

信道是信息传递的通道,承担信息的传输任务,是构成通信系统的重要组成部分。本章将讨论信道模块,并重点探讨离散信道及其所能传送的信息量——平均互信息量。

3.1 信道的模型和分类

一般来说,信道指传输信息的物理媒质,如电缆、光缆、电波、光波等。为了集中精力讨论信息理论,本章不研究具体的物理传输媒质的特性,而是研究由这些物理传输媒质及相应的调制解调器组成的**编码信道**的特性,或者更进一步,研究由编码信道与信道编译码器和信源编译码器组成的**等效信道**的特性。

3.1.1 信道的系统模型

图 3.1 描绘了关注信道的一般通信系统模型。图 3.1 中编码器包括信源编码器与信道编码器,而译码器包括信源译码器与信道译码器。根据不同的研究需要,图 3.1 可简化成图 3.2 和图 3.3。通过图 3.2,可研究各种编码信道的传信能力,包括它的**信息传输速率**和**信道容量**;通过图 3.3,可了解消息从信源输出到信宿接收这一过程的等效信道特性。

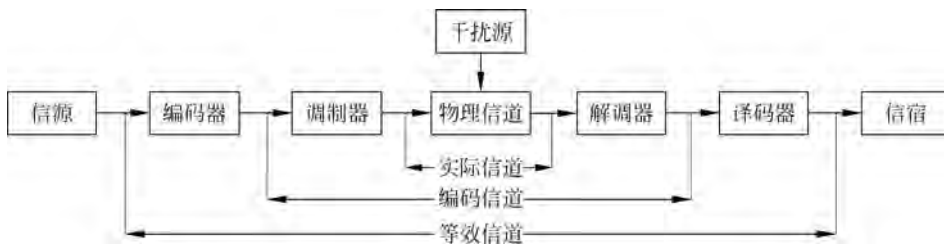


图 3.1 关注信道的一般通信系统模型



图 3.2 关注编码信道的通信系统简化模型



图 3.3 关注等效信道的通信系统简化模型

3.1.2 信道的分类

信道可以从不同的角度加以分类,我们所关注的等效信道归纳起来主要可分为以下五类。

1. 根据信道输入和输出空间的连续与否进行分类

(1) **离散信道**。输入和输出均为离散消息的集合,有时也称为数字信道。

(2) **连续信道**。输入和输出均为连续消息的集合,又称为模拟信道。波形信道是典型的连续信道。

(3) **半连续信道**。在输入和输出中,一个是离散消息集合,另一个是连续消息集合。

2. 根据信道输入、输出消息集合的个数分类

(1) **两端信道**。在信道的输入端和输出端中,每端只有一个消息集合,即只有一对用户。在信道两端进行单向通信,也称为单用户或单路信道。

(2) **多端信道**。在信道的输入端和输出端中,至少有一端具有一个以上的消息集合,也称为多用户信道,典型的多用户信道有多元接入信道和广播信道。

3. 根据信道的统计特性分类

(1) **恒参信道**。信道统计特性不随时间变化,又称为平稳信道,如有线信道。

(2) **随参信道**。信道统计特性随时间变化,如无线信道。

4. 根据信道的记忆特性分类

(1) **无记忆信道**。信道输出仅与当前输入有关,而与以前输入无关。目前离散无记忆信道的理论发展得比较成熟和完整,因此本章将会进行介绍。

(2) **有记忆信道**。信道输出不仅与当前输入有关,而与以前的输入有关。码间串扰信道和衰落信道都是有记忆信道。

5. 根据信道上是否存在干扰进行分类

(1) **无扰信道**。指信道上没有干扰,这是一种理想化的信道,如计算机与其外设之间的数据传输信道。

(2) **有扰信道**。指信道上存在干扰,实际上大部分信道都是有扰信道,特别是无线信道。

本章将侧重讨论恒参、无记忆、单路的离散信道。

3.1.3 离散信道的数学模型

从统计数学角度看,离散信道可以看成是一个随机变换,将输入的随机序列 x 变换成输出的随机序列 y 。由于干扰存在,输入 x 总是以一定的条件概率 $P(y|x)$ 变换成输出 y 。因此,离散信道的数学模型可以表示为 $\{X, P(y|x), Y\}$,如图 3.4 所示。其中条件概率 $P(y|x)$ 反映了信道的统计特征,称为信道转移概率或传递概率,具体描述如下:

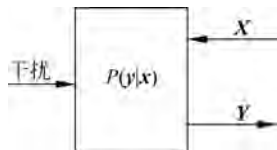


图 3.4 离散信道的数学模型

设离散信道输入符号集为 $A = \{a_1, a_2, \dots, a_n\}$, 相应的概率分布为 $\{P(a_i) = p_i, i = 1, 2, \dots, n\}$; 信道输出符号集为 $B = \{b_1, b_2, \dots, b_m\}$, 而其概率分布为 $\{P(b_j) = q_j, j = 1, 2, \dots, m\}$ 。

那么信道输入序列 $\mathbf{X} = (X_1, X_2, \dots, X_K)$ 可取值为 $\mathbf{x} = (x_1, x_2, \dots, x_K)$, 其中 $x_k \in \mathbf{A}$, $1 \leq k \leq K$; 而信道输出序列 $\mathbf{Y} = (Y_1, Y_2, \dots, Y_K)$ 可取值为 $\mathbf{y} = (y_1, y_2, \dots, y_K)$, 其中 $y_k \in \mathbf{B}$, $1 \leq k \leq K$ 。这样信道转移概率可表示为

$$P(\mathbf{y} | \mathbf{x}) = P(y_1 y_2 \cdots y_K | x_1 x_2 \cdots x_K) \quad (3.1)$$

对于理想的无扰信道, 输入的取值序列 $\mathbf{x}$ 与输出的取值序列 $\mathbf{y}$ 之间有一种确定的一一对应关系, 即 $\mathbf{y} = f(\mathbf{x})$ 。这样式(3.1)变为

$$P(\mathbf{y} | \mathbf{x}) = \begin{cases} 1, & \mathbf{y} = f(\mathbf{x}) \\ 0, & \mathbf{y} \neq f(\mathbf{x}) \end{cases}$$

实际信道中常常是有干扰的, 那么输入的取值序列 $\mathbf{x}$ 与输出的取值序列 $\mathbf{y}$ 之间不会有确定的一一对应关系。在实际有干扰的信道中, 无记忆信道是比较简单容易进行分析的信道。

1. 离散无记忆信道

对于离散无记忆信道, 其任意时刻输入符号只与对应时刻的输入符号有关, 而与以前时刻的输入符号、输出符号无关, 也与以后的输入符号无关。

命题 3.1 满足离散无记忆信道 $\{\mathbf{X}, P(\mathbf{y} | \mathbf{x}), \mathbf{Y}\}$ 的充分必要条件是

$$P(\mathbf{y} | \mathbf{x}) = \prod_{k=1}^K P(y_k | x_k) \quad (3.2)$$

证明 (必要性) 假定离散信道 $\{\mathbf{X}, P(\mathbf{y} | \mathbf{x}), \mathbf{Y}\}$ 是无记忆的, 则有

$$\begin{cases} P(y_1 | x_1 x_2 \cdots x_K) = P(y_1 | x_1) \\ P(y_2 | x_1 x_2 \cdots x_K y_1) = P(y_2 | x_2) \\ \vdots \\ P(y_K | x_1 x_2 \cdots x_K y_1 y_2 \cdots y_{K-1}) = P(y_K | x_K) \end{cases}$$

进而有

$$\begin{aligned} P(\mathbf{y} | \mathbf{x}) &= P(y_1 | x_1 x_2 \cdots x_K) P(y_2 | x_1 x_2 \cdots x_K y_1) P(y_3 | x_1 x_2 \cdots x_K y_1 y_2) \cdots \\ &\quad P(y_{K-1} | x_1 x_2 \cdots x_K y_1 y_2 \cdots y_{K-2}) \cdot P(y_K | x_1 x_2 \cdots x_K y_1 y_2 \cdots y_{K-1}) \\ &= \prod_{k=1}^K P(y_k | x_k) \end{aligned}$$

因此式(3.2)成立。

(充分性) 由于式(3.2)成立, 可推得

$$\begin{aligned} &P(y_K | x_1 x_2 \cdots x_K y_1 y_2 \cdots y_{K-1}) \\ &= \frac{P(y_1 y_2 \cdots y_K | x_1 x_2 \cdots x_K)}{P(y_1 y_2 \cdots y_{K-1} | x_1 x_2 \cdots x_K)} \\ &= \frac{\prod_{k=1}^K P(y_k | x_k)}{\sum_{y_K} P(y_1 y_2 \cdots y_K | x_1 x_2 \cdots x_K)} \\ &= \frac{\prod_{k=1}^K P(y_k | x_k)}{\left(\sum_{y_K} P(y_K | x_K) \right) \prod_{k=1}^{K-1} P(y_k | x_k)} \\ &= P(y_K | x_K) \end{aligned}$$

同理可推得

$$\begin{aligned} P(y_{K-1} | x_1 x_2 \cdots x_K y_1 y_2 \cdots y_{K-2}) &= P(y_{K-1} | x_{K-1}) \\ &\vdots \\ P(y_1 | x_1 x_2 \cdots x_K) &= P(y_1 | x_1) \end{aligned}$$

因此该信道是无记忆的。

对于离散无记忆信道,若还满足对任意的 i, j , 有

$$P(y_k = b_j | x_k = a_i) = P(y_1 = b_j | x_1 = a_i), \quad k = 1, 2, \dots, K \quad (3.3)$$

则称之为平稳的离散无记忆信道。平稳离散无记忆信道的转移概率不随时间变化,因此可用一维概率分布来描述。一般情况下,若无特殊声明,所讨论的离散无记忆信道均是平稳的。这样,在离散无记忆条件下,只需研究单个符号的传输。

2. 单符号离散信道

1) 一般单符号离散信道

单符号离散信道的数学模型变得特别简单,可表示为

$$\{X, P(y | x), Y\}$$

其转移概率简化为

$$P(y | x) = P(Y = b_j | X = a_i) = P(b_j | a_i)$$

并且满足

$$P(b_j | a_i) \geq 0, \quad \sum_j P(b_j | a_i) = 1$$

其中 $a_i \in \mathbf{A} = \{a_1, a_2, \dots, a_n\}$, $b_j \in \mathbf{B} = \{b_1, b_2, \dots, b_m\}$ 。

所有转移概率放在一起组成一个矩阵,称为信道矩阵,简记为 $\mathbf{P} = [P_{ij}]$, 其中 $P_{ij} = P(b_j | a_i)$ 。

2) 二进制对称信道

如令 $\mathbf{A} = \mathbf{B} = \{0, 1\}$, 则信道矩阵为

$$\begin{bmatrix} P(0 | 0) & P(0 | 1) \\ P(1 | 0) & P(1 | 1) \end{bmatrix}$$

进一步,如还有 $P(1 | 0) = P(0 | 1) = p$ 成立,则称这种信道为二进制对称信道,如图 3.5 所示。这时信道矩阵变为

$$\mathbf{P} = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix} \quad (3.4)$$

这里 $\bar{p} = 1 - p$ 。

3) 二进制删除信道

在 $\mathbf{A} = \{0, 1\}$ 和 $\mathbf{B} = \{0, ?, 1\}$ 时,如果还有 $P(1 | 0) = P(0 | 1) = p$

且 $P(x | 0) = P(x | 1) = q$ 成立,则称这种信道为二进制删除信道,如图 3.6 所示。二进制删除信道的信道矩阵表示为

$$\mathbf{P} = \begin{bmatrix} 1 - p - q & q & p \\ p & q & 1 - p - q \end{bmatrix} \quad (3.5)$$

进一步,如果 $p = 0$,那么如图 3.6 所示的模型将用如图 3.7 所示的模型代替,称此特殊的二进制删除信道为二进制纯删除信道,并称 q 为删除概率。二进制删除信道的信道矩阵简化为

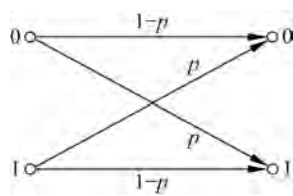


图 3.5 二进制对称信道

$$\mathbf{P} = \begin{bmatrix} 1-q & q & 0 \\ 0 & q & 1-q \end{bmatrix} \quad (3.6)$$

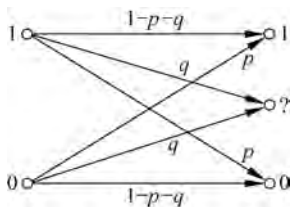


图 3.6 二进制删除信道

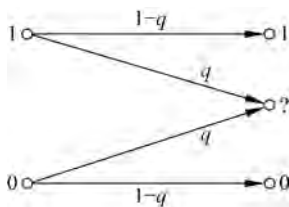


图 3.7 二进制纯删除信道

在编码通信技术中,二进制对称信道和二进制删除信道是两个典型的编码信道模型,具体将会在第 6 章进行讨论。

3.2 互信息量与平均互信息量

2.2 节研究了离散信源含有多少信息量的问题,本节将探讨通过离散信道能得到多少信息量。

3.2.1 互信息量

考虑一个简单的通信系统简化模型,如图 3.8 所示。 X 表示信源发出的符号消息集合,而 Y 表示信宿接收的符号消息集合。每个符号消息相当于一个随机事件, X 与 Y 均是由随机事件形成的集合,其概率空间分别为



图 3.8 通信系统简化模型

$$\begin{aligned} \begin{bmatrix} X \\ P \end{bmatrix} &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ P(x_1) & P(x_2) & \cdots & P(x_n) \end{bmatrix} \\ \begin{bmatrix} Y \\ P \end{bmatrix} &= \begin{bmatrix} y_1 & y_2 & \cdots & y_m \\ P(y_1) & P(y_2) & \cdots & P(y_m) \end{bmatrix} \end{aligned}$$

这里 x_i 表示一个随机事件,其出现的概率为 $P(x_i)$, $P(x_i)$ 也称为先验概率。信源发送 x_i 时所含有的自信息量为 $I(x_i) = -\log P(x_i)$, 它表示 x_i 的不确定度。这里 y_j 也表示一个随机事件,其出现的概率为 $P(y_j)$ 。假定信宿收到 y_j , 那么在已知 y_j 的情况下关于 x_i 可能出现的概率就是条件概率 $P(x_i|y_j)$, $P(x_i|y_j)$ 也称为后验概率。此时 x_i 的自信息量就变为 $I(x_i|y_j) = -\log P(x_i|y_j)$, 这是此时 x_i 的不确定度。条件自信息量 $I(x_i|y_j)$ 的大小能反映信道的好与坏。如果信道非常好,没有任何干扰,那么已知 y_j 即已知 x_i , 应有 $I(x_i|y_j) = 0$, x_i 的不确定度变为零; 反之, 如果信道很差, 干扰表现强烈, 导致 x_i 与 y_j 统计独立, 即没有获得任何关于 x_i 的信息, 那么应有 $I(x_i|y_j) = I(x_i)$, x_i 的不确定度没有变化。一般而言, 如果信道不错, 干扰较弱, 就能获得关于 x_i 的一些信息, 那么应有 $I(x_i|y_j) < I(x_i)$, x_i 的不确定度将会变小。综上所述, $I(x_i) - I(x_i|y_j)$ 能反映通过信道收到的信息量多少, 也能反映信道传信能力。为此, 引出如下定义。

定义 3.1 对于两个随机事件集合 X 与 Y , 事件 y_j 的出现给出关于事件 x_i 的信息量

定义为

$$\begin{aligned} I(x_i; y_j) &\stackrel{\text{def}}{=} I(x_i) - I(x_i | y_j) \\ &= +\log \frac{P(x_i | y_j)}{P(x_i)} \end{aligned} \quad (3.7)$$

并称为互信息量。

互信息量是一种消除不确定性的度量,即互信息量等于先验的不确定度减去尚存的不确定度。互信息量的单位与自信息量相同。另外,值得一提的是,互信息量的概念虽是在通信的背景下引出的,但它也适用于一般情况下两个给定随机事件的集合。

互信息量具有下述几个基本性质。

(1) 互信息量的互易性。

因为 $P(x_i, y_j) = P(x_i)P(y_j | x_i) = P(y_j)P(x_i | y_j)$, 所以 $\frac{P(y_j | x_i)}{P(y_j)} = \frac{P(x_i | y_j)}{P(x_i)}$ 。故 $I(x_i; y_j) = I(y_j; x_i)$, 即互易性成立。这表明,由事件 y_j 的出现给出关于事件 x_i 的信息量等于由事件 x_i 的出现给出关于事件 y_j 的信息量。互易性涉及信息反向传送的问题。

(2) 任何两个事件之间的互信息量不可能大于其中任一事件的自信息量。

既然 $I(x_i | y_j) \geq 0$, 由定义 3.1 中的式(3.7)可知, $I(x_i; y_j) \leq I(x_i)$ 成立。再由互信息量的互易性同理可得, $I(x_i; y_j) \leq I(y_j)$ 成立。这意味着即使信道条件再好,所获得的信息量也不会超过 $I(x_i)$ 和 $I(y_j)$ 。

(3) 互信息量既可为零也可等于自信息量。

当 x_i, y_j 统计独立时,由于 $P(x_i | y_j) = P(x_i)$, 故此时 $I(x_i; y_j) = 0$ 。这意味着此时不能从观测 y_j 获得 x_i 的任何信息。对于理想的无扰信道,输入 x_i 与其相应的输出 y_j 之间有一种一一对应的关系,因此 $P(x_i | y_j) = 1$, 进而 $I(x_i; y_j) = I(x_i)$ 。

(4) 互信息量既可正也可负。

显然有

$$\begin{cases} I(x_i; y_j) > 0, & P(x_i) < P(x_i | y_j) \\ I(x_i; y_j) < 0, & P(x_i) > P(x_i | y_j) \end{cases}$$

互信息量为正,意味着事件 y_j 的出现有助于肯定事件 x_i 的出现;反之,则是不利的。造成不利的原因是信道干扰。

例 3.1 考虑如图 3.5 所示的二进制对称信道,其信道矩阵为 $\mathbf{P} = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}$ 。假定信源发

出 0 和 1 的概率为 $P(x=0) = P(x=1) = \frac{1}{2}$, 计算如下四种情况下的互信息量 $I(x=0; y=0)$ 。

- (1) 当 $p=0$ 时, $P(x=0 | y=0) = 1$, 因此 $I(x=0; y=0) = \log \frac{1}{1/2} = 1(\text{bit})$;
- (2) 当 $p = \frac{1}{4}$ 时, $P(x=0 | y=0) = \frac{3}{4}$, 因此 $I(x=0; y=0) = \log \frac{3/4}{1/2} = 0.585(\text{bit})$;
- (3) 当 $p = \frac{1}{2}$ 时, $P(x=0 | y=0) = \frac{1}{2}$, 因此 $I(x=0; y=0) = \log \frac{1/2}{1/2} = 0(\text{bit})$;
- (4) 当 $p = \frac{3}{4}$ 时, $P(x=0 | y=0) = \frac{1}{4}$, 因此 $I(x=0; y=0) = \log \frac{1/4}{1/2} = -1(\text{bit})$ 。

注: 由于 $P(x_i, y_j) = P(x_i)P(y_j | x_i) = P(y_j)P(x_i | y_j)$, 因此

$$P(x_i | y_j) = \frac{P(x_i, y_j)}{P(y_j)}$$

因此, 为能计算出 $P(x_i | y_j)$, 先要计算出 $P(x_i, y_j) = P(x_i)P(y_j | x_i)$, 然后再计算出 $P(y_j) = \sum_i P(x_i, y_j)$ 。 ■

3.2.2 平均互信息量

上述互信息量描述了通过信道信宿从事件 y_j 获得关于事件 x_i 的信息量。可是, 它并不是信道传递多少信息整体测度。因此, 就像定义平均自信息量一样, 需要定义平均互信息量这一概念, 以从整体的角度刻画通过信道信宿获得关于信源的平均信息量。

定义 3.2 对于两个随机事件集合 X 与 Y , 已知事件 y_j 提供关于集合 X 的平均信息量定义为

$$\begin{aligned} I(X; y_j) &\stackrel{\text{def}}{=} \sum_i P(x_i | y_j) I(x_i; y_j) \\ &= \sum_i P(x_i | y_j) \log \frac{P(x_i | y_j)}{P(x_i)} \end{aligned} \quad (3.8)$$

称之为平均条件互信息量。进而定义集合 Y 提供关于集合 X 的平均信息量为

$$\begin{aligned} I(X; Y) &\stackrel{\text{def}}{=} \sum_j P(y_j) I(X; y_j) \\ &= \sum_{i,j} P(y_j) P(x_i | y_j) \log \frac{P(x_i | y_j)}{P(x_i)} \\ &= \sum_{i,j} P(x_i, y_j) \log \frac{P(x_i | y_j)}{P(x_i)} \\ &= \sum_{i,j} P(x_i, y_j) I(x_i; y_j) \end{aligned} \quad (3.9)$$

称之为平均互信息量。

平均互信息量表示信源的信息通过信道后传输到信宿的平均信息量。平均互信息量有以下几个基本性质。

1. 互易性

$$I(X; Y) = I(Y; X) \quad (3.10)$$

由互信息量与联合概率的互易性容易证明平均互信息量的互易性。

平均互信息量的互易性表明, 从 Y 中提取关于 X 的信息量等于从 X 中提取关于 Y 的信息量。

2. 非负性

$$I(X; Y) \geq 0 \quad (3.11)$$

由于 $P(y_j) \geq 0$, 从平均互信息量的定义式(3.9)可知, 只需证明 $I(X; y_j) \geq 0$ 成立即可。由式(3.8)得

$$I(X; y_j) = \sum_i P(x_i | y_j) \log \frac{P(x_i | y_j)}{P(x_i)}$$

令 $p_i = P(x_i | y_j)$ 和 $q_i = P(x_i)$, 显然

$$\sum_i p_i = \sum_i q_i = 1$$

这样应用引理 2.2 可知, $I(X; y_j) \geq 0$ 成立。

这个性质告诉我们, 通过一个信道获得的平均信息量不可能是负值。也就是说, 观察一个信道的输出, 从平均的角度来看总能消除一些不确定性, 从而获得一些信息。除非信道输入和输出是统计独立的, 才得不到任何信息。

例 3.2 考虑图 3.5 所示的二进制对称信道, 假定信源发出 0 和 1 的概率为 $P(x=0) = P(x=1) = \frac{1}{2}$, 而信道出错的概率为 $P(y=1|x=0) = P(y=0|x=1) = \frac{3}{4}$, 计算 $I(X; y=0)$ 。

这是例 3.1 的继续讨论。参考例 3.1 的“注”可计算出

$$P(x=1 | y=0) = \frac{3}{4}$$

$$P(x=0 | y=0) = \frac{1}{4}$$

因此

$$I(x=1; y=0) = \log \frac{3/4}{1/2} = 0.585(\text{bit})$$

$$I(x=0; y=0) = \log \frac{1/4}{1/2} = -1(\text{bit})$$

$$\begin{aligned} I(X; y=0) &= P(x=1 | y=0)I(x=1; y=0) + P(x=0 | y=0)I(x=0; y=0) \\ &= \frac{3}{4} \cdot 0.585 + \frac{1}{4} \cdot (-1) = 0.189 > 0 \end{aligned}$$

3. 极值性

$$I(X; Y) \leq H(X), \quad I(X; Y) \leq H(Y) \quad (3.12)$$

任何两个事件之间的互信息量不可能大于其中任一事件的自信息量, 即

$$I(x_i; y_j) \leq I(x_i), \quad I(x_i; y_j) \leq I(y_j)$$

那么对上述不等式两边做统计平均即可证得式(3.12)成立。

平均互信息量极值性表明: 不管信道条件多么好, 信宿所能获得的平均信息量不会超过集合 X 本身含有的信息量 $H(X)$, 也不会超过集合 Y 本身含有的信息量 $H(Y)$ 。

4. 与各种熵的关系

$$I(X; Y) = H(X) - H(X | Y) \quad (3.13)$$

$$I(X; Y) = H(Y) - H(Y | X) \quad (3.14)$$

$$I(X; Y) = H(X) + H(Y) - H(X, Y) \quad (3.15)$$

从平均互信息量的定义式(3.9)可知

$$\begin{aligned} I(X; Y) &= \sum_{i,j} P(x_i, y_j) \log \frac{P(x_i | y_j)}{P(x_i)} \\ &= - \sum_{i,j} P(x_i, y_j) \log P(x_i) + \sum_{i,j} P(x_i, y_j) \log P(x_i | y_j) \\ &= - \sum_i P(x_i) \log P(x_i) + \sum_{i,j} P(x_i, y_j) \log P(x_i | y_j) \\ &= H(X) - H(X | Y) \end{aligned}$$

因此推得式(3.13)。利用互易性和式(3.13)易证式(3.14)。再利用熵函数的可加性可证式(3.15)。

$H(X)$ 表示信宿在收到 Y 之前关于 X 的平均不确定度,而条件熵 $H(X|Y)$ 表示信宿在收到 Y 之后关于 X 的尚存的平均不确定度,其大小能反映信道好坏,故称为信道疑义度,有时也称为损失熵。至于条件熵 $H(Y|X)$,其大小完全为信道中的干扰或者说噪声强弱确定,故称为噪声熵。平均互信息量、信息熵、联合熵、噪声熵及损失熵之间的关系如图3.9所示。

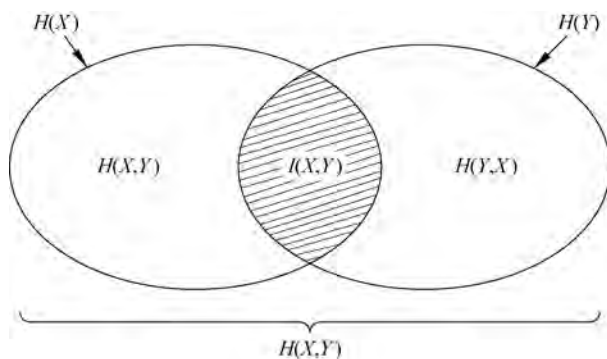


图 3.9 平均互信息量与各种熵之间关系图

例 3.3 考虑一个二进制删除信道。已知

$$P(x=0) = \frac{1}{3}, \quad P(x=1) = \frac{2}{3}$$

$$P(y=0 | x=0) = \frac{2}{3}, \quad P(y=? | x=0) = \frac{1}{6}, \quad P(y=1 | x=0) = \frac{1}{6}$$

$$P(y=0 | x=1) = \frac{1}{6}, \quad P(y=? | x=1) = \frac{1}{6}, \quad P(y=1 | x=1) = \frac{2}{3}$$

求 $H(X)$ 、 $H(Y)$ 、 $H(X, Y)$ 、 $H(X|Y)$ 、 $H(Y|X)$ 及 $I(X; Y)$ 。

解 由已知,显然有

$$\mathbf{P}_X = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \end{bmatrix}, \quad \mathbf{P}_{Y|X} = \begin{bmatrix} \frac{2}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{bmatrix}$$

则

$$\mathbf{P}_Y = \mathbf{P}_X \cdot \mathbf{P}_{Y|X} = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{12} \end{bmatrix}$$

于是有

$$H(X) = -\frac{1}{3} \log \frac{1}{3} - \frac{2}{3} \log \frac{2}{3} = 0.918$$

$$H(Y) = -\frac{1}{3} \log \frac{1}{3} - \frac{1}{6} \log \frac{1}{6} - \frac{1}{12} \log \frac{1}{12} = 1.459$$

另外,也有

$$\mathbf{P}_{XY} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{2}{3} \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{9} & \frac{1}{18} & \frac{1}{18} \\ \frac{1}{9} & \frac{1}{9} & \frac{4}{9} \end{bmatrix}$$

因此

$$H(X, Y) = - \sum_{x, y} P(x, y) \log P(x, y) = 2.170$$

这样应用各种熵之间关系式有

$$H(X | Y) = H(X, Y) - H(Y) = 0.711$$

$$H(Y | X) = H(X, Y) - H(X) = 1.252$$

$$I(X; Y) = H(X) - H(X | Y) = 0.207$$

5. 凸函数性

(1) 当信道固定即信道转移概率 $\{P(y|x)\}$ 固定时, 平均互信息量 $I(X; Y)$ 是关于信源概率分布 $\{P(x)\}$ 的上凸函数。

(2) 当信源固定即信源概率分布 $\{P(x)\}$ 固定时, 平均互信息量 $I(X; Y)$ 是关于信道转移概率 $\{P(y|x)\}$ 的下凸函数。

由式(3.14)可知

$$\begin{aligned} I(X; Y) &= \sum_{x, y} P(x, y) \log \frac{P(y|x)}{P(y)} \\ &= \sum_{x, y} P(x) P(y|x) \log \frac{P(y|x)}{P(y)} \end{aligned}$$

其中,

$$P(y) = \sum_x P(x, y) = \sum_x P(x) P(y|x)$$

因此平均互信息量 $I(X; Y)$ 是关于信源概率分布 $\{P(x)\}$ 与信道转移概率 $\{P(y|x)\}$ 的函数, 即

$$I(X; Y) = f(\{P(x)\}, \{P(y|x)\})$$

下面先举例说明信源概率分布与信道转移概率如何影响平均互信息量, 然后再证明平均互信息量的凸函数性质。

例 3.4 继续考虑如图 3.5 所示的二进制对称信道, 设其信道矩阵和先验概率分布分别为

$$\mathbf{P} = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}$$

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

其中, $\bar{p} = 1 - p, \bar{\omega} = 1 - \omega$ 。从式(3.14)知, $I(X; Y) = H(Y) - H(Y|X)$ 。因此, 为计算出 $I(X; Y)$, 需先分别计算出 $H(Y)$ 和 $H(Y|X)$ 。由于

$$P(y=0) = P(x=0, y=0) + P(x=1, y=0)$$

因此有 $P(y=0) = \omega \bar{p} + \bar{\omega} p$, 类似有 $P(y=1) = \omega p + \bar{\omega} \bar{p}$, 故得

$$H(Y) = - \sum_y P(y) \log P(y) = H(\omega \bar{p} + \bar{\omega} p)$$

此外有

$$\begin{aligned}
 H(Y|X) &= - \sum_{x,y} P(x,y) \log P(y|x) \\
 &= - \sum_{x,y} P(x) P(y|x) \log P(y|x) \\
 &= - \sum_x P(x) \sum_y P(y|x) \log P(y|x) \\
 &= -p \log p - \bar{p} \log \bar{p} \\
 &= H(p)
 \end{aligned}$$

因此

$$I(X; Y) = H(\omega \bar{p} + \bar{\omega} p) - H(p) \quad (3.16)$$

当信道固定即 p 固定时, $H(p)$ 为常数, 且有 $0 \leq H(p) \leq 1$ 。记 $q = \omega \bar{p} + \bar{\omega} p$ 。 $H(q)$ 关于 q 为上凸函数。鉴于 q 与 ω 的线性关系, 可以推出 $I(X; Y)$ 关于 ω 具有上凸性。此外, 当 $\omega = \bar{\omega} = \frac{1}{2}$ 时, $H(q) = H\left(\frac{1}{2}\right) = 1$ 达到最大, 当 $\omega = 0, 1$ 时, $H(q) = H(p)$ 达到最小。

注: 对于熵函数 $H(p_1, p_2, \dots, p_n)$, 当 $n=2$ 时, 简记 $H(p_1, p_2)$ 为 $H(p_1)$ 。

另外, 当信源固定, 即 ω 固定时, 可推知 $I(X; Y)$ 是关于 p 的下凸函数。特别地, 当 $\omega = 0, 1$ 时, $I(X; Y) = 0$; 当 $\omega = \frac{1}{2}$ 时, $I(X; Y) = 1 - H(p)$ 。二进制对称信道下的平均互信息量如图 3.10 所示。

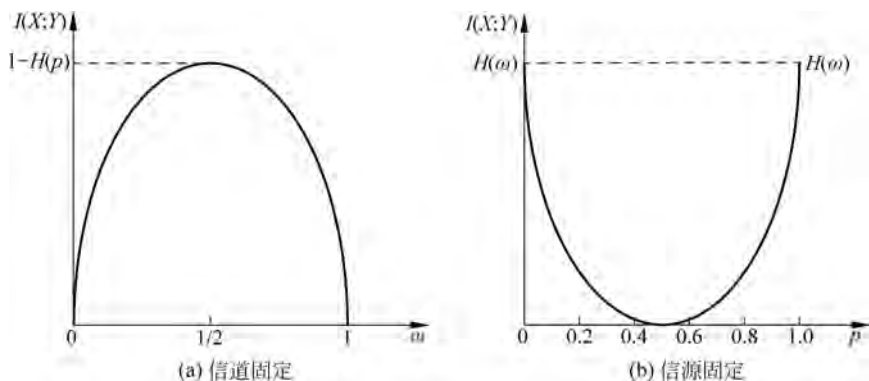


图 3.10 二进制对称信道下的平均互信息量

平均互信息量凸函数性质(1)意味着, 若固定某信道, 选择概率分布不同的信源与信道连接, 在信宿所获得的平均信息量可能是不同的。而且对于每个固定信道, 一定存在一种信源, 它的概率分布可使信宿所获得的平均信息量达到最大。而凸函数性质(2)意味着, 当信源固定后, 选择信道转移概率不同的信道来传输该信源发送的符号, 信宿所获得的平均信息量可能是不同的。而且对于每个固定信源, 一定存在一种信道, 它的信道转移概率可使信宿所获得的平均信息量达到最小。下面分别证明这两条凸函数性质。

证明 (1) 假定信道是固定的, 即 $\{P(y|x)\}$ 是已知且不变的, 而信源的概率分布即信源的 $\{P(x)\}$ 是可变的 (只要使 $\sum_x P(x) = 1$ 成立即可)。令 $\{P_1(x)\}$ 和 $\{P_2(x)\}$ 为信源的两不同分布, 相应的平均互信息量分别记为

$$I[P_1] = I(X_1; Y), \quad I[P_2] = I(X_2; Y)$$

先选取参数 $\theta \in (0, 1)$, 并记 $\bar{\theta} = 1 - \theta$ 。再选择另一种概率分布 $\{P(x)\}$ 满足

$$P(x) = \theta P_1(x) + \bar{\theta} P_2(x)$$

显然, 还有 $\sum_x P(x) = 1$ 成立。 $\{P(x)\}$ 相应的平均互信息量记为 $I[P] = I(X, Y)$ 。为证平均互信息量是关于信源概率分布的上凸函数, 需要证明如下不等式成立

$$\theta I[P_1] + \bar{\theta} I[P_2] \leq I[\theta P_1 + \bar{\theta} P_2]$$

或需要证明如下不等式成立

$$\theta I[P_1] + \bar{\theta} I[P_2] - I[P] \leq 0 \quad (3.17)$$

由式(3.14)可知有

$$\begin{aligned} & \theta I[P_1] + \bar{\theta} I[P_2] - I[P] \\ &= \theta \sum_{x,y} P_1(x, y) \log \frac{P(y|x)}{P_1(y)} + \bar{\theta} \sum_{x,y} P_2(x, y) \log \frac{P(y|x)}{P_2(y)} - \sum_{x,y} P(x, y) \log \frac{P(y|x)}{P(y)} \end{aligned} \quad (3.18)$$

由于

$$\begin{aligned} P(x, y) &= P(x)P(y|x) \\ &= [\theta P_1(x) + \bar{\theta} P_2(x)]P(y|x) \\ &= \theta P_1(x, y) + \bar{\theta} P_2(x, y) \end{aligned}$$

因此

$$\sum_{x,y} P(x, y) \log \frac{P(y|x)}{P(y)} = \sum_{x,y} [\theta P_1(x, y) + \bar{\theta} P_2(x, y)] \cdot \left(\log \frac{P(y|x)}{P(y)} \right)$$

这样式(3.18)变为

$$\begin{aligned} & \theta I[P_1] + \bar{\theta} I[P_2] - I[P] \\ &= \theta \sum_{x,y} P_1(x, y) \log \left(\frac{P(y|x)}{P_1(y)} \right) + \bar{\theta} \sum_{x,y} P_2(x, y) \log \left(\frac{P(y|x)}{P_2(y)} \right) - \sum_{x,y} P(x, y) \log \left(\frac{P(y|x)}{P(y)} \right) \\ &= \theta \sum_{x,y} P_1(x, y) \log \frac{P(y)}{P_1(y)} + \bar{\theta} \sum_{x,y} P_2(x, y) \log \frac{P(y)}{P_2(y)} - \sum_{x,y} P(x, y) \log \frac{P(y)}{P(y)} \end{aligned} \quad (3.19)$$

应用引理 2.2, 可知

$$\begin{aligned} \sum_{x,y} P_1(x, y) \log \frac{P(y)}{P_1(y)} &= \sum_y P_1(y) \log \frac{P(y)}{P_1(y)} \leq 0 \\ \sum_{x,y} P_2(x, y) \log \frac{P(y)}{P_2(y)} &= \sum_y P_2(y) \log \frac{P(y)}{P_2(y)} \leq 0 \end{aligned}$$

这样 $\theta I[P_1] + \bar{\theta} I[P_2] - I[P] \leq 0$, 从而式(3.17)成立, 即证得性质(1)。

(2) 假定信源概率分布 $\{P(x)\}$ 是固定的。令 $\{P_1(y|x)\}$ 与 $\{P_2(y|x)\}$ 表示两个不同的信道, 相应的互信息量分别记为 $I[P_1]$ 与 $I[P_2]$ 。选取 $\theta \in (0, 1)$, 并令

$$P(y|x) = \theta P_1(y|x) + \bar{\theta} P_2(y|x)$$

则 $\{P(y|x)\}$ 形成一个新信道, 其互信息量记为 $I[P]$ 。为证平均互信息量是关于信道转移概率的下凸函数, 需要证明如下不等式成立

$$I[P] = I[\theta P_1 + \bar{\theta} P_2] \leq \theta I[P_1] + \bar{\theta} I[P_2]$$

由式(3.14)有

$$\begin{aligned} & I[P] - \theta I[P_1] + \bar{\theta} I[P_2] \\ &= \sum_{x,y} P(x,y) \log \frac{P(y|x)}{P(y)} - \sum_{x,y} \theta P_1(x,y) \log \frac{P_1(y|x)}{P(y)} - \sum_{x,y} \bar{\theta} P_2(x,y) \log \frac{P_2(y|x)}{P(y)} \end{aligned} \quad (3.20)$$

既然

$$P(x,y) = P(x)P(y|x) = P(x)[\theta P_1(y|x) + \bar{\theta} P_2(y|x)]$$

并且

$$\begin{aligned} P(x,y) &= P(y|x)P(x) \\ P_1(x,y) &= P_1(y|x)P(x) \\ P_2(x,y) &= P_2(y|x)P(x) \end{aligned}$$

所以

$$P(x,y) = \theta P_1(x,y) + \bar{\theta} P_2(x,y)$$

故式(3.20)可再表示为

$$\begin{aligned} & I[P] - \theta I[P_1] + \bar{\theta} I[P_2] \\ &= \theta \sum_{x,y} P_1(x,y) \log \frac{P(y|x)}{P_1(y|x)} + \bar{\theta} \sum_{x,y} P_2(x,y) \log \frac{P(y|x)}{P_2(y|x)} \end{aligned} \quad (3.21)$$

再次应用引理 2.2, 可知

$$\begin{aligned} \sum_{x,y} P_1(x,y) \log \frac{P(y|x)}{P_1(y|x)} &= \sum_{x,y} P_1(x,y) \log \frac{P(x,y)}{P_1(x,y)} \leq 0 \\ \sum_{x,y} P_2(x,y) \log \frac{P(y|x)}{P_2(y|x)} &= \sum_{x,y} P_2(x,y) \log \frac{P(x,y)}{P_2(x,y)} \leq 0 \end{aligned}$$

从而 $I[P] - \theta I[P_1] + \bar{\theta} I[P_2] \leq 0$, 即证得性质(2)。 ■

3.3 信道容量

3.3.1 信道容量的定义

现在考虑二进制对称信道传送信息的能力。由例 3.4 可知, 当信道固定即 p 固定时, 平均互信息量 $I(X; Y) = H(\omega\bar{p} + \bar{\omega}p) - H(p)$ 是关于 ω 的上凸函数, 并当 $\omega = 1/2$ 时取得最大值

$$C = \max\{I(X; Y)\} = 1 - H(p)$$

这个最大值 C 反映该信道最大传输信息能力, 与信道输入 X 的概率分布 ω 无关, 只与信道传递概率 p 有关。因此, C 是信道特性参数, 称为信道容量。 C 与 p 的关系曲线如图 3.11 所示。容易发现, 信道容量 C 也是关于 p 的下凸函数。当 $p=0$ 时, 信道无干扰, 信道容量达到最大值 $C=1$; 当 $p=1/2$ 时, 信道干扰最大, 信道容量达到最小值 $C=0$, 即信道没有传送任何信息。

定义 3.3 对于一个离散信道, 其输入为 X , 输出为 Y 。该信道的信道容量定义为

$$C = \max_{\{P(x)\}} \{I(X; Y)\} \quad (3.22)$$

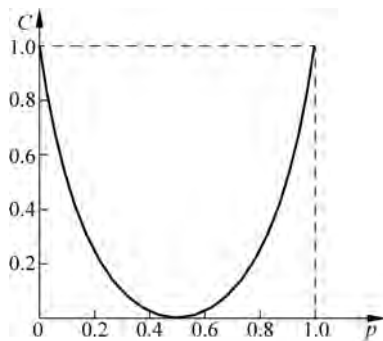


图 3.11 二进制对称信道的信道容量

信道容量 C 的单位一般也是比特(或比特/符号)。由平均互信息量的凸函数性质可知, $I(X; Y)$ 是关于 $\{P(x)\}$ 的上凸函数, 因此 C 是一直存在的。信道容量 C 是关于 $\{P(y|x)\}$ 的函数, 与 $\{P(x)\}$ 无关, 它表征了信道传送信息的最大能力。

此外, 也称信道中平均每个符号所能传送的信息量为信息传输率, 并记为 R 。实际上, 信息传输率就是平均互信息量, 因此有

$$R = I(X; Y) = H(X) - H(X|Y) \leq C$$

有时关心的是信道在单位时间内平均传输的信息量, 如果平均传输一个符号为 t 秒, 则信道每秒钟平均传输的信息量为

$$R_t = \frac{1}{t} I(X; Y) \quad (\text{bps})$$

一般称为信息传输速率。在单位时间平均传输的最大信息量则为

$$C_t = \frac{1}{t} \max_{\{P(x)\}} \{I(X; Y)\} \quad (\text{bps})$$

一般仍称之为信道容量。

例 3.5 考虑如图 3.7 所示的二进制纯删除信道, 设删除概率为 q , 并记 $\bar{q} = 1 - q$ 。假定信源概率分布为 $P(x=0) = \omega, P(x=1) = \bar{\omega} = 1 - \omega$ 。依据先验概率和信道转移概率, 可有

$$P(y=0) = \omega \cdot \bar{q} + \bar{\omega} \cdot 0 = \omega \bar{q}$$

$$P(y=1) = \omega \cdot 0 + \bar{\omega} \cdot \bar{q} = \bar{\omega} \bar{q}$$

$$P(y=?) = \omega \cdot q + \bar{\omega} \cdot q = q$$

于是有

$$\begin{aligned} H(Y) &= -[\omega \bar{q} \log(\omega \bar{q}) + \bar{\omega} \bar{q} \log(\bar{\omega} \bar{q}) + q \log(q)] \\ &= \bar{q} H(\omega) + H(q) \end{aligned}$$

又有

$$H(Y|X) = [P(x=0) + P(x=1)] \cdot H(q) = H(q)$$

故得

$$I(X; Y) = H(Y) - H(Y|X) = \bar{q} H(\omega)$$

显然, 当 $\omega = 1/2$ 时, $I(X; Y)$ 达到信道容量为

$$C = 1 - q$$

C 与 q 的关系曲线如图 3.12 所示。显然, 当 $q = 0$ 时, 信道无干扰, 信道容量达到最大值 $C = 1$; 当 $q = 1$ 时, 信道干扰最大, 信道容量达到最小值 $C = 0$, 即信道没有传送任何信息。

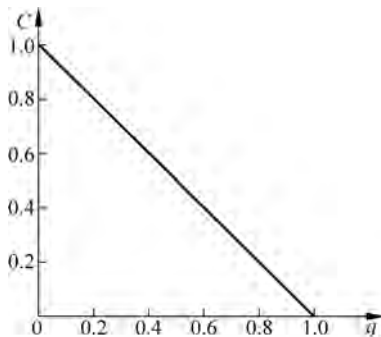


图 3.12 二进制纯删除信道的信道容量

3.3.2 无噪信道的信道容量

离散无噪信道的输出 Y 和输入 X 之间有着某种确定的关系, 按照 $H(X) \leq H(Y)$ 、 $H(X) \geq H(Y)$ 、 $H(X) = H(Y)$ 分为无损信道、确定信道、无损确定信道。

1. 无损信道

设输入符号集为 $\mathbf{A} = \{a_1, a_2, \dots, a_n\}$, 输出符号集为 $\mathbf{B} = \{b_1, b_2, \dots, b_m\}$ 。无损信道指一个输入对应多个互不相交的输出, 即

$$\begin{cases} a_i \leftrightarrow B_i, & B_i \subseteq \mathbf{B}, \quad |B_i| \geq 1 \\ B_i \cap B_j = \emptyset, & i \neq j \end{cases}$$

在无损信道的信道矩阵中,每列只有一个非零元素。例如,一个 $n=3$ 时的信道矩阵为

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{5} & \frac{3}{10} & \frac{1}{10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

对于无损信道,信宿接收到 Y 以后,必可知信源所发送的 X ,故信道的后验概率为

$$P(a_i | b_j) = \begin{cases} 0, & b_j \notin B_i \\ 1, & b_j \in B_i \end{cases}$$

因此信道疑义度(损失熵) $H(X|Y)=0$,故

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

这样

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

在这类信道中,因为信源发出符号 a_i ,并不一定能断定在信宿会收到哪个 b_j ,而是依一定概率接收 B_i 中的某一个 b_j ,因此噪声熵 $H(Y|X) \geq 0$ 。故应有

$$H(X) \leq H(Y)$$

显然,无损信道的信道容量应等于

$$C = \max_{\{P(x)\}} \{I(X; Y)\} = \max_{\{P(x)\}} H(X) = \log n$$

2. 确定信道

确定信道是指一个输出对应多个互不相交的输入,即

$$\begin{cases} A_j \leftrightarrow b_j, A_j \subseteq \mathbf{A}, \quad |A_j| \geq 1 \\ A_i \cap A_j = \emptyset, & i \neq j \end{cases}$$

其信道矩阵的每行只有一个 1,其余为 0。例如,一个 $m=2$ 时的信道矩阵为

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

确定信道的转移概率为

$$P(b_j | a_i) = \begin{cases} 0, & a_i \notin A_j \\ 1, & a_i \in A_j \end{cases}$$

因为发出某一个 a_i ,可以知道信道输出端接收到的是哪一个 b_j ,故噪声熵 $H(Y|X)=0$,且还有

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

因此

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

在这类信道中,信宿接收到某个 b_j 后,并不一定能断定信源发出的是哪个 a_i ,因而信道疑义度 $H(X|Y) \geq 0$ 。故

$$H(X) \geq H(Y)$$

确定信道的信道容量等于

$$C = \max_{\{P(x)\}} \{I(X; Y)\} = \max_{\{P(x)\}} H(Y) = \log m$$

即达到此类信道的信道容量的概率分布是使信道输出分布为等概率分布的输入分布。

3. 无损确定信道

无干扰信道也称为无损确定信道。无损确定信道的输入与输出是一一对应的关系,其信道矩阵每行每列只有一个非零元素 1。例如,当 $n=3$ 时,有 $m=3$,其信道矩阵为单位矩阵

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

一般地,对于无损确定信道,有

$$P(b_j | a_i) = P(a_i | b_j) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

$$H(Y | X) = H(X | Y) = 0$$

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

$$C = \max_{\{P(x)\}} \{I(X; Y)\} = \log n = \log m$$

3.3.3 对称信道的信道容量

考虑一个单符号离散信道,设其输入符号集为 $\mathbf{A} = \{a_1, a_2, \dots, a_n\}$,输出符号集为 $\mathbf{B} = \{b_1, b_2, \dots, b_m\}$ 。若在信道矩阵中,每行都是其他行的同一组元素不同排列,则称此类信道为输入对称信道。对于输入对称信道,应有

$$\begin{aligned} H(Y | X) &= - \sum_{i,j} P(a_i, b_j) \log P(b_j | a_i) \\ &= - \sum_i P(a_i) \sum_j P(b_j | a_i) \log P(b_j | a_i) \\ &= - \sum_j P(b_j | a_i) \log P(b_j | a_i) \\ &= H(Y | X = a_i) \end{aligned} \quad (3.23)$$

例如,信道矩阵为

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

的信道及如图 3.6 所示的二进制删除信道都是输入对称信道。另外,若在信道矩阵中,每列都是其他列的同一组元素的不同排列,则该类信道称为输出对称信道。例如,信道矩阵为

$$\mathbf{P} = \begin{bmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix}$$

的信道是输出对称信道。而信道矩阵为

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

和

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix} \quad (3.24)$$

的信道及二进制对称信道既是输入对称信道又是输出对称信道。

定义 3.4 若一个单符号离散信道的信道矩阵中,每行(列)都是其他行(列)的同一组元素的不同排列,则称该信道为对称信道。

如同二进制对称信道所显示的那样,对于一般的对称信道,当信道输入分布为等概率分布时,其输出分布也为等概分布。这是因为

$$\begin{aligned} P(b_j) &= \sum_{i=1}^n P(a_i, b_j) \\ &= \sum_{i=1}^n P(a_i) P(b_j | a_i) \\ &= \frac{1}{n} \sum_{i=1}^n P(b_j | a_i) \\ &= \frac{1}{n} \sum_{i=1}^n P_{ij} \end{aligned}$$

进一步地,由于信道矩阵 $\mathbf{P}=[P_{ij}]$ 中的第 j 列与其他 j' 列的所有元素之间只是重排关系,故有

$$\sum_{i=1}^n P_{ij} = \sum_{i=1}^n P_{ij'}, \quad j' \neq j$$

这样会有

$$P(b_j) = \frac{1}{m}, \quad j = 1, 2, \dots, m$$

这个性质很重要,可用于确定对称信道的信道容量。

命题 3.2 对于单符号离散对称信道,当其输入为等概率分布时,能达到信道容量

$$C = \log m - H(p'_1, p'_2, \dots, p'_m) \quad (3.25)$$

其中 $(p'_1, p'_2, \dots, p'_m)$ 为信道矩阵中的任一行。

证明 对称信道必然是输入对称的。故由式(3.23)知

$$H(Y | X) = H(Y | X = a_i) = H(p'_1, p'_2, \dots, p'_m)$$

其中 $(p'_1, p'_2, \dots, p'_m)$ 是信道矩阵 $\mathbf{P}$ 中的任一行。因此

$$\begin{aligned} I(X; Y) &= H(Y) - H(p'_1, p'_2, \dots, p'_m) \\ C = \max_{\{P(x)\}} I(X; Y) &= \max_{\{P(x)\}} H(Y) - H(p'_1, p'_2, \dots, p'_m) \end{aligned}$$

已知 $H(Y) \leq \log m$, 且在 $P(b_j) = \frac{1}{m}, j=1, 2, \dots, m$ 时等号成立。前面显示了对称信道有这样的性质: 若 $\{P(x)\}$ 等概率分布, 则 $\{P(y)\}$ 必为等概率分布, 故本命题成立。■

例如, 式(3.24)给出一个对称信道的信道矩阵。依据命题 3.2, 该对称信道的信道容量可计算如下

$$\begin{aligned} C &= \log m - H(p'_1, p'_2, \dots, p'_m) \\ &= \log 3 + \frac{1}{2} \log \frac{1}{2} + \frac{1}{3} \log \frac{1}{3} + \frac{1}{6} \log \frac{1}{6} \\ &= 0.126(\text{bit}) \end{aligned}$$

若信道输入符号和输出符号数目相同, 即 $n=m$, 则二进制对称信道的信道矩阵可一般化为

$$\mathbf{P} = \begin{bmatrix} \bar{p} & \frac{p}{m-1} & \frac{p}{m-1} & \dots & \frac{p}{m-1} \\ \frac{p}{m-1} & \bar{p} & \frac{p}{m-1} & \dots & \frac{p}{m-1} \\ \frac{p}{m-1} & \frac{p}{m-1} & \frac{p}{m-1} & \dots & \bar{p} \end{bmatrix}$$

称此类信道为强对称信道。强对称信道要求条件之强在于:

- (1) 一般信道的信道矩阵中各行之和为 1, 但各列之和不一定等于 1;
- (2) 一般对称信道其输入符号数不一定等于输出符号数;
- (3) 一般对称信道的总错误概率 p 不一定平均分配给 $m-1$ 个错误输出符号。

依据命题 3.2, 强对称信道的信道容量可计算为

$$C = \log n - p \log(n-1) - H(p)$$

3.3.4 一般信道的信道容量

现在讨论一般情况下单符号离散信道要到达信道容量其输入概率分布应满足的条件。由于信道固定, 平均互信息量 $I(X; Y)$ 是关于输入概率分布 $\{P(x)\}$ 的上凸函数, 因此信道容量 $C = \max_{\{P(x)\}} \{I(X; Y)\}$ 是一直存在的。 $I(X; Y)$ 是关于 $\{P(x)\}$ 的多元函数, 且有 $\sum_x P(x) = 1$ 成立。故可采用拉格朗日乘子法来求解这个极大值。

设输入符号集为 $\mathbf{A} = \{a_1, a_2, \dots, a_n\}$, 输出符号集为 $\mathbf{B} = \{b_1, b_2, \dots, b_m\}$ 。定义一个新函数

$$F(\{P(a_i)\}, \lambda) = I(X; Y) - \lambda \sum_i P(a_i)$$

其中, λ 为拉格朗日乘子, 是待定常数, 要由约束条件求解。那么最优解存在于下列方程之中:

$$\begin{cases} \frac{\partial F}{\partial P(a_k)} = \frac{\partial [I(X; Y) - \lambda \sum_i P(a_i)]}{\partial P(a_k)} = 0, & k=1, 2, \dots, n \\ \sum_i P(a_i) = 1 \end{cases} \quad (3.26)$$

式(3.26)第一个方程式中第二项等于

$$\frac{\partial}{\partial P(a_k)} \left[\lambda \sum_i P(a_i) \right] = \lambda$$

式(3.26)第一个方程式中第一项含有的平均互信息量可表示为

$$I(X; Y) = \sum_{i=1}^n \sum_{j=1}^m P(a_i) P(b_j | a_i) \log \frac{P(b_j | a_i)}{P(b_j)}$$

式中,

$$P(b_j) = \sum_{i=1}^n P(a_i) P(b_j | a_i)$$

对上式取对数并求导得

$$\begin{aligned} \frac{\partial}{\partial P(a_k)} \log P(b_j) &= \frac{1}{P(b_j)} \log e \cdot \left[\frac{\partial}{\partial P(a_k)} \cdot P(b_j) \right] \\ &= \frac{P(b_j | a_k)}{P(b_j)} \log e \end{aligned}$$

这里推导时要注意有

$$\frac{\partial}{\partial P(a_k)} P(a_i) = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}$$

因此,再对 $I(X; Y)$ 求导会有

$$\begin{aligned} \frac{\partial}{\partial P(a_k)} I(X; Y) &= \frac{\partial}{\partial P(a_k)} \sum_{i=1}^n \sum_{j=1}^m P(a_i) P(b_j | a_i) \log \frac{P(b_j | a_i)}{P(b_j)} \\ &= \sum_{j=1}^m \left[\frac{\partial}{\partial P(a_k)} \left[\sum_{i=1}^n P(a_i) P(b_j | a_i) \right] \log \frac{P(b_j | a_i)}{P(b_j)} \right] + \\ &\quad \sum_{i=1}^n \sum_{j=1}^m P(a_i) P(b_j | a_i) \left[\frac{\partial}{\partial P(a_k)} \log \frac{P(b_j | a_i)}{P(b_j)} \right] \\ &= \sum_{j=1}^m P(b_j | a_k) \log \frac{P(b_j | a_k)}{P(b_j)} - \sum_{i=1}^n \sum_{j=1}^m P(a_i) P(b_j | a_i) \frac{\partial}{\partial P(a_k)} \log P(b_j) \\ &= \sum_{j=1}^m P(b_j | a_k) \log \frac{P(b_j | a_k)}{P(b_j)} - \sum_{i=1}^n \sum_{j=1}^m P(a_i) P(b_j | a_i) \frac{P(b_j | a_k)}{P(b_j)} \log e \\ &= \sum_{j=1}^m P(b_j | a_k) \log \frac{P(b_j | a_k)}{P(b_j)} - \sum_{j=1}^m P(b_j | a_k) \log e \\ &= \sum_{j=1}^m P(b_j | a_k) \log \frac{P(b_j | a_k)}{P(b_j)} - \log e \end{aligned} \quad (3.27)$$

因此

$$\sum_{j=1}^m P(b_j | a_k) \log \frac{P(b_j | a_k)}{P(b_j)} - \log e - \lambda = 0$$

于是式(3.26)变换为

$$\begin{cases} \sum_{j=1}^m P(b_j | a_k) \log \frac{P(b_j | a_k)}{P(b_j)} = \log e + \lambda, & 1 \leq k \leq n \\ \sum_i P(a_i) = 1 \end{cases} \quad (3.28)$$

现设方程(3.28)解为 $\{\lambda, P(a_k), k=1, 2, \dots, n\}$ 。那么

$$\begin{aligned} C &= \sum_{k=1}^n \sum_{j=1}^m P(a_k) P(b_j | a_k) \log \frac{P(b_j | a_k)}{P(b_j)} \\ &= \sum_{k=1}^n P(a_k) (\log e + \lambda) = \log e + \lambda \end{aligned}$$

这样有

$$I(a_k; Y) = \sum_{j=1}^m P(b_j | a_k) \log \frac{P(b_j | a_k)}{P(b_j)} = C$$

另外,由式(3.27)可知

$$\frac{\partial}{\partial P(a_k)} I(X; Y) = I(a_k; Y) - \log e$$

因此

$$\frac{\partial}{\partial P(a_k)} I(X; Y) = \lambda, \quad k=1, 2, \dots, n$$

通过对上述优化求解表达式的推导,最终可以获得如下结果。

命题 3.3 对于一个输入符号集为 $\mathbf{A}$ 、输出符号集为 $\mathbf{B}$ 的单符号离散信道,当且仅当存在常数 C 使输入分布 $\{P(a_k), k=1, 2, \dots, n\}$ 满足

$$\begin{cases} I(a_k; Y) = C, & k \in \{i: p(a_i) \neq 0\} \\ I(a_k; Y) \leq C, & k \in \{i: p(a_i) = 0\} \end{cases} \quad (3.29)$$

时, $I(X; Y)$ 达极大值,此时 C 即为该信道的信道容量。式(3.29)可用式(3.30)取代,即

$$\begin{cases} \frac{\partial I(X; Y)}{\partial p(a_k)} = \lambda, & k \in \{i: p(a_i) \neq 0\} \\ \frac{\partial I(X; Y)}{\partial p(a_k)} \leq \lambda, & k \in \{i: p(a_i) = 0\} \end{cases} \quad (3.30)$$

证明 (充分性)简记 $P(a_k) = p_k, \mathbf{p} = \{p_k\}$ 。平均互信息量 $I(X; Y)$ 只是输入概率分布 $\mathbf{p} = \{p_k\}$ 的函数,故可简记为 $I(\mathbf{p})$ 。欲证输入概率分布 $\mathbf{p}$ 一定使 $I(\mathbf{p})$ 达到最大值,只要证对于任何其他输入分布 $\mathbf{q} = \{q_k\}$ 有 $I(\mathbf{q}) \leq I(\mathbf{p})$ 成立即可。

由于平均互信息 $I(\mathbf{p})$ 是 $\mathbf{p}$ 的上凸函数,若设 $0 < \theta < 1, \theta + \bar{\theta} = 1$, 则有

$$\theta I(\mathbf{q}) + \bar{\theta} I(\mathbf{p}) \leq I(\theta \mathbf{q} + \bar{\theta} \mathbf{p})$$

或者

$$I(\mathbf{q}) - I(\mathbf{p}) \leq \frac{1}{\theta} (I(\theta \mathbf{q} + \bar{\theta} \mathbf{p}) - I(\mathbf{p})) \quad (3.31)$$

若 $I(\mathbf{p}) = I(p_1, \dots, p_n)$, 则

$$\begin{aligned} I(\theta \mathbf{q} + \bar{\theta} \mathbf{p}) - I(\mathbf{p}) &= I[\mathbf{p} + \theta(\mathbf{q} - \mathbf{p})] - I(\mathbf{p}) \\ &= I[p_1 + \theta(q_1 - p_1), \dots, p_n + \theta(q_n - p_n)] - I(p_1, \dots, p_n) \\ &= [I(p_1 + \theta(q_1 - p_1), p_2 + \theta(q_2 - p_2), \dots, p_n + \theta(q_n - p_n)) - \\ &\quad I(p_1, p_2 + \theta(q_2 - p_2), \dots, p_n + \theta(q_n - p_n))] + \\ &\quad [I(p_1, p_2 + \theta(q_2 - p_2), p_3 + \theta(q_3 - p_3), \dots, p_n + \theta(q_n - p_n)) - \\ &\quad I(p_1, p_2, p_3 + \theta(q_3 - p_3), \dots, p_n + \theta(q_n - p_n))] + \dots + \\ &\quad [I(p_1, p_2, \dots, p_{n-1}, p_n + \theta(q_n - p_n)) - I(p_1, p_2, \dots, p_n)] \end{aligned}$$

注意到

$$\lim_{\Delta \rightarrow 0} \frac{f(x + \Delta v) - f(x)}{\Delta} = v \lim_{\Delta \rightarrow 0} \frac{f(x + \Delta v) - f(x)}{\Delta v} = v f'(x)$$

这样有

$$\begin{aligned} & \lim_{\theta \rightarrow 0} \frac{1}{\theta} [I(p_1, p_2, \dots, p_{k-1}, p_k + \theta(q_k - p_k), \dots, p_n + \theta(q_n - p_n)) - \\ & I(p_1, p_2, \dots, p_{k-1}, p_k, \dots, p_n + \theta(q_n - p_n))] \\ &= (q_k - p_k) \frac{\partial I(\mathbf{p})}{\partial p_k} \end{aligned}$$

故有

$$\lim_{\theta \rightarrow 0} \frac{1}{\theta} [I(\mathbf{p} + \theta(\mathbf{q} - \mathbf{p})) - I(\mathbf{p})] = \sum_{k=1}^n (q_k - p_k) \frac{\partial I(\mathbf{p})}{\partial p_k}$$

于是当 $\theta \rightarrow 0$ 时, 式(3.31)变为

$$I(\mathbf{q}) - I(\mathbf{p}) \leq \sum_{k=1}^n (q_k - p_k) \frac{\partial I(\mathbf{p})}{\partial p_k}$$

若假定概率分布 $\mathbf{p}$ 满足式(3.30), 那么上式可再写为

$$I(\mathbf{q}) - I(\mathbf{p}) \leq \lambda \left(\sum_{k=1}^n (q_k - p_k) \right) = 0$$

即 $I(\mathbf{q}) \leq I(\mathbf{p})$ 成立。

(必要性) 设概率分布 $\mathbf{p}$ 使 $I(\mathbf{p})$ 达到最大值, 下证明 $\mathbf{p}$ 满足式(3.30)。设另有一个分布 $\mathbf{q} = \{q_k\}$, 且有 $0 < \theta < 1, \theta + \bar{\theta} = 1$, 则 $\theta \mathbf{q} + \bar{\theta} \mathbf{p}$ 也是一个分布。故必有 $I(\theta \mathbf{q} + \bar{\theta} \mathbf{p}) - I(\mathbf{p}) \leq 0$ 或者 $\frac{1}{\theta} [I(\theta \mathbf{q} + \bar{\theta} \mathbf{p}) - I(\mathbf{p})] \leq 0$ 。由上面的充分性推导过程可知

$$\sum_{k=1}^n (q_k - p_k) \frac{\partial I(\mathbf{p})}{\partial p_k} \leq 0$$

因为 $\sum_{k=1}^n p_k = 1$, 故存在 l 使 $p_l \neq 1$ 。选择 $\mathbf{q} = \{q_k\}$ 满足

$$\begin{cases} q_l = p_l - \epsilon & k = l \\ q_j = p_j + \epsilon & k = j \\ q_k = p_k & k \neq l, k \neq j \end{cases}, \quad -p_j \leq \epsilon \leq p_l$$

这样

$$\begin{aligned} & -\epsilon \frac{\partial}{\partial p_l} I(\mathbf{p}) + \epsilon \frac{\partial}{\partial p_j} I(\mathbf{p}) \leq 0 \\ & \epsilon \frac{\partial}{\partial p_j} I(\mathbf{p}) \leq \epsilon \frac{\partial}{\partial p_l} I(\mathbf{p}) \end{aligned}$$

如令 $\frac{\partial}{\partial p_l} I(\mathbf{p}) = \lambda$, 必有

$$\epsilon \frac{\partial}{\partial p_j} I(\mathbf{p}) \leq \lambda \epsilon$$

若 $p_j = 0$, 则取 ϵ 为正数, 必有

$$\frac{\partial}{\partial p_j} I(\mathbf{p}) \leq \lambda$$

若 $p_j \neq 0$, 则 ϵ 可取正数也可取负数, 故有

$$\begin{cases} \frac{\partial}{\partial p_j} I(\mathbf{p}) \leq \lambda, & \epsilon > 0 \\ \frac{\partial}{\partial p_j} I(\mathbf{p}) \geq \lambda, & \epsilon < 0 \end{cases}$$

因此有

$$\frac{\partial}{\partial p_j} I(\mathbf{p}) = \lambda$$

说明 $\mathbf{p}$ 满足式(3.30)。

命题 3.3 表明, 当信道平均互信息量达到信道容量时, 信源符号集中除发生概率为零的符号外, 其他每个符号均对信宿提供相同的平均互信息量。在某给定的输入分布下, 若有一个输入符号 $x = a_i$ 对输出 Y 所提供的平均互信息量比其他输入符号提供的平均互信息量大, 则可以更多地使用这一符号来增大平均互信息量 $I(X; Y)$, 但是这将会改变输入符号的概率分布, 必然使这个符号的平均互信息 $I(a_i; Y)$ 减小, 而其他符号对应的平均互信息增加。所以, 经过不断调整输入符号的概率分布, 就可以使每个概率不为零的输入符号对输出 Y 提供相同的平均互信息量。命题 3.3 只给出了达到信道容量时, 最佳输入概率分布应满足的条件, 并没有给出输入符号的最佳概率分布值, 因而也没有给出信道容量的数值。另外, 命题 3.3 本身也隐含着: 达到信道容量的最佳输入分布并不一定是唯一的。在一些特殊情况下, 可利用这一命题来找出所求的输入概率分布和信道容量。

对于一个单符号离散信道的信道矩阵, 按照信道矩阵的列进行分组将矩阵分成几个子矩阵, 每个子矩阵中的每行(列)都是其他行(列)的同一组元素的不同排列, 则称这类信道为准对称信道。例如, 二进制删除信道就是典型的准对称信道。

例 3.6 证明准对称信道信道容量的输入分布为等概率分布。

证明 将准对称信道矩阵 $\mathbf{P}$ 分为一些子矩阵 $\mathbf{P}_l, l=1, 2, \dots, L$, 并保证: 在每个子矩阵 $\mathbf{P}_l$ 中, 其每行(列)都是其他行(列)的同一组元素的不同排列。在输出符号指标集中, 对应子矩阵 $\mathbf{P}_l$ 的指标子集记为 B_l 。显然, $\bigcup_{l=1}^L B_l = \{1, 2, \dots, m\}$ 。当信道输入为等概率分布即 $P(a_k) = \frac{1}{n}$ 时, 注意到

$$P(b_j) = \sum_{i=1}^n P(a_i) P(b_j | a_i) = \frac{1}{n} \sum_{i=1}^n P(b_j | a_i)$$

则有

$$\begin{aligned} I(a_k; Y) &= \sum_{j=1}^m P(b_j | a_k) \log \frac{P(b_j | a_k)}{P(b_j)} \\ &= \sum_{j=1}^m P(b_j | a_k) \log \frac{P(b_j | a_k)}{\frac{1}{n} \sum_{i=1}^m P(b_j | a_i)} \\ &= \sum_{l=1}^L \sum_{j \in B_l} P(b_j | a_k) \log \frac{P(b_j | a_k)}{\frac{1}{n} \sum_{i=1}^m P(b_j | a_i)} \end{aligned}$$

在子矩阵 $\mathbf{P}_l$ 中, 每列都是其他列的同一组元素的重排, 所以对于子矩阵 $\mathbf{P}_l$ 中的每个输出

b_j , 其概率 $P(b_j) = \frac{1}{n} \sum_{i=1}^n P(b_j | a_i)$ 都相等, 即

$$P(b_j) = \frac{1}{|B_l|} P(B_l), \quad j \in B_l$$

又因子矩阵 $\mathbf{P}_l$ 中的每行都是其他行的同一组元素的重新排列, 所以对任意 a_k , 集 $\{P(b_j | a_k), j \in B_l\}$ 与 a_k 无关, 故给定 l , 上述 $I(a_k; Y)$ 表达式中和式第 l 项

$$\sum_{j \in B_l} P(b_j | a_k) \log \frac{P(b_j | a_k)}{\frac{1}{m} \sum_{i=1}^m P(b_j | a_i)}$$

只与 B_l 有关, 与 j, k 无关, 记为 $f(B_l)$ 。故

$$I(a_i; Y) = \sum_{l=1}^L f(B_l) = C$$

为一常数。这满足命题 3.3 中的充要条件, 从而得证。 ■

例 3.7 一个单符号离散信道如图 3.13 所示, 求其信道容量。

由图 3.13 可知, 信道输入符号集为 $\mathbf{A} = \{0, 1, 2\}$, 输出符号集为 $\mathbf{B} = \{0, 1\}$, 信道矩阵为

$$\mathbf{P} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$$

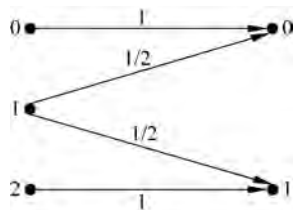


图 3.13 单符号离散信道

仔细考虑此信道, 可设想若输入符号 1 的概率等于零, 该信道就成了一一对应的信道, 接收到任何符号后对输入符号就是完全确定的。若输入符号 1 的概率不等于零, 就会增加不确定性。因此, 可先设输入分布为 $P(0) = P(2) = \frac{1}{2}, P(1) = 0$, 然后检查它是否满足命题 3.3 中的充要条件。显然

$$P(y) = \sum_{x=0}^2 P(x) P(y | x) = \frac{1}{2}$$

则有

$$I(0; Y) = \sum_{y=1}^2 P(y | 0) \log \frac{P(y | 0)}{P(y)} = \log 2 = 1$$

$$I(2; Y) = \sum_{y=1}^2 P(y | 2) \log \frac{P(y | 2)}{P(y)} = \log 2 = 1$$

$$I(1; Y) = \sum_{y=1}^2 P(y | 1) \log \frac{P(y | 1)}{P(y)} = 0$$

这满足命题 3.3 中的充要条件。因此, 信道容量为

$$C = \log 2 = 1 (\text{bit})$$

达到信道容量的最佳概率分布为

$$\left\{ P(0) = \frac{1}{2}, P(1) = 0, P(2) = \frac{1}{2} \right\}$$

对于一般的离散信道, 很难利用命题 3.3 来寻求信道容量和对应的输入概率分布, 因此

只能利用解方程组式(3.28)的方法,或者利用数值迭代算法进行求解。

3.3.5 信源与信道匹配

给定一个信道,其信道容量 C 是一定的,只有一定的信源才能使信道的信息传输率 $R = I(X; Y)$ 达到最大,即 R 等于 C 。一般情况下,信源与信道链接时,其信息传输率 R 并未到达 C 。当信源与信道链接时,若信源传输率 R 为 C ,则称信源与信道匹配,否则信道一定有冗余。为此定义

$$\begin{cases} \text{信道冗余度} = C - I(X; Y) \\ \text{信道相对冗余度} = \frac{C - I(X; Y)}{C} = 1 - \frac{I(X; Y)}{C} \end{cases}$$

这两个概念可用来描述信源与信道匹配情况。例如,在无损信道中, $C = \log n$, $I(X; Y) = H(X)$,其中 n 是信道输入符号的个数, $H(X)$ 是输入信道的信息熵。因而

$$\text{无损信道相对冗余度} = 1 - \frac{H(X)}{\log n}$$

与第2章介绍的信源冗余度比较,它实际上就是最大熵 H_0 对应的信源冗余度。所以提高无损信道信息传输率实际上就是减少信源冗余度。对于无损信道,可以通过信源编码来减少信源的冗余度,并使信息传输率到达信道容量,关于这部分内容将在第4章进行详细讨论。

3.4 离散无记忆信道

3.4.1 离散无记忆信道的数学描述

简单的离散无记忆信道数学模型如图3.14所示,用 $[X, P(b_j | a_i), Y]$ 表示。其特点是输入与输出为单个随机变量。输入 X 的取值符号集记为 $\mathbf{A} = \{a_1, a_2, \dots, a_n\}$, 输出 Y 的取值符号集记为 $\mathbf{B} = \{b_1, b_2, \dots, b_m\}$ 。简记信道传递概率 $P(b_j | a_i)$ 为 P_{ij} 。则信道矩阵可表示为 $\mathbf{P} = (P_{ij})$, 其中

$$\sum_{j=1}^m P_{ij} = 1, \quad i = 1, 2, \dots, n$$

一般离散无记忆信道的数学模型如图3.15所示,用 $[\mathbf{X}, P(\mathbf{y} | \mathbf{x}), \mathbf{Y}]$ 表示。其特点是输入与输出为随机序列。输入为 $\mathbf{X} = (X_1 X_2 \dots X_K)$, 输出 $\mathbf{Y} = (Y_1 Y_2 \dots Y_K)$ 。一般而言,不同的 X_i 可以取不同的输入符号集 $\mathbf{A}_i$, 不同的 Y_j 可以取不同的输出符号集 $\mathbf{B}_j$, 且即使相同,概率分布也可以不同。然而,如果满足

$$\begin{aligned} \mathbf{A}_i &= \mathbf{A}, \quad \mathbf{B}_i = \mathbf{B} \\ X_i &= X, \quad Y_i = Y \end{aligned}$$

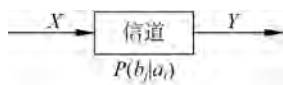


图 3.14 简单的离散无记忆信道数学模型



图 3.15 一般离散无记忆信道的数学模型

即输入符号集一样,输出符号集一样;输入概率分布一样,输出概率分布一样,那么称这类离散无记忆信道为离散无记忆 K 次扩展信道。离散无记忆扩展信道的数学模型如图 3.16 所示,用 $[X^K, P(\beta_s | \alpha_r), Y^K]$ 表示。

其中,

$$\alpha_r = (a_{r_1}, a_{r_2}, \dots, a_{r_K}), \quad r_k \in \{1, 2, \dots, r\}, \quad k = 1, 2, \dots, K$$

$$\beta_s = (b_{s_1}, b_{s_2}, \dots, b_{s_K}), \quad s_k \in \{1, 2, \dots, s\}, \quad k = 1, 2, \dots, K$$

$$\alpha_r: r = 1, 2, \dots, n^K$$

$$\beta_s: s = 1, 2, \dots, m^K$$

K 次扩展信道的信道矩阵表示为

$$\Pi = [\Pi_{rs}], \quad r = 1, 2, \dots, n^K; \quad s = 1, 2, \dots, m^K$$

其中,

$$\Pi_{rs} = P(\beta_s | \alpha_r) = \prod_{k=1}^K P(b_{s_k} | a_{r_k})$$

$$\sum_{s=1}^{m^K} \Pi_{rs} = 1, \quad r = 1, 2, \dots, n^K$$

例 3.8 设一个二进制对称信道的符号集与信道矩阵分别为

$$\mathbf{A} = \mathbf{B} = \{0, 1\}$$

$$\mathbf{P} = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}, \quad p + \bar{p} = 1$$

则其二次扩展信道的符号集与信道矩阵分别为

$$\mathbf{A}^2 = \mathbf{B}^2 = \{00, 01, 10, 11\}$$

$$\Pi = \begin{bmatrix} \bar{p}^2 & \bar{p}p & p\bar{p} & p^2 \\ \bar{p}p & \bar{p}^2 & p^2 & p\bar{p} \\ p\bar{p} & p^2 & \bar{p}^2 & \bar{p}p \\ p^2 & p\bar{p} & \bar{p}p & \bar{p}^2 \end{bmatrix} = \mathbf{P} \otimes \mathbf{P}$$

这里 $\otimes$ 表示 Kronecker 积, 定义为: 给定一个阶数为 $u \times v$ 矩阵 $\mathbf{W} = (w_{ij})$ 和一个阶数为 $p \times q$ 的矩阵 $\mathbf{V} = (v_{ij})$, 那么有阶数为 $(up) \times (vq)$ 的矩阵

$$\mathbf{W} \otimes \mathbf{V} = [w_{ij} \cdot \mathbf{V}]$$

3.4.2 离散无记忆信道的平均互信息量

单符号信道的平均互信息量定义容易推广到多符号信道。因此, 对一般离散无记忆信道, 有

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}) &= H(\mathbf{X}) - H(\mathbf{X} | \mathbf{Y}) = \sum_{r,s} P(\alpha_r, \beta_s) \log \frac{P(\alpha_r | \beta_s)}{P(\alpha_r)} \\ &= H(\mathbf{Y}) - H(\mathbf{Y} | \mathbf{X}) = \sum_{r,s} P(\alpha_r, \beta_s) \log \frac{P(\beta_s | \alpha_r)}{P(\beta_s)} \end{aligned} \quad (3.32)$$

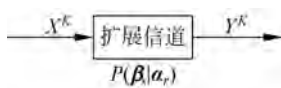


图 3.16 离散无记忆扩展信道的数学模型

对于 K 次扩展信道, 我们关心 $I(X^K; Y^K)$ 与 $I(X; Y)$ 之间的联系。直觉猜测的结论是 $I(X^K; Y^K) = KI(X; Y)$ 。这个结论对吗? 为了解决这一问题, 先考虑一般离散无记忆信道下 $I(\mathbf{X}; \mathbf{Y})$ 与 $I(X_k; Y_k)$ 之间的关系。

命题 3.4 对于一个如图 3.15 所示的离散无记忆信道 $[\mathbf{X}, P(\mathbf{y}|\mathbf{x}), \mathbf{Y}]$, 有

$$I(\mathbf{X}; \mathbf{Y}) \leq \sum_{k=1}^K I(X_k; Y_k) \quad (3.33)$$

证明 考虑信道无记忆性, 由式(3.32)可得

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y} | \mathbf{X}) \\ &= \sum_{r,s} P(\alpha_r, \beta_s) \log \frac{P(\beta_s | \alpha_r)}{P(\beta_s)} \\ &= \sum_{r,s} P(\alpha_r, \beta_s) \log \frac{\prod_{k=1}^K P(b_{s_k} | a_{r_k})}{P(\beta_s)} \end{aligned}$$

另外, 由于 $P(a_{r_k}, b_{s_k}) = \sum_{r_1, s_1; \dots; r_{k-1}, s_{k-1}; r_{k+1}, s_{k+1}; \dots; r_K, s_K} P(\alpha_r, \beta_s)$, 则有

$$\begin{aligned} \sum_{k=1}^K I(X_k; Y_k) &= \sum_{k=1}^K \sum_{r_k, s_k} P(a_{r_k}, b_{s_k}) \log \frac{P(b_{s_k} | a_{r_k})}{P(b_{s_k})} \\ &= \sum_{k=1}^K \sum_{r,s} P(\alpha_r, \beta_s) \log \frac{P(b_{s_k} | a_{r_k})}{P(b_{s_k})} \\ &= \sum_{r,s} P(\alpha_r, \beta_s) \log \left(\frac{\prod_{k=1}^K P(b_{s_k} | a_{r_k})}{\prod_{k=1}^K P(b_{s_k})} \right) \end{aligned}$$

因此, 有

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}) - \sum_{k=1}^K I(X_k; Y_k) &= \sum_{r,s} P(\alpha_r, \beta_s) \log \frac{\prod_{k=1}^K P(b_{s_k})}{P(\beta_s)} \\ &= \sum_s P(\beta_s) \log \frac{\prod_{k=1}^K P(b_{s_k})}{P(\beta_s)} \\ &= \sum_s P(\beta_s) \log \frac{Q(\beta_s)}{P(\beta_s)} \leq 0 \end{aligned}$$

最后的不等式可通过引理 2.2 推得。注意: 这里 $Q(\beta_s) = \prod_{k=1}^K P(b_{s_k})$, 且有

$$\sum_s Q(\beta_s) = \sum_{s_1 s_2 \dots s_K} \prod_{k=1}^K P(b_{s_k}) = 1$$

故有 $I(\mathbf{X}; \mathbf{Y}) \leq \sum_{k=1}^K I(X_k; Y_k)$ 。

命题 3.5 若信道 $[X, P(y|x), Y]$ 不一定是无记忆的, 而信源一定是无记忆的, 则

$$I(\mathbf{X}; \mathbf{Y}) \geq \sum_{k=1}^K I(X_k; Y_k) \quad (3.34)$$

证明 考虑信源无记忆性, 由式(3.32)可得

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}) &= H(\mathbf{X}) - H(\mathbf{X} | \mathbf{Y}) \\ &= \sum_{r,s} P(\alpha_r, \beta_s) \log \frac{P(\alpha_r | \beta_s)}{P(\alpha_r)} \\ &= \sum_{r,s} P(\alpha_r, \beta_s) \log \frac{P(\alpha_r | \beta_s)}{\prod_{k=1}^K P(a_{r_k})} \end{aligned}$$

另外, 由于 $P(a_{r_k}, b_{s_k}) = \sum_{r_1, s_1; \dots; r_{k-1}, s_{k-1}; r_{k+1}, s_{k+1}; \dots; r_K, s_K} P(\alpha_r, \beta_s)$, 则有

$$\begin{aligned} \sum_{k=1}^K I(X_k; Y_k) &= \sum_{k=1}^K \sum_{r_k, s_k} P(a_{r_k}, b_{s_k}) \log \frac{P(a_{r_k} | b_{s_k})}{P(a_{r_k})} \\ &= \sum_{k=1}^K \sum_{r,s} P(\alpha_r, \beta_s) \log \frac{P(a_{r_k} | b_{s_k})}{P(a_{r_k})} \\ &= \sum_{r,s} P(\alpha_r, \beta_s) \log \left(\frac{\prod_{k=1}^K P(a_{r_k} | b_{s_k})}{\prod_{k=1}^K P(a_{r_k})} \right) \end{aligned}$$

因此, 有

$$\begin{aligned} &-I(\mathbf{X}; \mathbf{Y}) + \sum_{k=1}^K I(X_k; Y_k) \\ &= \sum_{r,s} P(\alpha_r, \beta_s) \log \frac{\prod_{k=1}^K P(a_{r_k} | b_{s_k})}{P(\alpha_r | \beta_s)} \\ &= \sum_s P(\beta_s) \sum_r P(\alpha_r | \beta_s) \log \frac{\prod_{k=1}^K P(a_{r_k} | b_{s_k})}{P(\alpha_r | \beta_s)} \\ &= \sum_s P(\beta_s) \sum_r P(\alpha_r | \beta_s) \log \frac{Q(\alpha_r | \beta_s)}{P(\alpha_r | \beta_s)} \leq 0 \end{aligned}$$

最后的不等式可通过再次应用引理 2.2 推得。注意: 这里 $Q(\alpha_r | \beta_s) = \prod_{k=1}^K P(a_{r_k} | b_{s_k})$, 且有

$$\sum_r Q(\alpha_r | \beta_s) = \sum_{r_1 \dots r_K} \prod_{k=1}^K P(a_{r_k} | b_{s_k}) = 1$$

故有 $I(\mathbf{X}; \mathbf{Y}) \geq \sum_{k=1}^K I(X_k; Y_k)$ 。

推论 3.1 对于无扰的无记忆信道,有

$$H(\mathbf{X}) \leq \sum_{k=1}^K H(X_k) \quad (3.35)$$

证明 对于无扰的无记忆信道,有

$$\begin{aligned} H(\mathbf{X} | \mathbf{Y}) &= 0 \\ I(\mathbf{X}; \mathbf{Y}) &= H(\mathbf{X}) - H(\mathbf{X} | \mathbf{Y}) = H(\mathbf{X}) \end{aligned}$$

类似地,还有

$$\begin{aligned} H(X_k | Y_k) &= 0 \\ I(X_k; Y_k) &= H(X_k) - H(X_k | Y_k) = H(X_k) \end{aligned}$$

这样应用命题 3.4 即得证。 ■

推论 3.2 若信源和信道都是无记忆的,则

$$I(\mathbf{X}; \mathbf{Y}) = \sum_{k=1}^K I(X_k; Y_k) \quad (3.36)$$

推论 3.3 对于离散无记忆 K 次扩展信道,若信源是无记忆的,则

$$I(X^K; Y^K) = KI(X; Y) \quad (3.37)$$

此外,若信源是无记忆的,则离散无记忆 K 次扩展信道的信道容量等于

$$C_K = \max_{\{P(x^K)\}} \{I(X^K; Y^K)\} = K \max_{\{P(x)\}} I(X; Y) = KC$$

这里 $C = \max_{\{P(x)\}} I(X; Y) = \max_{\{P(x_k)\}} I(X_k; Y_k)$ 。即只有当输入信源是无记忆的,同时序列中每个分量 $X_k, k=1, 2, \dots, K$ 的分布各自达到最佳分布时, K 次扩展信道的信道容量才能达到 KC 。而一般情况下,信息序列在离散无记忆 K 次扩展信道中传输时,平均互信息量 $I(X^K; Y^K) \leq KC$ 。

3.5 串联信道的平均互信息量

前面讨论了单个离散信道,实际中常常会遇到两个或更多信道组合在一起使用的情况,有并联形式、串联形式,还有混合形式。如微波中继接力通信信道就是一种串联形式,本节侧重讨论信道串联的情况。

另外,我们常常需要在信道输出端对接收到的信号或数据进行适当的处理,这种处理称为数据处理。从广义上看,数据处理可视为一种信道,它与前面传输数据的信道是串联的关系。例如,将卫星上各种测得的科学数据编成由 0 和 1 组成的二进制码,然后以脉冲形式发送到地面,地面接收站收到的是一系列振幅不同的脉冲,将这些脉冲送入判决器,当脉冲振幅大于门限值时判为 1,当脉冲振幅小于门限值时判为 0。在这种情况下,从卫星到地面接收站可以看成是一个离散信道,其输入符号为 0 和 1 组成的二进制码,输出符号为一组不同幅度的数值。对于判决器,也可以将它看成另一个信道,其输入符号为一组不同幅度的数值,即前一信道的输出,而其输出为符号 0 和 1 组成的二进制码。实际上,这种判决器就是一种数据处理系统。因此,从卫星到判决器的输出可以看成两个信道的串联,下面将研究串联信道的平均互信息量问题。

假设有一离散单符号信道 I,其输入变量为 X ,取值 $\mathbf{A} = \{a_1, a_2, \dots, a_n\}$,输出变量为

Y , 取值 $\mathbf{B} = \{b_1, b_2, \dots, b_m\}$; 再设另有一离散单符号信道 II, 其输入变量为 Y , 输出变量为 Z , 取值 $\mathbf{C} = \{c_1, c_2, \dots, c_l\}$ 。这两个信道串联起来, 如图 3.17 所示。信道 I 的转移概率是 $P(y|x) = P(b_j|a_i)$, 信道 II 的转移概率一般与前面的符号 X 和 Y 都有关, 故 $P(z|x, y) = P(c_k|a_i, b_j)$ 。

显然, 这两个串接信道可以等效成一个总的离散信道, 如图 3.18 所示。此等效信道的输入为 X , 取值 $\mathbf{A} = \{a_1, a_2, \dots, a_n\}$, 输出为 Z , 取值 $\mathbf{C} = \{c_1, c_2, \dots, c_l\}$ 。此信道转移概率为

$$P(z|x) = \sum_y P(y|x) \cdot P(z|x, y)$$

则其信道矩阵为

$$\mathbf{P} = [P(z|x)]_{n \times l} = [P(y|x)]_{n \times m} \cdot [P(z|x, y)]_{m \times l}$$

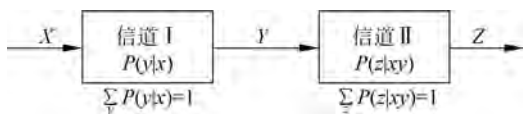


图 3.17 离散串联信道

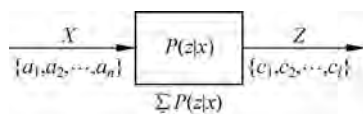


图 3.18 串联信道的等效信道

将二维联合随机变量当作一维随机变量看待, 很容易将二维随机变量形成的平均互信息量概念推广到三维。即容易给出三个随机事件集 X, Y, Z 下的平均互信息量表达式

$$\begin{aligned} I(XY; Z) &= H(Z) - H(Z|X, Y) \\ &= \sum_{x, y, z} P(x, y, z) \log \frac{P(z|x, y)}{P(z)} \end{aligned}$$

命题 3.6 串联信道中的平均互信息量满足以下关系。

$$(1) \quad I(XY; Z) \geqslant I(Y; Z)$$

其中等号成立的充要条件是对所有 x, y, z 有

$$P(z|x, y) = P(z|y)$$

$$(2) \quad I(XY; Z) \geqslant I(X; Z)$$

其中等号成立的充要条件是对所有 x, y, z 有

$$P(z|x, y) = P(z|x)$$

证明 (1) 从上述三维平均互信息量表达式可知

$$I(XY; Z) = \sum_{x, y, z} P(x, y, z) \log \frac{P(z|x, y)}{P(z)}$$

另外

$$\begin{aligned} I(Y; Z) &= \sum_{y, z} P(y, z) \log \frac{P(z|y)}{P(z)} \\ &= \sum_{x, y, z} P(x, y, z) \log \frac{P(z|y)}{P(z)} \end{aligned}$$

因此有

$$\begin{aligned} &I(Y; Z) - I(XY; Z) \\ &= \sum_{x, y, z} P(x, y, z) \log \frac{P(z|y)}{P(z|x, y)} \\ &= \sum_{x, y, z} P(x, y) P(z|x, y) \log \frac{P(z|y)}{P(z|x, y)} \end{aligned}$$

$$= \sum_{x,y} P(x,y) \sum_z P(z|x,y) \log \frac{P(z|y)}{P(z|x,y)}$$

因为

$$\begin{aligned} & \sum_z P(z|x,y) \log \frac{P(z|y)}{P(z|x,y)} \\ &= \sum_z W(z) \log \frac{V(z)}{W(z)} \leq 0 \end{aligned}$$

故有 $I(Y; Z) \leq I(XY; Z)$ 成立。注：这里最后一个不等式是应用引理 2.2 得出的；另外，从中定义了

$$W(z) = P(z|xy), \quad V(z) = P(z|y)$$

并有

$$\sum_z W(z) = 1, \quad \sum_z V(z) = 1$$

显然，从引理 2.2 可知，当且仅当 $V(z) = W(z)$ ，即 $P(z|x,y) = P(z|y)$ 时，

$$I(Y; Z) = I(XY; Z)$$

(2) 只需由 $Y \rightarrow X, X \rightarrow Y, y \rightarrow x, x \rightarrow y$ 即可类似证明出结论。■

假若信道 II 的传递概率使其输出 Z 只与输入 Y 有关，与前面的输入 X 无关，即满足 $P(z|y,x) = P(z|y)$ ，则称这两信道的输入和输出 X, Y, Z 序列构成马尔可夫链。此时有

$$P(z|x) = \sum_y P(y|x) \cdot P(z|y)$$

$$\mathbf{P} = [P(z|x)]_{n \times l} = [P(y|x)]_{n \times m} \cdot [P(z|y)]_{m \times l}$$

定理 3.1 若随机变量 X, Y, Z 构成一个马尔可夫链，则有

$$(1) \quad I(X; Z) \leq I(Y; Z)$$

$$(2) \quad I(X; Z) \leq I(X; Y)$$

证明 (1) 因为变量 X, Y, Z 构成一个马尔可夫链，故对于一切 x, y, z 有 $P(z|x,y) = P(z|y)$ ，因此由命题 3.5 可知

$$I(XY; Z) = I(Y; Z)$$

而 $I(XY; Z) \geq I(X; Z)$ ，故 $I(X; Z) \leq I(Y; Z)$ 。其中等号成立的充分必要条件是

$$P(z|x,y) = P(z|x)$$

这是因为 $I(X; Z) = I(Y; Z) \Leftrightarrow I(XY; Z) = I(X; Z)$ 。

(2) 因为 X, Y, Z 是马尔可夫链，可证 Z, Y, X 也是马尔可夫链，所以有 $P(x|y,z) = P(x|y)$ 。对于所有 x, y, z 通过类似于命题 3.5 证明过程，可推得：

① $I(ZY; X) \geq I(Y; X)$ ，当且仅当 $P(x|y,z) = P(x|y)$ 时，该式等号成立；

② $I(ZY; X) \geq I(Z; X)$ ，当且仅当 $P(x|y,z) = P(x|z)$ 时，该式等号成立。

再由①与②得

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

$$I(Y; X) \geq I(Z; X)$$

最后由平均互信息量的互易性可知

$$I(X; Y) \geq I(X; Z)$$

并且当且仅当 $P(x|y, z) = P(x|y) = P(x|z)$ 时等式成立。 ■

定理 3.1 表明,通过串联信道的传输一般只会丢失更多信息。可是,当

$$P(x|y) = P(x|z) \quad (3.38)$$

对所有 x, y, z 都成立,即串联信道的信道矩阵等于第一个信道的信道矩阵时,通过串联信道传输后不会增加信息的损失。当第二个信道是无损确定信道时,这个条件显然得到满足。如果第二个信道是数据处理系统,定理 3.1 表明,通过数据处理后,一般只会增加信息的损失,最多保持原来获得的信息,不可能比原来获得的信息有所增加。若要使数据处理后获得的平均互信息量保持不变,必须要使式(3.38)成立。总之,信息经数据处理具有不增性,称定理 3.1 为数据处理定理。

例 3.9 将两个二进制对称信道进行串联,如图 3.19 所示。已知第一个二进制对称信道的输入符号概率空间为

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

两个二进制对称信道的信道矩阵均为

$$\mathbf{P}_1 = \mathbf{P}_2 = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

并且 $0 < p < \frac{1}{2}$ 。试求 $I(X; Y)$ 与 $I(X; Z)$, 并对它们进行比较。如果有多个这样二进制对称信道进行串联,其最后总的等效信道如何?

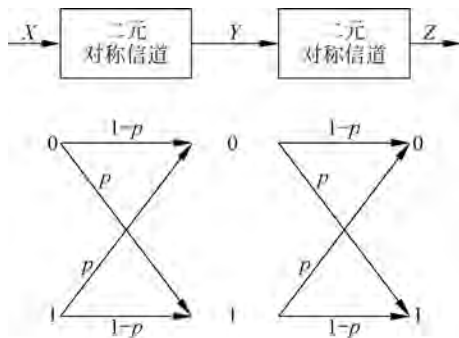


图 3.19 二进制对称信道的串联

解 串联信道的信道矩阵为

$$\mathbf{P} = \mathbf{P}_1 \cdot \mathbf{P}_2 = \begin{bmatrix} (1-p)^2 + p^2 & 2p(1-p) \\ 2p(1-p) & (1-p)^2 + p^2 \end{bmatrix}$$

这说明等效信道也是二进制对称信道,错误概率变为

$$p_2 = 2p(1-p)$$

依据平均互信息量的性质,可计算出

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

$$I(X; Z) = H(Z) - H(Z|X) = 1 - H(p_2)$$

因为 $0 < p < \frac{1}{2}$, $p_2 = 2p(1-p) > p$, 所以有

$$H(p_2) < H(p), \quad I(X; Y) > I(X; Z)$$

这表明二进制对称信道经串联后只会增加信息损失。

如果在两个二进制对称信道串联之后再串联一个同样的二进制对称信道, 并设信道输出为 W , 那么新的等效信道的平均互信息量为

$$I(X; W) = 1 - H(p_3)$$

其中错误概率为

$$p_3 = \frac{1}{2} [1 - (1 - 2p)^3]$$

一般地, 通过归纳法, 可以证明, 将 k 个同样的二进制对称信道进行串联, 并设信道输出为 V_k , 那么串联后等效信道的平均互信息量为

$$I(X; V_k) = 1 - H(p_k)$$

其中错误概率为

$$p_k = \frac{1}{2} [1 - (1 - 2p)^k]$$

图 3.20 给出了 $I(X; V_k)$ 与 p 之间的关系曲线。显然, 当串联级数增加时, 信息的损失也增加, 但损失的量逐渐减少。当 $k \rightarrow \infty$ 时, $p_k \rightarrow \frac{1}{2}$, 平均互信息量极限为

$$\lim_{k \rightarrow \infty} I(X; V_k) = 0$$

例 3.10 一个串联信道如图 3.21 所示, 设 X, Y, Z 形成马尔可夫链, 求串联信道的信道矩阵, 并讨论是否有信息损失。

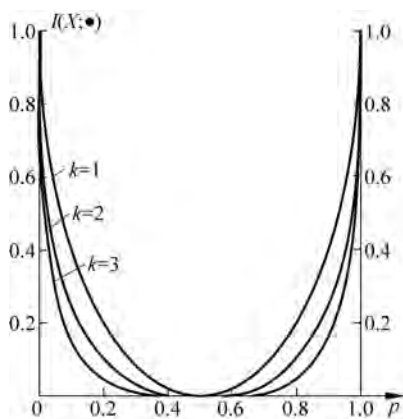


图 3.20 k 个二进制对称信道串联的平均互信息量

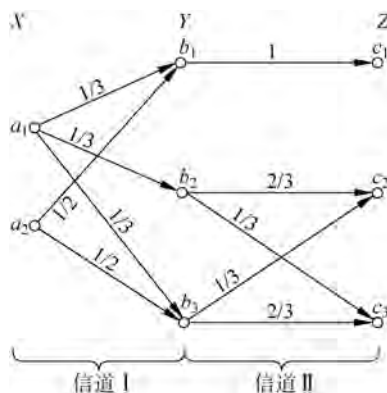


图 3.21 一个串联信道

解 由图 3.20 可知

$$\mathbf{P}_{Y|X} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}, \quad \mathbf{P}_{Z|Y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

因此

$$\mathbf{P}_{Z|X} = \mathbf{P}_{Y|X} \cdot \mathbf{P}_{Z|Y} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \end{bmatrix}$$

由于 $P(y|x) \neq P(z|x)$, 故经串联后信息有一定损失。

但是, 如果第一个信道的信道矩阵变为

$$\mathbf{P}_{Y|X} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

则有

$$\mathbf{P}_{Z|X} = \mathbf{P}_{Y|X}$$

进而

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

这表明经串联后信息没有进一步损失。这个没有信息损失的串联信道如图 3.22 所示。

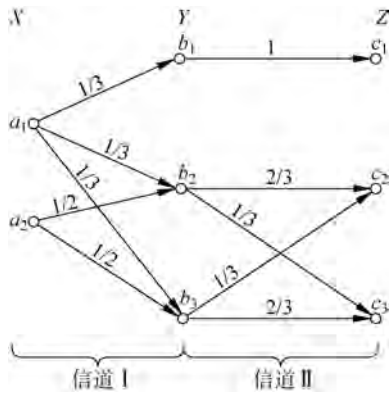


图 3.22 没有信息损失的串联信道

习题解答

3.1 考虑某二进制通信系统。已知信源 X 是离散无记忆的, 且只含有两个符号 x_0 和 x_1 , 设这两个符号出现概率分别为 $P(x_0)=1/4$ 和 $P(x_1)=3/4$ 。信宿 Y 的符号集为 $\{y_0, y_1\}$ 。已知信道转移概率为 $P(y_0|x_0)=\frac{9}{10}$, $P(y_1|x_0)=\frac{1}{10}$, $P(y_0|x_1)=\frac{1}{5}$ 和 $P(y_1|x_1)=\frac{4}{5}$ 。求:

- (1) $H(Y)$;
- (2) $I(X; Y)$;
- (3) $H(Y|X)$ 。

解 (1) 容易计算

$$H(X) = - \sum_{i=0}^1 P(x_i) \log P(x_i) = 0.811 (\text{比特 / 符号})$$

由于

$$P(x_0, y_0) = P(x_0)P(y_0|x_0) = 0.225$$

$$P(x_0, y_1) = P(x_0)P(y_1|x_0) = 0.025$$

$$P(x_1, y_0) = P(x_1)P(y_0|x_1) = 0.15$$

$$P(x_1, y_1) = P(x_1)P(y_1|x_1) = 0.60$$

$$P(y_0) = P(x_0, y_0) + P(x_1, y_0) = 0.375$$

$$P(y_1) = P(x_0, y_1) + P(x_1, y_1) = 0.625$$

因此

$$H(Y) = - \sum_{i=0}^1 P(y_i) \log P(y_i) = 0.955 (\text{比特 / 符号})$$

(2) 可计算

$$H(Y|X) = - \sum_{i,j} P(x_i, y_j) \log P(y_j | x_i) = 0.658 (\text{比特/符号})$$

因此

$$I(X; Y) = H(Y) - H(Y|X) = 0.297 (\text{比特/符号})$$

$$H(X, Y) = H(X) + H(Y|X) = 1.468 (\text{比特/符号})$$

(3) 由(1)和(2), 得

$$H(X|Y) = H(X) - I(X; Y) = 0.513 (\text{比特/符号})$$

$$H(Y|X) = - \sum_{i,j} P(x_i, y_j) \log P(y_j | x_i) = 0.658 (\text{比特/符号}) \quad \blacksquare$$

3.2 对某城市进行交通忙闲的调查, 并把天气分成晴雨两种状态, 气温分成冷暖两个状态。调查结果得到联合出现的相对频度如下:

忙	晴	冷	12	闲	晴	冷	8
		暖	8			暖	15
	雨	冷	27		雨	冷	4
		暖	16			暖	12

若把这些频度视为概率测度, 试求从天气状态和气温状态获得的关于忙闲的信息量。

解 记 X : 忙闲; Y : 晴雨; Z : 冷暖

(1) 以频率取代概率, 样本总数 102 天, 其中忙为 63 天, 闲为 39 天。因此

$$P(\text{忙}) = \frac{63}{102}, \quad P(\text{闲}) = \frac{39}{102}$$

$$H(X) = - \frac{63}{102} \log \frac{63}{102} - \frac{39}{102} \log \frac{39}{102} = 0.959 (\text{bit})$$

(2) 天气状况和冷暖状态形成概率分布: $P(yz)$

$$P(\text{晴冷}) = \frac{20}{102}, \quad P(\text{晴暖}) = \frac{23}{102}, \quad P(\text{雨冷}) = \frac{31}{102}, \quad P(\text{雨暖}) = \frac{28}{102}$$

已知上述条件下的忙闲条件分布: $P(x|yz)$

$$P(\text{忙} | \text{晴冷}) = \frac{12}{12+8} = \frac{3}{5}, \quad P(\text{闲} | \text{晴冷}) = \frac{2}{5}$$

$$P(\text{忙} | \text{晴暖}) = \frac{8}{15+8} = \frac{8}{23}, \quad P(\text{闲} | \text{晴暖}) = \frac{15}{23}$$

$$P(\text{忙} | \text{雨冷}) = \frac{27}{27+4} = \frac{27}{31}, \quad P(\text{闲} | \text{雨冷}) = \frac{4}{31}$$

$$P(\text{忙} | \text{雨暖}) = \frac{16}{16+2} = \frac{4}{7}, \quad P(\text{闲} | \text{雨暖}) = \frac{3}{7}$$

因此从天气状态和气温状态获得的关于忙闲的信息量

$$\begin{aligned} H(X|YZ) &= - \sum_{xyz} P(x, yz) \log P(x|yz) \\ &= - \sum_{xyz} P(yz) P(x|yz) \log P(x|yz) \\ &= 0.84 (\text{bit}) \quad \blacksquare \end{aligned}$$

3.3 设概率空间为 $\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ 0.6 & 0.4 \end{bmatrix}$, 信源通过一干扰信道,

接收符号为 $Y=[y_1, y_2]$, 信道传递概率如图 3.23 所示, 求:

- (1) 信道疑义度 $H(X|Y)$ 和噪声熵 $H(Y|X)$;
- (2) 接收到信息 Y 后获得的平均互信息量。

解 (1) 信道疑义度

$$H(X|Y) = - \sum_{i=1}^2 \sum_{j=1}^2 P(x_i) P(y_j | x_i) \log P(x_i | y_j)$$

由图 3.3 得

$$P(y_1 | x_1) = \frac{5}{6}, \quad P(y_2 | x_1) = \frac{1}{6}$$

$$P(y_1 | x_2) = \frac{3}{4}, \quad P(y_2 | x_2) = \frac{1}{4}$$

因为

$$P(x_1 | y_1) = \frac{P(x_1)P(y_1 | x_1)}{P(y_1)} = \frac{\frac{6}{10} \times \frac{5}{6}}{\frac{4}{5}} = \frac{5}{8}$$

$$P(x_2 | y_1) = \frac{P(x_2)P(y_1 | x_2)}{P(y_1)} = \frac{\frac{4}{10} \times \frac{3}{4}}{\frac{4}{5}} = \frac{3}{8}$$

$$P(x_1 | y_2) = \frac{P(x_1)P(y_2 | x_1)}{P(y_2)} = \frac{1}{2}$$

$$P(x_2 | y_2) = \frac{P(x_2)P(y_2 | x_2)}{P(y_2)} = \frac{1}{2}$$

所以

$$\begin{aligned} H(X|Y) &= -P(x_1)P(y_1 | x_1)\log P(x_1 | y_1) - P(x_2)P(y_1 | x_2)\log P(x_2 | y_1) - \\ &\quad P(x_1)P(y_2 | x_1)\log P(x_1 | y_2) - P(x_2)P(y_2 | x_2)\log P(x_2 | y_2) \\ &= -\frac{6}{10} \times \frac{5}{6} \log \frac{5}{8} - \frac{4}{10} \times \frac{3}{4} \log \frac{3}{8} - \frac{6}{10} \times \frac{1}{6} \log \frac{1}{2} - \frac{4}{10} \times \frac{1}{4} \log \frac{1}{2} \\ &\approx 0.3390 + 0.4245 + 0.1 + 0.1 \\ &\approx 0.9635 (\text{比特/符号}) \end{aligned}$$

噪声熵

$$\begin{aligned} H(Y|X) &= - \sum_{i=1}^2 \sum_{j=1}^2 P(x_i) P(y_j | x_i) \log P(y_j | x_i) \\ &= -P(x_1)P(y_1 | x_1)\log P(y_1 | x_1) - P(x_2)P(y_1 | x_2)\log P(y_1 | x_2) - \\ &\quad P(x_1)P(y_2 | x_1)\log P(y_2 | x_1) - P(x_2)P(y_2 | x_2)\log P(y_2 | x_2) \\ &= -\frac{6}{10} \times \frac{5}{6} \log \frac{5}{6} - \frac{4}{10} \times \frac{3}{4} \log \frac{3}{4} - \frac{6}{10} \times \frac{1}{6} \log \frac{1}{6} - \frac{4}{10} \times \frac{1}{4} \log \frac{1}{4} \\ &\approx 0.1315 + 0.2585 + 0.1245 + 0.2 \\ &\approx 0.7145 (\text{比特/符号}) \end{aligned}$$

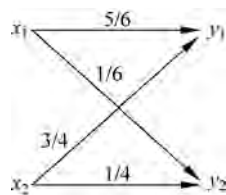


图 3.23 习题 3.3 图

(2) 接收到消息 Y 后获得的平均互信息量 $I(X; Y)$ 为

$$\begin{aligned}
 I(X; Y) &= \sum_{i=1}^2 \sum_{j=1}^2 P(x_i) P(y_j | x_i) \log \frac{P(x_i | y_j)}{P(x_i)} \\
 &= H(X) - H(X | Y) \\
 &= H(Y) - H(Y | X) \\
 &= \sum_{i=1}^2 \sum_{j=1}^2 P(x_i) P(y_j | x_i) I(x_i; y_j) \\
 &\approx 0.0075 (\text{比特/符号})
 \end{aligned}$$

3.4 设 8 个等概率分布的消息通过传递概率为 p 的二进制对称信道进行传送。8 个消息相应编成下述码字：

$$\begin{aligned}
 M_1 &= 0000, & M_2 &= 0101, & M_3 &= 0110, & M_4 &= 0011 \\
 M_5 &= 1001, & M_6 &= 1010, & M_7 &= 1100, & M_8 &= 1111
 \end{aligned}$$

试问：

- (1) 接收到第一个数字 0 与 M_1 之间的互信息量是多少？
- (2) 接收到第二个数字也是 0 时, 得到多少关于 M_1 的附加互信息量？
- (3) 接收到第三个数字仍是 0 时, 又增加多少关于 M_1 的互信息量？
- (4) 接收到第四个数字还是 0 时, 再增加了多少关于 M_1 的互信息量？
- (5) 接收到全部四个数字 0000 后, 获得多少关于 M_1 的互信息量？

解 信源

$$\begin{array}{cccccccc}
 M: & M_1 & M_2 & M_3 & M_4 & M_5 & M_6 & M_7 & M_8 \\
 P(M_i): & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8}
 \end{array}$$

对应码字: 0000 0101 0110 0011 1001 1010 1100 1111

信道为二元对称无记忆信道, 如图 3.24 所示。信道传递矩阵为

$$\begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}, \quad \bar{p} + p = 1$$

消息 M_i 与码字一一对应, 所以设

$$M_i = (x_{i_1} x_{i_2} x_{i_3} x_{i_4}), \quad x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4} \in \{0, 1\}, \quad i = 1, 2, \dots, 8$$

又设接收序列为

$$Y = (y_1 y_2 y_3 y_4), \quad y_1, y_2, y_3, y_4 \in \{0, 1\}$$

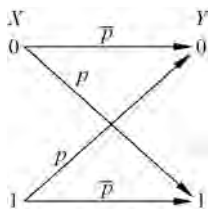


图 3.24 习题 3.4 图

(1) 接收到第一个数字为 0, 即 $y_1 = 0$ 。那么, 接收到第一个数字 0 与 M_1 之间的互信息量为

$$I(M_1; y_1 = 0) = \log \frac{P(y_1 = 0 | M_1)}{P(y_1 = 0)}$$

因为信道是无记忆信道,所以

$$\begin{aligned} P(y_1=0 | M_1) &= P(y_1=0 | x_{1_1}x_{1_2}x_{1_3}x_{1_4}=0000) \\ &= P(y_1=0 | x_{1_1}=0) = P(0 | 0) = \bar{p} \end{aligned}$$

同理,得

$$\begin{aligned} P(y_1=0 | M_i) &= P(y_1=0 | x_{i_1}x_{i_2}x_{i_3}x_{i_4}) = P(y_1=0 | x_{i_1}) \\ x_{i_1}x_{i_2}x_{i_3}x_{i_4} &\in \{0,1\}, \quad i=1,2,\dots,8 \end{aligned}$$

输出第一个符号 $y_1=0$ 时,有可能是 8 个消息中任意一个第一个数字传送来的。所以

$$\begin{aligned} P(y_1=0) &= \sum_{i=1}^8 P(M_i)P(y_1=0 | M_i) \\ &= \frac{1}{8}[P(y_1=0 | x_{1_1}=0) + P(y_1=0 | x_{2_1}=0) + P(y_1=0 | x_{3_1}=0) + \\ &\quad P(y_1=0 | x_{4_1}=0) + P(y_1=0 | x_{5_1}=1) + P(y_1=0 | x_{6_1}=1) + \\ &\quad P(y_1=0 | x_{7_1}=1) + P(y_1=0 | x_{8_1}=1)] \\ &= \frac{1}{8}[4P(0 | 0) + 4P(0 | 1)] \\ &= \frac{1}{8}[4(\bar{p} + p)] = \frac{1}{2} \end{aligned}$$

故得

$$I(M_1; y_1=0) = 1 + \log_2 \bar{p} \text{ (bit)}$$

(2) 接收到第二个数字也是 0 时($y_2=0$),得到关于 M_1 的附加互信息量为

$$I(M_1; y_2=0 | y_1=0) = I(M_1; y_1y_2=00) - I(M_1; y_1=0)$$

其中,

$$I(M_1; y_1y_2=00) = \log \frac{P(y_1y_2=00 | M_1)}{P(y_1y_2=00)}$$

同理,因为信道是无记忆信道,所以

$$\begin{aligned} P(y_1y_2=00 | M_i) &= P(y_1y_2=00 | x_{i_1}x_{i_2}x_{i_3}x_{i_4}) \\ &= P(y_1y_2=00 | x_{i_1}x_{i_2}) \\ &= P(y_1=0 | x_{i_1})P(y_2=0 | x_{i_2}) \end{aligned}$$

$$x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4} \in \{0,1\}, \quad i=1,2,\dots,8$$

得

$$\begin{aligned} P(y_1y_2=00 | M_1) &= P(y_1=0 | x_{1_1}=0)P(y_2=0 | x_{1_2}=0) \\ &= P(0 | 0)P(0 | 0) = \bar{p}^2 \end{aligned}$$

输出端出现第一个符号和第二个符号都为 0($y_1=0, y_2=0$)的概率为

$$\begin{aligned} P(y_1y_2=00) &= \sum_{i=1}^8 P(M_i)P(y_1y_2=00 | M_i) \\ &= \frac{1}{8}[P(y_1=0 | x_{1_1}=0)P(y_2=0 | x_{1_2}=0) + \\ &\quad P(y_1=0 | x_{2_1}=0)P(y_2=0 | x_{2_2}=1) + \end{aligned}$$

$$\begin{aligned}
& P(y_1=0 | x_{3_1}=0)P(y_2=0 | x_{3_2}=1) + \\
& P(y_1=0 | x_{4_1}=0)P(y_2=0 | x_{4_2}=0) + \\
& P(y_1=0 | x_{5_1}=1)P(y_2=0 | x_{5_2}=0) + \\
& P(y_1=0 | x_{6_1}=1)P(y_2=0 | x_{6_2}=0) + \\
& P(y_1=0 | x_{7_1}=1)P(y_2=0 | x_{7_2}=1) + \\
& P(y_1=0 | x_{8_1}=1)P(y_2=0 | x_{1_2}=1)
\end{aligned}$$

即

$$\begin{aligned}
P(y_1 y_2 = 00) &= \frac{1}{8} [P(0|0)P(0|0) + P(0|0)P(0|1) + \\
& P(0|0)P(0|1) + P(0|0)P(0|0) + \\
& P(0|1)P(0|0) + P(0|1)P(0|0) + \\
& P(0|1)P(0|1) + P(0|1)P(0|1)] \\
&= \frac{1}{8} [2\bar{p}^2 + 4\bar{p}p + 2p^2] \\
&= \frac{1}{4} [\bar{p}^2 + 2\bar{p}p + p^2] \\
&= \frac{1}{4} (\bar{p} + p)^2 \\
&= \frac{1}{4}
\end{aligned}$$

所以

$$I(M_1; y_1 y_2 = 00) = \log \frac{\bar{p}^2}{\frac{1}{4}} = 2(1 + \log_2 \bar{p}) \text{ (bit)}$$

得附加互信息量为

$$I(M_1; y_2 = 0 | y_1 = 0) = 1 + \log_2 \bar{p} \text{ (bit)}$$

(3) 接收到第三个数字仍是 $0(y_3=0)$ 时关于 M_1 的互信息又增加为

$$I(M_1; y_3 = 0 | y_1 y_2 = 00) = I(M_1; y_1 y_2 y_3 = 000) - I(M_1; y_1 y_2 = 00)$$

其中,

$$I(M_1; y_1 y_2 y_3 = 000) = \log \frac{P(y_1 y_2 y_3 = 000 | M_1)}{P(y_1 y_2 y_3 = 000)}$$

同理,得

$$\left. \begin{aligned}
& P(y_1 y_2 y_3 = 000 | M_i) \\
&= P(y_1 y_2 y_3 = 000 | x_{i_1} x_{i_2} x_{i_3} x_{i_4}) \\
&= P(y_1 y_2 y_3 = 000 | x_{i_1} x_{i_2} x_{i_3}) \\
&= P(y_1 = 0 | x_{i_1}) P(y_2 = 0 | x_{i_2}) P(y_3 = 0 | x_{i_3})
\end{aligned} \right\} x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4} \in \{0, 1\}, \quad i = 1, 2, \dots, 8$$

得

$$P(y_1 y_2 y_3 = 000 | M_1) = P(0|0)P(0|0)P(0|0) = \bar{p}^3$$

输出端出现三个符号都为 $0(y_1 y_2 y_3 = 000)$ 的概率为

$$\begin{aligned}
P(y_1 y_2 y_3 = 000) &= \sum_{i=1}^8 P(M_i) P(y_1 y_2 y_3 = 000 | M_i) \\
&= \frac{1}{8} [\bar{p}^3 + \bar{p}^2 p + \bar{p} p^2 + \bar{p}^2 p + \bar{p}^2 p + \bar{p} p^2 + \bar{p} p^2 + p^3] \\
&= \frac{1}{8} [\bar{p}^3 + 3\bar{p}^2 p + 3\bar{p} p^2 + p^3] \\
&= \frac{1}{8} (\bar{p} + p)^3 = \frac{1}{8}
\end{aligned}$$

所以

$$I(M_1; y_1 y_2 y_3 = 000) = 3(1 + \log_2 \bar{p}) (\text{bit})$$

而又增加的互信息量为

$$I(M_1; y_3 = 0 | y_1 y_2 = 00) = 1 + \log_2 \bar{p} (\text{bit})$$

(4) 接收到第四数字还是 0 时 (即 $y_1 y_2 y_3 y_4 = 0000$) 再增加了关于 M_1 的互信息为

$$\begin{aligned}
I(M_1; y_4 = 0 | y_1 y_2 y_3 = 000) \\
= I(M_1; y_1 y_2 y_3 y_4 = 0000) - I(M_1; y_1 y_2 y_3 = 000)
\end{aligned}$$

同理,有

$$\begin{aligned}
&P(y_1 y_2 y_3 y_4 = 0000 | M_i) \\
&= P(y_1 = 0 | x_{i_1}) P(y_2 = 0 | x_{i_2}) P(y_3 = 0 | x_{i_3}) P(y_4 = 0 | x_{i_4}) \\
&\quad x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4} \in \{0, 1\}, \quad i = 1, 2, \dots, 8 \\
&P(y_1 y_2 y_3 y_4 = 0000) = \sum_{i=1}^8 P(M_i) P(y_1 y_2 y_3 y_4 = 0000 | M_i) \\
&= \frac{1}{8} [\bar{p}^4 + 6\bar{p}^2 p^2 + p^4]
\end{aligned}$$

所以

$$\begin{aligned}
I(M_1; y_1 y_2 y_3 y_4 = 0000) &= \log \frac{P(y_1 y_2 y_3 y_4 = 0000 | M_1)}{P(y_1 y_2 y_3 y_4 = 0000)} \\
&= \log \frac{\bar{p}^4}{\frac{1}{8} [\bar{p}^4 + 6\bar{p}^2 p^2 + p^4]} \\
&= 3 + 4\log_2 \bar{p} - \log_2 (\bar{p}^4 + 6\bar{p}^2 p^2 + p^4) (\text{bit})
\end{aligned}$$

所以又增加了关于 M_1 的互信息量为

$$I(M_1; y_4 = 0 | y_1 y_2 y_3 y_4 = 0000) = \log_2 \bar{p} - \log_2 (\bar{p}^4 + 6\bar{p}^2 p^2 + p^4) (\text{bit})$$

(5) 接到全部四个数字 0000 后获得关于 M_1 的互信息量为

$$I(M_1; y_1 y_2 y_3 y_4 = 0000) = 3 + 4\log_2 \bar{p} - \log_2 (\bar{p}^4 + 6\bar{p}^2 p^2 + p^4) (\text{bit}) \quad \blacksquare$$

3.5 设有一批电阻,按阻值分,70%是 $2\text{k}\Omega$,30%是 $5\text{k}\Omega$;按功耗分,64%是 $1/8\text{W}$,其余是 $1/4\text{W}$ 。现已知 $2\text{k}\Omega$ 阻值的电阻中 80%是 $1/8\text{W}$ 。问通过测量阻值可以平均得到的关于瓦数的信息量是多少?

解 根据题意,设电阻的阻值为事件 X ,电阻的功率为事件 Y 。事件 X 、事件 Y 的信源概率空间分别为

$$\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} x_1 = 2\text{k}\Omega & x_2 = 5\text{k}\Omega \\ 0.7 & 0.3 \end{bmatrix}, \quad \begin{bmatrix} Y \\ P(y) \end{bmatrix} = \begin{bmatrix} y_1 = 1/8\text{W} & y_2 = 1/4\text{W} \\ 0.64 & 0.36 \end{bmatrix}$$

并已知 $P(y_1|x_1)=0.8$, 所以 $P(y_2|x_1)=1-P(y_1|x_1)=0.2$ 。

首先计算出 $P(y_1|x_2)$ 和 $P(y_2|x_2)$, 根据概率关系, 它们满足

$$\begin{cases} P(y_1) = P(x_1)P(y_1|x_1) + P(x_2)P(y_1|x_2) \\ P(y_2) = P(x_1)P(y_2|x_1) + P(x_2)P(y_2|x_2) \\ P(y_1|x_2) + P(y_2|x_2) = 1 \end{cases}$$

所以

$$\begin{cases} 0.64 = 0.7 \times 0.8 + 0.3P(y_1|x_2) \\ 0.36 = 0.7 \times 0.2 + 0.3P(y_2|x_2) \\ P(y_1|x_2) + P(y_2|x_2) = 1 \end{cases}$$

整理得

$$P(y_1|x_2) = \frac{8}{30} \approx 0.26$$

$$P(y_2|x_2) = \frac{22}{30} \approx 0.73$$

上面两值表示 $5\text{k}\Omega$ 阻值的电阻中约 73% 功率为 $1/4\text{W}$, 而 26% 是 $1/8\text{W}$ 。通过测量电阻阻值可以得到关于瓦数的平均信息量就是平均互信息量 $I(X; Y)$, 所以

$$\begin{aligned} I(X; Y) &= \sum_{i=1}^2 \sum_{j=1}^2 P(x_i)P(y_j|x_i) \log \frac{P(y_j|x_i)}{P(y_j)} \\ &= [p(x_1)P(y_1|x_1) \log \frac{P(y_1|x_1)}{P(y_1)} + p(x_1)P(y_2|x_1) \log \frac{P(y_2|x_1)}{P(y_2)} + \\ &\quad p(x_2)P(y_1|x_2) \log \frac{P(y_1|x_2)}{P(y_1)} + p(x_2)P(y_2|x_2) \log \frac{P(y_2|x_2)}{P(y_2)}] \\ &= \left[0.7 \times 0.8 \log \frac{0.8}{0.64} + 0.7 \times 0.2 \log \frac{0.2}{0.36} + 0.3 \times 0.26 \log \frac{0.26}{0.64} + \right. \\ &\quad \left. 0.3 \times 0.73 \log \frac{0.73}{0.36} \right] \\ &\approx 0.180 - 0.119 - 0.101 + 0.226 \\ &= 0.186(\text{bit}) \end{aligned}$$

3.6 若三个离散随机变量, 有如下关系: $X+Y=Z$, 其中 X 和 Y 相互统计独立。试证明:

- (1) $I(X; Z) = H(Z) - H(Y)$;
- (2) $I(XY; Z) = H(Z)$;
- (3) $I(X; YZ) = H(X)$;
- (4) $I(Y; Z|X) = H(Y)$;
- (5) $I(X; Y|Z) = H(X|Z) = H(Y|Z)$ 。

证明 离散随机变量 $Z = X + Y$, X 和 Y 统计独立, 如图 3.25 所示。

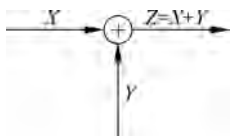


图 3.25 习题 3.6 图

设 $x \in X, y \in Y, z \in Z$, 满足 $z = x + y$, 因为

$$P_z(z) = \sum_Y P(y, z) = \sum_Y P(y, z = x + y) = \sum_Y P(y, x = z - y)$$

且 X 和 Y 统计独立, 所以

$$P_z(z) = \sum_Y P(y)P(x), \quad z = x + y \quad (3.39)$$

同理

$$P_z(z) = \sum_X P(x, z) = \sum_X P(x, z = x + y) = \sum_X P(x, x = z - y)$$

因为 X 和 Y 统计独立, 所以

$$P_z(z) = \sum_X P(x)P(y), \quad z = x + y \quad (3.40)$$

又因为 $P_z(z) = \sum_Y P(z | y)P(y)$, 与式(3.39)比较可得

$$P(z | y) = \begin{cases} P(x), & z = x + y \\ 0, & z \neq x + y \end{cases} \quad (3.41)$$

同理得

$$P(z | x) = \begin{cases} P(y), & z = x + y \\ 0, & z \neq x + y \end{cases} \quad (3.42)$$

因为要满足 $z = x + y$, 所以

$$P(z | xy) = \begin{cases} 1, & z = x + y \\ 0, & z \neq x + y \end{cases} \quad (3.43)$$

$$P(y | xz) = \begin{cases} 1, & z = x + y \\ 0, & z \neq x + y \end{cases} \quad (3.44)$$

$$P(x | yz) = \begin{cases} 1, & z = x + y \\ 0, & z \neq x + y \end{cases} \quad (3.45)$$

(1) 因为

$$H(Z | X) = - \sum_Z \sum_X P(x)P(z | x) \log P(z | x)$$

由式(3.42)得

$$\begin{aligned} H(Z | X) &= - \sum_Z \sum_X P(x)P(z | x) \log P(z | x) \\ &= - \sum_{z=x+y} \sum_X P(x)P(y) \log P(y) \\ &= - \sum_Y \sum_X P(x)P(y) \log P(y) \\ &= - \sum_Y P(y) \log P(y) = H(Y) \end{aligned}$$

所以

$$\begin{aligned} I(X; Z) &= H(Z) - H(Z | X) \\ &= H(Z) - H(Y) \end{aligned}$$

$$(2) I(XY; Z) = H(Z) - H(Z|XY)$$

根据式(3.43)可得

$$H(Z|XY) = 0$$

所以

$$I(XY; Z) = H(Z)$$

(3) 根据式(3.45)可得

$$H(X|YZ) = 0$$

所以

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

$$(4) I(Y; Z|X) = H(Y|X) - H(Y|XZ)$$

根据式(3.44)可得 $H(Y|XZ) = 0$, 又因 X 和 Y 统计独立, 所以

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

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

$$(5) I(X; Y|Z) = H(X|Z) - H(X|ZY) = H(X|Z)$$

因为

$$H(X|ZY) = 0$$

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

$$H(Y|XZ) = 0$$

所以

$$I(X; Y|Z) = H(X|Z) = H(Y|Z) \quad \blacksquare$$

3.7 设 X 和 Y 是两个相互统计独立的二元随机变量, 其取 0 或 1 的概率为等概率分布。定义另一个二元随机变量 Z , 而且 $Z = X \cdot Y$ (一般乘积)。试计算:

$$(1) I(X; Y), I(X; Z), I(Y; Z);$$

$$(2) I(X; Y|Z), I(Y; X|Z), I(Z; X|Y), I(Z; Y|X);$$

$$(3) I(XY; Z), I(X; YZ), I(Y; XZ)。$$

解 X 和 Y 是两个相互统计独立的二元随机变量, 其概率空间分别为

$$\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} x=0 & x=1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} Y \\ P(y) \end{bmatrix} = \begin{bmatrix} y=0 & y=1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

因为 $Z = X \cdot Y$ (一般乘积), 所以 Z 的概率空间为

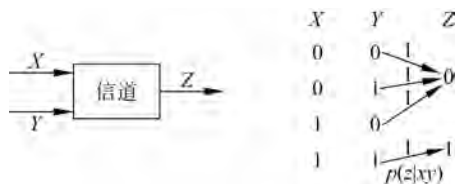
$$\begin{bmatrix} Z = XY \\ P(z) \end{bmatrix} = \begin{bmatrix} z=0 & z=1 \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}$$

X, Y 与 Z 之间关系相当的信道如图 3.26 所示, 其信道矩阵为

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

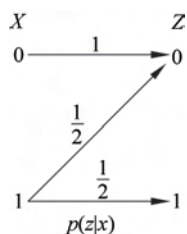
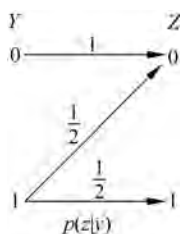
即

$$P(z|xy) = \begin{cases} 1, & z = xy \\ 0, & z \neq xy \end{cases}$$

图 3.26 X, Y 与 Z 之间关系相当的信道

由上面 $XY \rightarrow Z$ 的信道, 可得相当于 $X \rightarrow Z, Y \rightarrow Z$ 的信道如图 3.27 和图 3.28 所示。信道矩阵为

$$\mathbf{P}' = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad \mathbf{P}'' = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

图 3.27 $P(z|x)$ 图 3.28 $P(z|y)$

因为 $P(z) = \sum_x P(x)P(z|x)$, 而 $P(x)$ 和 $P(z)$ 已知, 所以也可以由此式求出 $P(z|x)$, 同理求出 $P(y|z)$ 。

又可由

$$P(x|z) = \frac{P(x)P(z|x)}{P(z)}, \quad x \in X, z \in Z = XY$$

$$P(y|z) = \frac{P(y)P(z|y)}{P(z)}, \quad y \in Y, z \in Z = XY$$

求得

$$\begin{cases} P(x=0|z=0) = \frac{2}{3} \\ P(x=0|z=1) = 0 \\ P(x=1|z=0) = \frac{1}{3} \\ P(x=1|z=1) = 1 \end{cases} \quad \begin{cases} P(y=0|z=0) = \frac{2}{3} \\ P(y=0|z=1) = 0 \\ P(y=1|z=0) = \frac{1}{3} \\ P(y=1|z=1) = 1 \end{cases}$$

又有

$$P(y|xz) = \frac{P(xyz)}{P(xz)} = \frac{P(xy)P(z|xy)}{P(x)P(z|x)}$$

因为 X 和 Y 统计独立, 所以

$$P(y|xz) = \frac{P(y)P(z|xy)}{P(z|x)}, \quad x \in X, y \in Y, z \in Z = XY$$

同理有

$$P(x | yz) = \frac{P(x)P(z | xy)}{P(z | y)}, \quad x \in X, y \in Y, z \in Z = XY$$

可求得

$$\left\{ \begin{array}{l} P(y=0 | x=0 \quad z=0) = \frac{1}{2} \\ P(y=1 | x=0 \quad z=0) = \frac{1}{2} \\ P(y=0 | x=1 \quad z=0) = 1 \\ P(y=1 | x=1 \quad z=1) = 1 \\ P(y | xz) = 0 \quad z \neq xy \end{array} \right\} \left\{ \begin{array}{l} P(x=0 | y=0 \quad z=0) = \frac{1}{2} \\ P(x=1 | y=0 \quad z=0) = \frac{1}{2} \\ P(x=0 | y=1 \quad z=0) = 1 \\ P(x=1 | y=1 \quad z=1) = 1 \\ P(x | yz) = 0 \quad z \neq xy \end{array} \right.$$

$$(1) H(X) = H\left(\frac{1}{2}\right) = 1 \quad (\text{比特/符号})$$

$$H(Y) = 1 \quad (\text{比特/符号})$$

$$H(Z) = H\left(\frac{3}{4}, \frac{1}{4}\right) \approx 0.811 \quad (\text{比特/符号})$$

因为 X 和 Y 统计独立, 所以

$$H(X | Y) = H(X) = 1 \quad (\text{比特/符号})$$

$$\begin{aligned} H(X | Z) &= - \sum_x \sum_{z=xy} P(x)P(z | x)P(x | z) \\ &= -\frac{1}{2} \times 1 \log \frac{2}{3} - \frac{1}{2} \times 0 \log 0 - \frac{1}{2} \times \frac{1}{2} \log \frac{1}{3} - \frac{1}{2} \times \frac{1}{2} \log 1 \\ &= \frac{3}{4} \log 3 - \frac{1}{2} \approx 0.689 \quad (\text{比特/符号}) \end{aligned}$$

同理

$$H(Y | Z) \approx 0.689 \quad (\text{比特/符号})$$

$$H(Z | X) = 0.5 \quad (\text{比特/符号})$$

$$H(Z | Y) = 0.5 \quad (\text{比特/符号})$$

因为 X 和 Y 统计独立, 则

$$I(X; Y) = H(X) - H(X | Y) = H(X) - H(X) = 0$$

$$I(X; Z) = H(Z) - H(Z | X) = H(X) - H(X | Z) \approx 0.311 \quad (\text{比特/符号})$$

$$I(Y; Z) = H(Z) - H(Z | Y) = H(Y) - H(Y | Z) \approx 0.311 \quad (\text{比特/符号})$$

(2) 因为

$$P(z | xy) = \begin{cases} 0, & z \neq xy \\ 1, & z = xy \end{cases}$$

所以

$$H(Z | XY) = 0$$

$$\begin{aligned} H(X | YZ) &= - \sum_x \sum_y \sum_{z=xy} P(xyz) \log P(x | yz) \\ &= - \sum_x \sum_y \sum_{z=xy} P(xy) P(z | xy) \log P(x | yz) \\ &= -\frac{1}{4} \times 1 \log \frac{1}{2} - \frac{1}{4} \times 1 \log \frac{1}{2} - \frac{1}{4} \times 1 \log 1 - \frac{1}{4} \times 1 \log 1 \\ &= 0.5 \quad (\text{比特/符号}) \end{aligned}$$

同理

$$H(Y | XZ) = 0.5 \quad (\text{比特/符号})$$

$$\begin{aligned} I(X; Y | Z) &= H(X | Z) - H(X | YZ) = H(Y | Z) - H(Y | XZ) \\ &\approx 0.189 \quad (\text{比特/符号}) \end{aligned}$$

又因为互信息的交互性

$$I(Y; X | Z) = I(X; Y | Z)$$

$$I(Z; X | Y) = H(Z | Y) - H(Z | XY) = H(Z | Y) = 0.5 \quad (\text{比特/符号})$$

$$I(Z; Y | X) = H(Z | X) - H(Z | XY) = H(Z | X) = 0.5 \quad (\text{比特/符号})$$

$$(3) I(XY; Z) = H(Z) - H(Z | XY) = H(Z) \approx 0.811 \quad (\text{比特/符号})$$

$$I(X; YZ) = H(X) - H(X | YZ) = 0.5 \quad (\text{比特/符号})$$

$$I(Y; XZ) = H(Y) - H(Y | XZ) = 0.5 \quad (\text{比特/符号})$$

3.8 有一个二元对称信道,其信道矩阵如图 3.29 所示。设该信道以 1500 个二元符号/秒的速度传输输入符号。现有一消息序列共有 14 000 个二元符号,并设在这消息中 $P(0) = P(1) = \frac{1}{2}$ 。问从信息传输的角度来考虑,10 秒内能否将这消息序列无失真地传送完。

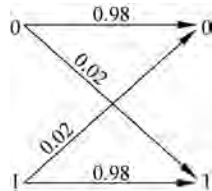


图 3.29 习题 3.8 图

解 消息是一个二元序列,这个二元符号是等概率分布,即 $P(0) = P(1) = \frac{1}{2}$,所以消息信源的熵 $H(X) = 1$ (比特/符号),即每个二元符号含有 1 比特信息量。那么这消息序列含有信息量 $= 14\,000 \text{ 符号} \times 1 \text{ (比特/符号)} = 1.4 \times 10^4 \text{ (bit)}$ 。

现计算这个二元对称信道的最大信息传输速率。此信道是二元对称信道,信道传递矩阵为

$$\mathbf{P} = \begin{bmatrix} 0.98 & 0.02 \\ 0.02 & 0.98 \end{bmatrix}$$

所以其信道容量(最大信息传输率)为

$$C = 1 - H(p) = 1 - H(0.98) \approx 0.8586 \quad (\text{比特/符号})$$

得最大信息传输速率为

$$R_t \approx 1500 \times 0.8586 \approx 1287.9 \approx 1.288 \times 10^3 \text{ (bps)}$$

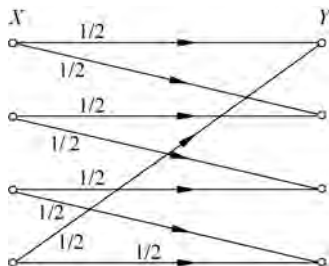


图 3.30 习题 3.9 图

此信道 10s 内能无失真传输的最大信息量为

$$10 \times R_t \approx 1.288 \times 10^4 \text{ (bit)}$$

可见,此信道 10s 内能无失真传输的最大信息量小于这消息序列所含有(携带)的信息量,所以从信道传输的角度来考虑,不可能在 10s 内将这消息无失真地传送完。

3.9 设一个有扰离散信道的传输情况如图 3.30 所示,试求出这种信道的信道容量。

解 信道传输矩阵如下

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

可以看出这是一个对称信道, $m=4$, 那么信道容量为

$$\begin{aligned} C &= \log m - H\left(\frac{1}{2}, \frac{1}{2}, 0, 0\right) \\ &= \log m + \sum_{j=1}^m p(y_j | x_i) \log p(y_j | x_i) \\ &= \log_2 4 + 2 \times \frac{1}{2} \log_2 \frac{1}{2} \\ &= \log_2 4 + \log_2 \frac{1}{2} \\ &= 1 \quad (\text{bit}) \end{aligned}$$

3.10 求图 3.31 中信道的信道容量及其最佳的输入概率分布。

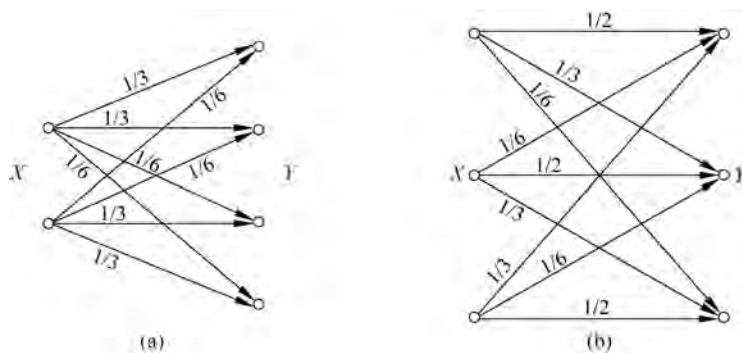


图 3.31 习题 3.10 图

解 (1) 信道传递矩阵为

$$\mathbf{P}_{Y|X} = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{3} \end{bmatrix}$$

$m=4$, 属于对称信道

$$\begin{aligned} C &= \log_2 4 + 2 \times \frac{1}{3} \log_2 \frac{1}{3} + 2 \times \frac{1}{6} \log_2 \frac{1}{6} \\ &= 0.0816 (\text{bit}) \end{aligned}$$

(2) 信道传递矩阵为

$$\mathbf{P}_{Y|X} = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$$

$m=3$, 属于对称信道

$$\begin{aligned} C &= \log m - H\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}\right) \\ &= \log_2 3 + \frac{1}{2} \log_2 \frac{1}{2} + \frac{1}{3} \log_2 \frac{1}{3} + \frac{1}{6} \log_2 \frac{1}{6} \\ &= 0.126(\text{bit}) \end{aligned}$$

3.11 若有一离散信道, 其信道转移概率如图 3.32 所示, 试求其信道容量。

解 由信道转移概率矩阵

$$\mathbf{P} = \begin{bmatrix} 1 - \epsilon_1 - \epsilon_2 & \epsilon_2 & \epsilon_1 \\ \epsilon_2 & 1 - \epsilon_1 - \epsilon_2 & \epsilon_1 \end{bmatrix}$$

可知此信道为准对称信道。因此, 当 $p_0 = p_1 = \frac{1}{2}$ 时达到信道容量
输出端分布为

$$\begin{aligned} q_0 &= \frac{1}{2}(1 - \epsilon_1 - \epsilon_2) + \frac{1}{2}\epsilon_2 = \frac{1}{2}(1 - \epsilon_1) \\ q_1 &= \frac{1}{2}\epsilon_2 + \frac{1}{2}(1 - \epsilon_1 - \epsilon_2) = \frac{1}{2}(1 - \epsilon_1) \\ q_\epsilon &= \frac{1}{2}\epsilon_1 \times 2 = \epsilon_1 \end{aligned}$$

所以

$$\begin{aligned} C &= \max I(X; Y) \\ &= \max [H(Y) - H(Y|X)] \\ &= [-q_0 \log q_0 - q_1 \log q_1 - q_\epsilon \log q_\epsilon] + \\ &\quad (1 - \epsilon_1 - \epsilon_2) \log(1 - \epsilon_1 - \epsilon_2) + \epsilon_1 \log \epsilon_1 + \epsilon_2 \log \epsilon_2 \\ &= -(1 - \epsilon_1) \log \frac{1}{2}(1 - \epsilon_1) + \epsilon_2 \log \epsilon_2 + (1 - \epsilon_1 - \epsilon_2) \log(1 - \epsilon_1 - \epsilon_2) \end{aligned}$$

3.12 一个离散信道, 其信道转移概率矩阵为

$$\mathbf{P} = \begin{bmatrix} 1 - p - \epsilon & p - \epsilon & 2\epsilon \\ p - \epsilon & 1 - p - \epsilon & 2\epsilon \end{bmatrix}$$

试求信道容量值 C 。

解 由信道转移概率矩阵

$$\mathbf{P} = \begin{bmatrix} 1 - p - \epsilon & p - \epsilon & 2\epsilon \\ p - \epsilon & 1 - p - \epsilon & 2\epsilon \end{bmatrix}$$

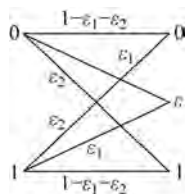


图 3.32 习题 3.11 图

可知此信道为准对称信道,当信道输入为等概率分布时达到信道容量。所以

$$\begin{aligned}
 C &= \max I(X; Y) = \max [H(Y) - H(Y | X)] \\
 &= \left[-2 \times \left(\frac{1}{2} - \epsilon \right) \log \left(\frac{1}{2} - \epsilon \right) - 2\epsilon \log 2\epsilon \right] + \\
 &\quad \frac{1}{2} \times 2[(1 - p - \epsilon) \log(1 - p - \epsilon) + (p - \epsilon) \log(p - \epsilon) + 2\epsilon \log 2\epsilon] \\
 &= (2\epsilon - 1) \log \left(\frac{1}{2} - \epsilon \right) + (1 - p - \epsilon) \log(1 - p - \epsilon) + (p - \epsilon) \log(p - \epsilon)
 \end{aligned}$$

3.13 一个离散信道,其信道转移概率分布如图 3.33 所示。

试求:

- (1) 达到信道容量 C 时输入概率分布 $\{P(x)\}$;
- (2) 信道容量值 C 。

解 (1) 信道转移概率矩阵为

$$\begin{bmatrix} 1 - \epsilon & \epsilon & 0 \\ 0 & 1 & 0 \\ 0 & \epsilon & 1 - \epsilon \end{bmatrix}$$

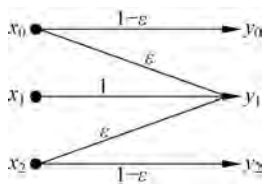


图 3.33 习题 3.13 图

则信道输出端分布为

$$\begin{aligned}
 q_0 &= p_0(1 - \epsilon) \\
 q_1 &= p_0\epsilon + p_1 + (1 - p_0 - p_1)\epsilon = p_1 + (1 - p_1)\epsilon \\
 q_2 &= (1 - \epsilon)(1 - p_0 - p_1)
 \end{aligned}$$

由信道容量定义知

$$\begin{aligned}
 C &= \max I(X; Y) = \max [H(Y) - H(Y | X)] \\
 &= \max [H(q_0, q_1, q_2) - p_0 H(\epsilon, 1 - \epsilon) - (1 - p_0 - p_1) H(\epsilon, 1 - \epsilon)] \\
 &= \max_{\{p(x)\}} [H(q_0, q_1, q_2) - (1 - p_1) H(\epsilon, 1 - \epsilon)]
 \end{aligned}$$

令 $\frac{\partial C}{\partial p_0} = 0$, 则有

$$-(1 - \epsilon) [\log p_0(1 - \epsilon) - \log(1 - p_0 - p_1)(1 - \epsilon)] = 0$$

所以

$$2p_0 = 1 - p_1$$

令 $\frac{\partial C}{\partial p_1} = 0$, 则有

$$-(1 - \epsilon) \{ \log [p_1 + (1 - p_1)\epsilon] - \log(1 - p_0 - p_1)(1 - \epsilon) \} + H(\epsilon) = 0$$

所以

$$\log \frac{p_1 + (1 - p_1)\epsilon}{(1 - p_0 - p_1)(1 - \epsilon)} = \frac{H(\epsilon, 1 - \epsilon)}{1 - \epsilon}$$

由联立方程组解得

$$\begin{cases} p_0 = \frac{1}{(1 - \epsilon) [2 + 2^{\frac{H(\epsilon)}{1 - \epsilon}}]} \\ p_1 = \frac{(1 - \epsilon) 2^{\frac{H(\epsilon)}{1 - \epsilon}} - 2\epsilon}{(1 - \epsilon) [2 + 2^{\frac{H(\epsilon)}{1 - \epsilon}}]} \\ p_2 = \frac{1}{(1 - \epsilon) [2 + 2^{\frac{H(\epsilon)}{1 - \epsilon}}]} \end{cases}$$

令

$$2 + 2^{\frac{H(\epsilon)}{1-\epsilon}} = A$$

则

$$\begin{cases} q_0 = \frac{1}{A} \\ q_1 = \frac{A-2}{A} \\ q_2 = \frac{1}{A} \end{cases}$$

(2) 信道容量为

$$\begin{aligned} C &= H(q_0, q_1, q_2) - (1-p_1)H(\epsilon, 1-\epsilon) \\ &= -\frac{2}{A} \log \frac{1}{A} - \frac{A-2}{A} \log \frac{A-2}{A} - \frac{2}{(1-\epsilon)A} H(\epsilon, 1-\epsilon) \end{aligned}$$

3.14 若已知信道输入分布为等概率分布,且有下列两个信道,其转移概率为

$$\mathbf{P}_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

试求这两个信道的信道容量,并判断这两个信道是否有噪声。

解 (1) 由信道 1 的转移概率矩阵可知其为对称信道,所以

$$C_1 = \log 4 - H\left(\frac{1}{2}, \frac{1}{2}\right) = 1(\text{bit})$$

因为

$$H(X) = \log_2 4 = 2 > C_1$$

所以有信息熵损失,信道有噪声。

(2) 由信道 2 的转移概率矩阵可知其为准对称信道,输入为等概分布时达到信道容量。

令此时的输出分布为

$$q_1 = q_2 = \cdots = q_8 = \frac{1}{8}$$

所以

$$\begin{aligned} C_2 &= H(Y) - H(Y | X) \\ &= \log 8 - H\left(\frac{1}{2}, \frac{1}{2}\right) \\ &= 2 \end{aligned}$$

因为

$$H(X) = \log 4 = C_2$$

所以信道 2 无噪声。

3.15 若有一个离散 Z 形信道,其信道转移概率如图 3.34 所示。试求:

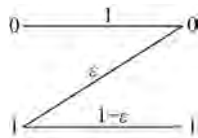


图 3.34 习题 3.15 图

- (1) 信道容量;
- (2) 若将两个同样 Z 形信道串接,求串接后的信道转移概率;
- (3) 求串接后的信道容量。

解 (1) 设信源输出 0、1 的先验概率分别为 p_0 、 p_1 , 且 $p_0 = 1 - p_1$, 则由该信道转移概率可知:

$$q_0 = p_0 + p_1 \cdot \epsilon = 1 - p_1(1 - \epsilon)$$

$$q_1 = p_1(1 - \epsilon)$$

$$\begin{aligned} I(X; Y) &= H(Y) - H(Y | X) = H(q_0, q_1) + \sum_i \sum_j p_i p_{ji} \log p_{ji} \\ &= -q_0 \log q_0 - (1 - q_0) \log(1 - q_0) - p_1 H(\epsilon) \end{aligned}$$

注:

$$H(\epsilon) = -[\epsilon \log \epsilon + (1 - \epsilon) \log(1 - \epsilon)]$$

由定义

$$C_1 = \max_{\{p(x)\}} I(X; Y)$$

令

$$\frac{dI(X; Y)}{dp_1} = 0$$

即

$$-\frac{dq_0}{dp_1} \log q_0 - \frac{dq_0}{dp_1} - \frac{dq_1}{dp_1} \log q_1 - \frac{dq_1}{dp_1} - H(\epsilon) = 0$$

$$\Rightarrow (1 - \epsilon) \log \frac{q_0}{q_1} = H(\epsilon)$$

$$\Rightarrow \frac{q_0}{q_1} = 2^{\frac{H(\epsilon)}{1 - \epsilon}}$$

令 $2^{\frac{H(\epsilon)}{1 - \epsilon}} = A$, $\left(\frac{H(\epsilon)}{1 - \epsilon} = \log A\right)$, 则有 $\frac{q_0}{1 - q_0} = A$ 。当 $q_0 = \frac{A}{A + 1}$ 时达到信道容量。所以

$$\begin{aligned} C_1 &= -\frac{A}{A + 1} \log \frac{A}{A + 1} - \frac{1}{A + 1} \log \frac{1}{A + 1} - \frac{1}{A + 1} \log A \\ &= -\left(\frac{A}{A + 1} + \frac{1}{A + 1}\right) \log \frac{A}{A + 1} = \log \frac{A}{A + 1} = \log \left[1 + 2^{\frac{H(\epsilon)}{1 - \epsilon}}\right] \\ &= \log [1 + (1 - \epsilon) \epsilon^{\frac{\epsilon}{1 - \epsilon}}] \end{aligned}$$

(2) 若将 Z 信道串接, 则信道转移概率矩阵为

$$\begin{bmatrix} 1 & 0 \\ \epsilon & 1 - \epsilon \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \epsilon & 1 - \epsilon \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2\epsilon - \epsilon^2 & (1 - \epsilon)^2 \end{bmatrix}$$

(3) 串接后的信道转移概率矩阵

$$\mathbf{p} = \begin{bmatrix} 1 & 0 \\ 2\epsilon - \epsilon^2 & (1 - \epsilon)^2 \end{bmatrix}$$

$$\text{令 } 2\epsilon - \epsilon^2 = \eta, \text{ 则 } \mathbf{p} = \begin{bmatrix} 1 & 0 \\ \eta & 1 - \eta \end{bmatrix}$$

由(1)的结论可知,信道容量为

$$\begin{aligned} C_2 &= \log [1 + 2^{\frac{H(\eta)}{1-\eta}}] = \log [1 + (1-\eta)\eta^{\frac{\eta}{1-\eta}}] \\ &= \log \left[1 + (1-\epsilon)^2 (2\epsilon - \epsilon^2) \frac{2\epsilon - \epsilon^2}{(1-\epsilon)^2} \right] \end{aligned}$$

3.16 若有一离散准对称信道,其信道转移概率如图 3.35 所示。

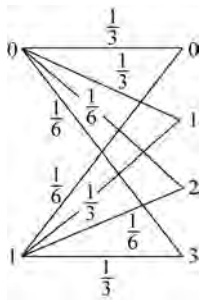


图 3.35 习题 3.16 图

试求:

- (1) 信道容量 C ;
- (2) 若将两个同样上述准对称信道串接,能否构成一个新的信道?

为什么?

解 (1) 信道转移概率矩阵为

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

当准对称信道输入端等概率分布时才达到信道容量值。所以

$$p_0 = p_1 = \frac{1}{2} \Rightarrow q_0 = \frac{1}{4}, q_1 = \frac{1}{3}, q_2 = \frac{1}{6}, q_3 = \frac{1}{4}$$

$$\begin{aligned} C &= \max_{p(x)} I(X; Y) = \max [H(Y) - H(Y|X)] = \left[-\frac{2}{4} \log \frac{1}{4} - \frac{1}{3} \log \frac{1}{3} - \frac{1}{6} \log \frac{1}{6} \right] + \\ &\quad \frac{1}{2} \times 2 \left[2 \times \frac{1}{3} \log \frac{1}{3} + 2 \times \frac{1}{6} \log \frac{1}{6} \right] = 0.041(\text{bit}) \end{aligned}$$

(2) 显然无法串接上述信道,原因是信道输出端符号数比信道输入端多。

3.17 设有一离散级联信道如图 3.36 所示。试求:

- (1) X 与 Y 间的信道容量;
- (2) Y 与 Z 间的信道容量;
- (3) X 与 Z 间的信道容量及其输入分布 $\{P(x)\}$ 。

解 (1) X 、 Y 间信道为强对称信道,所以

$$C_1 = 1 - H(\epsilon)$$

(2) Y 、 Z 间信道为准对称信道,信道容量为

$$C_2 = H\left(\frac{3}{8}, \frac{3}{8}, \frac{1}{4}\right) + 2 \cdot \frac{1}{2} \left(\frac{3}{4} \log \frac{3}{4} + \frac{1}{4} \log \frac{1}{4} \right) = 1.56 - 0.81 = 1.75(\text{bit})$$

(3) X 、 Z 间信道转移概率矩阵为

$$\begin{bmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{bmatrix} \cdot \begin{bmatrix} \frac{3}{4} & 0 & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{3}{4}(1-\epsilon) & \frac{3}{4}\epsilon & \frac{1}{4} \\ \frac{3}{4}\epsilon & \frac{3}{4}(1-\epsilon) & \frac{1}{4} \end{bmatrix}$$

它也属于准对称信道。当 $p_0 = p_1 = \frac{1}{2}$ 时达到信道容量。所以 $q_0 = \frac{3}{8}, q_1 = \frac{3}{8}, q_2 = \frac{1}{4}$, 信道容量为

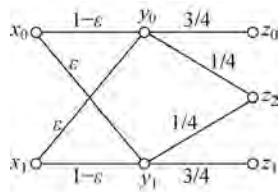


图 3.36 习题 3.17 图

$$\begin{aligned}
C_3 &= H\left(\frac{3}{8}, \frac{3}{8}, \frac{1}{4}\right) + \frac{3}{4}(1-\epsilon)\log\frac{3}{4}(1-\epsilon) + \frac{3}{4}\epsilon\log\frac{3}{4}\epsilon + \frac{1}{4}\log\frac{1}{4} \\
&= 1.06 + \frac{3}{4}(1-\epsilon)\log\frac{3}{4}(1-\epsilon) + \frac{3}{4}\epsilon\log\frac{3}{4}\epsilon \text{ (bit)}
\end{aligned}$$

3.18 若已知两信道的信道转移概率为

$$\mathbf{P}_1 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

试问：两信道串联的信道转移概率是多少？其信道容量是否发生变化？

解 两信道串联后的信道转移概率为

$$\mathbf{P} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

因为 $\mathbf{P}$ 不变, 所以信道容量不变。