第 6 章

离散信道及其容量

离散信道及其容量

★信道是信号的传输媒介，是传送信息的物理通道。

★研究信道主要目的是为了解决信息如何有效、可靠地传输的问题。

★本章重点解决某些特殊信道容量的计算问题。

本章主要内容

★ 概述

★ 单符号离散信道及其容量

★ 级联信道及其容量

★ 多维矢量信道及其容量

★ 信道容量的迭代计算

§6.1 概述

★ 信道的分类

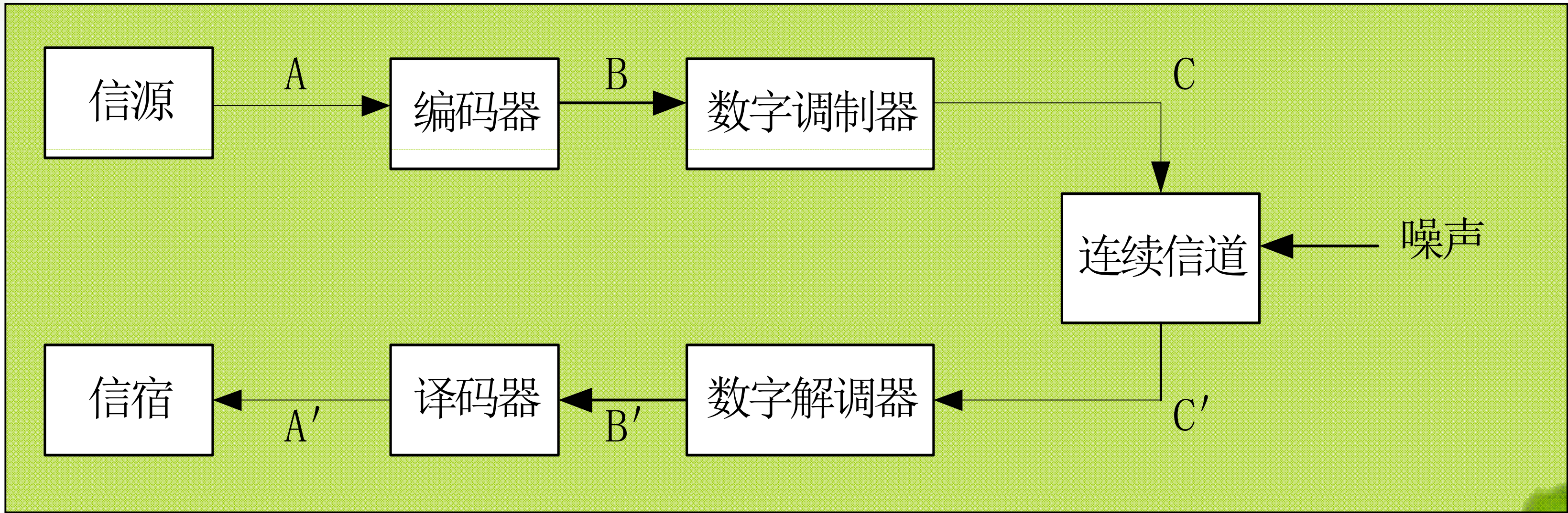
★ 离散信道的数学模型

★ 信道容量的定义

§6.1.1 信道的分类

★ 数字通信系统的基本模型

★ 依据不同的条件，不同模块之间的通道可以划分为不同的信道



§6.1.1 信道的分类

- ★ 按输入、输出集合的取值分类
- ★ 按输入、输出集合的个数分类
- ★ 按信道转移概率的性质分类
- ★ 根据信道统计特性划分
- ★ 根据信道噪声性质划分

§6.1.1 信道的分类

★ 按输入、输出集合的取值分类

1）离散信道：输入和输出均为离散集，如B-B’

2）连续信道：输入和输出均为连续集，也称波形信道，其特点是时间与取值都连续，如C-C’

3）半连续（或半离散）信道：输入和输出一个为连续、一个为离散，如B-C’或C-B’

4）时间离散连续信道：连续取值但时间离散，例如信道的输入和输出为模拟信号抽样的情况。

§6.1.1 信道的分类

★ 按输入、输出集合的个数分类

1） 单用户信道：X,Y中各有一个事件集，称单路或单端信道

2） 多用户信道：X,Y中至少有一端是多个事件集，也称多端信道。多用户信道包含两种特殊的信道，即多元接入信道和广播信道。

- ✓多元接入信道就是多个输入、单个输出的信道
- ✓广播信道就是单个输入、多个输出的信道

§6.1.1 信道的分类

★ 按信道转移概率的性质分类

1）无噪声信道
- 无损信道（每个输入对应多个输出）
- 确定信道（多个输入对应单个输出）
- 无扰信道（一个输入对应一个输出）

2）有噪声信道
- 无记忆信道 给定时间输出仅依赖于当前输入
- 有记忆信道 输出值同时依赖当前和以前的输入

§6.1.1 信道的分类

★ 根据信道统计特性划分

1）恒参信道 统计特性不随时间变化（也称平稳信道）
例如：卫星通信信道

2）变参信道 统计特性随时间变化。例如：短波，移动通信信道

§6.1.1 信道的分类

★ 根据信道噪声性质划分

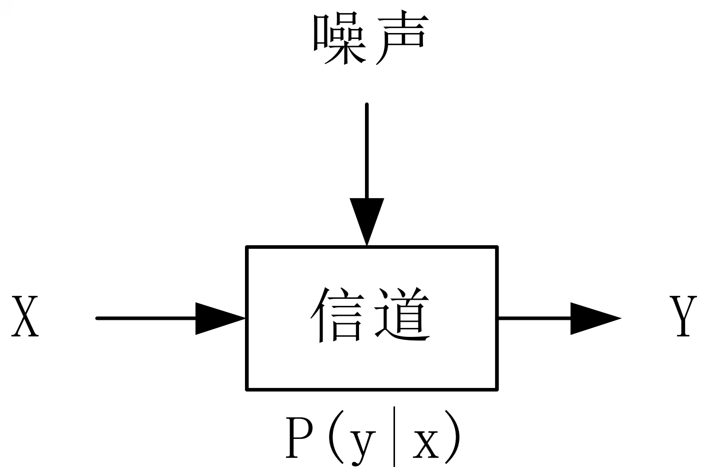
1）高斯噪声信道

信道噪声为高斯分布（白噪声或有色噪声）

2）非高斯噪声信道

信道噪声分布不是高斯分布

§6.1.2 离散信道的数学模型



离散无记忆信道

★ 一般的信道数学模型

★ 离散无记忆信道

★ 平稳（或恒参）信道

★ 单符号离散信道

一般信道的数学模型

★信道模型

$$\left\{X^N, p(\vec{y} \mid \vec{x}), Y^N\right\}$$

$$p(\vec{y} \mid \vec{x}) = p(y_1, y_2, \cdots y_N \mid x_1, x_2, \cdots x_N)$$

$$X^N = X_1 X_2 \cdots X_N \qquad \vec{x} = x_1, x_2, \cdots, x_N \qquad \vec{x} \in X^N$$

$$Y^N = Y_1 Y_2 \cdots Y_N \qquad \vec{y} = y_1, y_2, \cdots y_N \qquad \vec{y} \in Y^N$$

离散无记忆信道

★ 若信道的转移概率满足

$$p(\vec{y} \mid \vec{x}) = \prod_{n=1}^{N} p(y_n \mid x_n)$$

则称为此信道为离散无记忆信道（DMC），其数学模型为：

$$\{X, p(y_n \mid x_n), Y\}$$

利用给定时刻的输出符号仅依赖于当前输入符号的条件可以推出。

平稳（或恒参）信道

★ 如果对于任意正整数m、n，和 $a_i \in A, b_i \in B$

离散无记忆信道的转移概率满足：

$$p(y_n = b_j \mid x_n = a_i) = p(y_m = b_j \mid x_m = a_i)$$

则称为平稳或恒参信道

可见，对于平稳信道，$p(y_n \mid x_n)$ 不随时间变化。这样，平稳信道的模型就是

$$\{X, p(y \mid x), Y\}$$

无需规定时间

单符号离散信道

★ 对于离散平稳无记忆信道，可以用一维条件概率描述(和先前输入无关)

★ 这种用一维条件概率描述的信道为：单符号离散信道

其中，信道的输入X与输出Y都是一维随机变量集合，x∈X，取自字母表，$A=\{a_1,\ldots,a_r\}$。y∈Y，取自字母表 $B=\{b_1,\ldots,b_s\}$

★ 信道转移概率简记为：

$$p(y|x)=p(y=b_j|x=a_i)=P_{Y/X}(b_j|a_i)\equiv p_{ij}$$

单符号离散信道

★ 信道转移概率矩阵

$$[P]=\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1s} \\ p_{21} & p_{22} & \cdots & p_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ p_{r1} & p_{r2} & \cdots & p_{rs} \end{bmatrix}$$

p_{ij}是由a_i转移到b_j的概率

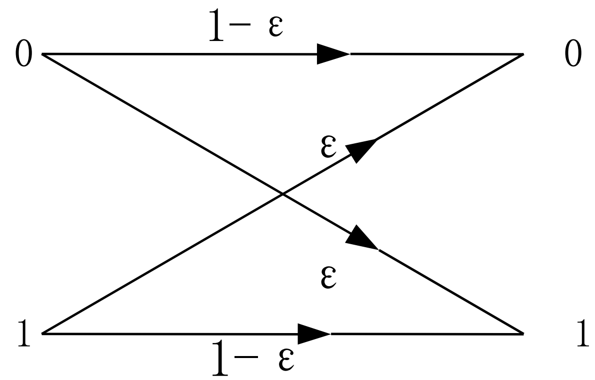
单符号离散信道

例 6.1.1 二元对称信道(BSC)，输入与输出符号集分别为 $A=\{0,1\},B=\{0,1\}$，信道转移概率p(y/x)满足 $P_{Y/X}(0|0)=P_{Y/X}(1|1)=1-\varepsilon$， ε 称为错误率。写出信道的转移概率矩阵并画出转移概率图。

解： 转移概率矩阵

$$[P]=\begin{bmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{bmatrix}$$

转移概率图



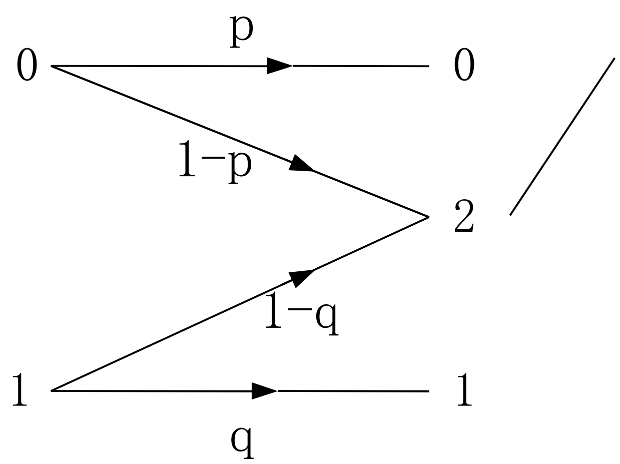
单符号离散信道

例 6.1.2　二元删除信道(BEC)：其中A={0, 1}，　B={0, 2, 1}
画出转移概率图和转移概率矩阵。

解：　转移概率矩阵

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

转移概率图



2为擦除点,符号0
以(1-p)的比例
在信道中丢失

单符号离散信道

例 6.1.3 四个等概消息，编成的码字为 $M_1=000$，$M_2=011$，$M_3=101$，$M_4=110$，当通过下图所示二进制对称无记忆信道传输时，求:

1）“接收到第一个数字为0”与“发M_1”的互信息

2）当“接收到第二个数字也为0”时，关于M_1的附加信息

3）当“接收到第三个数字也为0”时，又增加多少关于M_1的信息？

p(0|0)+p(0|0)+p(0|1)+p(0|1)

解： 1）$q(''0'')=\sum_{i=1}^{4}p(M_i)p(0|M_i)=\frac{1}{4}[2(1-\varepsilon)+2\varepsilon]=\frac{1}{2}$

$$\Rightarrow \text{互信息} I(M_1;''0'')=\log\frac{p(0|M_1)}{q(''0'')}=\log\frac{1-\varepsilon}{1/2}=\log[2(1-\varepsilon)]$$

单符号离散信道

2) $q("00") = \sum_{i=1}^{4} p(M_i)p(00|M_i) = \frac{1}{4}[p(0|0)p(0|0)+p(0|0)p(0|1)+p(0|1)p(0|0)+p(0|1)p(0|1)]$

$= \frac{1}{4}[(1-\varepsilon)^2 + 2\varepsilon(1-\varepsilon) + \varepsilon^2] = \frac{1}{4}$

$\Rightarrow$ 互信息$I(M_1;"00") = \log\frac{p(00|M_1)}{q("00")} = \log\frac{(1-\varepsilon)^2}{1/4} = 2\log[2(1-\varepsilon)]$

$\Rightarrow$ 附加信息 $= 2\log[2(1-\varepsilon)] - \log[2(1-\varepsilon)] = \log[2(1-\varepsilon)]$

3) $q("000") = \sum_{i=1}^{4} p(M_i)p(000|M_i) = \frac{1}{4}[(1-\varepsilon)^3 + 3(1-\varepsilon)\varepsilon^2] = \frac{1}{4}(1-\varepsilon)(4\varepsilon^2 - 2\varepsilon + 1)$

$\Rightarrow$ 互信息$I(M_1;"000") = \log\frac{p(000|M_1)}{q("000")} = 2\log[2(1-\varepsilon)] - \log(4\varepsilon^2 - 2\varepsilon + 1)$

又增加的信息 $-\log(4\varepsilon^2 - 2\varepsilon + 1)$

§6.1.3 信道容量的定义

★ 平稳离散无记忆信道的容量C定义为输入与输出平均互信息I(X;Y)的最大值：

$$C \equiv \max_{p(x)} I(X;Y)$$

决定无差错传输上限，C即为该上限，p(x)调整达到上限

1）单位为：比特/信道符号（奈特/信道符号）

2）当信道给定后，p(y|x)就固定，所以C 仅与p(y|x)有关，而与p(x)无关

3）C是信道无差错传输最大信息速率能力的度量

多维矢量信道

★若 X^N 和 Y^N 分别为信道的N维输入与输出随机矢量集合，则信道容量定义为：

$$C \equiv \max_{p(x_1 \cdots x_N)} I(\mathbf{X}^N;\mathbf{Y}^N)$$

其中， $p(x_1 \cdots x_N)$ 为信道输入矢量的联合概率

说明：

(1)信息论教材中,信道容量单位为bit/per channel use

(2) bit/per channel use 可以是单符号或多维矢量

§6.2 单符号离散信道及其容量

★ 离散无噪信道的容量

★ 离散对称信道的容量

★ 一般离散信道的容量

§6.2.1 离散无噪信道的容量

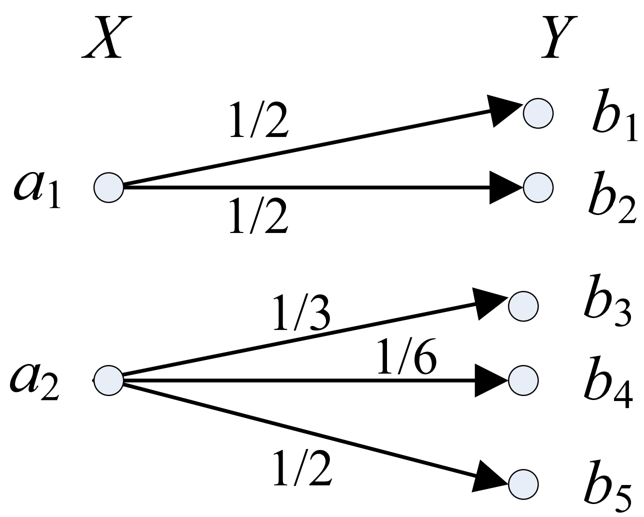
1. I(X;Y)=H(X)-H(X|Y)
   =H(X)
2. p(X)均匀分布时，H(X)最大

★ 无损信道：输出符号只对应一个输入符号

$$C = \max H(X) = \log r \ \ (\text{比特/符号})$$

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

其中r为输入符号集的大小



§6.2.1 离散无噪信道的容量

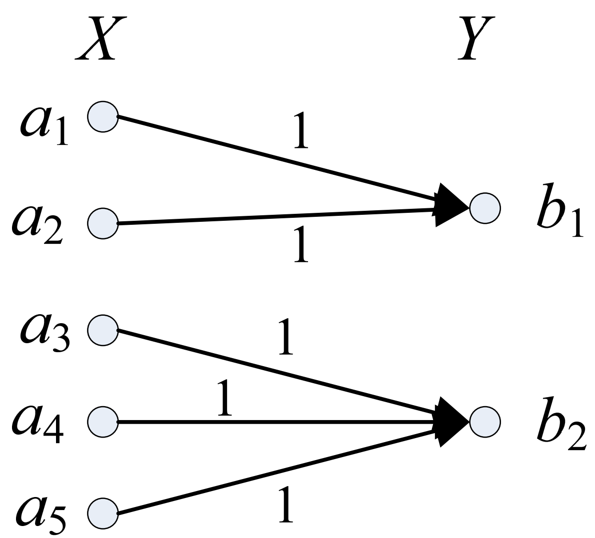
★ 确定信道：每个输入符号都对应一个输出符号

$$C = \max H(Y) = \log s \text{ (比特/符号)}$$

$$H(Y \mid X) = 0$$

1. I(X;Y)=H(Y)-H(Y|X)
2. H(Y|X)=0

其中s为输出符号集的大小



相当于一个信息有多个编码

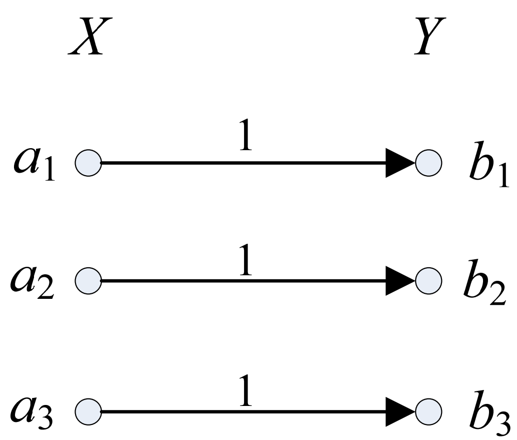
§6.2.1 离散无噪信道的容量

★ 无损确定信道：输入符号与输出符号是一一对应关系

$$C = \max H(Y) = \log s = \log r \text{ (比特符号)}$$

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

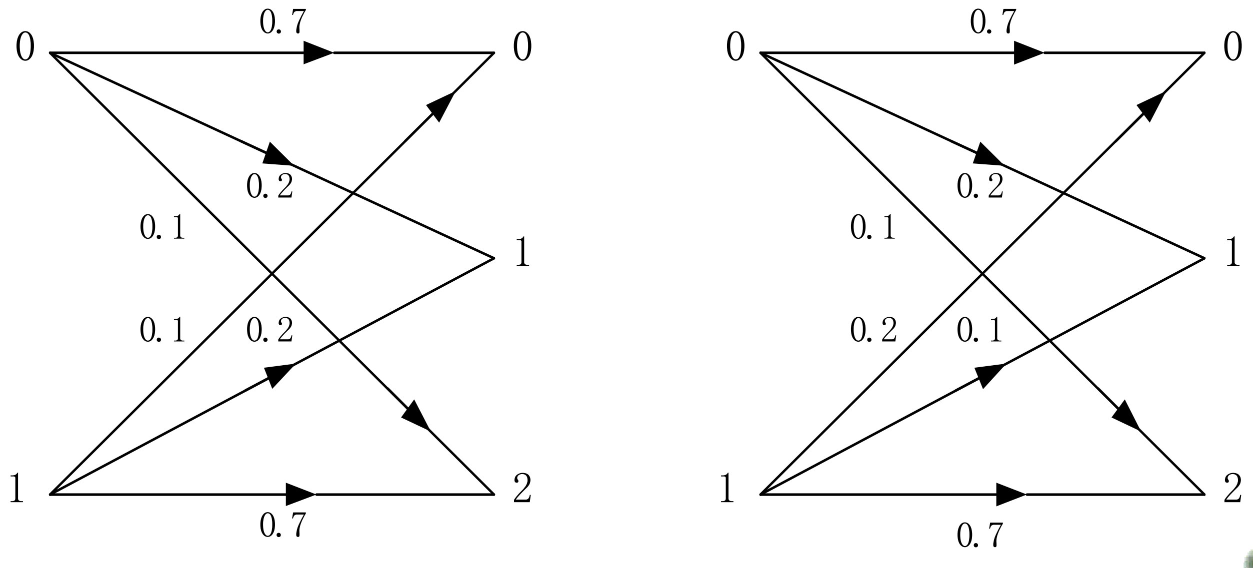
其中r、s为输入与输出字母表的大小，且r=s



§6.2.2 离散对称道的容量

若一个信道的转移概率矩阵按 **输出** 可分为若干子集，其中每个子集都有如下特性：即每一行是其他行的置换，每一列是其他列的置换，则信道称为 对称信道

例 6.2.1 分析下图信道的对称性



(a) (b)

§6.2.2 离散对称道的容量

解：(a) 可分成两个子矩阵

$$\begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.7 \end{bmatrix} \Rightarrow \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix} \begin{bmatrix} 0.7 & 0.1 \\ 0.1 & 0.7 \end{bmatrix} \Rightarrow \text{对称信道}$$

置换：每行或每列包含的元素相同

(b) 的概率转移矩阵为

$$\begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.1 & 0.7 \end{bmatrix}$$

所以不是对称信道

有时将转移概率矩阵可分成多个子集的对称信道为准对称或弱对称信道，而只有一个子集的对称信道称强对称信道

§6.2.2 离散对称道的容量

★ 定理6.2.1 对于离散对称信道，当输入等概率时达到信道容量：

注意：对称信道每行包含的元素一定相同

$$C = H(Y) - H(p_1, p_2, \ldots, p_s)$$

H(Y)为输入等概率时输出的熵

$H(p_1, p_2, \cdots, p_s)$为信道转移概率矩阵某行的元素

注释：对强对称信道,输入等概率时达到容量，此时输出等概率。

强对称，则任一行概率分布相同

$$C = \log s - H(p_{11}, p_{12}, \ldots, p_{1s})$$

§6.2.2 离散对称道的容量

例 6.2.2 一信道的转移概率矩阵如图，求信道容量和达到容量时的输出概率。

$$\begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 1/6 & 1/2 & 1/3 \\ 1/3 & 1/6 & 1/2 \end{bmatrix}$$

解： 设输出概率为 q_1, q_2, q_3 。由于信道为强对称信道，故当输入等概率时达到容量C，此时输出也等概率

$$\Rightarrow \quad q_1 = q_2 = q_3 = 1/3$$

$$C = \log 3 - H(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}) = 1.126 \text{ 比特/符号}$$

§6.2.2 离散对称道的容量

例 6.2.3 一信道的转移概率矩阵如图，求信道容量和达到容量时的输入概率。

$$\begin{bmatrix} 1-p & \frac{p}{r-1} & \cdots & \frac{p}{r-1} \\ \frac{p}{r-1} & 1-p & \cdots & \frac{p}{r-1} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{p}{r-1} & \frac{p}{r-1} & \cdots & 1-p \end{bmatrix}$$

解：设输入输出概率为 $p_i, q_j, i=1,2,\cdots,r$

由于信道为强对称信道，故当

$p_1 = \ldots = p_r = 1/r$ 时，达到容量。

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

$$= \log r - H(p) - p\log(r-1)$$

特别是，当r=2时，信道容量为C=1-H(p)比特/符号。

§6.2.2 离散对称道的容量

例 6.2.4 一信道的转移概率矩阵如图，求信道容量和达到容量时的输出概率。

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

解：设输出概率为 q_1, q_2, q_3, q_4

准对称信道，当输入等概率时达到信道容量，根据转移阵计算输出概率：

$$q_1=\frac{1}{2}[\frac{1}{6}+\frac{1}{3}]=\frac{1}{4} \qquad q_2=\frac{1}{2}[\frac{1}{3}+\frac{1}{3}]=\frac{1}{3} \qquad q_3=\frac{1}{2}[\frac{1}{6}+\frac{1}{6}]=\frac{1}{6} \qquad q_4=\frac{1}{4}$$

$$C=\underbrace{H(\frac{1}{4},\frac{1}{3},\frac{1}{6},\frac{1}{4})}_{H(Y)}-H(\frac{1}{3},\frac{1}{3},\frac{1}{6},\frac{1}{6})$$

$$=0.041\text{比特/符号}$$

§6.2.3 一般离散信道的容量

$I(X;Y)$为$p(x)$上凸函数 $\Rightarrow$ 极大值存在

$p(x)$要满足非负且归一化条件 $\Rightarrow$

求信道容量归结为求有约束极值的问题

$$p_i = p(x),\ p_{ij} = P(y|x),\ \mathrm{q}_\mathrm{j} = q(y)$$

§6.2.3 一般离散信道的容量

作为计算合理性验证

求 $I(X;Y)=\sum_{i,j} p_i p_{ij}\log\frac{p_{ij}}{q_j}$ 在约束 $\sum_i p_i=1, \boxed{p_i\ge 0}$ 下的极值

①利用拉格朗日乘子法，求函数 $J=I(X;Y)-\lambda\sum_i p_i$ 的极值

②计算 $\frac{\partial J}{\partial p_k}$ 并使其为0

$$\frac{\partial J}{\partial p_k}=\frac{\partial}{\partial p_k}\left[\sum_{i,j} p_i p_{ij}\log p_{ij}-\sum_j q_j\log q_j-\lambda\sum_i p_i\right]$$

$$=\sum_j p_{kj}\log p_{kj}-\sum_j(p_{kj}\log q_j+p_{kj}\log e)-\lambda=0$$

利用乘积求导公式

考虑到 $q_j=\sum_{i=1}^{r} p_i p_{ij}$ $\Rightarrow$ $\sum_j p_{kj}\log\frac{p_{kj}}{q_j}=\log e+\lambda$

§6.2.3 一般离散信道的容量

$$I(a_k;Y)=\sum_j p_{kj}\log\frac{p_{kj}}{q_j}\quad k=1,\dots,r$$

$$\sum_k p_k I(a_k;Y)=I(X;Y)$$

$$C=\sum_k p_k I(a_k;Y)=\sum_k p_k(\log e+\lambda)=\log e+\lambda$$

$$\sum_j p_{kj}\log\frac{p_{kj}}{q_j}=\log e+\lambda$$

§6.2.3 一般离散信道的容量

c=Σ(j)p_ij×c

$$\sum_j p_{ij}\log\frac{p_{ij}}{q_j}=c\Rightarrow\sum_j p_{ij}\log p_{ij}=\sum_j p_{ij}(c+\log q_j)$$

$$令\beta_j=c+\log q_j\Rightarrow\boxed{\sum_j p_{ij}\beta_j=\sum_j p_{ij}\log p_{ij}}$$

$$\begin{cases}\beta_j=c+\log q_j\\ \sum_j q_j=1\end{cases}\Rightarrow\sum_j 2^{\beta_j-c}=1\Rightarrow c=\log\sum_j 2^{\beta_j}$$

$$q_j=2^{\beta_j-c}$$

$$q_j=\sum_i p_i p_{ij}\Rightarrow\{p_i\}$$ $\Longrightarrow$ 验证C的正确性

§6.2.3 一般离散信道的容量

用矩阵表示

注：1. 假定信道模型已知，即 p_ij已知
2. s 个方程

$$\sum_j p_{ij}\beta_j = \sum_j p_{ij}\log p_{ij}$$

$$[P] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1s} \\ p_{21} & p_{22} & \cdots & p_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ p_{r1} & p_{r2} & \cdots & p_{rs} \end{bmatrix} = [p_{ij}]$$

$$h_i = -\sum_{j=1}^{s} p_{ij}\log p_{ij}$$

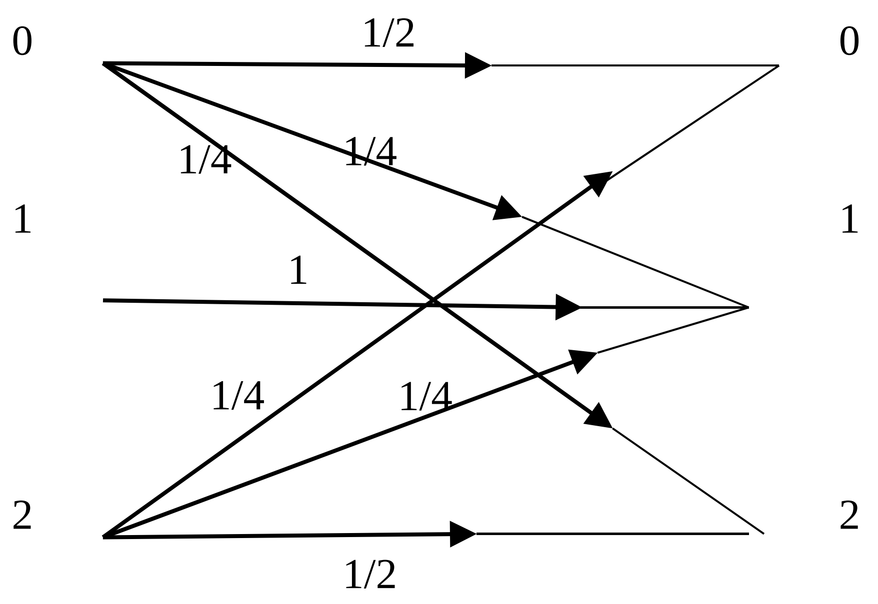
$$\Rightarrow [P][\beta] = -[h] \Rightarrow \boxed{[\beta] = -[P]^{-1}[h]}$$

§6.2.3 一般离散信道的容量

- 一般信道容量计算方法
  - 给定信道转移概率矩阵 $[P]=[p_{ij}]_{r\times s}$， 求其逆 $[P]^{-1}$
  - 计算 h ： $h_i=-\sum_{j=1}^{s} p_{ij}\log p_{ij}$
  - 计算 β ：$[\beta]=-[P]^{-1}[h]$
  - 计算信道容量： $C=\log\sum_j 2^{\beta_j}$
  - 验证：计算 $q_j=2^{\beta_j-c}$，利用$[P]^{-1}$ 计算 p_i ，如果所有 p_i >0，则结果正确；否则，设定某些 p_i 为0，重算

§6.2.3 一般离散信道的容量

例 一信道的转移概率如图所示，求信道容量和达到容量时的输出概率。



§6.2.3 一般离散信道的容量

解：

$$p=\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & \frac{1}{4}\\ 0 & 1 & 0\\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2}\end{bmatrix}$$

对称信道的基本条件：
每行包含的元素相同

$$\Rightarrow\begin{cases}\frac{1}{2}\beta_1+\frac{1}{4}\beta_2+\frac{1}{4}\beta_3=\frac{1}{2}\log\frac{1}{2}+\frac{1}{4}\log\frac{1}{4}+\frac{1}{4}\log\frac{1}{4}\\ \beta_2=0\\ \frac{1}{4}\beta_1+\frac{1}{4}\beta_2+\frac{1}{2}\beta_3=\frac{1}{2}\log\frac{1}{2}+\frac{1}{4}\log\frac{1}{4}+\frac{1}{4}\log\frac{1}{4}\end{cases}\Rightarrow\begin{cases}\beta_1=-2\\ \beta_2=0\\ \beta_3=-2\end{cases}$$

$$\Rightarrow c=\log(2^{-2}+2^{0}+2^{-2})=\log\frac{3}{2}=0.585\text{比特/符号}$$

$$\Rightarrow\begin{cases}q_0=q_2=\frac{1}{6}\\ q_1=\frac{2}{3}\end{cases}\Rightarrow\begin{cases}p_0=p_2=\frac{2}{9}\\ p_1=\frac{5}{9}\end{cases}$$

§6.2.3 一般离散信道的容量

例 利用 $[\beta]=-[P]^{-1}[h]$ 求例6.1.1中二元对称信道容量。

解：
$$\left.\begin{array}{l}\varepsilon\neq 1/2\\ \mathbf{P}=\begin{bmatrix}1-\varepsilon & \varepsilon\\ \varepsilon & 1-\varepsilon\end{bmatrix}\end{array}\right\}\Rightarrow [P]^{-1}=\frac{1}{1-2\varepsilon}\begin{bmatrix}1-\varepsilon & -\varepsilon\\ -\varepsilon & 1-\varepsilon\end{bmatrix}$$

$$\left.\begin{array}{l}[P]^{-1}=\frac{1}{1-2\varepsilon}\begin{bmatrix}1-\varepsilon & -\varepsilon\\ -\varepsilon & 1-\varepsilon\end{bmatrix}\\ [h]=-\begin{bmatrix}(1-\varepsilon)\log(1-\varepsilon)+\varepsilon\log\varepsilon\\ (1-\varepsilon)\log(1-\varepsilon)+\varepsilon\log\varepsilon\end{bmatrix}\end{array}\right\}\Rightarrow$$

$$[\beta]=-[P]^{-1}[h]=\frac{1}{1-2\varepsilon}\begin{bmatrix}1-\varepsilon & -\varepsilon\\ -\varepsilon & 1-\varepsilon\end{bmatrix}\cdot\begin{bmatrix}(1-\varepsilon)\log(1-\varepsilon)+\varepsilon\log\varepsilon\\ (1-\varepsilon)\log(1-\varepsilon)+\varepsilon\log\varepsilon\end{bmatrix}$$

$$=\begin{bmatrix}(1-\varepsilon)\log(1-\varepsilon)+\varepsilon\log\varepsilon\\ (1-\varepsilon)\log(1-\varepsilon)+\varepsilon\log\varepsilon\end{bmatrix}$$

§6.2.3 一般离散信道的容量

$\Longrightarrow$ 信道容量为

$$C=\log_2 2\cdot 2^{(1-\varepsilon)\log(1-\varepsilon)+\varepsilon\log\varepsilon}=1+(1-\varepsilon)\log(1-\varepsilon)+\varepsilon\log\varepsilon=1-H(\varepsilon)\ (\text{比特/符号})$$

$$q_j=2^{\beta_j-c} \qquad \Rightarrow \text{输出概率为}\ q_1=q_2=\frac{1}{2}$$

对应的输入概率为 $q_j=\sum_i p_i p_{ij} \Rightarrow \{p_i\}$

$$\Rightarrow \text{输入概率为}\ p_1=p_2=\frac{1}{2}$$

§6.2.3 一般离散信道的容量

★ 定理6.2.2 对于离散无记忆信道，当且仅当

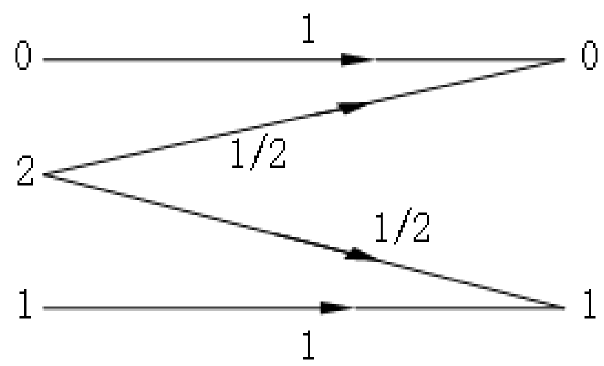
$$I(a_i;Y)=C\text{, 对于} p_i>0$$

$$I(a_i;Y)\le C\text{, 对于} p_i=0$$

I(X;Y)达到最大值，此时C为信道容量

§6.2.3 一般离散信道的容量

例 6.2.6 信道的转移概率如下图所示，求信道容量和达到容量时的输入/出概率。



解：设输入、输出概率为p_0, p_1, p_2. q_0, q_1

1）达到容量时，若输入概率全不为零

$$
\left.\begin{aligned}
-\log q_0 &= C \quad (p_0 > 0) \\
-1-\frac{1}{2}\log q_0 q_1 &= C \quad (p_2 > 0) \\
-\log q_1 &= C \quad (p_1 > 0)
\end{aligned}\right\}
$$

据 $I(a_k;Y)=\sum_j p_{kj}\log\frac{p_{kj}}{q_j}$

解得 $q_0=q_1=1/2$ ，C=1比特，但将结果代入第2式，使该式左边的值为0，出现矛盾（ $p_2>0$ ）。

§6.2.3 一般离散信道的容量

2）设p_2=0，p_0，p_1，不为零

1. 只使用方程1和3，得q0和q1
2. 利用P反推p1和p3

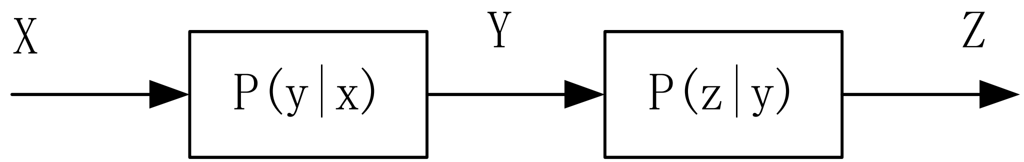
$$-\log q_0 = -\log q_1 = C \Longrightarrow q_0 = q_1 = 1/2, C = 1\text{比特}/\text{符号}$$

将结果代入第2式，该式左边的值为0<C。所以，信道容量C=1比特/符号，达到容量时的输入概率为概率为

$$p_0 = p_1 = \frac{1}{2}, p_2 = 0$$

§6.3 级联信道及其容量

★ 级联的含义是被连接的信道输入只依赖于前面相邻信道的输出而和前面的其它信道的输出无直接关系



★ 若随机变量集合(X, Y, Z)构成马氏链，则称信道X-Y与Y-Z构成级联信道。由于当Y给定时，Z不依赖于X，即 $P(z|y)=P(z|xy)$

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

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

§6.3 级联信道及其容量

★ 定理6.3.1 若X, Y, Z构成一马氏链，则：

$$\boxed{I(X;Z)\le I(X;Y)}$$

$$\boxed{I(X;Z)\le I(Y;Z)}$$

证

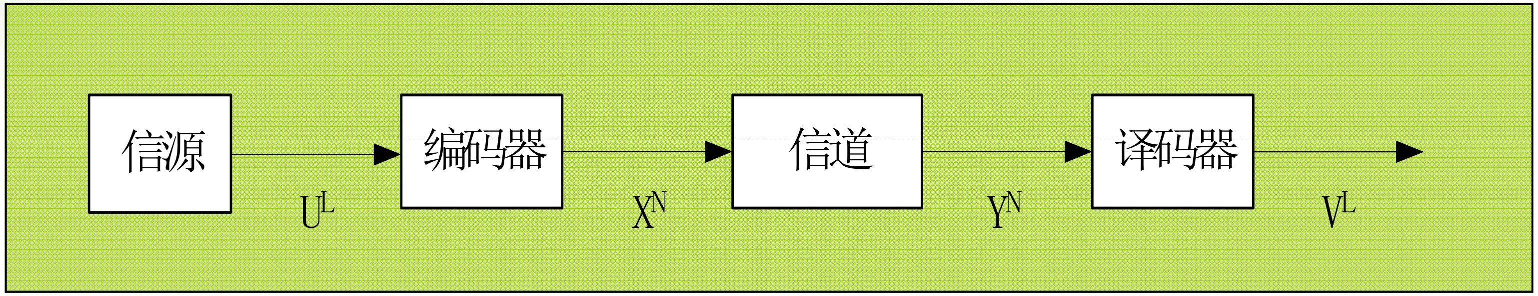
$$\left.\begin{array}{l} I(X;YZ)=I(X;Y)+I(X;Z|Y)=I(X;Z)+I(X;Y|Z) \\ \text{马氏链} \Rightarrow I(X;Z|Y)=0 \end{array}\right\} \Rightarrow$$

$I(X;Z)\le I(X;Y)$ 得证

同理可证： $I(X;Z)\le I(Y;Z)$

§6.3 级联信道及其容量

★ 通信系统模型各部分的级联



信道传输后，译码器收到N长序列为Y^N，译码后传给信宿的消息序列为V^L 。

§6.3 级联信道及其容量

★ 定理6.3.2(数据处理定理)

$$\boxed{I(U^L;V^L)\le I(X^N;Y^N)}$$

证 (U^L,X^N,Y^N)构成马氏链 $\Rightarrow$ $I(U^L;Y^N)\le I(X^N;Y^N)$

(U^L,Y^N,V^L)构成马氏链 $\Rightarrow$ $I(U^L;V^L)\le I(U^L;Y^N)$

★ 定理的含义：从信宿得到的关于信源的信息经过编译码器、信道的处理后会减少，而且处理的次数越多，减少得越多。

§6.3 级联信道及其容量

★ 级联信道的转移概率矩阵

级联信道为马氏链 $\Rightarrow$ 一级级联相当于状态的一步转移

$\Rightarrow$ 级联信道的转移概率矩阵为级联信道中各矩阵依次相乘

★ 级联信道的容量

得到级联信道转移概率阵后，按照前面给出的离散信道容量的计算方法即可计算级联信道容量。

§6.3 级联信道及其容量

例 6.3.1 给定二元对称信道其状态转移矩阵如下，计算两级级联信道的概率转移矩阵。如果信道输入0、1 等概率，求在两级级联和三级级联情况下输入与输出的平均互信息。

$$[P]=\begin{bmatrix}1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon\end{bmatrix}$$

解： 1）两级级联信道的概率转移矩阵

对称信道

$$[P][P]=\begin{bmatrix}1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon\end{bmatrix}\begin{bmatrix}1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon\end{bmatrix}=\begin{bmatrix}(1-\varepsilon)^2+\varepsilon^2 & 2\varepsilon(1-\varepsilon) \\ 2\varepsilon(1-\varepsilon) & (1-\varepsilon)^2+\varepsilon^2\end{bmatrix}$$

§6.3 级联信道及其容量

2)设原信道输入与输出集分别为X、Y，两级级联和三级级联情况下输出集合分别为Z、U

$$\begin{bmatrix} X \\ P_X \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} \Rightarrow \begin{bmatrix} Y \\ P_Y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} \quad \begin{bmatrix} Z \\ P_Z \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$$

$$I(X;Y) = H(Y) - H(Y|X) = 1 - H(\varepsilon)$$
$$I(X;Z) = 1 - H(Z|X) = 1 - H[2\varepsilon(1-\varepsilon)]$$

其中 $H(\varepsilon) = -\varepsilon\log\varepsilon - (1-\varepsilon)\log(1-\varepsilon)$

类似地,可计算三级级联信道的情况，注意也为对称信道

$$I(X;U) = 1 - H[3\varepsilon(1-\varepsilon)^2 + \varepsilon^3]$$

结论：信道级联后增加信息损失，级联级数越多，损失越大。

§6.3 级联信道及其容量

例 6.3.1（续） 设错误概率ε为1/3，计算两级级联信道的容量及达到容量时的输出概率。

解： 两级级联信道的转移矩阵为

$$P=\begin{pmatrix}2/3 & 1/3\\ 1/3 & 2/3\end{pmatrix}\begin{pmatrix}2/3 & 1/3\\ 1/3 & 2/3\end{pmatrix}=\begin{pmatrix}5/9 & 4/9\\ 4/9 & 5/9\end{pmatrix}$$

该级联信道是一个强对称信道，因此当输入等概时达到信道容量，此时输出也等概。所以

$$q_1=q_2=1/2$$

$$C=\log 2-H(\frac{5}{9},\frac{4}{9})=\log 2+(\frac{5}{9}\log\frac{5}{9}+\frac{4}{9}\log\frac{4}{9})=1.991\ \text{比特/符号}$$

§6.4 多维矢量信道及其容量

★多维矢量信道输入与输出的性质

★离散无记忆扩展信道及其容量

★并联信道及其容量

★和信道及其容量

§6.4.1 多维矢量信道输入与输出的性质

★ 对于多维矢量信道，输入与输出平均互信息为：

$$\boxed{I(X^N;Y^N)=H(X^N)-H(X^N|Y^N)}$$

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

$$=\sum_{X^N,Y^N} p(\vec{x}\vec{y})\log\frac{p(\vec{y}|\vec{x})}{q(\vec{y})}$$

§6.4.1 多维矢量信道输入与输出的性质

★ 引理6.4.1

设信道的输入输出分别为 X^N, Y^N ，其中 $X^N = X_1 \cdots X_N$，$Y^N = Y_1 \cdots Y_N$，则：

1）

$$H(X^N | Y^N) \le \sum_{i=1}^{N} H(X_i | Y_i)$$

注意：这个不是信道无记忆

当且仅当 $p(\vec{x} | \vec{y}) = \prod_{i=1}^{N} p(x_i | y_i)$时等式成立

2）

$$H(Y^N | X^N) \le \sum_{i=1}^{N} H(Y_i | X_i)$$

仅当信道无记忆时等式成立

§6.4.1 多维矢量信道输入与输出的性质

★ 定理6.4.1 对于离散无记忆信道，有：

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

证

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

信道独立

$$\sum_{i=1}^{N}I(X_i;Y_i)=\sum_{i=1}^{N}H(Y_i)-\sum_{i=1}^{N}H(Y_i|X_i)$$

$$H(Y^N)=H(Y_1...Y_N)\le\sum_{i=1}^{N}H(Y_i)$$

仅当输出独立时等式成立。

§6.4.1 多维矢量信道输入与输出的性质

★当输入独立时

$$p(\vec{x})=p(x_1)\ldots p(x_N)$$

$$p(\vec{y})=\sum_{X^N}p(\vec{x})p(\vec{y}|\vec{x})=\sum_{X_1}\cdots\sum_{X_N}\{\prod_{i=1}^{N}p(x_i)p(y_i|x_i)\}=\prod_{i=1}^{N}\{\sum_{X_i}p(x_i)p(y_i|x_i)\}=\prod_{i=1}^{N}p(y_i)$$

即当信源信道都无记忆时，等式成立。

§6.4.1 多维矢量信道输入与输出的性质

★ 定理6.4.2 对于无记忆信源，有：

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

证

$$I(X^N;Y^N)=H(X^N)-H(X^N|Y^N)=\sum_{i=1}^{N}H(X_i)-H(X^N|Y^N)$$

无记忆信源

较小

$$\sum_{i=1}^{N}I(X_i;Y_i)=\sum_{i=1}^{N}H(X_i)-\sum_{i=1}^{N}H(X_i|Y_i)$$

较大

等式成立条件： $p(\vec{x}|\vec{y})=\prod_{i=1}^{N}p(x_i|y_i)$

§6.4.1 多维矢量信道输入与输出的性质

★当信道无记忆时 信源无记忆：p(Y)=Σp(X)p(Y|X)=Σp(X1)p(Y1|X1)·
p(Xn)p(Yn|Xn)

$$p(\vec{x}\,|\,\vec{y})=\frac{p(\vec{x})p(\vec{y}\,|\,\vec{x})}{p(\vec{y})}=\frac{\prod_i\{p(x_i)p(y_i\,|\,x_i)}{\sum_{x_1,x_2,\cdots x_N} p(x_1)p(y_1\,|\,x_1)p(x_2)p(y_2\,|\,x_1)\ldots}$$

$$=\frac{\prod_i p(y_i)p(x_i\,|\,y_i)}{\prod_i p(y_i)}=\prod_{i=1}^{N}p(x_i\,|\,y_i)$$

即当信源信道都无记忆时，等式成立。

§6.4.1 多维矢量信道输入与输出的性质

★ 结论：

1）对于无记忆信源和无记忆信道，有：

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

2）对于平稳信源，因为X_i、Y_i同分布，因此：

$$I(X_i;Y_i)=I(X;Y)$$

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

⇒ 对于平稳离散无记忆信道（DMC）的N次扩展信道，当信源无记忆时，有

$$I(X^N;Y^N)=NI(X;Y)$$

§6.4.1 多维矢量信道输入与输出的性质

例 6.4.1　设无记忆信源X的5次扩展源为X_5，信道为下面矩阵所示的置换信道，其中第1行为输入的序号，第2行为信道输出的序号，例如X_1输出到Y_3，X_2输出到Y_2等。计算 $I(X^5;Y^5)$ 和 $\sum_{i=1} I(X_i;Y_j)$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix}$$

解：
$$I(X_1;Y_1)=I(X_1;X_3)=0$$
$$I(X_2;Y_2)=I(X_2;X_2)=H$$
$$I(X_3;Y_3)=I(X_4;Y_4)=I(X_5;Y_5)=0$$

$$\Rightarrow \sum_{i=1}^{5} I(X_i;Y_i)=H$$

$$I(X^5;Y^5)=H(X^5)=5H$$

§6.4.1 多维矢量信道输入与输出的性质

例 6.4.2 设离散无记忆信道的输入 $X^N = X_1 .. X_N$，输出 $Y^N = Y_1 .. Y_N$，且有 $X_1 = Y_1 = X_2 = Y_2, ..., X_N = Y_N = X$

计算 $I(X^N;Y^N)$ 和 $\sum_{i=1}^{N} I(X_i;Y_j)$

解：

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

$$= H(X_1 .. X_N) = H(X_1) + H(X_2 \mid X_1) + ... + H(X_N \mid X_1 ... X_{N-1})$$

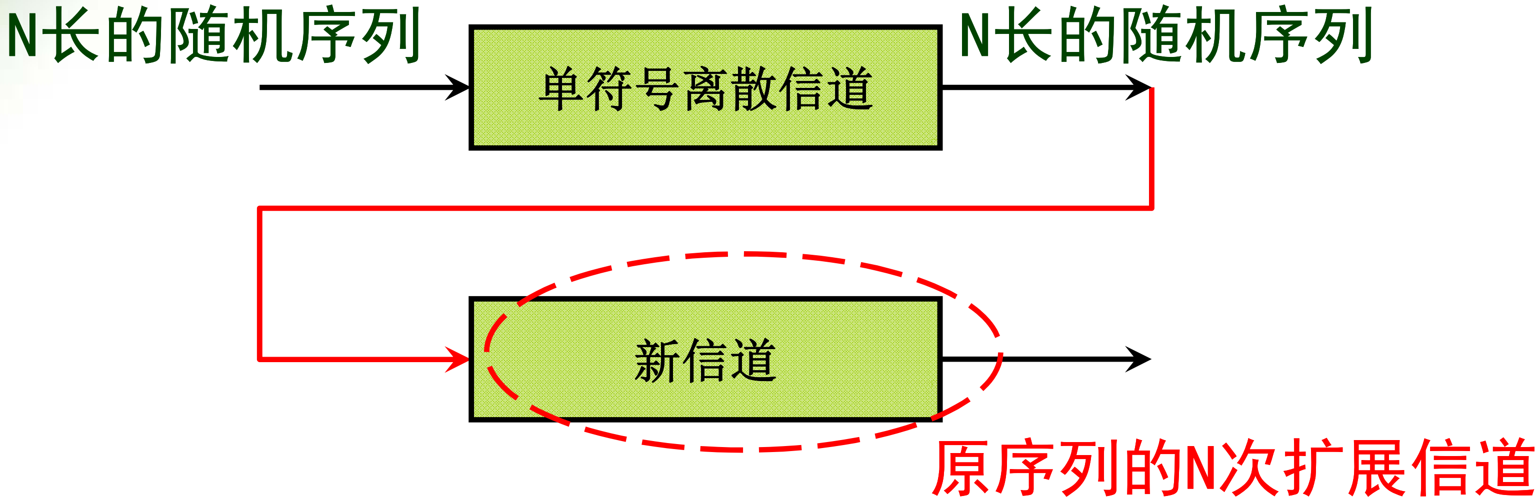
$$= H(X_1) = H(X) \le N H(X_1)$$

$$\sum_{i=1}^{N} I(X_i;Y_i) = \sum_{i=1}^{N} [H(X_i) - H(X_i \mid Y_i)] = \sum_{i=1}^{N} H(X_i) = N H(X)$$

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

§6.4.2 离散无记忆扩展信道及其容量

★ N次扩展信道



§6.4.2 离散无记忆扩展信道及其容量

- N次扩展信道的描述要满足一般的信道数学模型的描述，但符号集为同分布符号的扩展，即各X_i的分布都相同

- 信道的输入和输出集合分别为 $X^N = X_1X_2\cdots X_N$ 和 $Y^N = Y_1Y_2\cdots Y_N$ ，所包含的矢量分别为 $\vec{x}=x_1..x_N$, $\vec{y}=y_1..y_N$

- 信道可通过下式来计算

单符号信道<->无记忆信道

$$\pi_{kl} = P_{Y^N/X^N}(\vec{y}=\beta_l \mid \vec{x}=\alpha_k) = p(\vec{y}\mid\vec{x}) = p(y_1...y_N \mid x_1...x_N) = \prod_{i=1}^{N} p(y_i \mid x_i)$$

§6.4.2 离散无记忆扩展信道及其容量

★ 一个信道的N次扩展信道是N个原信道的Kroneck乘积

例 设输入与输出符号集的尺寸分别为r、s， 则N次扩展信道的输入与输出符号集的尺寸分别为 $R = r^N, S = s^N$

则，信道的转移概率为：

$$\prod = \begin{bmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1S} \\ \pi_{21} & \pi_{22} & \cdots & \pi_{2S} \\ \cdots & \cdots & \cdots & \cdots \\ \pi_{R1} & \pi_{R2} & \cdots & \pi_{RS} \end{bmatrix}$$

满足 $\sum_{l=1}^{s^N} \pi_{kl} = 1$, 对所有 $k = 1, \ldots, r^N$

§6.4.2 离散无记忆扩展信道及其容量

例 6.4.3 求错误概率为p的二元对称信道的二次扩展信道的转移概率矩阵。

解： $\alpha_k \in \{00,01,10,11\}$， $\beta_l = \{00,01,10,11\}$

$$\pi_{kl} = P_{Y^2|X^2}(\vec{y} = \beta_l \mid \vec{x} = \alpha_k) = p(y_1 \mid x_1)p(y_2 \mid x_2)$$

§6.4.2 离散无记忆扩展信道及其容量

2次扩展信道的转移概率矩阵：

$$\Pi = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \otimes \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} = \begin{bmatrix} (1-p)^2 & p(1-p) & p(1-p) & p^2 \\ p(1-p) & (1-p)^2 & p^2 & p(1-p) \\ p(1-p) & p^2 & (1-p)^2 & p(1-p) \\ p^2 & p(1-p) & p(1-p) & (1-p)^2 \end{bmatrix}$$

（1）⊗ ：张量乘积（tensor product）

（2） $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \otimes A = \begin{bmatrix} aA & bA \\ cA & dA \end{bmatrix}$

§6.4.2 离散无记忆扩展信道及其容量

★ 离散无记忆N次扩展信道的容量

$$C^N = \max_{p(\vec{x})}\{I(X^N;Y^N)\} \le \sum_{i=1}^{N}\max_{p(x_i)}\{I(X_i;Y_i)\} \le \sum_{i=1}^{N} C_i$$

$$C_i = \max_{p(x_i)}\{I(X_i;Y_i)\}$$

当信源为无记忆时，等式成立。假定条件是信道无记忆

$$C_i = C \quad \Longrightarrow \quad \boxed{C^N = NC}$$

§6.4.2 离散无记忆扩展信道及其容量

例 6.4.3(续) 求该二次扩展信道的容量。

解： 由例6.2.3可得，错误概率为p的二元对称信道的容量 ，根据式(6.4.6)，该信道的二次扩展信道容量为

$$C^2 = 2C = 2 - 2H(p)$$

§6.4.3 并联信道及其容量

★并联信道的定义

- 由若干并行的单符号子信道组成
- 在每单位时间，发送端都同时通过每个子信道发送不同符号集的消息
- 每子信道的输出仅与该子信道的输入有关

可能的情况是:如果X_i/Y_i与X_j/Y_j完全相同,则可减1个子信道

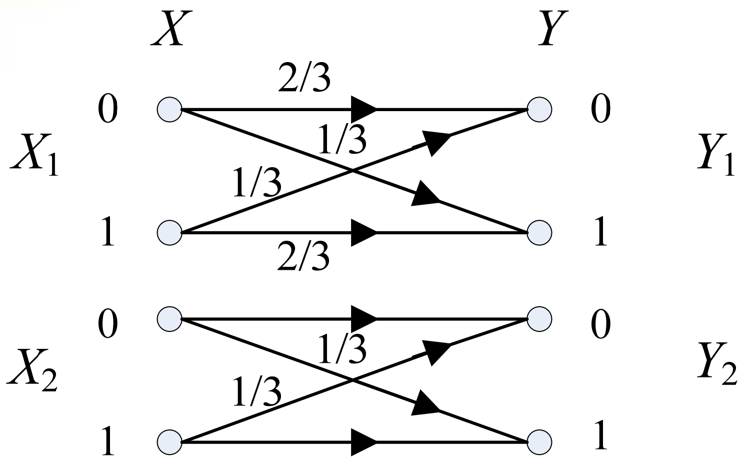
$$X^N = X_1 \ldots X_N \qquad Y^N = Y_1 \ldots Y_N$$

$$C = \max_{p(X_1 \ldots X_N)} \{I(X^N;Y^N)\} \le \max_{p(X_1,\cdots;X_N)} \sum_{i=1}^{N} I(X_i;Y_i) = \sum_{i=1}^{N} C_i$$

- 当各 X_i 独立时，上式等号成立

§6.4.3 并联信道及其容量

例 6.4.4　求下图信道的容量和达到容量时的输入概率分布。



解：　从信道的转移概率图可以看出，两个子信道是独立的，所以构成一个二维并联信道。所求信道容量为

$$C = C_1 + C_2 = 2C_1 = 2[1 - H(\frac{1}{3})] = 0.164 \quad \text{比特/符号}$$

达到容量时，输入 X_1X_2 相互独立，且均为等概率分布。

§6.4.4 和信道及其容量

★和信道的定义

- 一个信道分为若干子信道，且各子信道输入之间互不相交，输出之间也互不相交
- 信道总的输出与输入集合分别为各子信道输出与输入之并集
- 每次传输只能用一个子信道

§6.4.4 和信道及其容量

★定理6.4.3

对于和信道，信道容量为 $C=\log_2\sum 2^{C_i}$ 比特/符号，

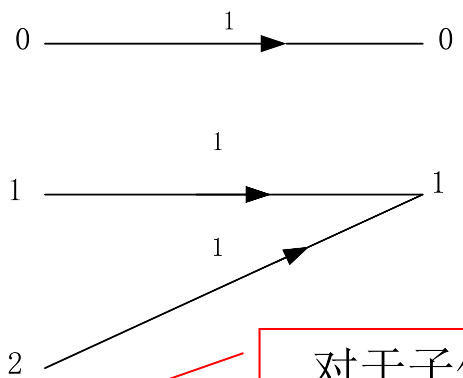
其中 Ci 为每个子信道的容量，第i个子信道使用概率

为 $r_i=2^{C_i-C}=\dfrac{2^{C_i}}{\sum_{i=1}^{N}2^{C_i}}$ ，达到容量的输入概率为各子信

道达到容量时的概率再乘以 r_i 。

§6.4.4 和信道及其容量

例 6.4.5 一信道的转移概率如图所示，求信道容量和达到容量时的输入概率。



对于子信道1，C=H(Y)-H(X|Y)=0-0=0
各子信道容量为0，但和信道容量1

解：

$$C_1 = C_2 = 0, r_1 = \frac{1}{2}, r_2 = \frac{1}{2}$$

$$C = \log_2 2 = 1bit \qquad p_0 = \frac{1}{2}, p_1 = \frac{p}{2}, p_2 = \frac{1-p}{2}$$

子信道2为确定信道
C=log(S)=log1=0

其中***p***为不大于**1**的正数

关于信道容量的注释

★ 信道达到容量时的输入概率分布不一定唯一。（对于对称信道是唯一的）

★ 达到容量时的输出概率严格为正。

★ 对应于信道容量的输出概率是唯一的。

★ 对于具有任意转移概率分布的离散信道，其容量要利用迭代算法进行计算（略，课后自己阅）。

本 章 小 结

1. 平稳离散无记忆信道模型： $\{X, p(y|x), Y\}$

2. 平稳离散无记忆信道的容量： $C \equiv \max_{p(x)} I(X;Y)$

3. 特殊离散无记忆信道的容量的计算

★对称信道：输入等概率时达到容量，且

$$C = H(Y) - H(p_1, p_2, \ldots, p_s)$$

★一般离散信道

[P]有逆阵时 $C = \log \sum_j 2^{\beta_j}$ $[\beta] = -[P]^{-1}[h]$

本 章 小 结

对于离散无记忆信道： $I(a_i;Y)=C$，对于$p_i>0$

$I(a_i;Y)\le C$，对于$p_i=0$

★ 和信道 $C=\log_2\sum 2^{C_i}$ 达到容量的输入概率为各子信道达到容量时的概率再乘以 r_i

★ 并联信道 $C=\sum_{i=1}^{N}C_i$

★ 离散平稳无记忆N次扩展信道

$C^N=NC$ 当信源无记忆时信道达到容量

★ 级联信道：转移概率矩阵为各级联信道矩阵的乘积，再计算容量。

课后习题

P.126

6.1, 6.4, 6.6, 6.13