

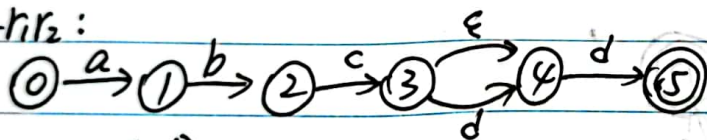
### 3.8.4.

- 算法:
- ① 从头开始扫描, 找到第一个 "/" 字符
  - ② 将 "/" 前部分作为  $r_1$ , "/" 后部分作为  $r_2$
  - ③ 分别对  $r_1$  与  $r_2$  进行解析, 使用 Thompson 算法  
由  $RE \rightarrow NFA$  (正则表达式  $\rightarrow$  NFA)
  - ④ 对  $r_1$  的 NFA 的每一个终态, 添加一条  $\epsilon$  边到  $r_2$  初态, 形成  $NFA - r_1 r_2$
  - ⑤ 对  $NFA - r_1 r_2$  使用 子集构造法, 得到未化简的 DFA.
  - ⑥ 对未化简 DFA 使用 Hopcroft 算法, 得到化简的 DFA.
  - ⑦ 对最小化 DFA 进行词法分析  
若输入字符串读入所有字符后, 当前状态为可接受, 则识别成功, 否则识别失败.

(1)  $(abcd|abc)/d$

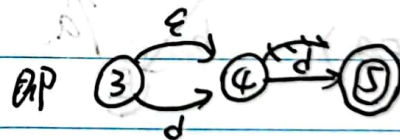
可转化为  $(abc + \epsilon/d) + d$ , 其中  $r_1 = abcd|abc$ ,  $r_2 = d$ .

$NFA - r_1 r_2$ :

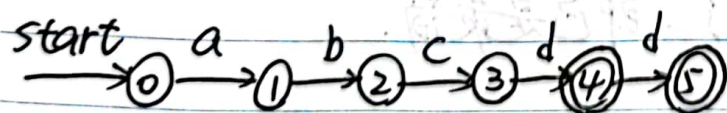


可对 3, 4, 5 进行最小化 DFA 如下

<u>I</u>	<u>I<sub>d</sub></u>	标记为	<u>I</u>	<u>I<sub>d</sub></u>
$\{3, 4\}$	$\{4, 5\}$	$\rightarrow$	A	B(终)
$\{4, 5\}$	$\{5\}$		B	C(终)



故得到最小化 DFA:



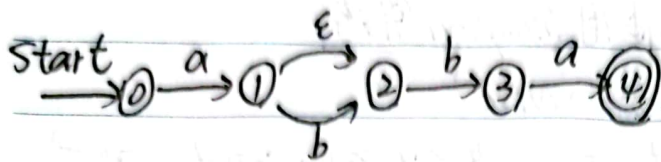
对输入字符串按照这个 DFA 进行状态转移.

若某状态下无法匹配则失败, 否则成功.

(2)  $(a|ab)/ba$

$r_1 = a|ab$      $r_2 = ba$

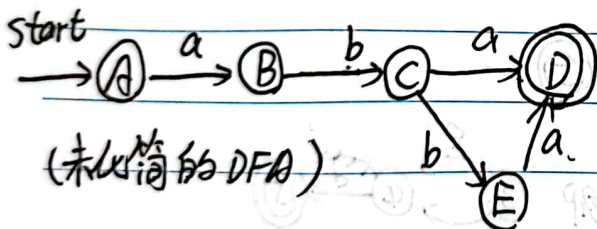
可画出其 NFA- $r_1 r_2$  为



做子集构造.

$I$	$I_a$	$I_b$	标记为 $I$	$I_a$	$I_b$
$\{0\}$	$\{1, 2\}$		$A = \{0\}$	$B$	
$\{1, 2\}$		$\{2, 3\}$	$B = \{1, 2\}$		$C$
$\{2, 3\}$	$\{4\}$	$\{3\}$	$C = \{2, 3\}$	$D$	$E$
$\{4\}$			(终结) $D = \{4\}$		
$\{3\}$	$\{4\}$		$E = \{3\}$	$D$	

得到未化简的 DFA

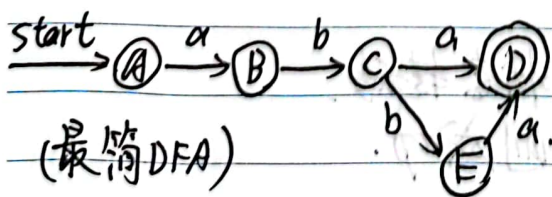


使用 Hopcroft 算法做最小化 DFA.

$I^{(1)} = \{A, B, C, E\}$      $I^{(2)} = \{D\}$     显然  $I_a^{(1)} = \{B, D\}$ . 继续分.

$I^{(1)} = \{A\}$      $I^{(2)} = \{C, E\}$      $I^{(2)} = \{D\}$     显然  $A$  与  $C$  与  $E$  不等价.

故分为  $\{A\} \{B\} \{C\} \{D\} \{E\}$ . 即已是最简.



对输入字符串按照这个 DFA 进行状态转移.  
若某状态下无法匹配则失败, 否则成功.



(3)  $aa^*/a^*$

$$r_1 = aa^*$$

$$r_2 = a^*$$

NFA- $r_1, r_2$

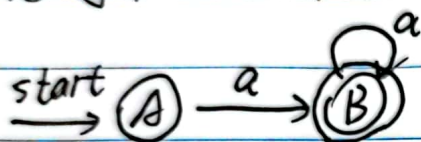


(为方便子集构造, 将  $\epsilon$  边尽可能压缩)

做子集构造.

$I$	$Ia$	标记为	$I$	$Ia$
$\{0\}$	$\{1, 2, 3\}$		$A$	$B$ (终态)
$\{1, 2, 3\}$	$\{1, 2, 3\}$		$B$	$B$

得到未化简 DFA



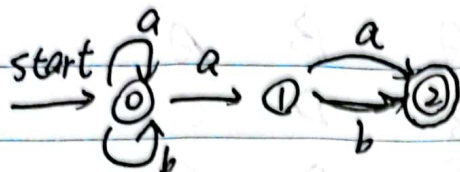
事实上, 这显然已经是最简的了.

对输入字符串按照这个 DFA 进行状态转移,  
若某状态下无法匹配则失败, 否则成功.

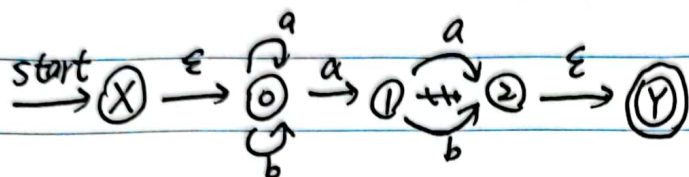
3.9.4

(1)  $(a|b)^* a (a|b)$

① NFA - M



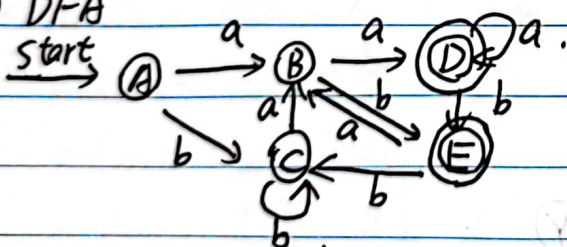
② NFA - M' (按<sup>PPT</sup>要求, 转为 NFA - M')



③ 使用子集构造法实现 NFA → DFA.

I	I <sub>a</sub>	I <sub>b</sub>		I	I <sub>a</sub>	I <sub>b</sub>
{X, 0}	{0, 1}	{0}		A	B	C
{0, 1}	{0, 1, 2, Y}	{0, 2, Y}	标记为	B	D	E
{0}	{0, 1}	{0}		C	B	C
{0, 1, 2, Y}	{0, 1, 2, Y}	{0, 2, Y}		D	D	E
{0, 2, Y}	{0, 1}	{0}		E	B	C

④ DFA



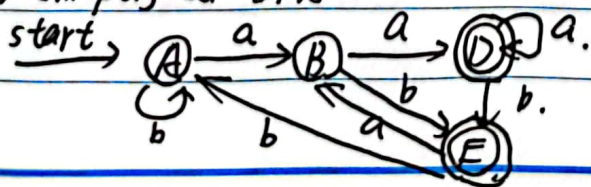
⑤ 通过 Hopcroft 算法 获得 simplified DFA.

$I^{(1)} = \{A, B, C\}$   $I^{(2)} = \{D, E\}$  发现  $I_a^{(1)} = \{B, D\}$  有问题, 重划为如下.

$I^{(1)} = \{A, C\}$   $I^{(2)} = \{B\}$   $I^{(2)} = \{D, E\}$ . 发现  $I_a^{(2)} = \{D, B\}$  有问题, 重划为

$I^{(1)} = \{A, C\}$   $I^{(2)} = \{B\}$   $I^{(2)} = \{D\}$   $I^{(2)} = \{E\}$  这是最终的了.

⑥ simplified DFA.



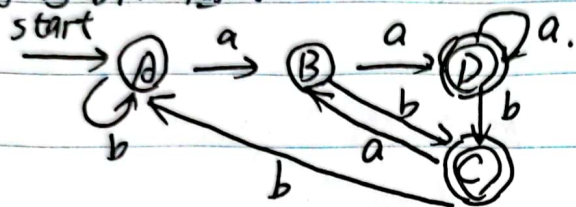


⑦ 对这一小问可以进一步思考。

若我们直接对  $NFA-M$  做子集构造法。

$I$	$I_a$	$I_b$		$I$	$I_a$	$I_b$
$\{0\}$	$\{0,1\}$	$\{0\}$	标记为	$A$	$B$	$A$
$\{0,1\}$	$\{0,1,2\}$	$\{0,2\}$		$B$	$D$	$C$
$\{0,1,2\}$	$\{0,1,2\}$	$\{0,2\}$		(终) $D$	$D$	$C$
$\{0,2\}$	$\{0,1\}$	$\{0\}$		(终) $C$	$B$	$A$

得到 DFA 图。



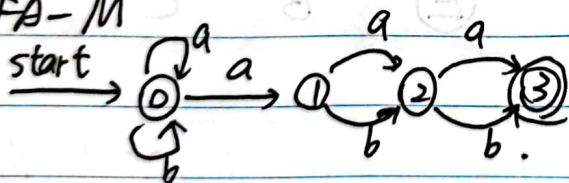
我们惊喜地发现，这就是我们要的最小 DFA。

跟前面的是一样的。

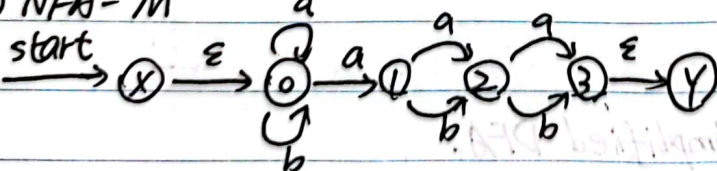
////////// 分割线 ////////// 分割线 ////////// 分割线 ////////// 分割线

(2)  $(a|b)^* a (a|b) (a|b)$

①  $NFA-M$



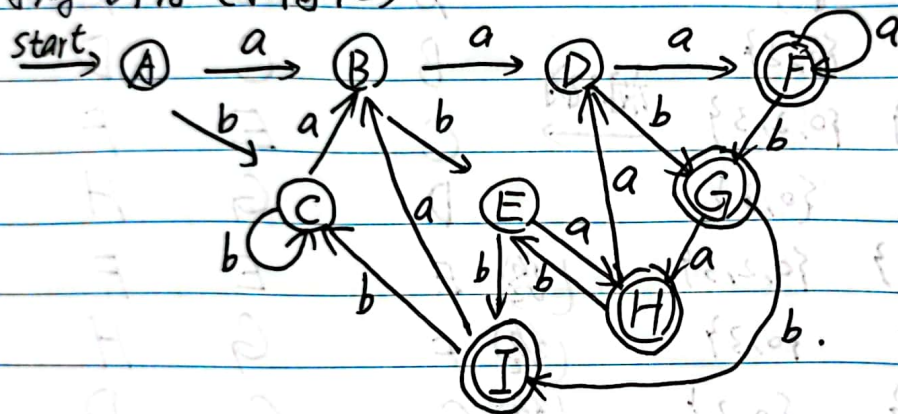
②  $NFA-M'$



③ 使用子集构造法:  $NFA \rightarrow DFA$

I	I <sub>a</sub>	I <sub>b</sub>		I	I <sub>a</sub>	I <sub>b</sub>
{x, 0}	{0, 1}	{0}		A	B	C
{0, 1}	{0, 1, 2}	{0, 2}		B	D	E
{0}	{0, 1}	{0}		C	B	C
{0, 1, 2}	{0, 1, 2, 3, Y}	{0, 2, 3, Y}	标记为	D	F	G
{0, 2}	{0, 1, 3, Y}	{0, 3, Y}		E	H	I
{0, 1, 2, 3, Y}	{0, 1, 2, 3, Y}	{0, 2, 3, Y}	(终)	F	F	G
{0, 2, 3, Y}	{0, 1, 3, Y}	{0, 3, Y}	(终)	G	H	I
{0, 1, 3, Y}	{0, 1, 2}	{0, 2}	(终)	H	D	E
{0, 3, Y}	{0, 1}	{0}	(终)	I	B	C

④ 可得 DFA (未简化)



⑤ 使用 Hopcroft 算法得到简化 DFA.

$I^{(1)} = \{A, B, C, D, E\}$      $I^{(2)} = \{F, G, H, I\}$      $I_a^{(1)} = \{B, D, E, F, H\}$  有问题  
 $I^{(11)} = \{A, B, C\}$      $I^{(12)} = \{D, E\}$      $I^{(2)} = \{F, G, H, I\}$      $I_a^{(11)} = \{B, D\}$  有问题.  
 $I^{(111)} = \{A, C\}$      $I^{(112)} = \{B\}$      $I^{(12)} = \{D, E\}$      $I^{(2)} = \{F, G, H, I\}$      $I_a^{(112)} = \{F, H, D, B\}$   
 $I^{(1111)} = \{A, C\}$      $I^{(1112)} = \{B\}$      $I^{(112)} = \{D, E\}$      $I^{(21)} = \{F, G\}$      $I^{(2)} = \{H, I\}$ .

此时  $I_a^{(112)} = \{F, H\}$  这要继续划分  $I_b^{(21)} = \{G, I\}$  也要继续划分.

$I_a^{(22)} = \{B, E\}$  也不存在, 要继续划分, 而  $I^{(111)} = \{A, C\}$  的  $I_a^{(111)}, I_b^{(111)}$

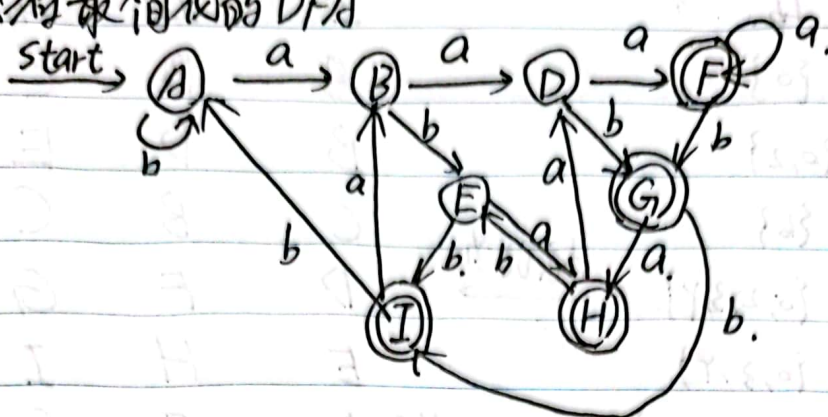
故划分为  $\{A, C\} \{B\} \{D\} \{E\} \{F\} \{G\} \{H\} \{I\}$ .

可取  $\{A\}$  代表  $\{A, C\}$

都无需再一步



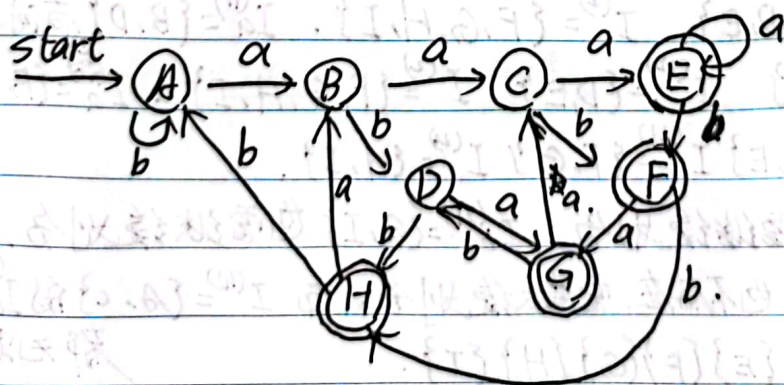
⑥ 求得最简化的 DFA



⑦ 对这一小问也可以做深入思考, 如果我们直接对 NFA-M ~~进行~~ 做子集构造.

I	I <sub>a</sub>	I <sub>b</sub>		I	I <sub>a</sub>	I <sub>b</sub>
{0}	{0, 1}	{0}		A	B	A
{0, 1}	{0, 1, 2}	{0, 2}		B	C	D
{0, 1, 2}	{0, 1, 2, 3}	{0, 2, 3}	标注为	C	E	F
{0, 2}	{0, 1, 3}	{0, 3}		D	G	H
{0, 1, 2, 3}	{0, 1, 2, 3}	{0, 2, 3}	(终)	E	E	F
{0, 2, 3}	{0, 1, 3}	{0, 3}	(终)	F	G	H
{0, 1, 3}	{0, 1, 2}	{0, 2}	(终)	G	C	D
{0, 3}	{0, 1}	{0}	(终)	H	B	A

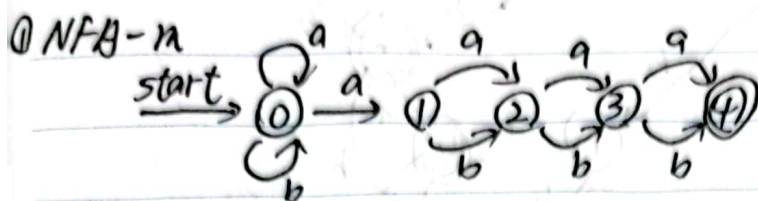
得到 DFA 图.



惊喜地发现, 这就是化简后的 DFA, 一模一样.

由此我大胆猜测, 若原 NFA 足够简洁, 其子集构造后的 DFA 也能足够简洁, 甚至就是最简的 DFA (仅为猜测).

(3)  $(a|b)^*a(a|b)(a|b)(a|b)$



这次直接上手转简法的NFA.

② 子集构造法:  $NFA \rightarrow DFA$ . 标记为,

I	Ia	Ib	标记为,	I	Ia	Ib
{0}	{0, 1}	{0}		A	B	A
{0, 1}	{0, 1, 2}	{0, 2}		B	C	D
{0, 1, 2}	{0, 1, 2, 3}	{0, 2, 3}		C	E	F
{0, 2}	{0, 1, 3}	{0, 3}		D	G	H
{0, 1, 2, 3}	{0, 1, 2, 3, 4}	{0, 2, 3, 4}		E	I	J
{0, 2, 3}	{0, 1, 3, 4}	{0, 3, 4}		F	K	L
{0, 1, 3}	{0, 1, 2, 4}	{0, 2, 4}		G	M	N
{0, 3}	{0, 1, 4}	{0, 4}		H	O	P
{0, 1, 2, 3, 4}	{0, 1, 2, 3, 4}	{0, 2, 3, 4}	(终)	I	I	J
{0, 2, 3, 4}	{0, 1, 3, 4}	{0, 3, 4}	(终)	J	K	L
{0, 1, 3, 4}	{0, 1, 2, 4}	{0, 2, 4}	(终)	K	M	N
{0, 3, 4}	{0, 1, 4}	{0, 4}	(终)	L	O	P
{0, 1, 2, 4}	{0, 1, 2, 3}	{0, 2, 3}	(终)	M	E	F
{0, 2, 4}	{0, 1, 3}	{0, 3}	(终)	N	G	H
{0, 1, 4}	{0, 1, 2}	{0, 2}	(终)	O	C	D
{0, 4}	{0, 1}	{0}	(终)	P	B	A

③ Hopcroft:  $DFA \rightarrow$  最简DFA : 未发现等价状态.

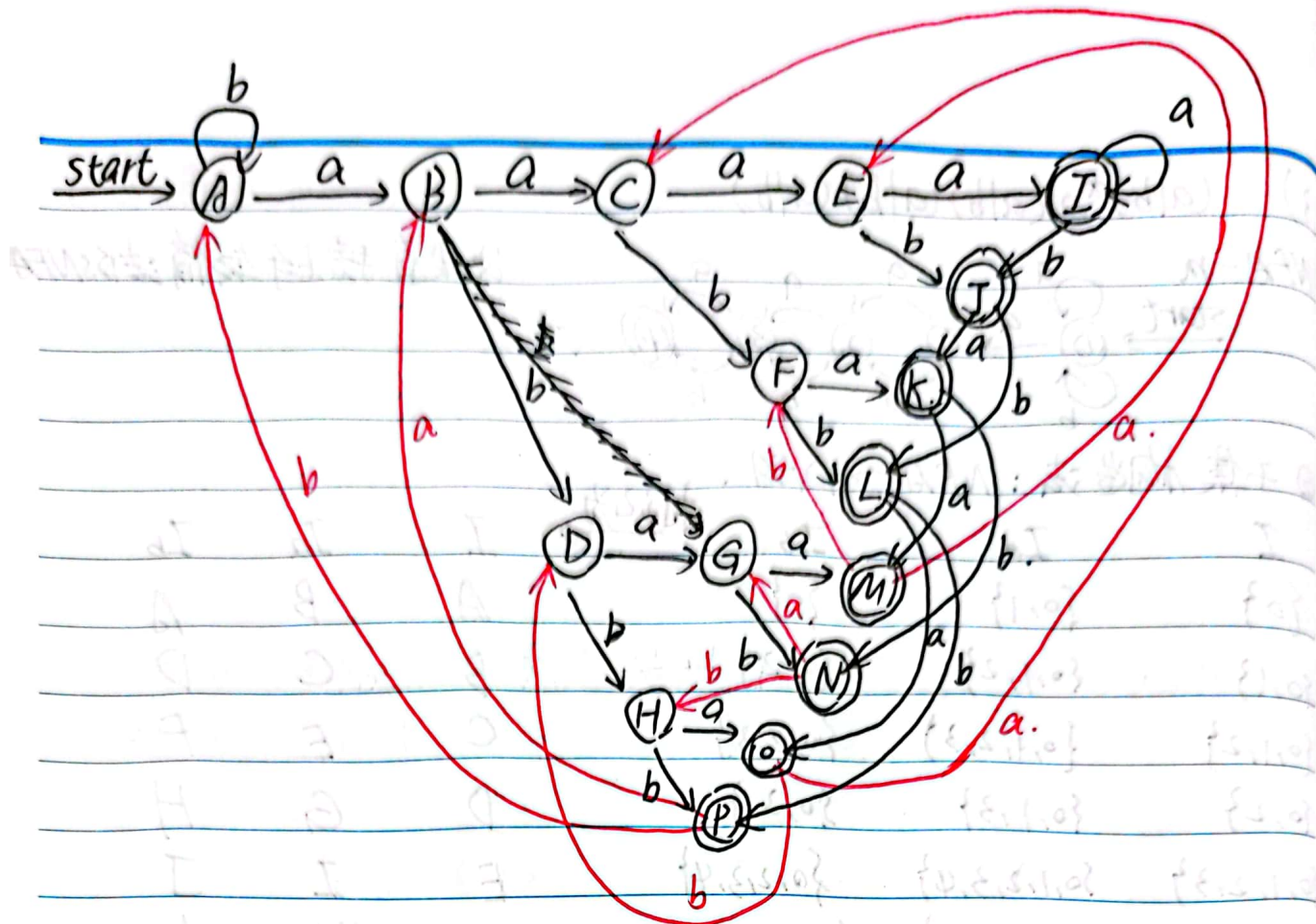
最终得到

{A} {B} {C} {D} {E} {F} {G} {H}

{I} {J} {K} {L} {M} {N} {O} {P}.

④ 得到最简DFA.





对这张图，对照状态转移表，可以发现，

前8个状态关系产生了所有终态，（确定了所有状态，很快完整）

后8个状态关系将这些非终态指向之前的非终态或互相联系

综合上述三题，可以发现终态占所有状态的一半（在最终DFA中）。