目录

面试常考的点：信息的概念、自信息、互信息、各种熵；信源（连续、离散、马尔科夫等）；信道（信道容量、信道编码）；香农三个定理（内容、解决了什么问题，我们只学了两个）

第一章　绪论

1.1 基本概念

一、信息的定义

1、香农对信息的定义："信息"就是消除不定性的东西

2、香农对通信过程的解释

- 收信者在收到消息之前是不知道消息的具体内容的
- 对于收信者来说，通信过程是消除事物状态的不确定性的过程。消除了不确定性，就获得了信息，原先的不确定性消除的越多，获得的信息就越多；

3、信息、消息、信号的辨析

- 信息：抽象的，非物理性的
- 消息：具体的，非物理性的（不是信息本身）
- 信号：具体的，物理性的（可在信道和自然界中传输）
- 信息传输过程：信息→消息→信号

二、信息的性质（参考）

1、存在的普遍性

2、有序性：信息可以用来消除系统的不确定性，增加系统的有序性。

3、相对性：同一个事物，不同的观察者获得的信息量可能不同。

4、可度量性：可以定量描述，通常用信息量表示。

5、可扩充性

6、可存储、传输与携带性

7、可压缩性

8、可替代性

9、可扩散性

10、可共享性

11、时效性：信息是有"寿命"的。

三、信息论的研究内容——研究信息传输和处理问题

1、信息论基础（香农信息论、狭义信息论）

a)　研究方向：从信源到信宿的全过程，收发两端联合优化问题

b)　具体内容：研究信息测度、信道容量、信源和信道编码理论、信息率失真函数

2、一般信息论（工程信息论）

a)　研究方向：重点是接收端，如何在接收端把消息从干扰中提取出来

b)　具体内容

- 香农信息论（上面那个）
- 维纳理论：噪声理论、信号滤波和预测、统计检测和估值理论、调制理论
- 其他：信息处理理论、保密理论

3、广义信息论

- 是一门综合的新兴学科，至今尚没有严格的定义。概括地说：凡是能用广义通信系统模型描述的过程和系统，都能用信息基本理论来研究。
- 不仅包括一般信息论所研究的内容，还包括医学、生物学、遗传学等有关信息的问题。

第二章 信源和信源熵

2.1 信息量与信息熵

一、信息量（自信息）

1、直观定义

$$\left.\begin{array}{l}\text{必然发生的事发生了}\Rightarrow\text{几乎没有带来新的信息}\\ \text{不可能发生的事发生了}\Rightarrow\text{带来了非常多新的信息}\end{array}\right\}\Rightarrow\text{信息量 }\boldsymbol{I}\propto\frac{\mathbf{1}}{\boldsymbol{P}}$$

2、数学定义

a) 公式：$I(x)=\log_a\frac{1}{p(x)}=-\log_a p(x)$

b) 单位及彼此间的换算关系

- a =2：比特(bit)
- a =e：奈特(nat)
- a =10：哈特莱（Hartley）

$1\text{net}=\log_2 e\approx 1.433\text{bit}$

$1\text{hart}=\log_2 10\approx 3.322\text{bit}$

3、信息量的特点

- 信息量和出现的概率成反比
- 必然事件的信息量为 0
- 可加性：两独立事件的联合信息量等于各自信息量之和。如以下说法信息量相同
  - 说法 1：办公室在"53 号"房间。
  - 说法 2，办公室在"5 楼"左数"第 3 个"房间。

二、信息熵

1、定义：信源所有符号自信息量的统计平均值

$$\boldsymbol{H}(\boldsymbol{X})=\boldsymbol{E}[\text{自信息}]$$

- 表示信源输出后每个消息/符号所提供的平均信息量
- 表示信源输出前，信源的平均不确定性
- 表征信源输出的随机性

2.2 离散信源与信源熵

一、离散信源

1、定义：信源输出的消息数是有限的或可数的

2、离散信源的表示

a) 单符号离散信源：每次只输出 1 个符号

$$\begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \cdots & a_q \\ p(a_1) & p(a_2) & \cdots & p(a_q) \end{bmatrix} \quad 且\sum_{i=1}^{q} p(a_i) = 1$$

b) N 次扩展信源：每次输出 N 个符号（和 N 维离散信源一个意思）

$$\begin{bmatrix} X^N \\ p(x) \end{bmatrix} = \begin{bmatrix} (a_1 a_1 \cdots a_1) & (a_1 a_1 \cdots a_2) & \cdots & (a_q a_q \cdots a_q) \\ p(a_1 a_1 \cdots a_1) & p(a_1 a_1 \cdots a_2) & \cdots & p(a_q a_q \cdots a_q) \end{bmatrix}$$

- 码字的种类（样本空间）：共有q^N 种元素
- 二元二次扩展信源：10 01 10 11 11 00 10
- 十元四次扩展信源：1901 3760 2181 9810 1234

二、离散信源熵的基本特性和定理

(1)对称性：$H(p) = H(p_1, p_2, \cdots, p_q) = H(p_2, p_3, \cdots, p_q, p_1) = \cdots$

$\Rightarrow$熵面对随机变量的总体概率分布,不能描述事物本身的具体含义

(2)确定性: $H(1,0) = H(1,0,0) = H(1,0,0,0) = \cdots = H(1,0,\cdots,0) = 0$

(3)非负性: $H(X) \geq 0$

(4)扩展性: $\lim_{\varepsilon \to 0} H_{q+1}(p_1, p_2, \cdots, (p_q - \varepsilon), \varepsilon) = H_q(p_1, p_2, \cdots, p_q)$

$\Rightarrow$增加有限个概率甚小的随机变量，熵不变

$\Rightarrow$虽然，概率很小的事件信息量较大。但从总体来考虑，
因为发生的概率小，所以它在熵的计算中占的比重很小

(5)强可加性：$H(XY) = H(X) + H(Y/X) = H(Y) + H(X/Y)$

若X与Y独立，则$H(XY) = H(X) + H(Y)$

推广——链规则：$H(X_1 X_2 \cdots X_N) = H(X_1) + H(X_2/X_1) + H(X_3/X_1 X_2) + \cdots + H(X_N/X_1 X_2 \cdots X_{N-1})$

(6)递增性：若原信源X中有一个符号元素划分成m 个符号，
且它们的概率之和等于原符号元素的概率，则新信源的熵增加

$$\begin{aligned} & H_{n+m-1}(p_1, p_2, \cdots, p_{n-1}, \underbrace{q_1, q_2, \cdots, q_m}_{\sum_{j=1}^{m} q_j = p_n}) \\ & = H_n(p_1, p_2, \cdots, p_{n-1}, p_n) + \underbrace{p_n H_m\left(\frac{q_1}{p_n}, \frac{q_2}{p_n}, \cdots, \frac{q_m}{p_n}\right)}_{增加量} \end{aligned}$$

(7)极值性：当 $p_1 = p_2 = \cdots = p_q = \frac{1}{q}$ 时，$H(X) = H_{\max}(X) = \log_2 q$

$\Rightarrow$离散信源各符号等概率出现时熵最大

三、几种离散信源的信息熵

1、单符号离散信源的信息熵

$$H(X)=-\sum_{i=1}^{q}P(a_i)\left[\log_2 P(a_i)\right](bit/\text{符号})$$

2、无记忆的 N 次扩展信源的信息熵

$$H(X^N)=NH(X)$$

3、N 维离散信源的联合信息熵（以二维为例）

a) 定义：$H(X_1X_2)=-\sum_{i=1}^{q}\sum_{j=1}^{q}p(a_ia_j)\log_2 p(a_ia_j)$

b) 如果还平稳，可以求一维的信源熵

$$H(X)=\frac{1}{2}H(X_1X_2)\qquad \text{其中：}H(X_1X_2)=H(X_1)+H(X_2/X_1)$$

2.3 离散平稳信源与熵

一、离散平稳信源的概念

1、完全平稳信源

用随机矢量$\boldsymbol{X}=\left(X_1,X_2,\cdots,X_i,\cdots\right)$描述信源发出的消息，其中$X_i$表示$t=i$时发出的符号

(1)一维平稳信源：$p\left(x_i\right)=p\left(x_j\right)\ \left(i\neq j\right)$ 即信源的各符号等概

(2)二维平稳信源：$\begin{cases} p\left(x_i\right)=p\left(x_j\right) & \left(i\neq j\right) \\ p\left(x_i,x_{i+1}\right)=p\left(x_j,x_{j+1}\right) & \left(i\neq j\right) \end{cases}$

(3)完全平稳信源：$\begin{cases} p\left(x_i\right)=p\left(x_j\right) & \left(i\neq j\right) \\ p\left(x_i,x_{i+1}\right)=p\left(x_j,x_{j+1}\right) & \left(i\neq j\right) \\ \cdots \\ p\left(x_i,x_{i+1},\cdots,x_{i+N}\right)=p\left(x_j,x_{j+1},\cdots,x_{j+N}\right) & \left(i\neq j\right) \end{cases}$

2、离散平稳信源:

a) 定义：各维联合概率分布均与时间起点无关的完全平稳离散信源

b) 另一充要条件：各维条件概率与时间起点无关，只与关联长度有关

- 联合概率与条件概率的关系

$$p\left(x_i,x_{i+1},\cdots,x_{i+N}\right)=p\left(x_i\right)p\left(x_{i+1}/x_i\right)\cdots p\left(x_{i+N}/x_ix_{i+1}\cdots x_{i+N-1}\right)$$

二、离散平稳信源的熵

1、二维联合熵：$H\left(X_1X_2\right)=-\sum_{i=1}^{q}\sum_{j=1}^{q}p\left(a_ia_j\right)\log_2 p\left(a_ia_j\right)$

a) 性质

$$H\left(X_1X_2\right)=H\left(X_1\right)+H\left(X_2/X_1\right)\leq H\left(X_1\right)+H\left(X_2\right)=2H\left(X\right)=H\left(X^2\right)$$

- 若二维信源前后符号不相关，X1X2 就是无记忆的二次扩展信源
- 若前后符号相关，则输出两个符号的联合熵总是小于信源熵的 2 倍

b) 相关公式——香农不等式：$H\left(X_2/X_1\right)\leq H\left(X_2\right)$

2、二维平稳信源的信息熵：$H\left(X\right)=\frac{1}{2}H\left(X_1X_2\right)$

3、平均符号熵：表示信源平均每发一个符号所提供的平均信息量

$$H_N\left(X\right)=\frac{1}{N}H\left(X_1X_2\cdots X_N\right)$$

4、离散平稳信源的熵的四个性质

条件：$H\left(X_1\right)<\infty$

(1)条件熵的非递增性：$H\left(X_N/X_1X_2\cdots X_{N-1}\right)\leq H\left(X_N/X_2\cdots X_{N-1}\right)$

(2)平均符号熵的非递增性：$H_N\left(X\right)\leq H_{N-1}\left(X\right)$

(3)条件熵与平均符号熵：$H\left(X_N/X_1X_2\cdots X_{N-1}\right)\leq H_N\left(X\right)$

(4)极限熵：$H_\infty=\lim_{N\to\infty}H_N\left(X\right)$存在，并且$H_\infty=\lim_{N\to\infty}H_N\left(X\right)=\lim_{N\to\infty}H\left(X_N/X_1X_2\cdots X_{N-1}\right)$

2.4 马尔科夫信源

一、基本概念

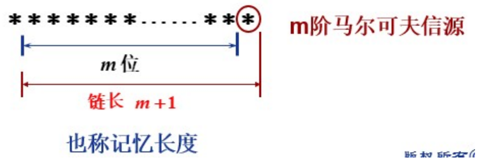
1、离散记忆信源与离散无记忆信源

a) 离散无记忆信源：信源发出的符号是彼此统计独立的，即

$$p(\boldsymbol{X})=\prod_{i=1}^{N}p\left(x_i=a_{k_i}\right)\qquad k_i=1,2,\cdots,q$$

b) 离散记忆信源：反之

2、m 阶马尔科夫信源：当前的符号与之前 m 个符号有关



- 时齐马尔可夫信源：条件概率与时间与起点 i 无关
- 马尔科夫链：当前符号加上之前 m 个符号，链长为 m+1

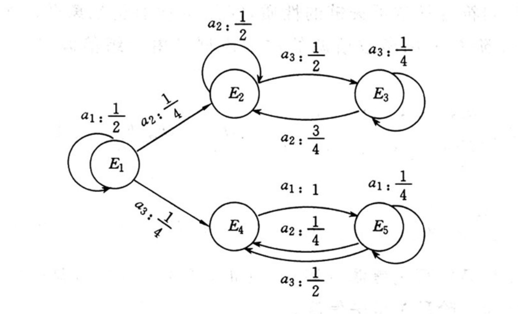
二、马尔科夫信源

1、马尔科夫信源的特征

$$\begin{cases}\text{信源状态}\boldsymbol{S}\in\boldsymbol{E}=\left[\boldsymbol{E}_1,\boldsymbol{E}_2,\cdots,\boldsymbol{E}_J\right]\\ \text{可能输出的符号}\boldsymbol{X}\in\boldsymbol{A}=\left[\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_q\right]\end{cases}\qquad\begin{cases}\text{信源所处的状态序列：}\ \boldsymbol{s}_1\boldsymbol{s}_2\cdots\boldsymbol{s}_{l-1}\boldsymbol{s}_l\cdots\\ \text{信源输出的符号：}\ \boldsymbol{x}_1\boldsymbol{x}_2\cdots\boldsymbol{x}_{l-1}\boldsymbol{x}_l\cdots\end{cases}$$

- 某一时刻信源输出符号只与此刻信源所处的状态有关，与以前的状态及以前的符号无关
- 此刻状态仅由前一状态和当前输出符号惟一确定

2、状态转移图与状态转移矩阵



状态转移矩阵：

$$\boldsymbol{P}=\left[p\left(E_j/E_i\right)\right]=\begin{array}{c|c} & E_1\quad E_2\quad E_3\quad E_4\quad E_5\\ \begin{matrix}E_1\\E_2\\E_3\\E_4\\E_5\end{matrix} & \begin{bmatrix}1/2 & 1/4 & 0 & 1/4 & 0\\ 0 & 1/2 & 1/2 & 0 & 0\\ 0 & 3/4 & 1/4 & 0 & 0\\ 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 3/4 & 1/4\end{bmatrix}\end{array}$$

$\boldsymbol{E_j}$ $\boldsymbol{E_i}$ $\boldsymbol{p\left(E_j/E_i\right)}$

三、马尔科夫信源的熵

- 讨论前提：时齐、平稳、遍历的马尔可夫信源

1、考虑哪些

联合熵：$\boldsymbol{H}\left(\boldsymbol{X}_1\boldsymbol{X}_2\cdots\boldsymbol{X}_N\right)$

平均符号熵：$\boldsymbol{H}_N\left(\boldsymbol{X}\right)=\dfrac{1}{N}\boldsymbol{H}\left(\boldsymbol{X}_1\boldsymbol{X}_2\cdots\boldsymbol{X}_N\right)$

极限熵：$\boldsymbol{H}_\infty=\lim\limits_{N\to\infty}\boldsymbol{H}_N\left(\boldsymbol{X}\right)=\lim\limits_{N\to\infty}\boldsymbol{H}\left(\boldsymbol{X}_N/\boldsymbol{X}_1\boldsymbol{X}_2\cdots\boldsymbol{X}_{N-1}\right)$

2、极限熵的求法（也即信源熵）：$\boldsymbol{H}_\infty=\boldsymbol{H}_{m+1}=\sum\limits_{i=1}^{J}\boldsymbol{p}\left(\boldsymbol{E}_i\right)\boldsymbol{H}\left(\boldsymbol{X}/\boldsymbol{E}_i\right)$

2.5 信源的相关度和剩余度

一、信源的相关性

1、几个熵

符号独立等概：$H_0 = \log m$

独立信源：$H_1 = H(X_1)$

一阶马氏信源：$H_2 = H(X_2/X_1)$

$\mathbf{N-1}$阶马氏信源：$H_N = H(X_N/X_1X_2\cdots X_{N-1})$

极限熵：$H_\infty = \lim\limits_{N\to\infty} H_N(X) = \lim\limits_{N\to\infty} H(X_N/X_1X_2\cdots X_{N-1})$

- 关系：$H_0 \ge H_1 \ge H_2 \ge \ldots \ge H_\infty$

2、熵的相对率：$\eta = \dfrac{H_\infty}{H_0}$　　　H_0又称为最大熵

3、信源剩余度：$\gamma = 1-\eta$

- 对于剩余度大的信源可进行压缩，从而提高传输效率
- 但增大信源剩余度，能提高传输中的抗干扰能力

2.6 连续信源

一、连续信源的概念

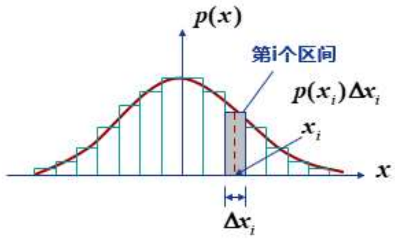
1、定义：信源输出消息的取值是连续的

2、单符号连续信源的表示

$$\begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} (a,b) \\ p(x) \end{bmatrix} 或 \begin{bmatrix} \mathbf{R} \\ p(x) \end{bmatrix} \qquad 且\int_a^b p(x)dx = 1 或 \int_R p(x)dx = 1$$

二、连续信源的熵

1、推导



$$由图可得 \boldsymbol{H(x) = -\sum_{i=-N}^{N} p(x_i)\Delta x_i \log\left[p(x_i)\Delta x_i\right]}$$

$$\boldsymbol{\Rightarrow \lim_{\substack{N\to\infty \\ \Delta x_i \to 0}} H(x) = \lim_{\substack{N\to\infty \\ \Delta x_i \to 0}} \sum_{i=-N}^{N} -p(x_i)\Delta x_i \log\left[p(x_i)\Delta x_i\right]}$$

$$\boldsymbol{= \lim_{\substack{N\to\infty \\ \Delta x_i \to 0}} \sum_{i=-N}^{N} -p(x_i)\log p(x_i) \bullet \Delta x_i + \lim_{\substack{N\to\infty \\ \Delta x_i \to 0}} -\sum_{i=-N}^{N} p(x_i)\Delta x_i \log \Delta x_i}$$

$$\boldsymbol{= \underbrace{-\int_{-\infty}^{\infty} p(x)\log p(x)dx}_{H_c(x)\ 相对熵} + \underbrace{\lim_{\Delta x_i \to 0}\left[-\log \Delta x_i\right]}_{无限大常数项（不计）}}$$

- 按照离散熵的概念，连续随机变量的熵应为无穷大，失去意义
- 香农定义取有限值的项为连续信源的信息熵，也称微分熵或差熵

2、连续信源不确定程度的相对度量——相对熵/差熵/微分熵

a) 表达式：$\boldsymbol{H_c(x) = -\int_{-\infty}^{\infty} p(x)\log p(x)dx}$

b) 和离散信源熵的区别

- 它是连续信源的熵，而不是连续信源输出的信息量，输出的信息量是整个 H(x)
- 离散信源输出信息量就是信源的熵

三、几种特殊的连续信源的熵

1、均匀分布

$$\boldsymbol{p(x) = \begin{cases} \dfrac{1}{b-a}, a \le x \le b \\ 0, \quad 其他 \end{cases} \Rightarrow H_c(x) = \log_2(b-a)}$$

$$\boldsymbol{(1)(b-a) > 1:\ H_c(x) > 0}$$

$$\boldsymbol{(2)(b-a) = 1:\ H_c(x) = 0}$$

$$\boldsymbol{(3)(b-a) < 1:\ H_c(x) < 0}$$

- 注意!!!：相对熵不具有非负性

2、高斯分布

$$\boldsymbol{p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-m)^2}{2\sigma^2}} \Rightarrow H_c(X) = \log\sqrt{2\pi e}\cdot\sigma}$$

3、指数分布

$$p(x)=\frac{1}{m}e^{-\frac{x}{m}}，\ (x\geq 0)\Rightarrow H_c(x)=\log_2 me$$

- 指数分布的信源的熵取决于均值 m

四、最大连续熵定理

1、限峰值功率的最大熵定理

若信源输出的幅度被限定在[a,b]区域内，当输出信号的概率密度函数服从均匀分布时，信源具有最大熵，为 log(b-a)

2、限平均功率的最大熵定理：此时高斯分布的信源熵最大

3、均值受限条件下的最大连续熵定理：此时指数分布的信源熵最大

第三章　信道和信道容量

3.1 信道的数学模型和分类

一、基本概念

1、信道的作用

- 在信息系统中信道主要用于传输与存储信息
- 在通信系统中信道主要用于传输信息

2、研究信道的目的：主要是为了描述、分析不同类型信道的特性，度量信道的极限传输能力，即信道容量

3、信道的分类

a) 根据用户数分类：单用户信道、多用户信道
b) 根据信道的统计特性分类：恒参信道、随参信道
c) 根据信道的记忆特性分类：无记忆信道、有记忆信道（当前输出符号是否与过去输入符号有关）

- 离散无记忆信道的充要条件

$$p(y/x)=p(y_1y_2\cdots y_N/x_1x_2\cdots x_N)=p(y_1/x_1)p(y_2/x_2)\cdots p(y_N/x_N)$$

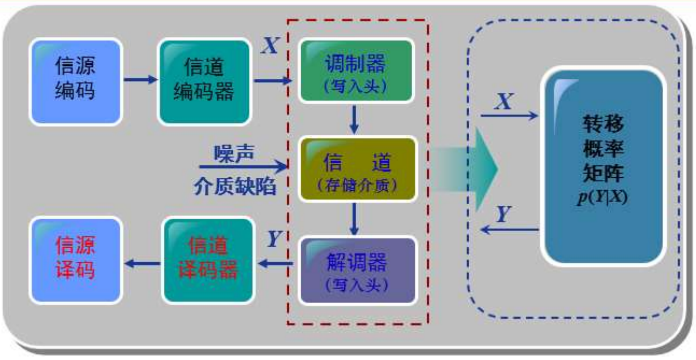
- 离散记忆信道的两种处理方法
  - 转化成无记忆信道：将记忆较强的 N 个符号当作一个矢量符号处理，而各矢量符号之间认为是无记忆的
  - 看成马尔可夫链

d) 根据输入和输出信号的特点：　见下表

	输出信号连续	输出信号离散
输入信号连续	连续信道	半离散（或半连续）信道
输入信号离散	半离散（或半连续）信道	离散信道（数字信道）
时间还连续的连续信道是波形信道		

二、离散信道的数学模型

1、通信系统的模型



- 把信道编、解码器之间的所有部件看成是一个“黑箱”，问题归结为输入、输出和转移概率矩阵三个要素。

2、离散信道的一般的数学模型

a)　数学模型

- 也可用概率空间表示：$\{X,p(y/x),Y\}$

b)　离散信道按信道转移概率分类

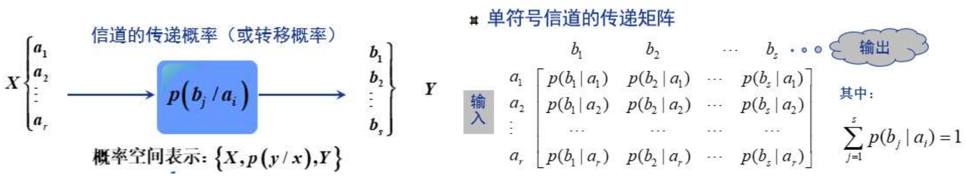
- 无干扰（无噪）信道：信道的输入和输出符号间有确定的一一对应关系

$$p(y/x)=\begin{cases}1 & y=f(x)\\ 0 & y\neq f(x)\end{cases}$$

- 有干扰信道：存在噪声和码间干扰，导致不存在…

3、单符号离散信道的数学模型

a)　数学模型及信道传递矩阵



b)　概率关系

输入输出的联合概率：$p(a_ib_j)=p(a_i)p(b_j/a_i)=p(b_j)p(a_i/b_j)$

(1) $p(a_i)$：输入符号的先验概率

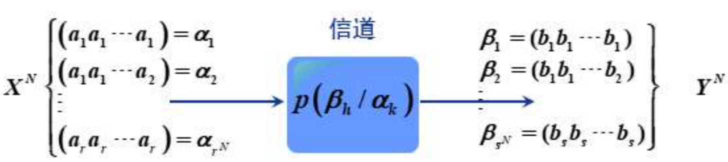
(2) $p(a_i/b_j)$：输入符号的后验概率

(3) $p(b_j/a_i)$：信道传递概率（前向概率）

(4) $p(b_j)=\sum_{i=1}^{r}p(a_i)p(b_j/a_i)$

4、多符号离散信道的数学模型（N 次扩展信道）

概念：N 维离散扩展信道的输入和输出都是长为 N 的消息序列



3.2 信道疑义度和平均互信息

一、信道疑义度

1、信道疑义度的引入

a) 回顾信源熵（先验熵）: $H(X)=\sum_{i=1}^{r}p(a_i)\log\frac{1}{p(a_i)}$

- 表示接收到 Y 以前关于输入变量 X 的先验不确定性的度量

b) 接收到 Y 后对于信源的不确定性

- 在理想信道下：无噪声，无码间干扰，输入输出一一对应。接收符号后就消除了对发送信号的先验不确定性
- 有噪声信道下：怎样度量接收到 Y 后关于 X 的不确定性？

2、信道疑义度的定义（后验熵、损失熵）

a) 定义

$$\begin{aligned}H(X/Y)&=E\left[H(X/b_j)\right]\\&=\sum_{j=1}^{s}p(b_j)H(X/b_j)\\&=\sum_{j=1}^{s}p(b_j)\sum_{i=1}^{r}p(a_i/b_j)\log\frac{1}{p(a_i/b_j)}=\sum_{j=1}^{s}\sum_{i=1}^{r}p(a_ib_j)\log\frac{1}{p(a_i/b_j)}\\&=\sum_{X,Y}p(xy)\log\frac{1}{p(x/y)}\end{aligned}$$

b) 含义

- 输入符号在有噪声、有干扰信道中传输丢失的信息量
- 当 X 和 Y 统计独立时，H(X|Y)=H(X)，信道上没能传送任何信息（损失全部信息）

二、平均互信息

1、定义：

a) 公式

$$\begin{aligned}I(X;Y)&=H(X)-H(X/Y)\\&=\sum_{X,Y}p(xy)\log\frac{p(x/y)}{p(x)}=\sum_{X,Y}p(xy)\log\frac{p(xy)}{p(x)p(y)}=\sum_{X,Y}p(xy)\log\frac{p(y/x)}{p(y)}\end{aligned}$$

其中：$p(xy)=p(x)p(y/x)=p(y)p(x/y)$

注意大小写的区别：$I(X;Y)\ =E\left[I(x;y)\right]$

b) 含义：输出端每收到一个符号，获得了多少关于信源的信息量

c) 平均互信息的基本特性

(1) 非负性：$I(X;Y)\geq 0$　　仅当 X 与 Y 独立等号才成立

注意：互信息 $I(x;y)$ 可以取负值

推论：离散信源通过有噪声信道传输，接收端总会收到一定的信息

(2)极值性：$I(X;Y)\leq H(X)$　　仅当理想信道等号成立

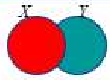
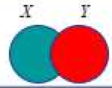
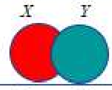
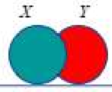
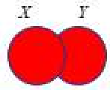
(3)对称性**(**交互性**)**：$I(X;Y)=I(Y;X)$

(4)熵的维拉关系

(5)凸函数性

2、维拉关系——平均互信息与其他熵的关系

a) 关系

名称	符号	关 系	图 示
无条件熵	$H(X)$	$H(X)\ge H(X/Y)$ $=H(X/Y)+I(X;Y)$ $H(X)=H(XY)-H(Y/X)$	
	$H(Y)$	$H(Y)\ge H(Y/X)$ $=H(Y/X)+I(X;Y)$ $H(Y)=H(XY)-H(X/Y)$	
条件熵	$H(X/Y)$	$H(X/Y)=H(XY)-H(Y)$ $=H(X)-I(X;Y)$	
条件熵	$H(Y/X)$	$H(Y/X)=H(XY)-H(X)$ $=H(Y)-I(X;Y)$	
联合熵	$H(XY)=H(YX)$	$H(XY)=H(X)+H(Y/X)$ $=H(Y)+H(X/Y)$ $=H(X)+H(Y)-I(X;Y)$ $=H(X/Y)+H(Y/X)+I(X;Y)$	
互信息	$I(X;Y)=I(Y;X)$	$I(X;Y)=H(X)-H(X/Y)$ $=H(Y)-H(Y/X)$ $=H(XY)-H(Y/X)-H(X/Y)$ $=H(X)+H(Y)-H(XY)$	 版权所有©赵蓉

b) 各个熵的含义

- H(X)：表示信源侧每个符号的平均信息量（信源熵）；
- H(Y)：表示信宿侧每个符号的平均信息量（信宿熵）；
- H(X|Y)：信道疑义度(损失熵)，表示在输出端接收到 Y 后，发送端 X 尚存的平均不确定性。这个对 X 尚存的不确定性是由于干扰引起的。
- H(Y|X)：信道散布度（噪声熵），表示已知 X 后，对于输出 Y 尚存的平均不确定性
- H(XY)：表示整个信息传输系统的平均不确定性

3、关于平均互信息的两个定理

(1) 对于固定信道，$I(X;Y)$是输入信源的概率$p(x)$ 的∩ 型凸函数——信道容量的基础

意义：对于某一个固定信道，一定存在一种信源使输出端获得的平均信息量最大

(2) 对于固定信源，$I(X;Y)$是信道传递概率$p(y/x)$ 的∪ 型凸函数——率失真函数的基础

意义：对每一种信源都存在一种最差的信道，此信道的干扰（噪声）最大，使输出端获得的信息量最小

3.3 离散信道的信道容量

一、信道容量的定义

1、信息传输率与信息传输速率

信息传输率：$R = I(X;Y) = H(X) - H(X/Y)(bit/符号)$

信息传输速率：$R_t = \frac{1}{t} I(X;Y) = \frac{1}{t} H(X) - \frac{1}{t} H(X/Y)(bit/s)$

其中t为传输一个符号所需的时间

2、信道容量

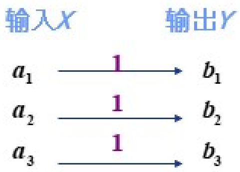
最大信息传输率：$C = \max\limits_{p(x)}[I(X;Y)](bit/符号)$

最大信息传输速率：$C_t = \frac{1}{t} C = \frac{1}{t} \max\limits_{p(x)}[I(X;Y)](bit/s)$

其中$p(x)$为使之取到最大值时的最佳分布

二、离散无噪信道的信道容量

1、无噪无损信道

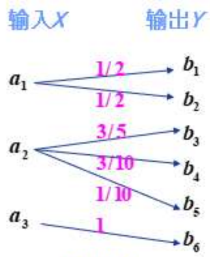


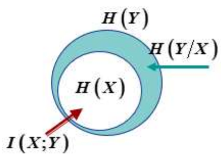
a) 信道特点：输入输出一一对应

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

b) 信道容量：$C = \max\limits_{p(x)}[I(X;Y)] = \max\limits_{p(x)} H(X) = \log r$

2、有噪无损信道



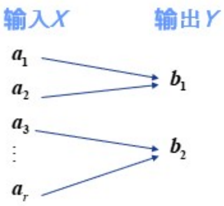


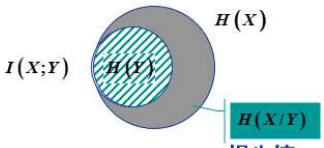
a) 信道特点：输入一个 X 值对应多个输出 Y 值

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

b) 信道容量：$C = \max\limits_{p(x)}[I(X;Y)] = \max\limits_{p(x)} H(X) = \log r$

3、无噪有损信道





a) 信道特点：r>s

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

b) 信道容量：$C = \max\limits_{p(x)}[I(X;Y)] = \max\limits_{p(x)} H(Y) = \log s$

4、有噪有损信道

三、对称离散信道的信道容量

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

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

$P = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.1 & 0.7 \end{bmatrix}$ 不对称

1、一般对称信道和信道容量

a) 对称条件：信道矩阵同时行对称和列对称

- 行对称：每一行元素都是另一行元素的一个排列（列对称同理）

b) 信道容量

$$C = \max_{p(x)}[I(X;Y)] = \max_{p(x)}[H(Y) - H(p_1', p_2', \cdots, p_s')]$$
$$= \log s - H(p_1', p_2', \cdots, p_s')(bit/符号)$$

- 对应的最佳分布是等概的情况

- 其中的 s 表示输出的符号数

2、强对称信道及其信道容量

a) 条件：输入符号和输出符号的个数相等的对称信道

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

b) 信道容量

$$C=\log r-H(\bar{p},\frac{p}{r-1},\frac{p}{r-1},\cdots,\frac{p}{r-1})=\log r-p\log(r-1)-H(p)$$

3、准对称信道

a) 定义：信道转移概率矩阵 P 可按输出符号集 Y 分成几个子集（子矩阵），而每一子集关于行、列都对称

$$P=\begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.2 & 0.1 & 0.7 \end{bmatrix}$$ 可以分成 $$P_1=\begin{bmatrix} 0.7 & 0.2 \\ 0.2 & 0.7 \end{bmatrix} \quad P_2=\begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}$$

b) 信道容量

$$C=\log r-H(p_1',p_2',\cdots,p_s')-\sum_{k=1}^{n}N_k\log M_k$$

N_k是第k个子矩阵中行元素之和

M_k是第k个子矩阵中列元素之和

四、一般离散信道的信道容量

1、定理

$C=\max\limits_{p(x)}\left[I(X;Y)\right]$取到极大值的充要条件是输入概率分布$p(a_i)$ 满足

$$\begin{cases} I(a_i;Y)=\mathrm{C} & for\ a_i\ whose\ p(a_i)\neq 0 \\ I(a_i;Y)\leq \mathrm{C} & for\ a_i\ whose\ p(a_i)=0 \end{cases}$$

其中：$I(x_i=a_i;Y)=\sum_{j=1}^{s}P(b_j/a_i)\log\frac{P(b_j/a_i)}{P(b_j)}$是输出端收到$Y$ 后获得$x=a_i$ 的信息量

2、对于不能变成一一对应的形式的离散信到的信道容量的求法

Step1：解方程求β_j（$r=s$时有解）。$\sum_{j=1}^{s}p(b_j|a_i)\beta_j=\sum_{j=1}^{s}p(b_j|a_i)\log p(b_j|a_i)$

Step2：求信道容量。$C=\log\sum_{j=1}^{s}2^{\beta_j}\quad(bit/符号)$

Step3：求对应的最佳概率分布。$p(b_j)=2^{\beta_j-C}\Rightarrow p(b_j)=\sum_{i=1}^{r}p(a_i)p(b_j/a_i)$

五、离散无记忆 N 次扩展信道的信道容量

1、相关定理

已知信道的输入序列为$X=(X_1X_2\cdots X_N)$,接收到的序列为$Y=(Y_1Y_2\cdots Y_N)$,

信道传递概率为$p(y\,|\,x)=\prod_{i=1}^{N}p(y_i\,|\,x_i)$，则

(1)信道无记忆$\Rightarrow I(X;Y)=I(X^N;Y^N)\le\sum_{i=1}^{N}I(X_i;Y_i)$

(2)信源无记忆$\Rightarrow I(X;Y)=I(X^N;Y^N)\ge\sum_{i=1}^{N}I(X_i;Y_i)$

(3)二者均无记忆$\Rightarrow I(X;Y)=I(X^N;Y^N)=\sum_{i=1}^{N}I(X_i;Y_i)$

2、离散无记忆 N 次扩展信道的信道容量

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

3、特殊情况：*Xi*取自同一信源符号集且具有同一种分布，*Yi*同理

则有$I(X_1;Y_1)=I(X_2;Y_2)=\cdots=I(X_N;Y_N)=I(X;Y)$

$\Rightarrow I\left(\mathrm{X;Y}\right)=I(X^N;Y^N)=NI(X;Y)$

$\Rightarrow C^N=NC$

六、组合信道的信道容量

1、独立并联信道的信道容量

a) 独立并联信道的概念：输入符号集、输出符号集、传递概率互不相关

- 每个信道的输出只与本信道的输入有关，而与其他信道的输入、输出无关

b) 联合平均互信息

$$I(X_1X_2\cdots X_N;Y_1Y_2\cdots Y_N)\le\sum_{i=1}^{N}I(X_i;Y_i)$$

- 联合平均互信息不大于各信道的平均互信息之和

c) 信道容量

一般：$C_{并}=\max_{p(x_1x_2\cdots x_N)}I(X_1X_2\cdots X_N;Y_1Y_2\cdots Y_N)=\sum_{i=1}^{N}\max_{p(x_i)}I(X_i;Y_i)=\sum_{i=1}^{N}C_i$

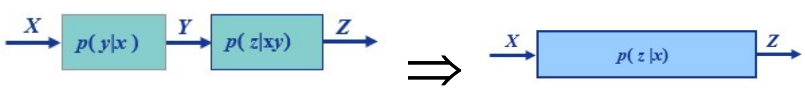
特例：$C_{并}=NC$（N个独立并联信道的信道容量都相同时）

2、级联信道的信道容量

a) 应用场景

- 微波中继接力通信就是一种级联信道。
- 信宿收到数据后再进行数据处理，数据处理系统可看成一种信道，它与前面传输数据的信道构成级联信道。

b) 求法

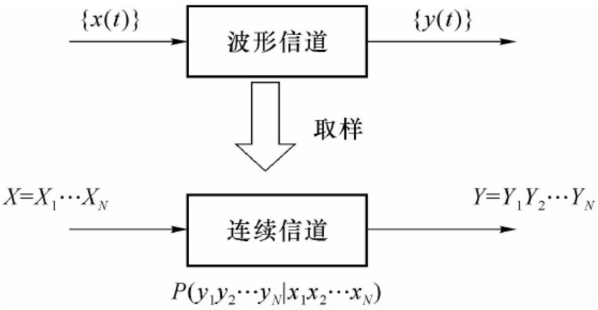


Step1：求总的信道转移矩阵——$\left[P(z\,|\,x)\right]_{r\times t}=\left[P(y\,|\,x)\right]_{r\times s}\left[P(z\,|\,y)\right]_{s\times t}$

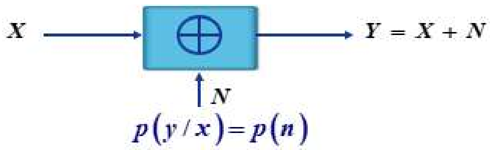
Step2：利用之前求信道容量的方法（通常是用对称信道的结论）

3.4 连续信道与波形信道

一、连续信道与波形信道的关系



二、连续信道的信道容量



1、加性连续信道

$$C=\frac{1}{2}\log_2\left(1+\frac{P_X}{P_N}\right)$$

2、限带高斯白噪声加性波形信道的信道容量（AWGN 信道、可加波形信道）

a) 什么是高斯加性连续信道

$$p(n)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{n^2}{2\sigma^2}\right)$$

b) 信道容量

$$\textbf{(1)}[0,T]\text{时刻内的信道容量：}C=BT\log_2\left(1+\frac{P_X}{n_0B}\right)\qquad bit/\text{样值}$$

$$\textbf{(2)}\text{单位时间内的信道容量：}C_t=B\log_2\left(1+\frac{P_X}{n_0B}\right)\quad bit/s$$

- 第二个为香农公式

c) 香农公式的指导意义

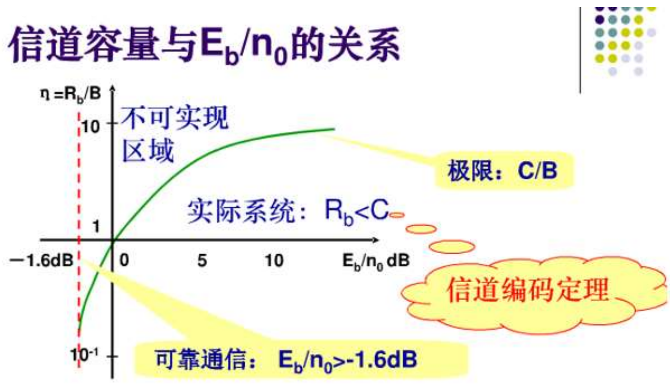
(1)提高信噪比能增加信道的信道容量

(2)当噪声功率$n_0\to 0$时，信道容量$C_t\to\infty$.

推论：无干扰连续信道的信道容量为无穷大

(3)通过增大带宽，达到极限传输速率：$\lim\limits_{B\to\infty}C_t\approx 1.44\dfrac{P_X}{n_0}$——香农极限

(4)当信道容量一定时，带宽、时间、信噪比可以互换



第四章　信源编码

4.1 编码器

一、概述

1、编码的意义：将信源发出的消息符号转换为适合信道传输的符号

2、通信的根本目的：有效而可靠地传输信息

3、几种编码

- 信源编码：信源消息确定后，如何提高信息传输速度，提高通信的有效性（提高编码效率或压缩比）
- 信道编码：信道传输速度一定，如何提高信息传输质量，提高通信的可靠性， 编码具有发现错误或纠正错误的抗干扰能力
- 加密编码：提高通信的安全性

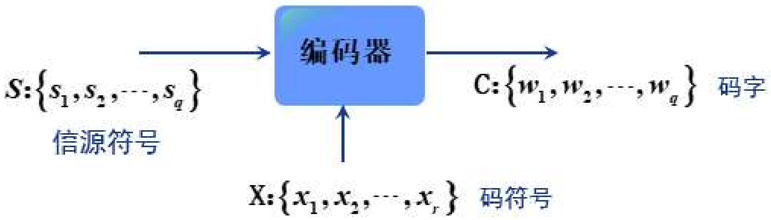
4、两个编码定理

- Shannon 第一编码定理（信源编码定理，有效性编码定理）
- Shannon 第二编码定理（信道编码定理，抗干扰编码定理）

二、编码器

1、数学模型

- 编码就是从信源符号到码符号的一种映射



2、基本概念

a)　等长码与变长码

b)　奇异码与非奇异码：信源符号与码字是否一一对应的（是则为非奇异）

- 非奇异码是正确译码的必要条件
- 非奇异的分组码并不能保证正确的译出，当码字排在一起时还可能出现奇异性。

c)　同价码：每个码符号所占的传输时间都相同

d)　N次扩展码：N个码字Wi 组成的码字

4.2 等长码和等长信源编码定理

一、对无失真编码的要求

1、要求：信源符号与码字是一一对应的（反变换也是唯一对应的），即是非奇异码

2、举例

信源符号	码1	码2
S_1	00	00
S_2	01	11
S_3	10	10
S_4	11	11

码1 是非奇异码,唯一可译；　　码2 是奇异码，非唯一可译。

二、等长码唯一可译的码长条件（无失真编码）

1、单符号信源

$$q \le r^l \Rightarrow l \ge \frac{\log q}{\log r}$$

$$\begin{cases} q\text{：信源符号的种类} \\ l\text{：码长} \\ r\text{：码符号个数} \end{cases}$$

2、N 次扩展信源

$$q^N \le r^l \Rightarrow l \ge N\frac{\log q}{\log r}$$

- 注意：没考虑符号出现的概率，编码效率，译码错误的概率等

3、一般信源——等长信源编码定理

a) 对码长的要求

$$\frac{l}{N} \ge \frac{H(S)+\varepsilon}{\log r} \quad \text{或} \quad \text{编码信息率} R' = \frac{l}{N}\log r \ge H(S)+\varepsilon$$

$$for \quad \forall \varepsilon > 0 \qquad \begin{cases} H(S)\text{：信源熵} \\ N\text{：信源长} \end{cases}$$

- 码长小于右边时，不可能实现无失真编码
- 无记忆、有记忆系统均适用，但是后者需要存在极限熵和极限方差
- 当信源长度 N 为有限时，用等长码有失真

b) 对信源长度的要求

$$N \ge \frac{D[I(S_i)]}{H^2(S)} \cdot \frac{\eta^2}{(1-\eta)^2\delta}$$ 允许错误概率

其中 $D[I(S_i)] = H(S^2) - [H(S)]^2$ ——自信息的方差

$H(S)$ ——信源S的熵

η ——编码效率

$$\eta = \frac{H(S)}{\text{编码后平均每个符号所携带的最大信息量 } R'}$$

4.3 变长码与变长信源编码定理

一、基本概念

1、为什么需要变长码

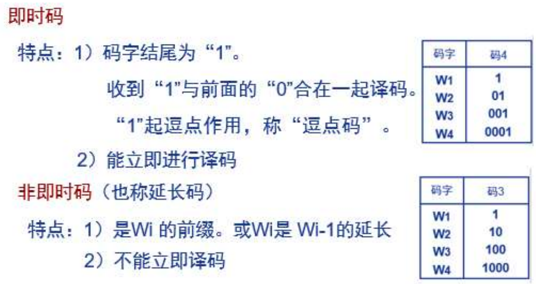
- 等长编码要想实现无失真编码，要求信源序列足够长，实际中很难实现。（对存储器和处理技术要求太高）
- 而变长码，码长 N 不很大时就能编出效率高且无失真的码

2、变长码的概念

a) 码的分类

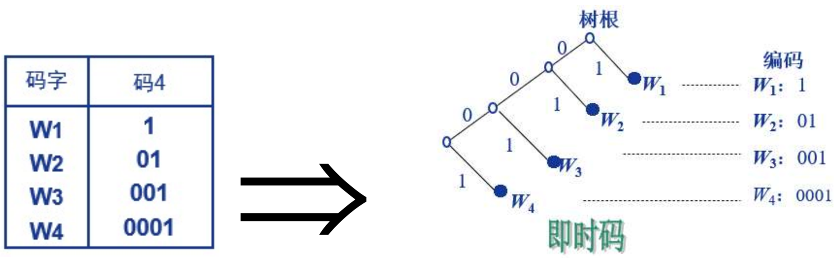
3、即时码与非即时码

a) 码的特点



b) 构造方法——码树图

- 举例



二、存在唯一可译码的充分必要条件——克拉夫特不等式

1、唯一可译的概念

信源符号	概率	码1	码2	码3	码4
s_1	1/2	0	0	1	1
s_2	1/4	11	10	10	01
s_3	1/8	00	00	100	001
s_4	1/8	11	01	1000	0001

非唯一可译　　唯一可译

2、不等式内容

$$\sum_{i=1}^{q} r^{-l_i} \leq 1$$

$$\begin{cases} r\text{是码的状态数（如二元码为2）} \\ l_i\text{是分别对应的码长} \end{cases}$$

码字	
W1	1
W2	01
W3	011
W4	0001

满足Kraft不等式但非唯一可译

- 唯一可译码一定满足不等式
- 满足不等式的码不一定是唯一可译码（但至少能构成一种唯一可译码）

三、变长码信源编码定理

1、紧致码

a) 平均码长的概念:

$$\bar{L} = \sum_{i=1}^{q} P(s_i) l_i$$ 其中：$s_1, s_2, \cdots, s_q$为信源符号，$l_1, l_2, \cdots, l_q$为各符号对应的码长

b) 紧致码：平均码长最小的唯一可译码

2、单符号变长编码定理——紧致码平均码长的理论界限

若一个离散无记忆信源的熵为$H(\mathrm{S})$，码符号$X = [x_1, x_2, \cdots, x_r]$ 则，

总可以找到一种无失真编码方法，构成唯一可译码，使其平均码长满足

$$\frac{H(S)}{\log r} \leq \bar{L} < 1 + \frac{H(S)}{\log r}$$

- 若平均码长$\bar{L} < \dfrac{H(S)}{\log r}$，则不存在唯一可译码

3、变长无失真信源编码定理——香农第一编码定理

离散无记忆信源S的N次扩展信源S^N，码符号$X = [x_1, x_2, \cdots, x_r]$，则

总可以找到一种编码方法，构成惟一可译码

使信源S中每个信源符号的平均码长满足

$$\frac{H(S)}{\log r} \leq \frac{\bar{L}_N}{N} < \frac{1}{N} + \frac{H(S)}{\log r}$$

4.4 几种变长码的编码方法

一、哈夫曼编码

1、编码方法（二元、r 元）

1）将信源消息符号按其出现的概率大小依次排列：

$p_1 \ge p_2 \ge \cdots \ge p_q$

2）取两个概率最小的字母分别用 0 和 1表示，并将这两个概率相加为一个新字母的概率，这时信源还剩q-1个符号，称缩减信源S1。

3）对缩减信源S1的符号仍按概率大小依次排列，在将最后两个概率最小的符号分别用0和1表示，然后，再合并成一个符号，形成了q-2个符号的缩减信源S2。

4）不断继续上述过程，直到最后两个符号用 0 和 1表示 为止。

5）从最后一级开始，向前返回得到各个信源符号所对应的码元序列，即相应的码字。

2、举例（二元、r 元）

3、特点

- 同一组信源符号的编码方式不唯一

【问题】为何霍夫曼编码方法得到的码并非是唯一的?

① 每次对信源缩减时，赋予信源最后两个概率最小的符号，用 0 和 1是可以任意的，所以可得到不同的霍夫曼码，但不影响码字长度。

② 对信源进行缩减时，两个概率最小的符号合并后的概率与其它信源符号的概率相同时，这两者在缩减信源中进行概率排序，其位置放置次序是可以任意的，故会得到不同的霍夫曼码。此时将影响码字的长度，一般将合并的概率放在上面，这样可获得较小的码方差。

- 是最佳变长码
- 码方差小的霍夫曼编码质量好

码方差： $\sigma_l^2 = E\left[\left(l_i - \bar{L}\right)^2\right] = \sum_{i=1}^{q} p(s_i)(l_i - \bar{L})^2$

方案一： $\sigma_{l1}^2 = 1.36$

方案二： $\sigma_{l2}^2 = 0.16$ 好

4、最佳编码方式

a) 尽可能减小码方差：一般将合并的概率放在上面，这样可获得较小的码方差

b) 充分利用短码，使平均码长最短：（需要满足下式）

$$q = (r-1)\theta + r \qquad q\text{为信源符号数，}\theta\text{为缩减次数}$$

$$\begin{cases} \text{二元码：} q = \theta + 2 \\ r\text{元码：q不满足上式，后面可附加概率为0的若干个符号 } S_{q+1}, S_{q+2}, \cdots \end{cases}$$

二、费诺码

1、编码方法

1. 将信源消息符号按其出现的概率大小依次排列：

$$p(x_1) \ge p(x_2) \ge \ \ldots \ \ge p(x_n)。$$

2. 将依次排列的信源符号按概率值分为两大组，使两个组的概率之和近于相同，并对各组赋予一个二进制码元 “0” 和 “1”。
3. 将每一大组的信源符号进一步再分成两组，使划分后的两个组的概率之和近于相同，并又赋予两个组一个二进制符号 “0’ 和 “1”。
4. 如此重复，直至每个组只剩下一个信源符号为止。
5. 信源符号所对应的码字即为费诺码。

2、举例

信源符号	概率					码字
s_1	1/4	0	0			00
s_2	1/4		1			01
s_3	1/8	1	0	0		100
s_4	1/8			1		101
s_5	1/16		1	0	0	1100
s_6	1/16				1	1101
s_7	1/16			1	0	1110
s_8	1/16				1	1111

紧致码

三、二者的评价

1、均考虑了信源概率分布特性——大概率用短码，因而平均码短，称为最佳编码

2、霍夫曼码最好，应用广泛

3、局限性——均需预知概率分布，主要用于无记忆信源

第五章　信道编码

5.1 错误概率和译码规则

一、承上启下

1、承上：第 4 章已经从理论上讨论了，对于无噪无损信道只要对信源进行适当的编码，总能以信道容量无差错的传递信息

2、启下：但是一般信道总会存在噪声和干扰，那么在有噪信道中进行无错传输可以达到的最大信息传输率是多少呢？

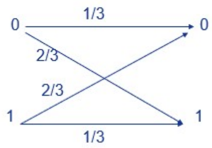
二、错误概率和译码规则

1、基本概念

a) 错误概率

- 影响因素：信道的统计特性、译码方法的选择
- 求法

例：有一个BSC信道，如图所示



若收到“0”译作“0”，收到“1”译作“1”，则平均错误概率为：

$$P_E = P(0)P_e^{(0)} + P(1)P_e^{(1)} = \frac{2}{3}$$

反之，若收到“0”译作“1”，收到“1”译作“0”，则平均错误概率为1/3，可见错误概率与译码准则有关。

b) 译码规则

$$\text{输入符号集}A = \{a_i\}\text{、输出符号集}B = \{b_i\} \Rightarrow \text{译码规则}F(b_j) = a_i$$

- 译码规则的选择应该有一个依据，一个自然的依据就是使平均错误概率最小

c) 平均错误概率

$$P_e = \sum_{j=1}^{m} p(b_j)P(e/b_j) = \sum_{j=1}^{s} p(b_j)(1 - P(a_i/b_j))$$

2、使平均错误概率最小的准则

a) 最小错误概率准则（最大后验概率准则）

$$\boldsymbol{p(F(y_j)\,|\,y_j) = \max_i p(x_i\,|\,y_j)}$$

- 优点：理想
- 缺点：后验概率不易得到且依赖于输入的分布

b) 最大似然译码准则（前者的特例，当输入等概时等价）

$$\boldsymbol{p(y_j\,|\,F(y_j)) = \max_i p(y_j\,|\,x_i)}$$

5.2 错误概率和编码方法

一、基本概念

1、引入

- 一般信道传输时都会产生错误，而选择译码准则并不会消除错误
- 可以通过编码方法减少错误概率呢

2、存在的两个问题

a) 问题一：对于传输同一个信息序列

- 可通过对信源进行 n 次扩展，然后取其中的 M 个，最后采用择多译码的方法进行译码，这样可以大大减少错误率。M 越小错误率越低
- M 越小，会导致传信率也减小，效率降低。

b) 问题二

二、汉明距离与码的最小距离

1、汉明距离

$D(C_i,C_j)$表示$C_i=(c_{i1},c_{i2},\cdots,c_{in})$，$C_j=(c_{j1},c_{j2},\cdots,c_{jn})$中$c_{ir}\neq c_{jr}$的个数

对于二元码：$D(C_i,C_j)=\sum_{k=1}^{n}c_{ik}\oplus c_{jk}$

2、码的最小距离

$$d_{\min}=\min\{D(C_i,C_j)\}\qquad C_i\neq C_j$$

- $d_{\min}$越大，错误概率越小

三、最小距离译码准则（上一节还有一个，一共三个）

1、内容（等价于最大似然译码准则）

在二元对称信道的情况下，译码规则可以如下：

选择 $F(y_j)=x^*$

使之满足 $D(x^*,y_j)\leq D(x_i,y_j)\qquad x_i\neq x^*$

四、编码准则

1、内容：尽量设法使选取的 M 个码字中任意两两不同码字的距离尽量大

5.3 信源编码定理

一、有噪信道信源编码定理（香农第二定理）

1、有噪信道编码定理

离散无记忆信道，信道容量为 C。当信息传输率 $R \leq C$ 时，只要码长足够长，总可以在输入符号集 X^n 中找到 M$(=2^{nR})$个码字组成的一组码$(2^{nR}, n)$和相应的译码准则，使信道输出端的平均错误译码概率达到任意小。

2、有噪信道编码逆定理

如一个离散无记忆信道，信道容量为 C。当信息传输率 $R>C$ 时，则无论码长 n 多长，都找不到一种编码方式使信道输出端的平均错误译码概率达到任意小。

3、总结：在任何信道中，信道容量都是可达的、最大的可靠信息传输率

二、联合信源信源编码定理（即对香农一二定理做总结）

1、香农一二定理的回顾

- 香农第一定理：无失真数据压缩必须 R' > H（有效性）
- 香农第二定理：可靠数据传输 R< C（可靠性）

2、联合定理提出的问题：若信源通过信道传输，要做到有效和可靠地传输，是否$H < C$是充分和必要的条件

3、分析

- 从香农第一、第二定理可以看出，要做到有效和可靠的传输信息，我们可以将通信系统设计成两部分的组合，即信源编码和信道编码两部分
- 信源编码：用尽可能少的信道符号来表达信源，尽可能减少编码后信源的数据的剩余率
- 然后针对信道，对信源编码后的数据独立的进行信道编码，适当增加一些剩余度，使能纠正和克服信道中引起的错误和干扰。
- 我们可以证明，只要满足香农第一定理和第二定理，用两步编码的方法传输信息和一步编码的方法传输信息其效果是一样的。