南京邮电大学 2011/2012 学年第 一 学期 7元

《信息论与编码》期末试卷（A）

院(系)_____ 班级_____ 学号_____ 姓名_____

题号	一	二	三	四	五	六	七	八	九	十	总 分
得分											

装 订 线 内 不 要 答 题

自觉遵守考试规则，诚信考试，绝不作弊

得 分

一、填空题（25 分，每空 2 分，最后一个 3 分）

1、1948 年，香农发表了题为 通讯的数学理论 的论文，奠定了信息论研究的基础。

2、信源编码的最终目的是 使信息量有效的传输，而信道编码的最终目的是 使信息无误码的传输

3、信道输入输出的平均互信息是输入概率分布的 凹 函数，是信道转移概率分布的 凸 函数。

4、对长度为 N 的信源符号序列进行等长编码，信源符号个数为 q，码符号个数为 r，则码长应满足 $L \geq N \frac{\log q}{\log r}$。

5、信道矩阵 $\begin{bmatrix} 1/4 & 3/4 \\ 3/4 & 1/4 \end{bmatrix}$ 代表的信道的信道容量等于 0.1887 bit/s，达到信道容量的条件是 信源等概率分布

6、设某二进制码{00011，10110，01101，11000，10010，10001}，则码的距离是 2，若通过二元对称信道接收码字为 01100，应译码为 01101。

7、对二元信源分布 $\begin{bmatrix} 0 & 1 \\ \omega & 1-\omega \end{bmatrix}$，对应的失真矩阵 $D=\begin{bmatrix} 0 & \alpha \\ \alpha & 0 \end{bmatrix}$，则 D_{min} 等于 0，D_{max} 等于 $\begin{cases} \omega & 0<\omega \leq \frac{1}{2} \\ 1-\omega & \frac{1}{2}<\omega \leq 1 \end{cases}$。

得 分

二、计算、证明题（55 分，第 1,2,3 题 15 分，第 4 题 10 分）

1. 有一阶平稳马氏信源 X，符号集是{0, 1, 2}，其中符号转移概率

9

为 P(1/0)=P(0/1)=P(0/2)=P(2/0)=P(1/2)=P(2/1)=0.25, 而 P(0/0)=P(1/1)=P(2/2)=0.5, (1)求这个一阶马尔科夫信源的信息熵; (2)对此马氏信源的二次扩展信源进行二元 huffman 编码, 并求编码后的平均码长和编码效率。

(1) $P=\begin{pmatrix}0.5 & 0.25 & 0.25\\ 0.25 & 0.5 & 0.25\\ 0.25 & 0.25 & 0.5\end{pmatrix}$

$\vec{W}=(W_1, W_2, W_3)$

$\vec{W}P=\vec{W}$

$\vec{W}=(\frac{1}{3},\frac{1}{3},\frac{1}{3})$

$$\begin{cases}0.5W_1+0.25W_2+0.25W_3=W_1\\ 0.25W_1+0.5W_2+0.25W_3=W_2\\ 0.25W_1+0.25W_2+0.5W_3=W_3\\ W_1+W_2+W_3=1\end{cases}$$

$H(X|S_1)=H(0.5, 0.25, 0.25)=\frac{3}{2}$

$H(X|S_2)=H(X|S_3)=\frac{3}{2}$

$H_\infty=\sum_i W_i H(X|S_i)=\frac{3}{2}$ bit/符号

(2)

X_1X_1	4/6	11 0
X_1X_2	1/12	1001
X_1X_3	1/12	011
X_2X_1	1/12	010
X_2X_2	1/6	111
X_2X_3	1/12	000
X_3X_1	1/12	001
X_3X_2	1/12	1000
X_3X_3	1/6	101

$\bar{K}=\sum_i P(i)K_i=\frac{19}{6}\approx 3.16$

$\frac{\bar{K}}{2}=1.58$

$\eta=\frac{H_\infty}{\bar{K}/2}=94.9\%$

2. 设一离散信道的信道矩阵为 $\begin{bmatrix}0.7 & 0.1 & 0.2\\ 0.2 & 0.1 & 0.7\end{bmatrix}$, 求: (1) 最佳概率分布和信道容量; (2)当 $P(x_1)=0.7$, $P(x_2)=0.3$ 时, 求 I(X;Y)和 H(X/Y); (3)当输入为等概率分布时, 试写出译码准则, 使平均译码错误率最小, 并求该值。

(1) 离散无损信道为准DMC信道 $\begin{bmatrix}0.7\\ 0.2\end{bmatrix}\begin{bmatrix}0.1\\ 0.1\end{bmatrix}\begin{bmatrix}0.2\\ 0.7\end{bmatrix}$

信道容量 $C=(\log 2)^2-H(0.7, 0.1, 0.2)=0.9\log 0.9-0.1\log 0.2$

$=0.2422$ bit/s

当且仅当等概率分布时, 才能达到概率分布

即 $P(x_1)=\frac{1}{2}$ $P(x_2)=\frac{1}{2}$

此时 $\begin{bmatrix}0.35 & 0.05 & 0.1\\ 0.1 & 0.05 & 0.35\end{bmatrix}$

$0.1+0.05+0.1=0.25$

(2) $P(x_1)=0.7$ $P(x_2)=0.3$ $P(y_j|x_i)=\begin{bmatrix}0.7 & 0.1 & 0.2\\ 0.2 & 0.1 & 0.7\end{bmatrix}$

联合概率矩阵为 $\begin{bmatrix}0.49 & 0.07 & 0.14\\ 0.06 & 0.03 & 0.21\end{bmatrix}$

$\therefore P(y_1)=0.55$ $P(y_2)=0.1$ $P(y_3)=0.35$

$P(x_1|y_1)=\frac{P(x_1)P(y_1|x_1)}{P(y_1)}=\frac{49}{55}$

同理可得 $P(x_1|y_2)=0.7$ $P(x_1|y_3)=0.4$

$P(x_2|y_1)=\frac{6}{55}$ $P(x_2|y_2)=0.3$ $P(x_2|y_3)=0.6$

$H(X|Y)=-\sum_i\sum_j P(x_i,y_j)\log P(x_i|y_j)=0.7214$ bit/符号

$H(X)=-0.7\log 0.7-0.3\log 0.3=0.8815$ bit/符号

$I(X;Y)=H(X)-H(X|Y)=0.1601$ bit/符号

③ $P=\begin{bmatrix}0.7 & 0.1 & 0.2\\ 0.2 & 0.1 & 0.7\end{bmatrix}$

∵输入为概率分布

∴ $p(x_1)=0.5$ $P(x_2)=0.5$

译码准则 $\begin{cases}F(b_1)=a_1=0.7\\ F(b_2)=a_2=0.1\\ F(b_3)=a_2=0.2\end{cases}$

误码率 $P_E=\sum_{ij}P(a_ib_j)-\sum(a^*b_j)$

$=\sum_{j\neq a^*}P(a_ib_j)$

$=0.5\times(0.2+0.2+0.1)$

$=0.25$

3. 某系统（8，4）码，其后4位校验位$v_i, i=0,1,2,3$与信息位$u_i, i=0,1,2,3$的关系式：$v_0=u_1+u_2+u_3$，$v_1=u_1+u_2+u_0$，$v_2=u_1+u_0+u_3$，$v_3=u_0+u_2+u_3$

求：该码的生成矩阵和校验矩阵，信息元(0, 1, 1, 0)对应的码字是什么？

∵ 后4位校验位 v_i 与信息位 u_i 之间关系满足

$$\begin{cases}v_0=u_1+u_2+u_3\\ v_1=u_1+u_2+u_0\\ v_2=u_1+u_0+u_3\\ v_3=u_0+u_2+u_3\end{cases}$$

∴生成矩阵 $G=[\quad]$

校验矩阵 $H=[\quad]$

信息元为 (0, 1, 1, 0)

$\therefore C=m\cdot G=(0110)\begin{pmatrix}\quad\end{pmatrix}$

$=(0110\,0011)$

∴信息元(0, 1, 1, 0)对应的码字为：(0110 0011)

4. 证明：H(X)是输入概率分布P(x)的凹函数。

得分

三、论述题（20 分）

1(8 分). 循环冗余码可以作为数据完整性校验的工具嘛？

~~循环冗余码之可以作为数据完整性校验工具~~
~~我理解的生成一个码字~~

2(12 分). 简要叙述信息论三大定理的内容。

(1) 无失真信源编码定理

码字平均长度不小于对信源熵时可以得到几乎无失真的编码

(2) 信道编码定理

输入信源熵不超过信道容量C时可以得到几乎无失真编码

(3) 限失真信源编码定理

信息率R>R(D)则可以保证失真度不会超过D

装订线内不要答题

自觉遵守考试规则，诚信考试，绝不作弊

南京邮电大学 2010/2011 学年第 1 学期

《信息论与编码》期末试卷（A）

1. 在无失真的信源中，信源输出由 $H(X)$ 来度量；在有失真的信源中，信源输出由 $R(D)$ 来度量。

2. 要使通信系统做到传输信息有效、可靠和保密，必须首先 信源 编码，然后 加密 编码，再 信道 编码，最后送入信道。

3. 带限 AWGN 波形信道在平均功率受限条件下信道容量的基本公式，也就是有名的香农公式是 $C = W\log(1+SNR)$；当归一化信道容量 C/W 趋近于零时，也即信道完全丧失了通信能力，此时 E_b/N_0 为 -1.6 dB，我们将它称作香农限，是一切编码方式所能达到的理论极限。

4. 保密系统的密钥量越小，密钥熵 $H(K)$ 就越 小，其密文中含有的关于明文的信息量 $I(M;C)$ 就越 大。

5. 已知 $n=7$ 的循环码 $g(x)=x^4+x^2+x+1$，则信息位长度 k 为 3，校验多项式 $h(x)=$ x^3+x+1。

6. 设输入符号表为 $X=\{0, 1\}$，输出符号表为 $Y=\{0, 1\}$。输入信号的概率分布为 $p=(1/2, 1/2)$，失真函数为 $d(0,0)=d(1,1)=0$，$d(0,1)=2$，$d(1,0)=1$，则 $D_{min}=$ 0，$R(D_{min})=$ 1bit/symbol，相应的编码器转移概率矩阵 $[p(y/x)]=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$；$D_{max}=$ 0.5，$R(D_{max})=$ 0，相应的编码器转移概率矩阵 $[p(y/x)]=\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}$。

7. 已知用户 A 的 RSA 公开密钥 $(e,n)=(3,55)$，$p=5, q=11$，则 $\phi(n)=$ 40，他的秘密密钥 $(d,n)=$ (27,55)。若用户 B 向用户 A 发送 $m=2$ 的加密消息，则该加密后的消息为 8。

二、判断题

1. 可以用克劳夫特不等式作为唯一可译码存在的判据。（√）
2. 线性码一定包含全零码。（√）
3. 算术编码是一种无失真的分组信源编码，其基本思想是将一定精度数值作为序列的编码，是以另外一种形式实现的最佳统计匹配编码。（×）
4. 某一信源，不管它是否输出符号，只要这些符号具有某些概率特性，就有信息量。（×）
5. 离散平稳有记忆信源符号序列的平均符号熵随着序列长度 L 的增大而增大。（×）
6. 限平均功率最大熵定理指出对于相关矩阵一定的随机矢量 X，当它是正态分布时具有最大熵。（√）
7. 循环码的码集中的任何一个码字的循环移位仍是码字。（√）
8. 信道容量是信道中能够传输的最小信息量。（×）
9. 香农信源编码方法在进行编码时不需要预先计算每个码字的长度。（×）
10. 在已知收码 R 的条件下找出可能性最大的发码 C_i 作为译码估计值，这种译码方法叫做最佳译码。（√）

三、计算题

某系统 (7, 4) 码

$c=(c_6\ c_5\ c_4\ c_3\ c_2\ c_1\ c_0)=(m_3\ m_2\ m_1\ m_0\ c_2\ c_1\ c_0)$ 其三位校验位与信息位的关系为：

$$\begin{cases} c_2 = m_3 + m_1 + m_0 \\ c_1 = m_3 + m_2 + m_1 \\ c_0 = m_2 + m_1 + m_3 \end{cases}$$

（1）求对应的生成矩阵和校验矩阵；

（2）计算该码的最小距离；

（3）列出可纠差错图案和对应的伴随式；

（4）若接收码字 R=1110011，求发码。

解：1. $G=\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$ $H=\begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$

2. d_{min}=3

3.

S	E
000	0000000
001	0000001
010	0000010
100	0000100
101	0001000
111	0010000
011	0100000
110	1000000

4. RH^T=[001]　接收出错

E=0000001　　$R+E=C$= 1110010 (发码)

四、计算题

已知(X,Y)的联合概率 $p(x,y)$ 为：

求 $H(X)$，$H(Y)$，$H(X,Y)$，$I(X;Y)$

X \ Y	0	1
0	1/3	1/3
1	0	1/3

解：$p(x=0)=2/3$　　$p(x=1)=1/3$

$p(y=0)=1/3$　　$p(y=1)=2/3$

$H(X)=H(Y)=H(1/3,2/3)=0.918$ bit/symbol

$H(X,Y)=H(1/3,1/3,1/3)=1.585$ bit/symbol

$I(X;Y)=H(X)+H(Y)-H(X,Y)=0.251$ bit/symbol

五、计算题

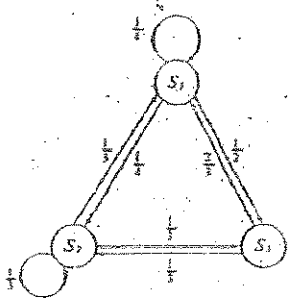
一阶齐次马尔可夫信源消息集 $X\in\{a_1,a_2,a_3\}$，

状态集 $S\in\{S_1,S_2,S_3\}$，且令 $S_i=a_i,i=1,2,3$，条件转移概率为

$[P(a_j/S_i)]=\begin{bmatrix}1/4 & 1/4 & 1/2\\ 1/3 & 1/3 & 1/3\\ 2/3 & 1/3 & 0\end{bmatrix}$，(1)画出该马氏链的状态转移图；

(2)计算信源的极限熵。

解：(1)



(2) $\begin{cases}\frac{1}{4}W_1+\frac{1}{3}W_2+\frac{2}{3}W_3=W_1\\ \frac{1}{4}W_1+\frac{1}{3}W_2+\frac{1}{3}W_3=W_2\\ \frac{1}{2}W_1+\frac{1}{3}W_2=W_3\\ W_1+W_2+W_3=1\end{cases}\rightarrow\begin{cases}W_1=0.4\\ W_2=0.3\\ W_3=0.3\end{cases}$

$H(X|S_1)=H(1/4,1/4,1/2)=1.5$ 比特/符号

$H(X|S_2)=H(1/3,1/3,1/3)=1.585$ 比特/符号

$H(X|S_3)=H(2/3,1/3)=0.918$ 比特/符号

$H_\infty=\sum_{i=1}^{3}w_iH(X|S_i)=0.4\times1.5+0.3\times1.585+0.3\times0.918=1.351$ 比特/符号

六、计算题

若有一信源 $\begin{bmatrix}X\\ P\end{bmatrix}=\begin{bmatrix}x_1 & x_2\\ 0.8 & 0.2\end{bmatrix}$，每秒钟发出 2.55 个信源符号。

将此信源的输出符号送入某一个二元信道中进行传输

（假设信道是无噪无损的；容量为 1bit/二元符号），

而信道每秒钟只传递 2 个二元符号。

（1） 试问信源不通过编码（即 $x_1\rightarrow0,x_2\rightarrow1$ 在信道中传输）

15

（2） 能否直接与信道连接？

（3） 若通过适当编码能否在此信道中进行无失真传输？

（4） 试构造一种哈夫曼编码(两个符号一起编码)。

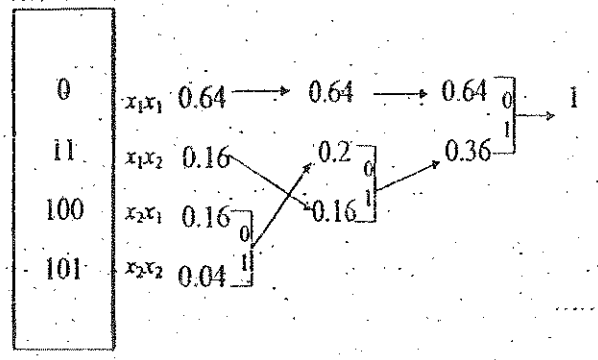
（5） 使该信源可以在此信道中无失真传输。

解：1.不能，此时信源符号通过 0，1 在信道中传输，2.55 二元符号/s>2 二元符号/s

2. 从信息率进行比较， $2.55^* H(0.8,0.2)=1.84<1^*2$

可以进行无失真传输

3.



$\overline{K}=\sum_{i=1}^{4} p_i K_i = 0.64+0.16*2+0.2*3=1.56$ 二元符号/2 个信源符号

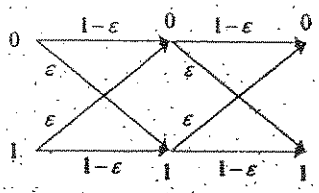
此时 1.56/2*2.55=1.989 二元符号/s < 2 二元符号/s

七、计算题

两个 BSC 信道的级联如右图所示：

（1）写出信道转移矩阵；

（2）求这个信道的信道容量。



解： (1)

$$P=P_1P_2=\begin{bmatrix}1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon\end{bmatrix}\begin{bmatrix}1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon\end{bmatrix}=\begin{bmatrix}(1-\varepsilon)^2+\varepsilon^2 & 2\varepsilon(1-\varepsilon) \\ 2\varepsilon(1-\varepsilon) & (1-\varepsilon)^2+\varepsilon^2\end{bmatrix}$$

(2) $C=\log 2-H((1-\varepsilon)^2+\varepsilon^2)$

12

南京邮电大学 2009/2010 学年第一学期
《信息论与编码》期末试卷（A）

班级_____ 姓名_____ 学号_____

题号	一	二	三	四	五	六	七	总分
得分								

一、填空题：（20 分）

1. 冗余度表征信源信息率多余程度的一个物理量。它来自两个方面，一是信源符号间_____，另一个是信源符号分布的_____。

2. 在信息传输理论中，互信息是输入符号概率分布和转移概率分布的函数，当_____一定时，选择_____使互信息达到最大值，这就是_____；当_____一定时，选择_____可使互信息达到最小值，这就是_____。

3. 通常，可将信息与通信中的基本问题归纳为三性：_____、_____和_____，分别通过_____、_____和_____编码来实现。（3 分）

4. （6，3）码的许用码字为_____，_____，_____，_____，_____，_____，_____，_____，000000，最小码距为_____。该码组若用于检错，能检出_____位错码；若用于纠错，能纠正_____位错码。

5. 假设两码字 A，B 相距 7（$d_{min}=7$），它的最大纠错能力为_____，若检错能力为 4，则纠错能力为_____。

6. 保密系统的密钥量越少，密钥熵 H(K)就越_____，其密文中含有的关于明文的信息量 I(M；C)就越_____。H(M)，H(K)及 I(M；C)三者间的关系为_____。

7. DES 的中文名称是_____，它采用的是_____密码体制。

8. 一个理想的保密系统，应满足对合法用户信宿能安全、保密地给出信源明文的全部信息，即 I(M；C)=_____；而对窃听者应使之得不到任何明文信息，即 I(M；C)=_____。

二、判断题：（对的写“√”，错的写“×”）。（10 分）

1. 只能用生成多项式来描述循环码。（×）

2. 一个密码体制的安全性，既依赖于其密钥的保密性，又依赖于其加密、解密算法的保密性。（×）

3. 在已知收码r的条件下找出可能性最大的发码Ci作为译码估计值，这种译码方法叫做最大似然译码。(最佳译码) P124 (X)

4. 信道中各码元是否出现差错，与其前后码元是否差错无关，每个码元独立地按一定的概率产生差错，称为突发差错。 (X)

5. 线性码一定包含全零码。 (√)

6. 在信道编码和密码中，符号之间的相关性越强越好。(越弱) (X)

7. 可以用克劳夫特不等式作为唯一可译码的判据 (X)

8. 信道编码的任务就是构造出以最小多余度代价换取最大抗干扰性能的"好码"。 (√)

9. 信源编、译码是寻求实现通信系统与信道统计特性相匹配的编、译码 (√)

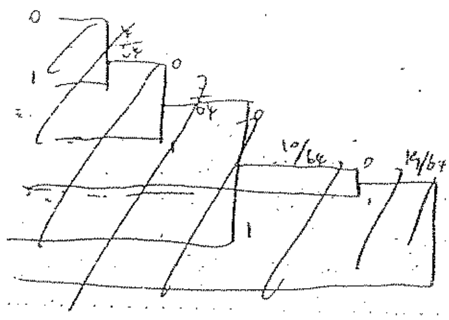
10. 算术编码是一种无失真的分组信源编码，其基本思想是将一定精度数值作为序列的编码，是以另外一种形式实现的最佳统计匹配编码。 ( )

三、设有一个无记忆信源发出符号A和B，已知 $p(A)=\frac{1}{4}$，$p(B)=\frac{3}{4}$，(1) 计算该信源熵；(2) 又设该信源改为发三重符号序列消息的信源，用霍夫曼编码方法进行二进制编码，求其编码效率。(10分)

(1) $H(X) = -\sum p(x)\log p(x) = \frac{1}{4}\text{lb}4 + \frac{3}{4}\text{lb}\frac{4}{3}$

(2) ∵无记忆 ∴ $p(xyz) = p(x)p(y)p(z)$

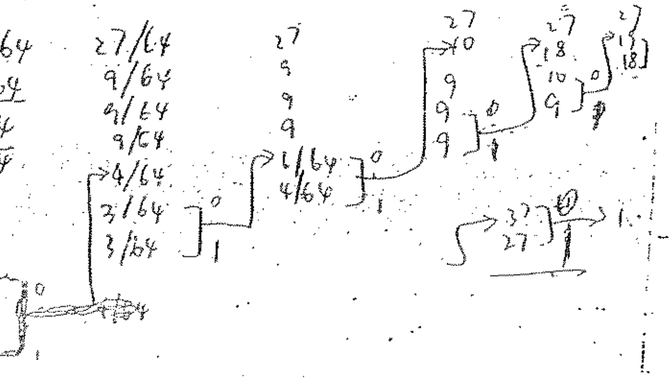
AAA	$\frac{1}{4}\times\frac{1}{4}\times\frac{1}{4}=\frac{1}{64}$
AAB	$\frac{1}{4}\times\frac{1}{4}\times\frac{3}{4}=\frac{3}{64}$
ABA	$\frac{3}{64}$
ABB	$\frac{1}{4}\times\frac{3}{4}\times\frac{3}{4}=\frac{9}{64}$
BAA	$\frac{3}{64}$
BAB	$\frac{9}{64}$
BBA	$\frac{9}{64}$
BBB	$\frac{27}{64}$



$\bar{K} = (27+3\times9+3\times9+3\times9+5\times3\times3+5)\cdot\frac{1}{64}$
$= (27+81+50)\frac{1}{64}$
$= \frac{158}{64} = \frac{79}{32}$

$\eta = \frac{H(X)}{\bar{K}} = \ldots$

0	BBB	27/64
101	BBA	9/64
110	BAB	9/64
111	ABB	9/64
10000	BAA	3/64
10001	ABA	3/64
10010	AAB	3/64
10011	AAA	1/64



四、(7, 4) 码的生成矩阵 $C=\begin{bmatrix}1&0&0&0&0&1&0\\0&1&0&0&1&1&1\\0&0&1&0&0&1&0\\0&0&0&1&1&0&1\end{bmatrix}$

(1) 求系统形式的生成矩阵 G 和校验矩阵 H；

(2) 计算该码的最小距离；

(3) 若 $m=1100$，求相应码字；

(4) 列出可纠单错图样表和伴随式；

(5) 求其缩短 (5, 2) 码的生成矩阵；

(6) 若接收码字 $R=1110000$，求发送及信息位。(20分)

五、发送者用 $(d_B, n_B)=(11,51)$ 和 $(e_A, n_A)=(3,33)$ 签署报文 B(B=02)，接收者收到密文后将加以认证。(18分)

计算签名：用户B用密钥计算签名 $S_i=(M_i)^{d_B} \bmod n_B$

再次用A的公开密钥计算签名 $S=(S_i)^{e_A} \bmod n_A$

验证签名：用户A用他的秘密密钥验证签名

再用用户B的公开密钥验证

签名：

设某语音信号{x(t)}，其最高频率为 4kHz，经取样、量化后编成等长码，设每个样本的分层数为 128。

(1) 求此语音信号的信息传输速率是多少（比特/秒）；

(2) 把这一语音信号送入一噪声功率谱为 5×10^{-6}mW/Hz，带宽为 4kHz 的高斯信道中传输，试求无差错传输时需要的最小输入功率。

解：

(1)考虑每一分层是等概率分布，得每一样本含有的信息量为

$R=\log 128=7bit/样本$

因为最高频率为$4Hz$,取样速率为$8Hz$，所以此语音信号的信息传输速率为

$R=8\times10^{3}\times7=5.6\times10^{4}bit/s$

(2)

$$R\le C=W\log(1+\frac{P}{N_0W})$$

$$5.6\times10^{4}=4\times10^{3}\log(1+\frac{P}{5\times10^{-9}\times4\times10^{3}})$$

$$P=(2^{14}-1)\times2\times10^{-5}=3.2\times10^{-2}W$$

已知（7，3）线性分组码的生成矩阵 $G=\begin{bmatrix}1110100\\1100011\\0111010\end{bmatrix}$，

（1）求该码系统形生成矩阵和校验矩阵；

(2) 该码的最小码距是多少？其纠能错力？

（3）列出可纠差错图案和对应的伴随式

（2）若收到 R＝0010101，请判断是否发生误码，计算其发码；

解：（1）

$$G_s=\begin{bmatrix}1&0&0&1&1&1&0\\0&1&0&1&1&0&1\\0&0&1&0&1&1&1\end{bmatrix}$$

$$H=\begin{bmatrix}1&1&0&1&0&0&0\\1&1&1&0&1&0&0\\1&0&1&0&0&1&0\\0&1&1&0&0&0&1\end{bmatrix}$$

（2）$d_{min}=4$

$e_t=1$

（3）

S	E
0000	0000000
0001	0000001
0010	0000010
0100	0000100
1000	0001000
0111	0010000
1101	0100000
1110	1000000

（2）R_1H^T=[0010101] H^T=[0010]≠[0000]

发码为：0010101+0000010=0010111

已知 $n=7$ 的循环码 $g(x)=x^4+x^3+x^2+1$。

(1) 求 k 和对应的 $h(x)$;

(2) 如信息多项式 $m(x)=x^2+x$，求该系统码多项式;

(3) 求对应的系统形式的生成矩阵 G'，校验矩阵 H;

(4) 计算该码的最小码距，纠错能力 t=?

(5) 列出可纠差错图案和对应的伴随式;

(6) 若接收码字 R=1110101，求发码及信息位（针对系统形式生成矩阵）。

解：(1)k=3 $h(x)=x^3+x^2+1$

(2) $m(x)*x^4=x^6+x^5$

$r(x)=(x^6+x^5) \bmod g(x)=x^3+1$

$c(x)=x^6+x^5+x^3+1$

(3) $G=\begin{bmatrix}1&1&1&0&1&0&0\\0&1&1&1&0&1&0\\0&0&1&1&1&0&1\end{bmatrix}$

$G'=\begin{bmatrix}1&0&0&1&1&1&0\\0&1&0&0&1&1&1\\0&0&1&1&1&0&1\end{bmatrix}$ $H=\begin{bmatrix}1&0&1&1&0&0&0\\1&1&1&0&1&0&0\\1&1&0&0&0&1&0\\0&1&1&0&0&0&1\end{bmatrix}$

1110101

(4 $d_{min}=4$,t=1

48

(5)

S	E
0000	0000000
0001	0000001
0010	0000010
0100	0000100
1000	0001000
1101	0010000
0111	0100000
1110	1000000

(6) $S = RH^T = [0 \quad 0 \quad 0 \quad 1]$,

E=[0 0 0 0 0 0 1],

C=R+E=[1 1 1 0 1 0 0]

信息位 M=[111]

南京邮电大学 2007/2008 学年第二学期

《信息论与编码》期末试卷（B）答案

院(系)_____ 班级_____ 学号_____ 姓名_____

题号	一	二	三	四	五	六	七	八	九	十	总分
得分											

1. 在无失真的信源中，信源输出由 $H(X)$ 来度量；在有失真的信源中，信源输出由 $R(D)$ 来度量。

2. 要使通信系统做到传输信息有效、可靠和保密，必须首先 信源 编码，然后 加密 编码，再 信道 编码，最后送入信道。

3. 带限 AWGN 波形信道在平均功率受限条件下信道容量的基本公式，也就是有名的香农公式是 $C = W\log(1+SNR)$；当归一化信道容量 C/W 趋近于零时，也即信道完全丧失了通信能力，此时 E_b/N_0 为 -1.6 dB，我们将它称作香农限，是一切编码方式所能达到的理论极限。

4. 保密系统的密钥量越小，密钥熵 $H(K)$ 就越 小，其密文中含有的关于明文的信息量 $I(M;C)$ 就越 大。

5. 已知 $n=7$ 的循环码 $g(x) = x^4 + x^3 + x + 1$，则信息位长度 k 为 3，校验多项式 $h(x)=$ x^3+x+1。

6. 设输入符号表为 $X=\{0, 1\}$，输出符号表为 $Y=\{0, 1\}$。输入信号的概率分布为 $p=(1/2, 1/2)$，失真函数为 $d(0,0)=d(1,1)=0$，$d(0,1)=2$，$d(1,0)\neq 1$，则 $D_{min}=$ 0，$R(D_{min})$ = 1 bit/symbol，相应的编码器转移概率矩阵 $[p(y/x)]=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$；$D_{max}=$ 0.5，$R(D_{max})$ = 0，相应的编码器转移概率矩阵 $[p(y/x)]=\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}$。

7. 已知用户 A 的 RSA 公开密钥 $(e,n)=(3,55)$，$p=5, q=11$，则 $\phi(n)=$ 40，他的秘密密钥 $(d,n)=$ (27,55)。若用户 B 向用户 A 发送 $m=2$ 的加密消息，则该加密后的消息为 8。

二、判断题

1. 可以用克劳夫特不等式作为唯一可译码存在的判据。（√）
2. 线性码一定包含全零码。（√）
3. 算术编码是一种无失真的分组信源编码，其基本思想是将一定精度数值作为序列的编码，是以另外一种形式实现的最佳统计匹配编码。（×）
4. 某一信源，不管它是否输出符号，只要这些符号具有某些概率特性，就有信息量。（×）
5. 离散平稳有记忆信源符号序列的平均符号熵随着序列长度 L 的增大而增大。（×）
6. 限平均功率最大熵定理指出对于相关矩阵一定的随机矢量 X，当它是正态分布时具有最大熵。（√）
7. 循环码的码集中的任何一个码字的循环移位仍是码字。（√）

8. 信道容量是信道中能够传输的最小信息量。 （×）

9. 香农信源编码方法在进行编码时不需要预先计算每个码字的长度。 （×）

10. 在已知收码 R 的条件下找出可能性最大的发码 C_i 作为译码估计值，这种译码方法叫做最佳译码。 （√）

三、计算题

某系统（7，4）码

$c=(c_6\ c_5\ c_4\ c_3\ c_2\ c_1\ c_0)=(m_3\ m_2\ m_1\ m_0\ c_2\ c_1\ c_0)$ 其三位校验位与信息位的关系为：

$$\begin{cases} c_2=m_3+m_1+m_0 \\ c_1=m_3+m_2+m_1 \\ c_0=m_2+m_1+m_0 \end{cases}$$

（1）求对应的生成矩阵和校验矩阵；

（2）计算该码的最小距离；

（3）列出可纠差错图案和对应的伴随式；

（4）若接收码字 R=1110011，求发码。

解：1. $G=\left[\begin{array}{ccccccc} 1&0&0&0&1&1&0 \\ 0&1&0&0&0&1&1 \\ 0&0&1&0&1&1&1 \\ 0&0&0&1&1&0&1 \end{array}\right]$

$H^T=\left[\begin{array}{ccc} 1&1&1 \\ 1&1&0 \\ 1&0&1 \\ 0&1&1 \\ 1&0&0 \\ 0&1&0 \\ 0&0&1 \end{array}\right]$

2. $d_{min}=3$

3.

$S_2S_1S_0$	E
000	0000000
001	0000001
010	0000010
100	0000100
101	0001000
111	0010000
011	0100000
110	1000000

$$\begin{cases} S_2=e_6+e_4+e_3+e_2 \\ S_1=e_6+e_5+e_4+e_1 \\ S_0=e_5+e_4+e_3+e_0 \end{cases}$$

4. RH^T=[001]　接收出错

E=0000001　　$R+E=C$= 1110010 (发码)

四、计算题

已知(X,Y)的联合概率$p(x,y)$为：

X \ Y	0	1
0	1/3	1/3
1	0	1/3

求$H(X)$，$H(Y)$，$H(X,Y)$，$I(X;Y)$

解：$p(x=0)=2/3$　　$p(x=1)=1/3$

$p(y=0)=1/3$　　$p(y=1)=2/3$

$H(X)=H(Y)=H(1/3,2/3)=0.918$ bit/symbol

$H(X,Y)=H(1/3,1/3,1/3)=1.585$ bit/symbol

$I(X;Y)=H(X)+H(Y)-H(X,Y)=0.251$ bit/symbol

五、计算题

一阶齐次马尔可夫信源消息集$X\in\{a_1,a_2,a_3\}$，

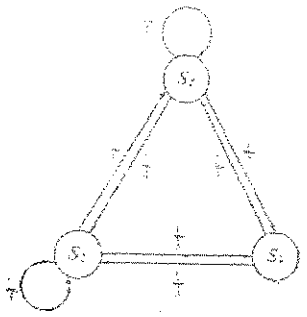
状态集$S\in\{S_1,S_2,S_3\}$，且令$S_i=a_i, i=1,2,3$，条件转移概率为

$$[P(a_j/S_i)]=\begin{bmatrix}1/4 & 1/4 & 1/2\\ 1/3 & 1/3 & 1/3\\ 2/3 & 1/3 & 0\end{bmatrix}$$

(1)画出该马氏链的状态转移图；

(2)计算信源的极限熵。

解：(1)



(2)
$$\begin{cases} \frac{1}{4}w_1+\frac{1}{3}w_2+\frac{2}{3}w_3=w_1 \\ \frac{1}{4}w_1+\frac{1}{3}w_2+\frac{1}{3}w_3=w_2 \\ \frac{1}{2}w_1+\frac{1}{3}w_2=w_3 \\ w_1+w_2+w_3=1 \end{cases} \rightarrow \begin{cases} w_1=0.4 \\ w_2=0.3 \\ w_3=0.3 \end{cases}$$

$H(X|S_1)=H(1/4,1/4,1/2)=1.5$ 比特/符号

$H(X|S_2)=H(1/3,1/3,1/3)=1.585$ 比特/符号

$H(X|S_3)=H(2/3,1/3)=0.918$ 比特/符号

$$H_\infty=\sum_{i=1}^{3}w_iH(X|S_i)=0.4\times1.5+0.3\times1.585+0.3\times0.918=1.351 \text{ 比特/符号}$$

六、计算题

若有一信源 $\begin{bmatrix} X \\ P \end{bmatrix}=\begin{bmatrix} x_1 & x_2 \\ 0.8 & 0.2 \end{bmatrix}$，每秒钟发出 2.55 个信源符号。

将此信源的输出符号送入某一个二元信道中进行传输

（假设信道是无噪无损的，容量为 1bit/二元符号），

而信道每秒钟只传送 2 个二元符号。

（1） 试问信源不通过编码（即 $x_1\rightarrow0, x_2\rightarrow1$ 在信道中传输）

（2） 能否直接与信道连接？

（3） 若通过适当编码能否在此信道中进行无失真传输？

（4） 试构造一种哈夫曼编码(两个符号一起编码)，

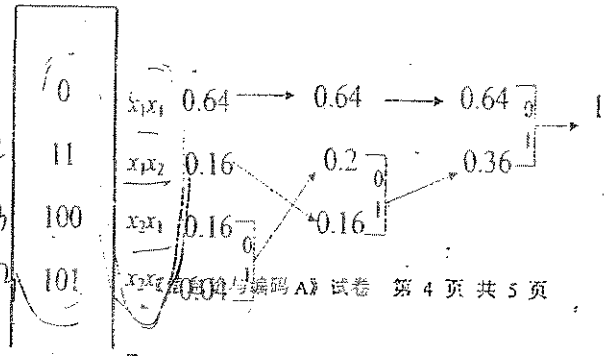
（5） 使该信源可以在此信道中无失真传输。

解：1.不能，此时信源符号通过 0，1 在信道中传输，2.55 二元符号/s>2 二元符号/s

2. 从信息率进行比较， $2.55*H(0.8,0.2)=1.84<1*2$

可以进行无失真传输

3.

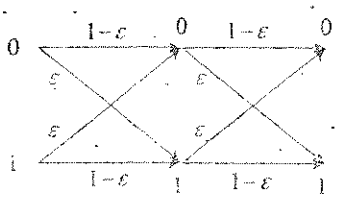


$$\overline{K}=\sum_{i=1}^{5}p_iK_i=0.64+0.16*2+0.2*3=1.56$$ 二元符号/2 个信源符号

此时 1.56/2*2.55=1.989 二元符号/s < 2 二元符号/s

七、计算题

两个 BSC 信道的级联如右图所示：



（1）写出信道转移矩阵；

（2）求这个信道的信道容量。

解：（1）

$$P=P_1P_2=\begin{bmatrix}1-\varepsilon & \varepsilon\\ \varepsilon & 1-\varepsilon\end{bmatrix}\begin{bmatrix}1-\varepsilon & \varepsilon\\ \varepsilon & 1-\varepsilon\end{bmatrix}=\begin{bmatrix}(1-\varepsilon)^2+\varepsilon^2 & 2\varepsilon(1-\varepsilon)\\ 2\varepsilon(1-\varepsilon) & (1-\varepsilon)^2+\varepsilon^2\end{bmatrix}$$

（2）$C=\log 2+H((1-\varepsilon)^2+\varepsilon^2)$

南京邮电大学 2006/2007 学年第二学期

《信息论与编码》期末试卷（B）答案

装订线内不要答题

自觉遵守考试规则，诚信考试，绝不作弊

院（系）__________ 班级__________ 学号__________ 姓名__________

题号	一	二	三	四	五	六	七	八	九	十	总分
得分											

得分

一、填空题：（20 分，每格 1 分）

1.汉明码是能纠正_1_个随机错误码元的完备码，它的码长 n 与监督码长度 m 间的关系为_$n=2^m-1$_。

2.在信息传输理论中，互信息是输入符号概率分布和转移概率分布的函数。当_$p(y|x)$_一定时，选择 $p(x)$ 可使互信息达到最大值，这就是_信道容量_；当_$p(x)$_一定时，选择 $p(y|x)$ 可使互信息达到最小值，这就是_率失真函数_。

3.一个平均功率受限的连续信道，其通频带为 2MHz，信道上存在高斯白噪声，已知信道上的信噪比为 8，则其信道容量为 6.33Mbps；若信道上的信噪比降至 4，则要达到相同的信道容量，其通频带应为 2.73MHz。

4.Shannon三大极限定理是：_无失真信源编码定理_、_信道编码定理_、_限失真信源编码定理_。

5.对于无限长的消息序列，(n_0, k_0, L) 卷积码的码率 $R=$_[illegible]_。

6.某线性分组码的最小码距为 5（$d_{min}=5$），该码若用于检错，能检出_4_位错码，若用于纠错，能纠正_1_位错误。若对该码扩展一位(采用偶校验)，此时最小码距为6，若对该码缩短一位，此时最小码距为5。

7.一个理想的保密系统，应确保合法用户的安全，保密地给出掩护明文的全部信息，即 $I(M;C,K)=$_$H(M)$_；而对窃听者截获之后不给任何明文信息，即 $I(M;C)=$_0_。

得分

二、判断题：（10 分，每题 1 分，对的写“√”，错的写“×”）

1.给定编码方法得到的码是唯一的。 （×）

2.线性码一定包含全零码。 （√）

3.对于一种编码方法，平均纠错能力依赖于最小码距而与译码规则没有依赖关系。 （×）

4.算术编码是一种无失真的分组信源编码，其基本思想是将一定精度数值作为序列的编码，是以另外一种形式实现的最佳统计匹配编码。 （×）

5.香农-费诺编码方法在进行编码时不需要预先计算每个码字的长度。 （×）

6.存在码字长度为(2,2,3,3,3,4,4)的二进制唯一可译码。 （√）

7.某一信源，不管它是否输出符号，只要这些符号具有某些概率特性，就有信息量。

（×）

8. 对称 DMC 信道，当信道输入符号等概分布时，信道输出符号也等概分布；反之，若信道输出符号等概分布，则信道输入符号的概率分布不定。

（×）

9. 信道中各码元是否出现差错，与其前后码元是否差错无关，每个码元独立地按一定的概率产生差错，称为随机差错。

（√）

10. 只要传信率 R 小于信道容量 C，总存在一种信道码，并以所要求的任意小的差错概率实现可靠的通信。

（√）

得分

三、计算（10 分）若采用 RSA 密码体制，用户 A 选择 $p=7, q=17$，公钥为(77,119)，用户 B 的公钥为（7，123），试计算用户 A 向用户 B 发送 $m=12$ 的签名保密消息。

解：对于用户 A

$\Phi(n)=(p-1)\times(q-1)=96$;

由 $(77\times d)\bmod 96=1$，可解得 $d=5$，

（4 分）

用户 A 用自己的私钥（5，119）计算签名：$S_1=12^5 \bmod 119=3$ （3 分）

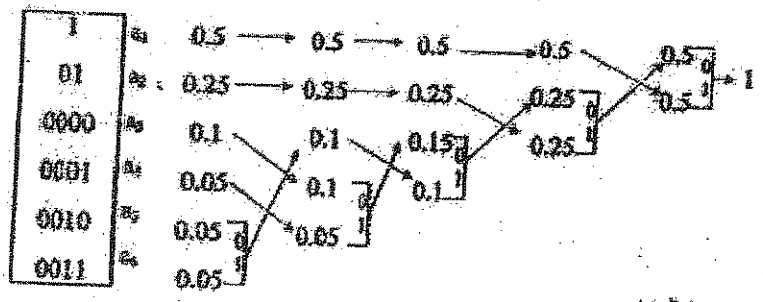
用户 A 用 B 的公钥（7，123）再次加密：$S=3^7 \bmod 123=96$ （3 分）

得分

四、计算（15 分）设有离散无记忆信源 $\begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ 0.5 & 0.25 & 0.1 & 0.05 & 0.05 & 0.05 \end{pmatrix}$。

1. 编出二进制符号的哈夫曼码，求其平均码长；
2. 编出三进制符号的哈夫曼码，求其平均码长；
3. 二进制信道中传输一个码字需要 1.8 元，三进制信道中传输一个码字需要 2.7 元。请问传输这个消息序列使用哪一种方案更划算。

解:1.

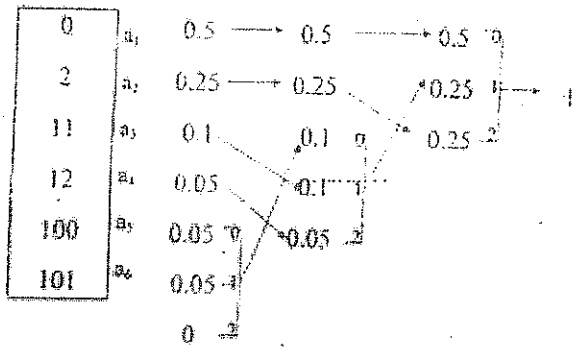


（5 分）

$\overline{K}=\sum_{i=1}^{6} p(a_i)K_i=2$

（1 分）

2.



（5 分）

$$\overline{K}=\sum_{i=1}^{6}p(a_i)K_i=1.35$$ （1 分）

3. 二进制 传输一个消息符号需 2*1.8=3.6 元

三进制 传输一个消息符号需 1.35*2.7=3.645 元

所以采用二进制信道更划算。 （3 分）

得分

五、计算（10 分）若有一信源 $\begin{bmatrix}X\\P\end{bmatrix}=\begin{bmatrix}a_1 & a_2\\0.5 & 0.5\end{bmatrix}$，每秒钟发出 2.62 个信源符号。将此信源的输出符号送入某二元无噪无损信道中进行传输，而信道每秒钟只传送 2 个二元符号。

1. 试问信源能否在此信道中进行无失真的传输，
2. 若此信源失真测度为汉明失真，可知率失真函数

$R(D)=\begin{cases}1-H(D,1-D) & 0\le D\le 0.5\\ 0 & D>0.5\end{cases}$，由编码定理，当传输速率 R_t 满足 $(R(D)bit/symbol\times 2.62symbol/s)\le R_t<C_t$ 时，在信道中传输后失真不超过失真限度 D，试问 D 至少应为多少？

解：

1. $H(X)=H(0.5)=1$ bit/symbol

每秒发出 2.62 个信源符号，

则信源输出 $H(X)*2.62=2.62$ bit/s （2 分）

而传输信道，二元无噪无损，

信道容量 $C=1$ bit/二元符号，

p 和 $H(p,1-p)$ 的对应表

p	$H(p,1-p)$
0.035	0.2189
0.037	0.2284
0.039	0.2377
0.041	0.2469
0.043	0.2559
0.045	0.2648

33

每秒传递 2 个二元符号，所以 Cr=2bit/s （2 分）

Cr<2.62，所以不能进行无失真的传输。 （2 分）

2. 当 $R(D)$<=Cr，信源在该信道中传输不会引入新的失真（1 分）

则 $2.62*(1-H(D))=2$

得 $H(D)\approx 0.2366$

易知 $D\approx 0.039$。 （3 分）

得 分

六、计算（20 分）已知 $n=7$ 的 $(7,\ k)$ 循环码 $g(x)=x^3+x+1$。

1. 如信息多项式 $m(x)=x^3+1$，求相应的码多项式；
2. 求对应的系统形式的生成矩阵 G'，校验矩阵 H；
3. 计算该码的最小距离，列出可纠差错图样和对应的伴随式；
4. 若接收码字 $R=1100100$，试问该收码是否有错？并译出发码估值 $\hat{C}$ 及对应的发送信息 $\hat{M}$（针对系统码）。

解：1. $c(x)=m(x)g(x)=x^6+x^3+x+1$ （4 分）

2. $G=\begin{bmatrix}1&0&1&1&0&0&0\\0&1&0&1&1&0&0\\0&0&1&0&1&1&0\\0&0&0&1&0&1&1\end{bmatrix}$ $G'=\begin{bmatrix}1&0&0&0&1&0&1\\0&1&0&0&1&1&1\\0&0&1&0&1&1&0\\0&0&0&1&0&1&1\end{bmatrix}$ （4 分）

$H=\begin{bmatrix}1&1&1&0&1&0&0\\0&1&1&1&0&1&0\\1&1&0&1&0&0&1\end{bmatrix}$ （2 分）

3. $d_{min}=3$ （2 分）

$\hat{M}$	$\hat{C}$
000	0000000
001	0000001
010	0000010
100	0000100
011	0001000
110	0010000
111	0100000
101	1000000

（4 分）

$S(x)$	$E(x)$
0	0
1	1
x	x
x^2	x^2
$x+1$	x^3
x^2+x	x^4
x^2+x+1	x^5
x^2+1	x^6

4. $2H^T=[110]$　接收出错（1分）

$E=0010000$　　$R+E=C=1110100$（发码）　　（2分）

$M=1110$　（1分）

……或 $R(x) \bmod g(x)=x^2+x=S(x)$　接收出错

$E(x)=x^4$　　$R(x)+E(x)=C(x)=x^6+x^5+x^4+x^2$

$m(x)=x^3+x^2+x$

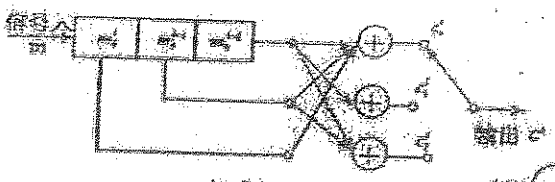
得分	

七、计算（15分）(3,1,2) 卷积码，已知转移函数矩阵是

$$G(D)=[1+D+D^2, D^2, D+D^2]$$

1. 画出该卷积码的编码器图；
2. 画出状态图；
3. 假如输入信息序列 $u=110$ 时，求输出码字 c，并在网格图上画出编码路径。

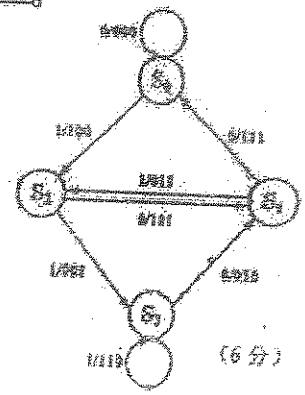
解：1、$G_0=(100)$　$G_1=(101)$　$G_2=(111)$



（4分）

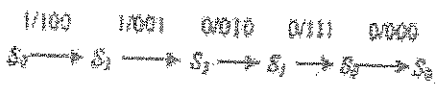
2.

	$m_2^{t-1}=0$		$m_2^{t-1}=1$	
S^t	S^{t+1}	C^t	S^{t+1}	C^t
$S_0(00)$	S_0	000	S_2	100
$S_1(01)$	S_0	111	S_2	011
$S_2(10)$	S_1	101	S_3	001
$S_3(11)$	S_1	010	S_3	110

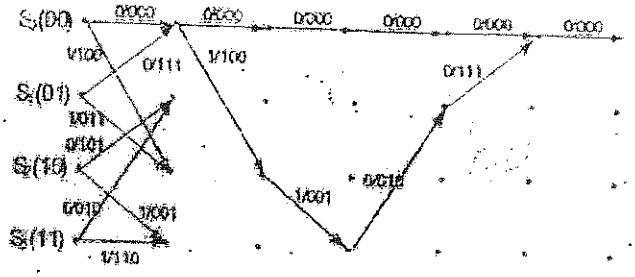


（6分）

4. 当 $u=110$ 时，输号流图如下



所以c=100，001，010，111，000　　（2分）　　（后两个没有写可酌情不扣分）



（3分）

南京邮电大学 2005/2006 学年 第二学期

期终 信息论与编码 试题(A)(答案)

班级_____ 姓名_____ 学号_____

题号	一	二	三	四	五	六	七	八	总分
得分									

一、填空题：(20 分)

1. 香农于 1948 年发表了《通讯的数学理论》，这是一篇关于现代信息论的开创性的权威论文，为信息论的创立作出了独特的贡献。
2. 冗余度来自两个方面，一是信源符号间的相关性，另一个方面是信源符号分布的不均匀性。
3. 信息率失真函数的物理意义是：对于给定信源，在平均失真不超过失真限度D的条件下，信息率容许压缩的最小值。
4. 码的纠、检错能力取决于码的最小距离。
5. 对于离散信源，无失真时R(0)= H(X)，对连续信源R(0)= ∞。
6. X为一个离散随机变量，Y是X的函数，记作Y=f(X)，则H(Y/X)= 0。
7. 为了使系统能够信源和发送最经济信道容量C，应使信道输入符号尽量均匀分布，信源尽量冗余。
8. 差错控制的途径：提高信道容量C、减小误率R、增加码长N。
9. 信源编码的主要目标是压缩每个信源符号的平均比特数或信源的码率，以提高通信系统的有效性。信道编码的主要目标是提高信息传输的可靠性。
10. x是一个离散随机变量，y1=exp(x)和y2=cosx，试确定H(y1)与H(y2)关系：H(y1)大于等于H(y2)。(小于、大于、等于、小于等于、大于等于)
11. 带限AWGN波形信道在平均功率受限条件下信道容量的基本公式，也就是有名的香农公式C=Wlog2(1+SNR) b/s；当归一化信道容量C/W趋近于零时，也即信道完全丧失了通信能力，此时E_b/N_0为 -1.6db，我们将它称作香农限，是一切编码方式所能达到的理论极限。
12. 为了使最大似然译码等同于最佳译码，必须使发码等概化和信道对称均值。

37

二、判断题：(10 分)（对的写“√”，错的写“×”）(每空 2 分)

1. 某一信源，不管它是否输出符号，只要这些符号具有某些概率特性，就有信息量。（×）

2. 对于 m 阶马尔可夫信源来说，在某一时刻出现的符号，取决于前面已出现的 m 个符号。（√）

3. 根据熵的可加性，序列的熵等于组成序列的各符号熵之和。（×）

4. 哈夫曼码在编码时充分考虑了信源符号分布的不均匀性和符号之间的相关性。（×）

5. 要使信息率小于 R(D)，平均失真一定会超过失真限度 D。（×）

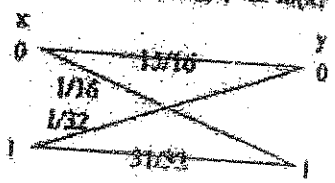
三、(6 分)设输入输出符号表为 X＝Y＝{0，1}，输入概率分布 p(x)={1/4，3/4}，失真矩阵为 $d=\begin{bmatrix} d(x_1,y_1) & d(x_1,y_2) \\ d(x_2,y_1) & d(x_2,y_2) \end{bmatrix}=\begin{bmatrix} 1/5 & ; \\ 1 & 1/4 \end{bmatrix}$，求 D_{max}。

解：$D_{max}=\min_j \sum_i p_i d_{ij}$

$$=\min\left\{\frac{1}{4}\times\frac{1}{5}+\frac{3}{4}\times 1,\ \frac{1}{4}\times 1+\frac{3}{4}\times\frac{1}{4}\right\}$$

$$=\frac{7}{16}=0.4375$$

四、(10 分)一个信源以相等的概率发 1000 码元/秒的速率把“0”和“1”码送入有噪声信道，由于信道中噪声的影响，发送为“0”接收为“1”的概率是 1/16，而发送为“1”接收为“0”的概率是 1/32，求收信者接收的信息率。(提示：信息率 R=I(x；y))



解：H(x)=-Σp(x)log₂(x)=1bit/s　(2 分)

利用贝叶斯公式得，

p(x=0/y=0)=p(x=0)p(y=0/x=0)/p(y=0)=p(x=0)p(y=0/x=0)/[p(x=0)p(y=0/x=0)+p(x=1)p(y=0/x=1)]

=1/2*(15/16)/[1/2*(15/16)+1/2*(1/32)]=30/31

同理：p(x=1/y=0)=1/32　　p(x=1/y=1)=31/33　　p(x=0/y=1)=2/33　　(4 分)

所以 H(x/y)=-Σ Σp(x,y)log p(x/y)

=-15/32log 0/31 -1/32 log2/33 -31/64 log 31/33 - 1/64 log 1/31

=0.269bit/s　　(2 分)

因此，编速率 R=α[H(x)-H(x/y)]=1000*(1-0.269)=731bit/s　　(2 分)

五、(9 分)已知一个信道的信道矩阵为 $\begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{bmatrix}$，传输一个符号所需的时间为 1 毫秒，求该信道能通过的最大速率。

解：C=H（Y）-H（Y/X）=1.585-（0.3+0.464+0.521）=1.585-1.485=0.1bit/sym

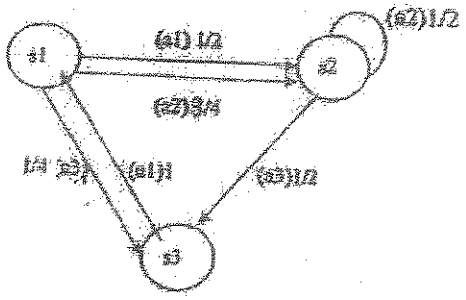
0.1×1000=100bit/s

六、(20 分)有号求可无信源 x∈{a1,a2,a3}，其状态转移概率如图。

(1)　求 H(x/s1),H(x/s2)和 H(x/s3)；

(2)　列各符号态，分别对 x 的符号编成最佳变长二进码；

(3)　求 H(x)。



解：(1) 状态转移矩阵

$$P=\left[P(s_j/s_i)\right]=\begin{bmatrix} 0 & 3/4 & 1/4 \\ 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \end{bmatrix}$$　　(1 分)

符号条件概率矩阵 $\left[P(a_i/s_j)\right]=\begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \end{bmatrix}$　(1 分)

H(x/s1)=-ΣP(ai/s1)logP(ai/s1)=3/2bit/symbol

H(x/s2)=1/2log2 +1/2log2=1bit/symbol

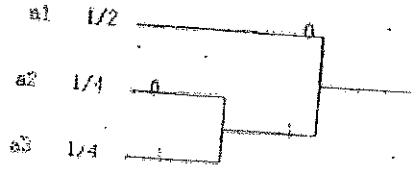
H(x/s3)=1log1=0bit/symbol

(3 分)

(2)

用哈夫曼编码

对状态 s1：

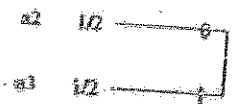


0

10

11 (3 分)

对状态 s2



不编码 (2 分)

对状态 s3

a1 0

a2 不编码

a3 不编码 (2 分)

(3) P(s1)=ΣP(sj)P(s1/sj)=P(s3)

P(s2)=3/4*P(s1) +1/2*P(s2)

P(s3)=1/4*P(s1)+1/2*P(s2)

ΣPi=1

(1 分)

推出：P(s1)=P(s3)=2/7

P(s2)=3/7

(3 分)

$H_\infty(X)$=-ΣP(si)H(X/si)=0.857bit/symbol

(2 分)

平均长度 k=2/7*(1/2*1 + 1/4*2 +1/4 2) +3/7(1/2*1+1/2*1)+1*1 +2/7=8/7 (2 分)

七、(10 分)在公开密钥密码算法 RSA 中，设令 p=5,q=11,e=27,试计算解密密钥 d，并加密 M=3。

解：$n=p\times q=55$ (1 分)

$\phi(n)=(p-1)\times(q-1)=4\times10=40$ (3 分)

$e\times d \bmod 40=1$，则 $d=3$ 。 (3 分)

$3^{27}\bmod 55=42$ (3 分)

八、(15 分)有两个同时输出消息的信源 X 和 Y，第一个信源 X 能输出 a,b,c 三个消息；第二个信源 Y 能输出 d,e,f,g 四个消息。信源 X 各消息出现的概率为 p(x)。信源 Y 各消息出现的条件概率为 p(y/x)。p(x)和 p(y/x)的具体数值如表所示。求联合信源的熵、条件熵及 H(x)、H(y)和 max H(xy)。

X		a	b	c
p(x)		1/2	1/3	1/6
p(y/x)	d	1/4	3/10	1/6
	e	1/4	1/5	1/2
	f	1/4	1/5	1/6
	g	1/4	3/10	1/6

解：信源 x，y 输出的每一对消息出现的联合概率为 $p(xy)=p(x)*p(y/x)$

计算结果如表所示

p(xy)	A	b	c
d	1/8	1/10	1/36
e	1/8	1/15	1/12
f	1/8	1/15	1/36
g	1/8	1/10	1/36

所以 $H(xy)=-\sum\sum p(xy)\log p(xy)$

$=-(3*1/8\log1/8+2*1/10\log1/10+2*1/15\log1/15+1/12\log1/12+3*1/36\log1/36$

$=3.417bit/symbol$ (3 分)

信源 x 本身的熵 $H(x)=-\sum p(x)\log p(x)=-(1/2\log1/2+1/3\log1/3+1/6\log1/6)$

$=1.461bit/symbol$ (3 分)

信源 y 的条件熵 $H(y/x)=-\sum\sum p(xy)\log p(y/x)$

$=-(4*1/8\log1/4+2*1/10\log3/10+2*1/15\log1/5+1/12\log1/2+3*1/36\log1/6$

$=1.955bit/symbol$ (3 分)

41

信源 y 输出符号 d 的概率 $p(d)=\sum p(xd)=\sum p(x)p(d/x)$

$=p(a)p(d/a)+p(b)p(d/b)+p(c)p(d/c)$

$=1/2*1/4+1/3*3/10+1/6*1/6=91/360$

同理可求出 $p(e)=1/8+1/15+1/12=33/120$

$p(f)=1/8+1/15+1/36=79/360$

$p(g)=1/8+1/10+1/36=91/360$

因此 $H(y)=-\sum p(y)\log p(y)$

$=-(91/360\log 91/360+33/120\log 33/120+79/360\log 79/360+91/360\log 91/360)$

$=1.997bit/symbol$

(3 分)

$H(xy)_{max}=H(x)+H(y)=3.458bit/symbol$

(3 分)

南京邮电大学 2004 / 2005 学年 第 学期

期 终《信息论与编码》试题(B)答案

学号_____ 姓名_____ 得分_____

题号	一	二	三	四	五	六	七
得分	20	10	10	15	17	13	15

一、填空题（每空 1 分，共 20 分）：

1. 在异前缀码中，任一种代表某信源符号的码字，其前缀部分都 不是 其它码字的前缀部分。

2. 码的纠、检错能力取决于 码字的最小距离（d_{min}） 。

3. 信源编码的目的是 减少冗余，提高编码效率 ，信道编码的目的是 提高信息传递的可靠性 。

4. 把信息组原封不动的搬到码字前k位的（n，k）码就叫做 系统码。

5. 卷积码与分组码的主要差异在于 分组码的各个分组之间没有任何联系，而卷积码在某一时刻的编码输出取决于本时刻及以前的L个分组。

6. 如果一线性分组码的最小距离为d_{min}，则其扩展码的最小距离是偶校验时， $d_{min}+1$（d_{min}是奇数） ，其缩短码的最小距离是 d_{min} 。

7. 为了使最大似然译码等同于最佳译码，必须使 构成码集的 2^r 个码字以相同的概率发送 和 $P(r)$ 对于任何 r 都有相同的值 。

8. 循环码中的任何一个码字的循环移位仍是码字，因此用一个基底就足以表示一个码的特征，所以描述循环码的常用数学工具是 码多项式 。

9. （4，2，3）卷积码指该编码器的 3 个输出不仅与此时段的 2 个输入有关，而且也与前 3 个输入（记忆器件中存储的）有关。

10. 香农信息论中的三大极限定理是： 无失真信源编码定理 、信道编码定理 、 限失真信源编码定理 。

11. 根据密钥的设置方法不同，密码学中最具代表性的两种密码算法是 数据加密标准（DSE） 和 公开密钥密码 。

12. 限失真信源编码定理与无失真信源编码定理的主要差别是 信源编码器输出的信息是否有失真 。

二、判断题：（对的写“√”，错的写“×”，错误题请说明理由）（判断每题 1 分，错误题说明每题 1 分，共 10 分）

1. 一般情况下，用变长编码要求的信源符号长度 L 比定长编码大得多。

（ × ）

用变长编码来达到相当高的编码效率，一般所要求的符号长度 L 可以比定长编码小很多

2. 由于构成同一空间的基底不是唯一的，所以不同的基底或生成矩阵有可能生成同一码集。（ √ ）

3. 可以用生成多项式来描述任何线性分组码。（ × ）

可以用生成多项式来描述循环码

4. 一个密码体制的安全性，既依赖于其密钥的保密性，又依赖于其加密、解密算法的保密性。（ × ）

一个密码体制的安全性，必须只依赖于密钥的保密性，而不依赖于其加密、解密算法的保密性。

5. 只要传信率R大于信道容量C，总存在一种信道码（及解码器），可以以所要求的任意小的差错概率实现可靠的通信。（ × ）

只要传信率R小于信道容量C，总存在一种信道码（及解码器），可以以所要求的任意小的差错概率实现可靠的通信。

6. 各码字的长度n_i应符合克劳夫特不等式，是唯一可译码存在的充分和必要条件。（ √ ）

三（10分）. 设某二元码为 C={11100，01001，10010，00111}

（1）（3分）求此码的最小距离d_{min}；

（2）（4分）采用最小距离译码准则，试问接收序列10000，01100和00100应译成什么码字？

（3）（3分）此码能纠正几位码元的错误？

45

解：（1）用穷举法可知，此码的最小距离d_{min}为 3

（2）若采用最小距离译码准则，接收序列 10000 应译为 10010，01100 应译为 11100，00100 应译为 11100 或 00111

（3）由（2）可知，此码只能纠正一位码元的错误

为 15.（6 分）在公开密钥密码算法 RSA 中，设 p＝5，q＝11，e＝27，求 d；（3 分）并将明文 23，5，33，25 加密为一英文单词（1＝A，2＝B，…，26＝Z）。（12 分）

解：n=p*q=5*11=55

$\Phi(n)=(p-1)\times(q-1)=(5-1)*(11-1)=4*10=40$

由$[e\times d]\bmod \Phi(n)]=1$，即$(27d)\bmod 40=1$，可得 d=3

根据公式$x=y^d \pmod n$，当密文为 23，5，33，25 时，

$x_{23}=23^d \pmod n=23^3 \bmod 55=12$

$x_5=5^d \pmod n=5^3 \bmod 55=15$

$x_{33}=33^d \pmod n=33^3 \bmod 55=22$

$x_{25}=25^d \pmod n=25^3 \bmod 55=5$

由对应法则（1＝A，2＝B，…，26＝Z）可知，12=L，15=O，22=V，5=E，

所以明文为 LOVE

五(17 分)、已知一信源包含 8 个消息符号，其出现的概率为 P(X)={0.1，0.18，0.4，0.05，0.06，0.1，0.07，0.04}。

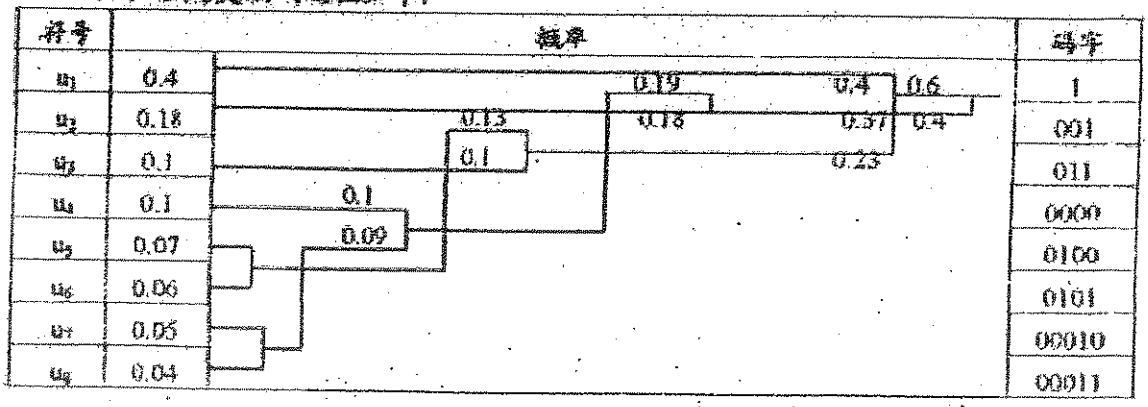
(1) 该信源在每秒钟内发出 1 个符号，求该信源的熵(2 分)及所需的信息传输速率(1 分)。

(2) 对该 8 个符号作哈夫曼编码，写出相应码字(8 分)（平均码长 2 分），并求出编码效率(2 分)，此时所需的信息传输速率是多少(2 分)?

解：（1）该信源的熵$H(X)=-\sum p(xi)\log_2 p(xi)=2.55$(bit/符号)

由于该信源在每秒钟内发出 1 个符号，所以其信息传输速率为 2.55 bps

（2）哈夫曼编码过程如下：



所以相应的码字为：1、001、011、0000、0100、0101、00110、00111

平均码长：

$\bar{k} = \sum p(x_i)k_i = 0.4 \times 1 + 0.18 \times 3 + 0.1 \times 3 + 0.1 \times 4 + 0.07 \times 4 + 0.06 \times 4 + 0.05 \times 5 + 0.04 \times 5 = 2.61$

编码效率：$R = H(x) / \bar{k} = 2.55 / 2.61 = 97.7\%$

此时所需的信息传输速率为 2.61 bps

六、(13 分)、多项式 $g(x) = x^4 + x^2 + x + 1$ 是一个 (7，3) 循环汉明码的生成多项式

(1) (3 分)计算其对偶码的生成多项式

(2) (4 分)用生成多项式 g(x)编出消息 (110) 的系统循环码字。

(3) (2 分)若接收码多项式为 $R(x) = x^6 + x^5 + 1$，判断是否为许用码多项式。

(4) (4 分)写出由此构造 (6，2) 缩短码的生成矩阵和校验矩阵。

解：(1) 其对偶码的生成多项式为 $h(x) = (x^7 + 1)/(x^4 + x^2 + x + 1) = x^3 + x + 1$

(2) 由于 m=110，所以 $m(x) = x^2 + x$

$x^{n-k}m(x) = x^4(x^2 + x) = x^6 + x^5$

由 $x^{n-k}m(x)$ 除以 g(x)，即 $(x^6 + x^5)/(x^4 + x^3 + x + 1)$ 可得余式 $r(x) = x^3 + 1$

所以 $c(x) = x^{n-k}m(x) + r(x) = x^6 + x^5 + x^3 + 1$

因此对应的系统循环码字为 (1100101)

(3) 由于 $R(x)h(x) \bmod (x^7 + 1) = (x^6 + x^5 + 1)(x^3 + x + 1) \bmod (x^7 + 1) = x^5 + x^3 + x^2 \neq 0$

所以 $R(x) = x^6 + x^5 + 1$ 不是许用码多项式

(4) 由生成多项式 $g(x) = x^4 + x^2 + x + 1$，校验多项式 $h(x) = x^3 + x + 1$ 可以写出 (7，3) 循环汉明码的

生成矩阵

$$G = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

校验矩阵

$$H = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

化为系统形式：

$$G=\left[\begin{array}{ccc:cccc}1&0&0&1&0&1&1\\0&1&0&1&1&1&0\\0&0&1&0&1&1&1\end{array}\right]$$

$$H=\left[\begin{array}{ccc:cccc}1&1&0&1&0&0&0\\0&1&1&0&1&0&0\\1&1&1&0&0&1&0\\1&0&1&0&0&0&1\end{array}\right]$$

则(6，2)新短码的生成矩阵为

$$G'=\left[\begin{array}{cccccc}1&0&1&1&1&0\\0&1&0&1&1&1\end{array}\right]$$（删去原来系统生成矩阵的第一行和第一列）

校验矩阵为

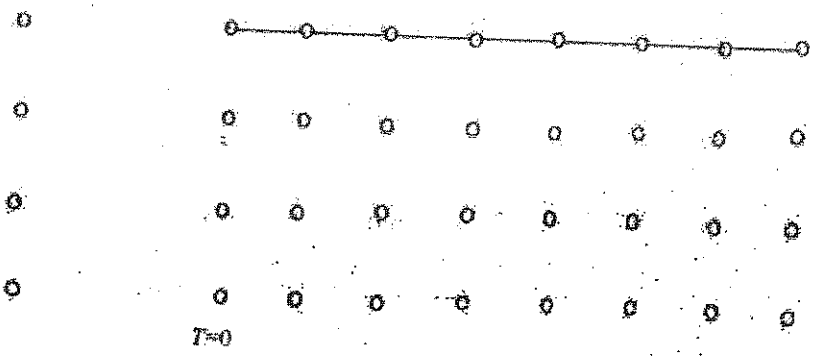
$$H'=\left[\begin{array}{cccccc}1&0&1&0&0&0\\1&1&0&1&0&0\\1&1&0&0&1&0\\0&1&0&0&0&1\end{array}\right]$$（删去原来系统校验矩阵的第一列）

七(15 分)、(3，1，2)卷积码，已知转移函数矩阵是$G(D)=[1+D+D^2,\ 1+D^2,\ D+D^2]$。

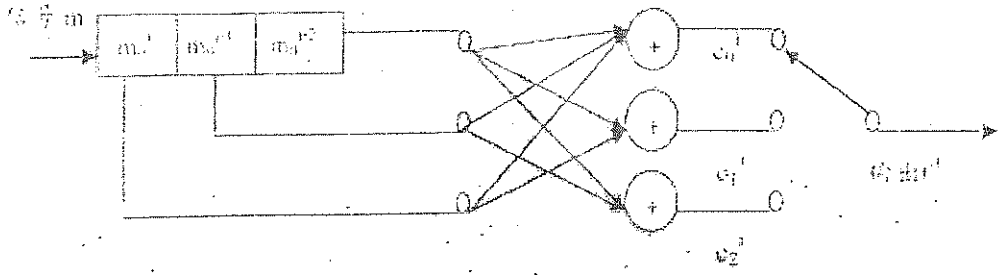
(1) (3 分)画出该卷积码编码器图；

(2) (6 分)画出状态图和网格图；

(3) (6 分)当 u=1110100 时，求 c，并画出编码路径。



解：（1）该卷积码的编码器图为



（2）首先列出编码器各种可能情况的汇总表

编码器状态的定义

状态	$m_0^{j-1}m_0^{j-2}$
S0	00
S1	01
S2	10
S3	11

不同状态与输入时输出的码字

状态 \ 输入	$m_0^j=0$	$m_0^j=1$
S0	000	110
S1	111	001
S2	101	011
S3	010	100

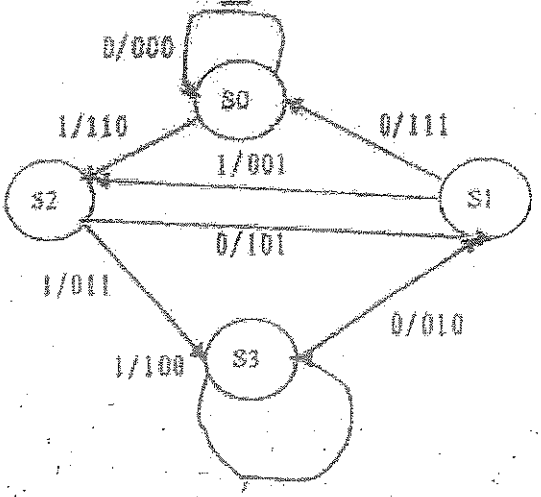
不同状态 Si 与输入时的下一状态 Si+1

状态 \ 输入	$m_0^j=0$	$m_0^j=1$
S0	S0	S2
S1	S0	S2
S2	S1	S3
S3	S1	S3

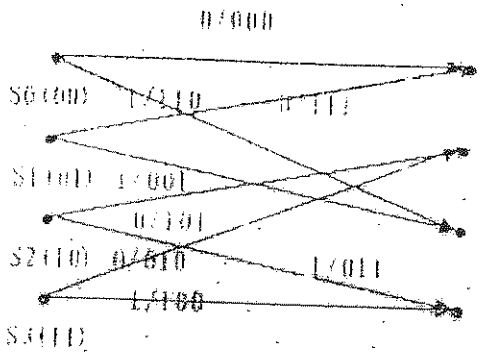
然后列出编码矩阵

$$C=\begin{array}{c} \\ S0 \\ S1 \\ S2 \\ S3 \end{array}\begin{array}{c} \begin{array}{cccc} S0 & S1 & S2 & S3 \end{array} \\ \begin{bmatrix} 000 & & 110 & \\ 111 & & 001 & \\ & 101 & & 011 \\ & 010 & & 100 \end{bmatrix} \end{array}$$

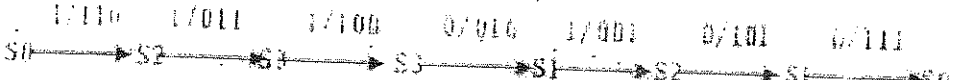
最后，画出状态图



网络图

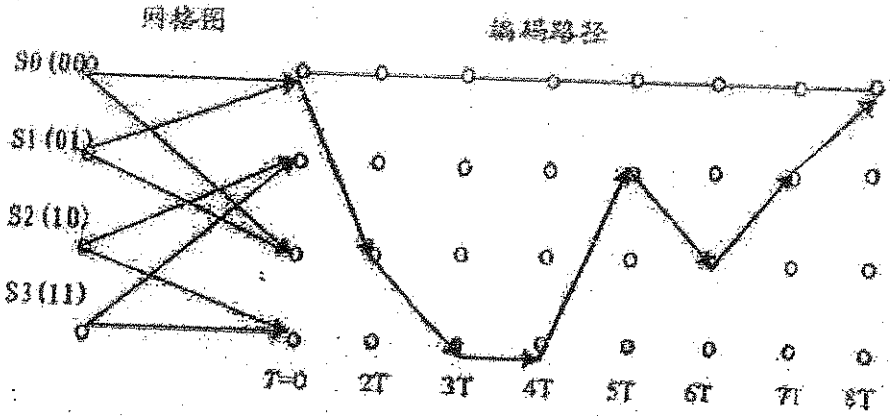


(3) 令 u=1101000 时，状态流图如下



所以 c=110，011，100，010，001，101，111

编码轨迹如下：



《信息论与编码理论》试卷

学生姓名	学号	班级	成绩

一、填空题（每空 1 分，共 30 分）

（1） 在现代通信系统中，信源编码主要用于解决信息传输中的 有效 性，信道编码主要用于解决信息传输中的 可靠 性 ，加密编码主要用于解决信息传输中的 安全 性。

（2） 不可能事件的自信息量是 ∞ ，必然事件的自信息是 0 。

（3） 离散平稳无记忆信源 X 的 N 次扩展信源的熵等于离散信源 X 的熵的 N 倍 。

（4） 在信息处理中，随着处理级数的增加，输入和输出消息之间的平均互信息量会 减少 。

（5） 若一离散无记忆信源的信源熵 H（X）等于 2.5，对信源进行等长的无失真二进制编码，则编码长度至少为 3 。

（6） 假设每个消息的发出都是等概率的，四进制脉冲所含信息量是二进制脉冲的 2 倍。

（7） 对于香农编码、费诺编码和霍夫曼编码，编码方法惟一的是 香农 编码。 霍夫曼 编码方法构造的是最佳码。

（8） 已知某线性分组码的最小汉明距离为 3，那么这组码最多能检测出 2 个码元错误，最多能纠正 1 个码元错误。

（9） 设有一个离散无记忆平稳信道，其信道容量为 C，只要待传送的信息传输率 R 小于 C（大于、小于或者等于），则存在一种编码，当输入序列长度 n 足够大，使译码错误概率任意小。

（10）平均错误概率不仅与信道本身的 统计 特性有关，还与 译码 规则和 编码 方法有关

（11）互信息 I(X;Y)与信息熵 H(Y)的关系为：I(X;Y) 小于 （大于、小于或者等于）H(Y)。

（12）克劳夫特不等式是唯一可译码 存在 的充要条件。{00，01，10，11}是否是唯一可译码？ 是 。

（13）差错控制的基本方式大致可以分为 前向纠错 、 反馈重发 和 混合纠错 。

（14） 如果所有码字都配置在二进制码树的叶节点，则该码字为 唯一可译 码 。

（15） 设信道输入端的熵为 H(X)，输出端的熵为 H(Y)，该信道为无噪有损信道，则该信道的容量为 Max H (Y)。

（16）某离散无记忆信源 X，其符号个数为 n，则当信源符号呈 等概 分布情况下，信源熵取最大值 log (n) 。

（17）平均互信息是输入信源概率分布的 上凸 函数；平均互信息是信道转移概率的 下凸 函数，平均互信息的最大值为 信道容量 。

二、简答题（共 4 题，每题 5 分）

1，简述离散信源和连续信源的最大熵定理。

1.答：离散无记忆信源，等概率分布时熵最大。连续信源，峰值功率受限时，均匀分布的熵最大。

平均功率受限时，高斯分布的熵最大。均值受限时，指数分布的熵最大。

2, 简述信源的符号之间的依赖与信源冗余度的关系。

当信源的符号之间有依赖时，信源输出消息的不确定性减弱。而信源冗余度正是反映信源符号依赖关系的强弱，冗余度越大，依赖关系就越大。

3, 简述香农第一编码定理的物理意义?

1.答：无失真信源编码，编码后尽可能等概率分布，使每个码元平均信息量最大。从而使信道信息传输率 R 达到信道容量 C，实现信源与信道理想的统计匹配。

4, 什么是最小码距，以及它和检错纠错能力之间的关系。

某一码书 C 中，任意两个码字之间汉明距离的最小值称为该码的最小码距 Dmin. 当已知某线性分组码的最小汉明距离为 Dmin，那么这组码最多能检测出 e =Dmin-1 个码元错误，最多能纠正 t =(Dmin-1) /2 个码元错误。

三、计算题（共 6 题，每题 10 分）

1, 有两个二元随机变量 X 和 Y，它们的联合概率为

Y \ X	$x_1=0$	$x_2=1$
$y_1=0$	1/8	3/8
$y_2=1$	3/8	1/8

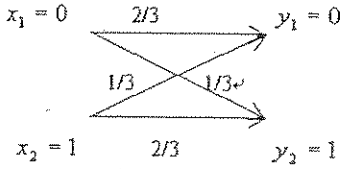
定义另一随机变量 $Z = XY$（一般乘积），试计算 $H(Z)=?$

解：$Z = XY$ 的概率分布如下：

$$\begin{bmatrix} Z \\ P(Z) \end{bmatrix} = \begin{Bmatrix} z_1=0 & z_2=1 \\ \frac{7}{8} & \frac{1}{8} \end{Bmatrix}$$

$$H(Z) = -\sum_{k}^{2} p(z_k) = -\left(\frac{7}{8}\log\frac{7}{8} + \frac{1}{8}\log\frac{1}{8}\right) = 0.544 \ bit/symbol$$

2, 二元对称信道如图。



1）若 $p(0)=\frac{3}{4}$，$p(1)=\frac{1}{4}$，求 $H(X)$、$H(X|Y)$ 和 $I(X;Y)$；

2）求该信道的信道容量。

解：1）共6分

$$H(X) = 0.8113 \ bit/符号$$

$$H(X|Y) = 0.749 bit/符号$$

$$I(X;Y) = 0.0616 \ bit/符号$$

2），$C = 0.082 \ bit/符号$ （3分）此时输入概率分布为等概率分布。（1分）

3, 求以下二个信道的信道容量：

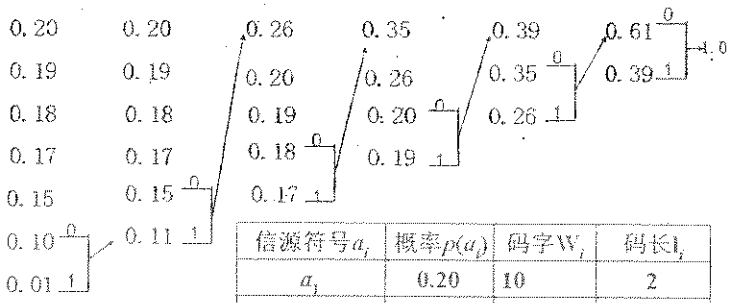
$$P_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

3答：P1为一一对应确定信道，因此有 $C_1=\max H(X)=\log_2 4=2$ *bit/*符号。

P2为具有归并性能的信道，因此有 $C_2=\max H(Y)=\log_2 3=1.5995$ *bit/*符号。

4, 信源空间为

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \\ 0.2 & 0.19 & 0.18 & 0.17 & 0.15 & 0.1 & 0.01 \end{bmatrix}$$，试构造二元霍夫曼码，计算其平均码长和编码效率（要求有编码过程）。



信源符号a_i	概率$p(a_i)$	码字W_i	码长l_i
a_1	0.20	10	2
a_2	0.19	11	2
a_3	0.18	000	3
a_4	0.17	001	3
a_5	0.15	010	3
a_6	0.10	0110	4
a_7	0.01	0111	4

$$\bar{L}=\sum_{i=1}^{7} p(a_i)l_i = 2.72$$

码元/符号

$$R=\frac{H(X)}{\bar{L}}=\frac{2.61}{2.72}=0.96$$

比特/符号

5, 已知一个高斯信道，输入信噪比(比率)为3，频带为3kHz，求最大可能传送的信息率。若信噪比提高到15，理论上传送同样的信息率所需的频带为多少？

5答:(1) 最大可能传送的信息率是

$Ct = w \log (1+ Px/Pn) = 3\times1000 \times \log (1+ 3) = 6\times1000$比特/秒

(2) 1.5kHZ

6, 设一线性分组码具有一致监督矩阵 $H=\begin{bmatrix}0&0&0&1&1&1\\0&1&1&0&0&1\\1&0&1&0&1&1\end{bmatrix}$

1）求此分组码 n=?,k=?共有多少码字？

2）求此分组码的生成矩阵 G。

3）写出此分组码的所有码字。

4）若接收到码字（101001），求出伴随式并给出翻译结果。

解：1）n=6,k=3,共有8个码字。（2分）

2）设码字 $\vec{C}=(C_5C_4C_3C_2C_1C_0)$ 由 $HC^T=0^T$ 得

$$\begin{cases} C_2\oplus C_1\oplus C_0=0\\ C_4\oplus C_3\oplus C_0=0\\ C_5\oplus C_3\oplus C_1\oplus C_0=0\end{cases}$$

令监督位为 $(C_2C_1C_0)$，则有

$$\begin{cases}C_2=C_5\oplus C_3\\ C_1=C_5\oplus C_4\\ C_0=C_4\oplus C_3\end{cases}$$

生成矩阵为 $\begin{bmatrix}1&0&0&1&1&0\\0&1&0&0&1&1\\0&0&1&1&0&1\end{bmatrix}$ （3分）

3）所有码字为000000，001101，010011，011110，100110，101011，110101，111000。（3分）

4）由 $S^T=HR^T$ 得

$S=(101)$，该码字在第5位发生错误，（101001）纠正为（101011），即译码为（101001）（2分）

一、概念简答题（每题5分，共40分）

1. 什么是平均自信息量与平均互信息，比较一下这两个概念的异同？

2. 简述最大离散熵定理。对于一个有m个符号的离散信源，其最大熵是多少？

3. 解释信息传输率、信道容量、最佳输入分布的概念，说明平均互信息与信源的概率分布、信道的传递概率间分别是什么关系？

4. 对于一个一般的通信系统，试给出其系统模型框图，并结合此图，解释数据处理定理。

5. 写出香农公式，并说明其物理意义。当信道带宽为5000Hz，信噪比为30dB时求信道容量。

6. 解释无失真变长信源编码定理。

7. 解释有噪信道编码定理。

8. 什么是保真度准则？对二元信源 $\begin{bmatrix} u \\ P(u) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \omega & 1-\omega \end{bmatrix}$，其失真矩阵 $D = \begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix}$，求a>0时率失真函数的 $D_{\min}$ 和 $D_{\max}$？

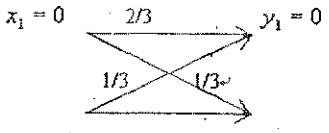
二、综合题（每题10分，共60分）

1. 黑白气象传真图的消息只有黑色和白色两种，求：

1） 黑色出现的概率为0.3，白色出现的概率为0.7。给出这个只有两个符号的信源X的数学模型。假设图上黑白消息出现前后没有关联，求熵 $H(X)$；

2） 假设黑白消息出现前后有关联，其依赖关系为：$P(\text{白}/\text{白}) = 0.9$，$P(\text{黑}/\text{白}) = 0.1$，$P(\text{白}/\text{黑}) = 0.2$，$P(\text{黑}/\text{黑}) = 0.8$，求其熵 $H_2(X)$；

2. 二元对称信道如图。



1）若 $P(0)=\frac{3}{4}$，$P(1)=\frac{1}{4}$，求 $H(X)$ 和 $I(X;Y)$；

2）求该信道的信道容量和最佳输入分布。

3. 信源空间为 $\begin{bmatrix} S \\ P(s) \end{bmatrix}=\begin{bmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ 0.4 & 0.2 & 0.1 & 0.1 & 0.05 & 0.05 & 0.05 & 0.05 \end{bmatrix}$，试分别构造二元和三元霍夫曼码，计算其平均码长和编码效率。

4. 设有一离散信道，其信道传递矩阵为 $\begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$，并设 $\begin{cases} P(x_1)=\frac{1}{4} \\ P(x_2)=\frac{1}{2} \\ P(x_3)=\frac{1}{4} \end{cases}$，试分别按最小错误概率准则与最大似然译码准则确定译码规则，并计算相应的平均错误概率。

5. 已知一（8，5）线性分组码的生成矩阵为 $\begin{bmatrix} 10000111 \\ 01000100 \\ 00100010 \\ 00010001 \\ 00001111 \end{bmatrix}$。

求：1）输入为全 00011 和 10100 时该码的码字；2）最小码距。

6. 设某一信号的信息传输率为 5.6kbit/s，在带宽为 4kHz 的高斯信道中传输，噪声功率谱 N0=5×10－6mw/Hz。试求：

（1）无差错传输需要的最小输入功率是多少？

（2）此时输入信号的最大连续熵是多少？写出对应的输入概率密度函数的形式。

……

一、 概念简答题（每题 5 分，共 40 分）

1. 答：平均自信息为 $$H(X) = -\sum_{i=1}^{n} p(x_i)\log p(x_i)$$

表示信源的平均不确定度，也表示平均每个信源消息所提供的信息量。

平均互信息 $$I(X;Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i y_j)\log \frac{p(x_i / y_j)}{p(x_i)}$$

表示从 Y 获得的关于每个 X 的平均信息量，也表示发 X 前后 Y 的平均不确定性减少的量，还表示通信前后整个系统不确定性减少的量。

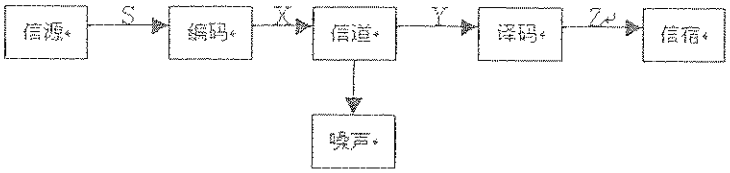
2. 答：最大离散熵定理为：离散无记忆信源，等概率分布时熵最大。

最大熵值为 $H_{\max} = \log_2 m$。

3. 答：信息传输率 R 指信道中平均每个符号所能传送的信息量。信道容量是一个信道所能达到的最大信息传输率。信息传输率达到信道容量时所对应的输入概率分布称为最佳输入概率分布。

平均互信息是信源概率分布的 ∩ 型凸函数，是信道传递概率的 U 型凸函数。

4. 答：通信系统模型如下：



数据处理定理为：串联信道的输入输出 X、Y、Z 组成一个马尔可夫链，且有 $I(X;Z) \le I(X;Y)$，$I(X;Z) \le I(Y;Z)$。说明经数据处理后，一般只会增加信息的损失。

5]

5. 答：香农公式为 $C_t = \lim_{T\to\infty}\frac{C}{T} = W\log_2\left(1+\frac{P}{N_0W}\right)\ bit/s$，它是高斯加性白噪声信道在单位时间内的信道容量，其值取决于信噪比和带宽。

由 $10\lg\frac{P}{N_0W} = 30\ dB$ 得 $\frac{P}{N_0W} = 1000$，则 $C_t = 5000\log_2(1+1000) = 49836\ bit/s$

6. 答：只要 $\frac{\overline{K_I}}{L} \geq \frac{H(X)}{\log_2 m}$，当 N 足够长时，一定存在一种无失真编码。

7. 答：当 R<C 时，只要码长足够长，一定能找到一种编码方法和译码规则，使译码错误概率无穷小。

8. 答：1）保真度准则为：平均失真度不大于允许的失真度。

2）因为失真矩阵中每行都有一个 0，所以有 $D_{\min} = 0$，而 $D_{\max} = \min\{(1-\omega)\alpha, \omega\alpha\}$。

二、综合题（每题 10 分，共 60 分）

1. 答：1）信源模型为 $\begin{bmatrix} a_1=黑 & a_2=白 \\ 0.3 & 0.7 \end{bmatrix}$

$$H(X) = -\sum_{i=1}^{2} P(a_i)\log_2 P(a_i) = 0.881\ \ bit/符号$$

2）由 $\begin{cases} P(a_i) = \sum_{j=1}^{2} P(a_j)P(a_i/a_j), & i = 1,2 \\ P(a_1) + P(a_2) = 1 \end{cases}$ 得 $\begin{cases} P(白) = \frac{2}{3} \\ P(黑) = \frac{1}{3} \end{cases}$

则 $H_2(X) = -\sum_{i=1}^{2}\sum_{j=1}^{2} P(a_i)P(a_j/a_i)\log_2 P(a_j/a_i) = 0.5533\ \ bit/符号$

2. 答：1）$H(X) = 0.8113\ bit/符号$

$$I(X;Y) = 0.0616\ \ bit/符号$$

2）$C = 0.082\ bit/符号$，最佳输入概率分布为等概率分布。

3. 答：1）二元码的码字依序为：10，11，010，011，1010，1011，1000，1001。

平均码长 $L_2 = 2.6\ bit/符号$，编码效率 $\eta_2 = 0.97$

2）三元码的码字依序为：1，00，02，20，21，22，010，011。

平均码长 $L_3 = 1.7\ bit/符号$，编码效率 $\eta_3 = 0.936$

4. 答：1）最小似然译码准则下，有 $\begin{cases} F(y_1) = x_1 \\ F(y_2) = x_2 \\ F(y_3) = x_3 \end{cases}$，$P_E = \frac{1}{2}$

2）最大错误概率准则下，有 $\begin{cases} F(y_1) = x_1 \\ F(y_2) = x_2 \\ F(y_3) = x_3 \end{cases}$，$P_E = \frac{11}{24}$

5. 答：1）输入为 00011 时，码字为 00011110；输入为 10100 时，码字为 10100101。

2）$d_{min} = 2$

6. 答：1）无错传输时，有 $R \le C = W\log_2(1+\frac{P}{N_0 W})$

即 $5.6\times10^3 = 4\times10^3\log_2(1+\frac{P}{5\times10^{-9}\times4\times10^3})$ 则 $P \ge 0.0328\ mw$

2）在 $P = 0.0328\ mw$ 时，最大熵 $H_c = \frac{1}{2}\log_2(2\pi eP) = -5.4\ bit/自由度$

对应的输入概率密度函数为 $p(x) = \frac{1}{\sqrt{0.206\times10^{-3}}} e^{-\frac{x^2}{0.0556\times10^{-3}}}$

信息论习题集

一、名词解释（每词2分）(25道)

1、“本体论”的信息（P3） 2、“认识论”信息（P3） 3、离散信源（11）

4、自信息量（12） 5、离散平稳无记忆信源（49） 6、马尔可夫信源（58）

7、信源冗余度 （66） 8、连续信源 （68） 9、信道容量 （95）

10、强对称信道 （99） 11、对称信道 （101-102）12、多符号离散信道（109）

13、连续信道 （124） 14、平均失真度 （136） 15、实验信道 （138）

16、率失真函数 （139） 17、信息价值率 （163） 18、游程序列 （181）

19、游程变换 （181） 20、L-D编码（184）、 21、冗余变换 （184）

22、BSC信道 （189） 23、码的最小距离 （193）24、线性分组码 （195）

25、循环码 （213）

二、填空（每空1分）(100道)

1、 在认识论层次上研究信息的时候，必须同时考虑到 形式、含义和效用 三个方面的因素。

2、 1948年，美国数学家 香农 发表了题为“通信的数学理论”的长篇论文，从而创立了信息论。

3、 按照信息的性质，可以把信息分成 语法信息、语义信息和语用信息 。

4、 按照信息的地位，可以把信息分成 客观信息和主观信息 。

5、 人们研究信息论的目的是为了 高效、可靠、安全 地交换和利用各种各样的信息。

6、 信息的 可度量性 是建立信息论的基础。

7、 统计度量 是信息度量最常用的方法。

8、 熵 是香农信息论最基本最重要的概念。

9、 事物的不确定度是用时间统计发生 概率的对数 来描述的。

10、单符号离散信源一般用随机变量描述，而多符号离散信源一般用 随机矢量 描述。

11、一个随机事件发生某一结果后所带来的信息量称为自信息量，定义为 其发生概率对数的负值 。

12、自信息量的单位一般有 比特、奈特和哈特 。

13、必然事件的自信息是 0 。

14、不可能事件的自信息量是 ∞ 。

15、两个相互独立的随机变量的联合自信息量等于 两个自信息量之和 。

16、数据处理定理：当消息经过多级处理后，随着处理器数目的增多，输入消息与输出消息之间的平均互信息量 趋于变小 。

17、离散平稳无记忆信源X的N次扩展信源的熵等于离散信源X的熵的 N倍 。

18、离散平稳有记忆信源的极限熵，$H_\infty = \lim_{N\to\infty} H(X_N / X_1 X_2 \cdots X_{N-1})$ 。

19、对于n元m阶马尔可夫信源，其状态空间共有 n^m 个不同的状态。

20、一维连续随即变量X在[a，b]区间内均匀分布时，其信源熵为 $\log_2(b-a)$ 。

21、平均功率为P的高斯分布的连续信源，其信源熵，$H_c(X) = \frac{1}{2}\log_2 2\pi eP$ 。

22、对于限峰值功率的N维连续信源，当概率密度 均匀分布 时连续信源熵具有最大值。

23、对于限平均功率的一维连续信源，当概率密度 高斯分布 时，信源熵有最大值。

24、对于均值为0，平均功率受限的连续信源，信源的冗余度决定于平均功率的限定值P和信源的熵功率 $\overline{P}$ 之比 。

25、若一离散无记忆信源的信源熵H（X）等于2.5，对信源进行等长的无失真二进制编码，则编码长度至少为 3 。

26、m元长度为 k_i，i=1，2，…n的异前置码存在的充要条件是：$\sum_{i=1}^{n} m^{-k_i} \le 1$ 。

27、若把掷骰子的结果作为一离散信源，则其信源熵为 $\log_2 6$ 。

28、同时掷两个正常的骰子，各面呈现的概率都为1/6，则“3和5同时出现”这件事的自信息量是 $\log_2 18$（1+2 $\log_2 3$）。

29、若一维随即变量X的取值区间是[0，∞]，其概率密度函数为 $p(x) = \frac{1}{m} e^{-\frac{x}{m}}$ ，其中：$x \ge 0$，m是X的数学期望，则X的信源熵 $H_C(X) = \log_2 me$ 。

30、一副充分洗乱的扑克牌（52张），从中任意抽取1张，然后放回，若把这一过程看作离散无记忆信源，则其信源熵为 $\log_2 52$ 。

31、根据输入输出信号的特点，可将信道分成离散信道、连续信道、半离散或半连续 信道。

32、信道的输出仅与信道当前输入有关，而与过去输入无关的信道称为 无记忆 信道。
33、具有一一对应关系的无噪信道的信道容量 C= $\log_2 n$ 。
34、强对称信道的信道容量 C= $\log_2 n - H_{mi}$ 。
35、对称信道的信道容量 C= $\log_2 m - H_{mi}$ 。
36、对于离散无记忆信道和信源的 N 次扩展，其信道容量 C^N= NC 。
37、对于 N 个对立并联信道，其信道容量 C_N = $\sum_{k=1}^{N} C_k$ 。
38、多用户信道的信道容量用 多维空间的一个区域的界限 来表示。
39、多用户信道可以分成几种最基本的类型： 多址接入信道、广播信道 和相关信源信道。
40、广播信道是只有 一个输入端和多个输出端 的信道。
41、当信道的噪声对输入的干扰作用表现为噪声和输入的线性叠加时，此信道称为 加性连续信道 。
42、高斯加性信道的信道容量 C= $\frac{1}{2}\log_2(1+\frac{P_X}{P_N})$ 。
43、信道编码定理是一个理想编码的存在性定理，即：信道无失真传递信息的条件是 信息率小于信道容量 。
44、信道矩阵 $\begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 代表的信道的信道容量 C= 1 。
45、信道矩阵 $\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ 代表的信道的信道容量 C= 1 。
46、高斯加性噪声信道中，信道带宽 3kHz，信噪比为 7，则该信道的最大信息传输速率 C_t= 9 kHz 。
47、对于具有归并性能的无燥信道，达到信道容量的条件是 $p(y_j)=1/m$ 。
48、信道矩阵 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 代表的信道，若每分钟可以传递 $6*10^5$ 个符号，则该信道的最大信息传输速率 C_t= 10kHz 。
49、信息率失真理论是量化、数模转换、频带压缩和 数据压缩 的理论基础。
50、求解率失真函数的问题，即：在给定失真度的情况下，求信息率的 极小值 。
51、信源的消息通过信道传输后的误差或失真越大，信宿收到消息后对信源存在的不确定性就 越大 ，获得的信息量就越小。
52、信源的消息通过信道传输后的误差或失真越大道传输消息所需的信息率 也越小 。
53、单符号的失真度或失真函数 $d(x_i, y_j)$ 表示信源发出一个符号 x_i，信宿再现 y_j 所引起的 误差或失真 。
54、汉明失真函数 $d(x_i, y_j)$ = $\begin{cases} 0 & i=j \\ 1 & i\neq j \end{cases}$ 。
55、平方误差失真函数 $d(x_i, y_j)$ = $(y_j - x_i)^2$ 。
56、平均失真度定义为失真函数的数学期望，即 $d(x_i, y_j)$ 在 X 和 Y 的 联合概率空间 P（XY）中 的统计平均值。
57、如果信源和失真度一定，则平均失真度是 信道统计特性 的函数。
58、如果规定平均失真度 $\overline{D}$ 不能超过某一限定的值 D，即：$\overline{D} \le D$ 。我们把 $\overline{D} \le D$ 称为 保真度准则 。
59、离散无记忆 N 次扩展信源通过离散无记忆 N 次扩展信道的平均失真度是单符号信源通过单符号信道的平均失真度的 N 倍。
60、试验信道的集合用 P_D 来表示，则 P_D= $\{p(y_j/x_i): \overline{D} \le D; i=1,2,\cdots,n, j=1,2,\cdots,m\}$ 。
61、信息率失真函数，简称为率失真函数，即：试验信道中的平均互信息量的 最小值 。
62、平均失真度的下限取 0 的条件是失真矩阵的 每一行至少有一个零元素 。
63、平均失真度的上限 D_{max} 取 $\{D_j: j=1, 2, \cdots, m\}$ 中的 最小值 。
64、率失真函数对允许的平均失真度是 单调递减和连续的 。
65、对于离散无记忆信源的率失真函数的最大值是 $\log_2 n$ 。
66、当失真度大于平均失真度的上限时 D_{max} 时，率失真函数 R（D）= 0 。
67、连续信源 X 的率失真函数 R（D）= $\inf_{p(y/x)\in P_D} I(X;Y)$ 。

68、当 $D \le \sigma^2$ 时，高斯信源在均方差失真度下的信息率失真函数为 $R(D) = \frac{1}{2}\log_2 \frac{\sigma^2}{D}$ 。

69、保真度准则下的信源编码定理的条件是 信源的信息率 R 大于率失真函数 R（D） 。

70、某二元信源 $\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{Bmatrix}$ 其失真矩阵 D= $\begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix}$，则该信源的 D_{max}= a/2 。

71、某二元信源 $\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{Bmatrix}$ 其失真矩阵 D= $\begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix}$，则该信源的 D_{min}= 0 。

72、某二元信源 $\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{Bmatrix}$ 其失真矩阵 D= $\begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix}$，则该信源的 R（D）= 1-H(D/a) 。

73、按照不同的编码目的，编码可以分为三类：分别是 信源编码、信道编码和安全编码 。

74、信源编码的目的是： 提高通信的有效性 。

75、一般情况下，信源编码可以分为 离散信源编码、连续信源编码和相关信源编码 。

76、连续信源或模拟信号的信源编码的理论基础是 限失真信源编码定理 。

77、在香农编码中，第 i 个码字的长度 k_i 和 $p(x_i)$ 之间有 $-\log_2 p(x_i) \le k_i < 1 - \log_2 p(x_i)$ 关系。

78、对信源 $\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\ 1/4 & 1/4 & 1/8 & 1/8 & 1/16 & 1/16 & 1/16 & 1/16 \end{Bmatrix}$ 进行二进制费诺编码，其编码效率为 1 。

79、对具有 8 个消息的单符号离散无记忆信源进行 4 进制哈夫曼编码时，为使平均码长最短，应增加 2 个概率为 0 的消息。

80、对于香农编码、费诺编码和哈夫曼编码，编码方法惟一的是 香农编码 。

81、对于二元序列 00111000001111100111000001111111，其相应的游程序列是 23652457 。

82、设无记忆二元序列中，“0”和“1”的概率分别是 p_0 和 p_1，则“0”游程长度 L（0）的概率为 $p[L(0)] = p_0^{L(0)-1} p_1$ 。

83、游程序列的熵 等于 原二元序列的熵。

84、若“0”游程的哈夫吗编码效率为 η_0，“1”游程的哈夫吗编码效率为 η_1，且 $\eta_0 > \eta_1$ 对应的二元序列的编码效率为 η，则三者的关系是 $\eta_0 \ge \eta \ge \eta_1$ 。

85、在实际的游程编码过程中，对长码一般采取 截断 处理的方法。

86、“0”游程和“1”游程可以分别进行哈夫曼编码，两个码表中的码字可以重复，但 C 码 必须不同。

87、在多符号的消息序列中，大量的重复出现的，只起占时作用的符号称为 冗余位 。

88、“冗余变换”即：将一个冗余序列转换成一个二元序列和一个 缩短了的多元序列 。

89、L-D 编码是一种 分帧传送冗余位序列 的方法。

90、L-D 编码适合于冗余位 较多或较少 的情况。

91、信道编码的最终目的是 提高信号传输的可靠性 。

92、狭义的信道编码即： 检、纠错编码 。

93、BSC 信道即： 无记忆二进制对称信道 。

94、n 位重复码的编码效率是 1/n 。

95、等重码可以检验 全部的奇数位错和部分的偶数位错 。

96、任意两个码字之间的最小汉明距离有称为码的最小距 d_{min}，则 d_{min}= $\min_{c \ne c'} d(c, c')$ 。

97、若纠错码的最小距离为 d_{min}，则可以纠正任意小于等于 t= $\left\lfloor \frac{d_{min} - 1}{2} \right\rfloor$ 个差错。

98、若检错码的最小距离为 d_{min}，则可以检测出任意小于等于 l= d_{min}-1 个差错。

99、线性分组码是同时具有 分组特性和线性特性 的纠错码。

100、循环码即是采用 循环移位特性界定 的一类线性分组码。

三、判断（每题 1 分）（50 道）

1、 必然事件和不可能事件的自信息量都是 0 。错

2、 自信息量是 $p(x_i)$ 的单调递减函数。对

3、 单符号离散信源的自信息和信源熵都具有非负性。对

4、 单符号离散信源的自信息和信源熵都是一个确定值。错
5、 单符号离散信源的联合自信息量和条件自信息量都是非负的和单调递减的。对
6、 自信息量、条件自信息量和联合自信息量之间有如下关系：

$$I(x_i y_j) = I(x_i) + I(y_j / x_i) = I(y_j) + I(x_i / y_j)$$ 对

7、 自信息量、条件自信息量和互信息量之间有如下关系：

$$I(x_i; y_j) = I(x_i) - I(x_i / y_j) = I(y_j) - I(y_j / x_i)$$ 对

8、 当随即变量 X 和 Y 相互独立时，条件熵等于信源熵。对
9、 当随即变量 X 和 Y 相互独立时，I（X；Y）=H（X）。错
10、信源熵具有严格的下凸性。错
11、平均互信息量 I（X；Y）对于信源概率分布 p（x_i）和条件概率分布 p（y_j/x_i）都具有凸函数性。 对
12、m 阶马尔可夫信源和消息长度为 m 的有记忆信源，其所含符号的依赖关系相同。 错
13、利用状态极限概率和状态一步转移概率来求 m 阶马尔可夫信源的极限熵。 对
14、N 维统计独立均匀分布连续信源的熵是 N 维区域体积的对数。 对
15、一维高斯分布的连续信源，其信源熵只与其均值和方差有关。 错
16、连续信源和离散信源的熵都具有非负性。 错
17、连续信源和离散信源都具有可加性。 对
18、连续信源和离散信源的平均互信息都具有非负性。 对
19、定长编码的效率一般小于不定长编码的效率。 对
20、若对一离散信源（熵为 H（X））进行二进制无失真编码，设定长码子长度为 K，变长码子平均长度为 $\overline{K}$，一般 $\overline{K}$>K。 错
21、信道容量 C 是 I（X；Y）关于 p（x_i）的条件极大值。 对
22、离散无噪信道的信道容量等于 $\log_2 n$，其中 n 是信源 X 的消息个数。 错
23、对于准对称信道，当 $p(y_j) = \frac{1}{m}$ 时，可达到信道容量 C。错
24、多用户信道的信道容量不能用一个数来代表。 对
25、多用户信道的信道容量不能用一个数来代表，但信道的信息率可以用一个数来表示。错
26、高斯加性信道的信道容量只与信道的信噪有关。 对
27、信道无失真传递信息的条件是信息率小于信道容量。对
28、最大信息传输速率，即：选择某一信源的概率分布（p（x_i）），使信道所能传送的信息率的最大值。 错
29、对于具有归并性能的无燥信道，当信源等概率分布时（p（x_i）=1/n），达到信道容量。 错
30、求解率失真函数的问题，即：在给定失真度的情况下，求信息率的极小值。对
31、信源的消息通过信道传输后的误差或失真越大，信宿收到消息后对信源存在的不确定性就越小，获得的信息量就越小。 错
32、当 p（x_i）、p（y_j/x_i）和 d（x_i，y_j）给定后，平均失真度是一个随即变量。 错
33、率失真函数对允许的平均失真度具有上凸性。对
34、率失真函数没有最大值。 错
35、率失真函数的最小值是 0 。对
36、率失真函数的值与信源的输入概率无关。错
37、信源编码是提高通信有效性为目的的编码。 对
38、信源编码通常是通过压缩信源的冗余度来实现的。 对
39、离散信源或数字信号的信源编码的理论基础是限失真信源编码定理。 错
40、一般情况下，哈夫曼编码的效率大于香农编码和费诺编码。 对
41、在编 m（m>2）进制的哈夫曼码时，要考虑是否需要增加概率为 0 的码字，以使平均码长最短。 对
42、游程序列的熵（“0”游程序列的熵与“1”游程序列的熵的和）大于等于原二元序列的熵。 错
43、在游程编码过程中，“0”游程和“1”游程应分别编码，因此，它们的码字不能重复。 错
44、L-D 编码适合于冗余位较多和较少的情况，否则，不但不能压缩码率，反而使其扩张。 对
45、狭义的信道编码既是指：信道的检、纠错编码。 对
46、对于 BSC 信道，信道编码应当是一对一的编码，因此，消息 m 的长度等于码字 c 的长度。 错
47、等重码和奇（偶）校验码都可以检出全部的奇数位错。 对
48、汉明码是一种线性分组码。对
49、循环码也是一种线性分组码。 对
50、卷积码是一种特殊的线性分组码。 错

四、简答（每题 4 分）（20 道）

1、 信息的主要特征有哪些？（4）
2、 信息的重要性质有哪些？（4）
3、 简述几种信息分类的准则和方法。（5）

4、 信息论研究的内容主要有哪些？（8）
5、 简述自信息的性质。（13）
6、 简述信源熵的基本性质。（23）
7、 简述信源熵、条件熵、联合熵和交互熵之间的关系。（48）
8、 信道的分类方法有哪些？（93-94）
9、 简述一般离散信道容量的计算步骤。（107）
10、简述多用户信道的分类。（115-116）
11、简述信道编码定理。（128）
12、简述率失真函数的性质。（140-145）……
13、简述求解一般离散信源率失真函数的步骤。（146-149）
14、试比较信道容量与信息率失真函数。（164）
15、简述编码的分累及各种编码的目的。（168）
16、简述费诺编码的编码步骤。（170）
17、简述二元哈夫曼编码的编码步骤。（173）
18、简述广义的信道编码的分类及各类编码的作用。（188）
19、简述线性分组码的性质。（196）
20、简述循环码的系统码构造过程。（221）

“信息论与编码”试题　　2007 级硕士研究生

2008 年 6 月 14 日

一、基本概念题（闭卷部分，每题 4 分，共 40 分；1 小时内完成并交卷）

1. 试证明 n 维随机变量的共熵，不大于它们各自的熵之和。

证明：

即证明 $H(\mathrm{X}_1,\mathrm{X}_2,\cdots,\mathrm{X}_n)\le\sum_{i=1}^{n}H(\mathrm{X}_i)$

因为　$0\le I(\mathrm{X};\mathrm{Y})=H(\mathrm{X})-H(\mathrm{X/Y})$,

所以　$H(\mathrm{X/Y})\le H(\mathrm{X})$。

由共熵的定义和熵的链接准则，有

$$H(\mathrm{X}_1,\mathrm{X}_2)=H(\mathrm{X}_1)+H(\mathrm{X}_2/\mathrm{X}_1)\le H(\mathrm{X}_1)+H(\mathrm{X}_2)$$

$$H(\mathrm{X}_1,\mathrm{X}_2,\mathrm{X}_3)=H(\mathrm{X}_1)+H(\mathrm{X}_2,\mathrm{X}_3/\mathrm{X}_1)$$

$$=H(\mathrm{X}_1)+H(\mathrm{X}_2/\mathrm{X}_1)+H(\mathrm{X}_3/\mathrm{X}_2,\mathrm{X}_1)\le H(\mathrm{X}_1)+H(\mathrm{X}_2)+H(\mathrm{X}_3)$$

…

$$H(\mathrm{X}_1,\mathrm{X}_2,\cdots,\mathrm{X}_n)=\sum_{i=1}^{n}H(\mathrm{X}_i/\mathrm{X}_{i-1},\cdots,\mathrm{X}_1)\le\sum_{i=1}^{n}H(\mathrm{X}_i)$$

证毕。

2. 请给出信源编码器的主要任务以及对信源编码的基本要求。解：信源编码器的主要任务是完成输入消息集合与输出代码集合之间的映射。

对信源编码有如下基本要求：

（1）选择合适的信道基本符号，以使映射后的代码适应信道。例如，ASCII 码选用了 16 进制数。

（2）寻求一种方法，把信源发出的消息变换成相应的代码组。这种方法就是编码，变换成的代码就是码字。

（3）编码应使消息集合与代码组集合中的元素一一对应。

3. 请给出平均码长界定定理及其物理意义。解：平均码长界定定理：若一个离散无记忆信源 X，具有熵 $H(\mathrm{X})$，对其编码用 D 种基本符号，则总可以找到一种无失真信源编码，构成单义可译码，使其平均码长满足

$$\frac{H(\mathrm{X})}{\log D} \le \bar{b} \le \frac{H(\mathrm{X})}{\log D} + 1$$

平均码长界定定理的物理意义：

编码所追求的，是在单义可译前提下寻求尽可能小的平均码长。平均码长界定定理指出，平均码长的下界值 $\bar{b}_{\min} = \frac{H(\mathrm{X})}{\log D}$。对于给定信源空间{X，$P(\mathrm{X})$}的离散信源，其熵 $H(\mathrm{X})$是确定的数值，如果信道基本符号也是确定的，即 D 也是给定的，则 $\bar{b}_{\min}$ 也就定了。这意味着，如果不改变信源的统计特性，减小 $\bar{b}$ 的潜力，到了其下界值也就到了极限了。因此，如果要进一步提高编码效率，必须对信源本身进行研究，例如改变信源本身的统计特性，对其进行扩展。

4. 请给出连续信源分别为均匀分布、高斯分布和指数分布时信源的相对熵。解：（1）均匀分布连续信源的相对熵为

$$\begin{aligned} h(x) &= -\int_a^b p(x)\log p(x)dx \\ &= \log(b-a) \end{aligned}$$

（2）高斯分布连续信源 X 的相对熵为

$$\begin{aligned} h(x) &= -\int_{-\infty}^{\infty} p(x)\ln p(x)dx \\ &= -\int_{-\infty}^{\infty} p(x)\ln\left\{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(x-m)^2}{2\sigma^2}\right]dx\right\} \\ &= -\int_{-\infty}^{\infty} p(x)\ln\frac{1}{\sqrt{2\pi\sigma^2}}dx + \int_{-\infty}^{\infty} p(x)\frac{(x-m)^2}{2\sigma^2}dx \\ &= \ln\sqrt{2\pi\sigma^2} + \frac{\sigma^2}{2\sigma^2} \qquad \text{中间步骤可以省略} \\ &= \frac{1}{2}\ln(2\pi\sigma^2) + \frac{1}{2} \\ &= \frac{1}{2}\ln(2\pi\sigma^2) + \frac{1}{2}\ln e \\ &= \frac{1}{2}\ln(2\pi e\sigma^2) \end{aligned}$$

（3）指数分布连续信源 X 的相对熵为

$$\begin{aligned} h(x) &= -\int_0^{\infty} p(x)\ln p(x)\,\mathrm{d}x \\ &= \int_0^{\infty}\left[\frac{1}{a}e^{-\frac{x}{a}}\right]\ln\left[\frac{1}{a}e^{-\frac{x}{a}}\right]\mathrm{d}x \\ &= \frac{1}{a}\ln a\cdot\int_0^{\infty} e^{-\frac{x}{a}}\,\mathrm{d}x + \frac{1}{a^2}\int_0^{\infty} xe^{-\frac{x}{a}}\,\mathrm{d}x \qquad \text{中间步骤可以省略} \\ &= \ln a + \ln e \\ &= \ln ae \end{aligned}$$

5. 请给出失真函数、平均失真度、保真度准则、信息率失真函数的定义。解：失真函数定义：对于有失真的信息传输系统，对应于每一对（a_i, b_j）（$n=1, 2, \ldots r$; $j=1, 2, \ldots, s$），定义一个非负实值函数

$$d(a_i, b_j) \geq 0 \qquad (i=1,2,\cdots r; j=1,2,\cdots s)$$

表示信源发出符号 a_i 而经信道传输后再现成信道输出符号集合中的 b_j 所引起的误差或失真，称之为 a_i 和 b_j 之间的失真函数（Distortion Function），简写为 d_{ij}。

平均失真度定义：若信源和信宿的消息集合分别为 X:$\{a_1, a_2, \ldots, a_r\}$和 Y:$\{b_1, b_2, \ldots, b_s\}$，其概率分别为 $P(a_i)$和 $P(b_j)$ ($i=1, 2, \ldots, r$; $j=1, 2, \ldots, s$)，信道的转移概率为 $P(b_j/a_i)$，失真函数为 $d(a_i, b_j)$，则称随机变量 X和 Y 的联合概率 $P(a_i b_j)$对失真函数 $d(a_i, b_j)$的统计平均值为该通信系统的平均失真度 $\overline{D}$。

保真度准则定义：从平均的意义上来说，信道每传送一个符号所引起的平均失真，不能超过某一给定的限定值 D，即要求 $\overline{D} \leq D$，称这种对于失真的限制条件为保真度准则。

信息率失真函数定义：用给定的失真 D 为自变量来描述的信息传输速率，称为信息率失真函数，用 $R(D)$表示。

6. 试证明(n, k)循环码的生成多项式 $g(x)$是 x^n+1 的因式。

证明：将生成多项式 $g(x)$乘以 x^k，得

$$x^k g(x) = g^{(k)}(x) + q(x)(x^n+1)$$

由于 $x^k g(x)$ 次数为 n，故上式中 $q(x)=1$，而 $g^{(k)}(x)$ 是 $g(x)$循环左移 k 次所得，它是 $g(x)$的倍式，设 $g^{(k)}(x)=u(x)g(x)$，故有 $x^n+1=[x^k+u(x)]g(x)=f(x)g(x)$

证毕。

7. 请给出域的定义并说明集合{0, 1, 2}可否构成域及其理由。

解：域的定义：非空元素集合 F，若在 F 中定义了加和乘两种运算，且满足

(1) F 关于加法构成 Abel 群，其加法恒元记为 0；

(2) F 中非零元素全体对乘法构成 Abel 群，其乘法恒元记为 1；

(3) 加法和乘法间有如下分配律：$a(b+c)=ab+ac$，$(b+c)a=ba+ca$，

则称 F 是一个域。

或者说，域是一个可换的、有单位元的、非零元素有逆元的环。

集合{0, 1, 2}可以构成域。对该集合中的元素定义模 3 加和模 3 乘这两种运算，完全符合域必须满足的 3 个条件。

8. 请给出本原多项式的定义，并用一个实例来说明它的性质。

解：本原多项式的定义：若 m 次既约多项式 $p(x)$除尽的 x^n+1 的最小正整数 n 满足 $n=2^m-1$，称 $p(x)$为本原多项式。

用实例来说明本原多项式有如下性质：

1）本原多项式一定是既约的（因为它是用既约多项式来定义的），但既约多项式不一定是本原的。

例如：4 次既约多项式 x^4+x+1 能除尽 $x^{15}+1$，但除不尽任何 $1 \leq n < 15$ 的 x^n+1，所以 x^4+x+1 是本原的；但同样是 4 次既约多项式 $x^4+x^3+x^2+x+1$，能除尽 $x^{15}+1$，但也能除尽 x^5+1，所以 $x^4+x^3+x^2+x+1$ 是既约的但不是本原的。

2）对于给定的 m，可能有不止一个 m 次本原多项式。

例如，对于 $m=5$，x^5+x^3+1 是本原多项式，x^5+x^2+1 也是。

9. 试说明(n, k)循环码对突发错误的检测能力。

解：（1）(n, k)循环码能检测长为 $n-k$ 或更短的任何突发错误，包括首尾相接突发错误。

（2）(n, k)循环码对 $n-k+1$ 位长的突发错误不能被检出所占的概率最大是 $2^{-(n-k+1)}$。

（3）如果 $l>n-k+1$，则 (n,k) 循环码不能检测长为 l 的突发错误所占的比值为 $2^{-(n-k)}$。

因此，循环码检测突发错误非常有效。

10. 请给出最佳自由距离卷积码的定义并简要说明如何获得具有最佳自由距离的卷积码。

解：最佳自由距离卷积码的定义：对于相同的码率 R 和相同的电路复杂性（存储单元总数 m 等）的各种卷积码，使得自由距离 d_f 最大的编码称为最佳自由距离（OFD，Optimal Free Distance）码。

为了得到各种 OFD 码，通常采用计算机搜索的方法，即对于给定的存储单元总数 m 所有可能的卷积码编码器，首先排除恶性卷积码，然后对应每一可能的卷积码编码器求其自由距离 d_f，逐一比较得到自由距离 d_f 最大者即为最佳自由距离卷积码编码器。

二、综合题（开卷部分，每题 10 分，共 60 分。闭卷部分交卷后方可参阅参考资料）

1. 某通信系统的信源输出仅有 2 个符号 a、b，拟采用 Lempel-Ziv 编码后送信道传输，若某次通信需传输的符号序列为“aaaaaaa bbbbbbb aaaaaaaa bbbbbbb a b aaaaaa bbbbbbb aaaaaaaaaaaaaaaab”，请给出其 Lempel-Ziv 编码结果并简要说明该编码的性能。

解：

编码结果

编码包	<0,7,b>	<1,6,a>	<15,15,b>	<15,14,a>	<1,15,b>
内容	aaaaaaa b	bbbbbb a	aaaaaaab bbbbbba b	aaaaaabbbbbbba a	aaaaaaaaaaaaaaab
码段符号数	8	7	16	15	16

如果把 a、b 看作为 1、0，对编码结果（即每个编码包）可以用 4+4+1=9-bit 表示，传输该序列用 5 个编码包即 45-bit，而该序列有 62-bit，因此该编码起到压缩作用。

2. 若题 1 信源符合 a、b 的出现概率分别为 0.9 和 0.1，拟对其采用 3 重扩展后再进行霍夫曼编码，请给出编码过程及结果，并求该种信源编码的效率。

解：

aaa：0.729　aab：0.081　aba：0.081　abb：0.009　baa：0.081　bab：0.009　bba：0.009　bbb：0.001

设 aaa、aab、…bbb 分别为 x_1、x_2、…x_9，按照概率大小依此排列，有

x_1

x_2

x_3

x_5

x_4

x_6

x_7

x_8

具体编码过程、结果、编码效率一略。

3. 为了在有噪信道中获得可靠的通信，拟对题 2 的霍夫曼编码结果再进行信道编码，若霍夫曼编码的输出序列为 aabb baaa baba bbaa bbba abbb abab …，试给出采用戈莱码（23，12）编码的第一个码字；如果信道编码不是采用戈莱码而是采用缩短的 BCH（120，78）编码，试给出构造该种编码的生成多项式的方法以及缩短的方法，分析其纠错和检错能力，简述其编码和译码过程。

解：

（1）对戈莱码（23，12）采用生成多项式为 $g(x)=x^{11}+x^{10}+x^6+x^5+x^4+x^2+1$，即 110001110101

令 a=1，b=0，则要编码的序列为 1100 0111 0101 0011 0001 1000 1010 …；由于戈莱码是非本原 BCH 码，其编码规则与 BCH 码相同，现采用系统码，第一个码字的编码过程如下：

110001110101　　00000000000

110001110101

－－－－－－－－－－－－－－－

00000000000

因此第一个戈莱码码字为 110001110101 00000000000，即监督位为全 0（11 位）。

（2）缩短的 BCH（120，78）原码为 BCH（127，85），构造该种编码的生成多项式，可以由

$$g(x)=\mathrm{LCM}\{\phi_1(x)+\phi_2(x)+\phi_3(x)x^6+\cdots+\phi_{2t}(x)\}$$

对于本题，$127=2^7$，$\phi_i(x)$是 $GF(2^7)$上的元素α^i的最小多项式，求出α、α^3、α^5、α^7的最小多项式，将它们相乘即得到该种编码能够纠 5 个错的生成多项式（次数为 127－85＝42）。

缩短的方法取原码 BCH（127，85）中的一个子集，其消息位的前 7 位均为 0，编码方法与原码相同，只是传输时前 7 位 0 不要传送。

由求其生成多项式的过程已知该种编码的原码能够纠 5 个随机错误，至少能够检测 10 个随机错误，由于其监督位的个数位 42，它能够检测 42 个突发错误并依概率检测大于 42 个突发错误。

其编码方法上面已经说明，可以采用系统码的编码方法，只是传输时前 7 位 0 不要传送。

解码时通常先补上缩短的 0 的位数，再按照原码的译码方法进行译码，通常采用伴随式译码方法。即：先求出接收序列的伴随式；然后根据伴随式求错误位置，本题用查表法将很复杂，拟采用错误位置多项式的方法来求错误位置；得到错误位置后纠正之。

4. 假设对题 2 的霍夫曼编码输出进行卷积编码，采用的卷积码编码器为（3，1，2），请给出你设计的卷积码编码器并说明其是最佳自由距离卷积码，用状态图方法给出输入序列（aabbabab …）的编码输出和用网格图方法给出接收序列（aba abb aab baa abb aba …）的译码输出。

解：

可以将教材中图 12.7 变成系统卷积码，说明它不是恶性卷积码并通过与其他各类抽头方式比较知它的自由距离为最大，故是最佳自由距离卷积码。

具体的编码器、状态图、网格图一略。

5. （1）试根据香农第二定理说明为什么交织虽然没有注入冗余度但却提高了纠突发错误能力；（2）试给出利用信息加密技术进行保密通信的系统框图（加密信道模型）并简述其工作原理，说明为什么非通信对像虽然收到了密文但在不能成功破译时其获得的信息量为 0。

解：（1）设交织深度为λ，则交织的结果等效为编码长度扩大了λ倍，其禁用码组与许用码组之比也扩大了λ倍，根据香农第二定理，在 $R<C$ 情况下，码长越长其纠、检错的能力应该越强。

（2）信息加密技术进行保密通信的系统框图如图所示（教材图 14.1）。

其工作原理为：明码文本 M 利用密钥，通过某种可逆变换 E_K 加密成密文 C，即 $C=E_K(M)$；密文通过不安全的或公共信道进行传输，在传输过程中可能出现密文截取（截取密

文又称为攻击或入侵），当合法用户得到 C 后，用逆变换 $D_K = E_K^{-1}$ 进行解密得到原来的明码文本消息，即 $D_K(C) = E_K^{-1}\left[E_K(M)\right] = M$；参数 K 是由码元或字符组成的密钥，它规定了密码变换集合中特定的一种加密变换 E_K。

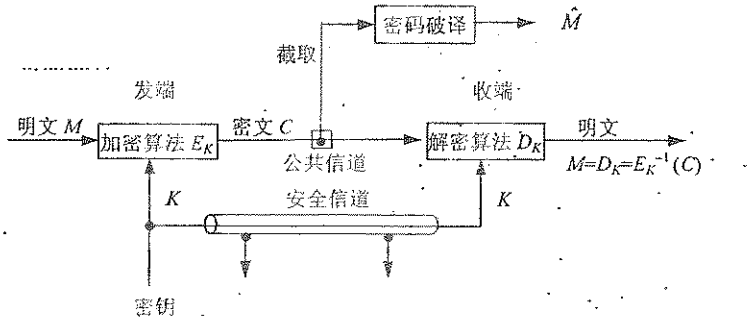


图 14.1　加密信道模型

非通信对像收到了密文但在不能成功破译时，其不确定性集合中的元素没有任何变化，根据自信息量和平均互信息量的定义，他所获得的信息量为 0。

6. 试从收、发双方联合优化的出发点来说明为什么信源编码通常其编码较为复杂而信道编码通常其译码较为复杂的原因。

解：

通常信源编码的信道基本符号数目较少，而根据收发联合优化的考虑，克服信道产生的影响（对数字通信而言主要表现在误码上）主要由信道编码完成，故对信源译码没有提出纠错或检错的要求，因此在收端有一个与发端完全相同的信道基本符号集合就足以完成译码；因为单译可译的编码工作在发端完成而收端只要按照编码规则对收到的码进行比对就可以了，因此通常编码比译码要复杂些。

对于信道编码，选定编码规则后，编码所要处理的主要是消息序列（把消息序列映射为码字），以分组码为例共 2^k 个；而译码需要处理的是所有可能接收到的序列（仍以分组码为例则共有 2^n 个），考虑到信道的复杂性，译码均依从最大似然准则，故其比对过程比编码的映射运算复杂很多。

一、（11’）填空题

（1） 1948 年，美国数学家 香农 发表了题为“通信的数学理论”的长篇论文，从而创立了信息论。

（2） 必然事件的自信息是 0 。

（3） 离散平稳无记忆信源 X 的 N 次扩展信源的熵等于离散信源 X 的熵的 N 倍 。

（4） 对于离散无记忆信源，当信源熵有最大值时，满足条件为 信源符号等概分布 。

（5） 若一离散无记忆信源的信源熵 H（X）等于 2.5，对信源进行等长的无失真二进制编码，则编码长度至少为 3 。

（6） 对于香农编码、费诺编码和霍夫曼编码，编码方法惟一的是 香农编码 。

（7） 已知某线性分组码的最小汉明距离为 3，那么这组码最多能检测出 2 个码元错误，最多能纠正 1 个码元错误。

（8） 设有一离散无记忆平稳信道，其信道容量为 C，只要待传送的信息传输率 R 小于 C（大于、小于或者等于），则存在一种编码，当输入序列长度 n 足够大，使译码错误概率任意小。

（9） 平均错误概率不仅与信道本身的统计特性有关，还与 译码规则 和 编码方法 有关

二、（9’）判断题

（1） 信息就是一种消息。 （ × ）

（2） 信息论研究的主要问题是在通信系统设计中如何实现信息传输、存储和处理的有效性和可靠性。 （ √ ）

（3） 概率大的事件自信息量大。 （ × ）

（4） 互信息量可正、可负亦可为零。 （ √ ）

（5） 信源剩余度用来衡量信源的相关性程度，信源剩余度大说明信源符号间的依赖关系较小。 （ × ）

（6） 对于固定的信源分布，平均互信息量是信道传递概率的下凸函数。 （ √ ）

（7） 非奇异码一定是唯一可译码，唯一可译码不一定是非奇异码。 （ × ）

（8） 信源变长编码的核心问题是寻找紧致码（或最佳码），霍夫曼编码方法构造的是最佳码。

（ √ ）

（9）信息率失真函数 R(D)是关于平均失真度 D 的上凸函数. （ × ）

三、（5′）居住在某地区的女孩中有 25%是大学生，在女大学生中有 75%是身高 1.6 米以上的，而女孩中身高 1.6 米以上的占总数的一半。

假如我们得知“身高 1.6 米以上的某女孩是大学生”的消息，问获得多少信息量？

解：设A表示“大学生”这一事件，B表示“身高1.60以上”这一事件，则

P(A)=0.25 p(B)=0.5 p(B|A)=0.75 （2分）

故 p(A|B)=p(AB)/p(B)=p(A)p(B|A)/p(B)=0.75*0.25/0.5=0.375 （2分）

I(A|B)=-log0.375=1.42bit （1分）

四、（5′）证明：平均互信息量同信息熵之间满足

$$\mathbf{I(X;Y)=H(X)+H(Y)-H(XY)}$$

证明：

$$I(X;Y)=\sum_X\sum_Y p(x_i y_j)\log\frac{p(x_i|y_j)}{p(x_i)}$$

$$=-\sum_X\sum_Y p(x_i y_j)\log p(x_i)-\left[-\sum_X\sum_Y p(x_i y_j)\log p(x_i|y_j)\right] \quad (2分)$$

$$=H(X)-H(X|Y)$$

同理

$$I(X;Y)=H(Y)-H(Y|X) \quad (1分)$$

则

$$H(Y|X)=H(Y)-I(X;Y)$$

因为

$$H(XY)=H(X)+H(Y|X) \qquad (1分)$$

故

$$H(XY)=H(X)+H(Y)-I(X;Y)$$

即

$$I(X;Y)=H(X)+H(Y)-H(XY) \qquad (1分)$$

五、（18'）.黑白气象传真图的消息只有黑色和白色两种，求：

1） 黑色出现的概率为 0.3，白色出现的概率为 0.7。给出这个只有两个符号的信源 X 的数学模型。假设图上黑白消息出现前后没有关联，求熵 $H(X)$；

2） 假设黑白消息出现前后有关联，其依赖关系为 $P(白/白)=0.9$，$P(黑/白)=0.1$，$P(白/黑)=0.2$，$P(黑/黑)=0.8$，求其熵 $H_\infty(X)$。

3）分别求上述两种信源的冗余度，比较它们的大小并说明其物理意义。

解：1）信源模型为

$$\begin{bmatrix} a_1=黑 & a_2=白 \\ 0.3 & 0.7 \end{bmatrix} \qquad (1分)$$

$$H(X)=-\sum_{i=1}^{2}P(a_i)\log_2 P(a_i)=0.881 \quad bit/符号 \qquad (2分)$$

2）由题意可知该信源为一阶马尔科夫信源。 （2分）

由

$$\begin{cases} P(a_i)=\sum_{j=1}^{2}P(a_j)P(a_i/a_j), \quad i=1,2 \\ P(a_1)+P(a_2)=1 \end{cases} \qquad (4分)$$

得极限状态概率

$$\begin{cases} P(白)=\frac{2}{3} \\ P(黑)=\frac{1}{3} \end{cases}$$

（2分）

$$H_2(X) = -\sum_{i=1}^{2}\sum_{j=1}^{2} P(a_i)P(a_j/a_i)\log_2 P(a_j/a_i) = 0.5533 \quad bit/符号 \text{.....} \text{（3分）}$$

3)

$$\gamma_1 = 1 - \frac{H(X)}{\log_2 2} = 0.119$$

（1分）

$$\gamma_2 = 1 - \frac{H_\infty(X)}{\log_2 2} = 0.447$$

（1分）

$\gamma_2 > \gamma_1$。说明：当信源的符号之间有依赖时，信源输出消息的不确定性减弱。而信源冗余度正是反映信源符号依赖关系的强弱，冗余度越大，依赖关系就越大。（2分）

六、（18’）信源空间为

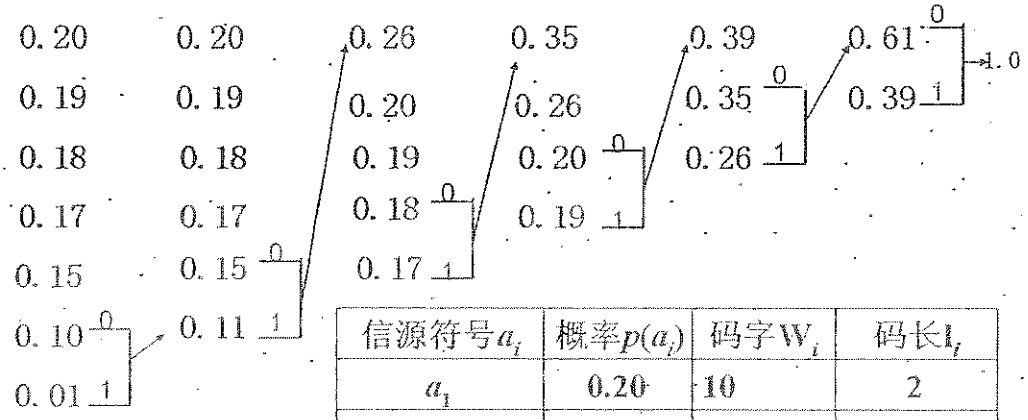
$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \\ 0.2 & 0.19 & 0.18 & 0.17 & 0.15 & 0.1 & 0.01 \end{bmatrix}$$

，试分别构造二元香农码和二元霍夫曼码，计算其平均码长和编码效率（要求有编码过程）。

信源消息符号a_i	符号概率(a_i)	累加概率P_i	$-\log p(a_i)$	码字长度l_i	码字
a_1	0.20	0	2.32	3	000
a_2	0.19	0.2	2.39	3	001
a_3	0.18	0.39	2.47	3	011
a_4	0.17	0.57	2.56	3	100
a_5	0.15	0.74	2.74	3	101
a_6	0.16	0.89	3.32	4	1110
a_7	0.01	0.99	6.64	7	1111110

$$\bar{L}=\sum_{i=1}^{7}p(a_i)l_i=3.14$$

$$R=\frac{H(X)}{\bar{L}}=\frac{2.61}{3.14}=0.831$$



信源符号a_i	概率$p(a_i)$	码字W_i	码长l_i
a_1	0.20	10	2
a_2	0.19	11	2
a_3	0.18	000	3
a_4	0.17	001	3
a_5	0.15	010	3
a_6	0.10	0110	4
a_7	0.01	0111	4

$$\bar{L}=\sum_{i=1}^{7}p(a_i)l_i=2.72$$

码元/符号

$$R=\frac{H(X)}{\bar{L}}=\frac{2.61}{2.72}=0.96$$

比特/符号

七（6'）.设有一离散信道，其信道传递矩阵为$\begin{bmatrix}1/2 & 1/3 & 1/6\\ 1/6 & 1/2 & 1/3\\ 1/3 & 1/6 & 1/2\end{bmatrix}$，并设$\begin{cases}p(x_1)=\frac{1}{4}\\ p(x_2)=\frac{1}{2}\\ p(x_3)=\frac{1}{4}\end{cases}$，试分别按最大后验概率准则与最大似然译码准则确定译码规则，并计算相应的平均错误概率。

1）（3分）最小似然译码准则下，有，

$$P_E=\frac{1}{2}$$

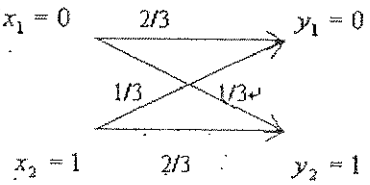
2）（3分）最大后验概率准则下，有，

$$\begin{cases}F(y_1)=x_1\\ F(y_2)=x_2\\ F(y_3)=x_2\end{cases}\qquad P_E=\frac{11}{24}$$

八（10′）.二元对称信道如图：……

1）若 $p(0)=\frac{3}{4}$，$p(1)=\frac{1}{4}$，求 $H(X)$、$H(X|Y)$和 $I(X;Y)$；

2）求



解：1）共6分

$H(X)=0.8113\ bit/符号$

$H(X|Y)=0.749bit/符号$

$I(X;Y)=0.0616\ bit/符号$

2），$C=0.082\ bit/符号$（3分）此时输入概率分布为等概率分布。（1分）

九、（18′）设一线性分组码具有一致监督矩阵 $H=\begin{bmatrix}0&0&0&1&1&1\\0&1&1&0&0&1\\1&0&1&0&1&1\end{bmatrix}$

1）求此分组码 n=?,k=?共有多少码字？

2）求此分组码的生成矩阵 G。

3）写出此分组码的所有码字。

4）若接收到码字（101001），求出伴随式并给出翻译结果。

解：1）n=6,k=3,共有8个码字。（3分）

2）设码字 $\vec{C}=(C_5C_4C_3C_2C_1C_0)$ 由 $HC^T=0^T$ 得

$$\begin{cases} C_2 \oplus C_1 \oplus C_0 = 0 \\ C_4 \oplus C_3 \oplus C_0 = 0 \\ C_5 \oplus C_3 \oplus C_1 \oplus C_0 = 0 \end{cases}$$ （3分）

令监督位为$(C_2C_1C_0)$，则有

$$\begin{cases} C_2 = C_5 \oplus C_3 \\ C_1 = C_5 \oplus C_4 \\ C_0 = C_4 \oplus C_3 \end{cases}$$ （3分）

生成矩阵为$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$ （2分）

3）所有码字为000000，001101，010011，011110，100110，101011，110101，111000。（4分）

4）由$S^T = HR^T$得

$S = (101)$，（2分）该码字在第5位发生错误，（101001）纠正为（101011），即译码为（101001）（1分）。