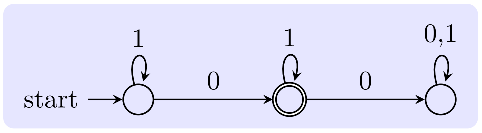
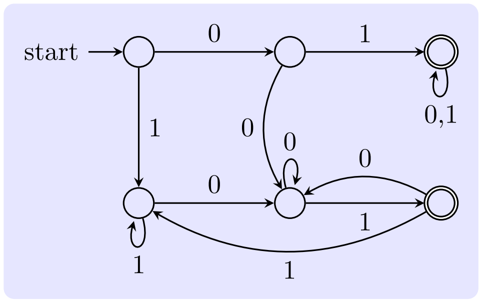
Homework 1 学号 姓名

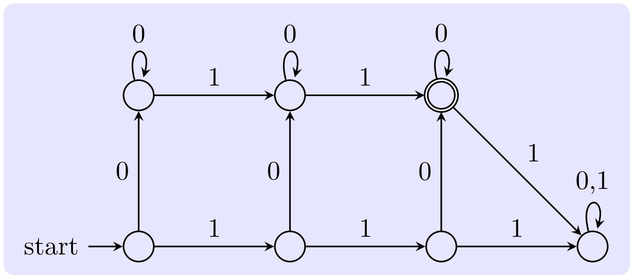
Give DFA's accepting the languages over the alphabet $\{0, 1\}$ (notice : give diagram notation of DFA)

1. The set of all strings with exactly one 0.
2. The set of strings that either begin or end with 01.

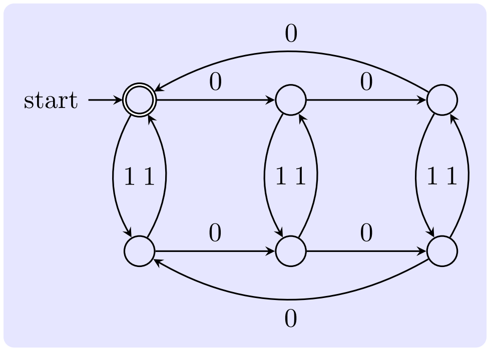




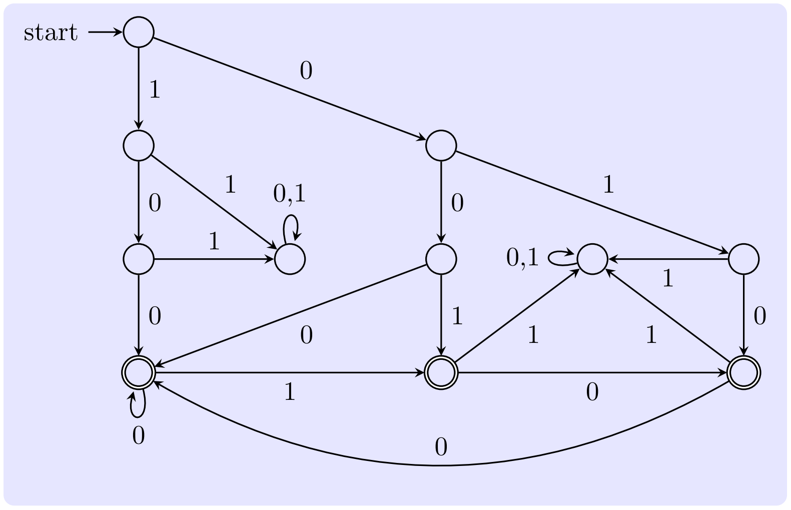
3. The set of all strings with at least one 0 and exactly two 1's.



4. The set of strings such that the number of 0's is divisible by three, and the number of 1's is divisible by two.



5. The set of all strings such that each block of three consecutive symbols contains at least two 0's.

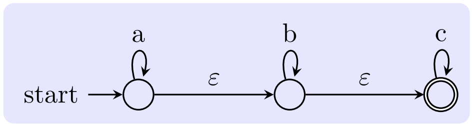


Homework 2

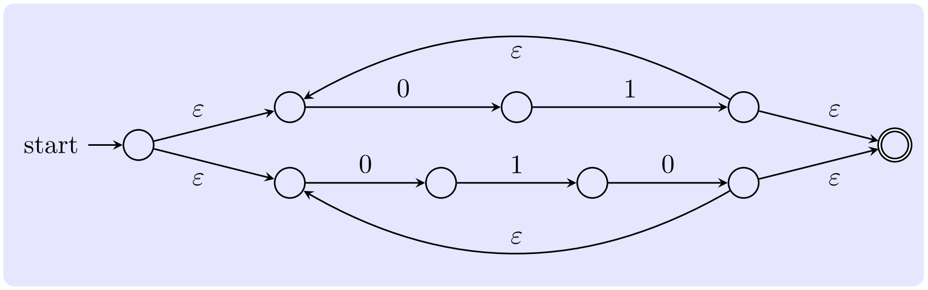
学号 姓名

1. Design ε-NFA's for the following languages. Try to use ε-transitions to simplify your design.

   a) The set of strings consisting of zero or more a's followed by zero or more b's, followed by zero or more c's.

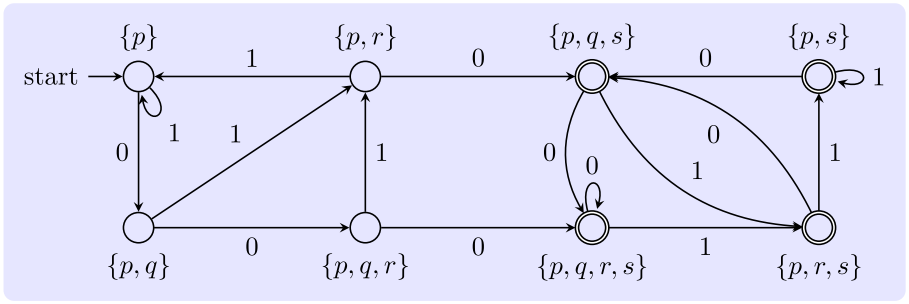


   b) The set of strings that consist of either 01 repeated one or more times or 010 repeated one or more times.



2. Convert to a DFA the following NFA:
   (Note: Give *transition table* or *transition diagram*, DFA states should be labeled with subset of the NFA states if you choose transition diagram.)

	0	1
$\rightarrow p$	$\{p,q\}$	$\{p\}$
q	$\{r\}$	$\{r\}$
r	$\{s\}$	$\emptyset$
$*s$	$\{s\}$	$\{s\}$



3. Give regular expression: the set of all strings of 0's and l's not containing 101 as a substring.

   $0^*(1+000^*)^*0^*$ or $(0+\varepsilon)(1+000^*)^*(0+\varepsilon)$ or $(0+\varepsilon)(1+00+000)^*(0+\varepsilon)$

4. Give regular expression: the set of all strings with an equal number of 0's and l's, such that no prefix has two more 0's than l's, nor two more l's than 0's.

   $(01+10)^*$

5. Give regular expression: the set of all strings of 0's and 1's, such that the number of 0's and the number of 1's are both even.

   $(00+11+(01+10)(00+11)^*(01+10))^*$

Homework 3

学号 姓名

1. Design a context-free grammar for $\{a^i b^j c^k \mid i \neq j \text{ or } j \neq k\}$, that is, the set of strings of a's followed by b's followed by c's, such that there are either a different number of a's and b's or a different number of b's and c's, or both.

   $S \to A_1C \mid A_2C \mid AB_1 \mid AB_2$
   $A_1 \to aA_1b \mid aA_1 \mid a$
   $A_2 \to aA_2b \mid A_2b \mid b$
   $C \to Cc \mid \varepsilon$
   $B_1 \to bB_1c \mid bB_1 \mid b$
   $B_2 \to bB_2c \mid B_2c \mid c$
   $A \to Aa \mid \varepsilon$

   (注意：$Cc \mid \varepsilon$ 若为 $Cc \mid c$ 则不能产生 a, c 同时为 0 个，或 b, c)

2. Design a context-free grammar for the set of all strings with twice as many 0's as 1's.

   $S \to 0S0S1S \mid 0S1S0S \mid 1S0S0S \mid \varepsilon$ 或
   $S \to 0S0S1S \mid 0S1S0S \mid 1S0S0S \mid 001 \mid 010 \mid 110$ (但此文法无法产生 ε) 或
   $S \to 0S0S1 \mid 0S1S0 \mid 1S0S0 \mid SS \mid \varepsilon$

   若仅为 $S \to 0S0S1 \mid 0S1S0 \mid 1S0S0 \mid \varepsilon$ (产生式中缺少 S 的)，则无法产生：001100 或开头结尾都是 1 的串.

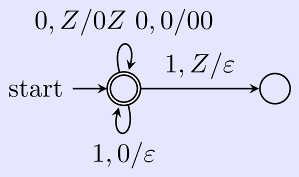
3. Design a context-free grammar for the language consisting of all strings over $\{a, b\}$ that are **not** of the form ww, for some string w. Explain how your grammar works. You needn't prove it's correctness formally.

   如果串长为奇数，显然不是 ww 形式 (对应下面文法中的 A 或 B)。而对于长度为偶数 $(2k)$ 的串，一定有在 j $(1 \leq j \leq k)$ 位置和 $(k+j)$ 位置不相同，且只要有一处即可。可以将偶数串拆分成两个奇数串，只要中间位置不同，分别通过 A 和 B 生成。

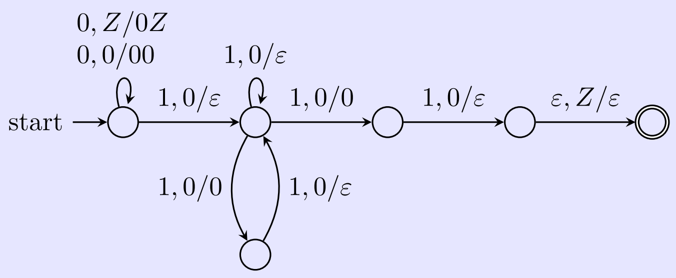
   $S \to A \mid B \mid AB \mid BA$
   $A \to XAX \mid a$
   $B \to XBX \mid b$
   $X \to a \mid b$

   任何偶数串，拆成两个奇数串，两个位于中间的字符不相同即可，如 $aaaabbbb = \underline{aa\check{a}ab}\ \underline{b\check{b}b}$ 或 $aabaaa = \underline{aa\check{b}aa}\ \underline{\check{a}}$.
   但 ww 形式无法生成，拆成任何两个奇数串，中间部分都相等，比如 $baabbaab = \underline{baa\hat{b}baa}\ \underline{\hat{b}} = \underline{b\hat{a}a}\ \underline{bb\hat{a}ab}$.

4. Design a PDA to accept the set of all strings of 0's and l's such that no prefix has more l's than 0's.



5. Design a PDA to accept: $\{0^n 1^m \mid n < m < 2n\}$. You may accept either by final state or by empty stack, whichever is more convenient.



Homework 4 学号 姓名

1. Use the CFL pumping lemma to show following language not to be context-free:
$\{a^i b^j c^k | i < j < k\}$.

   反证法。假设 L=$\{a^i b^j c^k | i < j < k\}$ 是 CFL，由 CFL 泵引理，存在正整数 N，使长度超过 N 的串符合 CFL 泵引理。
   取 $s = a^N b^{N+1} c^{N+2}$ 则 $s = uvwxy$ 中，因为 $|vwx| \le N$ vwx 可能几种分布:

   i) 都在 a 或 b 中，取 $i = 2$ 则 $s' = uv^i wx^i y$ 中 a 或 b 可能不小于 c
   ii) 在 c 中，取 $i = 0$
   iii) 在 ab 之间，取 $i = 2$
   iv) 在 bc 之间，取 $i = 0$

   无论何种情况，都与假设矛盾。得证

2. Consider the CFG G defined by productions:
$$S \to aS|Sb|a|b$$
   Prove by induction on the string length that no string in $L(G)$ has ba as a substring.
   对 $|w|$ 归纳，首先 $|w| = 1$ 显然成立
   假设，对所有 $|w| \le k - 1$ 命题成立
   考查 $|w| = k$ 有 $w = aw_1$ 或 $w = w_2 b$ 而由归纳假设知 w_1, w_2 都成立，显然无论如何无法增加 ba 子串，所以 w 成立

3. Convert the PDA $P = (\{p, q\}, (0, 1), \{X, Z_0\}, \delta, q, Z_0)$ to a CFG, if δ is given by:

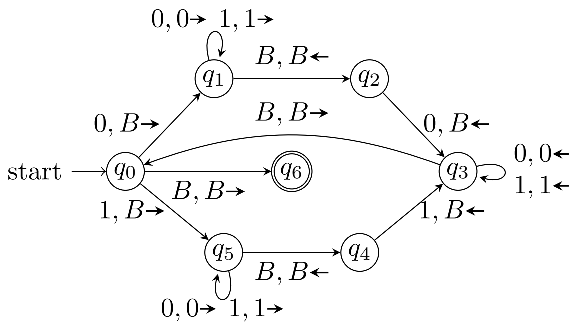
   (1) $\delta(q, 1, Z_0) = \{(q, XZ_0)\}$ (4) $\delta(q, \varepsilon, Z_0) = \{(q, \varepsilon)\}$
   (2) $\delta(q, 1, X) = \{(q, XX)\}$ (5) $\delta(p, 1, X) = \{(p, \varepsilon)\}$
   (3) $\delta(q, 0, X) = \{(p, X)\}$ (6) $\delta(p, 0, Z_0) = \{(q, Z_0)\}$

   1) $S \to [qZx]$ for each x in Q;
   2) $[qAr_n] \to a[pA_1r_1][pA_2r_2] \cdots [pA_nr_n]$ if $(p, A_1A_2 \cdots A_n) \in \delta(q, A, a)$ $n \ge 0$

0	$S \to [qZq]$		✓
	$S \to [qZp]$	$2[qZp]$	step 2, 消掉 $[qZp]$, 因与自己循环
1	$[qZq] \to 1[qXq][qZq]$	$4[qXq]$	
	$[qZq] \to 1[qXp][pZq]$		✓
	$[qZp] \to 1[qXq][qZp]$	$4[qXq]$	step 4, ...
	$[qZp] \to 1[qXp][pZp]$	$3[pZp]$	step 3, 因生成 step 2 中的 $[qZp]$,
2	$[qXq] \to 1[qXq][qXq]$	$4[qXq]$	
	$[qXq] \to 1[qXp][pXq]$	$1[pXq]$	step 1, 消掉 $[pXq]$, 因无此产生式
	$[qXp] \to 1[qXq][qXp]$	$4[qXq]$	
	$[qXp] \to 1[qXp][pXp]$		✓
3	$[qXq] \to 0[pXq]$	$1[pXq]$	
	$[qXp] \to 0[pXp]$		✓
4	$[qZq] \to \varepsilon$		✓
5	$[pXp] \to 1$		✓
6	$[pZp] \to 0[qZp]$	$3[pZp]$	
	$[pZq] \to 0[qZq]$		✓

$S \to [qZq]$	$S \to A$
$[qZq] \to 1[qXp][pZq]$	$A \to 1BC$
$[qXp] \to 1[qXp][pXp]$	$B \to 1BD$
$[qXp] \to 0[pXp]$	$B \to 0D$
$[qZq] \to \varepsilon$	$A \to \varepsilon$
$[pXp] \to 1$	$D \to 1$
$[pZq] \to 0[qZq]$	$C \to 0A$

4. Design Turing machine for the language: $\{ ww^R \mid w$ is any string of 0's and 1's $\}$.



1. 请给出集合｛ε, Φ, ｛Φ｝｝的幂集。

2. 请构造识别语言“由 0 和 1 组成、构造以 01 开头或结尾的字符串”的 DFA。

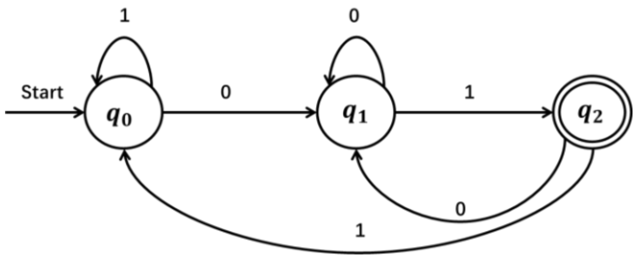
3. 请构造识别语言“由 0 和 1 组成、每三个符号串中至少含有两个 0 的字符串”的 DFA。

4. $\Sigma = \{0,1\}$，根据要求写出正则表达式

(a) $L = \{w|w \in \Sigma^*, w$中 0 和 1 的个数都是偶数$\}$

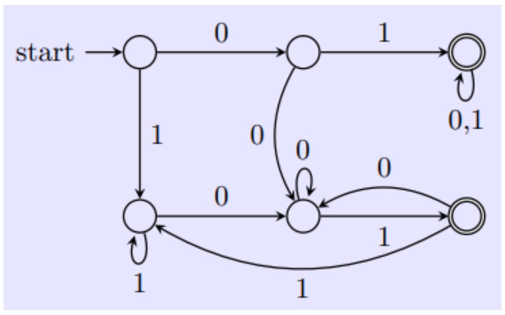
(b) $L = \{w|w \in \Sigma^*, w$不包含 101 子串$\}$

5. 将下图所示的 DFA 用递归构造法造出等价的正则表达式。

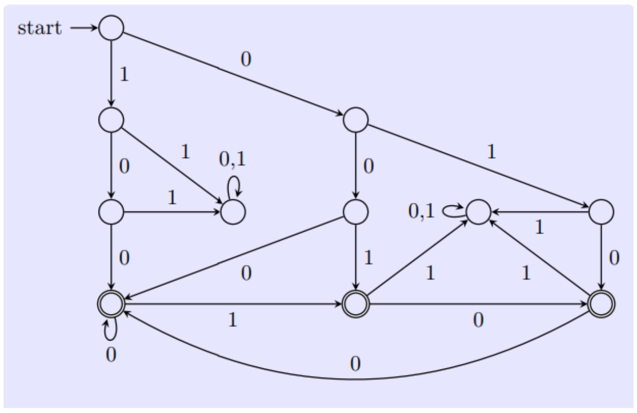

$$\begin{tikzcd}
\arrow[r, "Start"] & q_0 \arrow[loop above, "1"] \arrow[r, "0"] & q_1 \arrow[loop above, "0"] \arrow[r, "1"] & \boxed{q_2} \arrow[l, bend left, "0"] \arrow[ll, bend left=50, "1"]
\end{tikzcd}$$

1. 请给出集合｛ε, Φ, ｛Φ｝｝的幂集。
｛Φ, {ε}, ｛Φ}, ｛{Φ}}, {ε, Φ}, ｛ε, ｛Φ}｝, ｛Φ, ｛Φ｝}, {ε, Φ, ｛Φ｝}}

2. 请构造识别语言“由 0 和 1 组成、构造以 01 开头或结尾的字符串”的 DFA。



3. 请构造识别语言“由 0 和 1 组成、所有长度为 3 子串中至少含有两个 0 的字符串”的 DFA。



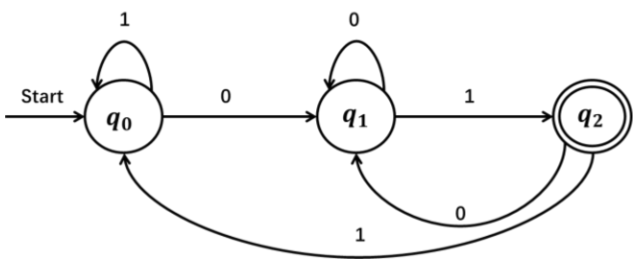
4. $\Sigma = \{0,1\}$，根据要求写出正则表达式
(a) $L = \{w|w \in \Sigma^*, w$中 0 和 1 的个数都是偶数}

$(00 + 11 + (01 + 10)(00 + 11)^* (01 + 10))^*$

(b) $L = \{w|w \in \Sigma^*, w$不包含 101 子串}

$0^*(1 + 000^*)^*0^*$ 或$(0 + \varepsilon)(1 + 000^*)^*(0 + \varepsilon)$或$(0 + \varepsilon)(1 + 00 + 000)^*(0 + \varepsilon)$

5. 将下图所示的 DFA 用递归构造法造出等价的正则表达式。



（1）递归构造法：

$$R^{(1)}_{13} = R^{(0)}_{13} + R^{(0)}_{11}\left(R^{(0)}_{11}\right)^* R^{(0)}_{13} = \emptyset + \emptyset = \emptyset$$

$$R^{(1)}_{12} = R^{(0)}_{12} + R^{(0)}_{11}\left(R^{(0)}_{11}\right)^* R^{(0)}_{12} = 0 + (\varepsilon + 1)(\varepsilon + 1)^*0 = 1^*0$$

$$R^{(1)}_{22} = R^{(0)}_{22} + R^{(0)}_{21}\left(R^{(0)}_{11}\right)^* R^{(0)}_{12} = (\varepsilon + 0) + \emptyset = \varepsilon + 0$$

$$R^{(1)}_{23} = R^{(0)}_{23} + R^{(0)}_{21}(R^{(0)}_{11})^* R^{(0)}_{13} = 1 + \emptyset = 1$$

$$R^{(1)}_{33} = R^{(0)}_{33} + R^{(0)}_{31}(R^{(0)}_{11})^* R^{(0)}_{13} = \varepsilon + 1(\varepsilon + 1)^*\emptyset = \varepsilon$$

$$R^{(1)}_{32} = R^{(0)}_{32} + R^{(0)}_{31}(R^{(0)}_{11})^* R^{(0)}_{12} = 0 + 1(\varepsilon + 1)^*0 = 0 + 11^*0$$

$$R^{(2)}_{13} = R^{(1)}_{13} + R^{(1)}_{12}\left(R^{(1)}_{22}\right)^* R^{(1)}_{23} = \emptyset + (1^*0)(\varepsilon + 0)^*1 = 1^*00^*1$$

$$R^{(2)}_{33} = R^{(1)}_{33} + R^{(1)}_{32}(R^{(1)}_{22})^* R^{(1)}_{23} = \varepsilon + (0 + 11^*0)(\varepsilon + 0)^*1 = \varepsilon + (0 + 11^*0)0^*1$$

$$R^{(3)}_{13} = R^{(2)}_{13} + R^{(2)}_{13}(R^{(2)}_{33})^* R^{(2)}_{33}$$

$$= 1^*00^*1 + (1^*00^*1)(\varepsilon + (0 + 11^*0)0^*1)^*(\varepsilon + (0 + 11^*0)0^*1)$$

作业二

1. 证明语言 $L = \{u^{2n}v^{3n}r^{n} | n \geq 0\}$不是正则语言。

2. 设字母表为$\Sigma = \{a\}$，证明如下语言

$$L = \{a^{n} \in \Sigma* \mid n \geq 2,\ 是一个素数\}$$

不是正则的

3. 构造下推自动机接收如下上下文无关文法定义的语言。

$$S \rightarrow aSbb|abb$$

4. 设计语言 $L = \{a^{i}b^{j}c^{k} \mid i = 2j 或 j = 2k\}$的上下文无关文法，并将其转化为下推自动机 PDA。

5. 构造一个下推自动机接收如下语言：

$$L = \{a^{n}b^{n+m}c^{m} | n \geq 0, m \geq 1\}，其中\sum = \{a, b, c\}$$

6. 给定一个下推自动机 $P = (\{q_0, q_1\}, \{0,1\}, \{X, z_0\}, \delta, q_0, z_0, \emptyset)$，其中

$$\delta(q_0, 0, z_0) = \{(q_0, Xz_0)\}$$

$$\delta(q_0, 0, X) = \{(q_0, XX)\}$$

$$\delta(q_0, 1, X) = \{(q_1, \varepsilon)\}$$

$$\delta(q_1, 1, X) = \{(q_1, \varepsilon)\}$$

$$\delta(q_1, \varepsilon, X) = \{(q_0, \varepsilon)\}$$

$$\delta(q_1, \varepsilon, z_0) = \{(q_0, \varepsilon)\}$$

请将 P 转化为 CFG。

第 1 题解答：

用反证法证明：

假设 L 是正则语言，从 L 中取字符串 $w = u^{2N}v^{3N}r^{N}$，

显然 $s \in L$ 并且$|S| = 6N \geq N$

根据泵引理，w 可以被写成 $w = xyz$，满足条件

1. $xy^{n}z \in L$ 对于任何 $n \geq 0$。
2. $|y| > 0$
3. $|xy| \leq N$

由于 w 的前 N 个字母都是 u，那么结合第三个条件，可以推出结论：x，y 只由 u 组成，所以 x，y，z 可以写为：

$$x = u^{j}\text{ , 对于 } j \geq 0$$

$$y = u^{k}\text{ , 对于 } k \geq 1$$

$$z = u^{m+N}v^{3N}r^{N}\text{ , 对于 } m \geq 0$$

同时 j, k, m 满足 $j + k + m + N = 2N$，即 $j + k + m = N$，所以 $j + k \leq N$。

根据上述的第一个条件，$xy^{2}z \in L$，而

$$xy^{2}z = u^{j}u^{k}u^{k}u^{m+N}v^{3N}r^{N} = u^{2N+k}v^{3N}r^{N}$$

由于 $k \geq 1$，上述的 $xy^{2}z$ 一定不属于 L，所以推出矛盾

因此，得证 $L = \{u^{2n}v^{3n}r^{n} | n \geq 0\}$不是正则语言。

第 2 题解答：

(a) Suppose L is regular and m is given. Let p be the smallest prime number such that $p \geq m$. Then we pick $w = a^p$ in L. The string y must then be a^k and the pumped strings will be

$$w_i = a^{p+(i-1)k} \in L,$$

for $i = 0, 1, \ldots$ However, $w_{p+1} = a^{p+pk} = a^{p(1+k)} \notin L$. Therefore L is not regular.

第 3 题解答：

Convert the grammar into Greibach normal form.

$$\begin{aligned} S &\rightarrow aSSSA|\lambda, \\ A &\rightarrow aB, \\ B &\rightarrow b. \end{aligned}$$

Following the construction of Theorem 7.1, we get the solution

$$\begin{aligned} \delta(q_0, \lambda, z) &= \{(q_1, Sz)\}, \\ \delta(q_1, a, S) &= \{(q_1, SSSA)\}, \\ \delta(q_1, a, A) &= \{(q_1, B)\}, \\ \delta(q_1, \lambda, S) &= \{(q_1, \lambda)\}, \\ \delta(q_1, b, B) &= \{(q_1, \lambda)\}, \\ \delta(q_1, \lambda, z) &= \{(q_f, z)\}. \end{aligned}$$

with $F = \{q_f\}$

注：λ为ε

第 4 题解答：

上下文无关文法如下：

$$S \rightarrow PC \mid AQ$$

$$P \rightarrow aaP\,b \mid \epsilon$$

$$C \rightarrow cC \mid \epsilon$$

$$A \rightarrow aA \mid \epsilon$$

$$Q \rightarrow bbQc \mid \epsilon$$

PDA 如下：

$$P = (\{q\}, \{a, b, c\}, \{a, b, c, A, B, P, Q, S\}, \delta, q, S, ?\,),$$

('?'表示只有一个状态，无终结状态集)

δ如下：

$$\delta(q, \epsilon, S) = \{(q, P\,C), (q, AQ)\}.$$

$$\delta(q, \epsilon, P) = \{(q, aaP\,b), (q, \epsilon)\}.$$

$$\delta(q, \epsilon, C) = \{(q, cC), (q, \epsilon)\}.$$

$$\delta(q, \epsilon, A) = \{(q, aA), (q, \epsilon)\}.$$

$$\delta(q, \epsilon, Q) = \{(q, bbQc), (q, \epsilon)\}.$$

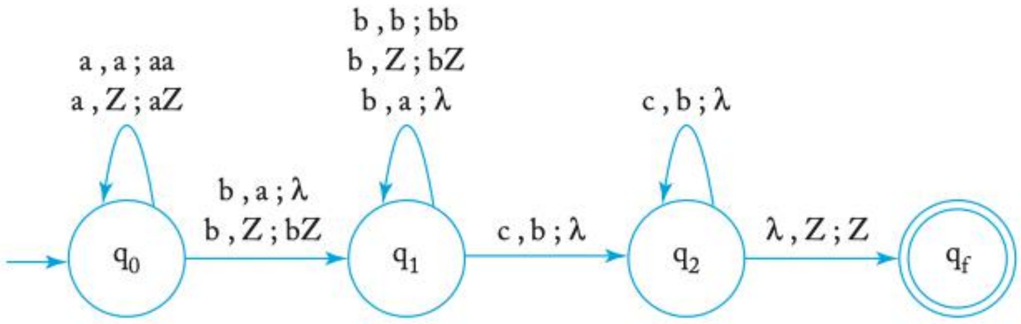
$$\delta(q, a, a) = \{(q, \epsilon)\}.$$

$$\delta(q, b, b) = \{(q, \epsilon)\}.$$

$$\delta(q, c, c) = \{(q, \epsilon)\}.$$

第 5 题解答：

Start with an initial z in the stack. Put a mark a on the stack when input symbol is an a, consume a mark a when input is a b. When all a's in the stack are consumed, put a mark b on the stack when input is a b. After input becomes c, eliminate a mark on the stack. The string will be accepted if the stack has only z left when the input is completed.



第 6 题解答：

首先，令
$V =$
$\{S, [q_0, X, q_0], [q_0, X, q_1], [q_1, X, q_0], [q_1, X, q_1], [q_0, z_0, q_0], [q_0, z_0, q_1], [q_1, z_0, q_0], [q_1, z_0, q_1]\}$
并且 $T = \{0, 1\}$.

δ	产生式	
(0)	$S \to [q_0, z_o, q_0]$	
	$S \to [q_0, z_o, q_1]$	
(1)	$[q_0, z_o, q_0] \to 0[q_0, X, q_0][q_0, z_0, q_0]$	
	$[q_0, z_o, q_0] \to 0[q_0, X, q_0][q_0, z_0, q_1]$	
	$[q_0, z_o, q_1] \to 0[q_0, X, q_0][q_0, z_0, q_1]$	
	$[q_0, z_o, q_1] \to 0[q_0, X, q_1][q_1, z_0, q_1]$	因为 $\delta(q_0, 0, z_0) = \{(q_0, X z_0)\}$
(3)	$[q_0, X, q_0] \to 0[q_0, X, q_0][q_0, X, q_0]$	
	$[q_0, X, q_0] \to 0[q_0, X, q_1][q_1, X, q_0]$	
	$[q_0, X, q_1] \to 0[q_0, X, q_0][q_0, X, q_1]$	
	$[q_0, X, q_1] \to 0[q_0, X, q_1][q_1, X, q_1]$	因为 $\delta(q_0, 0, X) = \{(q_0, XX)\}$
(4)	$[q_0, X, q_1] \to 1$	因为 $\delta(q_0, 1, X) = \{(q_1, \epsilon)\}$
(5)	$[q_1, z_0, q_1] \to \epsilon$	因为 $\delta(q_1, \epsilon, z_0) = \{(q_0, \epsilon)\}$
(6)	$[q_1, X, q_1] \to \epsilon$	因为 $\delta(q_1, \epsilon, X) = \{(q_0, \epsilon)\}$
(7)	$[q_1, X, q_1] \to 1$	因为 $\delta(q_1, 1, X) = \{(q_1, \epsilon)\}$

最终的生成式为

$S \to [q_0, z_0, q_1]$ ， $[q_1, z_0, q_1] \to \epsilon$ ， $[q_1, X, q_1] \to \epsilon$

$[q_0, z_o, q_1] \to 0[q_0, X, q_1][q_1, z_0, q_1]$ ， $[q_1, X, q_1] \to 1$

$[q_1, X, q_1] \to 0[q_0, X, q_1][q_1, X, q_1]$ ， $[q_0, X, q_1] \to 1$

作业 3：

1. 考虑以下两个语言:

L1 $=\{a^n b^{2n} c^m \mid n,m \geq 0\}$

L2 $=\{a^n b^m c^{2m} \mid n,m \geq 0\}$

a) 通过分别给出上述语言的文法来证明这些语言都是上下文无关的。

b) L1∩L2 是 CFG 吗?证明你的结论的正确性

2. 构造 TM M，使得$L(M) = \{x \mid x \in \{0,1\}^* \wedge x$中至多含有 3 个 1$\}$。

3. 设计识别给定的语言$L = \{a^i b^j c^k \mid i * j = k; i,j,k \geq 1\}$的图灵机，其中“a”、“b”和“c”的每个字符串都有一定数量的 a，然后是一定数量的 b，然后是一定数量的 c。条件是这三个符号中的每一个都应至少出现一次。'a' 和 'b' 可以出现多次，但 'c' 的出现次数必须等于 'a' 的出现次数和 'b' 的出现的次数乘积。假设字符串以“$”结尾。

示例：

输入： aabbbcccccc

这里 a = 2, b = 3, c = 2 * 3 = 6

输出：接受

输入：aabbccc

这里 a = 2, b = 2, c = 3 但 c 应该是 4

输出：不接受