4.1.2:

e) 由0和1构成的ww形式的串的集合，也就是某个串重复的串集合。

f) 由0和1构成的ww^R形式的串的集合，也就是由某个串后面跟着它的反转所构成的串的集合。（一个串的逆的形式化定义见4.2.2节。）

g) 由0和1构成的 $w\overline{w}$ 形式的串的集合，其中 $\overline{w}$ 是把w中所有的0都换成1同时把所有的1都换成0而得到的串，例如，$\overline{011}=100$，因此011100是该语言中的一个串。

h) 所有由0和1构成的$w1^n$形式的串的集合，其中w是由0和1构成的长度为n的串。

e) 存在与集合相关的常数n满足泵引理，
构造这样一个串 10^n110^n1
$|10^n110^n1| \geq n$
若L为正则语言，则该串可写作 xyz
∵ $|xy| \leq n$ ∴ y仅有两种可能，一是 10^x $(x<n)$ 二是 0^x $(x<n)$
则 xy^0z 为 $0^{n-x}110^n1$ 或 $10^{n-x}110^n1$ ∵ $|y|>0$
∴ 以上均不属于L，矛盾
故不是正则语言

f) 存在常数n，使得 $|ww^R| \geq n$ 使可写作 xyz
构造这样一个串 10^n110^n1
∴ $|10^n110^n1| \geq n$
若L为正则语言，则该串可写为 xyz $|y| \geq 1$ $|xy| \leq n$
∴ y仅有两种可能，一是 10^x $(x<n)$ 二是 0^x $(x<n)$
∴ xy^0z 可能为 $0^{n-x}110^n1$ 或 $10^{n-x}110^n1$
若要将其划分为 ww^R，则1的数量必须相同，$0^{n-x}110^n1$ 不可能
$10^{n-x}110^n1$ 只能划分到中间，但0的个数不同，所以不可能
∴ 矛盾，故不是正则语言

g) 存在常数n使得L满足泵引理
构造 1^n0^n 满足L，且 $|1^n0^n| \geq n$ ∴ 1^n0^n 可写作 xyz
∵ $|xy| \leq n$ ∴ y只能为 1^x $1 \leq x \leq n$
$xy^0z = 1^{n-x}0^n$ 不属于L，产生矛盾
∴ L不是正则语言

h) 存在常数n使得L满足泵引理

构造 0^n1^n 满足L，$|0^n1^n|\ge n$　∴ 0^n1^n 可写作 xyz

∵ y可能是0^x　$1\le|y|\le n$

$xy^2z = 0^{n+x}1^n$ 不属于L

∴ 矛盾，L不是正则语言

4.1.3:

‼ 习题4.1.3　证明下列语言都不是正则的：

a) 所有满足以下条件的串的集合：由0和1构成，开头的是1，并且当我们把该串看作是一个整数时该整数是一个素数。

b) 所有满足以下条件的0^i1^j形式的串的集合：i和j的最大公约数是1。

存在一常数n使得L满足泵引理

取满足$|w|\ge n$的串w，w可写作xyz

考察素数$q=x\cdot 2^{m+k}+y\cdot 2^k+z$，$w_q$，$|w_q|\ge n$为$q$对应的二进制串，则$|y|=m$　$|z|=k$

∵ L是正则的

∴ $xy^qz\in L$

∴ $p=x\cdot 2^{qm+k}+y\cdot 2^k\cdot\sum_{j=0}^{q-1}2^{jm}+z$ 也为素数

∵ $2^{q-1}\equiv 1 \pmod q$

∴ $2^{qn}=2^{(q-1)n}\cdot 2^n\equiv 2^n \pmod q$

∴ $2^{qn}-1\equiv 2^n-1 \pmod q$

∵ $2^n-1<q$

∴ $\frac{2^{qn}-1}{2^n-1}=1+2^n+\dots+2^{(q-1)n}\equiv 1 \pmod q$

∴ $p\equiv x\cdot 2^{m+k}+y\cdot 2^k+z=q\equiv 0 \pmod q$

即 $q|p$　∴ p不为素数，矛盾

∴ L不是正则的

b) 存在常数n使得L满足泵引理

构造 0^p1^q p为大于等于n的最小质数，q为(p-1)!

则w可写作 xyz $|xy| \le n$ 若 $|y| = m$，

则 $xy^{p-m}z$ 也应属于L

∵ $|xy| \le n$ ∴ y应全为0，x应全为0

相比于 0^p1^q 增加的0的个数为 $(p-m-1)m$

∴现在的0的个数为 $mp - m^2 - m + p = (m+1)(p-m)$ 为合数

和q有公因数 $(p-m)$ ∴矛盾，故不是正则语言

4.2.2:

***! 习题4.2.2** 如果L是一个语言，a是一个符号，则L/a（称作L和a的商）是所有满足如下条件的串w的集合：wa属于L。例如，如果$L = \{a, aab, baa\}$，则$L/a = \{\varepsilon, ba\}$，证明：如果$L$是正则的，那么$L/a$也是。提示：从$L$的DFA出发，考虑接受状态的集合。

根据L的DFA可构造L/a的DFA：

将L的DFA的接受态设置为不可接受，同时将输入a可到达L的DFA的接受态的状态设置为可接受，这样构造出L/a的DFA。

该DFA可以识别L/a，因为~~[illegible]~~其接受状态再输入a被L的DFA识别，符合描述。

∴L/a可以被DFA识别，∴L/a是正则语言

4.2.7:

! 习题4.2.7 如果$w = a_1a_2\cdots a_n$和$x = b_1b_2\cdots b_n$是同样长度的串，定义$alt(w, x)$是把w和x交叉起来且以w开头所得到的串，即$a_1b_1a_2b_2\cdots a_nb_n$。如果$L$和$M$是语言，定义$alt(L, M)$是所有形式为$alt(w, x)$的串的集合，其中$w$是$L$中的任意串，而$x$是$M$中与$w$等长的任意串。证明：如果$L$和$M$都是正则的，那么$alt(L, M)$也是。

设语言L对应的自动机 $A_L=(Q_L,\Sigma,\delta_L,q_L,F_L)$

设语言M对应的自动机 $A_M=(Q_M,\Sigma,\delta_M,q_M,F_M)$

构造这样一个自动机：

① $Q=\{(0,l,m)\mid l\in Q_L, m\in Q_M\}\cup$
$\{(1,l,m)\mid l\in Q_L, m\in Q_M\}$

② $\Sigma=\Sigma$

③ $\delta((0,l,m),a)=(1,l',m)$ $\delta_L(l,a)=l'$

$\delta((1,l,m),a)=(0,l,m')$ $\delta_M(m,a)=m'$

④ $q=(0,q_L,q_M)$

⑤ $F=\{(0,l,m)\mid l\in F_L, m\in F_M\}$

该自动机可识别交叉串：

$\hat{\delta}((0,q_L,q_M),a_1b_1a_2b_2\cdots a_nb_n)=(0,l,m)$

$l=\hat{\delta}_L(q_L,a_1a_2\cdots a_n)\in F_L$

$m=\hat{\delta}_M(q_M,b_1b_2\cdots b_n)\in F_M$

$\therefore (0,l,m)\in F$

若串不是 $a_1b_1a_2b_2\cdots a_nb_n$ 则不能识别

$\therefore$ 语言能被自动机识别 即该语言正则

4.2.8:

***!! 习题4.2.8** 设L是一个语言，定义$half(L)$是所有L中串的前一半构成的集合，即$\{w \mid$ 对于某个满足$|x| = |w|$的x，wx属于$L\}$。例如，如果$L = \{\varepsilon, 0010, 011, 010110\}$，则$half(L) = \{\varepsilon, 00, 010\}$。注意，长度为奇数的串对于$half(L)$没有贡献。证明：如果$L$是正则的，那么$half(L)$也是。

设L对应的自动机DFA $A=(Q_A, \Sigma, \delta_A, q_0, F_A)$

构造自动机B如下：

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

$\delta_B([q,S],a]=[\delta_A(q,a),T] \quad a\in\Sigma \quad T=\{p\in Q_A \mid \forall b\in\Sigma, \delta_A(p,b)\in S\}$

$F_B=\{[q,S]\in Q_B \mid F_A\cap S\neq\phi\}$

这样构造出了具有如下性质的DFA：

该DFA能识别这样的字符串：能被A识别的串的前半部分子串，因为，再输入一个相同长度的串(存在不是任意)即可被自动机A识别，故，~~half(A~~ $half(L)$ 是正则的

4.2.9:

!! 习题4.2.9 我们把习题4.2.8推广到能够决定取走串中多大部分的一系列函数。如果f是一个整数函数，定义$f(L)$为$\{w \mid$ 对某个满足$|x| = f(|w|)$的x，wx属于$L\}$。例如，和运算$half$对应的f是恒等函数$f(n) = n$，因为$half(L)$的定义中有$|x| = |w|$。证明：如果L是正则的，那么对于以下的f，$f(L)$也是正则的：

a) $f(n) = 2n$（也就是取走串的前三分之一）。

b) $f(n) = n^2$（也就是取走的长度是没取走部分长度的平方根）。

c) $f(n) = 2^n$（也就是取走的长度是剩下长度的对数）。

另L对应的DFA $A=(Q_A, \Sigma, \delta_A, q_A, F_A)$，构造如下的DFA B

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

$\delta_B([q,S],a)=[\delta_A(q,a),T]$，$a\in\Sigma$，$T=\{p\in Q_A \mid \forall b\in\Sigma, \delta_A(p,b)\in S\}$

$F_B=\{[q,S]\in Q_B \mid q\in S\}$

构造出的自动机具有这样的性质，接受这样的w，

ww'被自动机A接受 $|w'|=f(|w|)$

所以自动机B识别 $f(L)$

$\therefore f(L)$是正则的

补充 1:

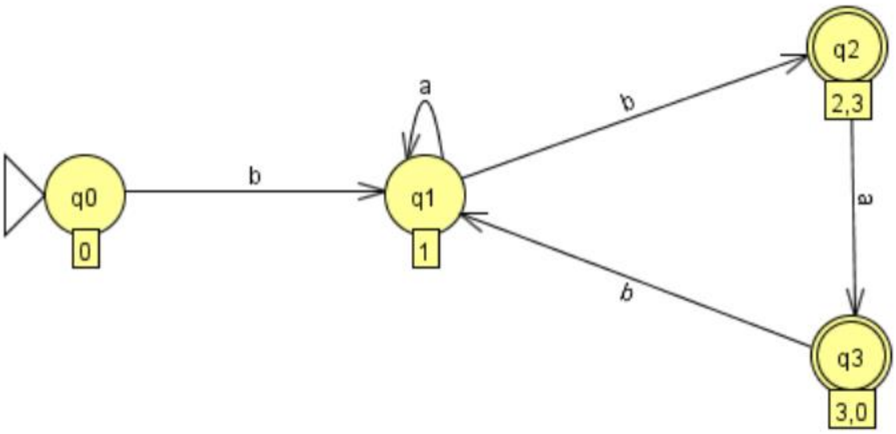
1. 给出如下的正则文法G，求出对应的DFA M，使得L(M) = L(G)。

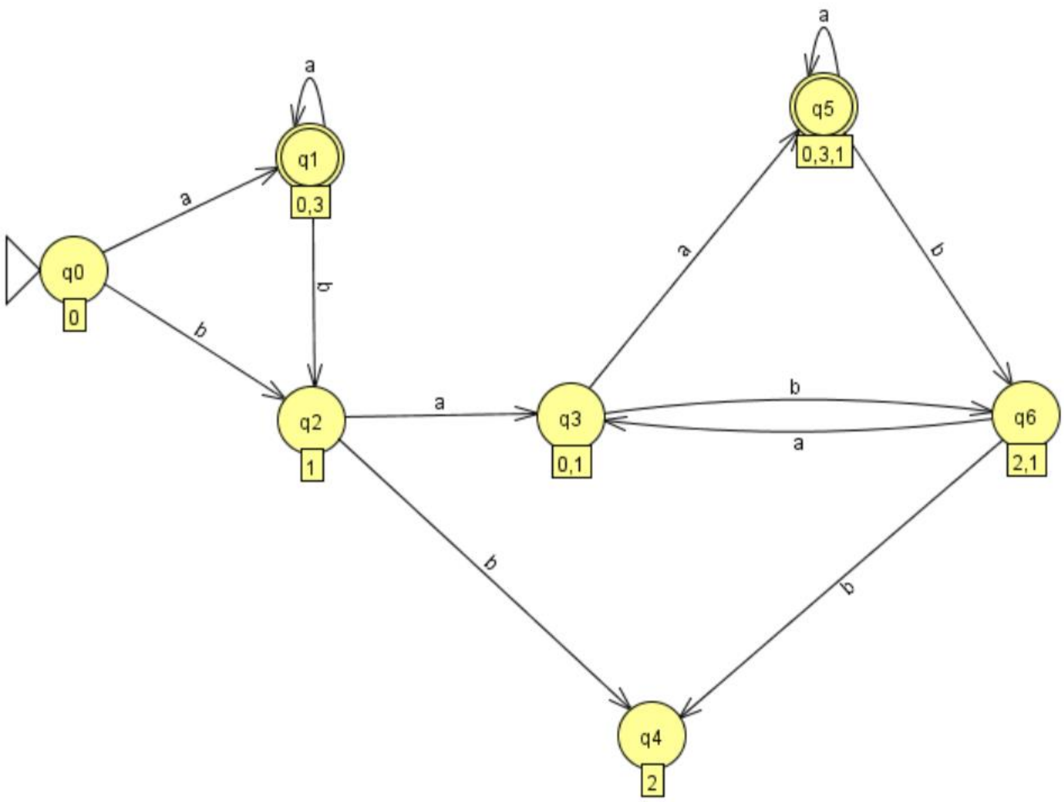
- G_1=（V，T，P_1，S）

P_1: S→bB, B→aB | bA | b, A→a | aS

- G_2=（V，T，P_2，S）

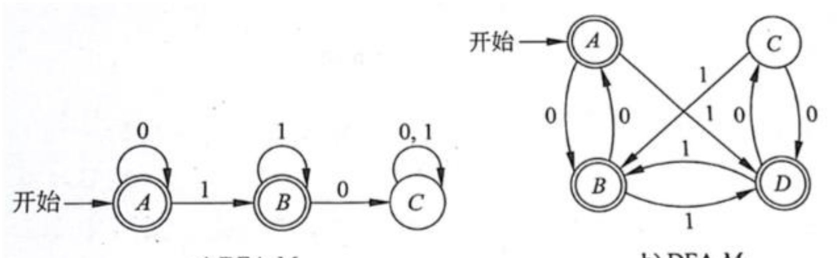
P_2： S→aS | bB | a, B→bA | aB | aS





补充 2:

2. 给出下图描述的两个DFA M，分别求出对应的正则文法G，使得L(G)=L(M)。



a) DFA M_1 b) DFA M_2

G1=(V,T,P1,S)
P1:A->0A|1B|0|1|ε,B->1B|0C|1|ε,C->0C|1C

G2=(V,T,P1,S)
P2=A->0B|1D|0|1|ε,B->0A|1D|0|1|ε,C->0D|1B|0|1,D->0C|1B|1|ε
补充 3:

3. Let L1 ⊆ {0, 1, 2}* be a regular language, we can consider L1 as a subset of integers under base 3, let L2 be the corresponding set of L1 over {0, 1}* (i.e. under base 2), for example if L1 = {11, 12, 121}, then L2 = {100, 101, 10000}. Question: is L2 a regular language?