形式语言与自动机　作业 4

正则语言的性质 2

1951112　林日中

习题 1　利用 Myhill-Nerode 定理证明下列语言是否正则语言，如果是正则语言，请构造其 FA、RE 及 RG.

a) $\{x|x = x^R, x \in \{0,1\}^+\}$

b) $\left\{x\middle|x\text{ 中 0 的个数不少于 1 的个数}, x \in \{0,1\}^+\right\}$

c) $\{xx^Rw|x, w \in \{0,1\}^+\}$

证明.

a)　该语言 L_1 不是正则语言．证明如下：

L_1 是由 0 和 1 组成的回文串的集合.

任取 $x, y \in \{0,1\}^+, x \neq y, |x| = |y|$, 由 Myhill-Nerode 定理, 令 $z = x^R$, 有 $xz = xx^R \in L_1, yz = yx^R \notin L_1$，即任意长度相等的串均属于不同等价类.

因此，R_{L_1} 的指数是无穷的，因此 L_1 不是正则语言．□

b)　该语言 L_2 不是正则语言．证明如下：

由该语言的句子特征可以得到如下一些等价类.

接受状态：

- $[0]$：0 的个数与 1 的个数相等的串所在的等价类；
- $[1]$：0 的个数比 1 的个数多 1 的串所在的等价类；
- $[2]$：0 的个数比 1 的个数多 2 的串所在的等价类；
- ……
- $[n]$：0 的个数与 1 的个数多 n 的串所在的等价类；
- ……

非接受状态：

- $[-1]$：0 的个数比 1 的个数少 1 的串所在的等价类；
- $[-2]$：0 的个数比 1 的个数少 2 的串所在的等价类；
- ……
- $[-n]$：0 的个数与 1 的个数少 n 的串所在的等价类；

- ……

综上所述，R_{L_2} 的指数是无穷的.

因此，由 Myhill-Nerode 定理，L_2 不是正则语言. □

c) 该语言 L_3 不是正则语言. 证明如下：

定义 $x_n = 10^n 1$，任取 $i, j \in \mathbb{N}, i < j$，由 Myhill-Nerode 定理，令 $z = {x_i}^R 1$，有 $x_i z \in L_3, x_j z \notin L_3$. 故对于 $\forall i, j \in \mathbb{N}, i \neq j$，$x_i, x_j$ 均属于不同的等价类.

因此，R_{L_3} 的指数是无穷的. 故 L_3 不是正则语言. □

习题 2 判断下列命题，并证明你的结论.

a) 正则语言的任意子集都是正则语言.

b) 正则语言的补也是正则语言.

c) 无穷多个正则语言的并不一定是正则语言.

证明.

a) 该命题是假命题. 证明如下：

已知 $\{x | x = \{0,1\}^*\}$ 为正则语言，$\{xx^R w | x, w \in \{0,1\}^+\}$ 为其子集. 由习题 1c) 得该子集不是正则语言. □

b) 该命题是真命题. 证明如下：

记正则语言为 L. 因为 L 是正则的，那么存在 DFA $A = (Q, \Sigma, \delta, s, F)$ 使得 $L(A) = L$.

构造 DFA $B = (Q, \Sigma, \delta, s, Q - F)$，即 A 和 B 的接受状态互补，也即它们接受的句子在 Σ^* 上互补. 故对于 L 的补语言 $\bar{L}$，有 $\bar{L} = \Sigma^* - L = L(B)$.

因此，$\bar{L}$ 是正则语言. □

c) 该命题是真命题. 证明如下：

取无穷多个正则语言 L_x：

- $L_0 = \{\varepsilon\}$
- $L_1 = \{0,1\}$
- $L_2 = \{00,11\}$
- $L_3 = \{000,111\}$
- ……
- $L_n = \{0^n, 1^n\}$

- ……

记语言 $L = L_0 \cup L_1 \cup L_2 \cup \cdots \cup L_n \cup \cdots = \{x | x = 0^n 1^n, n \geq 0\}$. L 的句子的主要特点有两个：

i. 句子中所含的 0 的个数与所含的 1 的个数相同.

ii. 所有的 0 都在所有的 1 的前面.

那么，可以得到如下的等价类：

- $[0]$：ε 所在的等价类；
- $[1]$：0 所在的等价类；
- $[2]$：00 所在的等价类；
- $[3]$：000 所在的等价类；
- ……
- $[n]$：0^n 所在的等价类；
- ……

故 R_L 的指数是无穷的. L 不是正则语言. □

习题 3 设 L 是正则语言，字母表是 Σ，定义 $L_{1/3} = \{w \in \Sigma^* | \exists x, y \in \Sigma^*, wxy \in L, |w| = |x| = |y|\}$. 试证明 $L_{1/3}$ 是否正则语言吗？

证明. $L_{1/3}$ 是正则语言，证明如下.

设 f 是一个整数函数，定义 $f(L)$ 为 $\left\{w \middle| 对某个满足\ |x| = f(|w|)\ 的\ x，wx\ 属于\ L\right\}$.

要证原命题，只需证：如果 L 是正则的，那么对于 $f(n) = 2n$，$f(L)$ 也是正则的.

因为 L 是正则的，那么存在 DFA $A = (Q_A, \Sigma, \delta_A, q_0, F_A)$ 使得 $L(A) = L$. 对于 $f(L)$ 构造 DFA $B = (Q_B, \Sigma, \delta_B, [q_0, F_A], F_B)$ 如下：

- $Q_B =$

$$\left\{[q, S] \in Q_A \times P(Q_A) \middle| \forall w \in \Sigma^*, q = \hat{\delta}_A(q_0, w), S = \left\{p \in Q_A \middle| \hat{\delta}_A(p, w') \in F_A, w' \in \Sigma^{f(|w|)}\right\}\right\}$$

- $\delta_B([q, S], a) = [\delta_A(q, a), T]$，其中 $a \in \Sigma$，$T = \{p \in Q_A | \forall b \in \Sigma, \delta_A(p, b) \in S\}$
- $F_B = \{[q, S] \in Q_B | q \in S\}$

以这种构造方式，我们得到：状态 $[q, S]$ 是某个串 w 可以让 A 从初始状态 q_0 转移到状态 q，并且 S 中的状态均可以在接收某个长度为 $f(|w|)$ 的串后转移至 A 的接受状

态.

那么，F_B 中的所有元素都具有“DFA A 接收某个长度的串到达状态 q 后，可再次接收另一个长度为 $f(n)$ 的串到达接受状态”的性质. 因此，$L(B)=f(L)$.

故如果 L 是正则的，那么 $f(L)$ 也是正则的.

因此，若 L 是正则语言，$L_{1/3}$ 也是正则语言. □

习题 4 用正则语言的**扩充泵引理**证明语言 $\{0^n1^m0^m|n,m\geq 1\}$ 不是正则的.

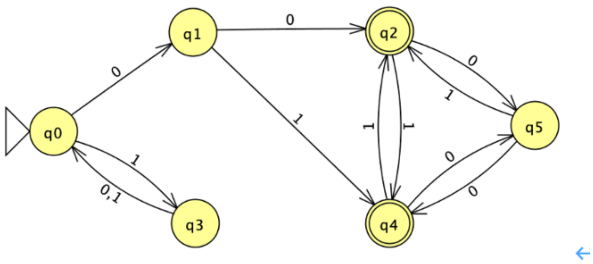
证明. 假设这个语言 L 是正则的. 那么由扩充泵引理，存在一常数 $k>0$，对于语言 L 中每个满足 $|y|\geq k$ 的字符串 $s=xyz$，存在一组 u,v,w 使得 $y=uvw$ 且有 $|uv|\leq k,|v|\geq 1$，并且有 $\forall i\geq 0, xuv^iwz\in L$.

考察串 $s=xyz=0^n1^m0^m, x=0^n1^{m-k}, y=1^k, z=0^m$，显然有 $|y|\geq k$. 按照扩充泵引理对 $y=1^k$ 生成一组分解 u,v,w，其中 v 必仅由 1 构成，且 $|v|\geq 1$.

那么，由 L 是正则的，有 $xuwz\in L$，其中 $xuwz=0^n1^{m-|v|}0^m$；显然，有 $xuwz\notin L$.

因此，存在语言 L 中的串不满足扩充泵引理，这与假设矛盾. 故这个语言不是正则的. □

习题 5 对下图给出的 DFA，求出它的极小状态 DFA，**要求给出主要的求解步骤**.



解答. 使用极小化算法：

① 为所有状态对 $(p,q),p,q\in Q$ 画一张表，开始时表中每个格子内均为空白（未做任何标记）.

② 对 $p\in F,q\notin F$ 的一切状态对 (p,q)，在相应的格子内做标记（画一个“×”），表示 (p,q) 是可以区分的.

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

③ 重复下述过程，直到表中内容不再改变为止：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 $\left(r=\delta(p,a),\ s=\delta(q,a)\right)$ 已做了标记，则在 (p,q) 相应的格子内做标记（红色“×”）.

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

对于 $0 \in \Sigma$，$\delta(q_0,0)=q_1$，$\delta(q_1,0)=q_2$，(1,2) 已标记，故标记 (0,1).

对于 $0 \in \Sigma$，$\delta(q_0,0)=q_1$，$\delta(q_3,0)=q_0$，(0,1) 已标记，故标记 (0,3).

对于 $0 \in \Sigma$，$\delta(q_0,0)=q_1$，$\delta(q_5,0)=q_4$，(1,4) 已标记，故标记 (0,5).

对于 $0 \in \Sigma$，$\delta(q_1,0)=q_2$，$\delta(q_3,0)=q_0$，(0,2) 已标记，故标记 (1,3).

对于 $0 \in \Sigma$，$\delta(q_3,0)=q_0$，$\delta(q_5,0)=q_4$，(1,4) 已标记，故标记 (3,5).

考察完成，(1,5) 和 (2,4) 无法标记.

④ 在完成 ①②③ 之后，所有未被标记的状态对 (p,q) 都是等价的，即 $p \equiv q$，状态 p 和状态 q 可以合并. 可得状态 1 与状态 5 等价，状态 2 与状态 4 等价.

构造极小状态 DFA 如下.