第 4 次作业 正则语言的性质 2---参考答案

习题 1：试用 Myhill-Nerode 定理，证明语言 $L=\{xwx^T \mid x, w\in\Sigma^+\}$是正则语言。

方法一：运用 MN 定理

语言 L 是正则语言，因为它具有有穷指数，证明如下：

1) 当 x 的长度小于 3 时，有 7 个等价类：

   [ε]: x=ε

   [0]: x=0

   [1]: x=1

   [0 x]: x=00 或 x=01

   [1 x]: x=10 或 x=11

2) 当 x 的长度 ≥3 时，有 4 个等价类：

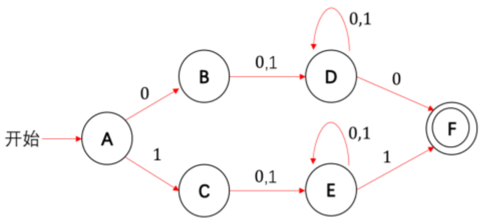
   [1x x]: x = {|x|≥ 3 && 首字母为 1}

   [0x x]: x = {|x|≥ 3 && 首字母为 0 }

3) 上述 7 个等价类两两相交均为空集，7 个等价类的并为Σ^*，所以语言 L 的指数为 7。根据 MH 定理，该语言一定是正则语言。

方法二：非 MN 定理

存在 NFA M（见下图）识别该语言 L，所以语言 L 是正则语言。



习题 2：试用 Myhill-Nerode 定理，证明语言 L={x | x 中 0 的个数不等于 1 的个数，$x\in (0, 1)^*$ }不是正则语言。

我们给出R_L的等价划分：

$[n]$：字符串中0的个数减去1的个数等于n

首先我们证明这是一个划分：

对于任何一个字符串x，x可以数出0的个数和1的个数，相减必然等于某个整数。并且这个整数只有一个。所以这是一个划分。

下面我们来证明$x, y \in [n], \forall z \in L$，$xR_Ly$:

对于$x, y \in [n], \forall z \in L$，$xz, yz$的0的个数和1的个数的差值是相同的，要么都是0，要么都不是0，因此，xz, yz要么都不属于L，要么都属于L。

下面证明$\forall x \in [n_1], \forall y \in [n_2](n_1 \neq n_2)$，$xR_Ly$恒不成立:

对于$\forall x \in [n_1], \forall y \in [n_2](n_1 \neq n_2)$，找到$z \in [-n_1]$，那么$xz \in [0]$，$yz \in [n_2 - n_1]$，因为$n_2 \neq n_1$，所以$n_2 - n_1 \neq 0$，所以$xz \notin L$，$yz \in L$。

由上述证明可知，我们给出的是R_L所确定的等价划分，因此R_L具有无穷指数，L不是正则语言。

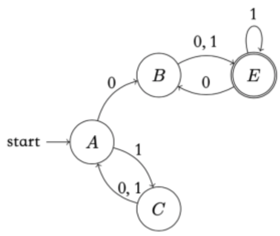
【注】：本题的参考答案摘自于李润中同学的作业。

习题 4：教材 p100 5.6

通过等价状态计算，当表格内容不再改变时为以下情况：

A					
X	B				
X	X	C			
X	-	X	D		
X	X	X	X	E	
X	X	X	X	-	F

所以合并 BD 和 EF 状态。极小状态 DFA 如下图所示：



习题 3：教材 p100 5.7

第 1 小题：L={ 0^i10^j | i , j≥1 }

根据 R_L 的定义，R_L 的等价类完全是由语言 L 来确定的，所以，需要根据 L 的结构特征来寻找 R_L 的等价类。

语言 L 具有明显特征：只有一个 1。

所以，RL 可将 $\sum^*$ 划分成 3 个等价类：

[0] = {x|x 包含 0 个 1} U {ε}

[1] = {x|x 包含 1 个 1}

[2] = {x|x 包含多个 1}

第 2 小题： L = { 0^n1^n | n≥1 }

问题：可以直接使用习题 2 的证明过程吗？

根据 R_L 的定义，R_L 的等价类完全是由语言 L 来确定的，所以，需要根据 L 的结构特征来寻找 R_L 的等价类。

语言 L 具有两个特征：（1）0 的个数和 1 的个数相等；（2）0 在前，1 在后；

因为 $01 \in L$，$001 \notin L$，所以，0 和 00 不在同一个等价类。由此启发，不同个数的 0 属于不同的等价类：

[ε]: 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；

$[0] = \{0^h1^h | h\geq1, m\geq1\}$

$[01]=\{0^h1^m | h=m+1, h\geq1, m\geq1\}$

$[02]=\{0^h1^m | h=m+3, h\geq1, m\geq1\}$

…

$[0n]= \{0^h1^m | h=m+n, h\geq1, m\geq1\}$

…

$[10]=\{0^n1^m | m\geq n+1, m\geq1\} \cup \{x | x$ 中含有子串 $10\}$

显然，这样的等价类有无限多个，即 RL 的指数是无穷的。因此，L 不是正则语言。