

信道及信道容量

信道是传送信息的物理通道,是构成信息通信系统的重要组成部分,其任务是以信号形式传输和存储信息。例如,二人面对面聊天,语言靠声波通过空气来传送,因此二人间的空气就是信道。再如,手机等无线通话的信道是电波传播所通过的空间,而有线电话的信道就是连接电话机的双绞线电缆。在物理信道一定的情况下,人们总是希望传输的信息越多越好,例如在载波电话中,一个电话机作为发出信息的信源,接听的另一方就是接收信息的信宿,而它们之间的设施就是一条信道,而这条信道可以通过采用不同频率载波的方式被多组用户所共享。一条信道能传输多少信息,不仅与物理信道本身的特性有关,还与载荷信息的信号形式和信源输出信号的统计特性有关。本章主要讨论所谓的信道容量问题,它反映了信道所能传输的最大信息量,即在什么条件下通过信道的信息量最大。

3.1 信道的数学模型和分类

3.1.1 信道的分类

信息论不研究信号在信道中传输的物理过程,它假定信道的传输特性是已知的,这样信道就可以用如图 3.1 所示的抽象模型来描述,其符号表示为 $\{XP(X/Y)Y\}$ 。

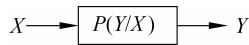


图 3.1 一般信道的数学模型

信号在信道中传输,不可避免地会受到噪声的干扰,从而产生错误和失真,所以信道的输入和输出之间一般不是确定的函数关系,而是统计依赖关系。只要知道输入输出信号之间的依赖关系,信道的特性也就确定了。根据实际应用的需求,信道可以有多种分类方法。

(1) 根据输入输出信号在时间和幅度上的取值是否连续来分类可分为离散信道,输入输出随机变量都取离散值;连续信道,输入输出随机变量都取连续值;半离散/半连续信道,输入变量取离散值而输出变量取连续值,或反之。

(2) 根据输入输出随机变量个数的多少分类可分为单符号信道,输入和输出端都只

用一个随机变量来表示；多符号信道，输入和输出端用随机变量序列/随机矢量来表示。

(3) 根据输入输出个数分类可分为单用户信道，只有一个输入和输出的信道；多用户信道，有多个输入和输出的信道。

(4) 根据信道上有无干扰分类可分为有干扰信道和无干扰信道。

(5) 根据信道有无记忆特性分类可分为有记忆信道和无记忆信道。

(6) 根据输入和输出之间有无反馈可分为有反馈信道和无反馈信道。

实际信道的带宽总是有限的，所以输入和输出信号总可以分解成随机序列来研究。一个实际信道可同时具有多种属性，其中最简单的信道是单符号离散信道。

3.1.2 单符号离散信道的数学模型

单符号离散信道是指信道的输入输出都取值于离散集合并且都用一个随机变量来表示的信道。设单符号离散信道的输入随机变量为 X ，其所有可能的取值为 $a_i, i=1, 2, 3, \dots, n$ ，输出随机变量为 Y ，其所有可能的取值为 $b_j, j=1, 2, 3, \dots, m$ ，则其信道模型如图 3.2 所示。

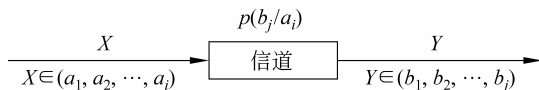


图 3.2 离散单符号信道模型

由于信道中存在干扰，因此输入的符号会在传输中产生错误，这种信道干扰对传输的影响可以用传递概率 $p(b_j/a_i)$ 来表示：

$$p(b_j/a_i) = p(Y = b_j/X = a_i), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m \quad (3-1)$$

这时信道传递概率实际上是一个传递概率矩阵，称为信道矩阵，记为：

$$\begin{matrix} & \begin{matrix} b_1 & b_2 & \cdots & b_m \end{matrix} \text{ (输出)} \\ \begin{matrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{matrix} \text{ (输入)} & \begin{bmatrix} p(b_1/a_1) & p(b_2/a_1) & \cdots & p(b_m/a_1) \\ p(b_1/a_2) & p(b_2/a_2) & \cdots & p(b_m/a_2) \\ \vdots & \vdots & \ddots & \vdots \\ p(b_1/a_n) & p(b_2/a_n) & \cdots & p(b_m/a_n) \end{bmatrix} \end{matrix} = \mathbf{P} \quad (3-2)$$

其中满足 $\sum_{j=1}^m p(b_j/a_i) = 1$ ，即信道矩阵中每个元素均为非负，每一行元素之和为 1。由式(3-2)所示的信道矩阵，可以得出输入和输出符号的联合概率为 $p(X = a_i, Y = b_j) = p(a_i b_j)$ ，则有

$$p(a_i b_j) = p(a_i) p(b_j/a_i) = p(b_j) p(a_i/b_j) \quad (3-3)$$

由全概公式可求得输出符号的概率为：

$$p(b_j) = \sum_{i=1}^n p(a_i) p(b_j/a_i) \quad (\text{对 } j = 1, 2, \dots, m \text{ 都成立}) \quad (3-4)$$

由贝叶斯公式可得后向概率公式为：

$$p(a_i/b_j) = \frac{p(a_i b_j)}{p(b_j)} \quad (p(b_j) \neq 0)$$

$$= \frac{p(a_i)p(b_j/a_i)}{\sum_{i=1}^n p(a_i)p(b_j/a_i)} \quad i = 1, 2, \dots, n \quad j = 1, 2, \dots, m \quad (3-5)$$

其中 $\sum_{i=1}^n p(a_i/b_j) = 1, j = 1, 2, \dots, m$ 。

最常见的信道是二元对称信道(Binary Symmetric Channel, BSC), 发送端发送 0、1 数据分别用 a_1 、 a_2 表示, 接收端收到 0、1 数据分别用 b_1 、 b_2 表示, 如图 3.3 所示。

其信道的传递概率为:

$$p(b_1/a_1) = p(0/0) = 1 - p = \bar{p}$$

$$p(b_2/a_2) = p(1/1) = 1 - p = \bar{p}$$

$$p(b_1/a_2) = p(0/1) = p$$

$$p(b_2/a_1) = p(1/0) = p$$

其中, p 表示单个符号传输发生错误的概率, $1-p$ 表示单个符号无错误传输的概率。其信道矩阵也可表示为:

$$\mathbf{P} = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \quad (3-6)$$

式(3-6)满足 $\sum_{j=1}^2 p(b_j/a_1) = \sum_{j=1}^2 p(b_j/a_2) = 1$ 。

另一种常见的二元信道是删除信道, 如图 3.4 所示。

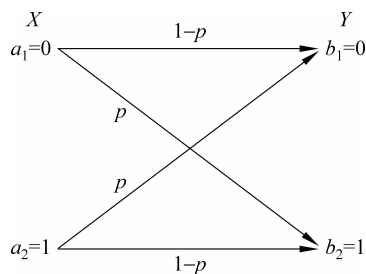


图 3.3 二元对称信道

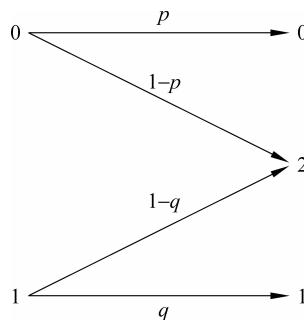


图 3.4 二元删除信道

其信道矩阵可表示为:

$$\mathbf{P} = \begin{bmatrix} p & 1-p & 0 \\ 0 & 1-q & q \end{bmatrix} \quad (3-7)$$

3.2 单符号离散信道中的熵速率与信道容量

3.2.1 离散无噪声信道的熵速率和信道容量

1. 离散无噪声信道的熵速率

离散无噪声信道的熵速率(信息速率)可以定义为信源在单位时间内输出的熵(或平均信息量)。离散信道的熵也就是离散信道每个消息(符号)的平均信息量, 它完全由信源

的概率空间来描述。即如果离散信道每秒钟输出的消息(符号)数为 t 个,则离散信源的熵速率可以表示为:

$$R = tH(X) = -t \sum_{i=1}^n p(a_i) \lg p(a_i) \quad (3-8)$$

2. 离散无噪声信道的信道容量

知道了离散无噪声信道的熵速率,那么离散无噪声信道的信道容量就可以定义为:信道对信源一切可能的概率分布(先验分布)而言能够传送的最大熵速率,即:

$$C = \max_{P(a_i)} R \quad (3-9)$$

这里需要指出,信道容量通常可以认为是一条信道内可以传输数据的最大值,换言之信道容量主要研究的是在什么条件下,通过信道的信息量最大。

假定某个离散信源具有 n 种状态,离散信道每秒钟输出的消息(符号)数为 t 个。则当信源等概时,输出最大熵为:

$$H_{\max}(X) = \lg n \quad (3-10)$$

此时,该无噪声离散信道的信道容量为:

$$C = t \lg n \text{ 比特/符号} \quad (3-11)$$

例如,某个离散信源 $n=2$,则根据式(3-11)可知其信道容量 $C=t$ 比特/符号,如果 $n=4$,这时的 $C=2t$ 比特/符号。

在实际的信源中,由于信源符号不一定等概,同时符号间很可能具有相关性,因此,实际信源的熵速率往往远远小于信道容量,从而使信道的实际传信率大大下降,但信道可能的最大熵速率(信道容量)却比实际值要大得多。

因此,信源输出的消息在被送入信道之前往往要通过编码技术将其转换为其他形式的码,原因之一就是希望将所要发送的消息变成能使信源的熵速率接近于信道容量的信号来传送。也就是使信源和信道相匹配,这就是所谓的信源最佳编码或匹配编码。

3.2.2 几种特殊离散信道的容量

下面来讨论几种常见的离散信道,并讨论离散信道的信道容量的计算问题。

1. 具有一一对应关系的离散无噪信道

假定信道输入和信道输出的符号集分别表示为

$$x \in \{a_1, a_2, \dots, a_n\} \quad (3-12)$$

$$y \in \{b_1, b_2, \dots, b_m\} \quad (3-13)$$

(1) 当输入 X 和输出 Y 符号集的元素个数相等,并且一一对应时,即有 $m=n$ 且 $a_i = b_i$,这时符号间的对应关系如图 3.5 所示。

其信道矩阵可以表示为:

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

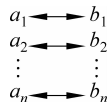


图 3.5 一一对应的无噪信道

(2) 当输入 X 和输出 Y 符号集的元素个数相等, 并且一一对应时, 并且满足关系式 $a_i = b_{m+1-i}$ 且 $m = n$, 这时符号间的对应关系如图 3.6 所示。

其信道矩阵可以表示为:

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

根据信道容量的定义, 这种 X, Y 一一对应的无噪信道, 有噪声熵 $H(Y/X) = 0$, 且收到 Y 后 X 也不存在不确定性, 信道疑义度 $H(X/Y) = 0$ 。此时输入端符号熵等于输出端符号熵, $H(X) = H(Y)$ 。当信源呈等概率分布时, 具有一一对应确定关系的无噪信道达到信道容量

$$C = \max_{p(a_i)} tH(X) = t \ln n \text{ (比特/符号)} \quad (3-14)$$

式(3-14)表明当信源呈等概分布时, 具有一一对应关系的无噪信道会达到信道容量其值就是信源 X 的最大熵值。同时还说明信道容量只决定于信道的输入符号数 n , 与信源无关, 是表征信道特性的一个参量。

2. 具有扩展功能的无噪信道

对于一个输入对应多个输出的信道, 一般称为具有扩展性能的无损信道, 如图 3.7 所示。

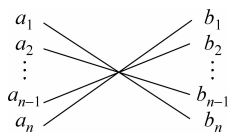


图 3.6 交叉对应的无噪信道

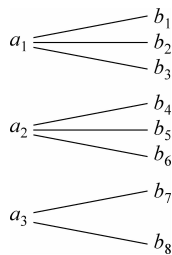


图 3.7 具有扩展性的无噪信道

其信道矩阵可以表示为:

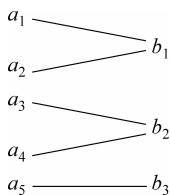
$$\begin{bmatrix} p(b_1/a_1) & p(b_2/a_1) & p(b_3/a_1) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p(b_4/a_2) & p(b_5/a_2) & p(b_6/a_2) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & p(b_7/a_3) & p(b_8/a_3) \end{bmatrix}$$

虽然信道矩阵中的元素不再是一一对应的关系, 但每列中只有一个非 0 元素, 也就是说在已知 Y 的条件下, X 就完全确定了, 即有信道疑义度 $H(X/Y) = 0$ 。例如, 输出端收到 b_2 后可以确定输入端发送的是 a_1 , 收到 b_7 后可以确定输入端发送的是 a_3 , 等等。与一一对应信道不同的是, 此时输入端符号熵小于输出端符号熵, $H(X) < H(Y)$ 。当信源呈等概率分布时, 具有扩展性能的无噪信道达到信道容量

$$C = \max_{p(a_i)} tI(X;Y) = \max_{p(a_i)} tH(X) = t \ln n \text{ (比特/符号)} \quad (3-15)$$

3. 具有归并性的信道

如果无噪信道是一个输出对应多个输入,则该类信道称为具有归并性的信道,如图 3.8 所示。



其信道矩阵可以表示为:

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

图 3.8 具有归并性的无噪信道

信道矩阵中的元素每行仅有一个非 0 元素,但是每列的非 0 元素个数大于 1,也就是说在已知 X 的条件下,对应的 Y 是完全确定的。但是收到某一个 b_j 后,对应的 a_i 不完全确定,信道疑义度 $H(X/Y) \neq 0$ 。此时输入端符号熵大于输出端符号熵,即有 $n > m$, $H(X) > H(Y)$ 。只有当 Y 集合等概率分布时,才能达到信道容量

$$C = \max_{p(a_i)} I(X;Y) = \max_{p(a_i)} I H(Y) = \text{tlbm(比特/符号)} \quad (3-16)$$

对于图 3.8 所示的例子,由信道矩阵有

$$p(b_1) = p(a_1) \times 1 + p(a_2) \times 1$$

$$p(b_2) = p(a_3) \times 1 + p(a_4) \times 1$$

$$p(b_3) = p(a_5)$$

$$p(b_1) = p(b_2) = p(b_3) = \frac{1}{3}$$

这时 $C = \text{lb}3 = 1.585$ (比特/符号)。此时必须满足 $p(b_1) = p(b_2) = p(b_3)$ 这一条件,可见信源的 $p(a_i)$ 的分布是不唯一的。由此,得到结论:无噪信道的信道容量 C 只决定于信道的输入符号数 n ,或输出符号数 m ,与信源无关。

例 3.1 有一无噪信道如图 3.9 所示,写出其信道矩阵。

解: 由图 3.9,容易得到该信道的信道矩阵为

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

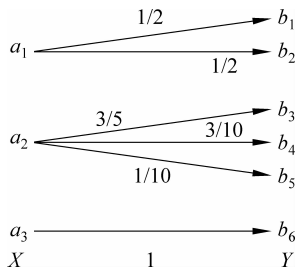


图 3.9 例 3.1 无噪信道

3.2.3 对称离散信道的信道容量

若信道转移矩阵中所有列都是第一列的一种排列,就称信道输出为对称的。若矩阵中每行都是同一集合 $p = \{p_1, p_2, \dots, p_n\}$ 中元素的不同排列,称行可排列。若矩阵中每列都是同一集合 $Q = \{q_1, q_2, \dots, q_n\}$ 中元素的不同排列,称列可排列。若行、列都可排列,则矩阵可排列,这时它所表示的信道称对称离散信道。

例如:

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

都是对称信道。

1. 强对称(均匀)离散信道及其信道容量

若对称信道中输入符号和输出符号个数相同,并且信道中总的错误概率为 p ,对称地平均分配给 $n-1$ 个输出符号,其中 n 表示输入输出符号的个数,则称此信道为强对称信道(也称均匀信道),其信道矩阵为

$$P = \begin{bmatrix} \bar{p} & \frac{p}{n-1} & \frac{p}{n-1} & \vdots & \frac{p}{n-1} \\ \frac{p}{n-1} & \bar{p} & \frac{p}{n-1} & \vdots & \frac{p}{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{p}{n-1} & \frac{p}{n-1} & \frac{p}{n-1} & \vdots & \bar{p} \end{bmatrix}, \quad p + \bar{p} = 1 \quad (3-17)$$

对于一般的信道,其信道矩阵中各行之和为 1,各列之和不一定等于 1,而均匀信道中各行、列之和均为 1 且元素相同。例如,二元对称信道就是 $n=2$ 的均匀信道。

例 3.2 求如图 3.10 所示二元强对称离散信道的信道容量。

$$\begin{aligned} \text{解: } H(X|Y) &= - \sum_{i=0}^1 \sum_{j=0}^1 p(a_i b_j) \lg p(a_i | b_j) \\ &= - p \lg p - (1-p) \lg (1-p) \\ H(X) &= - \sum_{i=0}^1 p(a_i) \lg p(a_i) \end{aligned}$$

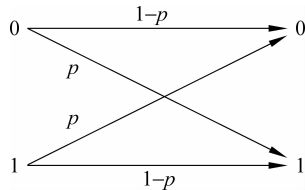


图 3.10 二元强对称离散信道

设信道每秒传输的消息数为 t ,则接收熵速率 $R(X;Y)$ 为:

$$R = t[H(X) - H(X|Y)]$$

该信道的信道容量为:

$$C = \max_{p(x)} \{R\}$$

可见只有当 $H(X)$ 最大, $H(X|Y)$ 最小的时候,该信道具有最大熵速率(信道容量)。即当 $p(a_1) = p(a_2) = 1/2$ 时,有

$$C = 1 + p \lg p + (1-p) \lg (1-p)$$

根据平均互信息的定义

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

$$H(Y/X) = \sum_{i=1}^n p(a_i) \sum_{j=1}^m p(b_j/a_i) \lg \frac{1}{p(b_j/a_i)} = \sum_{i=1}^n p(a_i) H(Y/X = a_i)$$

$$H(Y/X = a_i) = \sum_{j=1}^m p(b_j/a_i) \lg \frac{1}{p(b_j/a_i)}$$

由于信道的对称性, $H(Y/X=a_i)$ 与 a_i 无关, 且 $H(Y/X=a_i) = H(p'_1, p'_2, \dots, p'_m)$, 这里 $\{p'_1, p'_2, \dots, p'_m\}$ 为信道矩阵行矢量, 所以有

$$I(X;Y) = H(Y) - H(p'_1, p'_2, \dots, p'_m)$$

根据信道容量的定义

$$C = \max_{p(x)} [H(Y) - H(p'_1, p'_2, \dots, p'_m)]$$

当输出 Y 为等概分布时, $H(Y)$ 达到最大值 $\lg m$, 这时信道达到信道容量, 即

$$C = \lg m - H(p'_1, p'_2, \dots, p'_m) \quad (\text{比特/符号}) \quad (3-18)$$

另外, 需要指出的一点是对于对称信道, 当输入符号达到等概率分布时, 输出符号一定也达到等概率分布。

例 3.3 设某离散对称信道的信道矩阵为

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

求该信道的信道容量。

解: 由于该信道为对称信道, 因此其信道容量可用式 (3-18) 求得, 即将信道矩阵中的各参数代入式 $C = \lg m - H(p'_1, p'_2, \dots, p'_m)$ 中有

$$\begin{aligned} C &= \lg 4 - H\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right) = 2 + \left[\frac{1}{3} \lg \frac{1}{3} + \frac{1}{3} \lg \frac{1}{3} + \frac{1}{6} \lg \frac{1}{6} + \frac{1}{6} \lg \frac{1}{6}\right] \\ &= 0.0817 \quad (\text{比特/符号}) \end{aligned}$$

在这个对称信道中, 每个符号平均能够传输的最大信息量为 0.0817 比特, 而且只有在信道输入符号等概的条件下才能达到这个最大值。

例 3.4 离散均匀信道的信道矩阵为

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

求该信道的信道容量。

解: 均匀信道中输入输出符号数相等, 有 $n=m$ 成立, 所以

$$\begin{aligned} C &= \lg n - H\left(\bar{p}, \frac{p}{n-1}, \frac{p}{n-1}, \dots, \frac{p}{n-1}\right) \\ &= \lg n + \underbrace{\bar{p} \lg \bar{p} + \frac{p}{n-1} \lg \frac{p}{n-1} + \dots + \frac{p}{n-1} \lg \frac{p}{n-1}}_{\text{共}(n-1)\text{项}} \\ &= \lg n + \bar{p} \lg \bar{p} + p \lg \frac{p}{n-1} \\ &= \lg n - p \lg(n-1) - H(p) \end{aligned}$$

其中, p 是错误传输概率, $\bar{p}$ 是正确传输概率。因此, 得到结论: 均匀信道的信道容量为 $C = \lg n - p \lg(n-1) - H(p)$ 。

2. 准对称离散信道及其信道容量

若信道矩阵 $\mathbf{Q}$ 的列可以划分成若干个互不相交的子集 B_k , 即 $B_1 \cap B_2 \cap \cdots \cap B_K = \emptyset$, $B_1 \cup B_2 \cup \cdots \cup B_K = Y$, 由 B_k 为列组成的矩阵 $\mathbf{Q}_k$ 是对称矩阵, 则称信道矩阵 $\mathbf{Q}$ 所对应的信道为准对称信道。

也可以理解为: 若矩阵行可排列, 列不可排列, 但列的子集构成的子矩阵可排列, 则矩阵表示准对称信道。

例如, 信道

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \vdots & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{2} & \vdots & \frac{1}{8} & \frac{1}{8} \end{bmatrix}$$

具有行可排列性, 列则不具有可排列性, 但是如果把前两列和后两列分成互不相交的子集, 构成两个子矩阵:

$$\mathbf{P}_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix} \quad \text{和} \quad \mathbf{P}_2 = \begin{bmatrix} \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} \end{bmatrix}$$

两个子矩阵的行和列都是可排列的, 这时就称该信道为准排列信道。

准对称信道的信道容量为

$$C = \lg n - H(p'_1, p'_2, \dots, p'_m) - \sum_{k=1}^K N_k \lg M_k \quad (3-19)$$

其中, n 表示输入符号集的个数, $\{p'_1, p'_2, \dots, p'_m\}$ 表示准对称信道矩阵中的行元素。设矩阵可划分成 n 个互不相交的子集, N_k 为第 k 个子矩阵 $\mathbf{Q}_k$ 中行元素之和, M_k 为第 k 个子矩阵 $\mathbf{Q}_k$ 中列元素之和。即 $N_k = \sum_j P(b_j | a_i)$, $M_k = \sum_i P(b_j | a_i)$, $y \in Y_k$, $(k=1, 2, \dots, K)$ 。可以证明, 当输入为等概分布时, 可以达到信道容量。

例 3.5 二元删除信道的信道矩阵如下, 求该信道的信道容量。

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

解: 该信道是一个准对称信道, 有

$$\begin{aligned} N_1 &= 1-q, & N_2 &= q \\ M_1 &= 1-q, & M_2 &= 2q \end{aligned}$$

所以其信道容量为

$$\begin{aligned} C &= \lg n - H(p'_1, p'_2, \dots, p'_m) - \sum_{k=1}^K N_k \lg M_k \\ &= \lg 2 - H(1-p-q, q, p) - (1-q) \lg(1-q) - q \lg 2q \\ &= p \lg p + (1-p-q) \lg(1-p-q) + (1-q) \lg \frac{2}{1-q} \end{aligned}$$

当满足 $p+q=1$ 时, 上例的信道矩阵变为

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

这时的信道容量为

$$\begin{aligned} C &= p \lg p + (1-p-q) \lg(1-p-q) + (1-q) \lg \frac{2}{1-q} \\ &= 1-q \text{ (比特/符号)} \end{aligned}$$

例 3.6 已知信道矩阵

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \end{bmatrix}$$

求其信道容量。

解：将 $\mathbf{P}$ 分成可排列子矩阵

$$\mathbf{P}_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix} \quad \text{和} \quad \mathbf{P}_2 = \begin{bmatrix} \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} \end{bmatrix}$$

由于每个子集都只有两个元素即 $m_1 = m_2 = 2$, 所以 $p(a_1) = p(a_2) = \frac{1}{2}$

$$\bar{p}(b_1) = \frac{\sum_{i=1}^2 p(a_i) \sum_{p(b_j) \in M_1} p(b_j/a_i)}{m_1} = \frac{\sum_{i=1}^2 \sum_{j=1}^2 p(a_i) p(b_j/a_i)}{m_1} = \frac{1}{2} \left[\frac{1}{2} + \frac{1}{4} \right] = \frac{3}{8}$$

同理：

$$\bar{p}(b_2) = \frac{\sum_{i=1}^2 \sum_{j=3}^4 p(a_i) p(b_j/a_i)}{m_2} = \frac{1}{2} \left[\frac{1}{8} + \frac{1}{8} \right] = \frac{1}{8}$$

相应的,该准对称信道的信道容量为

$$\begin{aligned} C &= - \sum_{k=1}^2 M_k \bar{p}(b_k) \lg \bar{p}(b_k) - H(q_1, q_2, \dots, q_m) \\ &= - \left[2 \cdot \frac{3}{8} \lg \frac{3}{8} + 2 \cdot \frac{1}{8} \lg \frac{1}{8} \right] - \left[-\frac{1}{2} \lg \frac{1}{2} - \frac{1}{4} \lg \frac{1}{4} - 2 \cdot \frac{1}{8} \lg \frac{1}{8} \right] \\ &= 0.061\,278 \text{ (比特/符号)} \end{aligned}$$

3.2.4 离散信道容量的一般计算方法

研究信道的主要目的之一便是讨论信道中平均每个符号所能传送的信息量,即信道的信息传输率 R 。前文已经提到过,信道容量即是在固定信道的条件下,对所有可能的输入概率分布 $P(a_i)$ 求平均互信息的极大值。信道的信息传输率 R 可以表示为

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

考虑到信道传输信息时的时间因素,则有

$$R_t = \frac{1}{t} I(X; Y) = \frac{1}{t} H(X) - \frac{1}{t} H(X | Y) \text{ (比特/秒)}$$

R_i 称为信息传输速率,表征信道在单位时间内可以传输的信息量。 $I(X;Y)$ 是输入信源概率分布 $P(a_i)$ 的凸函数,对于一个固定的信道,总存在一种信源(某种概率分布 $P(a_i)$),使传输每个符号平均获得的信息量最大,即每个固定信道都有一个最大的信息传输率。定义这个最大的信息传输率为信道容量 C ,即 $C = \max_{P(a_i)} \{I(X;Y)\}$ (比特/符号),相应的输入概率分布称为最佳输入分布。信道容量 C 是信道传输概率函数,只与信道的统计特性有关。它是一个能够完全描述信道特性的参量,是信道能够传输的最大信息量。

由前面的讨论可知, $I(X;Y)$ 是 n 个变量 $\{P(a_1), P(a_2), \dots, P(a_n)\}$ 的多元函数,并满足 $\sum_{i=1}^n P(a_i) = 1$ 。所以可以用拉格朗日乘子法来求这个条件极大值,构造新函数

$$\Phi = I(X;Y) - \lambda \sum_i P(a_i) \quad (3-20)$$

其中, λ 为拉格朗日乘子(待定常数)。解方程组:

$$\frac{\partial \Phi}{\partial P(a_i)} = \frac{\partial [I(X;Y) - \lambda [\sum_{i=1}^n p(a_i) - 1]]}{\partial P(a_i)} = 0 \quad (3-21)$$

有 $I(X;Y)$ 的表达式

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

并且有

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

将以上两式代入式(3-21)得

$$\frac{\partial}{\partial P(a_i)} \lg P(b_j) = \left[\frac{\partial}{\partial P(a_i)} \ln P(b_j) \right] \lg e = \frac{P(b_j | a_i)}{P(b_j)} \lg e$$

求偏导

$$\sum_{j=1}^m P(b_j | a_i) \lg \frac{P(b_j | a_i)}{P(b_j)} - \sum_{i=1}^n \sum_{j=1}^m P(a_i) P(b_j | a_i) \lg \frac{P(b_j | a_i)}{P(b_j)} \lg e - \lambda = 0$$

上式对 $i=1, 2, \dots, n$ 均成立,同时因为 $\sum_{i=1}^n P(a_i) P(b_j | a_i) = P(b_j)$ 且有 $\sum_{j=1}^m P(b_j | a_i) = 1$, 整理得到

$$\sum_{j=1}^m P(b_j | a_i) \lg \frac{P(b_j | a_i)}{P(b_j)} = \lambda + \lg e \quad (3-22)$$

上式左边即为平均互信息量的极大值 C ,因此可以得到信道容量的表达式为:

$$C = \lambda + \lg e \quad (3-23)$$

令 $I(a_i;Y) = \sum_{j=1}^m P(b_j | a_i) \lg \frac{P(b_j | a_i)}{P(b_j)}$, 该式表示输出端接收到 Y 后获得关于 $X=a_i$ 的信息量,即信源符号 $X=a_i$ 对输出端 Y 平均提供的互信息。要使 $I(a_i;Y) = C (i=1, 2, \dots, n)$, 可得到如下定理:

定理 3.1 一般离散信道的平均互信息 $I(X;Y)$ 达到极大值(即等于信道容量)的充要条件是输入概率分布 $\{p(a_i)\}$ 满足

$$\begin{cases} I(x_i; Y) = C, & \text{对所有 } x_i \text{ 其 } p_i \neq 0 \\ I(x_i; Y) \leq C, & \text{对所有 } x_i \text{ 其 } p_i = 0 \end{cases} \quad (3-24)$$

或

$$\begin{cases} \frac{\partial I(X; Y)}{\partial p(a_i)} = \lambda, & \text{对所有 } a_i \text{ 其 } p(a_i) \neq 0 \\ \frac{\partial I(X; Y)}{\partial p(a_i)} \leq \lambda, & \text{对所有 } a_i \text{ 其 } p(a_i) = 0 \end{cases} \quad (3-25)$$

证明: (为表达方便, 以下过程中 $I(X; Y)$ 简写成 $I(P)$)

首先来证明其充分性。假设有一输入概率分布 $\{p(a_i)\} = P$ 满足式 (3-24) 和式 (3-25), 证明分布 $\{p(a_i)\}$ 一定使平均互信息 $I(P)$ 达到极大值, 即证明对于任何其他输入概率分布 $Q = \{q(a_i)\}$, 有 $I(Q) \leq I(P)$ 。

设 $0 < \theta < 1$, 则有

$$\theta I(Q) - (1 - \theta) I(P) \leq I[\theta Q + (1 - \theta) P]$$

整理可得

$$I(Q) - I(P) \leq \frac{\{I[\theta Q + (1 - \theta) P] - I(P)\}}{\theta}$$

当 $\theta \rightarrow 0$ 时, 有 $I(Q) - I(P) \leq \sum_{i=1}^n (q_i - p_i) \frac{\partial}{\partial p_i} I(P)$ 成立。

因为 P 满足式 (3-25), 即当 $p_i \neq 0$ 时, $\frac{\partial I(P)}{\partial p_i} = \lambda$; 而当 $p_i = 0, (q_i - p_i) = q_i \geq 0$ 时, $\frac{\partial I(P)}{\partial p_i} \leq \lambda$ 。所以得: $I(Q) - I(P) \leq \lambda \sum_{i=1}^r (q_i - p_i) = \lambda \left[\sum_{i=1}^r q_i - \sum_{i=1}^r p_i \right] = 0 \Rightarrow I(Q) \leq I(P)$ 。

下面再来证明必要性。假设有一输入概率分布 P 使 $I(X; Y)$ 达到极大值 $I(P)$, 证明概率分布 P 满足条件式 (3-24) 和式 (3-25)。

设存在任意其他概率分布 $Q = \{q_i\}$, 满足: $I[\theta Q + (1 - \theta) P] - I(P) \leq 0, 0 < \theta < 1$ 。将该式除以 θ , 并取 $\theta \rightarrow 0$ 的极限, 可以得到: $\sum_{i=1}^r (q_i - p_i) \frac{\partial}{\partial p_i} I(P) \leq 0$ 。对于 $\{p_i\}$, 因为

$\sum_{i=1}^r p_i = 1$, 所以至少有一分量 $p_l \neq 0$, 令 $p_l \neq 0$ 。选择另一种概率分布 $Q = \{q_i\}$ 满足

$$\begin{cases} q_l = p_l - \epsilon \\ q_i = p_i, & (i \neq l, j), \text{ 其中 } \epsilon \text{ 为任意数, 满足 } -p_j \leq \epsilon \leq p_l \\ q_j = p_j + \epsilon \end{cases}$$

就变为 $-\epsilon \frac{\partial I}{\partial p_l} + \epsilon \frac{\partial I}{\partial p_j} \leq 0$ 。如令 $\frac{\partial I}{\partial p_l} = \lambda$, 会得到 $\epsilon \frac{\partial I}{\partial p_j} \leq \lambda \epsilon$ 。

因为 $p_l > 0$, 所以当 $p_j = 0, \epsilon$ 总取正数, 得 $\frac{\partial I}{\partial p_j} \leq \lambda$ 。

若 $p_j \neq 0, \epsilon$ 可取正数, 也可取负数。若取正数, 得 $\frac{\partial I}{\partial p_j} \leq \lambda$, 若取负数, 得 $\frac{\partial I}{\partial p_j} \geq \lambda$ 。故当 $p_j \neq 0, \frac{\partial I}{\partial p_j} = \lambda$, 满足定理条件。由此可以得到以下结论: 当信道平均互信息达到信道容量时, 输入信源符号集中每一个信源符号 x 对输出端 Y 提供相同的互信息, 只是概率为

零的符号除外。

对于一般的离散信道,很难利用上述定理来寻求信道容量和对应的输入概率分布。因此仍只能采用求解方程组的方法。即

$$\begin{cases} \sum_{j=1}^m P(b_j | a_i) \lg \frac{P(b_j | a_i)}{P(b_j)} = \lambda + \lg e \\ \sum_{i=1}^n P(a_i) = 1 \end{cases}$$

$$\sum_{j=1}^m P(b_j | a_i) \lg P(b_j | a_i) - \sum_{j=1}^m P(b_j | a_i) \lg P(b_j) = C, \quad (i = 1, 2, \dots, n)$$

$$\sum_{j=1}^m P(b_j | a_i) [C + \lg P(b_j)] = \sum_{j=1}^m P(b_j | a_i) \lg P(b_j | a_i), \quad (i = 1, 2, \dots, n)$$

令 $\beta_j = C + \lg P(b_j)$, 得到:

$$\sum_{j=1}^m P(b_j | a_i) \beta_j = \sum_{j=1}^m P(b_j | a_i) \lg P(b_j | a_i) \quad (3-26)$$

设 $n=m$, 信道传递矩阵 $\mathbf{P}$ 是非奇异矩阵, 则此方程组有解。这时就可以求出 β_j 的数值,

然后再根据 $P(b_j) = 2^{\beta_j - C}$, $(j=1, 2, \dots, m)$ 和 $\sum_{i=1}^m P(b_j) = 1$ 求得信道容量:

$$\sum_j 2^{\beta_j - C} = 1$$

所以有

$$C = \lg \left(\sum_{j=1}^m 2^{\beta_j} \right) \quad (3-27)$$

对一般离散信道而言,求信道容量,就是在固定信道的条件下,对所有可能的输入概率分布,求其平均互信息量的极大值。

下面来总结一下计算信道容量 C 的步骤:

(1) 利用公式 $\sum_{j=1}^m p(b_j/a_i) \lg p(b_j/a_i) = \sum_{j=1}^m p(b_j/a_i) \beta_j$, 求 β_j ;

(2) 利用公式 $C = \lg \left(\sum_{j=1}^m 2^{\beta_j} \right)$, 求 C ;

(3) 利用公式 $P(b_j) = 2^{\beta_j - C}$, 求 $P(b_j)$;

(4) 利用公式 $p(b_j) = \sum_{i=1}^n p(a_i) p(b_j/a_i)$, 求 $p(a_i)$ 。

例 3.7 设某离散信道如图 3.11 所示,输入符号集为 $[a_1, a_2, a_3]$, 输出符号集为 $[b_1, b_2, b_3]$, 求其信道容量。

解: 该信道转移矩阵为

$$\mathbf{P} = \begin{bmatrix} 0.7 & 0.3 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0.3 & 0.7 \end{bmatrix}$$

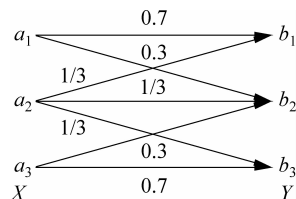


图 3.11 例 3.7 中的离散信道

因此有

$$I(a_1; Y) = \sum_{j=1}^3 p(b_j/a_1) \lg \frac{p(b_j/a_1)}{p(b_j)} = 0.7 \lg \frac{0.7}{0.35} + 0.3 \lg \frac{0.3}{0.3} = 0.7$$

$$I(a_3; Y) = \sum_{j=1}^3 p(b_j/a_3) \lg \frac{p(b_j/a_3)}{p(b_j)} = 0.7 \lg \frac{0.7}{0.35} + 0.3 \lg \frac{0.3}{0.3} = 0.7$$

$$I(a_2; Y) = \sum_{j=1}^3 p(b_j/a_2) \lg \frac{p(b_j/a_2)}{p(b_j)} = 0$$

可以得到

$$\begin{cases} I(a_i; Y) = 0.7 & p(a_i) \neq 0 \text{ 的所有 } a_i \\ I(a_i; Y) < 0.7 & p(a_i) = 0 \text{ 的所有 } a_i \end{cases}$$

故

$$C = 0.7 \text{ (比特/符号)}$$

例 3.8 设某离散信道如图 3.12 所示, 求其信道容量。

解: 该信道转移矩阵为

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

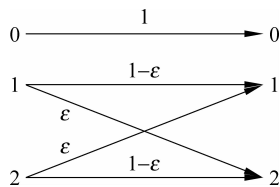


图 3.12 例 3.8 中的离散信道

根据式(3-25)

$$\sum_{j=1}^m p(b_j/a_i) \beta_j = \sum_{j=1}^m p(b_j/a_i) \lg p(b_j/a_i), \quad i = 1, 2, \dots, n$$

可得方程组

$$\begin{cases} \beta_1 = 0 \\ (1-\epsilon)\beta_2 + \epsilon\beta_3 = (1-\epsilon)\lg(1-\epsilon) + \epsilon\lg\epsilon \\ \epsilon\beta_2 + (1-\epsilon)\beta_3 = \epsilon\lg\epsilon + (1-\epsilon)\lg(1-\epsilon) \end{cases}$$

解方程组有

$$\beta_1 = 0$$

$$\beta_2 = \beta_3 = \epsilon \lg \epsilon + (1-\epsilon) \lg (1-\epsilon)$$

因此, 该信道的信道容量为

$$\begin{aligned} C &= \lg \sum_{j=1}^3 2^{\beta_j} = \lg [2^0 + 2 \times 2^{\epsilon \lg \epsilon + (1-\epsilon) \lg (1-\epsilon)}] \\ &= \lg [1 + 2^{1-H(\epsilon)}] \\ &= \lg [1 + 2(1-\epsilon)^{1-\epsilon} \epsilon^\epsilon] \text{ (比特/符号)} \end{aligned}$$

再利用 $P(b_j) = 2^{\beta_j - C}$ 式, 可得

$$\begin{aligned} p(b_1) &= 2^{\beta_1 - C} = 2^{-C} = \frac{1}{1 + 2(1-\epsilon)^{1-\epsilon} \epsilon^\epsilon} \\ p(b_2) &= 2^{\beta_2 - C} = \frac{(1-\epsilon)^\epsilon \epsilon^\epsilon}{1 + 2(1-\epsilon)^{1-\epsilon} \epsilon^\epsilon} \end{aligned}$$

$$p(b_3) = 2^{\beta_3 - C} = \frac{(1 - \epsilon)^\epsilon \epsilon^\epsilon}{1 + 2(1 - \epsilon)^{1 - \epsilon} \epsilon^\epsilon} = p(b_2)$$

最后利用 $p(b_j) = \sum_{i=1}^3 p(a_i) p(b_j/a_i)$, $j = 1, 2, 3$ 式, 可求得 $p(a_i)$

$$p(a_1) = \frac{1}{1 + 2(1 - \epsilon)^{1 - \epsilon} \epsilon^\epsilon}$$

$$p(a_2) = p(a_3) = \frac{(1 - \epsilon)^\epsilon \epsilon^\epsilon}{1 + 2(1 - \epsilon)^{1 - \epsilon} \epsilon^\epsilon}$$

当 $\epsilon = \frac{1}{2}$ 时, 则有

$$C = \text{lb}2(\text{比特/符号})$$

$$p(a_1) = \frac{1}{2}$$

$$p(a_2) = p(a_3) = \frac{1}{4}$$

从以上讨论可知, 信道容量 C 不仅与信源有关, 还与信道特性有关, 是信道转移概率的函数, 不同的信道就有不同的信道容量。它反映了信道本身的传信能力。离散信道容量的计算, 就是在各种情况下, 计算平均互信息量的最大值。

3.3 多符号离散信道及其信道容量

上一节讨论了单个随机变量通过信道传递的情况, 而实际离散信道的输入和输出常常是随机变量序列。如果在不同时刻有多个来自同一信源的随机变量通过离散信道传输, 则称这种信道为多符号离散信道。

多符号离散信道相当于单符号离散信道在 N 个不同时刻连续运用了 N 次, 所以也称为单符号离散信道的 N 次扩展信道, 其数学模型如图 3.13 所示。

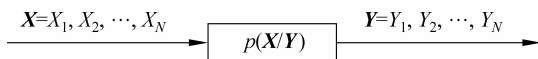


图 3.13 多符号离散信道模型

实际信道往往是记忆信道, 为了简单起见, 下面主要讨论离散无记忆信道。无记忆信道主要表现出两个性质:

(1) 无记忆性, 即 k 时刻的输出只与 k 时刻的输入有关, 与之前的输入输出无关。

(2) 无预感性, 即 k 时刻之前的输出随机变量序列只与 k 时刻之前的输入随机变量序列有关, 与以后的第 k 时刻的输入随机变量无关。

设离散无记忆信道的输入符号集 $A = \{a_1, a_2, \dots, a_n\}$, 输出符号集 $B = \{b_1, b_2, \dots, b_m\}$, 其信道矩阵为:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1m} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2m} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \cdots & p_{nm} \end{bmatrix} \quad \text{且} \quad \sum_{j=1}^m p_{ij} = 1 \quad (i = 1, 2, \dots, n)$$

该信道矩阵是一个 $n^N \times m^N$ 的矩阵,其输入随机序列 X 可能的取值有 n^N 个,输出随机序列 Y 的可能取值有 m^N 个。再根据信道无记忆的特性,可知无记忆信道 N 次扩展后的平均互信息为:

$$\begin{aligned} I(X;Y) &= I(X^N;Y^N) = H(X^N) - H(X^N | Y^N) = H(Y^N) - H(Y^N | X^N) \\ &= \sum_{X^N, Y^N} P(\alpha_k \beta_h) \lg \frac{P(\alpha_k | \beta_h)}{P(\alpha_k)} \\ &= \sum_{X^N, Y^N} P(\alpha_k \beta_h) \lg \frac{P(\beta_h | \alpha_k)}{P(\beta_h)} \end{aligned} \quad (3-28)$$

定理 3.2 若信道的输入随机序列为 $X=(X_1 X_2 \cdots X_N)$,通过信道传输,接收到的随机序列为 $Y=(Y_1 Y_2 \cdots Y_N)$ 。假设信道无记忆,即信道传递概率满足 $P(\beta_h | \alpha_k) = \prod_{i=1}^N P(b_{h_i} | a_{k_i})$, $(k=1,2,\cdots,n^N \quad h=1,2,\cdots,m^N)$, 则存在

$$I(X;Y) \leq \sum_{i=1}^N I(X_i;Y_i) \quad (3-29)$$

证明: 设信道输入和输出序列 X 和 Y 的一个取值为:

$$\alpha_k = (a_{k_1} a_{k_2} \cdots a_{k_N}), \quad \text{其中 } a_{k_i} \in \{a_1, a_2, \cdots, a_n\}, (i=1,2,\cdots,N)$$

$$\beta_h = (b_{h_1} b_{h_2} \cdots b_{h_N}), \quad \text{其中 } b_{h_i} \in \{b_1, b_2, \cdots, b_m\}, (i=1,2,\cdots,N)$$

根据式(3-27)给出的无记忆信道 N 次扩展后的平均互信息计算公式可得

$$\begin{aligned} I(X;Y) &= \sum_{X,Y} P(\alpha_k \beta_h) \lg \frac{P(\beta_h | \alpha_k)}{P(\beta_h)} = E \left[\lg \frac{P(\beta_h | \alpha_k)}{P(\beta_h)} \right] \\ &= E \left[\lg \frac{P(b_{h_1} | a_{k_1}) P(b_{h_2} | a_{k_2}) \cdots P(b_{h_N} | a_{k_N})}{P(\beta_h)} \right] \end{aligned}$$

输入输出变量的平均互信息之和为

$$\begin{aligned} &\sum_{i=1}^N I(X_i;Y_i) \\ &= \sum_{i=1}^N \sum_{X_i, Y_i} P(a_{k_i} b_{h_i}) \lg \frac{P(b_{h_i} | a_{k_i})}{P(b_{h_i})} \\ &= \sum_{X_1, Y_1} P(a_{k_1} b_{h_1}) \lg \frac{P(b_{h_1} | a_{k_1})}{P(b_{h_1})} + \sum_{X_2, Y_2} P(a_{k_2} b_{h_2}) \lg \frac{P(b_{h_2} | a_{k_2})}{P(b_{h_2})} + \cdots \\ &\quad + \sum_{X_N, Y_N} P(a_{k_N} b_{h_N}) \lg \frac{P(b_{h_N} | a_{k_N})}{P(b_{h_N})} \\ &= \sum_{X_1, Y_1} \cdots \sum_{X_N, Y_N} P(a_{k_1} \cdots a_{k_N} b_{h_1} \cdots b_{h_N}) \lg \frac{P(b_{h_1} | a_{k_1}) \cdot P(b_{h_2} | a_{k_2}) \cdots P(b_{h_N} | a_{k_N})}{P(b_{h_1}) \cdot P(b_{h_2}) \cdots P(b_{h_N})} \\ &= E \left[\lg \frac{P(b_{h_1} | a_{k_1}) \cdot P(b_{h_2} | a_{k_2}) \cdots P(b_{h_N} | a_{k_N})}{P(b_{h_1}) \cdot P(b_{h_2}) \cdots P(b_{h_N})} \right] \end{aligned}$$

所以有

$$\begin{aligned}
I(X;Y) &= \sum_{i=1}^N I(X_i;Y_i) \\
&= E \left[\lg \frac{P(b_{h_1} | a_{k_1}) \cdot P(b_{h_2} | a_{k_2}) \cdots P(b_{h_N} | a_{k_N})}{P(\beta_h)} \right. \\
&\quad \left. - \lg \frac{P(b_{h_1} | a_{k_1}) \cdot P(b_{h_2} | a_{k_2}) \cdots P(b_{h_N} | a_{k_N})}{P(b_{h_1}) \cdot P(b_{h_2}) \cdots P(b_{h_N})} \right] \\
&= E \left[\lg \frac{P(b_{h_1}) \cdot P(b_{h_2}) \cdots P(b_{h_N})}{P(\beta_h)} \right] \\
&\leq E \left[\lg \frac{P(b_{h_1}) \cdot P(b_{h_2}) \cdots P(b_{h_N})}{P(\beta_h)} \right] \\
&= \lg \sum_{X,Y} P(\alpha_k \beta_h) \frac{P(b_{h_1}) \cdot P(b_{h_2}) \cdots P(b_{h_N})}{P(\beta_h)} \\
&= \lg \sum_{X,Y} P(\alpha_k | \beta_h) P(b_{h_1}) \cdot P(b_{h_2}) \cdots P(b_{h_N}) \\
&= \lg \sum_Y P(b_{h_1}) \cdot P(b_{h_2}) \cdots P(b_{h_N}) = \lg 1 = 0
\end{aligned}$$

因此可得: $I(X;Y) \leq \sum_{i=1}^N I(X_i;Y_i)$, 证毕。

式(3-29)表明,对于无记忆离散信道,其平均互信息小于输入和输出序列中所有对应时刻随机变量的平均互信息之和,只有当信源满足无记忆时,等号成立。下面来证明等号成立的情况,即若信道的输入随机序列为 $X=(X_1 X_2 \cdots X_N)$,通过信道传输,接收到的随机序列为 $Y=(Y_1 Y_2 \cdots Y_N)$,而信道的传递概率为 $P(Y_j | X_i)$,假若信源是无记忆的,则

存在 $I(X;Y) = \sum_{i=1}^N I(X_i;Y_i)$ 。

证明: 平均互信息为

$$I(X;Y) = \sum_{X,Y} P(\alpha_k \beta_h) \lg \frac{P(\alpha_k | \beta_h)}{P(\alpha_k)} = E \left[\lg \frac{P(\alpha_k | \beta_h)}{P(\alpha_k)} \right]$$

其中, $\alpha_k = (a_{k_1} a_{k_2} \cdots a_{k_N})$, $a_{k_i} \in \{a_1, a_2, \cdots, a_r\}$, $(i=1, 2, \cdots, N)$; $\beta_h = (b_{h_1} b_{h_2} \cdots b_{h_N})$, $b_{h_i} \in \{b_1, b_2, \cdots, b_s\}$, $(i=1, 2, \cdots, N)$ 。同时,由于 $P(\alpha_k) = P(a_{k_1}) P(a_{k_2}) \cdots P(a_{k_N})$, 所以有

$$I(X;Y) = E \left[\lg \frac{P(\alpha_k | \beta_h)}{P(a_{k_1}) P(a_{k_2}) \cdots P(a_{k_N})} \right]$$

则

$$\begin{aligned}
&\sum_{i=1}^N I(X_i;Y_i) \\
&= \sum_{i=1}^N \sum_{X_i, Y_i} P(a_{k_i} b_{h_i}) \lg \frac{P(a_{k_i} | b_{h_i})}{P(a_{k_i})} \\
&= \sum_{X_1, Y_1} \cdots \sum_{X_N, Y_N} P(a_{k_1} \cdots a_{k_N} b_{h_1} \cdots b_{h_N}) \lg \frac{P(a_{k_1} | b_{h_1}) \cdot P(a_{k_2} | b_{h_2}) \cdots P(a_{k_N} | b_{h_N})}{P(a_{k_1}) \cdot P(a_{k_2}) \cdots P(a_{k_N})} \\
&= E \left[\lg \frac{P(a_{k_1} | b_{h_1}) \cdot P(a_{k_2} | b_{h_2}) \cdots P(a_{k_N} | b_{h_N})}{P(a_{k_1}) \cdot P(a_{k_2}) \cdots P(a_{k_N})} \right]
\end{aligned}$$

两式相减得

$$\begin{aligned}
 \sum_{i=1}^N I(X_i; Y_i) - I(X; Y) &= E \left[\lg \frac{P(a_{k_1} | b_{h_1}) \cdot P(a_{k_2} | b_{h_2}) \cdots P(a_{k_N} | b_{h_N})}{P(\alpha_k | \beta_h)} \right] \\
 &\leq \lg E \left[\frac{P(a_{k_1} | b_{h_1}) \cdot P(a_{k_2} | b_{h_2}) \cdots P(a_{k_N} | b_{h_N})}{P(\alpha_k | \beta_h)} \right] \\
 &= \lg \sum_{X,Y} P(\alpha_k \beta_h) \frac{P(a_{k_1} | b_{h_1}) \cdot P(a_{k_2} | b_{h_2}) \cdots P(a_{k_N} | b_{h_N})}{P(\alpha_k | \beta_h)} \\
 &= \lg \sum_{X,Y} P(\beta_h) P(a_{k_1} | b_{h_1}) \cdot P(a_{k_2} | b_{h_2}) \cdots P(a_{k_N} | b_{h_N}) \\
 &= \lg \sum_Y P(\beta_h) = \lg 1 = 0
 \end{aligned}$$

这时有 $I(X; Y) = \sum_{i=1}^N I(X_i; Y_i)$, 即当信源和信道都是无记忆时, 式(3-29)中的等号成立。

同时, 对于离散无记忆 N 次扩展信道, 若信道输入序列 $X = (X_1 X_2 \cdots X_N)$ 中 X_i 取自同一信源符号集, 并具有同一概率分布, 且通过相同的信道传送到输出端 (即输出序列 $Y = (Y_1 Y_2 \cdots Y_N)$ 中 Y_i 也取自于同一符号集), 则满足:

$$I(X_1, Y_1) = I(X_2, Y_2) = \cdots = I(X_N, Y_N) = I(X; Y) \Rightarrow \sum_{i=1}^N I(X_i; Y_i) = NI(X; Y) \quad (3-30)$$

式(3-30)表明, 对于离散无记忆信道的 N 次扩展信道, 若信源也是无记忆的, 则其平均互信息等于单符号信道的平均互信息的 N 倍。

前面已经证明过, 对于一般的离散无记忆信道的 N 次扩展信道满足

$$I(X; Y) \leq \sum_{i=1}^N I(X_i; Y_i)$$

则其信道容量:

$$C^N = \max I(X; Y) = \max \sum_{i=1}^N I(X_i; Y_i) = \sum_{i=1}^N \max I(X_i; Y_i) = \sum_{i=1}^N C_i \quad (3-31)$$

因为输入随机序列在同一个信道中传输, 所以其通过离散无记忆信道传输的最大信息量都相同, 即有 $C_i = C, i = 1, 2, \cdots, N$, 所以有

$$C^N = NC \quad (3-32)$$

这说明离散无记忆的 N 次扩展信源的信道容量等于原来原单符号离散信道的信道容量的 N 倍。信道容量只有当信源也是无记忆离散信源且每时刻的输入分布达到最佳输入分布时, 才能达到该信道容量。一般情况下, 消息序列在离散无记忆的 N 次扩展信道中传输的平均互信息量 $I(X; Y) \leq NC$ 。

3.4 组合信道的信道容量

前面分析了单符号离散信道和离散无记忆信道, 实际中常常会遇到多个信道组合在一起的情况, 也就是令信道和输入和输出随机变量序列中的各随机变量分别取值于不同

的符号集合,就构成了独立并联信道。有时由于某种原因,消息会依次通过一个不同的信道,这就构成一种级联信道,成为串行信道。本节将讨论组合信道及其信道容量问题。

3.4.1 串联信道及其信道容量

串联信道作为最基本的信道组合方式,应用非常广泛,如无线电中继信道等。一般的串联信道如图 3.14 所示,设信道 I 的输入随机变量为 X ,其所有可能的取值为 $a_i, i=1, 2, 3, \dots, n$,输出随机变量为 Y ,其所有可能的取值为 $b_j, j=1, 2, 3, \dots, m$ 。信道 II 的输入随机变量为信道 I 的输出随机变量 Y ,输出随机变量为 Z ,其所有可能的取值为 $c_k, k=1, 2, 3, \dots, l$ 。

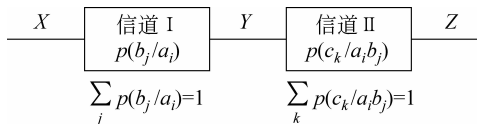


图 3.14 串联信道

信道 I 中,输入变量为 X ,输出变量为 Y 。信道 II 中,输入变量为 Y ,输出变量为 Z 。信道 I 的输出是信道 II 的输入,所以 Y 与 X 统计相关,而信道 II 的输出 Z 与输入 Y 统计相关,一般来说 Z 将与 X 统计相关,但串联结构又决定了在给定 Y 之后, Z 的取值就不与 X 有关了。所以, X, Y, Z 组成了一个马尔科夫链,根据马尔科夫链的性质,串联信道总的信道矩阵等于这两个串联信道的信道矩阵的乘积。当得到串联信道总的信道矩阵后,其信道容量就可以用求离散单符号信道容量的方法完成。

下面来分析串联信道的输出与输入过程,串联信道总的信道输出为两个串联信道的乘积,即

$$p(c_k/a_i) = \sum_j p(b_j/a_i) p(c_k/a_i b_j) \quad (3-33)$$

其中, $p(c_k/a_i b_j) = p(c_k/b_j)$ 对所有的 X, Y, Z 的取值均成立,则式(3-33)可写为

$$p(c_k/a_i) = \sum_j p(b_j/a_i) p(c_k/b_j)$$

则有

$$\begin{aligned} I(X; Z/Y) &= H(X/Y) - H(X/YZ) = H(Z/Y) - H(Z/XY) \\ &= H(X/Y) - H(XZ/Y) + H(Z/Y) = 0 \\ I(X; YZ) &= I(X; Y) + I(X; Z/Y) = I(X; Z) + I(X; Y/Z) \end{aligned}$$

所以

$$\begin{aligned} I(X; Y) &= I(X; Z) + I(X; Y/Z) \\ I(X; Y) &\geq I(X; Z) \end{aligned}$$

同理

$$\begin{aligned} I(XY; Z) &= I(X; Z) + I(Y; Z/X) = I(Y; Z) + I(X; Z/Y) \\ I(Y; Z) &\geq I(X; Z) \end{aligned}$$

所以 $p(b_j/a_i) = p(c_k/a_i)$ 对所有 a_i, b_j, c_k 都成立。

例 3.9 两个串联信道的信道矩阵分别为

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad \text{和} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

求该串联信道总的信道矩阵。

解：因为

$$p(c_k/a_i) = \sum_j p(b_j/a_i) p(c_k/b_j)$$

同时 $p(b_j/a_i) = p(c_k/a_i)$ 对所有 a_i, b_j, c_k 都成立, 则有

$$p(c_k/a_i) = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

此时 $I(X;Z) = I(X;Y)$ 。

若将串联信道从两个子信道的推广到多个子信道, 容易证明

$$P_{\text{总}} = P_1 P_2 \cdots P_N = \prod_{i=1}^N P_i \quad (3-34)$$

例 3.10 设离散二元对称信道的信道矩阵为

$$\mathbf{P}_0 = \begin{bmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{bmatrix}$$

求 N 个相同的该信道串联后的信道容量。

解：由公式(3-34)可知

$$\mathbf{P} = \mathbf{P}_0^N$$

要求该信道的信道容量, 首先构造

$$\mathbf{T}^{-1} \mathbf{P}_0 \mathbf{T} = \mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1-2\epsilon \end{bmatrix}$$

其中

$$\mathbf{T} = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \mathbf{T}^{-1}$$

则

$$\begin{aligned} \mathbf{P} = \mathbf{P}_0^N &= \mathbf{T} \mathbf{A}^N \mathbf{T}^{-1} = \mathbf{T} \begin{bmatrix} 1 & 0 \\ 0 & (1-2\epsilon)^N \end{bmatrix} \mathbf{T}^{-1} \\ &= \frac{1}{2} \begin{bmatrix} 1 + (1-2\epsilon)^N & 1 - (1-2\epsilon)^N \\ 1 - (1-2\epsilon)^N & 1 + (1-2\epsilon)^N \end{bmatrix} \end{aligned}$$

这时的串联信道仍旧是一个二元对称信道, 其信道容量可表示为

$$C_N = 1 - H\left(\frac{1 - (1-2\epsilon)^N}{2}\right)$$

当 N 趋近于无穷大时

$$\lim_{N \rightarrow \infty} \mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

这时信道容量为