 划分 $Q_1=\{3\}$， $Q_2=\{4\}$

即 Q_{NE}= { {1}，{3}，{4}， {5, 6, 7} }

④ 取 {5, 6, 7}： 同理，可划分成 $Q_1=\{5\}$， $Q_2=\{6, 7\}$；

最后： Q_{NE}= { {1}，{3}，{4}， {5}，{6, 7} }

DFA：

		a	b
+	1	6	3
	3	1	5
−	6	4	
−	5	7	3
	4	4	6
−	7	4	

※ （接上页）：

DFA:

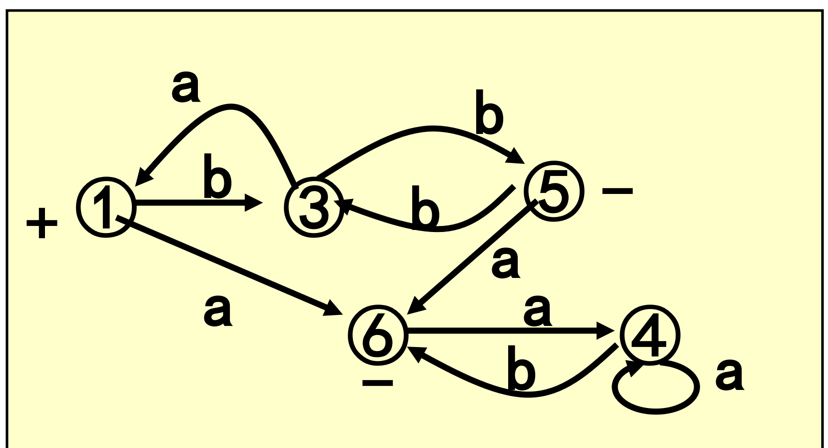
		a	b
+	1	6	3
	3	1	5
−	6	4	
−	5	7	3
	4	4	6
~~−~~	~~7~~	~~4~~	

合并等价状态：
{6，7}

6替换7

		a	b
+	1	6	3
	3	1	5
−	6	4	
−	5	6	3
	4	4	6

最小的
DFA ！



谢谢收看！