

21 世纪高等学校规划教材·电子信息

信息论与编码技术（MATLAB 实现）

习

题

详

解

朱春华 编著

清华大学出版社

北 京

第一章

1. 信息、消息和信号的区别与联系是什么？

答：消息(message)指的是信号要传递的内容，是本质。指文字、符号、数据、语言、图像等能够被人们感觉器官所感知的内容。信号(signal)是消息传递的形式，是把消息转换成适合信道传输的物理量，比如是电信号、光信号等，是载体。信息(information)是抽象的，能消除受信者的某些不确定性。

消息是信息的形式，消息中包含信息，信息是消息的内容，信号则是消息的表现形式，是消息的运载工具。

2. 通信系统的组成与各部分作用是什么？

通信系统主要由五部分组成：

(1) 信源：信源是产生消息(或消息序列)的源（人或机器）。

(2) 编码器：编码器的作用是把消息转换成信号，并将信源和信道进行匹配。可分为信源编码器和信道编码器，编码器还应包括换能、调制、发射等各种变换处理。信源编码器是对信源输出的消息进行适当的变换和处理，具体包含两种形式：将信源发出的消息变换为二进制码元或多进制码元组成的代码组（数字序列，通常为二进制数字序列）；通过信源编码压缩信源的冗余度（即剩余度），以提高通信系统的传输效率。信道编码要在信源编码器输出的数据码流中有目的地增加一些监督码元，从而达到在接收端进行判错和纠错的目的，增加通信的可靠性。除了信源编码和信道编码，为了提高信息在传输和处理过程中的安全保密问题，还需要进行加密编码。

(3) 信道：信道是指传输信号的媒质或通道，如架空明线、电线、射频波束、人造卫星等。根据噪声和干扰的统计特性，信道有多种模型。

(4) 译码器：译码就是把信道输出的编码信号进行反变换，从编码信号中最大程度地提取有关信源输出的信息。译码器也分为信源译码、信道译码和解密。

(5) 信宿：信宿是消息传送的对象，即接收信息的人或机器。

3. 通信系统的指标都有什么，如何提高？

答：通信系统的指标：有效性，可靠性，安全性。

通过信源编码提高有效性；通过信道编码提高可靠性；通过加密编码提高安全性。

4. 通信过程获得的信息量与不确定性、概率之间的关系？

答：通信以前，收信者对发送的信息存在“疑义”和“不知”，即不确定性。数学上，不确定性就是随机性，具有不确定性的事件就是随机事件。事件或消息发生的概率越小，不确定性越大，所以不确定性与概率呈反比。通信过程中，通信前接收者对发送的消息的不确定性减去通信后收信者还存在的 uncertainty，就是收信者收到该消息后获得的信息量。

第二章

1. 设有一个 4 行 8 列的棋形方格, 有一个质点 A 分别以等概率落入任一方格内, 求 A 落入任一小格的自信息量为多少?

解: $p(A) = \frac{1}{32}$

$$I(A) = -\log p(A) = \log 32 = 5 \quad \text{bit}$$

2. 若某单符号离散无记忆信源的数学模型为

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$$

该信源 10 秒内发出的消息为(11110010110011000011), 求

- (1) 此消息的自信息量是多少?
(2) 此消息中每个符号携带的信息量是多少?

解: 消息序列中, “0” 个数为 $n_1=9$, “1” 个数为 $n_2=11$.

消息序列总长为 $N = n_1 + n_2 = 20$ (个符号)

- (1) 消息序列的自信息量:

$$I = \sum_{i=1}^2 n_i I(x_i) = -n_1 \log p(0) - n_2 \log p(1) = 9 \log 4 + 11 \log \frac{4}{3} = 22.565$$

- (2) 平均每个符号携带的信息量为:

$$\frac{I}{N} = \frac{22.5650}{20} = 1.1283 \quad \text{bit}$$

3. 设有一非均匀骰子, 若其任一面出现的概率与该面上的点数成正比,

- (1) 分别求各点出现时所给出的信息量;
(2) 求扔一次平均得到的信息量。

解: $p(i) = \frac{i}{21}$

$$(1) \quad I(1) = -\log \frac{1}{21} = 4.3923 \quad I(2) = -\log \frac{2}{21} = 3.3923 \quad I(3) = -\log \frac{3}{21} = 2.8074$$

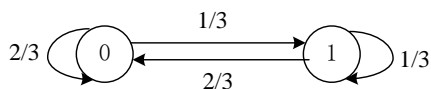
$$I(4) = -\log \frac{4}{21} = 2.3923 \quad I(5) = -\log \frac{5}{21} = 2.0704 \quad I(6) = -\log \frac{6}{21} = 1.8074$$

- (2) 扔一次平均得到的信息量

$$H(X) = H\left(\frac{1}{21}, \frac{2}{21}, \frac{3}{21}, \frac{4}{21}, \frac{5}{21}, \frac{6}{21}\right) = 2.3983$$

4. 若二元一阶平稳马尔可夫信源的符号转移概率为 $p(1/0) = \frac{1}{3}$, $p(1/1) = \frac{1}{3}$, 画出马尔可夫状态图, 并求状态平稳分布概率。

解: (1) 马尔可夫状态图



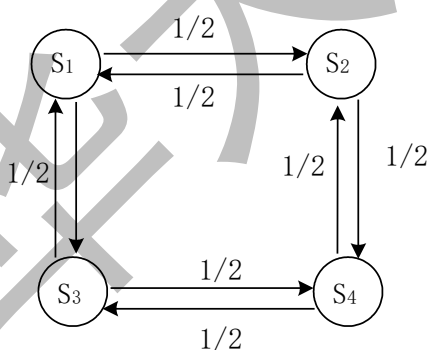
(2) 马尔可夫信源状态概率为 w_1 和 w_2 , 则

$$\begin{cases} w_j = \sum_{i=1}^2 w_i p_{ij} \\ \sum_{j=1}^2 w_j = 1 \end{cases}$$

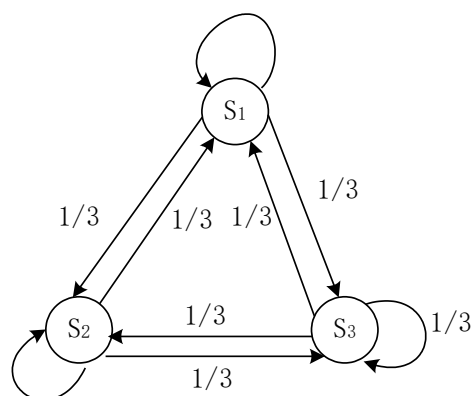
代入状态转移概率,
$$\begin{cases} w_0 = \frac{2}{3} w_0 + \frac{2}{3} w_1 \\ w_1 = \frac{1}{3} w_0 + \frac{1}{3} w_1 \\ w_0 + w_1 = 1 \end{cases}$$

解得 $w_0 = \frac{2}{3}$ $w_1 = \frac{1}{3}$

5. 某马尔可夫信源状态图如图 2-1(a)(b),



(a) 四状态马氏链



(b) 三状态马氏链

图 2-1 习题 5 马尔可夫信源状态图

分别判断图 2-1(a)(b)所代表的四状态马尔可夫信源和三状态马尔可夫信源是否具有各态遍历性。

解: 一个不可约的、非周期的、状态有限的马尔可夫信源具有各态遍历性。

图(a)中, $P_{ii}^{(k)} > 0$ 的 $k=2, 4, 6, 8 \dots$ 是周期的马尔可夫链

$P_{ij}^{(k)} > 0$ 成立, 满足不可约性

因此, 图(a) 四状态马尔可夫信源不具有各态历经性。

图(b)中, $P_{ii}^{(k)} > 0$ 的 $k=1, 2, 3, 4, \dots$ 满足周期性

$P_{ij}^{(k)} > 0$ 成立, 满足不可约性

因此, 图(b) 三状态马尔可夫信源具有各态历经性。

6. 一个两状态的一阶马尔可夫链, 其状态转移图为:

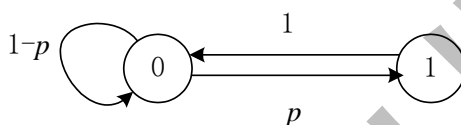


图 2-2 习题 6 状态转移图

(1) 写出其一步转移概率矩阵;

(2) 求信源平稳后的概率分布 $p(0), p(1)$;

(3) p 取何值时, H_∞ 最大? 用 matlab 程序画出 H_∞ 随 p 变化的曲线。

解: (1) $P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 1-p & p \\ 1 & 0 \end{bmatrix}$

(2)
$$\begin{cases} w_j = \sum_{i=1}^2 w_i p_{ij} \\ \sum_{j=1}^2 w_j = 1 \end{cases} \quad \text{解得} \quad w_1 = \frac{1}{p+1} \quad w_2 = \frac{p}{p+1}$$

即 $p(0) = \frac{1}{p+1} \quad p(1) = \frac{p}{p+1}$

(3) $H_\infty = \sum_{i=1}^2 w_i H(X / s_i)$

$$H(X / s_i) = - \sum_{j=1}^2 p(x_j / s_i) \log p(x_j / s_i) = -p_{i1} \log p_{i1} - p_{i2} \log p_{i2}$$

所以

$$H_\infty = w_1 H(X / s_1) + w_2 H(X / s_2) = w_1 H(p_{11}, p_{12}) + w_2 H(p_{21}, p_{22})$$

$$= w_1 H(1-p, p) + w_2 H(1, 0) = w_1 H(1-p, p)$$

$$= \frac{1}{p+1} [-(1-p)\log(1-p) - p\log p]$$

$$\text{当 } p=1/2 \text{ 时, } H_{\infty} \text{ 最大, } H_{\infty} = \frac{2}{3}$$

7. 设某一信源输出信号 X , 在传输过程中受到噪声 Z 的影响, 导致收到的信号为 $Y=X+Z$ 。其中 X 与 Z 为同分布的高斯随机变量, 均值为 0, 方差为 σ^2 的高斯分布, X 与 Z 独立。

(1) 求 X 的相对熵;

(2) 求 Y 的方差及相对熵。

解: (1) X 为标准正态分布, 其相对熵为

$$H(X) = \log \sqrt{2\pi e \sigma^2}$$

(2) Y 为标准正态分布, 其均值为 0, 方差为 $2\sigma^2$, Y 的相对熵为

$$H(Y) = \log \sqrt{4\pi e \sigma^2}$$

8. 有一信源发出恒定宽度, 但不同幅度的脉冲, 幅度值 x 处于 a_1 和 a_2 之间, 此信源连至某信道接收端, 接收脉冲的幅度 y 处于 b_1 和 b_2 之间, 已知随机变量 x 和 y 相互独立, 计算 $H(X)$, $H(Y)$, $H(XY)$ 和 $I(X;Y)$ 。

解: 随机变量 x 的概率密度为:

$$p_X(x) = \frac{1}{(a_2 - a_1)} \quad a_1 \leq x \leq a_2$$

随机变量 y 的概率密度为:

$$p_Y(y) = \frac{1}{(b_2 - b_1)} \quad b_1 \leq y \leq b_2$$

随机变量 x 和 y 的联合概率密度为

$$p_{XY}(xy) = \frac{1}{(a_2 - a_1)(b_2 - b_1)}$$

$$\text{则 } H(X) = \log(a_2 - a_1)$$

$$H(Y) = \log(b_2 - b_1)$$

$$H(XY) = \log[(a_2 - a_1)(b_2 - b_1)]$$

由于 x 和 y 相互独立, 所以

$$I(X;Y) = 0$$

9. X 表示掷骰子试验,

(1) 如果骰子是正常的, 求熵 $H_n(X)$;

(2) 如果骰子有一面不正常, 且其出现的概率为 $\frac{1}{2}$, 求熵 $H_u(X)$ 。

解: (1) $H_n(X) = \log_2 6 = 2.585$

$$(2) H_u(X) = H\left(\frac{1}{2}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}\right) = \frac{1}{2} \log 2 + 5 \times \frac{1}{10} \log 10 = 2.1610$$

10. 单符号离散信源 X 和 Y 的数学模型为

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0.99 & 0.01 \end{bmatrix}, \quad \begin{bmatrix} Y \\ P \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

编写 matlab 程序比较信源 X 和 Y 的熵及信源冗余度。

解: 信源 X 和 Y 的熵分别为

$$H(X) = -0.99 \log 0.99 - 0.01 \log 0.01 = 0.08(\text{bit/s})$$

$$H(Y) = -0.5 \log 0.5 - 0.5 \log 0.5 = 1(\text{bit/s})$$

可见 $H(Y) > H(X)$, 即信源 Y 的平均不确定度要大。

11. 黑白气象传真图的消息只有黑色和白色两种, 即信源 $X = \{\text{黑}, \text{白}\}$ 。设黑色出现的概率为 $P(\text{黑}) = 0.4$, 白色的出现概率 $P(\text{白}) = 0.6$ 。

(1) 假设图上黑白消息出现前后没有关联, 求熵 $H(X)$;

(2) 假设消息前后有关联, 其依赖关系为 $P(\text{白}/\text{白}) = 0.8$, $P(\text{黑}/\text{白}) = 0.2$, $P(\text{白}/\text{黑}) = 0.3$, $P(\text{黑}/\text{黑}) = 0.7$, 求此一阶马尔可夫信源的熵 $H_2(X)$;

(3) 分别求上述两种信源的剩余度, 比较 $H(X)$ 和 $H_2(X)$ 的大小, 并说明其物理意义。

解: (1) 信源模型为 $\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} a_1 = \text{黑} & a_2 = \text{白} \\ 0.4 & 0.6 \end{bmatrix}$

$$H(X) = -\sum_{i=1}^2 p(a_i) \log p(a_i) = 0.971 \text{ bit/符号}$$

(2) 求稳态概率 $W_i = p(s_i)$, 可写出状态转移概率矩阵

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} 0.7 & 0.3 \\ 0.2 & 0.8 \end{bmatrix}$$

$$\text{计算方程: } \begin{cases} \sum_{j=1}^2 W_j = 1 \\ \sum_{i=1}^2 W_i P_{ij} = W_j \end{cases} \quad \text{即: } \begin{cases} W_1 = W_1 P_{11} + W_2 P_{21} \\ W_2 = W_1 P_{12} + W_2 P_{22} \\ W_1 + W_2 = 1 \end{cases} \quad \begin{cases} W_1 = 0.7W_1 + 0.2W_2 \\ W_2 = 0.3W_1 + 0.8W_2 \\ W_1 + W_2 = 1 \end{cases}$$

$$\text{得 } W_2 = P(\text{白}) = 0.6 \quad W_1 = P(\text{黑}) = 0.4$$

一阶马尔可夫信源的熵为

$$\begin{aligned} H_2 &= \sum_{i=1}^2 p(s_i) H(X/s_i) = W_1 H(P_{11}, P_{12}) + W_2 H(P_{21}, P_{22}) \\ &= 0.4 \times 0.8813 + 0.6 \times 0.7219 = 0.7856 \end{aligned}$$

$$(3) \text{ 信源(1) 的冗余度为 } \gamma = 1 - \frac{H(X)}{H_{\max}} = 1 - \frac{0.971}{\log 2} = 1 - 0.971 = 0.029$$

$$\text{信源(2) 的冗余度为 } \gamma = 1 - \frac{H_2(X)}{H_{\max}} = 1 - 0.786 = 0.214$$

这种结果的原因是信源 2 是有记忆信源，信源符号间的相关性造成了信源冗余度较大。信源 1 的符号不是等概分布，也造成了冗余度。

12. 一平稳信源，它在任意时间，不论以前发出过什么符号，都按 $p(a_1) = 1/4$, $p(a_2) = 1/4$,

$p(a_3) = 1/2$ 发出符号，

(1) 判断该信源类型；

(2) 求信源 X 的二次扩展信源 $X^2 = (X_1 X_2)$ 的序列熵 $H(X^2)$ ，并写出 X^2 信源中可能有的所有符号序列；

(3) $H(X^2)$ 与 $H(X_1)$ 和 $H(X_2)$ 具有什么关系？

(4) 求 $H(X_3/X_1 X_2)$ 。

解：

(1) 该信源为单符号无记忆平稳信源。

$$(2) H(X) = H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right) = \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 + \frac{1}{4} \log_2 4 = \frac{3}{2} \text{ 比特/符号}$$

$$H(X^2) = 2H(X) = 3 \text{ 比特/符号}$$

$$X^2 = \{a_1 a_1, a_1 a_3, a_1 a_3, a_2 a_1, a_2 a_2, a_2 a_3, a_3 a_1, a_3 a_2, a_3 a_3\}$$

$$(3) H(X^2) = H(X_1) + H(X_2)$$

$$(4) H(X_3 / X_1 X_2) = H(X) = \frac{3}{2} \text{ 比特/符号}$$

13. 每帧电视图像可以认为是由 3×10^5 个像素组成，每个像素均是独立变化，若每个像素可取 128 个不同的亮度电平并设亮度电平等概出现。

问每帧图像含有多少信息量？若有一广播员在约 10000 个汉字的字中选出 1000 个字来口述电视图像，试问广播员描述此图像所广播的信息量是多少（假设字汇是等概率分布的，并彼此无依赖）？若要适当的描述此图像，广播员在口述中至少需要多少汉字？

解：

∵ 亮度电平等概率出现

∴ 每个像素所含的信息量为 $H(X) = \log 128 = 7 \text{ bit/像素}$ 。

而每个像素均是独立变化的

∴ 每帧电视图像所包含的信息量为 $H(X) = 3 \times 10^5 H(X) = 2.1 \times 10^6 \text{ bit}$

∴ 假设汉字字汇是等概率分布

∴ 每个汉字出现的概率均为 $\frac{1}{10000}$

从而每个汉字携带的信息量为 $\log 10000 = 13.2877 \text{ bit/字}$

∴ 汉字间彼此无依赖，广播员口述的 1000 个汉字所广播的信息量为

$$1000 \times 13.2877 = 13287.7 \text{ bit}$$

若要恰当地描述图像，广播员在口述中至少需要的汉字数为 $\frac{2.1 \times 10^6}{13.2877} \approx 158041$ 个汉字。

14. 为了传输一个有字母 A, B, C, D 组成的符号集，把每个字母编码成有两个二进制码组成的脉冲序列，以 00 代表 A，01 代表 B，10 代表 C，11 代表 D。每个二进制码脉冲宽度为 5ms。

(1) 不同字母等概率出现时，计算传输的平均信息速率。

(2) 若每个字母出现的概率分别为 $\{1/5, 1/4, 1/4, 3/10\}$ ，试计算传输的平均信息速率？

解：(1) 由题可知，当不同字母等概率出现时，平均自信息量为：

$$H(x) = \log 4 = 2 \text{ (比特/字母)}$$

又因为每个二进制脉冲宽度为 5ms，故一个字母的脉冲宽度为 10ms，则字母的传输速率为 100 字母/秒，故传输的平均信息速率为：200 比特/秒

(2) 当每个字母分别以题中的概率出现时，平均自信息量为：

$$H(x) = - \sum P(a_i) \log P(a_i) = (1/5) * \log 5 + 2 * (1/4) * \log 4 + (3/10) * \log (10/3) = 1.98 \text{ (比特/字母)}$$

同样字母的传输速率为 100 个/秒，故传输的平均信息速率为：198 比特/秒。

15. 某一无记忆信源的符号集为{0,1}，已知 $p(0) = 1/4$ ， $p(1) = 3/4$ 。

(1) 求信源的熵。

(2) 由100个符号构成的序列，求某一特定的序列（例如有 m 个 0 和 $100-m$ 个 1）的自信息量表达式。

(3) 计算 (2) 中序列的熵。

$$\text{解: (1) } H(X) = - \sum_i \log p(x_i) = - \left(\frac{1}{4} \log \frac{1}{4} + \frac{3}{4} \log \frac{3}{4} \right) = 0.811 \text{ bit}$$

$$(2) \quad p(x_i) = \left(\frac{1}{4} \right)^m \times \left(\frac{3}{4} \right)^{100-m} = \frac{3^{100-m}}{4^{100}}$$

$$I(x_i) = -\log(x_i) = -\log \frac{3^{100-m}}{4^{100}} = 41.5 + 1.585m \text{ bit}$$

$$(3) \quad H(X^{100}) = 100H(X) = 100 \times 0.811 = 81.1 \text{ bit}$$

第三章

1. 在一个二进制信道中, 信源消息集 $X = \{0,1\}$ 且 $P(1) = P(0)$, 信宿的消息集 $Y = \{0,1\}$, 信道传

输概率 $p(y=1|x=0) = 1/4$, $p(y=0|x=1) = 1/8$. 求:

(1) 在接收端收到 $y=0$ 后, 所提供的关于传输消息 x 的平均条件互信息 $I(X; y=0)$;

(2) 该情况下所能提供的平均互信息量 $I(X; Y)$ 。

解: (1) $X=\{0,1\}, Y=\{0,1\}$

$$p(x_0) = p(x_1), p(y_1/x_0) = 1/4, p(y_0/x_1) = 1/8$$

$$\because p(x_0) + p(x_1) = 1, p(x_0) = p(x_1) \Rightarrow p(x_0) = p(x_1) = 1/2$$

$$p(y_0/x_0) = 1 - p(y_1/x_0) = 3/4$$

$$p(y_1/x_1) = 1 - p(y_0/x_1) = 7/8$$

$$P(x_0 y_0) = P(y_0/x_0)P(x_0) = \frac{3}{8} \quad P(x_0 y_1) = P(y_1/x_0)P(x_0) = \frac{1}{8}$$

$$P(x_1 y_0) = P(y_0/x_1)P(x_1) = \frac{1}{16} \quad P(x_1 y_1) = P(y_1/x_1)P(x_1) = \frac{7}{16}$$

$$p(y_j) = \sum_{i=0}^1 p(x_i y_j)$$

$$P(y_0) = \sum_{i=0}^1 P(x_i y_0) = \frac{3}{8} + \frac{1}{16} = \frac{7}{16}, \quad P(y_1) = \frac{9}{16},$$

$$\text{所以 } P(y_j/x_i) = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{8} & \frac{7}{8} \end{pmatrix} \quad P(x_i/y_j) = \begin{pmatrix} \frac{3}{8} & \frac{1}{8} \\ \frac{1}{16} & \frac{7}{16} \end{pmatrix} \quad P(x_i/y_j) = \begin{pmatrix} \frac{6}{7} & \frac{1}{7} \\ \frac{2}{9} & \frac{7}{9} \end{pmatrix}$$

$$I(x_0; y_0) = \log \frac{p(y_0/x_0)}{p(y_0)} = \frac{3}{4} \times \log \frac{16}{7} = \frac{12}{7}$$

$$I(x_1; y_0) = \log \frac{p(y_0/x_1)}{p(y_0)} = \frac{1}{8} \times \log \frac{16}{7} = \frac{2}{7}$$

$$I(X; y_0) = P(x_0/y_0)I(x_0; y_0) + P(x_1/y_0)I(x_1; y_0) = \frac{6}{7} \times \log \frac{12}{7} + \frac{1}{7} \times \log \frac{2}{7} = 0.4083 \text{ bit/s}$$

$$I(X; y_1) = P(x_0/y_1)I(x_0; y_1) + P(x_1/y_1)I(x_1; y_1) = \frac{2}{9} \log \frac{14}{9} + \frac{7}{9} \log \frac{4}{9} = 0.2358 \text{ bit/s}$$

$$(2) I(X; Y) = \sum_{j=0}^1 P(y_j)I(X; y_j)$$

$$= P(y_0)I(X; y_0) + P(y_1)I(X; y_1) = \frac{7}{16} \times 0.4083 + \frac{9}{16} \times 0.2358 = 0.3113 \text{ 比特/符号}$$

2. 一信道的输入和输出分别为 X 和 Y , 其中 X 等概率取值为 $+1, -1$, $Y=X+Z$, Z 是均值为零, 方差为 σ^2 的高斯分布。

(1) 画出 Y 的概率分布 $P_Y(y)$ 与 σ^2 关系的曲线。

(2) 画出信道输入与输出之间的平均互信息 $I(X;Y)$ 与 σ^2 关系的曲线。

解: (1) Y 服从均值为 $x_i = \pm 1$, 方差为 $\sigma_Y^2 = \sigma_X^2 + \sigma_Z^2 = 1 + \sigma^2$ 的正态分布, 其概率密度函数为

$$P_Y(y) = \frac{1}{\sqrt{2\pi(1+\sigma^2)}} e^{-\frac{(y-x_i)^2}{2}}$$

画图略

(2)

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y/X) = H(Y) - H(Z) \\ &= \log 2\pi e \sigma_Y^2 - \log 2\pi e \sigma^2 = \log \left(1 + \frac{1}{\sigma^2} \right) \end{aligned}$$

3. 某二元信道如图 3.1 所示, 请编写程序画出 $p(x_0) = \frac{1}{2}$ 时互信息 $I(x_0; y_0)$ 和 $I(X;Y)$ 关于错误概率 p 的关系图, 并进行解释。

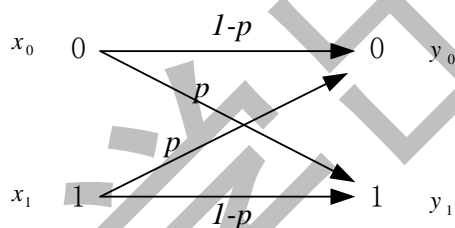


图 3.1 二元对称信道

解: 信道转移概率矩阵为 $P = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$

令 $p(x_0) = p_0$, $p(x_1) = p_1 = 1 - p_0$ 则, 联合概率矩阵为

$$P_{XY} = \begin{bmatrix} p_0(1-p) & p_0p \\ p_1p & p_1(1-p) \end{bmatrix}$$

$$p(y_0) = \sum_{i=0}^1 p(x_i y_0) = p_0(1-p) + p_1p = p_0 + p - 2p_0p$$

$$p(y_1) = \sum_{i=0}^1 p(x_i y_1) = p_0p + p_1(1-p) = 1 - p_0 - p + 2p_0p$$

$$I(x_0; y_0) = \log \frac{p(y_0/x_0)}{p(y_0)} = \log \frac{1-p}{p_0 + p - 2p_0p}$$

$$I(X;Y) = H(Y) - H(Y/X) = H(p_0 + p - 2p_0p, 1 - p_0 - p + 2p_0p) - H(1-p, p)$$

当 $p_0 = \frac{1}{2}$ 时,

$$I(x_0; y_0) = \log 2(1-p)$$

$$I(X; Y) = 1 - H(1-p, p)$$

4. 一离散无记忆信道转移概率图如图 3.1 所示, 信道输入、输出分别为 X, Y 。

- (1) 写出该信道的转移概率矩阵 \mathbf{P} 。
- (2) 求信道容量 C 。
- (3) 求达到容量时的输出概率分布。
- (4) 求达到容量时的输入概率分布。

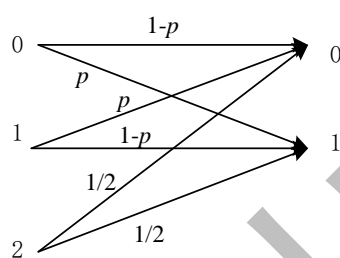


图 3.1

解: (1) $\mathbf{P} = \begin{bmatrix} 1-p & p \\ p & 1-p \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

(2) $I(0; Y) = \log 2 - H(p, 1-p)$

$$I(1; Y) = \log 2 - H(p, 1-p)$$

$$I(2; Y) = \log 2 - H\left(\frac{1}{2}, \frac{1}{2}\right) = 0$$

信道容量:

$$C = I(0; Y) = I(1; Y) = \log 2 - H(p, 1-p)$$

- (3) 输出等概。
- (4) 输入 0, 1 符号等概。输入 2 概率为零。

5. 二元对称信道转移矩阵为 $\mathbf{P} = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$, 其中

(1) 若输入 $p(x_0) = \frac{3}{4}$, $p(x_1) = \frac{1}{4}$, 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) p 取何值时, 信道容量达到最大值, 用 matlab 编程画出该信道的信道容量随 p 的变化曲线。

解: (1) $H(X) = H(\frac{3}{4}, \frac{1}{4}) = 0.81 \text{ bit/symbol}$

$$P_{XY} = \begin{bmatrix} \frac{3}{4}(1-p) & \frac{3}{4}p \\ \frac{1}{4}p & \frac{1}{4}(1-p) \end{bmatrix}$$

$$p(y_0) = \sum_{i=0}^1 p(x_i, y_0) = \frac{3}{4}(1-p) + \frac{1}{4}p = \frac{3}{4} - \frac{1}{2}p$$

$$p(y_1) = \sum_{i=0}^1 p(x_i, y_1) = \frac{3}{4}p + \frac{1}{4}(1-p) = \frac{1}{4} + \frac{1}{2}p$$

$$I(X;Y) = H(Y) - H(Y/X) = H(\frac{3}{4} - \frac{1}{2}p, \frac{1}{4} + \frac{1}{2}p) - H(1-p, p)$$

(2) $C = \log 2 - H(1-p, p)$, 当 $p=1$ 或 0 时, 信道容量达到最大值。此时输入符号等概率分布

6. 某信道的概率转移矩阵为

$$P_1 = \begin{bmatrix} 1/4 & 3/4 \\ 3/4 & 1/4 \end{bmatrix} \quad P_2 = \begin{bmatrix} 1/4 & 1/8 & 5/8 \\ 1/8 & 1/4 & 5/8 \end{bmatrix}$$

分别求上述两个信道的信道容量及其达到信道容量时的输入概率分布。

解: 信道 1 是对称的 DMC 信道, 其信道容量为

$$C = \log 2 - H(\frac{3}{4}, \frac{1}{4}) = 0.19 \text{ bit}$$

信道 2 是准对称的 DMC 信道, 可以分成两个子信道

$$\begin{bmatrix} 1/4 & 1/8 \\ 1/8 & 1/4 \end{bmatrix} \quad \begin{bmatrix} 5/8 \\ 5/8 \end{bmatrix}$$

其信道容量为

$$C = \log 2 - (\frac{1}{4} + \frac{1}{8}) \log(\frac{1}{4} + \frac{1}{8}) - \frac{5}{8} \log(\frac{5}{8} + \frac{5}{8}) = 1.3294$$

7. 一个高斯白噪声信道, 接收机前端的带通滤波器带宽为 1MHz, 信道上的信号与噪声的平均功率之比为 30.1dB, 求该信道的信道容量。

解: 香农公式为 $C_t = \lim_{T \rightarrow \infty} \frac{C}{T} = W \log_2 \left(1 + \frac{P}{N_0 W} \right) \text{ bit/s}$

$$\text{由 } 10 \lg \frac{P}{N_0 W} = 30.1 \text{ dB 得 } \frac{P}{N_0 W} = 1023$$

$$C_t = 1 \times 10^6 \log_2 (1 + 1023) = 10 \text{ Mbit/s}$$

8. 设有三个信道的概率转移矩阵分别为

$$P_{ij}^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad P_{ij}^{(2)} = \begin{bmatrix} 1/4 & 3/4 \\ 3/4 & 1/4 \end{bmatrix} \quad P_{ij}^{(3)} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

试比较上述三个信道的好坏。

解：信道 1 是无干扰 BSC 信道， $C = \log 2 = 1 \text{ bit/符号}$

信道 2 是有干扰 BSC 信道， $C = \log m - H(1/4, 3/4) = 1 - 0.81 = 0.19 \text{ bit/符号}$

信道 3 是全损 BSC 信道， $C = 0$ 比特/符号。

由以上信道容量对比可知，信道 1 最优，信道 2 次之，信道 3 最差。

9. 设某信源发送端符号集为 $X \in \{x_1, x_2\}$ ， $p(x_1) = a$ ，接收端符号集为 $Y = \{y_1, y_2, y_3\}$ ，信道转移

矩阵为 $P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 \end{pmatrix}$ ，求该信道的信道容量及其达到信道容量时的输入概率分布。

解：由 $P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 \end{pmatrix}$ 得联合概率矩阵

$$P_{XY} = \begin{pmatrix} \frac{1}{2}a & \frac{1}{2}a & 0 \\ \frac{1}{2}(1-a) & \frac{1}{4}(1-a) & \frac{1}{4}(1-a) \end{pmatrix}$$

$$p(y_0) = \sum_{i=0}^1 p(x_i y_0) = \frac{1}{2}a + \frac{1}{2}(1-a) = \frac{1}{2}$$

$$p(y_1) = \sum_{i=0}^1 p(x_i y_1) = \frac{1}{2}a + \frac{1}{4}(1-a) = \frac{1}{4}(1+a)$$

$$p(y_2) = \sum_{i=0}^1 p(x_i y_2) = \frac{1}{4}(1-a)$$

可以求得：

$$H(Y) = H\left(\frac{1}{2}, \frac{1}{4}(1+a), \frac{1}{4}(1-a)\right) = \frac{3}{2} - \frac{1}{4}(1+a)\log(1+a) - \frac{1}{4}(1-a)\log(1-a)$$

$$\begin{aligned} H(Y/X) &= \frac{1}{2}a \log 2 + \frac{1}{2}a \log 2 + \frac{1}{2}(1-a) \log 2 + \frac{1}{4}(1-a) \log 4 + \frac{1}{4}(1-a) \log 4 \\ &= \frac{3}{2} - \frac{1}{2}a \end{aligned}$$

因此

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y/X) = \frac{3}{2} - \frac{1}{4}(1+a)\log(1+a) - \frac{1}{4}(1-a)\log(1-a) - \frac{3}{2} + \frac{1}{2}a \\ &= \frac{1}{2}a - \frac{1}{4}(1+a)\log(1+a) - \frac{1}{4}(1-a)\log(1-a) \end{aligned}$$

$$\text{令 } \frac{\partial I(X;Y)}{\partial a} = 0 \quad \text{得} \quad \frac{1}{4} \cdot \frac{-\ln(1+a) + \ln(1-a) + 2\ln 2}{\ln 2} = 0$$

得

$$a = \frac{3}{5}$$

$$\text{所以, } C = \max_{p(x_i)} I(X;Y) = I(X;Y) \Big|_{a=\frac{3}{5}} = 0.16$$

10. 设某一信号的信息传输率为 5.6 kbit/s , 在带宽为 4 kHz 的高斯信道中传输, 噪声功率谱 $N_0=5 \times 10^{-9}\text{ W/Hz}$ 。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。

解: (1) 无差错传输时, 有

$$R \leq C = W \log_2 \left(1 + \frac{P}{N_0 W} \right)$$

即

$$5.6 \times 10^3 \leq 4 \times 10^3 \log_2 \left(1 + 5 \frac{P}{\times 10^{-9} \times 4 \times 10^3} \right)$$

得

$$P \geq 0.0328 \text{ mW}$$

(2) $P = 0.0328 \text{ mW}$ 时, 输入信号的最大熵

$$H_c = \log_2 \sqrt{2\pi e P} = -5.4$$

对应的输入概率密度函数为

$$p(x) = \frac{1}{\sqrt{0.206 \times 10^{-3}}} e^{-\frac{x^2}{0.0656 \times 10^{-3}}}$$

11. 一个二元对称信道, 其信道转移如图 3.2 所示, 其中 $P=0.1$ 。

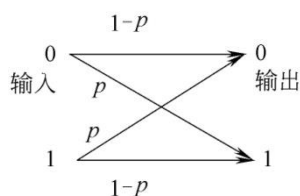


图 3.2 二元对称信道

(1) 设信道以 1800 个二元符号/秒的速度输入符号。现有一消息序列共有 11000 个二元符号, 符号间无统计关联性, 且每一符号取值概率分布 $p(x=0)=p(x=1)=1/2$, 从信息传输的角度考虑,

10 秒内能否将这消息序列无失真地传送完?

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

解: (1) $C = \log_2 \frac{1}{H(0.9, 0.1)} = 0.469 \text{ bit/symbol}$ 此时输入符号概率等概分布。

信源传输的消息序列的信息量为 $13000 \times \log_2 2 = 13000 \text{ bit}$

信道每秒传输的最大信息量为 $1800 \times C = 844.2 \text{ bit}$,

信道 10 秒内最大传输量为 $844.2 \times 10 = 8442 \text{ bit}$ 小于 11000 bit , 即 10 秒内信道传不完。

- (2) 信源传输的消息序列的信息量为 $11000 \times H(0.7, 0.3) = 11000 \times 0.8813 = 9694.3 \text{ bit}$

信道 10 秒内最大传输量为 8442 bit 小于 9694.3 bit , 即 10 秒内信道传不完。

无失真传输 9694.3 bit 需要的时间最少是:

$$9694.3 / 844.2 = 11.4834 \text{ 秒}$$

12. 一信道的转移概率如图 3.3 所示。

- (1) 求信道容量。
(2) 若将两个同样的信道串接, 求串接后信道的转移概率矩阵。
(3) 求(2)中串接信道的容量和达到容量时的输入概率分布。

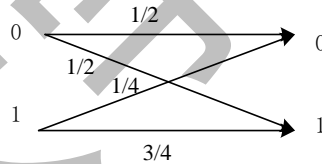


图 3.3

解: (1) 此信道转移概率矩阵为

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$$

信道容量 $C_1 = 0.0487 \text{ bit / 符号}$

- (3) 串接后的信道转移概率矩阵为

$$\mathbf{P} = \begin{bmatrix} \frac{3}{8} & \frac{5}{8} \\ \frac{5}{16} & \frac{11}{16} \end{bmatrix}$$

串接后信道的信道容量 $C_2 = 0.0033 \text{ bit / 符号}$

13. 电视图像由 30 万个像素组成, 对于适当的对比度, 一个像素可取 10 个可辨别的亮度电平。

假设各个像素的 10 个亮度电平都已等概率出现,实时传送的电视图像每秒发送 30 帧图像。
为了获得满意的图像质量,要求信号与噪声的平均功率比值为 30dB。试计算在这些条件下
传送电视的视频信号所需的带宽。

解: $H = \log_2 n = \log_2 10 = 3.322 \text{ bit/symbol}$

$I = NH = 3 \times 10^5 \times 3.322 = 9.966 \times 10^5 \text{ bit}$

$C_r = 9.966 \times 10^5 \times 30 = 2.99 \times 10^7 \text{ bit/s}$

$C_t = W \log(1 + P_X/P_N)$

$W = C_t / \log(1 + P_X/P_N) = 2.99 \times 10^7 / \log(1 + 1000) = 2.99 \text{ MHz}$

第四章

1. 信源 $X = \{0, 1, 2, 3, 4, 5\}$ ，编码输出 $Y = \{0, 1, 2\}$ ，若失真函数用输入输出符号幅度之差的

绝对值表示，即绝对失真： $d(x_i, y_j) = |x_i - y_j|$ ，试求失真矩阵。

解：

$$D = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \\ 4 & 3 & 2 \\ 5 & 4 & 3 \end{bmatrix}$$

2. 一个四元对称信源 $\begin{bmatrix} U \\ P(u) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 3 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$ ，接收符号为 $V = \{0, 1, 2, 3\}$ ，其失真函数为

$$d(u_i, v_j) = \begin{cases} 0 & i = j \\ 1 & i \neq j \end{cases}, \text{试求:}$$

- (1) 失真矩阵;
- (2) D_{\min} 和 D_{\max} 及 $R(D_{\min})$ 、 $R(D_{\max})$;
- (3) 信源的 $R(D)$ 函数，并编写 matlab 程序画出 $R(D)$ 函数曲线。

解：

- (1) 失真矩阵

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

- (2) 离散信源 $D_{\min} = 0$

$$R(D_{\max}) = 0$$

$$\begin{aligned}
 D_{\max} &= \min_{j=1,2,3,4} \sum_{i=1}^4 p_i d_{ij} \\
 &= \min_j \{p_1 d_{1j} + p_2 d_{2j} + p_3 d_{3j} + p_4 d_{4j}, p_1 d_{12} + p_2 d_{22} + p_3 d_{32} + p_4 d_{42}, \\
 &\quad p_1 d_{13} + p_2 d_{23} + p_3 d_{33} + p_4 d_{43}, p_1 d_{14} + p_2 d_{24} + p_3 d_{34} + p_4 d_{44}\} \\
 &= \min_j \left\{ \frac{1}{4}(0+1+1+1), \frac{1}{4}(1+0+1+1), \frac{1}{4}(1+1+0+1), \frac{1}{4}(1+1+1+0) \right\} \\
 &= \frac{3}{4}
 \end{aligned}$$

(3) 离散对称信源的信息率失真函数

$$R(D) = \log n - H(D, 1-D) - D \log(n-1)$$

带入 $n=4$ ，得四元对称信源的信息率失真函数为

$$R(D) = 2 - H(D, 1-D) - D \log 3$$

3. 对二元信源 $\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}$ ，其失真矩阵 $D = \begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix}$ ，求 $a>0$ 时信息率失真函数的 D_{\min}

和 D_{\max} 及 $R(D_{\min})$ 、 $R(D_{\max})$ ，并求选择何种信道可达到该 D_{\min} 和 D_{\max} 的失真（给出信道转移概率矩阵）。

解：离散信源 $D_{\min} = 0$

当 $D_{\min} = 0$ 时为无失真编码，信源符号与码符号一一对应，即

$$\begin{aligned}
 0 &\rightarrow 0 \\
 1 &\rightarrow 1
 \end{aligned}$$

达到该失真的假想信道转移概率矩阵为

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

$$\begin{aligned}
 D_{\max} &= \min_{j=0,1} \sum_{i=0}^1 p_i d_{ij} \\
 &= \min_j \{p_0 d_{00} + p_1 d_{10}, p_0 d_{01} + p_1 d_{11}\} \\
 &= \min_j \left\{ \frac{3}{4}(0+a), \frac{1}{4}(a+0) \right\} \\
 &= \frac{3}{4}a
 \end{aligned}$$

$j=0$ 时失真达到 D_{\max} ，即 $p_{i0} = p_i$ ，所有信源符号均对应码符号 0，即

$$\begin{array}{l} 0 \rightarrow 0 \\ 1 \nearrow \end{array}$$

达到该失真的假想信道转移概率矩阵为

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

则 $p(b_1)=1, p(b_2)=p(b_3)=p(b_4)=0$

4. 对于语音信源, 假设它的概率分布为高斯分布, 其信号功率为 σ^2 , 编码后失真为 D , 且失真度为平方失真: $d(x_i, y_j) = (x_i - y_j)^2$, 当数字电话中要求输入信噪比 σ^2/D 为 26dB 时, 试求它的最小传输信息率。

解:

$$\sigma^2/D = 10^{\frac{26}{10}} \approx 400$$

$$\text{高斯信源信息率失真函数 } R(D) = \frac{1}{2} \log \left(\frac{\sigma^2}{D} \right) = \frac{1}{2} \log_2 400 = 4.322 \text{ bit/符号}$$

实际的数字量化技术所采用的 A/D 芯片, 最少需要 8bit 采样量化。

5. 当连续信源呈现非高斯分布, 且其率失真函数难于求解时, 通常采用

$$H(X) - \log \sqrt{2\pi e D} \leq R(D) \leq \log \frac{\sigma}{\sqrt{D}}$$

进行保守估计。工程设计时则直接采用高斯分布代替非高斯分布。试编写 matlab 程序分析该近似估计的误差。

6. 一个二元信源 $\begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0.5 & 0.5 \end{bmatrix}$ 每秒钟发出 2.66 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), 二信道每秒钟传递二个二元符号。

- (1) 试问信源能否此信道中进行无失真的传输?
- (2) 若此信源失真度测定为汉明失真, 问允许信源平均失真多大时, 此信源就可以信道中传输?

解: (1) 信源 $\begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0.5 & 0.5 \end{bmatrix}$

其信源熵

$$H(X) = -\sum_{i=1}^2 p(x_i) \log p(x_i) = 1 \text{ 比特/符号}$$

而其每秒钟发 2.66 个信源符号，所以信源输出的信息速率为

$$R_t = 2.66H(X) \approx 2.66 \text{符号/秒} \times 1 \text{比特/符号} = 2.66 \text{比特/秒}$$

送入一个二元无噪无损信道，此信道的最大信息传输率(信道容量) $C=1$ 比特/符号。而信道每秒钟只传送两个二元符号，所以信道的最大信息传输速率

$$C_t = 2C \text{符号/秒} \times 1 \text{比特/符号} = 2 \text{比特/秒}$$

可见 $R_t > C_t$ ，根据无噪信道编码定理（即无失真信源编码定理），因为 $R_t > C_t$ ，所以不论进行任何编码此信源都不可能在此信道中实现无失真的传输，所以信源在此信道中传输会引起错误和失真。

(2) 若此信源失真测度定义为汉明失真，因为是二元信源，输入是等概分布，所以信源地信息率失真函数为

$$R(D) = 1 - H(D) \text{ 比特/信源符号}$$

$$R_t(D) = 2.66R(D) = 2.66(1 - H(D)) \text{ 比特/秒}$$

若 $C_t > R_t(D)$

则此信源在此信道中传输时不会引起错误。也就是不会因为信道而增加信源新的失真。总的信源的失真是信源压缩编码所造成的允许失真 D 。

所以有

$$C_t = R_t(D) = 2.66R(D) = 2.66(1 - H(D)) = 2$$

$$2.66H(D) = 0.66$$

$$H(D) = 0.2481$$

故 $D=0.0415$

允许信源平均失真为 $D=0.0415$ 时，此信源就可以在此信道中传输

7. 有一个二元、等概、平稳、无记忆信源 $X \in \{0,1\}$ ，接收符号集为 $Y \in \{0,1,2\}$ 且失真矩阵为

$$\mathbf{d} = \begin{bmatrix} 0 & \infty & 1 \\ \infty & 0 & 1 \end{bmatrix}$$

求率失真函数 $R(D)$ 。

解：

$$D_{\min} = \sum_x p(x) \min_y d(x, y) = 0$$

$$D_{\max} = 0 = \min_y \sum_x p(x) d(x, y) = 1$$

由于信源等概分布，失真函数具有对称性，因此，存在着与失真矩阵具有同样对称性的转移概率分布达到率失真 $R(D)$ ，该转移概率矩阵可写为

$$P = [p(y/x)] = \begin{bmatrix} \alpha & \beta & \gamma \\ \beta & \alpha & \gamma \end{bmatrix} \quad \alpha + \beta + \gamma = 1$$

由于 $d(0,1) = d(1,0) = \infty$ ，因此对于任何有限平均失真，必须 $\beta = 0$ ，于是转移概率矩阵为

$$P = [p(y/x)] = \begin{bmatrix} \alpha & 0 & \gamma \\ 0 & \alpha & \gamma \end{bmatrix} \quad \alpha + \gamma = 1$$

对应转移概率矩阵的平均失真：

$$D = \sum_x p(x) p(y/x) d(x, y) = 1 - \alpha$$

因此

$$\alpha = 1 - D$$

可求得此时的互信息为

$$\begin{aligned} R(D) &= I(X;Y) = H(Y) - H(Y/X) \\ &= H\left(\frac{1-D}{2}, D, \frac{1-D}{2}\right) - H(1-D, D) \\ &= 1 - D \quad 0 \leq D \leq 1 \end{aligned}$$

8. 有一个 n 元、等概率、平稳、无记忆信源 $x = \{0, 1, \dots, n\}$ ，接收符号集为 $Y = \{0, 1, \dots, n\}$ ，且

规定失真矩阵为

$$[d] = \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 \end{bmatrix}$$

求率失真函数 $R(D)$ 。

解：

$$D_{\min} = 0, \quad R(0) = H(X) = \log n$$

$$D_{\max} = \frac{n-1}{n}, \quad R(D_{\max}) = 0$$

任取一假想信道使 $\bar{D} = D$ ，该信道的信息传输率

$$R(D) = I(X;Y) = H(X) - H(X/Y) = \log n - H(X/Y)$$

费诺不等式给出了信道疑义度 $H(X/Y)$ 与信道平均差错概率的关系

$$H(X/Y) \leq H(P_e, 1-P_e) + P_e \log(n-1)$$

因此

$$I(X;Y) \geq \log n - H(P_e, 1-P_e) - P_e \log(n-1)$$

因 $P_e = D$ ，且当上式取等号时，对应 $R(D)$ 的最小值，也即满足失真度上限为 D 的信源编码方法的信息传输率下限为

$$R(D) = \log n - H(D, 1-D) - D \log(n-1) \quad)$$

其对应的假想信道为原假想信道的反向信道，

$$H(X/Y) = H(P_e, 1-P_e) + P_e \log(n-1)$$

反向信道与正向信道均为离散对称信道，其转移概率分布为

$$p(x_i/y_i) = 1-D$$

$$p(x_i/y_j) = \frac{D}{n-1}$$

为强对称信道。据此可得

$$p(y_i) = \frac{1}{n}$$

经计算，该反向信道中，平均失真度

$$\bar{D} = D$$

且满足 $\bar{D} \leq D$ 的信源的编码信息率下限，即信息率失真函数为

$$R(D) = \begin{cases} \log n - H(D, 1-D) - D \log(n-1) & 0 \leq D \leq \frac{n-1}{n} \\ 0 & D > \frac{n-1}{n} \end{cases}$$

第五章

1. 数据压缩的一个基本问题是“我们要压缩什么？”；你对此如何理解？

答：数据压缩，就是指不丢失有用信息的前提下，以最少的符号表示信源所发出的信号，减少容纳给定消息集合或数据采样集合的信号空间。所谓的信号空间就是压缩的对象，即物理空间，如存储器和 U 盘等数据存储介质；时间空间，如传输给定消息集合所需的时间；频带空间，如传输给定消息所要求的带宽等。衡量一种数据压缩技术的好坏有三个重要的指标：压缩比要大；恢复效果要好，要尽可能地恢复原始数据；实现压缩的算法要简单，压缩、解压缩速度快，尽可能做到实时压缩、解压。

2. 某信源 $X \in \{a_1, a_2, a_3, a_4\}$, 各符号概率为: $p(a_1) = 0.6$, $p(a_2) = 0.2$, $p(a_3) = 0.1$,

$p(a_4) = 0.1$ 。其对应的三组码 A、B、C 如下：

码组 A: 00 01 10 11

码组 B: 1 01 110 101

码组 C: 0 10 110 111

(1) 试判断这些码中哪些是唯一可译码，为什么？

(2) 对所有唯一可译码，求其平均码长。

解：(1) 码组 A 是等长码，各码字各不相同，是唯一可译码。

码组 B、C 的码长分别是 1, 2, 3, 3, $\sum_{i=1}^4 2^{-k_i} = 1$, 满足 kraft 不等式。

码组 B 的码字是前缀码，1 是 110, 101 的前缀，其后缀 10, 01 中 01 是码组中单独的码字，所以码组 B 不是唯一可译码。

码组 C 是异前缀码，一定是唯一可译码。

(2) 码组 A 是等长码，其码长恒等于 2。

码组 C 的平均码长：

$$K_c = \sum_{i=1}^4 k_i p(x_i) = 1 \times 0.6 + 2 \times 0.2 + 3 \times 0.1 + 3 \times 0.1 = 1.6 \text{ 码元/信源符号}$$

3. 有线电报通信采用的摩尔斯编码把英语的 26 个字母和 5 个标点符号通过简单的编码表达出来。

编码单元的长度与字母和标点符号出现的频度有关。统计发现英文字母的出现频率如表 5-1 所

示。

表 5-1 英文字母的出现频率

字母	A	B	C	D	E	F	G	H	I	J	K	L	M
出现频率(%)	8.167	1.492	2.782	4.253	12.7.2	2.228	2.015	6.094	6.966	0.153	0.772	4.025	2.406
字母	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
出现频率(%)	6.749	7.507	1.929	0.095	5.987	6.327	9.056	2.758	0.978	2.360	0.15	1.974	0.074

(1) 若信源符号统计独立，求每个电报符号二元无失真编码的最佳码长；

(2) 考虑符号间的相关性，且每个电报符号携带的平均信息量是 $H_{\infty} = 1.4 \text{ bit}$ ，则应该如何进行二元无失真编码？与(1)相比，编码效率有什么变化？

4. 已知信源 $X \in \{x_1, x_2, x_3, x_4, x_5, x_6\}$ ，若 $p(x_1) = 0.37$ ， $p(x_2) = 0.25$ ， $p(x_3) = 0.18$ ， $p(x_4) = 0.10$ ， $p(x_5) = 0.07$ ， $p(x_6) = 0.03$ 。

(1) 分别用香农编码、费诺编码和哈夫曼编码写出各符号的二元码字，并计算编码效率和信息传输率；

(2) 从(1)中，能得出什么结论？

(3) 若该信源每秒钟发出 2 个符号，求信源的信息传输速率。

解：

(1) 香农编码

信源符号 x_i	$p(x_i)$	$P = \sum_{k=1}^{i-1} p(x_k)$	$-\log_2 p(x_i)$	码长 k_i	码字
x_1	0.37	0	1.4344	2	00
x_2	0.25	0.37	2	2	01
x_3	0.18	0.62	2.4379	3	100
x_4	0.10	0.80	3.3219	4	1100
x_5	0.07	0.90	3.8365	4	1110
x_6	0.03	0.97	5.0589	6	111110

$$\bar{K} = \sum_{i=1}^6 p(x_i) k_i = 0.37 \times 2 + 0.25 \times 2 + 0.18 \times 3 + 0.1 \times 4 + 0.07 \times 4 + 0.03 \times 6 = 2.64 \quad \text{码元/符号}$$

$$H(X) = -\sum_{i=1}^6 p(x_i) \log p(x_i) = 2.2221$$

$$\eta = \frac{H(X)}{\bar{K}} = 84.17\%$$

$$R = \frac{H(X)}{\bar{K}} = 0.8417 \text{ 比特/码元}$$

费诺编码:

x_1	0.37	0	0			00	
x_2	0.25		1			01	
x_3	0.18	1	0			10	
x_4	0.10		1	0		110	
x_5	0.07			1	0		1110
x_6	0.03				1		

$$\bar{K} = \sum_{i=1}^4 p(x_i)k_i = 0.37 \times 2 + 0.25 \times 2 + 0.18 \times 2 + 0.1 \times 3 + 0.07 \times 4 + 0.03 \times 4 = 2.3 \text{ 码元/符号}$$

$$\eta = \frac{H(X)}{\bar{K}} = 96.61\%$$

$$R = \frac{H(X)}{\bar{K}} = 0.9661 \text{ 比特/码元}$$

哈夫曼编码:

信源符号 x_i	$p(x_i)$						码字 码长	
x_1	0.37	0.37	0.37	0.38	0.62	0	00	2
x_2	0.25	0.25	0.25	0.37	0.38	1	01	2
x_3	0.18	0.18	0.20	0.25			11	2
x_4	0.10	0.10	0.18				101	3
x_5	0.07	0.10					1000	4
x_6	0.03						1001	4

$$\bar{K} = \sum_{i=1}^4 p(x_i)k_i = 0.37 \times 2 + 0.25 \times 2 + 0.18 \times 2 + 0.1 \times 3 + 0.07 \times 4 + 0.03 \times 4 = 2.3 \text{ 码元/符号}$$

$$\eta = \frac{H(X)}{\bar{K}} = 96.61\%$$

$$R = \frac{H(X)}{\bar{K}} = 0.9661 \text{ 比特/码元}$$

(2) 与香农编码相比, 费诺编码和哈夫曼编码的平均码长较小, 消息传输率大, 编码效率高。

$$(3) R_t = \frac{H(X)}{2} = 1.1111 \quad \text{比特/秒} \quad (2 \text{ 分})$$

5. 给定信源 $\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} s_1 & s_2 \\ 0.8 & 0.2 \end{bmatrix}$,

(1) 对该信源进行哈夫曼编码, 求所编出的码字和编码效率;

(2) 对该信源的二次扩展信源 X^2 进行哈夫曼编码, 求所编出的码字和编码效率;

(3) 如对该信源的三次扩展信源 X^3 进行哈夫曼编码, 预测其编码效率与(1)(2) 相比会发生什么变化, 为什么?

解: (1) 信源熵 $H(S) = -\sum_{i=1}^2 p(s_i) \log p(s_i) \approx 0.722 \text{ 比特/符号}$

$$s_1 \rightarrow 1, s_2 \rightarrow 0, \bar{k} = 1 \text{ 二元符号/信源符号}, \eta = \frac{H(S)}{\bar{K}} = 72.2\%$$

(2) $N=2$, 对二次扩展信源进行哈夫曼编码, 得

$$\begin{bmatrix} X^2 \\ p(s) \end{bmatrix} = \begin{bmatrix} s_1s_1 & s_1s_2 & s_2s_1 & s_2s_2 \\ 0.64 & 0.16 & 0.16 & 0.04 \end{bmatrix}$$

哈夫曼码 0 10 110 111

得 $\bar{k}_2 = 1.56 \text{ 二元符号/信源序列}$

$$\bar{k} = 0.78 \text{ 二元符号/信源符号}$$

$$\eta = \frac{H(S)}{\bar{K}} = \frac{0.722}{0.78} = 92.6\%$$

(3) 可以预测, 随着 N 的增加, 所编码的单符号平均码长逐渐减小, 编码效率逐渐增大, 趋近于 1。

6. 已知信源 $X \in \{x_1, x_2, x_3, x_4\}$, 若 $p(x_1) = 1/2$, $p(x_2) = 1/4$, $p(x_3) = 1/8$, $p(x_4) = 1/8$, 用香农编码、费诺编码和哈夫曼编码写出各符号的二元码字, 并计算编码效率和信息传输率。

解: 编码过程略

香农编码码字: 0,10,110,111

费诺编码码字: 0,10,110,111

哈夫曼编码码字: 1,01,000,001

三种编码方法的平均码长和编码效率相同, 为

平均码长：
$$\bar{K} = \sum_{i=1}^4 k_i p(x_i) = 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 3 \times \frac{1}{8} = \frac{7}{4} \text{ 码元/符号}$$

信源熵 $H(X) = -\sum_{i=1}^4 p(x_i) \log p(x_i) = 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 3 \times \frac{1}{8} = \frac{7}{4} \text{ 比特/符号}$

编码效率 $\eta = \frac{H(X)}{\bar{K}} = 100\%$

7. 有二元独立信源，已知 $p(0)=0.9$ ， $p(1)=0.1$ ，求信源的符号熵。当用哈夫曼编码时，以二个二元符号合成一个新符号，求这个新符号平均码长和编码效率。设输入二元符号的速率是每秒 100 个，若信道码率已规定为 80 bit/s，存储器容量（比特数）将如何选择？

解：(1) 信源熵 $H(X) = -\sum_{i=1}^2 p(x_i) \log p(x_i) = H(0.9, 0.1) = 0.469 \text{ 比特/符号}$

(2) $N=2$ ，对二次扩展信源进行哈夫曼编码，得

$$\begin{bmatrix} X^2 \\ P \end{bmatrix} = \begin{bmatrix} 00 & 01 & 10 & 11 \\ 0.81 & 0.09 & 0.09 & 0.01 \end{bmatrix}$$

哈夫曼编码过程：

信源序列	序列概率		码字	码长
00	0.81	0.81	0	1
01	0.09	0.1	11	2
10	0.09	0.09	100	3
11	0.01	0.01	101	3

平均码长

$$\bar{K}_2 = \sum_{i=1}^4 k_i p(\alpha_i) = 1 \times 0.81 + 2 \times 0.09 + 3 \times (0.09 + 0.01) = 1.29 \text{ 符号/序列}$$

$$\bar{K} = \frac{\bar{K}_2}{2} = 0.645 \text{ 二元符号/信源符号}$$

编码效率：

$$\eta = \frac{H(X)}{\bar{K}} = 72.7\%$$

编码信息率：

$$R = \frac{H(X)}{K} = 0.727 \text{ 比特/码元}$$

编码后码符号的信息传输速率为

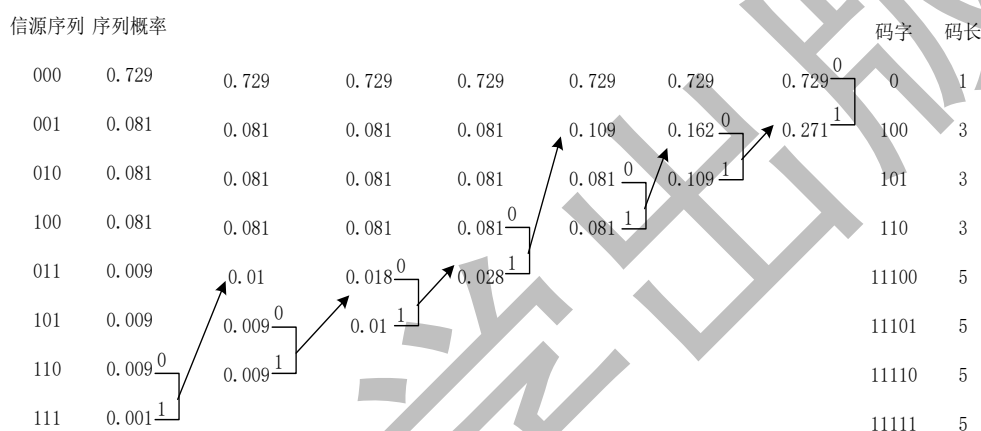
$$R_t = 100R = 72.7 \text{ 比特/秒}$$

信道容量 $C_t = 80$ 比特/秒，因 $C_t < R_t$ ，因此还需要对信源进行压缩。

(3) $N=3$ ，对三次扩展信源进行哈夫曼编码，得

$$\begin{bmatrix} X^3 \\ P \end{bmatrix} = \begin{bmatrix} 000 & 001 & 010 & 100 & 011 & 101 & 110 & 111 \\ 0.729 & 0.081 & 0.081 & 0.081 & 0.009 & 0.009 & 0.009 & 0.001 \end{bmatrix}$$

哈夫曼编码过程：



平均码长

$$\overline{K}_3 = \sum_{i=1}^8 k_i p(\alpha_i)$$

$$= 1 \times 0.729 + 3 \times (0.081 + 0.081 + 0.081) + 5 \times (0.009 + 0.009 + 0.009 + 0.001) = 1.598 \text{ 符号/序列}$$

$$\overline{K} = \frac{\overline{K}_3}{3} = 0.5327 \text{ 二元符号/信源符号}$$

编码信息率：

$$R = \frac{H(X)}{\overline{K}} = 0.88 \text{ 比特/码元}$$

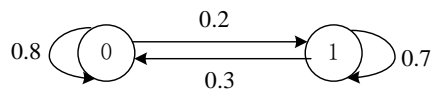
编码后码符号的信息传输速率为

$$R_t = 100R = 88 \text{ 比特/秒}$$

因 $C_t > R_t$ ，所以满足传输要求。存储器容量最少为 3 比特。

8. 有二元平稳马氏链，已知 $p(0/0) = 0.8$ ， $p(1/1) = 0.7$ ，求它的符号熵。用二个符号合成一个新符号来编哈夫曼码，求这个新符号的平均码长和编码效率。

(1) 马尔可夫状态图



(2) 马尔可夫信源状态概率为 w_1 和 w_2 ，则

$$\begin{cases} w_j = \sum_{i=1}^2 w_i p_{ij} \\ \sum_{j=1}^2 w_j = 1 \end{cases}$$

代入状态转移概率，

$$\begin{cases} w_0 = 0.8w_0 + 0.3w_1 \\ w_1 = 0.2w_0 + 0.7w_1 \\ w_0 + w_1 = 1 \end{cases}$$

解得 $w_0 = \frac{3}{5}$ $w_1 = \frac{2}{5}$

(3) 马尔可夫信源熵

$$H_\infty = \sum_{i=0}^1 w_i H(X / s_i)$$

$$H(X / s_i) = - \sum_{j=1}^2 p(x_j / s_i) \log p(x_j / s_i)$$

所以

$$\begin{aligned} H_\infty &= w_0 H(X / s_0) + w_1 H(X / s_1) \\ &= w_0 H(p(0/0), p(1/0)) + w_1 H(p(0/1), p(1/1)) \\ &= \frac{3}{5} H(0.8, 0.2) + \frac{2}{5} H(0.3, 0.7) \\ &= \frac{3}{5} \times 0.722 + \frac{2}{5} \times 0.881 \\ &= 0.786 \end{aligned}$$

(4) $N=2$ ，对二次扩展马氏信源进行哈夫曼编码，得

$$\begin{bmatrix} X^2 \\ P \end{bmatrix} = \begin{bmatrix} 00 & 01 & 10 & 11 \\ 0.48 & 0.12 & 0.12 & 0.28 \end{bmatrix}$$

其中

$$\begin{aligned} p(00) &= p(0)p(0/0) = 0.48 \\ p(01) &= p(0)p(1/0) = 0.12 \\ p(10) &= p(1)p(0/1) = 0.12 \\ p(11) &= p(1)p(1/1) = 0.28 \end{aligned}$$

哈夫曼编码过程：

信源序列	序列概率			码字	码长
00	0.48	0.48	0.52 $\begin{matrix} 0 \\ 1 \end{matrix}$	1	1
11	0.28	0.28 $\begin{matrix} 0 \\ 1 \end{matrix}$	0.48 $\begin{matrix} 0 \\ 1 \end{matrix}$	00	2
01	0.12 $\begin{matrix} 0 \\ 1 \end{matrix}$	0.24 $\begin{matrix} 0 \\ 1 \end{matrix}$		010	3
10	0.12 $\begin{matrix} 0 \\ 1 \end{matrix}$			011	3

平均码长

$$\overline{K}_2 = \sum_{i=1}^4 k_i p(\alpha_i)$$

$$= 1 \times 0.48 + 2 \times 0.28 + 3 \times (0.12 + 0.12) = 1.76 \text{ 符号/序列}$$

$$\overline{K} = \frac{\overline{K}_2}{2} = 0.88 \quad \text{二元符号/信源符号}$$

编码效率:

$$\eta = \frac{H_{\infty}(X)}{\overline{K}} = 89\%$$

9. 设二元无记忆信源 $X \in \{0,1\}$, 其中 $p(0) = \frac{1}{4}$, $p(1) = \frac{3}{4}$ 。对二元序列 11111100 做算术编码。

解: 根据算术码编码规则, 可得

$$p(s = 11111100) = p^2(0)p^6(1) = (1/4)^2(3/4)^6$$

$$\log \frac{1}{p(s)} = 2 \log 4 + 6 \log \frac{4}{3} = 8 \log 4 - 6 \log 3 = 6.49$$

$$l = \left\lceil \log \frac{1}{p(s)} \right\rceil = 7$$

根据累积分布函数的定义, 信源符号序列 s 的累积分布函数:

$$F(s) = \sum_{y < s} p(y) = \sum_{s \text{ 左侧的所有节点 } T} p(T)$$

显然

$$\begin{aligned} F(s) &= p(0) + p(10) + p(110) + p(1110) + p(11110) + p(111110) \\ &= 1 - \sum_{y \geq s} p(y) \\ &= 1 - p(1111111) - p(1111110) - p(11111101) - p(11111100) \\ &= 1 - p(111111) = 1 - (3/4)^6 = 0.82202 \\ &= 0.110100100111 \end{aligned}$$

得 $C = 0.1101010$ (将符号序列的累积分布函数写成二进制的小数, 取小数点后 l 位, 若后面有尾数, 就进位到第 l 位, 得到一个数 C), 则 s 对应的码字为 1101010。

编码效率为

$$\eta = \frac{H(X)}{\overline{K}} = \frac{0.811}{7/8} = 92.7\%$$

10. 一信源发出的数字有 1,2,3,4,5,6,7, 对应的概率分别为 $p(1)=p(2)=1/3$, $p(3)=p(4)=1/9$, $p(5)=p(6)=p(7)=1/27$, 在二进制或三进制无噪声信道中传输, 若二进制信道中传输一个码符号需要 1.8 元, 三进制信道中传输一个码符号需要 2.7 元。

- (1) 编出二进制符号的霍夫曼码, 求其编码效率。
- (2) 编出三进制符号的费诺码, 求其编码效率。
- (3) 根据(1)和(2)的结果, 确定在何种信道中传输可得到较小的花费。

11. 设有一页传真文件, 其中某一扫描行上的像素点如下所示:

|← 73 白 →|← 7 黑 →|← 白 →|← 18 黑 →|← 1619 白 →|

- (1) 该扫描行的 MH 码。
- (2) 编码后该行的总比特数。
- (3) 本行编码压缩比(原码元总数: 编码后码元总数)。

解:(1)根据 MH 码表可得, 该扫描行的 MH 码为

64 白 + 19 白 9 黑 12 白 5 黑 1600 白 + 19 白 EOL

11011+0001100 000100 001000 0011 010011010+0001100 000000000001\$编码后该行(2)

总比特数为 56 位。

(3)这一行编码压缩比为 $(1728-56)/1728 \approx 96.8\%$ 。

第六章

1. 设有一离散信道，其信道传递矩阵为

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$$

并设信源符号概率为 $p(x_1) = \frac{1}{4}$, $p(x_2) = \frac{1}{2}$, $p(x_3) = \frac{1}{4}$, 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。

解：(1) 采用最小错误概率译码准则，则联合概率矩阵为：

$$[p(a_i b_j)] = [p(a_1) \ p(a_2) \ p(a_3)] \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$$

当信源不等概分布时， $P(a_1) = \frac{1}{4}$, $P(a_2) = \frac{1}{2}$, $P(a_3) = \frac{1}{4}$ ，此时联合矩阵为

$$[P(a_i b_j)] = \begin{bmatrix} \frac{1}{8} & \frac{1}{12} & \frac{1}{24} \\ \frac{1}{12} & \frac{1}{4} & \frac{1}{6} \\ \frac{1}{12} & \frac{1}{24} & \frac{1}{8} \end{bmatrix}$$

按照最小错误概率译码准则，所得译码函数为：B: $\begin{cases} F(b_1) = a_1 \\ F(b_2) = a_2 \\ F(b_3) = a_2 \end{cases}$

平均错误率为：

$$P_E = \sum_{Y, X=a^*} p(a_i b_j) = (\frac{1}{12} + \frac{1}{12}) + (\frac{1}{12} + \frac{1}{24}) + (\frac{1}{8} + \frac{1}{24}) = \frac{11}{24}$$

(2) 根据最大似然准则，译码函数为 A: $\begin{cases} F(b_1) = a_1 \\ F(b_2) = a_2 \\ F(b_3) = a_3 \end{cases}$

平均错误率为：

$$P_E = \sum_{Y, X=a^*} p(a_i b_j) = (\frac{1}{12} + \frac{1}{12}) + (\frac{1}{12} + \frac{1}{24}) + (\frac{1}{6} + \frac{1}{24}) = \frac{1}{2}$$

2. 发送端有 3 种等概率符号 $\{x_1, x_2, x_3\}$, $p(x_i) = 1/3$, 接收端收到三种符号 (y_1, y_2, y_3) , 信道转移概率矩阵为

$$P = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.3 & 0.3 \\ 0.1 & 0.9 & 0 \end{bmatrix}$$

- (1) 接收端收到一个符号后得到的信息量 $H(Y)$ 。
- (2) 计算噪声熵 $H(Y/X)$ 。
- (3) 计算接收端收到一个符号 y_2 的错误概率。
- (4) 计算从接收端看的平均错误概率。
- (5) 计算从发送端看的平均错误概率。
- (6) 从转移矩阵能看出信道的好坏吗?
- (7) 计算收到的 $H(X)$ 和 $H(X/Y)$ 。

解:(1) 信道转移概率

$$[P(y_j / x_i)] = \begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.3 & 0.3 \\ 0.1 & 0.9 & 0 \end{pmatrix}$$

根据 $p(y_j) = \sum_{i=1}^3 p(x_i) p(y_j / x_i)$, 可得输出符号概率分布

$$P(y_0) = \frac{1}{3}, P(y_1) = \frac{1}{2}, P(y_2) = \frac{1}{6}$$

因此

$$H(Y) = \frac{1}{3} \log 3 + \frac{1}{2} \log 2 + \frac{1}{6} \log 6 = 1.459$$

(2) 联合概率矩阵和后验条件概率矩阵分别为

$$[P(x_i y_j)] = \begin{pmatrix} \frac{1}{6} & \frac{1}{10} & \frac{2}{5} \\ \frac{2}{15} & \frac{1}{10} & \frac{3}{5} \\ \frac{1}{30} & \frac{1}{10} & 0 \end{pmatrix} \quad [P(x_i / y_j)] = \begin{pmatrix} \frac{1}{2} & \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} & \frac{3}{5} \\ \frac{1}{10} & \frac{3}{5} & 0 \end{pmatrix}$$

因此

$$H(Y/X) = \frac{1}{10} \log \frac{10}{3} + \frac{1}{10} \log \frac{10}{3} + \frac{1}{30} \log 10 + \frac{3}{10} \log \frac{10}{9} + \frac{1}{6} \log 2 + \frac{1}{10} \log \frac{10}{3} + \frac{1}{15} \log 5 + \frac{2}{15} \log \frac{5}{2} = 1.173$$

(3) 当接收为 y_2 , 发为 x_1 时正确, 如果发的是 x_1 和 x_3 为错误, 各自的概率为:

$$P(x_1/y_2) = \frac{1}{5}, P(x_2/y_2) = \frac{1}{5}, P(x_3/y_2) = \frac{3}{5}$$

其中错误概率为:

$$P_e = P(x_1/y_2) + P(x_3/y_2) = 0.8$$

(4) 平均错误概率为

$$\frac{2}{15} + \frac{1}{30} + \frac{1}{10} + \frac{3}{10} + \frac{1}{15} + \frac{1}{10} = 0.733$$

(5) 仍为 0.733

(6) 此信道不好

原因是信源等概率分布, 从转移信道来看

正确发送的概率 x_1-y_1 的概率 0.5 有一半失真

x_2-y_2 的概率 0.3 有失真严重

x_3-y_3 的概率 0 完全失真

(7)

$$H(X) = \log 3 = 1.585$$

$$H(X/Y) = \frac{1}{6} \log 2 + \frac{1}{10} \log 5 + \frac{1}{15} \log \frac{5}{2} + \frac{2}{15} \log \frac{5}{2} + \frac{1}{10} \log 5 + \frac{1}{10} \log \frac{5}{3} + \frac{1}{30} \log 10 + \frac{3}{10} \log \log \frac{5}{3} = 1.301$$

3. 已知(8,5)线性分组码的生成矩阵为

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

求: (1) 输入为全 00011 和 10100 时该码的码字; (2) 最小码距。

解: 根据 $C=mG$, 输入 00011 时输出的码字是

$$(00010001) + (00001111) = (00011110)$$

输入 10100 时输出的码字是

$$(10000111)+(00100010)=(10100101)$$

(2)最小码距等于码集中码字的最小重量，因此最小距离为2。

4. 二进制线性分组码的生成矩阵为 $G = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$ ，试求：

(1) 生成矩阵的系统形式；

(2) 给出校验矩阵 H 。

解：生成矩阵 G 中的第一行和第三行交换顺序，得

$$G_{sys} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix} = [I_3 | P]$$

校验矩阵：

$$H = [P | I_4] = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

5. 某分组码的校验矩阵 $H = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$ ，求：

(1) 请分别给出信息位 k 和码字长度 n 的值；

(2) 求该码的生成矩阵；

(3) 试判断码字 010111 和 100011 是否是码字？并给出理由。

解：(1) $k=3$, $n=6$

$$(2) H = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} = [P^T | I_3]$$

$$G = [I_3 | P] = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

$$(3) (010111) \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}^T = (010) \neq 0, 010111 \text{ 不是许用码字；}$$

$$(100011) \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}^T = (000), 010111 \text{ 是许用码字。}$$

6. 已知一个线性分组码的校验矩阵 $H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$, 试给出当输入信息序列为

1101, 1000, 1001 时, 编码器的输出码字。

$$\text{解: } H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} = [P^T | I_3]$$

$$G = [I_4 | P] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

输入信息序列为 1101, 1000, 1001 时, 编码器的输出码字分别为

1101——1101001

1000——1000011

1001——1001100

7. 二进制 (15,8) 循环码共有多少码字? 能否由一个码字循环产生所有的码字? 给出理由。

解: (1) 码字数量为 $2^8=256$ 。(2) 不能, 因为一个码字的长度是 15, 一个循环的码字数量最多是 15。

8. 已知 (15,7) 循环码的生成多项式为 $g(x) = x^8 + x^7 + x^6 + 1$ 。试求出该循环码的生成矩阵和校验矩阵, 并给出其标准形式。

$$\text{解: 生成矩阵与生成多项式的关系为: } G(x) = \begin{bmatrix} x^{k-1}g(x) \\ x^{k-2}g(x) \\ \vdots \\ xg(x) \\ g(x) \end{bmatrix}$$

因此生成矩阵为

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

生成矩阵的系统（标准）形式为：

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = [I_7 | P]$$

因此校验矩阵为

$$H = [P^T | I_8] = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

9. 已知 (3,1,2) 卷积码，生成矩阵为 $G^0 = [111]$ ， $G^1 = [101]$ ， $G^2 = [011]$ 。试：

- (1) 画出该码的状态图。
- (2) 画出该码的网格图。
- (3) 如果输入序列为 1011，写出编码后的序列，并给出 MATLAB 实现。

解：(3,1,2) 卷积码共有 4 个状态，记为 $S_0 = (00)$ ， $S_1 = (01)$ ， $S_2 = (10)$ ， $S_3 = (11)$

编码方程为

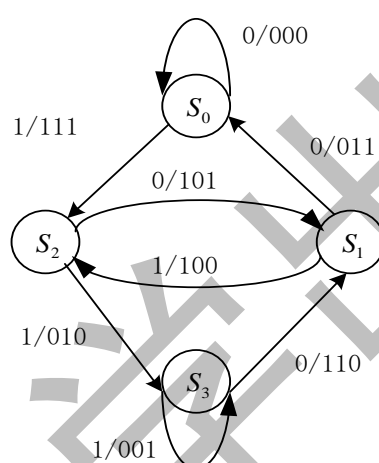
$$\begin{cases} c_0^i = m_0^i + m_0^{i-1} \\ c_1^i = m_0^i + m_0^{i-2} \\ c_2^i = m_0^i + m_0^{i-1} + m_0^{i-2} \end{cases}$$

记忆阵列状态和输入输出码元的关系

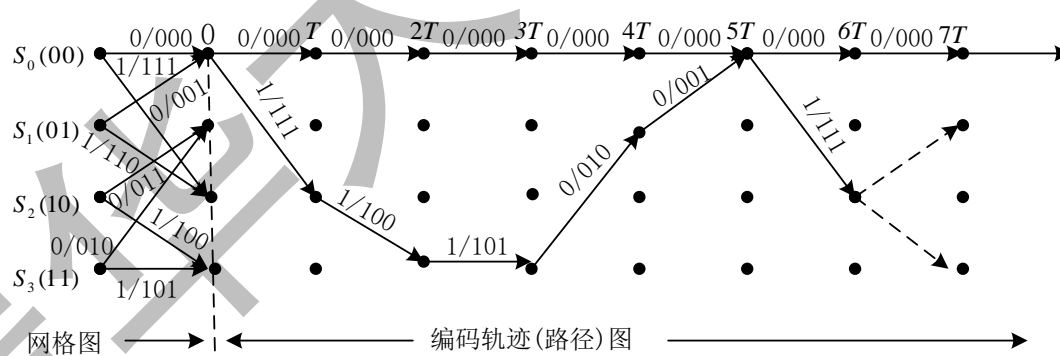
移存器前一状态	当前输入信息位	输出码元	移存器下一状态
---------	---------	------	---------

$m_0^{i-1}m_0^{i-2}$	m_0^i	$C^i = (c_0^i c_1^i c_2^i)$	$m_0^i m_0^{i-1}$
$S_0(00)$	0	000	$S_0(00)$
	1	111	$S_2(10)$
$S_1(01)$	0	011	$S_0(00)$
	1	100	$S_2(10)$
$S_2(10)$	0	101	$S_1(01)$
	1	010	$S_3(11)$
$S_3(11)$	0	110	$S_1(01)$
	1	001	$S_3(11)$

(1) (3,1,2) 卷积码状态流程图



(2) 卷积码网格图。



(3)

编码输出为

$$C = [1011] \begin{bmatrix} 111 & 101 & 011 \\ & 111 & 101 & 011 \\ & & 111 & 101 & 011 \\ & & & 111 & 101 & 011 \end{bmatrix} = [111, 101, 100, 010]$$

10. 请简要介绍 Turbo 码和 LDPC 码的基本编译码原理。(略)