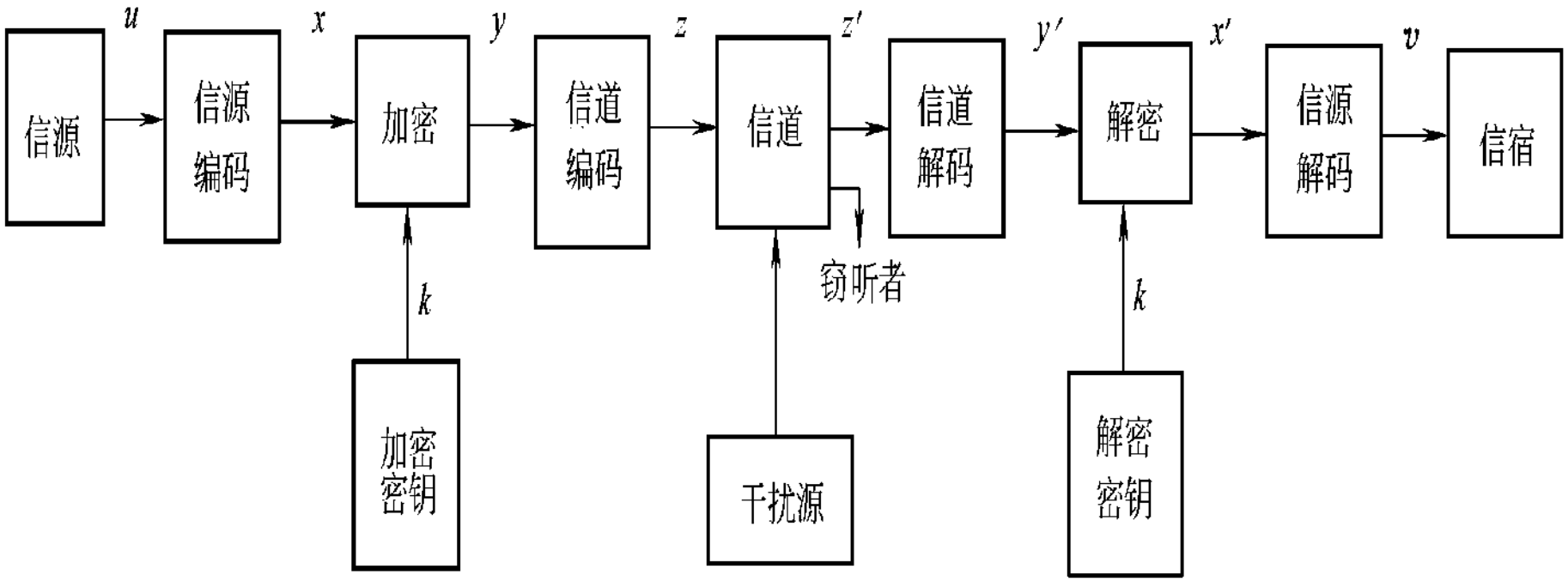
第一章 绪论内容

- 1.1 信息论的形成和发展
  - 信息、信息论和信息技术
  - 信息论研究的基本问题
  - 信息的基本概念（概念、特征及性质）
  - 信息论的发展

- 1.2 通信系统模型
  - 通信系统模型
  - 通信系统的性能指标
  - 通信系统的问题—编码问题
    （信源编码、信道编码、密码）

一、完整的通信系统的模型

- 下图是目前较常用的、也是较完整的通信系统模型 .



- 信源：向通信系统提供消息的人和机器。
- 信宿：消息传递的对象。
- 信道：传递消息的通道。
- 干扰源：整个通信系统中各个干扰的集中反映，用以表示消息在信道中传输时遭受干扰的情况。
- 密钥源：产生密钥 k 源。信源编码器输出信号 x 经过 k 的加密运算后，就把明文 x 变换为密文 y 。

二、通信系统的性能指标

- □　消息从信源通过信道到信宿，如何有效、可靠地传输，是通信系统要解决的根本问题。

- **通信系统的性能指标主要是有效性、可靠性、安全性和经济性。通信系统优化就是使这些指标达到最佳。**

三、通信中的问题—编码问题

- **根据信息论的各种编码定理和上述通信系统的指标，编码问题可分解为三类：信源编码、信道编码和密码。**

- 信源编码器：把信源发出的消息变换成由二进制码元（或多进制码元）组成的代码组以提高通信系统传输消息的效率。

- 信道编码器：在信源编码器输出的代码组上有目的地增加一些监督码元，使之具有检错或纠错的能力。

- 密码学：研究如何隐蔽消息中的信息内容，使它在传输过程中不被窃听，提高通信系统的安全性。

信源编码器的作用

- 通过信源编码可以压缩信源的冗余度以提高通信系统传输消息的效率；以提高传输消息的有效性。

信道编码器的作用

- 在信源编码器输出的代码组上有目的地增加一些监督码元，使之具有检错或纠错的能力；以提高传输消息的可靠性。

第 2 章 信息的度量

1. 自信息　平均自信息、信息熵
2. 熵的性质　　联合熵、条件熵
3. 互信息　　平均互信息
4. 熵的关系 (两个图)
5. 数据处理定理

一、自信息　平均自信息、信息熵

- 1 信息度量的思路
- 2 自信息量的定义
- 3 信息量与不确定度
- 4 平均自信息量

$$I(x_i) \overset{def}{=} -\log p(x_i) = \log \frac{1}{p(x_i)}$$

定义 2.3 随机变量 X 的每一个可能取值的自信息 $I(x_i)$ 的统计平均值定义为随机变量 X 的**平均自信息量**。又称为**信息熵**、**信源熵**，简称**熵**。

$$H(X) = E[I(x_i)] = -\sum_{i=1}^{q} p(x_i)\log p(x_i) \qquad (2.3)$$

熵函数的性质

1. 对称性

2. 确定性

3. 非负性

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

5. 连续性 $\lim_{\varepsilon \to 0} H \ (p_1, \ p_2, \cdots, p_{q-1} - \varepsilon \ , \ p_q + \varepsilon) = H(p_1, \ p_2, \cdots, p_q)$

6. 递推性 $H(p_1, \ p_2, \cdots, p_{n-1}, q_1, \ q_2, \cdots, q_m) = H(p_1, \ p_2, \cdots, p_n) + p_n H\left(\frac{q_1}{p_n}, \ \frac{q_2}{p_n}, \cdots, \frac{q_m}{p_n}\right)$

7. 极值性（最大熵定理） $H(p_1, \ p_2, \cdots, p_n) \le H\left(\frac{1}{n}, \ \frac{1}{n}, \cdots, \frac{1}{n}\right) = \log n$

8. 上凸性

联合熵和条件熵的关系

1. 联合熵与信息熵、条件熵的关系：

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

这个关系可以方便地推广到N个随机变量的情况：

$$H(X_1X_2\cdots X_N)=H(X_1)+H(X_2|X_1)+\cdots+H(X_N|X_1X_2\cdots X_{N-1})$$

称为**熵函数的链规则**。

推论： 当二维随机变量X，Y相互独立时，联合熵等于X和Y各自熵之和：

$$H(XY)=H(X)+H(Y)$$

2．条件熵与信息熵的关系：

$$H(X|Y)\le H(X),H(Y|X)\le H(Y)$$

3．联合熵和信息熵的关系：

$$H(XY)\le H(X)+H(Y)$$

当X、Y相互独立时等号成立。

返回

互信息、平均互信息

- 1 互信息的定义
- 2 平均互信息的定义
- 3 平均互信息性质
- 4 通信系统中的互信息

1 互信息

定义 2.2 一个事件y_i所给出关于另一个事件x_i的信息定义为互信息，用$I(x_i;y_j)$表示。

$$I(x_i;y_j)=I(x_i)-I(x_i\mid y_j)=\log\frac{p(x_i\mid y_j)}{p(x_i)} \qquad (2.2)$$

定义：后验概率与先验概率比值的对数．

$$I(x_i;y_j)=\log\frac{p(x_i/y_j)}{p(x_i)}$$

互信息是已知事件 Yj 后所消除的关于事件 Xi 的不确定性，它等于事件 Xi 本身的不确定性 I（ xi ）减去已知事件 Yj 后对 Xi 仍然存在的不确定性 I（ xi|yj ）

2 平均互信息

为了从整体上表示从一个随机变量Y所给出关于另一个随机变量X的信息量，我们定义互信息$I(x_i;y_j)$在的XY联合概率空间中的统计平均值为随机变量X和Y间的**平均互信息**：

定义2.6

$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)I(x_i;y_j)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{p(x_i\mid y_j)}{p(x_i)}$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{1}{p(x_i)}-\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{1}{p(x_i\mid y_j)}$$

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

3 平均互信息性质

1. 非负性： $I(X;Y) \geq 0$

平均互信息是非负的，说明给定随机变量Y后，一般来说总能消除一部分关于X的不确定性。

2. 互易性（对称性）： $I(X;Y) = I(Y;X)$

对称性表示Y从X中获得关于的信息量等于X从Y中获得关于的信息量。

3.平均互信息和各类熵的关系：

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

当X,Y统计独立时，$I(X;Y) = 0$

4. 极值性： $I(X;Y) \le H(X), I(X;Y) \le H(Y)$

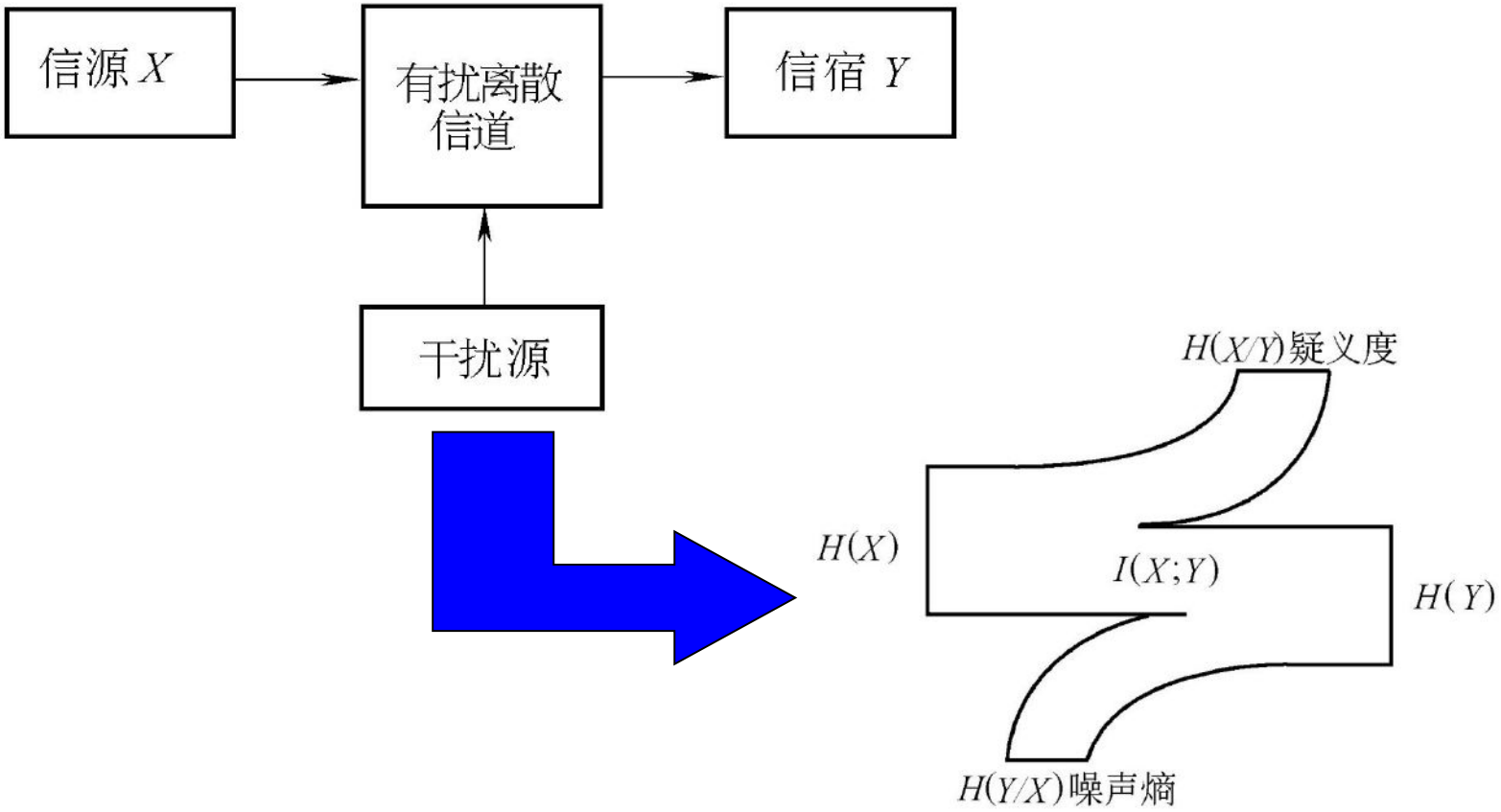
极值性说明从一个事件提取关于另一个事件的信息量，至多只能是另一个事件的平均自信息量那么多，不会超过另一事件本身所含的信息量。

5. 凸函数性：

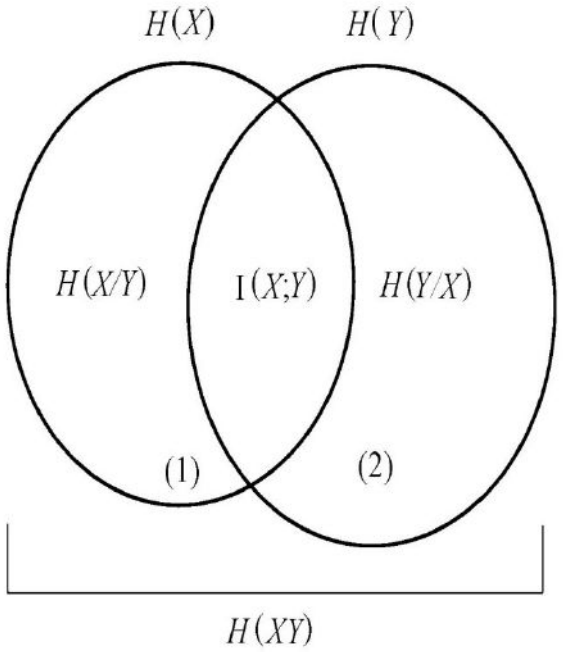
定理2.1 当条件概率分布 $\{p(y_j \mid x_i)\}$给定时，平均互信息 $I(X;Y)$ 是输入分布$\{p(x_i)\}$的上凸函数。

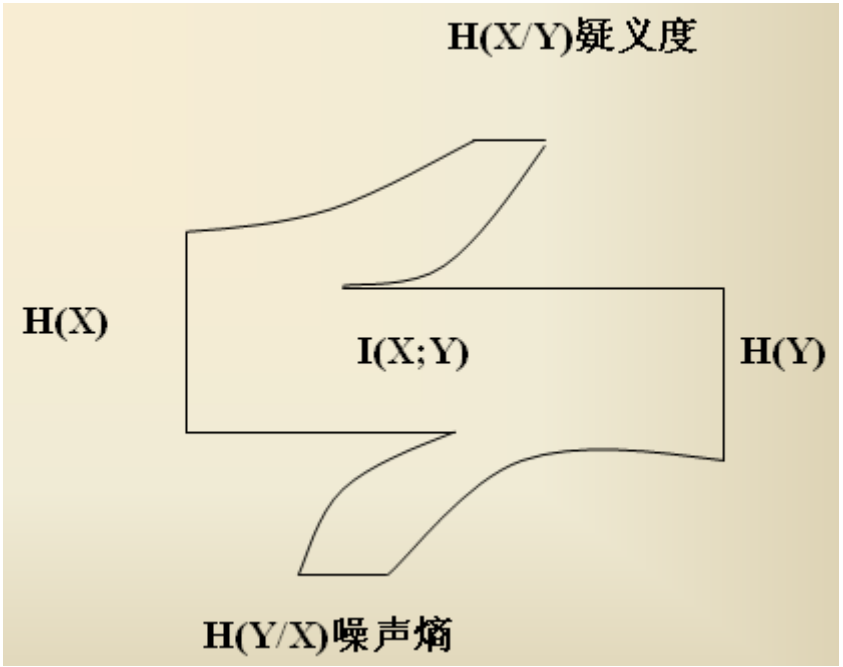
定理2.2 对于固定的输入分布 $\{p(x_i)\}$，平均互信息量$I(X;Y)$ 是条件概率分布 $\{p(y_j \mid x_i)\}$的下凸函数。

4　通信系统中的平均互信息



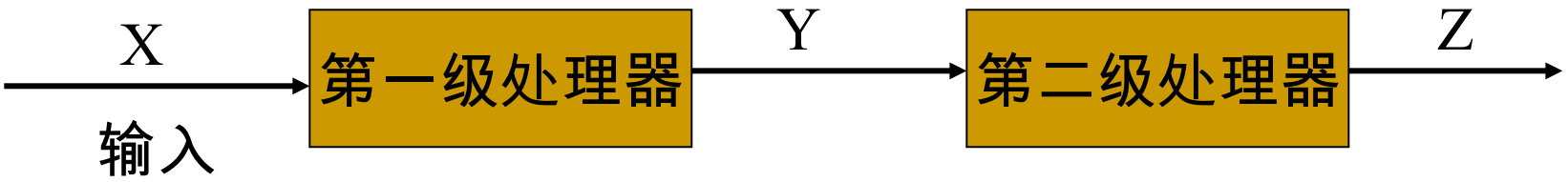
各种熵之间的关系图示





数据处理定理

当消息通过多级处理器时，随着处理器数目的增多，输人消息与输出消息之间的平均互信息量趋于变小。



级联处理器

结论：数据处理过程中只会失掉一些信息，绝不会创造出新的信息，所谓信息不增性。

习题重点：第二章所留的作业

第 3 章　信源及信源熵

1. 信源的分类及数学模型
2. 离散单符号信源
3. 离散多符号信源
4. 马尔卡夫信源
5. 冗余度

信源(Information Source)是信息的来源

信源的主要问题：

1. 如何描述信源（信源的数学建模问题）
2. 怎样计算信源所含的信息量
3. 怎样有效的表示信源输出的消息，也就是信源编码问题

一、信源的分类及数学模型

1. 信源的定义
2. 不同分类模式获得的信息源分类
3. 基于可研究的信源分类方式
4. 信源的数学描述方式

二、离散单符号信源

1 离散单符号信源

2 离散信源的特例 --- 二元信源

三、离散平稳多符号信源

1. 离散多符号信源
2. 离散多符号信源的平均不确定度—熵率
3. 离散多符号无记忆信源定义及熵率计算
4. 离散平稳有记忆信源
5. 离散平稳有记忆信源的熵率计算问题

1 离散多符号信源

定义3.1：对于随机变量序列 $X_1, X_2, \cdots, X_n, \cdots$，在任意两个不同时刻 i 和 j（i 和 j 为大于1的任意整数）信源发出消息的概率分布完全相同，即对于任意的 $N = 0,\ 1,\ 2,\ \cdots$，$X_i X_{i+1} \cdots X_{i+N} \cdots$ 和 $X_j X_{j+1} \cdots X_{j+N} \cdots$ 具有相同的概率分布。也就是

$$P(X_i) = P(X_j)$$

$$P(X_i X_{i+1}) = P(X_j X_{j+1})$$

$$\vdots$$

$$P(X_i X_{i+1} \cdots X_{i+N}) = P(X_j X_{j+1} \cdots X_{j+N})$$

即各维联合概率分布均与时间起点无关的信源称为离散平稳信源。

2 熵率（极限熵）

对于离散多符号信源，我们引入**熵率**的概念，它表示信源输出的符号序列中，平均每个符号所携带的信息量。

定义3.2 随机变量序列中，对前N个随机变量的联合熵求平均：

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

称为**平均符号熵**。如果当$N\to\infty$时上式极限存在，则$\lim_{N\to\infty}H_N(X)$称为**熵率**，或称为**极限熵**，记为

$$H_\infty \overset{def}{=} \lim_{N\to\infty}H_N(X)$$

3 离散平稳无记忆信源熵率的计算

离散平稳无记忆信源输出的符号序列是平稳随机序列，并且符号之间是无关的，即是统计独立的。假定信源每次输出的是N长符号序列，则它的数学模型是N维离散随机变量序列：$X = X_1 X_2 \cdots X_N$，并且每个随机变量之间统计独立。同时，由于是平稳信源，每个随机变量的统计特性都相同，我们还可以把一个输出N长符号序列的信源记为：

$$X = X_1 X_2 \cdots X_N = X^N$$

根据统计独立的多维随机变量的联合熵与信息熵之间的关系，可以推出：

$$H(X) = H(X^N) = N\,H(X)$$

离散平稳无记忆信源的熵率：

$$H_\infty = \lim_{N\to\infty} H_N(X) = \lim_{N\to\infty} \frac{1}{N} NH(X) = H(X)$$

4 离散平稳有记忆信源

实际信源往往是有记忆信源。对于相互间有依赖关系的N维随机变量的联合熵存在以下关系（熵函数的链规则）：

$$H(\boldsymbol{X}) = H(X_1X_2\cdots X_N)$$
$$= H(X_1)+H(X_2|X_1)+H(X_3|X_1X_2)+\cdots+H(X_N|X_1X_2\cdots X_{N-1})$$

定理3.1 对于离散平稳信源，有以下几个结论：

（1）条件熵$H(X_N|X_1X_2\cdots X_{N-1})$随$N$的增加是递减的；

（2）N给定时平均符号熵大于等于条件熵，即

$$H_N(\boldsymbol{X}) \geq H(X_N|X_1X_2\cdots X_{N-1})$$

（3）平均符号熵$H_N(\boldsymbol{X})$随N的增加是递减的；

（4）如果$H(X_1)<\infty$，则$H_\infty = \lim_{N\to\infty} H_N(\boldsymbol{X})$存在，并且

$$H_\infty = \lim_{N\to\infty} H_N(\boldsymbol{X}) = H(X_N|X_1X_2\cdots X_{N-1})$$

四、马尔卡夫信源

1. 马尔科夫特性
2. 马尔科夫信源
3. 马尔科夫信源的计算条件
4. 马尔可夫信源序列熵的计算

马尔可夫信源

有一类信源，信源在某时刻发出的符号仅与在此之前发出的有限个符号有关，而与更早些时候发出的符号无关，这称为马尔可夫性，这类信源称为马尔可夫信源。马尔可夫信源可以在N不很大时得到H_∞。如果信源在某时刻发出的符号仅与在此之前发出的m个符号有关，则称为m阶马尔可夫信源，它的熵率：

$$H_\infty = \lim_{N\to\infty} H(X_N \mid X_1 X_2 \cdots X_{N-1})$$

$$= \lim_{N\to\infty} H(X_N \mid X_{N-m} X_{N-m+1} \cdots X_{N-1}) \quad \text{（马尔可夫性）}$$

$$= H(X_{m+1} \mid X_1 X_2 \cdots X_m) \quad \text{（平稳性）}$$

M 阶马尔可夫信源熵率的计算

下面计算遍历的m阶马尔可夫信源的熵率。

当时间足够长后，遍历的马尔可夫信源可以视作平稳信源来处理，又因为m阶马尔可夫信源发出的符号只与最近的m个符号有关，所以极限熵H_∞等于条件熵 H_{m+1}。

对于齐次遍历的马尔可夫链，其状态s_i由$x_{i_1}x_{i_2}\cdots x_{i_m}$唯一确定，因此有 $p(s_j \mid s_i) = p(x_{i_{m+1}} \mid x_{i_1}x_{i_2}\cdots x_{i_m}) = p(x_{i_{m+1}} \mid s_i)$

- 所以：

$$H_{m+1} = H(X_{m+1} \mid X_1 X_2 \cdots X_m) = E\left[p(x_{i_{m+1}} \mid x_{i_1} x_{i_2} \cdots x_{i_m})\right]$$

$$= E\left[p(x_{i_{m+1}} \mid s_i)\right] = -\sum_{i=1}^{q^m} \sum_{i_{m+1}=1}^{q} p(s_i) p(x_{i_{m+1}} \mid s_i) \log p(x_{i_{m+1}} \mid s_i)$$

$$= \sum_i p(s_i) H(X \mid s_i) == -\sum_i \sum_j p(s_i) p(s_j \mid s_i) \log p(s_j \mid s_i)$$

其中，$p(s_i)$是马尔可夫链的平稳分布或称状态极限概率；$H(X \mid s_i)$表示信源处于某一状态s_i时发出下一个符号的平均不确定性；$p(s_j \mid s_i)$表示下一步状态转移概率.

马尔可夫信源序列熵的计算步骤

1、获得信源的状态转移概率矩阵
(*判断 遍历性)

2、利用 $W_j = \sum_i W_i P_{ij}, \quad \sum_j W_j = 1,$ 求解 Wj

3、利用公式 $H_{m+1} = \sum_i p(s_i) H(X_i / s_i)$ 求 H_{m+1}

4、$H_\infty(X) = \lim_{L \to \infty} H(X_L / X_1 X_2 \cdots X_{L-1})$

$= H(X_{m+1} / X_1 X_2 \cdots X_m) = H_{m+1}(X)$

五、信源的相关性和剩余度—冗余度

1 信息的冗余度

2 通信效率与可靠性

信源的相关性和剩余度

根据定理3.1可得 $\log q = H_0 \geq H_1 \geq H_2 \geq \cdots \geq H_{m+1} \geq \cdots \geq H_\infty$

由此看出，由于信源输出符号间的依赖关系也就是信源的相关性使信源的实际熵减小。信源输出符号间统计约束关系越长，信源的实际熵越小。当信源输出符号间彼此不存在依赖关系且为等概率分布时，信源的实际熵等于最大熵 H_0。

定义3.3 一个信源的熵率（极限熵）与具有相同符号集的最大熵的比值称为**熵的相对率**：

$$\eta = \frac{H_\infty}{H_0}$$

信源剩余度为：$\gamma = 1 - \eta = 1 - \frac{H_\infty}{H_0} = 1 - \frac{H_\infty}{\log q}$

信源的相关性和剩余度

信源的**剩余度**来自两个方面，一是信源符号间的**相关性**，相关程度越大，符号间的依赖关系越长，信源的实际熵越小，另一方面是信源符号分布的**不均匀性**使信源的实际熵越小。

为了更经济有效的传送信息，需要尽量压缩信源的剩余度，**压缩剩余度的方法就是尽量减小符号间的相关性，并且尽可能的使信源符号等概率分布。**

从提高信息传输效率的观点出发，人们总是希望尽量去掉剩余度。但是从提高抗干扰能力角度来看，却希望增加或保留信源的剩余度，因为剩余度大的消息抗干扰能力强。

信源编码是减少或消除信源的剩余度以提高信息的传输效率，而信道编码则通过增加冗余度来提高信息传输的抗干扰能力。

- 习题重点：课后习题及课堂上的练习题