《信息论基础》参考答案

一、填空题

1、信源编码的主要目的是提高有效性，信道编码的主要目的是提高可靠性。

2、信源的剩余度主要来自两个方面，一是信源符号间的相关性，二是信源符号的统计不均匀性。

3、三进制信源的最小熵为 0，最大熵为 $\log_2 3$ bit/符号。

4、无失真信源编码的平均码长最小理论极限制为信源熵（或 H(S)/logr= H_r(S)）。

5、当 R=C 或（信道剩余度为 0）时，信源与信道达到匹配。

6、根据信道特性是否随时间变化，信道可以分为恒参信道和随参信道。

7、根据是否允许失真，信源编码可分为无失真信源编码和限失真信源编码。

8、若连续信源输出信号的平均功率为 σ^2，则输出信号幅度的概率密度是高斯分布或正态分布或 $f(x)=\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{x^2}{2\sigma^2}}$ 时，信源具有最大熵，其值为值 $\frac{1}{2}\log 2\pi e\sigma^2$。

9、在下面空格中选择填入数学符号“$=,\geq,\leq,\rangle$”或“$\langle$”

（1）当 X 和 Y 相互独立时，H（XY）=H(X)+H(X/Y)=H(Y)+H(X)。

（2）$H_2(X)=\frac{H(X_1X_2)}{2}\geq H_3(X)=\frac{H(X_1X_2X_3)}{3}$

（3）假设信道输入用 X 表示，信道输出用 Y 表示。在无噪有损信道中，H(X/Y)≥ 0, H(Y/X)=0, I(X;Y)≤H(X)。

三、已知信源

$$\begin{bmatrix} S \\ P \end{bmatrix}=\begin{bmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.1 & 0.1 \end{bmatrix}$$

（1）用霍夫曼编码法编成二进制变长码；（6 分）

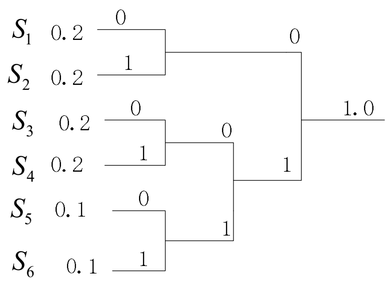
（2）计算平均码长 $\overline{L}$；（4 分）

（3）计算编码信息率 R'；（2 分）

（4）计算编码后信息传输率 R；（2 分）

（5）计算编码效率 η。（2 分）

（1）



编码结果为：

$$S_1=00$$
$$S_2=01$$
$$S_3=100$$
$$S_4=101$$
$$S_5=110$$
$$S_6=111$$

（2）$\overline{L}=\sum_{i=1}^{6}P_i\rho_i=0.4\times2+0.6\times3=2.6$ 码元/符号

（3）$R'=\overline{L}\log \mathrm{r}=2.6$ bit/符号

（4）$R=\dfrac{H(S)}{\overline{L}}=\dfrac{2.53}{2.6}=0.973$ bit/码元 其中，$H(S)=H(0.2,0.2,0.2,0.2,0.1,0.1)=2.53$ bit/符号

（5）$\eta=\dfrac{H(S)}{\overline{L}\log r}=\dfrac{H(S)}{\overline{L}}=0.973$

四、某信源输出 A、B、C、D、E 五种符号，每一个符号独立出现，出现概率分别为 1/8、1/8、1/8、1/2、1/8。如果符号的码元宽度为 0.5 μs。计算：

（1）信息传输速率 R_t。

（2）将这些数据通过一个带宽为 B=2000kHz 的加性白高斯噪声信道传输，噪声的单边功率谱密度为 $n_0=10^{-6}\,W/Hz$。试计算正确传输这些数据最少需要的发送功率 P。

解：

（1）$R_t=\dfrac{1}{t}\left[H(X)-H\left(X/Y\right)\right]$

$$\begin{aligned}H(X)&=-\frac{1}{8}\log\frac{1}{8}\times4-\frac{1}{2}\log\frac{1}{2}\\&=\frac{1}{2}\log8+\frac{1}{2}\log2\\&=\frac{3}{2}\log2+\frac{1}{2}\log2\\&=2\log2\\&=2\mathrm{bit}\end{aligned}$$

$$R_t=\frac{2bit}{0.5\mu s}=4\times10^6 bps$$

$$4\times10^6=2\times10^6\log\left(1+\frac{P}{10^{-6}\times2\times10^6}\right)$$

（2）$1+\dfrac{P}{2}=2^2$

$P=6W$

五、一个一阶马尔可夫信源，转移概率为

$P(S_1|S_1)=\dfrac{2}{3},P(S_2|S_1)=\dfrac{1}{3},P(S_1|S_2)=1,P(S_2|S_2)=0$。

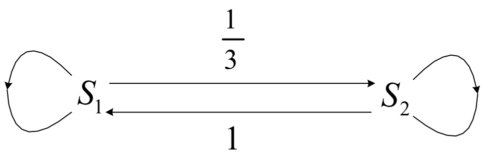
(1) 画出状态转移图。

(2) 计算稳态概率。

(3) 计算马尔可夫信源的极限熵。

(4) 计算稳态下 H_1,H_2 及其对应的剩余度。

解：(1)



(2)由公式 $P(S_i)=\sum_{j=1}^{2}P\left(S_i|S_j\right)P\left(S_j\right)$

有
$$\begin{cases} P(S_1)=\sum_{i=1}^{2}P(S_1\,|\,S_i)P(S_i)=\frac{2}{3}P(S_1)+P(S_2) \\ P(S_2)=\sum_{i=1}^{2}P(S_2\,|\,S_i)P(S_i)=\frac{1}{3}P(S_1) \\ P(S_1)+P(S_2)=1 \end{cases}$$

得
$$\begin{cases} P(S_1)=\frac{3}{4} \\ P(S_2)=\frac{1}{4} \end{cases}$$

(3)该马尔可夫信源的极限熵为：

$$\begin{aligned} H_\infty &= -\sum_{i=1}^{2}\sum_{j=1}^{2}P(S_i)P(S_j\,|\,S_i)\log P(S_j\,|\,S_i) \\ &= -\frac{3}{4}\times\frac{2}{3}\times\log\frac{2}{3}-\frac{3}{4}\times\frac{1}{3}\times\log\frac{1}{3} \\ &= \frac{1}{2}\times 0.578+\frac{1}{4}\times 1.599 \\ &= 0.681\,bit/\text{符号} \\ &= 0.472\,nat/\text{符号} \\ &= 0.205\,hart/\text{符号} \end{aligned}$$

(4)在稳态下：

$$=-\sum_{i=1}^{2}P(x_i)\log P(x_i)=-\left(\frac{3}{4}\times\log\frac{3}{4}+\frac{1}{4}\times\log\frac{1}{4}\right)=0.811\,bit/\text{符号}$$

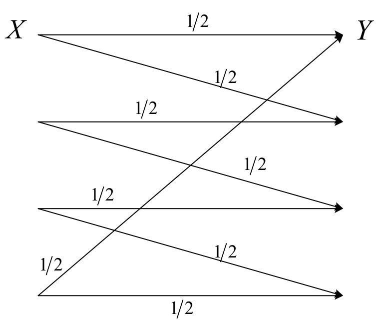
$$H_2=H_\infty=0.205\,hart/\text{符号}=0.472\,nat/\text{符号}=0.681\,bit/\text{符号}$$

对应的剩余度为

$$\eta_1=1-\frac{H_1}{H_0}=1-\frac{0.811}{-\left(\frac{1}{2}\log\left(\frac{1}{2}\right)+\frac{1}{2}\log\left(\frac{1}{2}\right)\right)}=0.189$$

$$\eta_2=1-\frac{H_2}{H_0}=1-\frac{0.681}{-\left(\frac{1}{2}\log\left(\frac{1}{2}\right)+\frac{1}{2}\log\left(\frac{1}{2}\right)\right)}=0.319$$

六、设有扰信道的传输情况分别如图所示。试求这种信道的信道容量。



解：信道传输矩阵如下

$$P_{Y|X}=\begin{bmatrix}\frac{1}{2} & \frac{1}{2} & 0 & 0\\ 0 & \frac{1}{2} & \frac{1}{2} & 0\\ 0 & 0 & \frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & 0 & 0 & \frac{1}{2}\end{bmatrix}$$

可以看出这是一个对称信道，L=4,那么信道容量为

$$\begin{aligned}C&=\log 4-H\left(\frac{1}{2},\frac{1}{2},0,0\right)\\&=\log L+\sum_{j=1}^{L}p\left(y_j\mid x_i\right)\log p\left(y_j\mid x_i\right)\\&=\log 4+2\times\frac{1}{2}\log\frac{1}{2}\\&=1bit\end{aligned}$$

七、设 X、Y 是两个相互独立的二元随机变量，其取 0 或 1 的概率相等。定义另一个二元随机变量 Z=XY(一般乘积)。试计算

(1) $H(X),H(Z)$;

(2) $H(XY),H(XZ)$;

(3) $H(X\mid Y),H(Z\mid X)$;

(4) $I(X;Y),I(X;Z)$;

解：(1)

Z	0	1	
P(Z)	3/4	1/4	

$$H(X)=H\left(\frac{1}{2},\frac{1}{2}\right)=1bit$$

$$H(2)=H\left(\frac{3}{4},\frac{1}{4}\right)=0.8113bit$$

(2) $H(XY)=H(X)+H(Y)=1+1=2\,bit/$对

$$H(XZ)=H(X)+H(Z|X)=1+\frac{1}{2}H(1,0)+\frac{1}{2}H\left(\frac{1}{2},\frac{1}{2}\right)=1.5bit/\text{对}$$

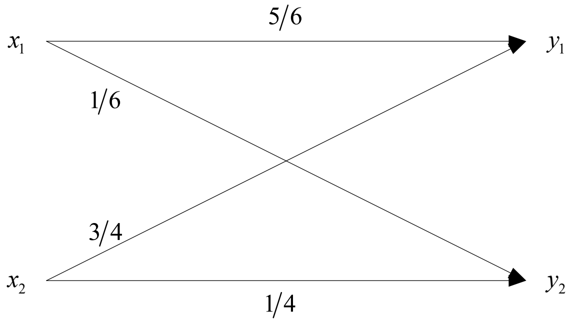
(3) $H(X\mid Y)=H(X)=1bit$

$$H(Z\mid X)=\frac{1}{2}H(1,0)+\frac{1}{2}H\left(\frac{1}{2},\frac{1}{2}\right)=0.5bit$$

(4) $I(X,Y)=H(Y)-H(Y\mid X)=H(Y)-H(Y)=0$

$$I(X,Z)=H(Z)-H(Z\mid X)=0.8113-0.5=0.3113bit$$

八、设离散无记忆信源的概率空间为$\begin{bmatrix} X \\ P \end{bmatrix}=\begin{bmatrix} x_1 & x_2 \\ 0.8 & 0.2 \end{bmatrix}$，通过干扰信道，信道输出端的接收符号集为$Y=[y_1,y_2]$，信道传输概率如下图所示。



(1) 计算信源 X 中事件 x_1 包含的自信息量；

(2) 计算信源 X 的信息熵；

(3) 计算信道疑义度 $H(X\mid Y)$；

(4) 计算噪声熵 $H(Y\mid X)$；

(5) 计算收到消息 Y 后获得的平均互信息量。

解：

(1) $I(x_1)=-\log 0.8=0.322bit=0.0969hart=0.223nat$

(2) $H(X)=H(0.8,0.2)=0.722bit/符号=0.5nat/符号=0.217hart/符号$

(3)转移概率：

x \ y	y_1	y_2
x_1	5/6	1/6
x_2	3/4	1/4

联合分布：

x \ y	y_1	y_2	
x_1	2/3	12/15	4/5
x_1	3/20	1/20	1/5
	49/60	11/60	1/5

$$\begin{aligned} H(XY)&=H\left(\frac{2}{3},\frac{2}{15},\frac{3}{20},\frac{1}{20}\right)\\ &=1.404\,bit/符号\\ &=0.973\,nat/符号\\ &=0.423\,hart/符号 \end{aligned}$$

$$H(Y)=H(49/60,11/60)=0.687bit/符号=0.476nat/符号=0.207hart/符号$$

$$H(X\mid Y)=H(XY)-H(Y)=0.717bit/符号=0.497nat/符号=0.216hart/符号$$

(4)

$$H(Y\mid X)=H(XY)-H(X)=0.682bit/符号=0.473nat/符号=0.205hart/符号$$

(5)

$$I(X;Y)=H(X)-H(X|Y)=0.00504bit/\text{符号}=0.00349nat/\text{符号}=0.00152hart/\text{符号}$$

( )

1、 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。
2、 信息的 可度量性 是建立信息论的基础。
3、 统计度量 是信息度量最常用的方法。
4、 熵 是香农信息论最基本最重要的概念。
。
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，$H_\infty=\lim\limits_{N\to\infty}H(X_N/X_1X_2\cdots X_{N-1})$。
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。
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：$\sum\limits_{i=1}^{n}m^{-k_i}\le 1$。
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 $\log_2 18$（$1+2\log_2 3$）。
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= $\log_2 n-H_{ni}$ 。
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 。
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 。
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）= $\begin{cases}0 & i=j\\ 1 & i\ne j\end{cases}$。
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= $\left\{p(y_j/x_i):\overline{D}\le D;i=1,2,\cdots,n,j=1,2,\cdots,m\right\}$ 。
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，•••，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）= $\underset{p(y/x)\in P_D}{Inf}\ I(X;Y)$ 。

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 0011100000011111001111000001111111，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 $\eta_0>\eta>\eta_1$ 。
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。