第五章

正则语言的性质

例5-1 判断下列语言是否是正则语言？

1. $L = \{0^m 1^n | m, n \geq 0\}$
2. $L = \{0^m 1^n | m \geq 2, n \geq 4\}$
3. $L = \{0^n 1^n | n \geq 0\}$

1. $0^* 1^*$
2. $000^* 11111^*$
3. 直观理解：需要有穷自动机在扫描0的时候记住其数量，从而匹配1的数量，而有穷自动机无法记住任意数量的0，所以不是正则语言

- $L = \{0^n1^n | n \geq 0\}$不是RL不是因为某人构造不出产生该语言的RG、FA或者RE，而是从RL的性质出发，可以证明该语言不具有RL的性质，所以不存在这些形式的描述。RL的泵引理及其应用；RL关于并、乘积、闭包、补、交、正则代换、同态、逆同态等运算的封闭性。

- 另一个问题是对于给定的RL，存在多个产生该语言的正则文法、FA和多个RE，由于他们描述的语言相同，所以他们“本质上”应该是一样的，因此，人们希望能找到一种方法，构造出接受这个RL的状态最少的确定性有穷状态自动机—最小DFA。右不变的等价关系、 Myhill-Nerode定理、DFA极小化等。

章节目录

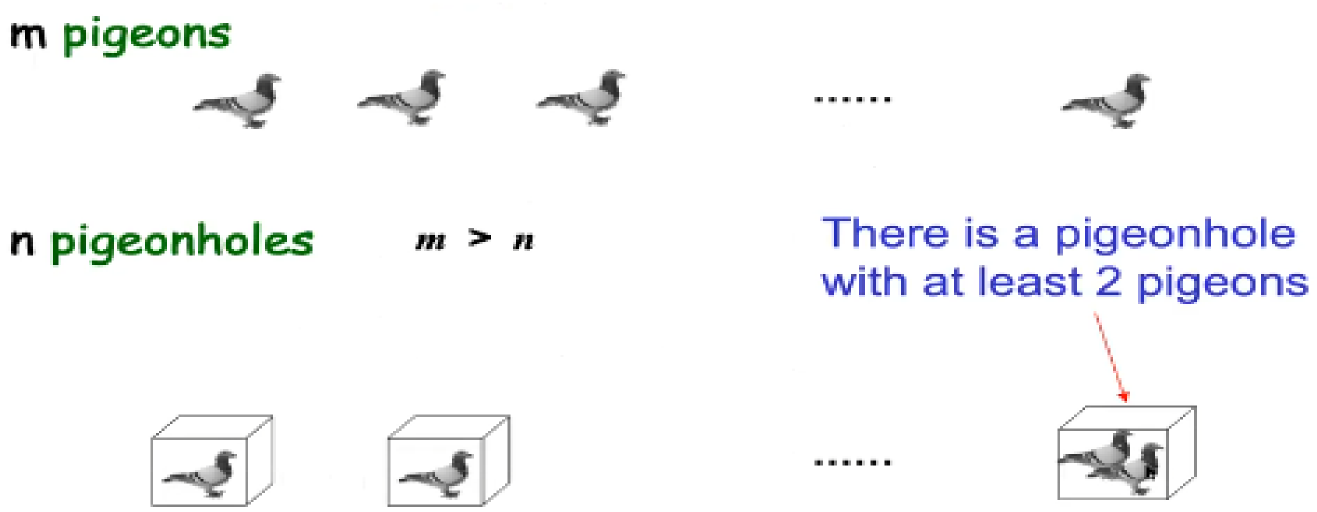
5.1 正则语言的泵引理

- 任何有穷语言都是RL，所以，非RL一定是无穷语言。因此本节只讨论无穷语言是否为RL的判定问题。

- DFA是RL的识别模型，一个DFA只有有穷个状态，也就是说，当该DFA识别的语言L是无穷语言时，L中必定存在一个足够长的句子，使得DFA在识别该句子的过程中，肯定要重复地经过某些状态。

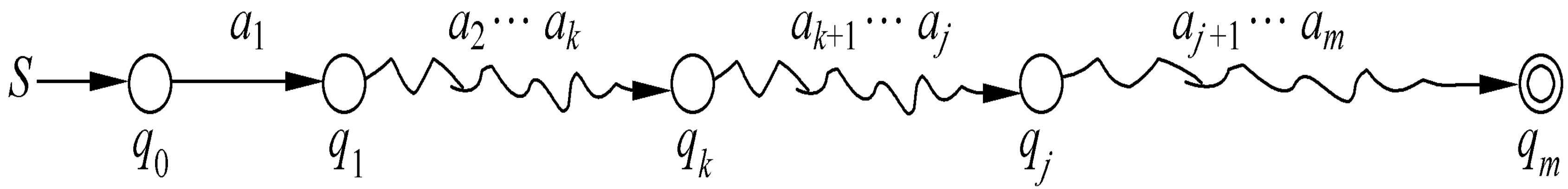
- 例如句子$z = a_1 \cdots a_m \in L$，不妨假设DFA在识别它的过程中需要经过的状态数依次为$q_1, q_2, \cdots, q_m$，即$q_i = \delta(q_0, a_1 \cdots a_i)$，当$m$大于或等于DFA所有可达状态的个数时($m \geq N$)，由鸽巢原理，所以在DFA的状态序列中至少有一对是重复的。

鸽巢原理：m只鸽子，n个鸽笼，且满足$m > n$, 则一定存在装了至少两个鸽子的鸽笼。

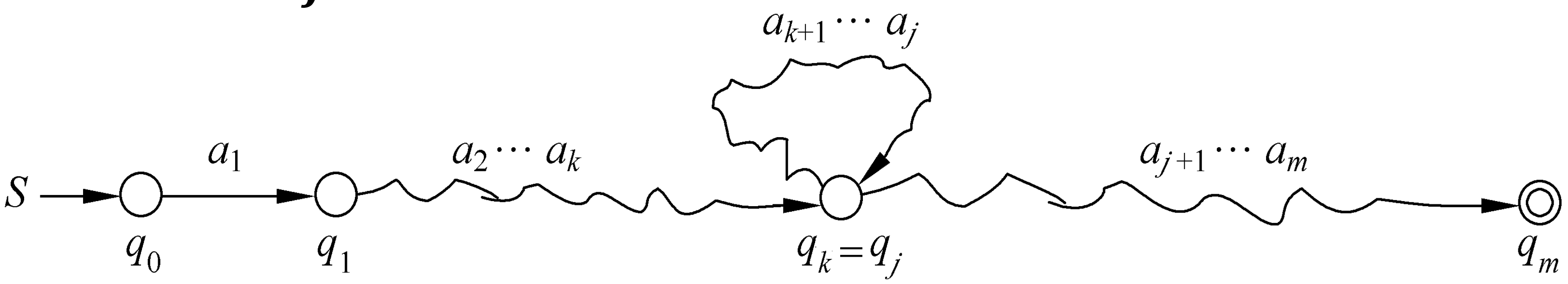


5.1 正则语言的泵引理

- 取$u = a_1 a_2 \cdots a_k$，$v = a_{k+1} \cdots a_j$，$w = a_{j+1} \cdots a_m$



- 不妨设q_k与q_j是最早出现的相同状态，即$q_k = q_j$，显然$k < j \le N$
- 由于$q_k = q_j$，则上图可画为



- 对于任意的整数$i \ge 0$，有$uv^i w \in L$。
- 因为$k < j \le N$，故$a_{k+1} \cdots a_j \ne \varepsilon$，即$|v| \ge 1$；$|uv| \le N$。
- 再注意到讨论中的DFA M的任意性，该结论对最小DFA也成立。

引理5-1 正则语言的泵引理：如果语言$\boldsymbol{L}$是正则的, 则存在仅依赖于$\boldsymbol{L}$的正整数$\boldsymbol{N}$, 对于$\forall \boldsymbol{z} \in \boldsymbol{L}$, 如果$|\boldsymbol{z}| \geq \boldsymbol{N}$, 则存在$\boldsymbol{u}$，$\boldsymbol{v}$，$\boldsymbol{w}$，满足：

1. $\boldsymbol{z} = \boldsymbol{uvw}$
2. $|\boldsymbol{uv}| \leq \boldsymbol{N}$
3. $|\boldsymbol{v}| \geq \mathbf{1}$
4. 对于任意整数$\forall \boldsymbol{i} \geq \mathbf{0}, \boldsymbol{uv^iw} \in \boldsymbol{L}$
5. N不大于接受L的最小DFA M的状态数

- 中间的v可以用来泵出新串，所泵的新串还在这个语言中

■DFA在处理一个足够长的句子的过程中，必定会重复地经过某一个状态。

■换句话说，在DFA的状态转移图中，必定存在一条含有回路的从启动状态到某个终止状态的路。

■由于是回路，所以，DFA可以根据实际需要沿着这个回路循环运行，相当于这个回路中弧上的标记构成的非空子串可以重复任意多次。

如何应用泵引理证明语言的非正则性?

例5-2 证明$L = \{0^n 1^n | n \geq 0\}$不是正则语言。

证明：• **反证法，假设是正则语言，利用泵引理推出矛盾。**

1. 假设$L = \{0^n 1^n | n \geq 0\}$是正则的；
2. 则存在$\boldsymbol{N} \in \boldsymbol{Z}^+$，对$\forall \boldsymbol{z} \in \boldsymbol{L}$ $(|\boldsymbol{z}| \geq \boldsymbol{N})$满足泵引理；
3. 从L中取一个特定的字符串$\boldsymbol{z} = 0^N 1^N$，显然$\boldsymbol{z} \in \boldsymbol{L}$且$|\boldsymbol{z}| = \boldsymbol{2N} > \boldsymbol{N}$；
4. 则$\boldsymbol{z}$可被分为$\boldsymbol{z} = \boldsymbol{uvw}$，且$|\boldsymbol{uv}| \leq \boldsymbol{N}$和$|\boldsymbol{v}| \geq \boldsymbol{1}$；
5. 因此，$\boldsymbol{v}$只能是$\boldsymbol{0^k}$，且k> $\boldsymbol{0}$；
6. 那么$\boldsymbol{uv^2w} = 0^{N+k} 1^N \notin L$，而由泵引理$\boldsymbol{uv^2w} \in \boldsymbol{L}$，矛盾；
7. 故假设不成立，$\boldsymbol{L}$不是正则的。

如何应用泵引理证明语言的非正则性?

例5-3 证明$L=\{0^n1^m2^{n+m}|m,n\geq 1\}$不是正则语言。

证明:

1. 假设L是正则语言;
2. 则存在$\boldsymbol{N}\in\boldsymbol{Z}^+$,对$\forall \boldsymbol{z}\in L(|\boldsymbol{z}|\geq \boldsymbol{N})$满足泵引理;
3. 从L中取一个特定的字符串$\boldsymbol{z}=0^N1^N2^{2N}$,显然$\boldsymbol{z}\in\boldsymbol{L}$且$|\boldsymbol{z}|=\boldsymbol{4N}>\boldsymbol{N}$;
4. 则$\boldsymbol{z}$可被分为$\boldsymbol{z}=\boldsymbol{uvw}$,且$|\boldsymbol{uv}|\leq\boldsymbol{N}$和$|\boldsymbol{v}|\geq\boldsymbol{1}$;

如何应用泵引理证明语言的非正则性?

例5-3 证明$L = \{0^n 1^m 2^{n+m} | m, n \geq 1\}$不是正则语言。

证明:

5. 则$\boldsymbol{v}$只能是由0组成的非空串，不妨设$\boldsymbol{v} = \boldsymbol{0^k}, \boldsymbol{k} \geq \boldsymbol{1}$，此时有$\boldsymbol{u} = \boldsymbol{0^{N-k-j}}$，$\boldsymbol{w} = \boldsymbol{0^j 1^N 2^{2N}}$。
6. 那么$\boldsymbol{uv^0w} = 0^{N-k-j} 0^j 1^N 2^{2N} = 0^{N-k} 1^N 2^{2N}$，由于$N - k + N \neq 2N$，故$\boldsymbol{uv^0w} \notin \boldsymbol{L}$，矛盾;
7. 故假设不成立，$\boldsymbol{L}$不是正则的。

如何应用泵引理证明语言的非正则性?

例5-4 证明$L = \{0^n | n\text{为素数}\}$不是正则语言。

证明：
1. 假设L是正则语言；
2. 则存在$\boldsymbol{N} \in \boldsymbol{Z}^+$，对$\forall \boldsymbol{z} \in \boldsymbol{L}$ $(|\boldsymbol{z}| \geq \boldsymbol{N})$满足泵引理；
3. 从L中取一个特定的字符串$\boldsymbol{z} = 0^{N+p}$，其中$N + p$是素数，显然$\boldsymbol{z} \in \boldsymbol{L}$且$|\boldsymbol{z}| = \boldsymbol{N} + \boldsymbol{p} > \boldsymbol{N}$；
4. 则$\boldsymbol{\omega}$可被分为$\boldsymbol{z} = \boldsymbol{uvw}$，且$|\boldsymbol{uv}| \leq \boldsymbol{N}$和$|\boldsymbol{v}| \geq \boldsymbol{1}$；
5. 不妨设$\boldsymbol{v} = \boldsymbol{0^k}, \boldsymbol{k} \geq \boldsymbol{1}$，设$\boldsymbol{w} = \boldsymbol{0^j}$，则$\boldsymbol{u} = \boldsymbol{0^{N+p-k-j}}$。
6. 那么$\boldsymbol{uv^iw} = 0^{N+p-k-j}0^{ki}0^j = 0^{N+p+(i-1)k}$

如何应用泵引理证明语言的非正则性？

例5-4 证明$L = \{0^n | n为素数\}$不是正则语言。

证明：7. 令$i = N + p + 1$，有$N + p + (i - 1)k = N + p + (N + p + 1 - 1)k = (N + p)(k + 1)$，由于$\boldsymbol{k \geq 1}$，所以$(N + p)(k + 1)$不是素数，则$\boldsymbol{uv^{N+p+1}w \notin L}$

8. 而由泵引理$\boldsymbol{uv^iw \in L}$，矛盾。

9. 故假设不成立，$\boldsymbol{L}$不是正则的。

如何应用泵引理证明语言的非正则性?

例5-5 证明$L = \{a^{k^2} | k \ge 1\}$不是正则语言。

证明：1. 假设L是正则语言

2. 则存在$\boldsymbol{N} \in \boldsymbol{Z}^+$，对$\forall \boldsymbol{z} \in L(|\boldsymbol{z}| \ge \boldsymbol{N})$满足泵引理
3. 从L中取一个特定的字符串$\boldsymbol{z} = a^{N^2}$，显然$|\boldsymbol{z}| = \boldsymbol{N}^2 \ge \boldsymbol{N}$
4. 则$\boldsymbol{z}$可被分为$\boldsymbol{z} = \boldsymbol{uvw}$，且$|\boldsymbol{uv}| \le \boldsymbol{N}$和$|\boldsymbol{v}| \ge \mathbf{1}$
5. 则$\boldsymbol{v}$只能是由a组成的非空串，不妨设$\boldsymbol{v} = \boldsymbol{a}^{\boldsymbol{k}}, \mathbf{1} \le \boldsymbol{k} \le \boldsymbol{N}$;
6. 那么$\boldsymbol{uv}^2\boldsymbol{w} = \boldsymbol{a}^{N^2}\boldsymbol{a}^{\boldsymbol{k}} = \boldsymbol{a}^{N^2+k}$，其中$\boldsymbol{N}^2 + \boldsymbol{k} \le \boldsymbol{N}^2 + \boldsymbol{N} < (\boldsymbol{N} + \mathbf{1})^2$，故$\boldsymbol{uv}^2\boldsymbol{w} \notin \boldsymbol{L}$;
7. 而由泵引理$\boldsymbol{uv}^2\boldsymbol{w} \in \boldsymbol{L}$，矛盾;
8. 故假设不成立，$\boldsymbol{L}$不是正则的。

几点注意：

除了证明一个语言不是RL外，有时也希望证明一个语言是RL，最直接的方法是给出该语言的正则文法、或者FA、RE。

泵引理是用来证明一个语言不是 RL的。因为它只是说RL必定满足这些条件，并没有说满足条件的语言是RL。在使用泵引理证明一个给定语言不是RL时，需要注意以下几方面问题：

■由于泵引理给出的是 RL 的必要条件，所以，在用它证明一个语言不是 RL 时，我们使用反证法。

■泵引理中提到的仅依赖于L的正整数N不用给出具体的数值，只用符号N来表示即可。

几点注意：

- 在选择z时，应该选择既可以找出矛盾，又能使证明尽可能简单的特殊的z，在用泵引理证明一个语言不是RL的过程中，z的选取是最困难的。一旦选出了恰当的z，证明基本上就可以顺利地进行下去，因为剩余的叙述基本上是“按部就班”地进行。
- 当一个特意被选择用来“发现矛盾”的z确定以后，就必须说明满足条件$|uv| \leq N$和$|v| \geq 1$的所有v都不能使$uv^iw \in L$对所有$i \geq 0$成立（对“存在u, v, w”的否定）
- 与选z时类似，在寻找i时，也仅需要找到一个表明矛盾的“具体”值就可以了（对“所有i”的否定）

章节目录

<u>定义5-1</u> 封闭性：如果任意的、属于同一语言类的语言在某一特定运算下所得的结果仍然是该类语言，则称该语言类对此运算是封闭的，并称该语言类对此运算具有封闭性。

Callback:

- 设L和M是两个语言
- ■并：$L \cup M$
- ■乘积：$L \cdot M$
- ■闭包：L^*
- ■补：$\overline{L}$
- ■交：$L \cap M$

定理 5-1 RL在并、乘积、闭包运算下是封闭的。

- 根据正则表达式的定义直接得到

Callback:

$\boldsymbol{r}$与$\boldsymbol{s}$的“和”($\boldsymbol{r}+\boldsymbol{s}$)是Σ上的正则表达式，表示语言$R\cup S$

$\boldsymbol{r}$与$\boldsymbol{s}$的“乘积”($\boldsymbol{rs}$)是Σ上的正则表达式，表示语言RS；

$\boldsymbol{r}$的克林闭包$\boldsymbol{r}^*$是Σ上的正则表达式，表示语言R^*；

定理 5-2 RL在补运算下是封闭的。

即如果L是正则语言，那么$\overline{L} = \Sigma^* - L$也是正则语言。

证明：

① 设接受语言L的DFA为$M = (Q, \Sigma, \delta, q_0, F)$，满足$L(M) = L$，构造DFA $M' = (Q, \Sigma, \delta, q_0, Q - F)$

② 显然对于任意的$x \in \Sigma^*$：

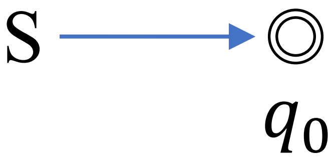
$$x \in \overline{L} \Leftrightarrow \delta(q_0, x) \notin F \Leftrightarrow \delta(q_0, x) \in Q - F \Leftrightarrow x \in L(M')$$

③ 即$x \in \overline{L} \Leftrightarrow x \in L(M')$，即$\overline{L} = L(M')$。

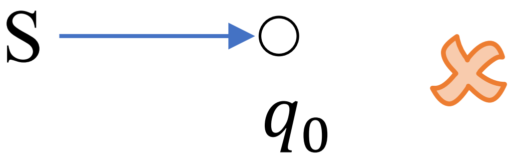
④ 所以RL在补运算下是封闭的，定理得到证明。

5.2 正则语言的封闭性

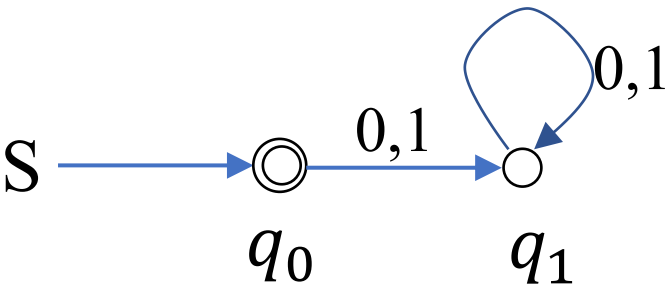
例5-6 若$\Sigma = \{0, 1\}$，$L = \{\varepsilon\}$的DFA如图，请给出的$\overline{L}$ DFA。

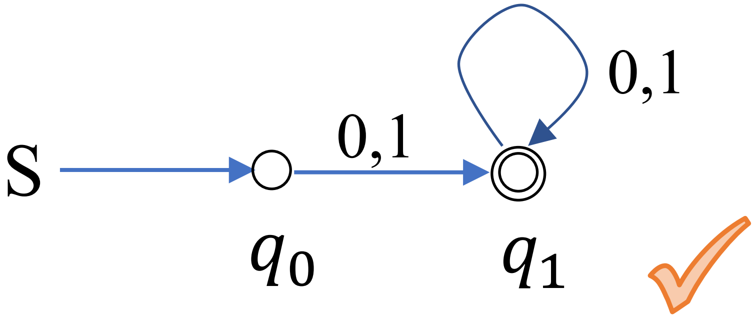


直接翻转状态？



应使用下面完整的DFA去求补





例5-7 证明 $L=\{\omega|\omega$由数量不相等的0和1构成$\}$ 不是正则语言。

- 由泵引理不易直接证明L不是正则的，因为无论如何取z，将其分成三部分$z=uvw$时，都不易产生L之外的串；
- 而证明$\bar{L}$ 不是正则的很容易（参见例5-2）
  $\bar{L}=\{0^n1^n|n\geq 0\}$
- 由补运算的封闭性，可得到L不是正则的。

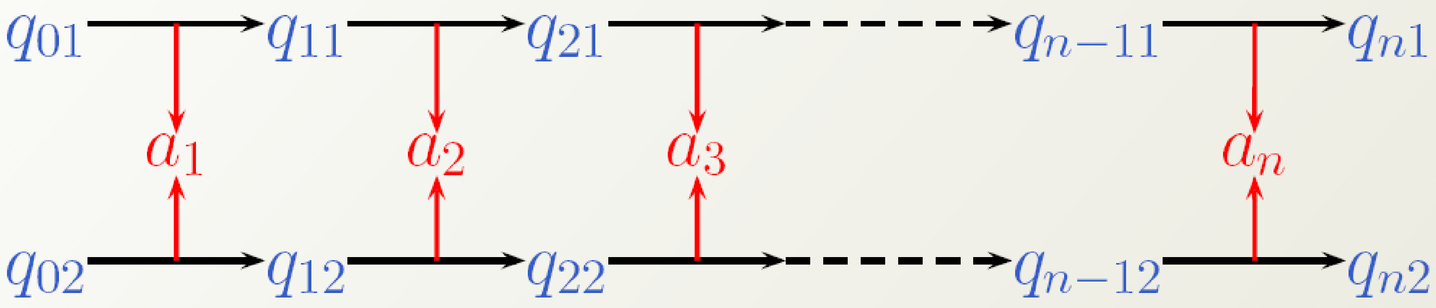
定理 5-3 RL在交运算下封闭。

证明思路：

设L_1, L_2为任意给定的RL，相应的DFA为

$M_1=(Q_1,\Sigma,\delta_1,q_{01},F_1), M_2=(Q_2,\Sigma,\delta_2,q_{02},F_2)$

那么，如果一个串$\omega=a_1a_2\cdots a_n\in L_1\cap L_2$当且仅当$\omega$同时被$M_1$和$M_2$接受



并行计算

$M=(Q_1\times Q_2,\Sigma,\delta,(q_{01},q_{02}),F_1\times F_2)$

$\delta((q_1,q_2),a)=(\delta_1(q_1,a),\delta_2(q_2,a))$

■定义5-2 代换(substitution)

设Σ、Δ是两个字母表，映射$f:\Sigma \to 2^{\Delta^*}$被称为是从$\Sigma$到$\Delta$的代换。如果对于$\forall a \in \Sigma$，$f(a)$是$\Delta$上的**RL**，则称$f$为正则代换。

■先将$\boldsymbol{f}$的定义域扩展到$\boldsymbol{\Sigma}^*$上：

$$\boldsymbol{f}:\boldsymbol{\Sigma}^* \to \boldsymbol{2}^{\Delta^*}$$

(1)f(ε)={ε}；

(2) f(xa)=f(x)f(a)。

■再将$\boldsymbol{f}$的定义域扩展到$\boldsymbol{2}^{\Sigma^*}$：

$$\boldsymbol{f}:\boldsymbol{2}^{\Sigma^*} \to \boldsymbol{2}^{\Delta^*}$$

对于$\forall \boldsymbol{L} \subseteq \boldsymbol{\Sigma}^*$，$\boldsymbol{f}(\boldsymbol{L}) = \bigcup_{\boldsymbol{x}\in \boldsymbol{L}} \boldsymbol{f}(\boldsymbol{x})$

例5-8 设$\Sigma = \{0,1\}, \Delta = \{a, b\}$, ，正则代换$f$定义为： $f(0) = a$, $f(1) = b^*$，则:

- $f(010) = f(0)f(1)f(0) = ab^*a$
- $f(\{11,00\}) = f(11) \cup f(00)$
  $= f(1)f(1) \cup f(0)f(0)$
  $= b^*b^* + aa \ = b^* + aa$
- $f(0^*(0+1)1^*) = L(a^*(a+b^*)(b^*)^*)$
  $= L(a^*(a+b^*)b^*)$
  $= L(a^*ab^* + a^*b^*b^*)$
  $= L(a^*b^*)$

定义5-3 Σ、Δ是两个字母表， $f:\Sigma \to 2^{\Delta^*}$是正则代换，则

① $f(\emptyset)=\emptyset$;

② $f(\varepsilon)=\varepsilon$;

③ 对于$\forall a \in \Sigma$， $f(a)$是Δ上的RE;

④ 如果r，s是Σ上的RE，则以下均是Δ上的RE。

- $f(r+s)=f(r)+f(s)$
- $f(rs)=f(r)f(s)$
- $f(r^*)=f(r)^*$

<u>定理 5-4</u> 设L是Σ上的一个RL，$f: \Sigma \rightarrow 2^{\Delta^*}$是正则代换，则$f(L)$也是 RL。

证明：描述工具：RE

对r中运算符的个数n施以归纳，证明f(r)是表示f(L)的RE。

当n=0时,结论成立。

当n≤k时定理成立，即当r中运算符的个数不大于k时：

f(L(r))=L(f(r))。

当n=k+1时，分别进行讨论。

<u>定理 5-4</u> 设L是Σ上的一个RL，$f:\Sigma \to 2^{\Delta^*}$是正则代换，则$f$(L)也是 RL。

(1) $r=r_1+r_2$。

$f(L)=f(L(r))$	
$=f(L(r_1+r_2))$	
$=f(L(r_1)\cup L(r_2))$	RE的定义
$=f(L(r_1))\cup f(L(r_2))$	正则代换的定义
$=L(f(r_1))\cup L(f(r_2))$	归纳假设
$=L(f(r_1)+f(r_2))$	RE的定义
$=L(f(r_1+r_2))$	RE的正则代换的定义
$=L(f(r))$	

定理 5-4 设L是Σ上的一个RL，$f:\Sigma \rightarrow 2^{\Delta^*}$是正则代换，则$f$(L)也是 RL。

(2) $r=r_1r_2$。

$f(L)=f(L(r))$	
$=f(L(r_1r_2))$	
$=f(L(r_1)\ L(r_2))$	RE的定义
$=f(L(r_1))\ f(L(r_2))$	正则代换的定义
$=L(f(r_1))\ L(f(r_2))$	归纳假设
$=L(f(r_1)\ f(r_2))$	RE的定义
$=L(f(r_1r_2))$	RE的正则代换的定义
$=L(f(r))$	

定理 5-4 设L是Σ上的一个RL，$f:\Sigma \rightarrow 2^{\Delta^*}$是正则代换，则$f(L)$也是 RL。

(3) $r=r_1^*$。

$f(L)=f(L(r))$	
$=f(L(r_1^*))$	
$=f(L(r_1)^*)$	RE的定义
$=(f(L(r_1)))^*$	正则代换的定义
$=(L(f(r_1)))^*$	归纳假设
$=L(f(r_1)^*)$	RE的定义
$=L(f(r_1^*))$	RE的正则代换的定义
$=L(f(r))$	

例5-9 设$\Sigma=\{0,1,2\}, \Delta=\{a,b\}$，正则代换$f$定义为：$f(0)=ab$，$f(1)=b^*a^*$，$f(2)=a^*(a+b)$，则

- $f(00)=f(0)f(0)=abab$
- $f(010)=f(0)f(1)f(0)=abb^*a^*ab=ab^+a^+b$
- $f\left((0+1+2)^*\right)=(ab+b^*a^*+a^*(a+b))^*=(a+b)^*$
- $f(0(0+1+2)^*)=ab(ab+b^*a^*+a^*(a+b))^*$
  $=ab(a+b)^*$
- $f(012)=abb^*a^*a^*(a+b)=a\,bb^*a^*(a+b)$
- $f\left((0+1)^*\right)=(ab+b^*a^*)^*=(a+b)^*$

Callback:

■代换 $f:\Sigma\rightarrow 2^{\Delta^*}$

■**定义5-4** 同态映射(**Homomorphism**)

设Σ、Δ是两个字母表，映射$f:\Sigma\rightarrow\Delta^*$，如果对于$\forall x,y\in\Sigma^*$，$f(xy)=f(x)f(y)$，则称$f$为从Σ到$\Delta^*$的同态映射。

同态映射可以理解为正则代换的一种特例。

- 对于$\forall L\subseteq\Sigma^*$，**L**的同态像$f(L)=\cup_{x\in L}f(x)$
- 对于$\forall w\in\Delta^*$，**w**的同态原像是一个集合

$$f^{-1}(w)=\{x|f(x)=w\&x\in\Sigma^*\}$$

- 对于$\forall L\subseteq\Delta^*$，**L**的同态原像是一个集合

$$f^{-1}(L)=\{x|f(x)\in L\}$$

5.2 正则语言的封闭性

例5-10 设$\Sigma = \{0,1\}, \Delta = \{a,b\}$，同态映射$f$定义为：$f(0) = aa$，$f(1) = aba$，则

1. $f(01) = aaaba$
2. $f\left((01)^*\right) = (aaaba)^*$
3. $f^{-1}(aab) = \emptyset$
4. $f^{-1}(aa) = \{0\}$
5. $f^{-1}(\{aaa, aba, abaaaaa, abaaaaaa\}) = \{1, 100\}$
6. $f^{-1}\left((ab + ba)^* a\right) = \{1\}$
7. $f(f^{-1}\left((ab + ba)^* a\right)) = f(\{1\}) = \{aba\}$

令$L = (ab + ba)^* a$，上述7表明$f(f^{-1}(L)) = \{aba\} \neq L$，但是$f(f^{-1}(L)) \subseteq L$

推论 5-1 RL的同态像是 RL。

证明：

同态映射是正则代换的特例，可以直接得到此结论。
该定理表明， RL 在同态映射下是封闭的。

定理 5-5 RL的同态原像是 RL。

■<u>定义5-5</u> 商(Quotient)

设$L_1, L_2 \subseteq \Sigma^*$，$L_1$除以$L_2$的商定义为：

$$L_1/L_2 = \{x \mid \exists y \in L_2, \text{使得} xy \in L_1\}$$

■计算语言的商主要是考虑语言句子的后缀。只有当L_1的句子的后缀在L_2中时，其相应的前缀才属于L_1/L_2。

例5-11 取$L_1=\{000\}, L_2=\{\varepsilon\}, L_3=\{\varepsilon,0\}, L_4=\{\varepsilon,0,00\}, L_5=\{\varepsilon,0,00,000\}$，则

$$
\begin{aligned}
L_1/L_2 &= \{000\} = L_1 \\
L_1/L_3 &= \{000,00\} \\
L_1/L_4 &= \{000,00,0\} \\
L_1/L_5 &= \{000,00,0,\varepsilon\}
\end{aligned}
$$

定理 5-6 设$L_1, L_2 \subseteq \Sigma^*$，如果$L_1$是**RL**，则$L_1/L_2$也是**RL**。

章节目录

■对给定 RL L，DFA M接受L，M不同，由M确定的Σ^*上的等价类也可能不同。

■每个DFA M决定上的一个等价分类：set(q)

■如果M是最小DFA，则M所给出的等价类的个数应该是最少的。

■最小DFA是不是唯一的？如果是，如何构造？

■最小DFA的状态对应的集合与其他DFA的状态对应的集合有什么样的关系？用这种关系是否能从一般的DFA出发，求出最小DFA？

5. 3 Myhill-Nerode定理与DFA的极小化

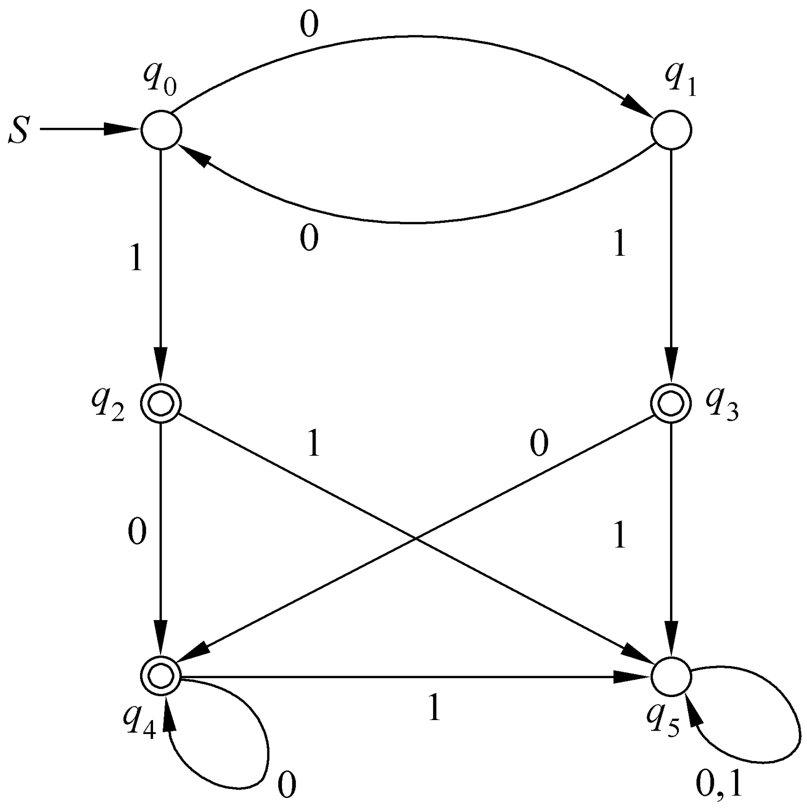
定义5-6 等价关系

DFA M确定的等价关系R_M定义为：对于$\forall x,y \in \Sigma^*$，$xR_My \Leftrightarrow \delta(q_0,x)=\delta(q_0,y)$。

显然，$xR_My \Leftrightarrow \exists q \in Q, x,y \in set(q)$

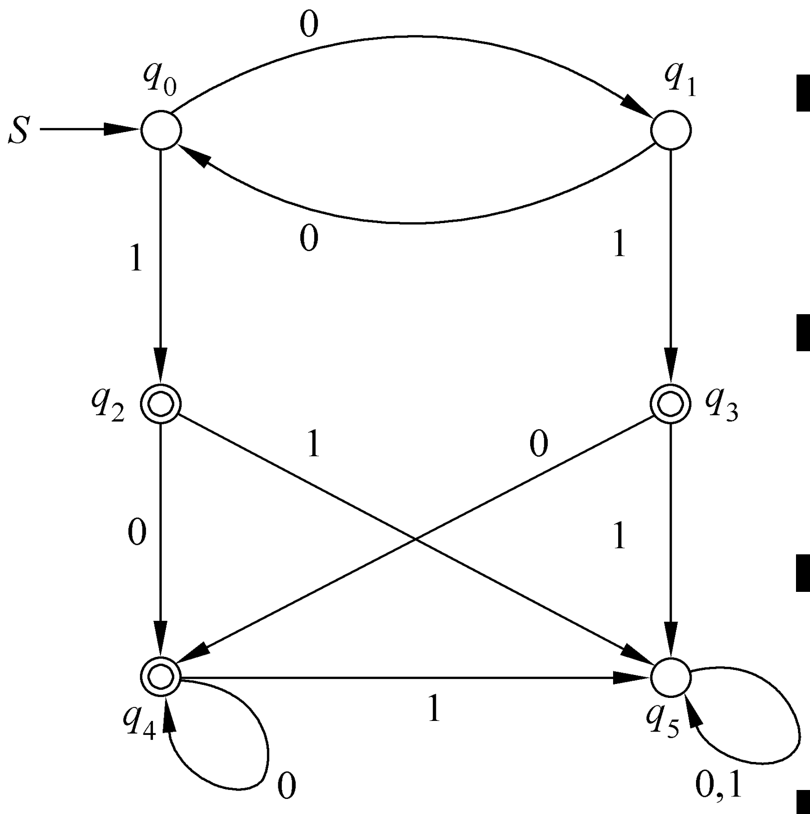
xR_My的直观含义：从初始状态q_0出发，x和y都能把自动机**M**引导到相同的状态q，$q \in Q$。

例5-12 设L = 0^*10^*，它对应的DFA M如下图。



1. 00 R_M 0000 ■ q_0
2. 000 R_M 0 ■ q_1
3. 001 R_M 00001 ■ q_2
4. 0001 R_M 01 ■ q_3
5. 0010 R_M 000010 ■ q_4
6. 00010 R_M 010 ■ q_4
7. 00101 R_M 0000101 ■ q_5
8. 00011 R_M 011 ■ q_5
9. 00110 R_M 00001100 ■ q_5
10. 00001100 R_M 011110 ■ q_5

例5-12 设$L = 0^*10^*$，它对应的DFA M如下图。



- **对应于q_0:**

  $(00)^n R_M (00)^m \quad n, m \geqslant 0;$

- **对应于q_1:**

  $0(00)^n R_M 0(00)^m \quad n,m \geqslant 0;$

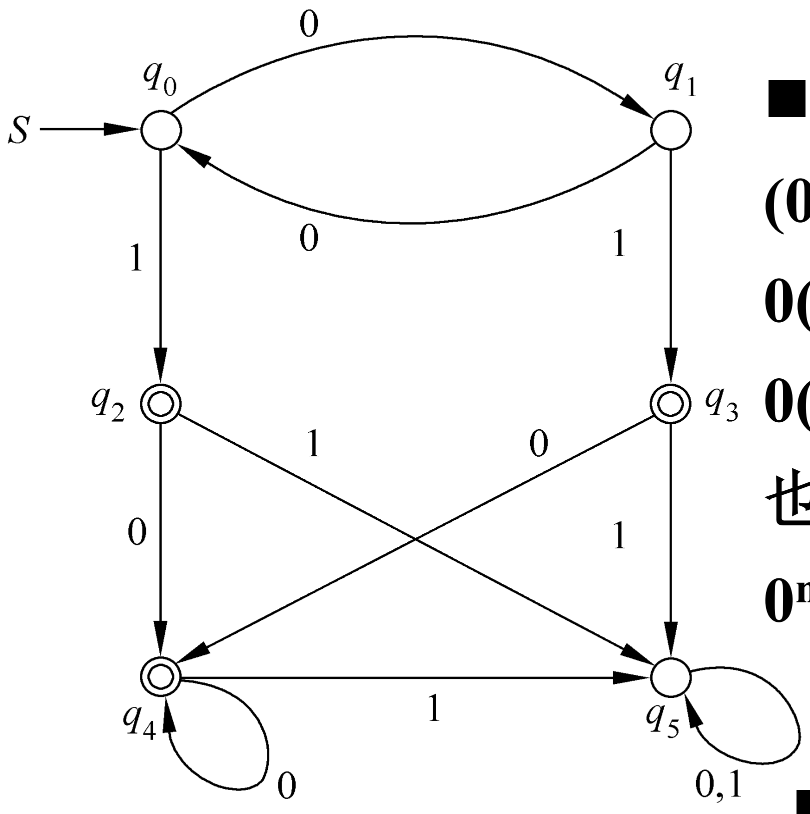
- **对应于q_2:**

  $(00)^n1 R_M (00)^m1 \quad n,m \geqslant 0;$

- **对应于q_3:**

  $0(00)^n 1R_M 0(00)^m1 \quad n,m \geqslant 0;$

例5-12 设$L = 0^*10^*$，它对应的DFA M如下图。



■ 对应于q_4:

$(00)^n10^k R_M (00)^m10^h \quad n, m \geq 0, k, h \geq 1$;

$0(00)^n 10^k R_M 0(00)^m10^h \quad n,m \geq 0, k, h \geq 1$;

$0(00)^n 10^k R_M (00)^m10^h \quad n, m \geq 0, k, h \geq 1$;

也就是:

$0^n 10^k R_M 0^m 10^h \quad n, m \geq 0, \ k, h \geq 1$

■ 对应于q_5:

$x R_M y$—x，y为至少含两个1的串。

定义5-7 语言L确定的Σ^*上的关系R_L：

对于$\forall x, y \in \Sigma^*$，$xR_Ly \Leftrightarrow (\forall z \in \Sigma^*, xz \in L \Leftrightarrow yz \in L)$

即：对于$\forall x, y \in \Sigma^*$，如果$xR_Ly$，则$x$和$y$后无论接$\Sigma^*$中的任何串$z$，$xz$和$yz$要么都是L的句子，要么都不是L的句子。

这里的语言L不一定是正则语言。

5.3.1 Myhill-Nerode定理

■关系R_M：设DFA $\mathbf{M} = (Q, \Sigma, \delta, q_0, F)$，M所确定的$\boldsymbol{\Sigma}^*$上的关系$R_M$定义为：对于$\forall x, y \in \Sigma^*$，$xR_My \Leftrightarrow \delta(q_0, x) = \delta(q_0, y)$。

-- M从q_0开始读入x和y进入同一个状态。

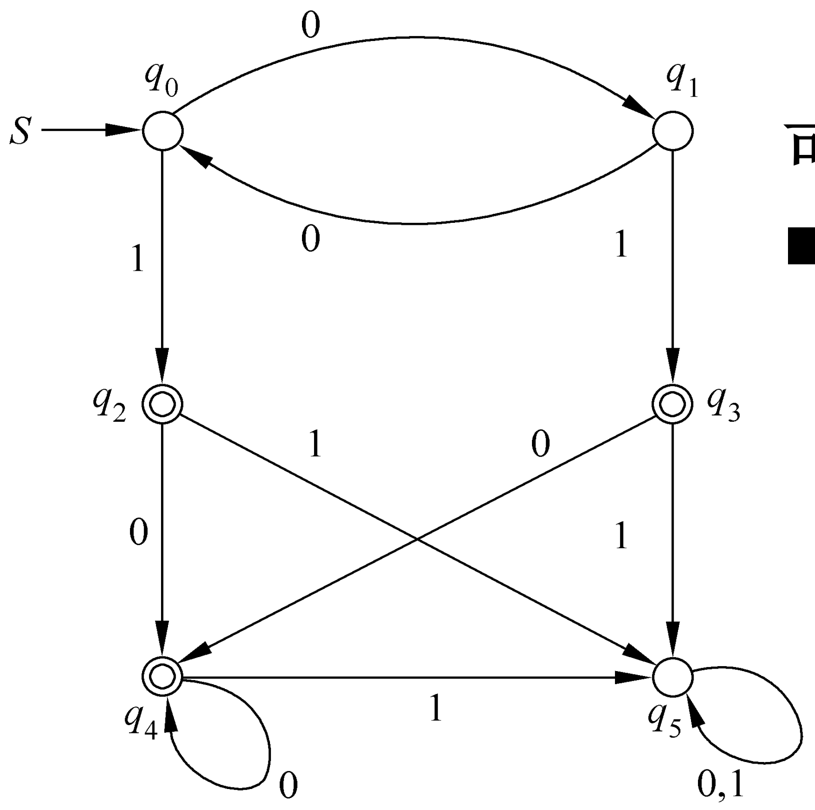
■关系R_L：设$\mathbf{L} \subseteq \boldsymbol{\Sigma}^*$，L确定的$\boldsymbol{\Sigma}^*$上的关系$R_L$定义为：对于$\forall x, y \in \Sigma^*$，$xR_Ly \Leftrightarrow (\forall z \in \Sigma^*, xz \in L \Leftrightarrow yz \in L)$

-- 如果xR_My，则必有$xR_{L(M)}y$；

-- 如果$xR_{L(M)}y$，则不一定有xR_My。

5.3.1 Myhill-Nerode定理

续例5-12 设L = 0^*10^*， $\mathbf{R_M}$和$\mathbf{R_{L(M)}}$的举例。



可以看出：

- **$0R_{L(M)}00$，但是没有$0R_M00$。**

R_M是$R_{L(M)}$的“加细” (refinement)

■**<u>定义5-8</u>** 右不变的等价关系

设**R**是Σ^*上的等价关系，对于$\forall x, y \in \Sigma^*$，如果$x\mathbf{R}y$，则必有$xz\mathbf{R}yz$ 对于$\forall z \in \Sigma^*$成立，则称**R**是右不变的等价关系。

■**命题5-1** 对于任意DFA M $= (Q, \Sigma, \delta, q_0, F)$，M所确定的$\Sigma^*$上的关系$R_M$为右不变的等价关系。

证明：

(1) R_M是等价关系。

自反性显然 。

对称性：$\forall x,y \in \sum^*$,

$x\ R_M\ y \Leftrightarrow \delta(q_0,x)=\delta(q_0,y)$　　根据R_M的定义；

$\Leftrightarrow \delta(q_0,y)=\delta(q_0,z)$　　“=”的对称性；

$\Leftrightarrow y\ R_M\ x$　　根据R_M的定义。

5.3.1 Myhill-Nerode定理

■**<u>命题5-1</u>** 对于任意DFA M $= (Q, \Sigma, \delta, q_0, F)$，M所确定的$\Sigma^*$上的关系$R_M$为右不变的等价关系。

证明： 传递性：设x R_M y,y R_M z。

由于x R_M y,$\delta(q_0,x)=\delta(q_0,y)$

由于y R_M z,$\delta(q_0,y)=\delta(q_0,z)$

由“=”的传递性知,

$\delta(q_0,x)=\delta(q_0,z)$

再由R_M的定义得：

x R_M z

即R_M是等价关系。

■**命题5-1** 对于任意DFA M $= (Q, \Sigma, \delta, q_0, F)$，M所确定的$\Sigma^*$上的关系$R_M$为右不变的等价关系。

证明： (2) R_M 是右不变的

设x R_M y。则$\delta(q_0,x)=\delta(q_0,y)=q$

所以,对于$\forall z\in\sum^*$,

$$
\begin{aligned}
\delta(q_0,xz)&=\delta(\delta(q_0,x),z))\\
&=\delta(q,z)\\
&=\delta(\delta(q_0,y)\ ,z)\\
&=\delta(q_0,yz)
\end{aligned}
$$

这就是说,$\delta(q_0,xz)=\delta(q_0,yz)$,再由$R_M$的定义,$xzR_Myz$

所以,R_M 是右不变的等价关系。

■**命题5-2** 对于任意$L\subseteq \Sigma^*$，L所确定的Σ^*上的关系 R_L为右不变的等价关系 。

证明：

(1)R_L是等价关系。

自反性显然。

对称性：不难看出：$x\ R_L\ y \Leftrightarrow$ (对$\forall z\in\Sigma^*$,$xz\in L \Leftrightarrow yz\in L$) $\Leftrightarrow y\ R_L\ x$

■**命题5-2** 对于任意$L \subseteq \Sigma^*$，L所确定的Σ^*上的关系R_L为右不变的等价关系 。

证明：传递性：设$x\ R_L\ y$, $y\ R_L\ z$。

$$x\ R_L\ y \Leftrightarrow (\text{对}\forall w \in \Sigma^*, xw \in L \Leftrightarrow yw \in L)$$

$$y\ R_L\ z \Leftrightarrow (\text{对}\forall w \in \Sigma^*, yw \in L \Leftrightarrow zw \in L)$$

所以,

$$(\forall w \in \Sigma^*, xw \in L \Leftrightarrow yw \in L \quad \text{且} yw \in L \Leftrightarrow zw \in L)$$

即：

$$(\text{对}\forall w \in \Sigma^*, xw \in L \Leftrightarrow zw \in L)$$

故：

$$x\ R_L\ z$$

即R_L是等价关系。

■**命题5-2** 对于任意L⊆ Σ^*，L所确定的Σ^*上的关系 R_L为右不变的等价关系 。

证明： (2) R_L 是右不变的。

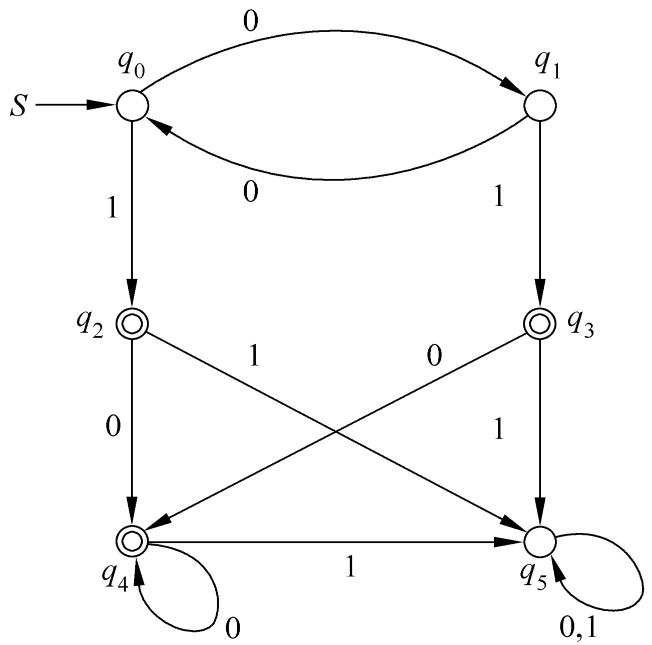
设xR_Ly。由R_L的定义,对$\forall w,v\in\sum^*$,$xwv\in L \Leftrightarrow zwv\in L$,注意到v的任意性,知,$xwR_Lyw$。

所以,R_L是右不变的等价关系。

定义5-9 指数(index)

设R是Σ^*上的等价关系，则称$|\Sigma^*/R|$是R关于Σ^*的指数，简称为R的指数。

下图DFA所确定的右不变的等价关系的指数为6，对应于M的6个状态，将Σ^*分成6个等价类。



- $set(q_0)=\{(00)^n \mid n\geq 0\}$;
- $set(q_1)=\{0(00)^n \mid n\geq 0\}$;
- $set(q_2)=\{(00)^n1 \mid n,\ m\geq 0\}$;
- $set(q_3)=\{0(00)^n 1 \mid n\geq 0\}$;
- $set(q_4)=\{0^n 10^k \mid n\geq 0,\ k\geq 1\}$;
- $set(q_5)=\{x \mid x$为至少含两个1的串$\}$

定理5-7 （Myhill-Nerode定理）

如下3个命题等价：

① $L \subseteq \Sigma^*$是正则语言RL；

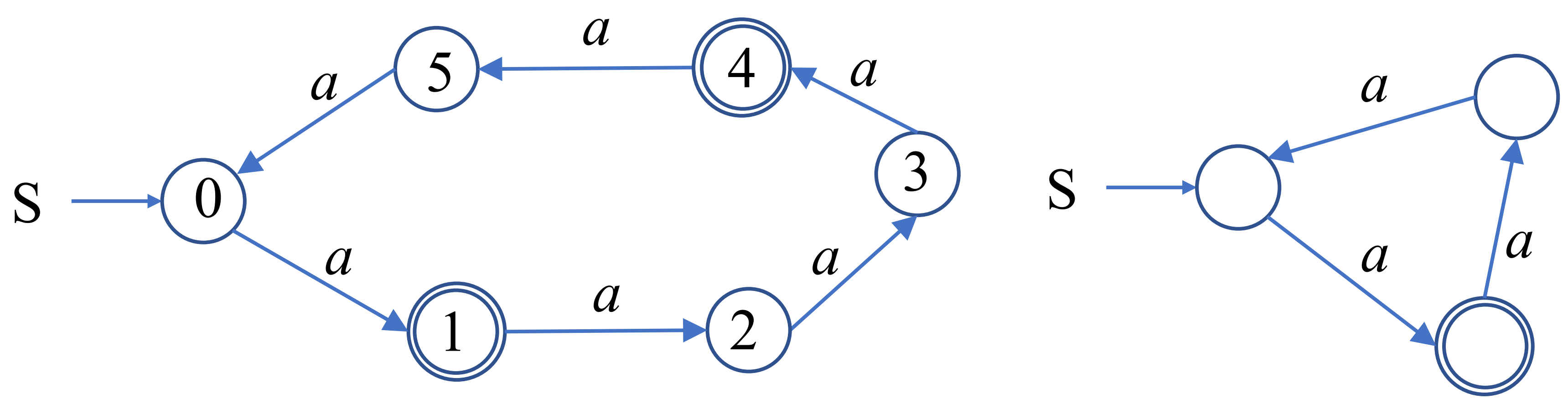
② L是Σ^*上的某一个具有有穷指数的右不变等价关系R的某些等价类的并；

③ R_L具有有穷指数。

Myhill-Nerode定理：判断语言是否正则的充分必要条件。

5. 3 Myhill-Nerode定理与DFA的极小化

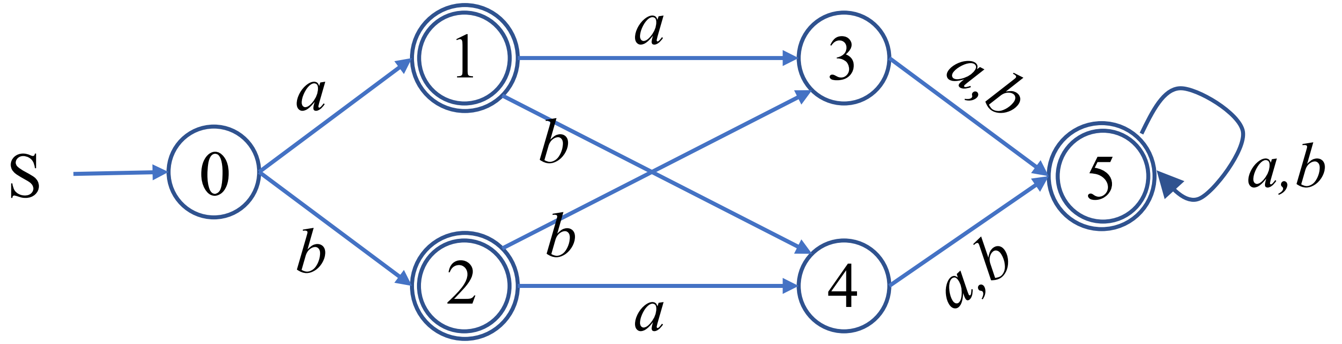
例5-13 接收语言$\{a^m | m \bmod 3 = 1\}$的DFA。



■上图的DFA共有6个状态，是否能合并？

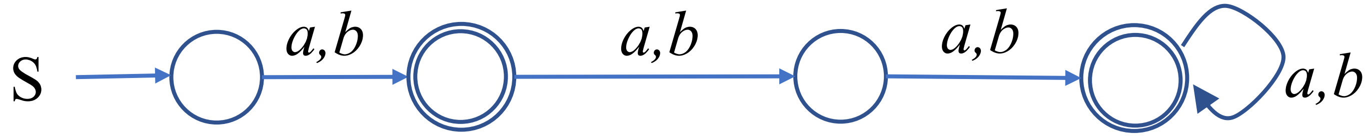
■将状态1和4合并，状态2和5合并，状态0和3合并，所得到的DFA只有3个状态，如右图所示。

例5-14 接收语言$\{x|x \in \{a,b\}^+, 且|x| \neq 2\}$的DFA如下所示。



■状态1和2的功能相同，表示读入串的长度为1时接收。状态3和4的功能相同，表示读入串的长度为2时不接收。

■将状态1和2合并，状态3和4合并，如下图所示。



5.3.2 DFA的极小化

■对于任意的RL L，接受L的最小DFA是唯一的。

■按照L决定的等价关系R_L的等价类来设立状态和状态之间的转移，能够构造出接受L的最小DFA。但是用这种方法构造最小DFA的过程中，需要根据L的特点确定等价类，所以难以自动化。

■选取识别L的任意一个DFA M，可以通过合并R_M的等价类来求出R_L的等价类。而R_M的等价类与M的状态一一对应，所以实际上是要判断哪些状态可以合并。

定义5-10 可以区分的状态对

设DFA $\mathbf{M} = (Q, \Sigma, \delta, q_0, F)$，如果$\exists x \in \Sigma^*$，对Q中的两个状态$p$和$q$，使得$\delta(p, x) \in F$和$\delta(q, x) \in F$中有且仅有一个成立，则称$p$和$q$是可以区分的。

否则，称q和p等价，并记作$p \equiv q$。

5.3.2 DFA的极小化

■判断哪些状态对不可区分（可以合并）：

- 对于M的两个不同状态p和q, 分别从set(p)和set(q)中取一个字符串x和y;
- 研究对于任意的$z \in \Sigma^*$，xz和yz是否同时属于L或者同时不属于L；
- 根据泵引理，z只要取长度不超过M的状态数的字符串就可以了。

虽然需要进行的判断是有限的，但是这种方法的时间复杂度较高。

5.3.2 DFA的极小化

■判断哪些状态对可区分（不可以合并）：

- ➢终止状态和非终止状态不可以合并；
- ➢如果两个状态读入同一个字符所进入的状态不能合并，那么这两个状态不能合并；
- ➢如此循环，当找到所有的不可合并的状态对后，剩余的状态对就是可合并的了。

算法5-1 DFA的极小化算法

- 算法思想：扫描所有的状态对，找出所有的可区分的状态对，不可区分的状态对一定是等价的。

➢对$p \in F$，$q \notin F$所有的状态对(p, q)，在相应的格子内标记×，表示p和q是可以区分的；

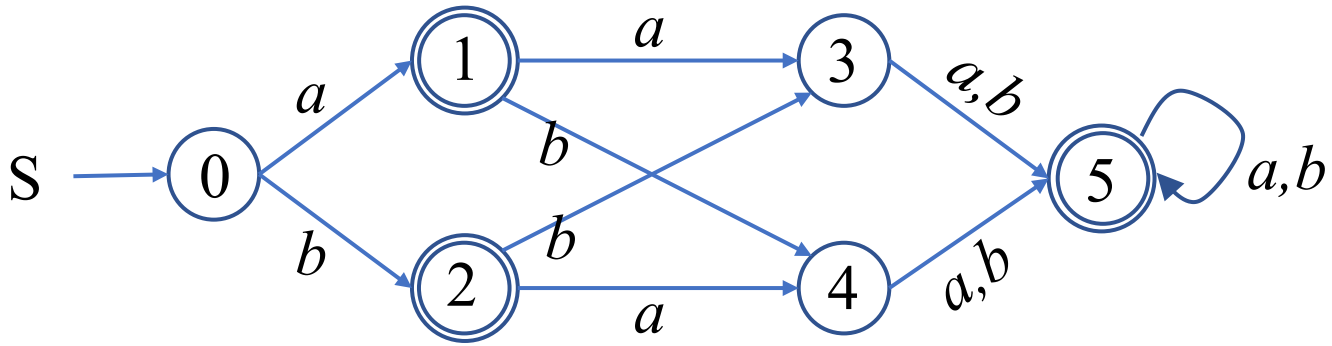
➢重复下面的过程，直到表中内容不再改变：

- 如果存在一个未被标记的状态对(p, q)，且对于某个$a \in \Sigma$，如果$(r = \delta(p, a), s = \delta(q, a))$已做了标记，则在$(p, q)$相应的格子内做标记。

➢在完成上述步骤后，所有未被标记的状态对(p, q)都是等价的，则状态p和状态q可以合并。

5.3.2 DFA的极小化

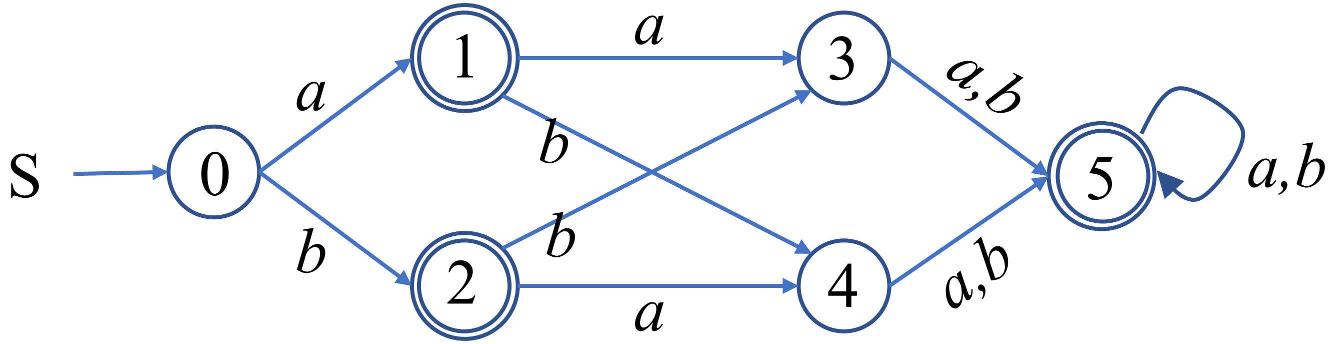
续例5-14 将接收语言$\{x|x \in \{a, b\}^{+}, 且|x| \neq 2\}$的DFA进行极小化。



1. 接收状态与非接收状态可区分
   {0,3,4}×{1,2,5}所得的任意状态对可区分

1	×				
2	×				
3		×	×		
4		×	×		
5	×			×	×
	0	**1**	**2**	**3**	**4**

续例5-14 将接收语言$\{x|x \in \{a,b\}^+, 且|x| \neq 2\}$的DFA进行极小化。



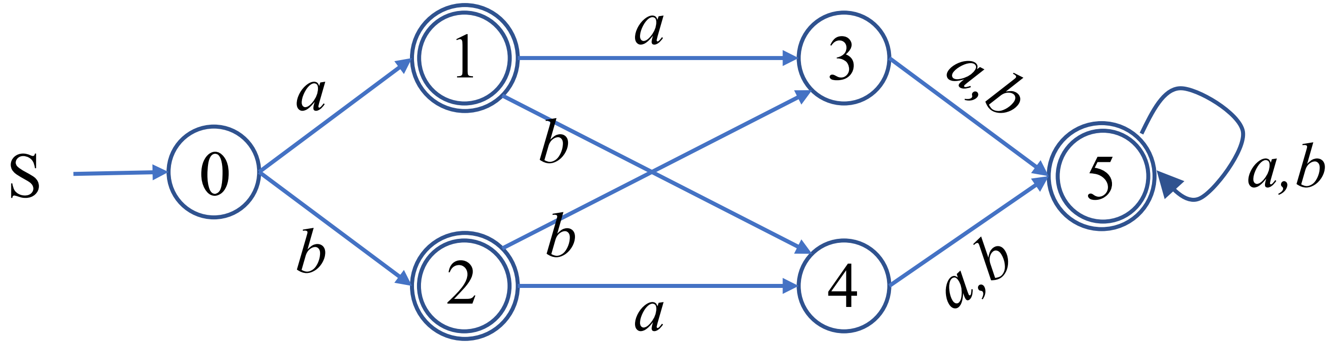
2. 一个未被标记的状态对(p,q)，且对于某个$a \in \Sigma$，如果$(r = \delta(p,a), s = \delta(q,a))$已做了标记，则在$(p,q)$相应的格子内做标记。

1	×				
2	×				
3		×	×		
4		×	×		
5	×			×	×
	0	**1**	**2**	**3**	**4**

- **(0,3)输入a和b后, 变为不可区分的(1,5)和(2,5)**
- **(0,4)输入a和b后, 变为不可区分的(1,5)和(2,5)**
- **(1,2)输入a和b后, 变为不可区分的(3,4)和(4,3)**

5.3.2 DFA的极小化

续例5-14 将接收语言$\{x|x \in \{a,b\}^+, 且|x| \neq 2\}$的DFA进行极小化。



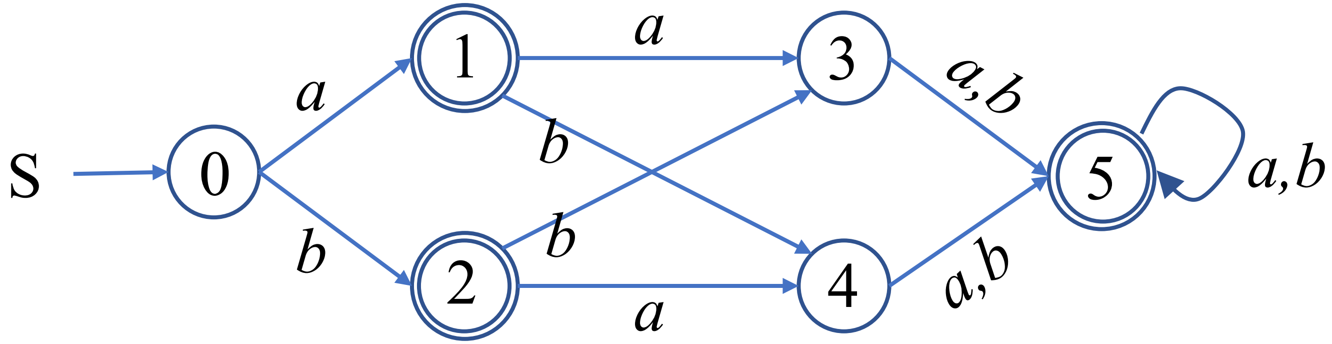
2. 一个未被标记的状态对(p,q)，且对于某个$a \in \Sigma$，如果$(r = \delta(p,a), s = \delta(q,a))$已做了标记，则在$(p,q)$相应的格子内做标记。

1	×				
2	×				
3		×	×		
4		×	×		
5	×	×	×	×	×
	0	**1**	**2**	**3**	**4**

- **(1,5)输入a后, 变为可区分的(3,5)**
- **(2,5)输入a后, 变为可区分的(4,5)**
- **(3,4)输入a和b后, 变为不可区分的(5,5)和(5,5)**

5.3.2 DFA的极小化

续例5-14 将接收语言$\{x|x \in \{a,b\}^+, 且|x| \neq 2\}$的DFA进行极小化。

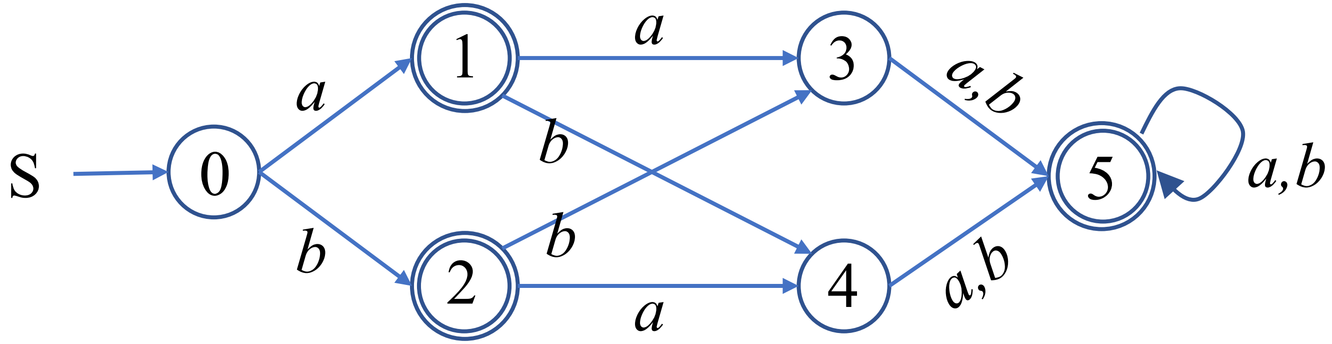


3. 根据新表，再开始下一轮考察。

- **(0,3)输入a后，变为可区分的(1,5)**
- **(0,4)输入a后，变为可区分的(1,5)**
- **(1,2)输入a和b后, 变为不可区分的(3,4)和(4,3)**
- **(3,4)输入a和b后, 变为不可区分的(5,5)和(5,5)**

1	×				
2	×				
3	×	×	×		
4	×	×	×		
5	×	×	×	×	×
	0	**1**	**2**	**3**	**4**

续例5-14 将接收语言$\{x|x \in \{a,b\}^+, 且|x| \neq 2\}$的DFA进行极小化。



4. 根据新表，再开始下一轮考察。

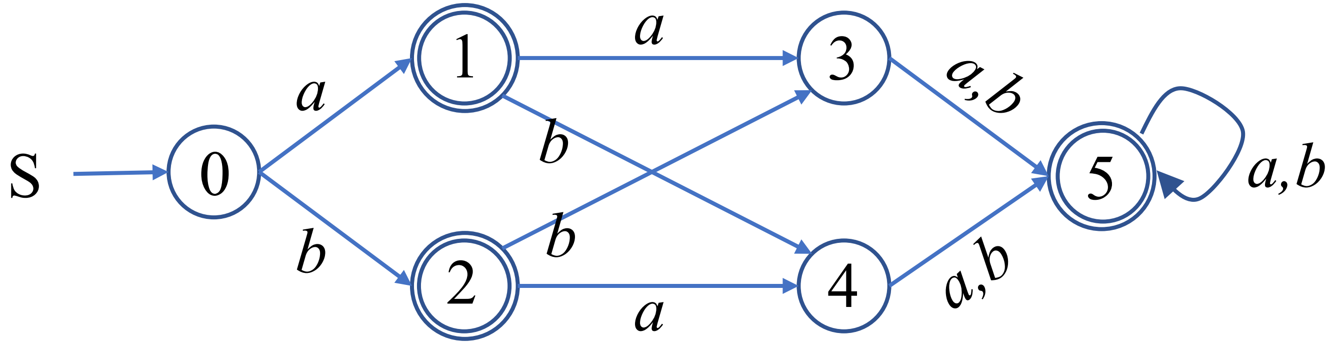
1	×				
2	×				
3	×	×	×		
4	×	×	×		
5	×	×	×	×	×
	0	**1**	**2**	**3**	**4**

- **(1,2)输入a和b后, 变为不可区分的(3,4)和(4,3)**
- **(3,4)输入a和b后, 变为不可区分的(5,5)和(5,5)**

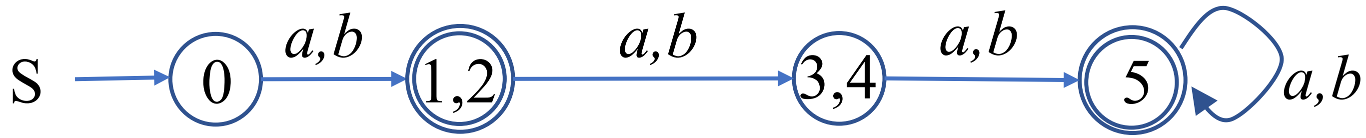
此时，发现表格不再变化，则(1,2)等价，(3,4)等价。

5.3.2 DFA的极小化

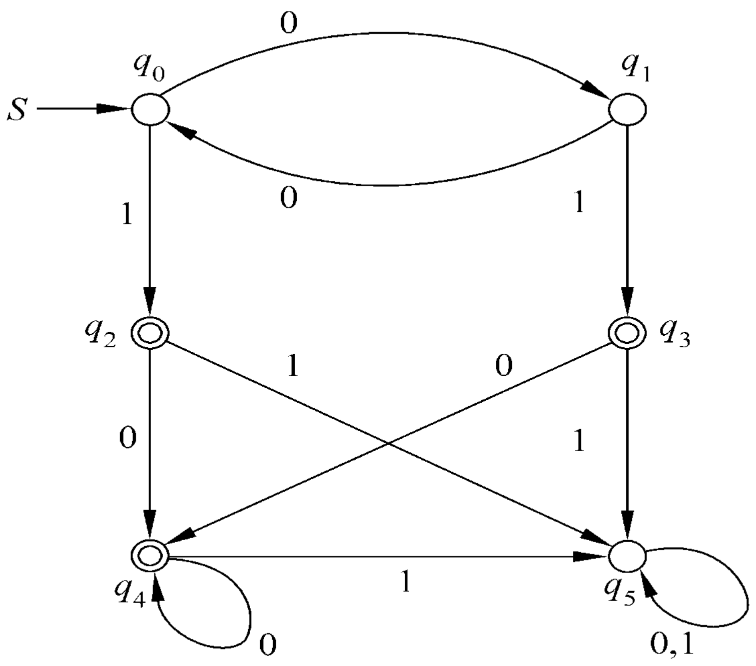
续例5-14 将接收语言$\{x|x \in \{a,b\}^+, 且|x| \neq 2\}$的DFA进行极小化。



5. 根据状态的等价性，画出极小化后的DFA。



1	×				
2	×				
3	×	×	×		
4	×	×	×		
5	×	×	×	×	×
	0	**1**	**2**	**3**	**4**



例5-15 将下面DFA进行极小化。

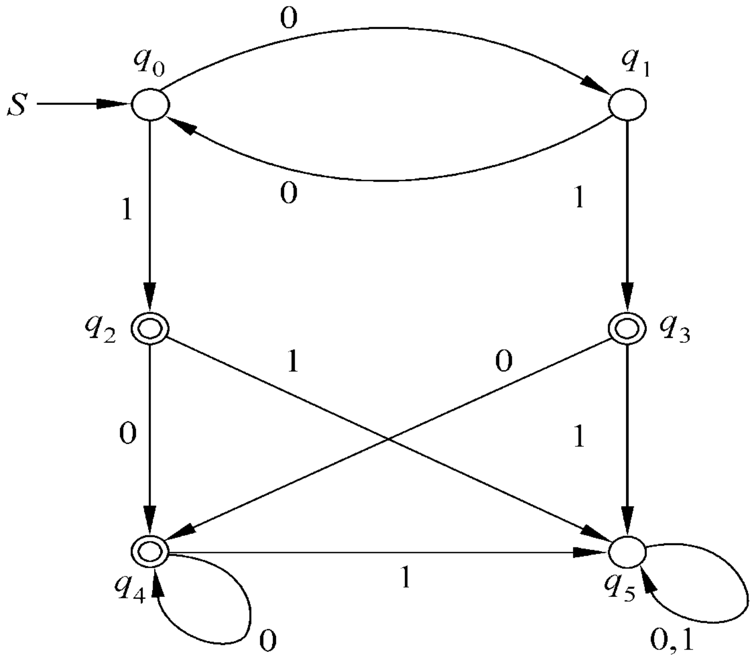
1. 接收状态与非接收状态可区分 $\{q_0,q_1,q_5\}\times\{q_2,q_3,q_4\}$的任意状态对可区分

2. 对表中未标记位置依次进行第一次完善

   - (q_0,q_1)输入0和1后，变为不可区分的(q_1,q_0)和(q_2,q_3)
   - (q_0,q_5)输入1后，变为可区分的(q_2,q_5)
   - (q_1,q_5)输入1后，变为可区分的(q_3,q_5)

q_1					
q_2	×	×			
q_3	×	×			
q_4	×	×			
q_5	×	×	×	×	×
	q_0	q_1	q_2	q_3	q_4

5.3.2 DFA的极小化

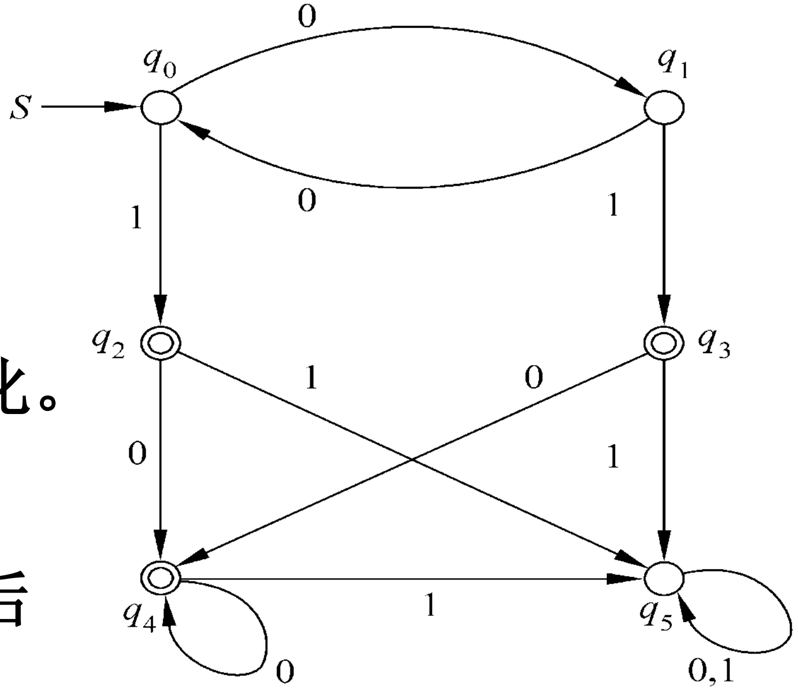


例5-15 将下面DFA进行极小化。

- (q_2, q_3)输入0和1后，变为不可区分的(q_4, q_4)和(q_5, q_5)
- (q_2, q_4)输入0和1后，变为不可区分的(q_4, q_4)和(q_5, q_5)
- (q_3, q_4)输入0和1后，变为不可区分的(q_4, q_4)和(q_5, q_5)

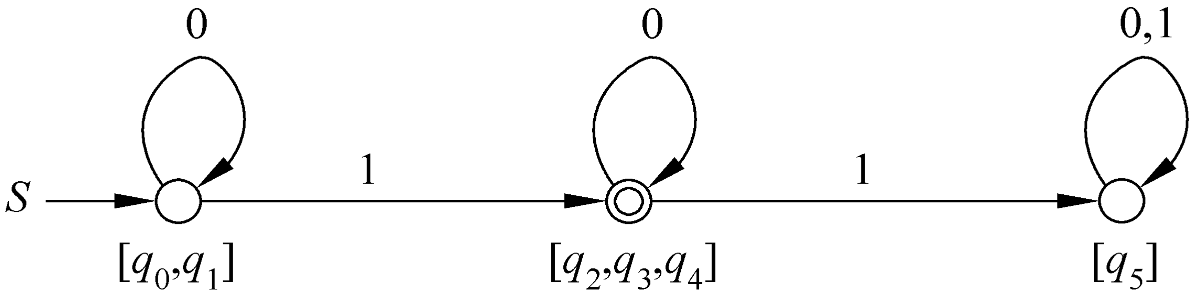
q_1					
q_2	×	×			
q_3	×	×			
q_4	×	×			
q_5	×	×	×	×	×
	q_0	q_1	q_2	q_3	q_4

例5-15 将下面DFA进行极小化。



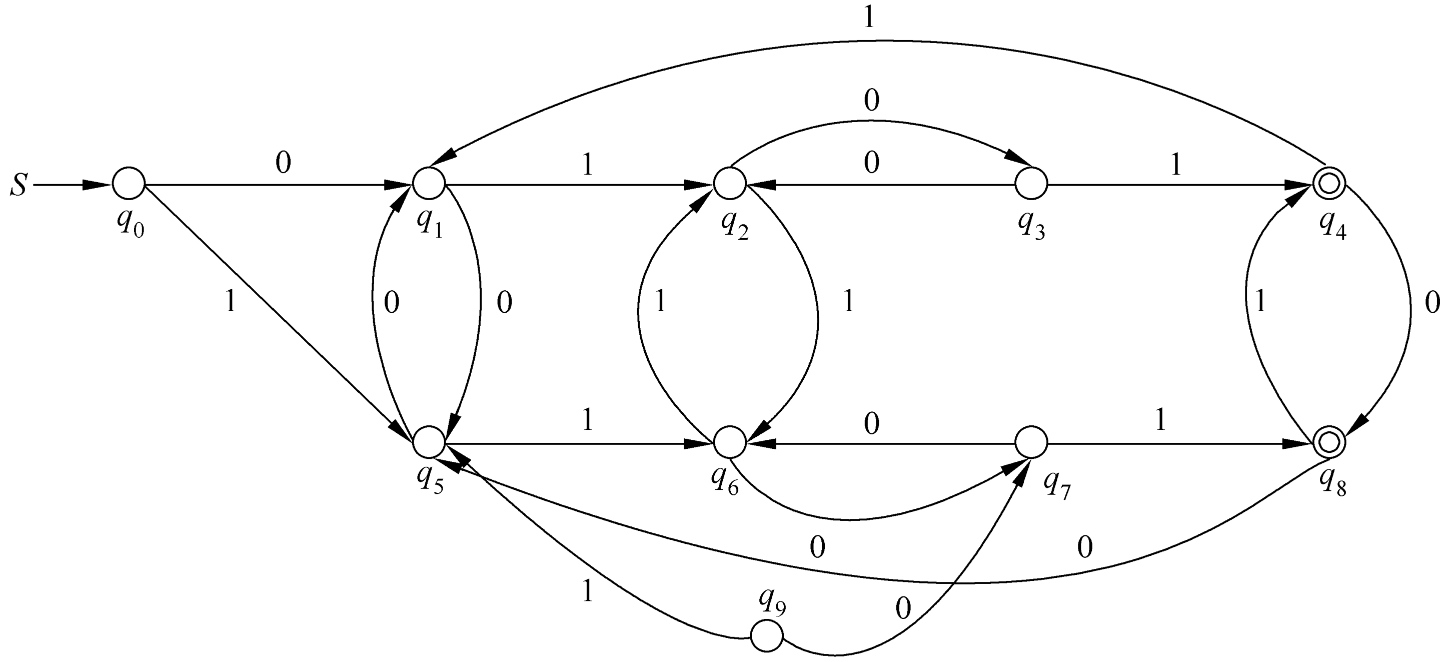
3. **再开始下一轮考察，发现不再变化。因此，$q_0 \equiv q_1$，$q_2 \equiv q_3 \equiv q_4$。**

4. **根据状态的等价性，画出极小化后的DFA。**

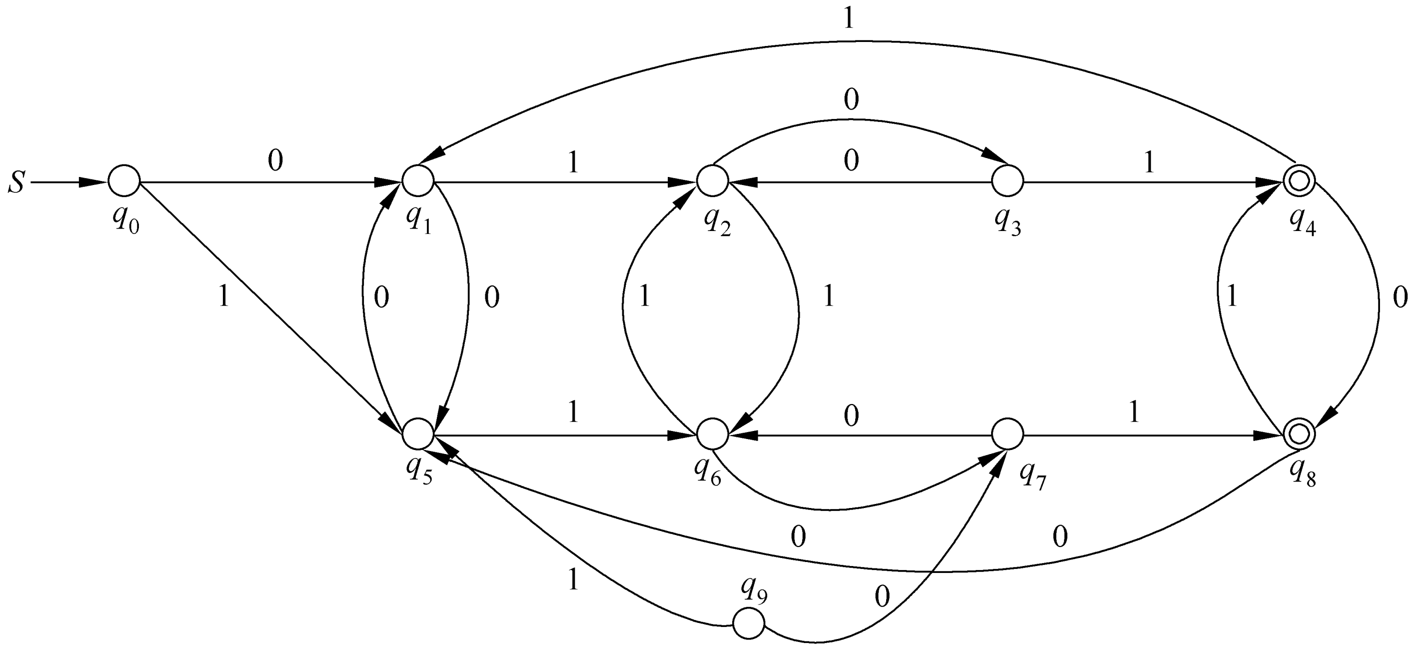


q_1					
q_2	×	×			
q_3	×	×			
q_4	×	×			
q_5	×	×	×	×	×
	q_0	q_1	q_2	q_3	q_4

例5-16 对下图的DFA进行极小化。



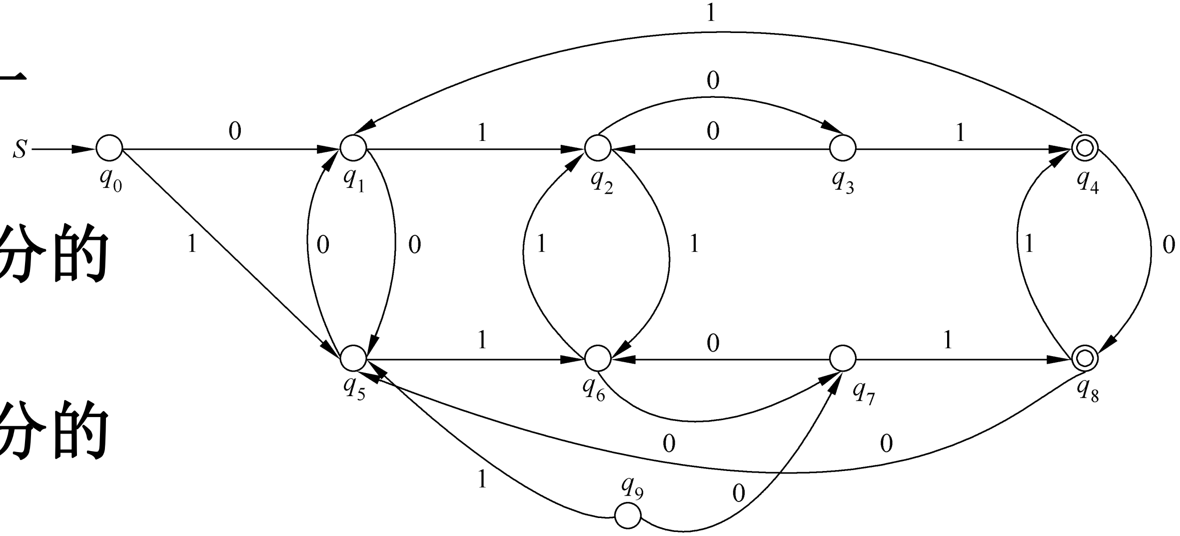
注意：q_9是不可达状态，需先删掉它。



q_1								
q_2								
q_3								
q_4	×	×	×	×				
q_5					×			
q_6					×			
q_7					×			
q_8	×	×	×	×		×	×	×
	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7

1. $\{q_0, q_1, q_2, q_3, q_5, q_6, q_7\} \times \{q_4, q_8\}$ 的任意状态对可区分

2. 对表中未标记位置依次进行第一次完善

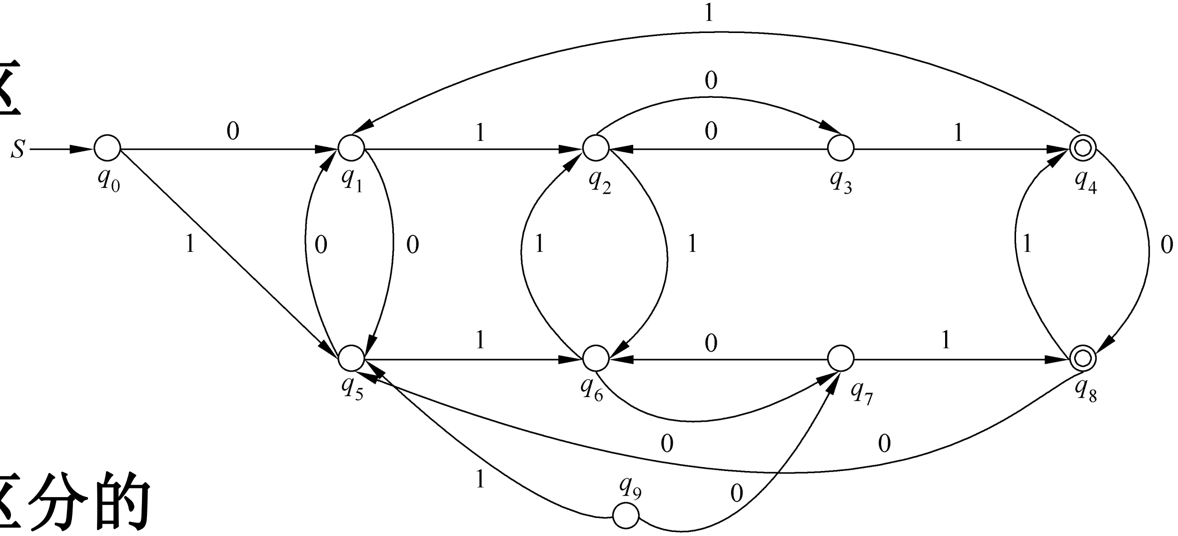


- (q_0, q_1)输入0和1后，变为不可区分的(q_1, q_5)和(q_5, q_2)
- (q_0, q_2)输入0和1后，变为不可区分的(q_1, q_3)和(q_5, q_6)
- (q_0, q_3)输入1后，变为可区分的(q_5, q_4)

q1								
q2								
q3	×							
q4	×	×	×	×				
q5					×			
q6					×			
q7	×				×			
q8	×	×	×	×		×	×	×
	q0	q1	q2	q3	q4	q5	q6	q7

- (q_0, q_5)输入0和1后，变为不可区分的(q_1, q_1)和(q_5, q_6)
- (q_0, q_6)输入0和1后，变为不可区分的(q_1, q_7)和(q_5, q_2)
- (q_0, q_7)输入1后，变为可区分的(q_5, q_8)

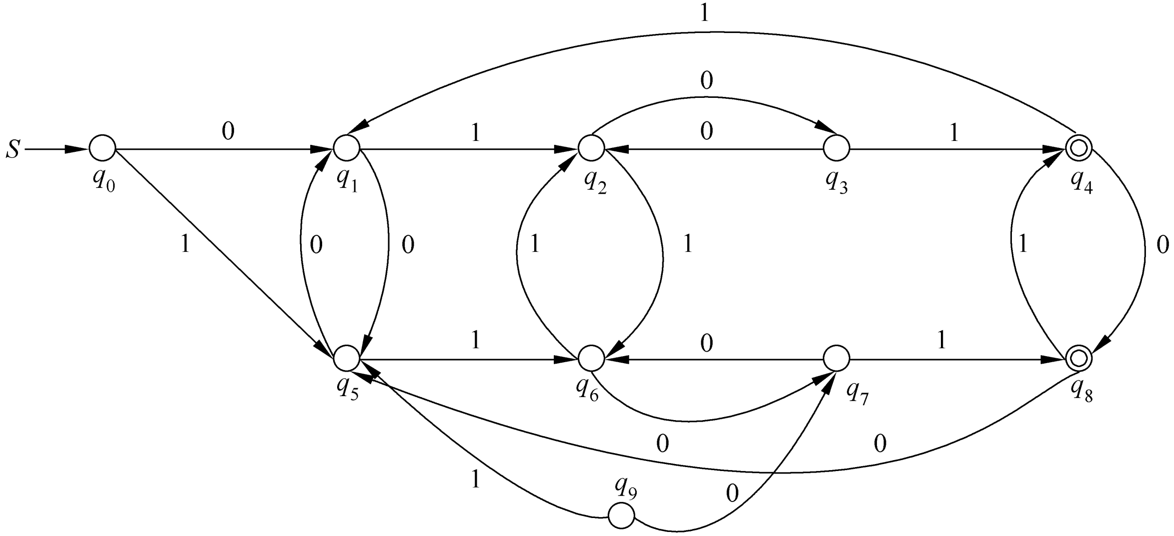
- (q_1,q_2)输入0和1后，变为不可区分的(q_5,q_3)和(q_2,q_6)
- (q_1,q_3)输入1后，变为可区分的(q_2,q_4)
- (q_1,q_5)输入0和1后，变为不可区分的(q_5,q_1)和(q_2,q_6)



	q0	q1	q2	q3	q4	q5	q6	q7
q1								
q2								
q3	×	×	×					
q4	×	×	×	×				
q5			×		×			
q6					×			
q7	×	×			×			
q8	×	×	×	×		×	×	×

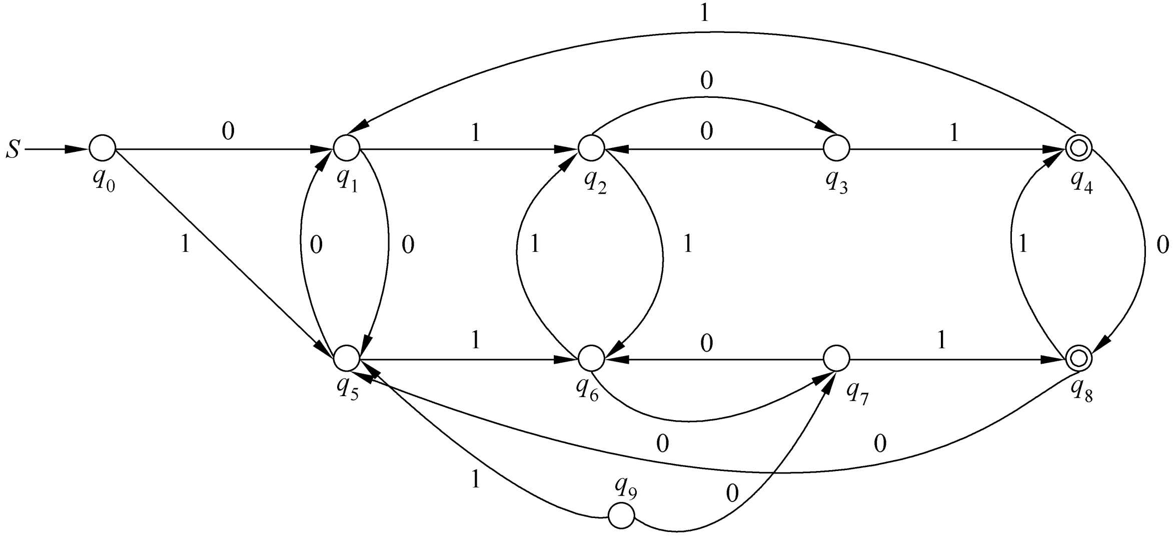
- (q_1,q_6)输入0和1后，变为不可区分的(q_5,q_7)和(q_2,q_2)
- (q_1,q_7)输入1后，变为可区分的(q_2,q_8)
- (q_2,q_3)输入1后，变为可区分的(q_6,q_4)
- (q_2,q_5)输入0后，变为可区分的(q_3,q_1)

- (q_2, q_6)输入0和1后，变为不可区分的(q_3, q_7)和(q_6, q_2)
- (q_2, q_7)输入1后，变为可区分的(q_6, q_8)



q1								
q2								
q3	×	×	×					
q4	×	×	×	×				
q5			×	×	×			
q6				×	×			
q7	×	×	×		×			
q8	×	×	×	×		×	×	×
	q0	q1	q2	q3	q4	q5	q6	q7

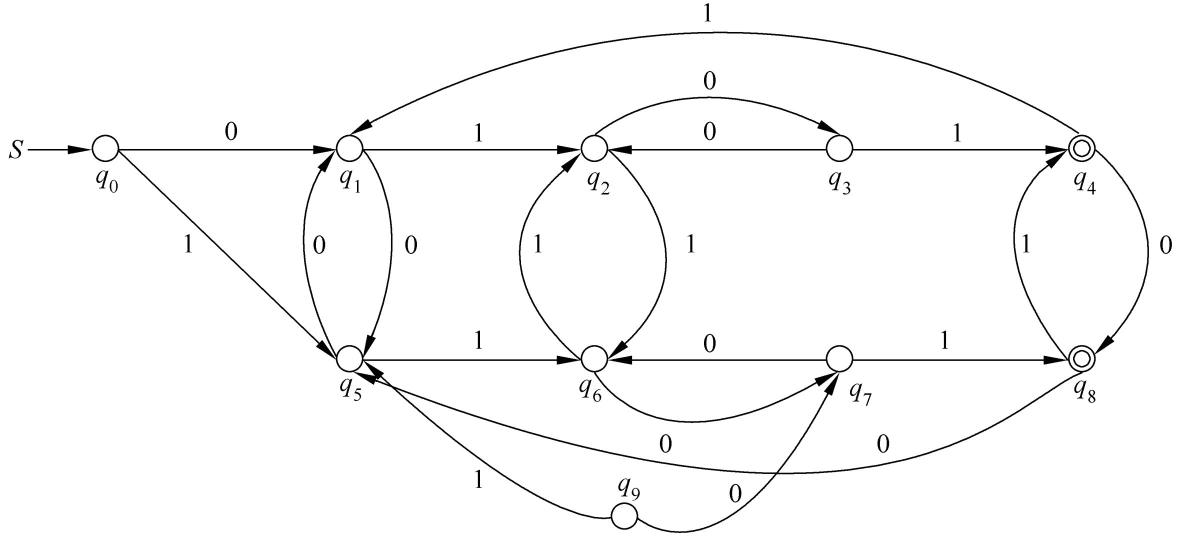
- (q_3, q_5)输入1后，变为可区分的(q_4, q_6)
- (q_3, q_6)输入1后，变为可区分的(q_4, q_2)
- (q_3, q_7)输入0和1后，变为不可区分的(q_2, q_6)和(q_4, q_8)



	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_1								
q_2								
q_3	×	×	×					
q_4	×	×	×	×				
q_5			×	×	×			
q_6				×	×	×		
q_7	×	×	×		×	×	×	
q_8	×	×	×	×	×	×	×	×

- (q_4, q_8)输入0后，变为可区分的(q_8, q_5)
- (q_5, q_6)输入0后，变为可区分的(q_1, q_7)
- (q_5, q_7)输入1后，变为可区分的(q_6, q_8)
- (q_6, q_7)输入1后，变为可区分的(q_2, q_8)

3. 对表中未标记位置依次进行第二次完善

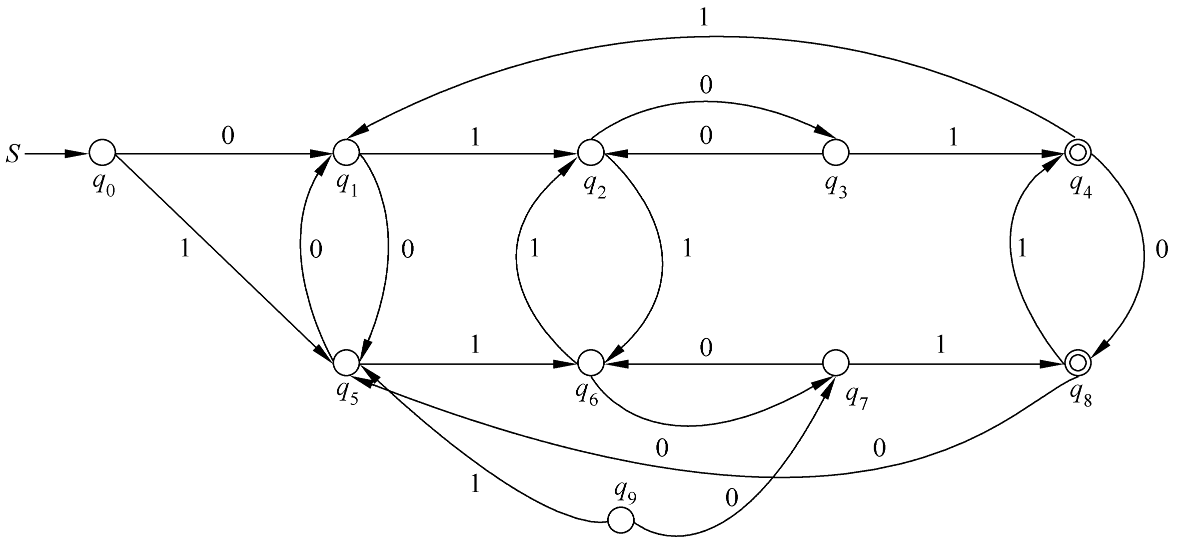


- (q_0,q_1)输入1后，变为可区分的(q_5,q_2)
- (q_0,q_2)输入0后，变为可区分的(q_1,q_3)
- (q_0,q_5)输入1后，变为可区分的(q_5,q_6)
- (q_0,q_6)输入0后，变为可区分的(q_1,q_7)

q1	×							
q2	×	×						
q3	×	×	×					
q4	×	×	×	×				
q5	×		×	×	×			
q6	×	×		×	×	×		
q7	×	×	×	×	×	×	×	
q8	×	×	×	×	×	×	×	×
	q0	q1	q2	q3	q4	q5	q6	q7

- (q_1,q_2)输入0后，变为可区分的(q_5,q_3)
- (q_1,q_5)输入0和1后，变为不可区分的(q_5,q_1)和(q_2,q_6)
- (q_1,q_6)输入0后，变为可区分的(q_5,q_7)
- (q_2,q_6)输入0和1后，变为不可区分的(q_3,q_7)和(q_6,q_2)
- (q_3,q_7)输入1后，变为可区分的(q_4,q_8)

4. 对表中未标记位置依次进行第三次完善



- (q_1, q_5)输入0和1后，变为不可区分的(q_5, q_1)和(q_2, q_6)
- (q_2, q_6)输入0后，变为可区分的(q_3, q_7)

5. 对表中未标记位置依次进行第四次完善

- (q_1, q_5)输入1后，变为可区分的(q_2, q_6)

因此，该DFA各个状态均不可区分，极小化DFA为去除不可达状态q_9后的它自己。

q1	×							
q2	×	×						
q3	×	×	×					
q4	×	×	×	×				
q5	×	×	×	×	×			
q6	×	×	×	×	×	×		
q7	×	×	×	×	×	×	×	
q8	×	×	×	×	×	×	×	×
	q0	q1	q2	q3	q4	q5	q6	q7

章节目录

- “给定DFA接受的语言是空集吗？”
- “给定DFA接受的语言是有穷集（无穷集）吗？”
- “给定两个DFA，他们接受同一个集合吗？（两个DFA是否等价）”
- “给定一个DFA M和一个字符串ω，M能接受ω吗？”

对于以上问题，是否存在一个算法能回答？如果存在，则称该算法为判定算法，相应的问题为可判定的；否则，该问题称为不可判定的。

一个问题能否被一个算法解决，也就是这个问题是否可计算。

定理 5-7 设DFA $M=(Q,\Sigma,\delta,q_0,F)$，L=L(M)非空的充分必要条件是：存在 $x\in\Sigma^*,|x|<|Q|,\delta(q_0,x)\in F$。

存在从 q_0 到某个终止状态且无重复状态的路，且 $|x|<|Q|$。

定理 5-8 设DFA $M=(Q,\Sigma,\delta,q_0,F)$，L=L(M)为无穷的充分必要条件是：存在$x\in\Sigma^*,|Q|\le|x|<2|Q|,$
$\delta(q_0,x)\in F$。

存在“回路”

<u>定理 5-9</u> 设DFA $\mathrm{M}_1=(Q_1,\Sigma,\delta_1,q_{01},F_1)$， DFA $\mathrm{M}_2=(Q_2,\Sigma,\delta_2,q_{02},F_2)$，则存在判定$\mathrm{M}_1$与$\mathrm{M}_2$是否等价的算法。

判定两个DFA的极小DFA是否同构即可判定它们是否等价。

<u>定理 5-10</u> 设L是字母表Σ上的 RL，对任意 $x\in\Sigma^*$，存在判定x是不是L的句子的算法。

一定意义上，接受L的DFA M即判定x是否是L的句子的“算法”。

章节目录

5.5 本章小结

本章讨论了RL的性质。包括：RL的泵引理，RL运算的封闭性，Myhill-Nerode定理与FA的极小化等。

- **泵引理：** 泵引理是用RL的必要条件来用来证明一个语言不是RL的。它不能用来证明一个语言是RL，而且是采用反证法。
- **封闭性：** RL在并、乘、闭包、补、交、正则代换、同态映射运算下是有效封闭的。RL 的同态原像是RL。设L_1、$L_2 \subseteq \Sigma^*$，如果L_1是RL，则L_1/L_2也是 RL 。
- **DFA的极小化：** 如果L是RL，则根据R_L确定的Σ^*的等价类可以构造出接受L的最小DFA。更方便的方法是通过确定给定DFA状态的可区分性构造出等价的最小DFA。
- **判定算法：** 存在判定L(M)是非空、M_1与 M_2是否等价、L(M)是否无穷、x是不是RL L的句子的算法。

Thanks!