

北方民族大学试卷

课程代码: 01100622 课程: 信息理论及编码 B 卷答案

说明: 此卷为《信息理论及编码》B 卷答案

一、概念简答题 (每小题 6 分, 共 30 分)

1、比较平均自信息 (信源熵) 与平均互信息的异同。

答: 平均自信息为 $H(X) = -\sum_{i=1}^q P(a_i) \log P(a_i)$, 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。.....(3 分)

平均互信息 $I(X; Y) = \sum_{x,y} P(xy) \log \frac{P(y/x)}{P(y)}$ 。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。.....(3 分)

2、简述香农第一定理。

答: 对于离散信源 S 进行 r 元编码, 只要其满足 $\frac{\bar{L}_N}{N} \geq \frac{H(s)}{\log r}$,(3 分)

当 N 足够长, 总可以实现无失真编码。.....(3 分)

3、简述唯一可译变长码的判断方法?

答: 将码 C 中所有可能的尾随后缀组成一个集合 F , 当且仅当集合 F 中没有包含任一码字时, 码 C 为唯一可译变长码。构成集合 F 的方法:(2 分)

首先, 观察码 C 中最短的码字是否是其他码字的前缀。若是, 将其所有可能的尾随后缀排列出。而这些尾随后缀又可能是某些码字的前缀, 再将由这些尾随后缀产生的新的尾随后缀列出。依此下去, 直至没有一个尾随后缀是码字的前缀或没有新的尾随后缀产生为止。.....(2 分)

接着, 按照上述步骤将次短的码字直至所有码字可能产生的尾随后缀全部列出, 得到尾随后缀集合 F 。.....(2 分)

4、简述最大离散熵定理。

答：最大离散熵定理为：对于离散无记忆信源，当信源等概率分布时熵最大。……(3 分)

对于有 m 个符号的离散信源，其最大熵为 $\log m$ 。……(3 分)

5、什么是汉明距离：两个二元序列 $\alpha_i = 1230210, \beta_j = 0210210$ ，求其汉明距离。

答：长度相同的两个码字之间对应位置上不同的码元的个数，称为汉明距离。……(3 分)

$$D(\alpha_i, \beta_j) = 5 \dots\dots\dots (3 \text{ 分})$$

二、判断题（每小题 2 分，共 10 分）

1、信息是事物运动状态或存在方式的不确定性的描述。 (☒)

2、当信源与信道连接时，若信息传输率达到了信道容量，则称此信源与信道达到匹配。 (☒)

3、二元霍夫曼码一定是最佳即时码。 (☒)

4、单符号离散信源的自信息和信源熵都是一个确定值。 (☒)

5、平均错误概率只与译码准则有关系。 (☒)

三、计算题（每小题 10 分，共 40 分）

1、设离散无记忆信源 $\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} a_1=0 & a_2=1 & a_3=2 & a_4=3 \\ \frac{3}{8} & \frac{1}{8} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$ 。其发生的消息为

(202123021302201320200231201320)，求：(1) 此消息的自信息是多少？(2) 在此消息中平均每个符号携带的自信息是多少？

答：(1) 自信息量 $I(a) = -9\log P(0) - 5\log P(1) - 11\log P(2) - 5\log P(3) \dots\dots\dots (5 \text{ 分})$
 $\approx 59.74(\text{比特})$

$$(2) I_2 = \frac{59.74}{30} \approx 1.99(\text{比特/符号}) \dots\dots\dots (5 \text{ 分})$$

2、计算以下 2 个信道的信道容量 C:

$$P_1 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{3} \end{bmatrix}, \quad P_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

答： P_1 为准对称信道，其信道容量 $C_1 = \log r - H(p'_1, p'_2, \dots, p'_s) - \sum_{k=1}^n N_k \log M_k \dots\dots\dots (2 \text{ 分})$

其中 3 个子矩阵为 $B_1 = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} \end{bmatrix}$, $B_2 = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$, $B_3 = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{6} \end{bmatrix}$, 所以 $N_1 = \frac{1}{2}, N_2 = \frac{1}{3}, N_3 = \frac{1}{6}$

$$M_1 = \frac{1}{2}, M_2 = \frac{2}{3}, M_3 = \frac{1}{3},$$

$$\text{则 } C_1 = \log 2 - H\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right) - \left(\frac{1}{2} \log \frac{1}{2} + \frac{1}{3} \log \frac{2}{3} + \frac{1}{6} \log \frac{1}{3}\right) = 0.0407(\text{bit} / \text{symbol}) \cdots \cdots (4 \text{ 分})$$

$$P_2 \text{ 为一一对应信道, 其信道容量 } C_2 = \max H(X) = \log_2 4 = 2(\text{bit} / \text{symbol}) \cdots \cdots (4 \text{ 分})$$

3、信源空间为 $\begin{bmatrix} S \\ P(s) \end{bmatrix} = \begin{bmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ 0.4 & 0.2 & 0.1 & 0.1 & 0.05 & 0.05 & 0.05 & 0.05 \end{bmatrix}$, 试构造二元霍夫曼码, 并

计算其平均码长和编码效率。

答: 霍夫曼编码如下表所示

s_i	概率	w_i (1 上 0 下)	w_i (0 上 1 下)	l_i
s1	0.4	11	00	2
s2	0.2	00	11	2
s3	0.1	011	100	3
s4	0.1	010	101	3
s5	0.05	1011	0100	4
s6	0.05	1010	0101	4
s7	0.05	1001	0110	4
s8	0.05	1000	0111	4

.....(6 分)

$$\text{平均码长 } \bar{L} = 0.05 \times 4 \times 4 + 0.1 \times 3 \times 2 + 0.2 \times 2 + 0.4 \times 2 = 2.6(\text{码符号} / \text{信源符号}) \cdots \cdots (2 \text{ 分})$$

$$\text{编 码 效 率 } \eta = \frac{H(S)}{\bar{L}} = (0.4 \log \frac{1}{0.4} + 0.2 \log \frac{1}{0.2} + 2 \times 0.1 \log \frac{1}{0.1} + 4 \times 0.05 \log \frac{1}{0.05}) / 2.6 = 0.97 \quad (2$$

分)

4、设某二元码为 $C = \{00000 \quad 01101 \quad 10111 \quad 11010\}$ 。(1) 计算此码的最小距离 d_{\min} ; (2) 计算此码的码率 R , 假设码字等概率分布; (3) 采用最小距离译码准则, 接收到序列 00100, 01010, 00111 和 00101 应译成什么码字。

答: (1) $d_{\min} = 3 \cdots \cdots (3 \text{ 分})$

$$(2) \quad R = \frac{\log m}{n} = \frac{\log 4}{5} = 0.4 \left(\frac{\text{比特}}{\text{码符号}} \right) \dots\dots\dots (3 \text{ 分})$$

(3) 序列 00100 译成 00000, $\dots\dots\dots$ (1 分)

序列 01010 译成 11010, $\dots\dots\dots$ (1 分)

序列 00111 译成 10111, $\dots\dots\dots$ (1 分)

序列 00101 译成 01101 $\dots\dots\dots$ (1 分)

四、综合题 (20 分)

1、二元对称信道的信道矩阵为 $P = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$, 信道传输速度为 1500 二元符号/秒, 设信源为等概率

分布, 信源消息序列共有 13000 个二元符号, 问:

(1) 试计算能否在 10 秒内将信源消息序列无失真传送完?

(2) 若信源概率分布为 $P(0) = 0.7, P(1) = 0.3$, 求无失真传送以上信源消息序列至少需要多长时间?

答: (1) 此信道为二元对称信道, 信源等概分布时, 信道传输率达最大 C 。信源熵 $H(S) = \log 2 = 1(\text{bit} / \text{symbol}) \dots\dots\dots$ (1 分)

$$C = 1 - H(P) = 1 - 0.9 \log_2 0.9 - 0.1 \log_2 0.1 = 1 - 0.469 = 0.531(\text{bit} / \text{symbol}) \dots\dots\dots (1 \text{ 分})$$

$$1500 \times 10 \times 0.531 = 7965(\text{bit}) < 13000 \times 1(\text{bit}), \text{ 所以 } 10 \text{ 秒内不能传完。} \dots\dots\dots (3 \text{ 分})$$

(2) 信源不等概分布, 信息传输率达不到 C

$$H(S) = 0.7 \log \frac{1}{0.7} + 0.3 \log \frac{1}{0.3} = 0.3602 + 0.52109 = 0.88129(\text{bit} / \text{symbol}) \dots\dots\dots (1 \text{ 分})$$

$$130000 \text{ 个信源符号所含信息量为: } 130000 \times H(S) = 11456.77(\text{bit}) \dots\dots\dots (2 \text{ 分})$$

$$P(y) = \begin{bmatrix} 0.7 & 0.3 \end{bmatrix} \times \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} = \begin{bmatrix} 0.66 & 0.34 \end{bmatrix} \dots\dots\dots (1 \text{ 分})$$

$$P(xy) = \begin{bmatrix} 0.9 \times 0.7 & 0.1 \times 0.7 \\ 0.1 \times 0.3 & 0.9 \times 0.3 \end{bmatrix} = \begin{bmatrix} 0.63 & 0.07 \\ 0.03 & 0.27 \end{bmatrix} \dots\dots\dots (1 \text{ 分})$$

信息传输率:

$$\begin{aligned} R = I(X; Y) &= \sum_{xy} P(x_i y_j) \log_2 \frac{P(y_j / x_i)}{P(y_j)} \\ &= 0.63 \log \frac{0.9}{0.66} + 0.07 \log \frac{0.1}{0.34} + 0.03 \log \frac{0.1}{0.66} + 0.27 \log \frac{0.9}{0.34} = 0.4558(\text{bit} / \text{symbol}) \end{aligned} \dots\dots\dots (5 \text{ 分})$$

要无失真传送 130000 个二元符号的信源消息序列至少需要的时间:

$$T \geq \frac{11456.77}{0.4558 \times 1500} = 16.757s \dots\dots\dots(5 \text{ 分})$$