形式语言与自动机期末速通

6. RL的性质

6.1 RL的封闭性质

[**定义6.1.1**]

(1)若对任意的、属于同一语言类的语言在某一特定运算下所得的结果仍属于该类语言,则称该类语言对该运算**封闭**.

(2)有效封闭性:给定一个语言类的若干个语言的描述,若存在一个算法,它可构造出这些语言在给定运算下所得的运算结果的相应形式的语言描述,则称该语言类对该运算**有效封闭**.

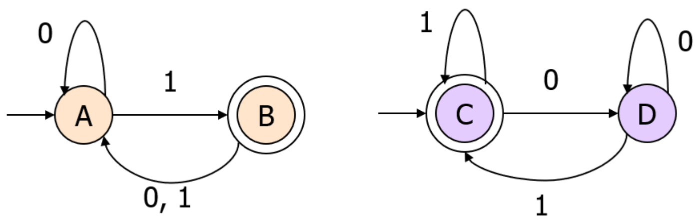
[**定理6.1.1**] RL对并、乘积、闭包运算封闭.

[**定理6.1.2**] 设L和M都是RL,则$L\backslash M$表示在L中但不在M中的串.

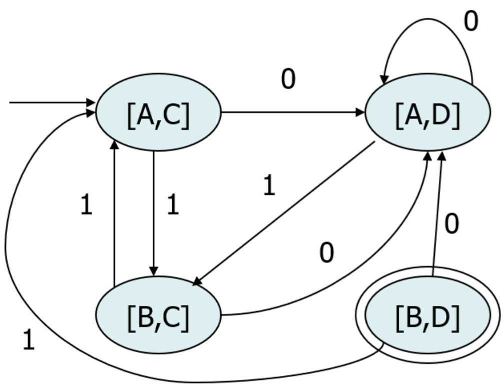
[**证**] 设A和B分别为L和M对应的DFA,下面构造A和B的乘积DFA C.

选取C的接受状态为A接受但B不接受的状态即可.

[**例6.1.1**] 设RL L和M对应的DFA分别如下图所示,求$L\backslash M$对应的DFA.



[**解**] 如下图所示.事实上,$L\backslash M$为空语言.



[**定理6.1.3**] RL对补运算封闭.

[**证**] 设RL L对应的FA为$M=(Q,\Sigma,\delta,q_0,F)$.取DFA $M'=(Q,\Sigma,\delta,q_0,Q\backslash F)$,

注意到对$\forall\in\Sigma^*$,有$\delta(q_0,x)=f\in F\Leftrightarrow\delta(q_0,x)=f\notin Q\backslash F$,即$x\in L(M)\Leftrightarrow x\notin L(M')$.

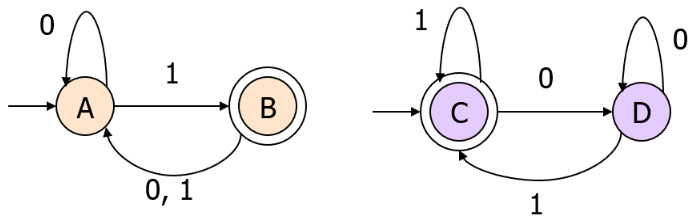
故$L(M')=\sum^*-L(M)$,而Σ^*和$L(M)$都是RL,故证.

[**定理6.1.4**] RL对交运算封闭.

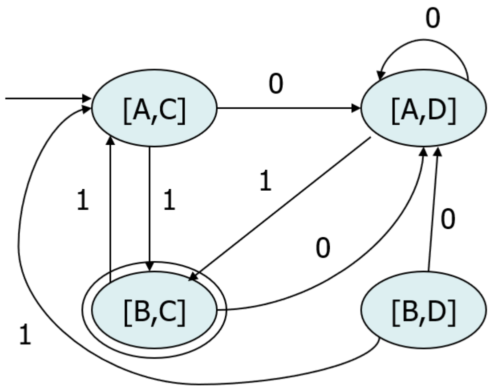
[**证**] 设A和B分别为L和M对应的DFA,下面构造A和B的乘积DFA C.

选取C的接受状态为A接受且B接受的状态即可.

[**例6.1.2**] 设RL L和M对应的DFA分别如下图所示,求$L\bigcap M$对应的DFA.



[**解**]



[**定义6.1.2**] 对字母表Σ和Δ,称映射$f:\Sigma\rightarrow 2^{\Delta^*}$为$\Sigma$到$\Delta$上的**代换**.若对$\forall a\in\Sigma$,有$f(a)$是$\Delta$上的RL,则称$f$为从$\Sigma$到$\Delta$的**正则代换**.

(1)将f的定义域扩展为Σ^*.对$\forall x\in\Sigma^*,\forall a\in\Sigma$,$f:\Sigma^*\rightarrow 2^{\Delta^*}$定义为:

①$f(\varepsilon)=\{\varepsilon\}$.

②$f(xa)=f(x)f(a)$.

(2)将f的定义域扩展为2^{Σ^*}.对$\forall L\subset\Sigma^*$,$f:2^{\Sigma^*}\rightarrow 2^{\Delta^*}$定义为:$f(L)=\bigcup\limits_{x\in L}f(x)$.

[**例6.1.3**] 设字母表$\Sigma=\{0,1\}$,字母表$\Delta=\{a,b\}$,代换$f(a)$满足:①$f(0)=a$;②$f(1)=b^*$.则:

(1)$f(010)=f(0)f(1)f(0)=ab^*a$.

(2)$f(\{11,00\})=f(11)\bigcup f(00)=f(1)f(1)\bigcup f(0)f(0)=b^*b^*+aa=b^*+aa$.

(3)$f(L(0^*(0+1)1^*))=L(a^*(a+b^*)(b^*)^*)=L(a^*(a+b^*)(b^*))$

$=L(a^*ab^*+a^*b^*b^*)=L(a^*b^*)$.

[**定理6.1.5**] 设f是从字母表Σ到字母表Δ的正则代换,则:

(1)$f(\varPhi)=\varPhi$.

(2)$f(\varepsilon)=\varepsilon$.

(3)对$\forall a\in\Sigma$,有$f(a)$是Δ上的RE.

(4)设r和s是Σ上的RE,则下列都是Δ上的RE:

①$f(r+s)=f(r)+f(s)$.

②$f(rs)=f(r)f(s)$.

③$f(r^*)=f(r)^*$

[**定理6.1.6**] 设字母表Σ和Δ.设L是Σ上的RL,$f:\Sigma\rightarrow 2^{\Delta^*}$是正则变换,则$f(L)$是RL.

[**证**] 设L对应的RE为r.下面对r中的运算符的个数n归纳:

(1)$n=0$时,结论成立.

(2)假设$n\leq k$时结论成立,即r中的运算符个数不超过k时,有$f(L(r))=L(f(r))$.

(3)$n=k+1$时,

①设$r=r_1+r_2$,其中r_1和r_2中的运算符个数不超过k.

$f(L)=f(L(r))$

$=f(L(r_1+r_2))$

$=f(L(r_1)\bigcup L(r_2))$ *RE的定义

$=f(L(r_1))\bigcup f(L(r_2))$ *正则代换的定义

$=L(f(r_1))\bigcup L(f(r_2))$ *归纳假设

$=L(f(r_1)+f(r_2))$ *RE的定义

$=L(f(r_1+r_2))$ *正则代换的定义

$=L(f(r))$.

②设$r=r_1r_2$,其中r_1和r_2中的运算符个数不超过k.

$f(L)=f(L(r))$

$=f(L(r_1r_2))$

$=f(L(r_1)L(r_2))$ *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

$=L(f(r_1r_2))$ *正则代换的定义

$=L(f(r))$.

③设$r=r_1^*$,其中r_1中的运算符个数不超过k.

$f(L)=f(L(r))$

$=f(L(r_1^*))$

$=f(L(r_1)^*)$ *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

$= L(f(r_1^*))$ *正则代换的定义

$= L(f(r))$.

[例6.1.4] 设字母表$\Sigma = \{0,1,2\}$,字母表$\Delta = \{a,b\}$,正则代换f满足:①$f(0) = ab$;②$f(1) = b^*a^*$;③$f(2) = a^*(a+b)$.则:

(1)$f(00) = abab$.

(2)$f(010) = abb^*a^*ab = ab^+a^+b$.

(3)$f((0+1+2)^*) = (ab + b^*a^* + a^*(a+b))^*$.

(4)$f(0(0+1+2)^*) = ab(ab + b^*a^* + a^*(a+b))^*$.

(5)$f(012) = abb^*a^*a^*(a+b) = ab^+a^*(a+b)$.

(6)$f((0+1)^*) = (ab + b^*a^*)^* = (ab + b + a + b^*a^*)^* = (a+b)^*$.

[定义6.1.3]

(1)设字母表Σ和Δ,映射$f : \Sigma \rightarrow \Delta^*$.若对$\forall x, y \in \Sigma^*$,有$f(xy) = f(x)f(y)$,则称$f$为从$\Sigma$到$\Delta^*$的**同态映射**.

(2)对$\forall L \subseteq \Sigma^*$,$L$的同态像定义为:$f(L) = \bigcup\limits_{x \in L} f(x)$.

(3)对$\forall w \subseteq \Delta^*$,$w$的同态原像定义为:$f^{-1}(w) = \{x \mid f(x) = w \wedge x \in \Sigma^*\}$.

(4)对$\forall L \subset \Delta^*$,$L$的同态原像定义为:$f^{-1}(L) = \{x \mid f(x) \in L\}$.

[注] 同态映射是正则代换的特例.

[例6.1.5] 设字母表$\Sigma = \{0,1\}$和字母表$\Delta = \{a,b\}$,同态映射f满足:①$f(0) = aa$;②$f(1) = aba$.则:

(1)$f(01) = aaaba$.

(2)$f((01)^*) = (aaaba)^*$.

(3)$f^{-1}(aab) = \varPhi$.

(4)$f^{-1}(aa) = \{0\}$.

(5)$f^{-1}(\{aaa, aba, abaaaaa, abaaaaaa\}) = \{1, 100\}$.

(6)$f^{-1}((ab+ba)^*a) = \{1\}$.

(7)$f(f^{-1}((ab+ba)^*a)) = f(\{1\}) = \{aba\}$.

设语言$L = (ab+ba)^*a$,则上式表明:$f(f^{-1}(L)) \neq L$.事实上,$f(f^{-1}(L)) \subset L$.

[定理6.1.7] RL的同态像是RL.

[证] 注意到同态映射是正则代换的特例即证.

[**推论**] RL对同态映射封闭.

[**定理6.1.8**] RL的同态原像为RL.

[**析**] 设RL对应的FA为M.下面构造接受该RL的同态原像的FA M',使得M'用一个移动模拟M处理$f(a)$所用的一系列移动.

具体地,对Σ中的任意字符a,若M从状态q开始处理$f(a)$,且当它处理完$f(a)$时到达状态p,则让M'在状态q读入字符a后转移到状态p.

M'与M状态相同,且在M'的状态转移图中,从状态q到状态p有一条标记为a的弧当且仅当在M的状态转移图中,有一条从状态q到状态p的标记为$f(a)$的路径.

[**证**] 设f是从字母表Σ到字母表Δ的同态映射.

(1)设DFA $M=(Q,\Delta,\delta,q_0,F)$识别的语言为$L$.

取DFA $M'=(Q,\Sigma,\delta',q_0,F)$,其中$\delta'(q,a)=\delta(q,f(a))$.

(2)下面证明$L(M')=f^{-1}(L(M))$.

对x的串长归纳,下证:对$\forall x\in\Sigma^*$,有$\delta'(q_0,x)=\delta(q_0,f(x))$.

①$|x|=0$时,结论成立.

②假设$|x|=k$时结论成立.

③下证$|x|=k=1$时结论成立.不妨设$x=ya$,其中$|y|=k$.

$$
\begin{aligned}
\delta'(q_0,x)=\delta'(q_0,ya)\\
&=\delta'(\delta'(q_0,y),a)\\
&=\delta'(\delta(q_0,f(y)),a)\ \text{*归纳假设}\\
&=\delta(\delta(q_0,f(y)),f(a))\ \text{*}\delta'\text{的定义}\\
&=\delta(q_0,f(y)f(a))\ \text{*扩展的}\delta\\
&=\delta(q_0,f(ya))\ \text{*同态映射的性质}\\
&=\delta(q_0,f(x)).
\end{aligned}
$$

综上,对$\forall x\in\Sigma^*$,有$\delta'(q_0,x)=\delta(q_0,f(x))$.

进而$\delta'(q_0,x)\in F\Leftrightarrow\delta(q_0,f(x))\in F$,故$L(M')=f^{-1}(L(M))$.

[**定义6.1.4**] 对语言$L_1,L_2\subseteq\Sigma^*$,定义L_1除以L_2的**商**为:$L_1/L_2=\{x\mid\exists y\in L_2\ s.t.\ xy\in L_1\}$.

[**注**] 语言的商主要考察语言句子的后缀.

只有L_1的句子的后缀都在L_2中时,其对应的前缀才属于L_1/L_2.故$\varepsilon\in L_2$时,有$L_1\subseteq L_1/L_2$.

[**例6.1.6**] 设语言$L_1=\{000\},L_2=\{\varepsilon\},L_3=\{\varepsilon,0\},L_4=\{\varepsilon,0,00\},L_5=\{\varepsilon,0,00,000\}$.则:

(1)$L_1/L_2=\{000\}=L_1$.

(2)$L_1/L_3=\{000,00\}$.

(3)$L_1/L_4=\{000,00,0\}$.

(4)$L_1/L_5=\{000,00,0,\varepsilon\}$.

[定理6.1.9] 设语言$L_1,L_2\subseteq\Sigma^*$.若L_1是RL,则L_1/L_2也是RL.

[证] 设RL $L_1\subseteq\Sigma^*$对应DFA为$M=(Q,\Sigma,\delta,q_0,F)$.

取DFA $M'=(Q,\Sigma,\delta,q_0,F')$,其中$F'=\{q\mid\exists y\in L_2,\delta(q,y)\in F\}$.

显然$(M')=L_1/L_2$,故证.

6.2 Myhill-Nerode定理

[定理6.2.1] DFA $M=(Q,\Sigma,\delta,q_0,F)$确定了等价关系:对$\forall x,y\in\Sigma^*$,有$x\ R_M\ y\Leftrightarrow\delta(q_0,x)=\delta(q_0,y)$.

[证] (1)自反性显然.

(2)对称性:对$\forall x,y\in\Sigma^*$,

$x\ R_M\ y$

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

$\Leftrightarrow\delta(q_0,y)=\delta(q_0,x)$

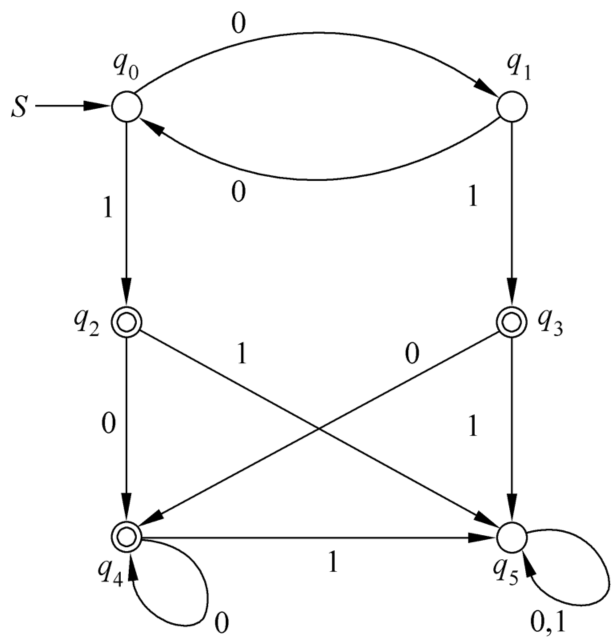
$\Leftrightarrow y\ R_M\ x.$ $*R_M$的定义

(3)传递性:设$x,y,z\in\Sigma^*$,且$x\ R_M\ y,y\ R_M\ z$,则$\delta(q_0,x)=\delta(q_0,y),\delta(q_0,y)=\delta(q_0,z)$,

进而$\delta(q_0,x)=\delta(q_0,z)$,即$x\ R_M\ z$.

[推论] $xR_My\Leftrightarrow q\in Q\ s.t.\ x,y\in set(q)$.

[例6.2.1] RL $L=0^*10^*$对应的DFA M如下图所示:



(1)对q_0,有$(00)^n\ R_M\ (00)^m\ \ (n,m\geq 0)$.

(2)对q_1,有$0(00)^n\ R_M\ 0(00)^m\ \ (n,m\geq 0)$.

(3)对q_2,有$(00)^n1\ R_M\ (00)^m1\ \ (n,m\geq 0)$.

(4)对q_3,有$0(00)^n1\ R_M\ 0(00)^m1\ \ (n,m\geq 0)$.

(5)对q_4,有:

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

②$(00)^n10^k\ R_M(00)^m10^h\ \ (n,m\geq 0;k,h\geq 1)$.

③$0(00)^n10^k\ R_M(00)^m10^h\ \ (n,m\geq 0;k,h\geq 1)$.

可统一表示为:$0^n10^k\ R_M\ 0^m10^h\ \ (n,m\geq 0;k,h\geq 1)$.

(6)对q_5,有$x\ R_M\ y$表示x和y都是至少含有两个1的串.

[**定理6.2.2**] 字母表Σ上的语言L确定了Σ^*上的关系R_L.对$\forall x,y\in\Sigma^*$,$x\ R_L\ y\Leftrightarrow$(对$\forall z\in\Sigma^*,xz\in L\Leftrightarrow yz\in L$),即对$\forall x,y\in\Sigma^*$,若$x\ R_L\ y$,则在$x$后和$y$后接$\forall z\in\Sigma^*$,串$xz$和串$yz$要么都是$L$的句子,要么都不是$L$的句子.

[**证**] 对$\forall x,y\in set(q)$,有$\delta(q_0,x)=\delta(q_0,y)=q$.

对$\forall z\in\Sigma^*$,$\delta(q_0,xz)=\delta(q(q_0,x),z)=\delta(q,z)=\delta(\delta(q_0,y),z)=\delta(q_0,yz)$.

这表明:$\delta(q_0,xz)\in F\Leftrightarrow\delta(q_0,yz)\in F$,即对$\forall z\in\Sigma^*$,有$xy\in L\Leftrightarrow yz\in L$.

[**定义6.2.1**] 设字母表Σ,R是Σ^*上的等价关系.对$\forall x,y\in\Sigma^*$,若$x\ R_L\ y$,则对$\forall z\in\Sigma^*$,有$xz\ R_L\ yz$,则称R为**右不变的**等价关系.

[**定理6.2.3**] 对DFA $M=(Q,\Sigma,\delta,q_0,F)$,$M$确定的$\Sigma^*$上的关系$R_M$是右不变的等价关系.

[**证**] 设$x,y\in\Sigma^*$,且$x\ R_M\ y$.设$\delta(q_0,x)=\delta(q_0,y)=q$.

对$\forall z\in\Sigma^*$,$\delta(q_0,xz)=\delta(q(q_0,x),z)=\delta(q,z)=\delta(\delta(q_0,y),z)=\delta(q_0,yz)$,即$xz\ R_M\ yz$.

[**定理6.2.4**] 设字母表Σ,则$\forall L\subseteq\Sigma^*$确定的$\Sigma^*$上的关系$R_L$是右不变的等价关系.

[**证**]

(1)下证R_L是等价关系.

①自反性显然.

②对称性:显然$x\ R_L\ y\Leftrightarrow$(对$\forall z\in\Sigma^*$,有$xz\in L\Leftrightarrow yz\in L$)$\Leftrightarrow y\ R_L\ x$.

③传递性:设$x,y,z\in\Sigma^*$,且$x\ R_L\ y,y\ R_L\ z$.

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

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

则对$\forall w\in\Sigma^*$,有$xw\in L\Leftrightarrow yw\in L$且$yw\in L\Leftrightarrow zw\in L$,即$xw\in L\Leftrightarrow zw\in L$,故$x\ R_L\ z$.

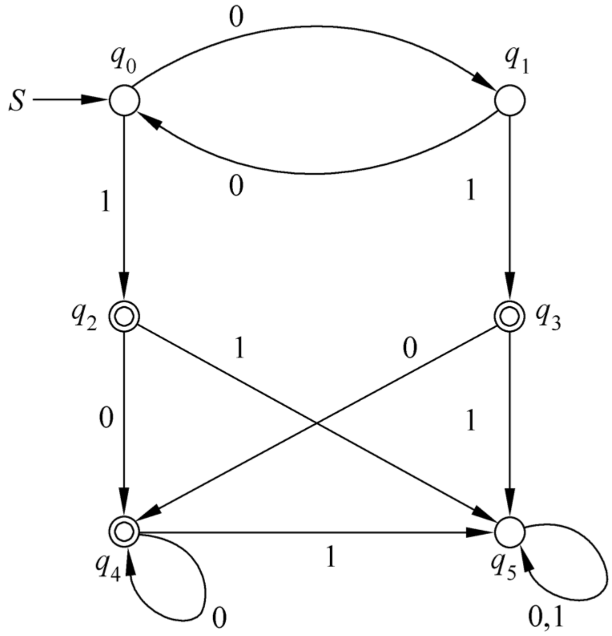
(2)下证R_L是右不变的.

设$x,y\in\Sigma^*$且$x\ R_L\ y$,则对$\forall w,v\in\Sigma^*$,有$xwv\in L\Leftrightarrow ywv\in L$.

由v的任意性:$xw\ R_L\ yw$,故证.

[**定义6.2.2**] 设字母表Σ,R是Σ^*上的等价关系,则称$|\Sigma^*/R|$是R关于Σ^*的**指数**,简称R的指数.Σ^*/R中的每个元素是R的一个等价类.

[**例6.2.2**] 下图所示的DFA M的所确定的等价关系R_M的指数为6.



R_M将Σ^*分为6个等价类:

(1)$set(q_0)=\{(00)^n\mid n\geq 0\}$.

(2)$set(q_1)=\{0(00)^n\mid n\geq 0\}$.

(3)$set(q_2)=\{(00)^n1\mid n\geq 0\}$.

(4)$set(q_3)=\{0(00)^n1\mid n\geq 0\}$.

(5)$set(q_4)=\{0^n10^k\mid n\geq 0,k\geq 1\}$.

(6)$set(q_5)=\{x\mid x\text{至少包含两个}1\}$.

[**定义6.2.3**] 设字母表Σ.对$\forall x,y\in\Sigma^*$,若$x\ R_M\ y$,有$x\ R_{L(M)}\ y$,即对$\forall$DFA $M=(Q,\Sigma,\delta,q_0,F)$,有$|\Sigma^*/R_{L(M)}|\leq|\Sigma^*/R_M|\leq|Q|$.按等价关系$R_M$分类,被分在同一等价类的串,按等价关系$R_{L(M)}$分类也会分在同一等价类.$R_M$对$\Sigma^*$的划分比$R_{L(M)}$对$\Sigma^*$的划分"更细",称$R_M$是$R_{L(M)}$的**加细**.

[**定理6.2.5**] [**Myhill-Nerode定理**] 下列三个命题等价:

(1)$L\subseteq\Sigma^*$是RL.

(2)L是Σ^*上某一有有穷指数的右不变的等价关系R的某些等价类的并.

(3)R_L有有穷指数.

[**证**] (1)⇒(2) 因L是RL,则$\exists$DFA $M=(Q,\Sigma,\delta,q_0,F)\ s.t.\ L(M)=L$.

因R_M是Σ^*上的右不变的等价关系,且$|\Sigma^*/R_M|\leq|Q|$,则R_M有有穷指数.

因$L=\displaystyle\bigcup_{q\in F}set(q)$,则$L$是$\Sigma^*$上某一有有穷指数的右不变的等价关系$R_M$的、对应于$M$的终止状态的等价类的并.

(2)⇒(3) 设$x,y\in\Sigma^*$且$x\ R\ y$.因R是右不变的,则$\forall z\in\Sigma^*$,有$xz\ R\ yz$.

因L是Σ^*上某一有有穷指数的右不变的等价关系R的某些等价类的并,

则$xz \in L \Leftrightarrow yz \in L$,即$x\ R_L\ y$.

故R是R_L的加细,而R有有穷指数,故R_L有有穷指数.

(3)⇒(1) 取DFA $M' = (\Sigma^*/R_L, \Sigma, \delta', [\varepsilon], \{[x] \mid x \in L\})$,

其中$[\varepsilon]$表示ε所在的等价类对应的状态,$[x]$表示x所在的等价类对应的状态.

注意到对$\forall([x], a) \in (\Sigma^*/R_L) \times \Sigma$,有$\delta'([x], a) = [xa]$,故$L(M') = L$.

[**例6.2.3**] 求证:语言$L = \{0^n1^n \mid n \geq 0\}$不是RL.

[**析**] 按L中的句子的特征求R_L的等价类.注意到L中的句子所含0的个数与所含1的个数相等,且所有0都在1之前.

[**证**] R_L有如下等价类:

①$[\varepsilon]$为ε所在的等价类.

②$[1]$为串0所在的等价类.

③$[2]$为串00所在的等价类.

$\vdots$

④$[n]$为串0^n所在的等价类.

故R_L的指数是无穷的,进而不是RL.

6.3 DFA的极小化

[**定理6.3.1**] 对RL L,若DFA $M = (Q, \Sigma, \delta, q_0, F)\ s.t.\ L(M) = L$,则$|\Sigma^*/R_L| \leq |Q|$.

[**推论**] 对$\forall$DFA $M = (Q, \Sigma, \delta, q_0, F)$,有$|Q| \geq |\Sigma^*/R_{L(M)}|$.

[**定义6.3.1**] 若两个DFA的状态间一一对应,且状态的一一对应也保持了状态转移的一一对应,则称这两个DFA**同构**.具体的,若状态q与状态$[w]$对应,状态p与状态$[z]$对应,则$\delta(q, a) = p$时,有$\delta([w], a) = [z]$.

设映射f满足:$f(q) = f(\delta(q_0, x)) = \delta'([\varepsilon], x) = [x] \Leftrightarrow \delta(q_0, x) = q$.

(1)下证f是Q与Σ^*/R_L间的一一对应.

[**证**] ①若$\delta(q_0, x) = \delta(q_0, y)$,则$x\ R_M\ y$,而$R_M$是$R_L$的加细,则$x\ R_L\ y$,

进而$[x] = [y]$,即$\delta'([\varepsilon], x) = \delta'([\varepsilon], y)$.

②若$\delta(q_0, x) \neq \delta(q_0, y)$,则$\delta'([\varepsilon], x) \neq \delta'([\varepsilon], y)$,即$[x] \neq [y]$.

若不然,则$|\Sigma^*/R_M| > |\Sigma^*/R_L|$,与$R_M$是$R_L$的加细矛盾.

(2)下证:若$\delta(q, a) = p$,$f(q) = [x]$,则$f(p) = [xa]$.这表明:若DFA M在状态q读入字符a时进入状态p,则M'在q对应的状态$f(\delta(q_0, x)) = [x]$读入字符a时进入p对应的状态$f(\delta(q_0, xa)) = [xa]$,即$f$是$M$与$M'$间的同构映射.

[**证**] 对$\forall q \in Q$,若$f(q) = f(\delta(q_0, x)) = [x]$,

则对$\forall a \in \Sigma$,若$p = \delta(q, a) = \delta(\delta(q_0, x), a) = \delta(q_0, xa)$,

则$f(p) = f(\delta(q, a)) = f(\delta(\delta(q_0, x), a)) = f(\delta(q_0, xa)) = [xa]$.

[**定理6.3.2**] 对RL L,在同构意义下接受L的最小DFA唯一.

[**证**] 设接受L的最小DFA为$M=(Q,\Sigma,\delta,q_0,F)$,则$M$的状态数等于$R_L$的指数.这表明:

M的状态数与Myhill-Nerode定理的证明中构造的DFA $M'=(\Sigma^*/R_L,\Sigma,\delta',[\varepsilon],\{[x]\mid x\in L\})$的状态数相等.

[**定义6.3.2**] 对DFA $M=(Q,\Sigma,\delta,q_0,F)$,若对$Q$中的两个状态$q$和$p$,$\exists x\in\Sigma^*\ s.t\ \delta(q,x)\in F$和$\delta(p,x)\in F$中有且只有一个成立,则称$q$与$p$是**可区分的**,否则称$q$与$p$**等价**,记作$q\equiv p$.

[**定理6.3.3**] [**DFA的极小化算法**]

(1)思想:遍历所有状态对,求所有可区分的状态对,则不可区分的状态对都等价.

(2)输入:给定的DFA.

(3)输出:可区分状态表.

(4)主要数据结构:状态对的关联链表、可区分状态表.

(5)步骤:

①遍历所有$(q,p)\in F\times(Q\backslash F)$,标记可区分状态表中的表项$(q,p)$.

②遍历所有$(q,p)\in F\times F\bigcup(Q-F)\times(Q-F)$,若$q\neq p$,则做如下操作:

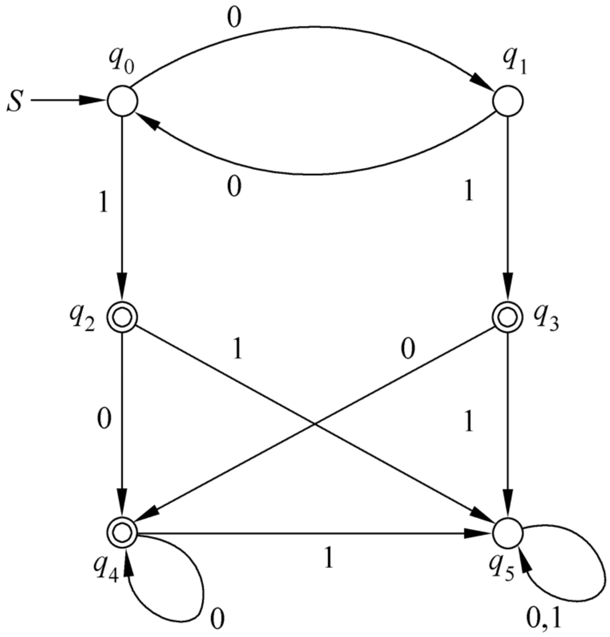
(i)if $\exists a\in\Sigma\ s.t.$ 可区分状态表中的表项$(\delta(q,a),\delta(p,a))$已被标记,则标记可区分状态表中的表项$(q,p)$,并递归地标记本次被标记的状态对的关联链表上的各个状态对在可区分状态表中的对应表项.

(ii)else 遍历所有$a\in\Sigma$,若$\delta(q,a)\neq\delta(p,a)$且$(q,p)$与$(\delta(q,a),\delta(p,a))$非同一状态对,则将$(q,p)$插入$(\delta(q,a),\delta(p,a))$的关联链表.

[**定理6.3.4**] 对DFA $M=(Q,\Sigma,\delta,q_0,F)$,$Q$中的两个状态$q$与$p$是可区分的的充要条件是:$(q,p)$在DFA的极小化算法中被标记.

[**定理6.3.5**] 由DFA极小化算法构造出的DFA去掉不可达状态后是最小DFA.

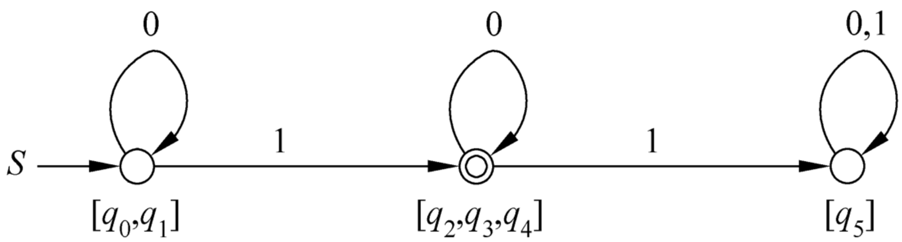
[**例6.3.1**] 极小化下图所示的DFA.



[解] 可区分状态表:

q_1					
q_2	$\times$				
q_3	$\times$	$\times$			
q_4	$\times$	$\times$			
q_5	$\times$	$\times$	$\times$	$\times$	$\times$
	q_0	q_1	q_2	q_3	q_4

最小DFA:



6.4 RL的判定性质

6.4.1 成员判定

[成员判定] 判断一个串w是否是RL L的句子时,若L可用DFA M描述,则模拟输入为w时M的状态转移即可.

[定理6.4.1] 设L是字母表Σ上的RL,则对$\forall x \in \Sigma^*$,存在判定x是否为L的句子的算法.

[注] 一定意义上,接受L的DFA即判断x是否为L的句子的算法.

6.4.2 空否判定

[**空否判定**] 判断一个RL L是否为空时,先作接受该语言的DFA的状态转移图,再求从开始状态q_0出发能到达的状态的集合.若至少存在一个可达的接收状态,则L非空,否则L为空.

[**定理6.4.2**] 设DFA $M=(Q,\Sigma,\delta,q_0,F)$,则语言$L=L(M)$非空的充要条件是:$\exists x\in\Sigma^*\ s.t.\ |x|<|Q|$,且$\delta(q_0,x)\in F$.

[**证**] (充) 充分性显然.

(必) 若L非空,则M的状态转移图中存在一条从状态q_0到某一终止状态q_f且无重复状态的路径,

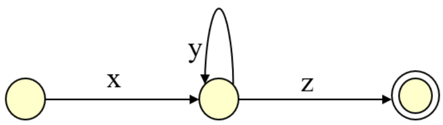
该路径的状态数$n\leq|Q|$,则该路径的标记x满足$|x|\leq n-1$.

而$\delta(q_0,x)\in F$,则x是L的长度$<|Q|$的句子.

6.4.3 无穷判定

[**定理6.4.3**] 若L包含长度$\geq n$的字符串,则也包含长度在$[n,2n-1]$范围内的串.

[**证**] 设接受L的DFA如下图,$w=xyz$,其中y为路径的第一个环.



因$|x|<n,1\leq|y|\leq n,|z|<n$,则$|xz|<2n$.

①若$|xz|\geq n$,则xz即为所求.

②若$|xz|<n$,中间添加若干个y直至$|xz|\geq n$.

[**无穷判定**] 判断一个RL L是否为空时:

(1)若接受L的DFA包含n个状态,但L包含长度$\geq n$的串,则L无穷,否则L有穷.

但有无穷多个长度$\geq n$的串,无法枚举验证.

(2)由**定理6.4.3**:检查所有长度在$[n,2n-1]$范围内的句子是否属于L,若存在属于L的句子,则L无穷,否则L有穷.

(3)作接受L的DFA的状态转移图,检查从开始状态到接收状态的路径上是否有环,若有则L无穷,否则L有穷.

6.4.4 泵引理

[**定理6.4.4**] [**泵引理**] 设字母表Σ,L是Σ^*RL,则$\exists$仅依赖于L的正整数$N\ s.t.$ 对$\forall z\in L$,若$|z|\geq N$,则$\exists u,v,w\in\Sigma^*$满足下列条件:

①$z=uvw$.

②$|uv|\leq N$.

③$|v|\geq 1$.

④对$\forall$整数$i>0$,有$uv^iw\in L$.

⑤N不大于接受L的最小DFA M的状态数.

[**注1**] 只能用泵引理证明一个语言不是RL,不能用来证明一个语言是RL.

[**注2**] 用泵引理证明一个语言不是RL时一般无需用到条件⑤.

[**注3**] 条件②不是必须的,但该限制可简化证明.

[**例6.4.1**] 求证:语言$L=\{0^n1^n\mid n\geq 1\}$不是RL.

[**证**] 若L是RL,取$z=0^N1^N\in L,v=0^k\ \ (k\geq 1)$,则$u=0^{N-k-j},w=0^j1^N$.

由泵引理:对$\forall$整数$i>0$,有$uv^iw=0^{N-k-j}(0^k)^i0^j1^N=0^{N+(i-1)k}1^N$.

因$i=2$时,$uv^2w=0^{N+k}1^N$,而$k\geq 1$,则$N+k\neq N$,进而$uv^2w\notin L$,矛盾.

[**例6.4.2**] 求证:语言$L=\{0^n\mid n\in primes\}$不是RL.

[**证**] 若L是RL,取$z=0^{N+p}\in L,v=0^k\ \ (k\geq 1)$,则$uv^iw=0^{N+p-k-j}(0^k)^i0^j=0^{N+p+(i-1)k}$.

因$i=N+p+1$时,$N+p+(i-1)k=(N+p)(k+1)$,

而$k\geq 1$,则$(N+p)(k+1)\notin primes$,进而$uv^{N+p+1}w\notin L$,矛盾.

6.4.5 等价性判定

[**等价性判定**] 判断RL L与M是否等价时,构造接受L的DFA和接受M的DFA的乘积DFA,用乘积DFA的状态同时模拟两DFA的移动.具体地,设接受L、M的DFA的状态集分别为Q、R,则乘积DFA的状态集为$Q\times R$,即对$\forall q\in Q,\forall r\in R$,$[q,r]$是乘积DFA的一个状态.

①开始状态:$[q_0,r_0]$.

②转移函数:$\delta([q,r],a)=[\delta_L(q,a),\delta_M(r,a)]$.

[**定理6.4.5**] 设DFA $M_1=(Q_1,\Sigma,\delta_1,q_{01},F_1)$,DFA $M_2=(Q_2,\Sigma,\delta_2,q_{02},F_2)$,则存在判定$M_1$与$M_2$是否等价的算法.
