第四次作业：正则语言的性质 2

习题1：

(1) 该语言(L_1)不是正则语言，理由如下：

任取两字符串 $x, y \in \{0,1\}^+$，满足 $x \neq y$ 且 $|x| = |y|$

取 $z = x^R$，则 $xz = xx^R \in L_1$ 且 $yz = yx^R \notin L_1$

因此 $\{0,1\}^+$ 上所有长度相等的不同字符串均属于 R_{L_1} 中不同的等价类

R_{L_1} 具有无穷指数，根据 Myhill-Nerode 定理，该语言不是正则语言

(2) 该语言(L_2)不是正则语言，理由如下：

分析该语言特征可得：可根据 $\{0,1\}^+$ 上字符串中0与1个数的差值来进行等价类的划分

如：[0]：0的个数与1的个数相同的字符串(包含ε)所在的等价类

[1]：0的个数比1的个数多1的字符串所在的等价类

[-1]：0的个数比1的个数少1的字符串所在的等价类

由于0与1个数的差值有无穷多种，因此 R_{L_2} 具有无穷指数，根据 Myhill-Nerode 定理，该语言不是正则语言

(3) 该语言(L_3)不是正则语言，理由如下：

对于 $x = 0^i1^i$，字符串 $z = x^R1 = 1^i0^i1$ 能使 $xz \in L_3$，对 $y = 0^j1^j$ $(i<j)$

$yz = 0^j1^j1^i0^i1 \notin L_3$，故对任意奇数 $i \neq j$，有 $x_i = 0^i1^i$ 与 $x_j = 0^j1^j$ 属于不同等价类

因此 R_{L_3} 具有无穷指数，由 Myhill-Nerode 定理，语言不是正则语言

习题2：

(1) 错误，$\{x \mid x = \{0,1\}^*\}$ 是正则语言，因为其能被正则表达式 0^*+1^* 所描述

由习题1知：(1)(2)(3)均为其子集，但都不是正则语言

(2) 正确 对正则语言 L，由 Myhill-Nerode 定理可得 R_L 具有有穷指数

记 L 的补为 $\bar{L}$，则 $\forall x \in \Sigma^*$，$x \in L \Leftrightarrow x \notin \bar{L}$。因此对于满足等价关系 R_L 的字符串

$x, y \in \Sigma^*$，$xR_Ly \Leftrightarrow (\forall z \in \Sigma^*, xz \in L \Leftrightarrow yz \in L) \Leftrightarrow (\forall z \in \Sigma^*, xz \notin \bar{L} \Leftrightarrow yz \notin \bar{L}) \Leftrightarrow (\forall z \in \Sigma^*, xz \in \bar{L} \Leftrightarrow yz \in \bar{L})$

$\Leftrightarrow xR_{\bar{L}}y$，故等价关系 R_L 与 $R_{\bar{L}}$ 相同，$R_{\bar{L}}$ 具有有穷指数，$\bar{L}$ 是正则语言

(3) 正确 取无穷多个正则语言：$L_0 = \{\epsilon\}$，$L_1 = \{01\}$，$L_2 = \{0011\}$，$L_3 = \{000111\}$

易得：语言 $L = \{0^i1^i \mid i \geq 0\}$ 是上述无穷多个正则语言的并

假设该语言是正则的，且 L 的泵长度为 p，考察字符串 $s = 0^p1^p \in L$，由泵引理可知：$|xy| \leq p$

故设 $y = 0^k$，$0^{p+k}1^p \notin L$ 与假设矛盾，故该语言不是正则语言。

由此得：无穷多个正则语言并不一定是正则语言

习题3：

解：设存在DFA $M_A = (Q_A, \Sigma_A, \delta_A, q_{0A}, F_A)$，使对于语言L，有$L(M_A)=L$。下构造$M=(Q,\Sigma,\delta,q_0,F)$识别$L_{1/3}$

其中Q代表M中的状态 如(q,S)的形式，$q\in Q_A$，用于记录在M中输入某个字符串后对应M_A所在的状态

S表示在M_A中接受了当前读入字符串两倍长度的字符串后，能够到达接受状态的所有状态

转移函数满足$\delta((q,S),a)=(\delta_A(q,a),T)$，其中T表示接受两个字符后能够到达S中任意状态的$M_A$中状态的集合

此外$q_0=(q_{0A},F_A)$对应在A中的起始状态，且接受字符串长为0，即集合对应M_A中的F_A

终止状态F对应所有满足$q'\in S$的状态(q',S)，即代表当前状态到接受状态的长度是当前读入字符串的长度的两倍

易得：上述M满足$L(M)=L_{1/3}$，故$L_{1/3}$是正则语言

习题4：

解：根据扩充泵引理，存在只依赖于L的正整数k，对于任何串x, y, z $(xyz\in L)$

只要$|y|\geq k$，就有y写成$y=uvw$ $(v\neq\varepsilon, |uv|\leq k)$，使对任意$i\geq 0$，有$xuv^iwz\in L$

对字符串$0^n1^k0^k$，其中$x=0^n$，$y=1^k$，$z=0^k$，将y分解为$y=uvw$

则v是由1构成的非空字符串

故$xuv^iwz=0^n1^{k+(i-1)|v|}0^k$，当$i\neq 1$时，$xuv^iwz\notin L$

由此：语言$\{0^n1^m0^m, n,m\geq 1\}$不是正则的

习题5：

解：运用极小化算法进行DFA化简，步骤如下：

(1) 对所有状态对(p,q) $(p,q\in Q)$画表，开始时表中每个格子内均为空白。

(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

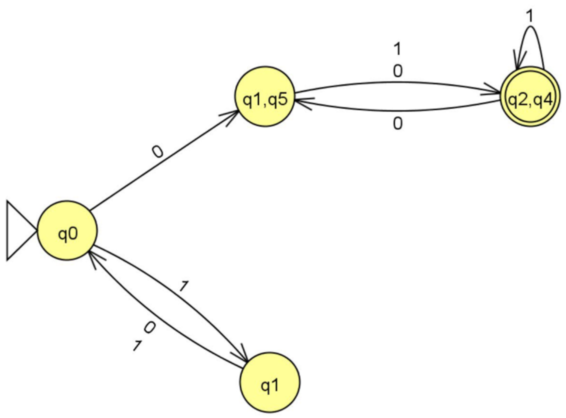
表示(p,q)是可以区分的，对接受状态和非接受状态的状态对的格子内做标记

(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$a\in\Sigma$，若$(r=\delta(p,a), s=\delta(q,a))$已做了标记，则在$(p,q)$相应的格子内做标记

(4) 完成(1)(2)(3)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，将q_1,q_3及q_2,q_4合并后可得DFA图


q0
q1,q5
q2,q4
q1
0
1
0
0
1
0
1
0
1