

数据通信原理

信息论基础

全宇晖

二零一九年秋

信源信息量度量

信息量

【讨论】什么是信息量？

通信系统任务的形式化

“形式化”说明一个物理的通信系统的通信过程，一般与其传送内容无关。

所以，一个二元的通信系统的任务是什么？

离散信源信息的建模

X	x_1	x_2	\cdots	x_N
$P(X)$	$p(x_1)$	$p(x_2)$	\cdots	$p(x_N)$

$$\sum_{i=1}^N p_i(x_i) = 1$$

[讨论] 怎么认为 x_i 的信息量比 x_j 多或少?

信息量的度量函数

X	x_1	x_2	\dots	x_N
$P(X)$	$p(x_1)$	$p(x_2)$	\dots	$p(x_N)$

信息量度量函数: $I: X \rightarrow \mathcal{R}$

f 应该满足什么特性?

1. 与 $p(x_i)$ 成反比
2. 可叠加性

信息量的度量

1. 与 $p(x_i)$ 成反比
2. 可叠加性

这两个条件意味着怎样的数学条件？

1. $I(x_1, x_2) = f[p(x_1, x_2)]$
 $I(x_1, x_2) = I(x_1) + I(x_2)$
 $f[p(x_1, x_2)] = f[p(x_1)] + f[p(x_2)]$
2. f/I 是一个单调非增函数

有哪些函数满足这些数学条件？

离散信源和消息的信息量

$$I(x_i) = \log_2 \left(\frac{1}{p(x_i)} \right) = -\log_2 p(x_i)$$

X	$x_1 = 0$	$x_2 = 1$	$x_2 = 3$	$x_4 = 4$
$P(X)$	$1/2$	$1/16$	$3/8$	$1/16$

离散信源的信息量 $I(X) = ?$

离散消息的信息量 $I(S = 131400) = ?$

信息量的单位是？

离散信源的信息熵

X	x_1	x_2	\dots	x_N
$P(X)$	$p(x_1)$	$p(x_2)$	\dots	$p(x_N)$

$$H(X) = - \sum_{i=1}^N \log_2(p(x_i))$$

熵是离散信源中每个符号所含的统计平均信息量

熵的单位是？

离散信源的信息熵

- 求以下两种情况的熵

X	x_1	x_2
$P(X)$	$1/2$	$1/2$

X	x_1	x_2
$P(X)$	1	0

- 什么时候熵最大？

离散信源的最大熵定理

当信源 X 去均匀分布时，其熵 $H(X)$ 最大。

有什么启发？

对信源的符号/码元进行某种变化，使得变换后的符号集具有等概或近似等概的特性，可使得每个符号/码元可携带最大的信息量。

离散信源的条件熵和联合熵

假设随机变量

$$\{XY : x_i y_j, i = 1, 2, \dots, M; j = 1, 2, \dots, N\}$$

服从概率场

	$x_1 y_1, p(x_1 y_1)$	$x_1 y_2, p(x_1 y_2)$...	$x_1 y_N, p(x_1 y_N)$
$XY :$	$x_2 y_1, p(x_2 y_1)$	$x_2 y_2, p(x_2 y_2)$...	$x_2 y_N, p(x_2 y_N)$
$p(XY) :$
	$x_M y_1, p(x_M y_1)$	$x_M y_2, p(x_M y_2)$...	$x_M y_N, p(x_M y_N)$

且满足条件

$$\sum_{i=1}^M \sum_{j=1}^N p(x_i y_j) = 1$$

离散信源的条件熵和联合熵

- 联合熵

$$H(XY) = -\sum_{i=1}^M \sum_{j=1}^N p(x_i y_j) \log p(x_i y_j) \quad (\text{比特/符号})$$

如果 X, Y 独立, 则有

$$\begin{aligned} H(XY) &= -\sum_{i=1}^M \sum_{j=1}^N p(x_i y_j) \log p(x_i y_j) \\ &= -\sum_{i=1}^M \sum_{j=1}^N p(x_i) p(y_j) \log [p(x_i) p(y_j)] \\ &= -\sum_{i=1}^M p(x_i) \log p(x_i) - \sum_{j=1}^N p(y_j) \log p(y_j) \\ &= H(X) + H(Y) \end{aligned}$$

离散信源的条件熵和联合熵

- 条件熵

$$H(X/Y) = -\sum_{i=1}^M \sum_{j=1}^N p(x_i y_j) \log p(x_i / y_j)$$

$$H(Y/X) = -\sum_{i=1}^M \sum_{j=1}^N p(x_i y_j) \log p(y_j / x_i)$$

条件熵不等式

$$H(X/Y) \leq H(X) \quad H(Y/X) \leq H(Y) \quad H(XY) \leq H(X) + H(Y)$$

说明了什么？

1. 一个随机变量的出现总有助于降低另一随机变量的不确定性。
2. 混在一起的随机变量的不确定性不大于各个随机变量的不确定性之和。（为什么？）

熵的应用

- 特征选择
- 决策树

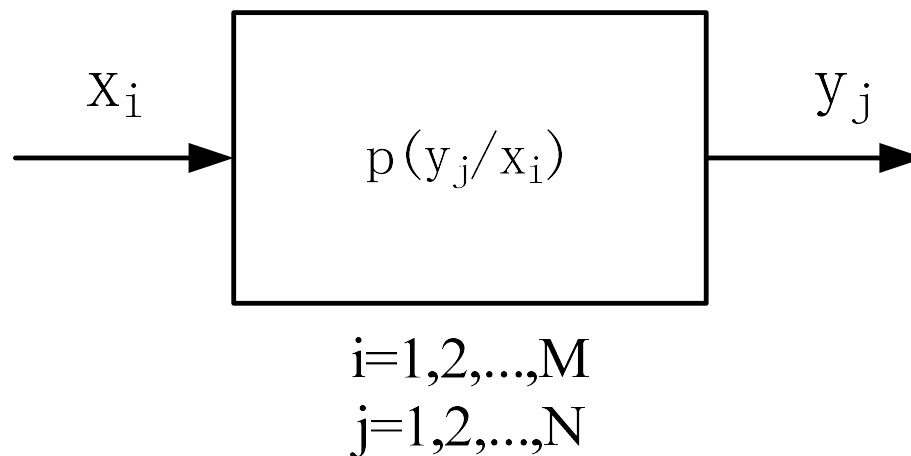
离散信道及容量

离散信道及容量

- 离散信道可用概率转移来描述

信道的输入 $\{X : x_i, i = 1, 2, \dots, M\}$

信道的输出 $\{Y : y_j, j = 1, 2, \dots, N\}$



概率转移矩阵与后验概率矩阵

- 概率转移矩阵

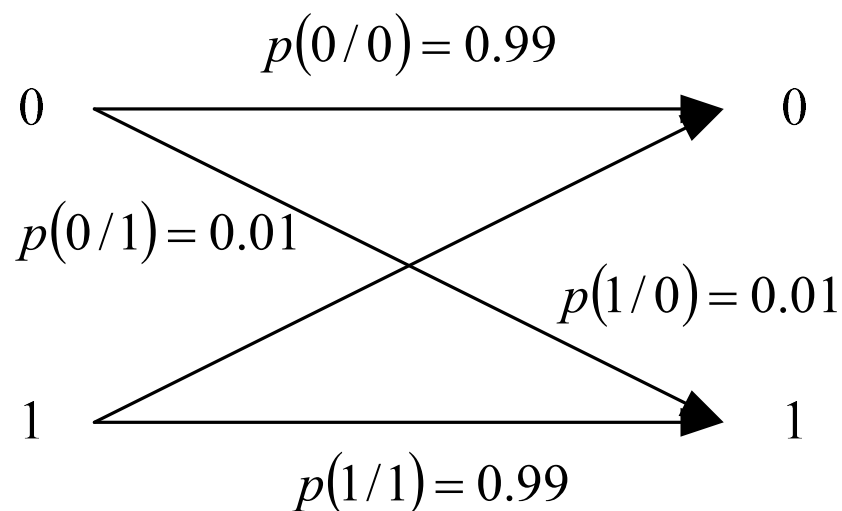
$$[p(Y/X)] = \begin{bmatrix} p(y_1/x_1) & p(y_2/x_1) & \dots & p(y_N/x_1) \\ p(y_1/x_2) & p(y_2/x_2) & \dots & p(y_N/x_2) \\ \dots & \dots & \dots & \dots \\ p(y_1/x_M) & p(y_2/x_M) & \dots & p(y_N/x_M) \end{bmatrix}$$

- 后验概率矩阵

$$[p(X/Y)] = \begin{bmatrix} p(x_1/y_1) & p(x_2/y_1) & \dots & p(x_M/y_1) \\ p(x_1/y_2) & p(x_2/y_2) & \dots & p(x_M/y_2) \\ \dots & \dots & \dots & \dots \\ p(x_1/y_N) & p(x_2/y_N) & \dots & p(x_M/y_N) \end{bmatrix}$$

示例

- 二元的离散无记忆信道发“0”和发“1”时能正确接收的概率为0.99，错误的概率为0.01。



求概率转移矩阵和后验概率矩阵

互信息量

- 接收端收到 y_j 后，对于发送端发出的是 x_i 这件事的不确定性可以表达为：

$$I(x_i / y_j) = \log \frac{1}{p(x_i / y_j)}$$

- 互信息量定义为收到 y_j ，对关于 x_i 的不确定性的抵消量：

$$I(x_i; y_j) = I(x_i) - I(x_i / y_j)$$

互信息具有对称性吗？

互信息的性质

互信息量具有对称性

$$\begin{aligned} I(x_i; y_j) &= I(x_i) - I(x_i / y_j) = \log \frac{1}{p(x_i)} - \log \frac{1}{p(x_i / y_j)} \\ &= \log \frac{p(x_i / y_j)}{p(x_i)} = \log \frac{p(x_i y_j)}{p(x_i) p(y_j)} = \log \frac{p(y_j / x_i)}{p(y_j)} = I(y_j; x_i) \end{aligned}$$

互信息量的性质

- (1) 若 $p(x_i / y_j) = 1 \rightarrow I(x_i; y_j) = I(x_i)$
- (2) 若 $p(x_i) < p(x_i / y_j) < 1 \rightarrow 0 < I(x_i; y_j) < I(x_i)$
- (3) 若 $p(x_i / y_j) = p(x_i) \rightarrow I(x_i; y_j) = 0$
- (4) 若 $0 < p(x_i / y_j) < p(x_i) \rightarrow I(x_i; y_j) < 0$

平均互信息量

- 平均互信息量定义为:

$$I(X;Y) = \sum_{i=1}^M \sum_{j=1}^N p(x_i y_j) I(x_i; y_j)$$

- 平均互信息量具有非负性:

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

说明了什么？

从统计上来说，两相关联的随机变量集，其中一个集中符号的出现总是有利于提供有关另外一个集的信息。

平均互信息量与熵

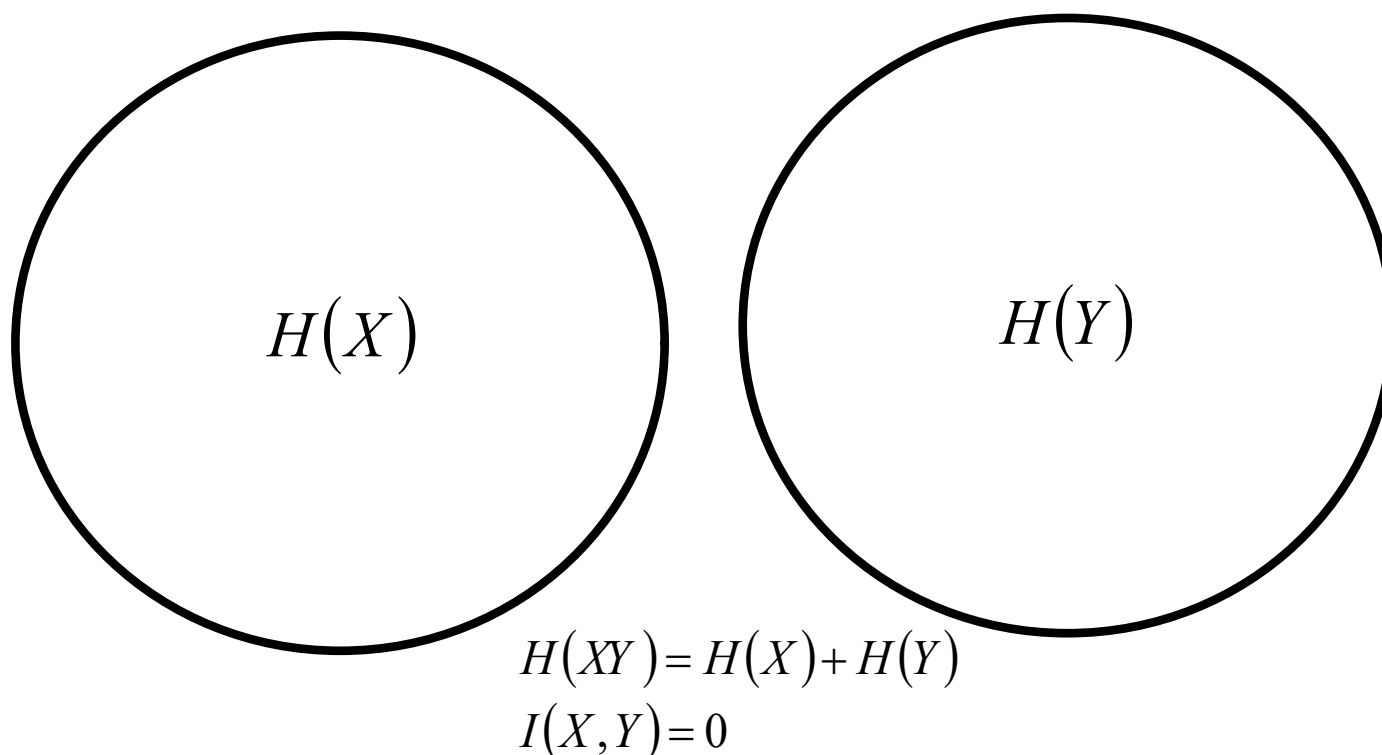
$$I(X;Y) = H(X) - H(X/Y)$$

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

进一步我们可以推导出什么？

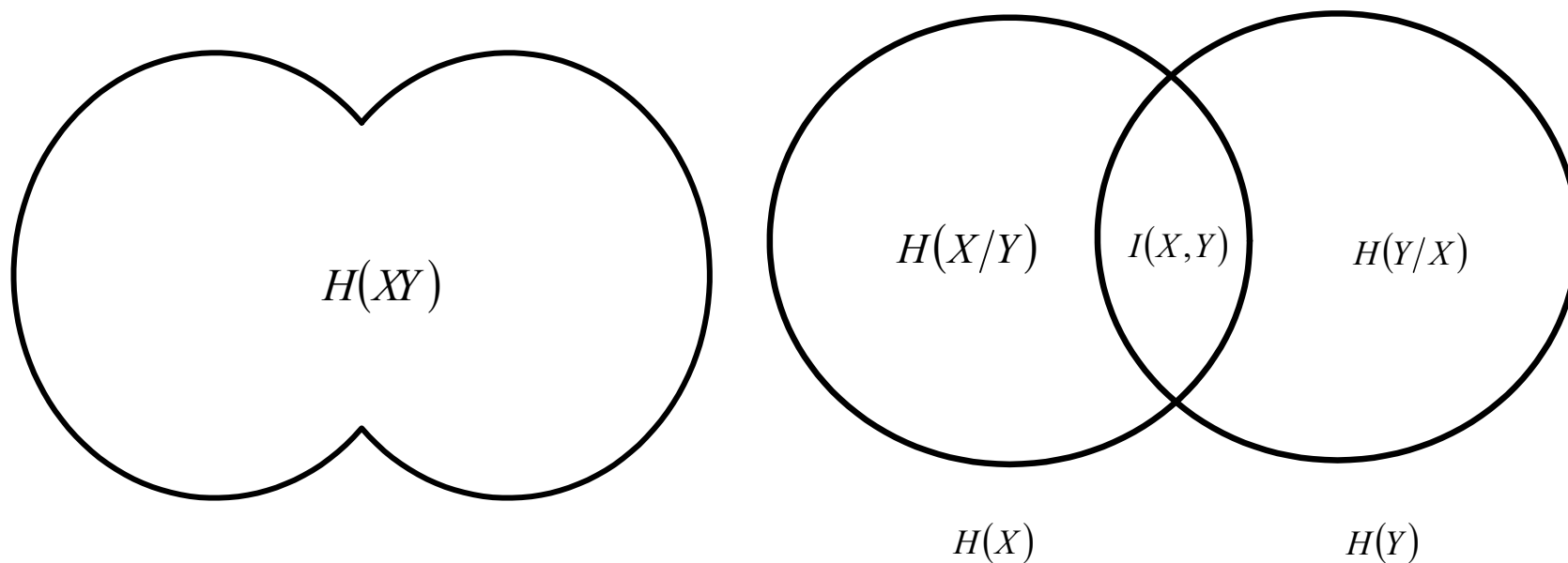
平均互信息量与熵

- 两个信息相互独立情况



平均互信息量与熵

- 两个信息相互不独立情况



系统的平均信息速率

$$\begin{aligned} R_I &= I(X;Y) \times R_S = \left[\sum_{i=1}^M \sum_{j=1}^N p(x_i y_j) \log \frac{p(x_i / y_j)}{p(x_i)} \right] \times R_S \\ &= \left[\sum_{i=1}^M \sum_{j=1}^N p(x_i / y_j) p(x_i) \log \frac{p(x_i / y_j)}{p(x_i)} \right] \times R_S \end{aligned}$$

符号传输速率： R_S

系统信息速率： R_I

信道容量

信道容量=离散信道的最大信息传输速率

$$C = \max_{p(x_i), i=1,2,\dots,M} R_I = \max_{p(x_i), i=1,2,\dots,M} [I(X;Y) \times R_s]$$

匹配信源的概念

- 信道容量由信道特性和信源的统计特性共同决定。
- 信道特性确定后，其容量由信源的统计特性决定
- 能使单位时间内信道可传输的平均信息量达到信道容量的信源即匹配信源。

匹配信源的相关优化

已知信道转移概率，匹配信源统计特性的求解

约束条件 $\sum_{i=1}^M p(x_i) = 1$

求 $p(x_i), i = 1, 2, \dots, N$ 使得 $I(X; Y)$ 达到最大。

由拉格朗日求极值的原理：

定义辅助方程 $F[p(x_1), p(x_2), \dots, p(x_M)] = I(X; Y) - \lambda \left(\sum_{i=1}^M p(x_i) - 1 \right)$

令

$$\frac{\partial F[p(x_1), p(x_2), \dots, p(x_M)]}{\partial p(x_i)} = \frac{\partial I(X; Y)}{\partial p(x_i)} - \lambda = 0 \quad i = 1, 2, \dots, M$$

可得信源分布特性应满足的条件

$$\begin{cases} \sum_{j=1}^N p(y_j / x_i) \log \frac{p(y_j / x_i)}{p(y_j)} = \lambda & i = 1, 2, \dots, M \\ \sum_{i=1}^M p(x_i) = 1 \end{cases}$$

匹配信源的相关优化

若取得其解为： $[p_m(x_1), p_m(x_2), \dots, p_m(x_M)]$
则可得最大的平均互信息量为

$$\begin{aligned} I_{\max} &= \max_{\{p(x_i), i=1, \dots, M\}} I(X; Y) = \sum_{i=1}^M \sum_{j=1}^N p(y_j / x_i) p_m(x_i) \log \frac{p(y_j / x_i)}{p(y_j)} \\ &= \sum_{i=1}^M p_m(x_i) \sum_{j=1}^N p(y_j / x_i) \log \frac{p(y_j / x_i)}{p(y_j)} = \sum_{i=1}^M p_m(x_i) \lambda = \lambda \end{aligned}$$

由上述两组关系式可得

$$\sum_{j=1}^N p(y_j / x_i) \log \frac{p(y_j / x_i)}{p(y_j)} = I_{\max}, \quad i = 1, 2, \dots, M$$

整理可得

$$\begin{aligned} \sum_{j=1}^N p(y_j / x_i) [\log p(y_j / x_i) - \log p(y_j)] &= I_{\max}, \quad i = 1, 2, \dots, M \\ \sum_{j=1}^N p(y_j / x_i) \log p(y_j / x_i) &= \sum_{j=1}^N p(y_j / x_i) [I_{\max} + \log p(y_j)], \quad i = 1, 2, \dots, M \end{aligned}$$

匹配信源的相关优化

记 $\beta_j = I_{\max} + \log p(y_j)$, $j = 1, 2, \dots, N$

可得 $\sum_{j=1}^N p(y_j / x_i) \log p(y_j / x_i) = \sum_{j=1}^N p(y_j / x_i) \beta_j$, $i = 1, 2, \dots, M$

若已知信道转移概率

可求解 $p(y_j / x_i)$, $i = 1, 2, \dots, M; j = 1, 2, \dots, N$ $\beta_j, j = 1, 2, \dots, N$

由此可进一步得

$$p(y_j) = 2^{\beta_j - I_{\max}}, j = 1, 2, \dots, N$$

利用概率约束条件: $\sum_{j=1}^N p(y_j) = \sum_{j=1}^N 2^{\beta_j - I_{\max}} = 1$

$$I_{\max} = \log \left(\sum_{j=1}^N 2^{\beta_j} \right)$$

匹配信源的相关优化

由此可得关于信源分布特性 $p(x_i), i = 1, 2, \dots, M$ 的方程组：

$$p(y_j) = \sum_{i=1}^M p(y_j x_i) = \sum_{i=1}^M p(y_j / x_i) p_m(x_i), j = 1, 2, \dots, N$$

求解该方程组，即可得能够达到信道容量的信源统计特性。

$$p_m(x_i), i = 1, 2, \dots, M$$

匹配信源统计特性的求解步骤

(1) 解方程组

$$\sum_{j=1}^N p(y_j / x_i) \log p(y_j / x_i) = \sum_{j=1}^N p(y_j / x_i) \beta_j \quad i = 1, 2, \dots, M$$

求解得 $\beta_j, \quad i = 1, 2, \dots, M$

(2) 求最大平均互信息量：

$$I_m = \log \left(\sum_{j=1}^N 2^{\beta_j} \right)$$

(3) 求相应后验概率：

$$p(y_j) = 2^{\beta_j - I_{\max}} \quad j = 1, 2, \dots, N$$

(4) 解方程组，

$$p(y_j) = \sum_{i=1}^M p(y_j x_i) = \sum_{i=1}^M p(y_j / x_i) p_m(x_i) \quad j = 1, 2, \dots, N$$

确定匹配信源的分布特性

$$p(x_i), i = 1, 2, \dots, N$$

匹配信源统计特性的求解步骤

示例：求匹配信源的统计特性。已知信道转移概率

$$[p(Y/X)] = \begin{bmatrix} 1/2 & 1/4 & 0 & 1/4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1/4 & 0 & 1/4 & 1/2 \end{bmatrix}$$

(1) 解方程组得中间结果参数：

$$\beta_1 = -2, \quad \beta_2 = 0, \quad \beta_3 = 0, \quad \beta_4 = -2$$

(2) 求最大平均互信息量：

$$I_m = \log\left(\sum_{j=1}^N 2^{\beta_j}\right) = \log(2^{-2} + 2^0 + 2^0 + 2^{-2}) = \log(5/2)$$

(3) 求相应后验概率：

$$p(y_1) = 2^{-2-\log(5/2)} = 1/10, \quad p(y_2) = 2^{0-\log(5/2)} = 2/5$$

$$p(y_3) = 2^{0-\log(5/2)} = 2/5, \quad p(y_4) = 2^{-2-\log(5/2)} = 1/10$$

匹配信源统计特性的求解步骤

(4) 获得匹配信源统计特性：

$$p_m(x_1)=4/30, \quad p_m(x_2)=11/30, \quad p_m(x_3)=11/30, \quad p_m(x_4)=4/30$$

(5) 信道容量为：

$$C = I_m R_s = \log(5/2) \times 1000 = 1.32 \times 1000 = 1320 (\text{比特/秒})$$

注：若求解过程中出现 $p_m(x_i) < 0$ ，可在方程组中令 $p_m(x_i) = 0$ ，重新求解。

离散无记忆信道

$$[p(Y/X)] = \begin{bmatrix} p(y_1/x_1) & p(y_2/x_1) & \dots & p(y_N/x_1) \\ p(y_1/x_2) & p(y_2/x_2) & \dots & p(y_N/x_2) \\ \dots & \dots & \dots & \dots \\ p(y_1/x_M) & p(y_2/x_M) & \dots & p(y_N/x_M) \end{bmatrix}$$

$$\sum_{i=1}^M p(y_j/x_i) \log p(y_j/x_i) = \text{常数} \quad j=1,2,\dots,N$$

$$\sum_{j=1}^N p(y_j/x_i) \log p(y_j/x_i) = \text{常数} \quad i=1,2,\dots,M$$

离散无记忆信道的特性

(1) 离散无记忆对称信道的条件熵满足：

$$H(Y/X) = -\sum_{j=1}^N p(y_j/x_i) \log p(y_j/x_i)$$

因为：

$$\begin{aligned} H(Y/X) &= -\sum_{i=1}^M \sum_{j=1}^N p(x_i y_j) \log p(y_j/x_i) \\ &= -\sum_{i=1}^M \sum_{j=1}^N p(x_i) p(y_j/x_i) \log p(y_j/x_i) \\ &= -\sum_{i=1}^M p(x_i) \sum_{j=1}^N p(y_j/x_i) \log p(y_j/x_i) \\ &= -\left[\sum_{j=1}^N p(y_j/x_i) \log p(y_j/x_i) \right] \cdot \left[\sum_{i=1}^M p(x_i) \right] \\ &= -\sum_{j=1}^N p(y_j/x_i) \log p(y_j/x_i) \end{aligned}$$

同理可得

$$H(X/Y) = -\sum_{i=1}^M p(x_i/y_j) \log p(x_i/y_j)$$

条件熵取值仅由信道特性决定，与信源的统计特性无关。

离散无记忆信道的特性

(2) 若输入信道的信源符号等概

$$p(x_i) = \frac{1}{M}, \quad i = 1, 2, \dots, M$$

则信道的输出符号也等概

$$\begin{aligned} p(y_j) &= \sum_{i=1}^M p(x_i y_j) = \sum_{i=1}^M p(x_i) p(y_j / x_i) = \sum_{i=1}^M \frac{1}{M} p(y_j / x_i) \\ &= \frac{1}{M} \sum_{i=1}^M p(y_j / x_i) = \frac{1}{N}, \quad j = 1, 2, \dots, N \end{aligned}$$

离散无记忆信道的特性

信道容量：

$$\begin{aligned} C &= \max_{p(x_i), i=1,2,\dots,M} [I(X;Y) \times R_S] = \max_{p(x_i), i=1,2,\dots,M} [H(X) - H(X/Y)] \times R_S \\ &= \max_{p(x_i), i=1,2,\dots,M} [H(X) - \text{常数}] \times R_S = \left[\max_{p(x_i), i=1,2,\dots,M} H(X) - \text{常数} \right] \times R_S \\ &= (\log M - \text{常数}) \times R_S \end{aligned}$$

对于离散无记忆对称信道，若要使信息传输速率达到信道容量，要求信源的符号等概分布。

对于非等概的信源，可设法对其输出的符号进行适当的组合/变换，使得重新构建后的符号(信源)，具有近似等概的分布特性。

连续信源的信息熵

$$\begin{aligned} H(X) &= \lim_{\Delta \rightarrow 0, n \rightarrow \infty} H(X_n) = \lim_{\Delta \rightarrow 0, n \rightarrow \infty} \left\{ - \sum_{i=1}^n p(x_i) \Delta \log[p(x_i) \Delta] \right\} \\ &= - \lim_{\Delta \rightarrow 0, n \rightarrow \infty} \left[\sum_{i=1}^n p(x_i) \Delta \log p(x_i) \right] - \lim_{\Delta \rightarrow 0, n \rightarrow \infty} \left[\sum_{i=1}^n p(x_i) \Delta \log \Delta \right] \\ &= - \lim_{\Delta \rightarrow 0, n \rightarrow \infty} \left[\sum_{i=1}^n p(x_i) \log p(x_i) \right] \Delta - \lim_{\Delta \rightarrow 0, n \rightarrow \infty} \log \Delta \\ &= - \int_a^b p(x) \log p(x) dx - \lim_{\Delta \rightarrow 0, n \rightarrow \infty} \log \Delta \\ &= - \int_a^b p(x) \log p(x) dx + \infty \end{aligned}$$

有什么启示？

可见连续信源的熵无限大。该熵称为连续信源的绝对熵。

连续信源的绝对熵，无法确切地定义。

连续信源的相对熵

相对熵定义

$$h(X) = -\int_a^b p(x) \log p(x) dx$$

[示例] 某信号的相对熵如下

$$h_1(X) = -\int_a^b p(x) \log p(x) dx = -\int_{-1}^1 \frac{1}{2} \log \frac{1}{2} dx = 1$$

若把信号放大两倍，相对熵变成多少？

这说明什么问题？

信号的简单放大并没有增加任何新的信息，但其相对熵发生了增大的变化，这说明相对熵已经不再具有信源平均信息量的内涵。

连续信源的条件熵

$$\begin{aligned} H(X/Y) &= \lim_{\Delta \rightarrow 0, \delta \rightarrow 0} H(X_n/Y_m) \\ &= \lim_{\Delta \rightarrow 0, \delta \rightarrow 0} \left\{ - \sum_{i=1}^n \sum_{j=1}^m p(x_i y_j) [\log p(x_i / y_j)] \Delta \delta - \log \Delta \right\} \\ &= - \int_{a_X}^{b_X} \int_{a_Y}^{b_Y} p(xy) \log p(x/y) dx dy - \lim_{\Delta \rightarrow 0, \delta \rightarrow 0} \log \Delta \\ &= - \int_{a_X}^{b_X} \int_{a_Y}^{b_Y} p(y) p(x/y) \log p(x/y) dx dy - \lim_{\Delta \rightarrow 0} \log \Delta \end{aligned}$$

可见？

连续信源的条件熵取值无限大，同样无法确切定义。但通常上式的第一项是一个有限取值的量。

连续信源的熵和条件熵均取值无限大，说明要在一个容量有限的通信系统中传递连续信源的全部信息是不可能的。

从离散信源的到连续信源

离散信道可以看作一种抽象的信道，对符号进行判决后统计的结果，实际的物理信道是一种连续的信道。

若已知随机信号 $X(t)$ 幅度取值的概率密度函数： $p(x)$

取值在任意小区间 $[x_i, x_i + \Delta x_i] = [a_X + (i-1)\Delta, a_X + i\Delta]$ 内的概率

$$P(x_i) = \int_{a_X + (i-1)\Delta}^{a_X + i\Delta} p(x) dx \approx p(x_i) \Delta \quad i = 1, 2, \dots, n$$

连续信源转变为具有 n 个随机变量的信源，且有

$$\sum_{i=1}^n P(x_i) = \sum_{i=1}^n p(x_i) \Delta \approx \sum_{i=1}^n \int_{a_X + [i-1]\Delta}^{a_X + i\Delta} p(x) dx = \int_a^b p(x) dx = 1$$

利用离散随机变量熵的定义，得

$$H(X_n) = -\sum_{i=1}^n P(x_i) \log P(x_i) = -\sum_{i=1}^n p(x_i) \Delta \log [p(x_i) \Delta]$$

连续信源的相对条件熵

$$h(X/Y) = - \int_{a_X}^{b_X} \int_{a_Y}^{b_Y} p(y) p(x/y) \log p(x/y) dx dy$$

相对熵和相对条件熵的差值与普通的熵和条件熵的差值一样，仍然等于平均互信息量：

$$\begin{aligned} I(X;Y) &= H(X) - H(X/Y) = \left[h(X) - \lim_{\Delta \rightarrow 0} \log \Delta \right] - \left[h(X/Y) - \lim_{\Delta \rightarrow 0} \log \Delta \right] \\ &= h(X) - h(X/Y) \end{aligned}$$

同理可以导出：

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

信源编码

信源编码的目的是什么？

提高传输效率

1. 去除消息中的冗余度，使传输的符号尽可能都是独立的，没有多余的成分(如语音、图像信号压缩)；
2. 使传输的符号所含的信息最大化。例如，通过编码使符号以等概分布的形式出现，使每个符号可能携带的信息量达最大；
3. 提高编码/传输效率。采用不等长编码，让出现概率大的符号用较短的码元序列表示，对概率小的符号用较长的码元序列；
4. 在允许一定失真的条件下，如何实现高效率的编码。

离散无记忆信源

离散无记忆信源的输出序列：

$\dots X_{-2} X_{-1} X_0 X_1 X_2 \dots X_n \dots$ 各个符号间彼此独立

其中

$$X_i \in S \quad \{S : S_i, i = 1, 2, \dots, L\}$$

反之，若输出的各符号间有一定的相关性，则其作为一种有记忆的信源。

有记忆的信源，经过处理后，有可能变为一种无记忆的信源。如有记忆的信源，经过理想的、完全去除冗余的压缩编码后产生的输出。

离散无记忆信源的编/译码

若将信源输出的符号按每 J 个为一组进行编码，则任意的第 m 个分组可以表示为

$$\bar{X}_J^{(m)} = (X_1^{(m)} X_2^{(m)}, \dots, X_J^{(m)}) \quad X_k^{(m)} \in \{S : S_i, i = 1, 2, \dots, L\} \text{ (信源符号集)}$$

编码输出

$$\bar{Y}_{n_J}^{(m)} = C(\bar{X}_J^{(m)})$$

$$\text{其中 } \bar{Y}_{n_J}^{(m)} = (Y_1^{(m)} Y_2^{(m)} \dots Y_{n_J}^{(m)}) \quad Y_k^{(m)} \in \{C : C_i, i = 1, 2, \dots, D\}$$

$\{C : C_i, i = 1, 2, \dots, D\}$ 为输出的码元集。

接收端的译码输出

$$\bar{X}_J^{(m)} = C^{-1}(\bar{Y}_{n_J}^{(m)}) \quad \text{怎么定义等长编/译码?}$$

离散无记忆信源的编/译码



待编码码组简单记为

$$\bar{X}_J = (X_1 X_2, \dots, X_J) \quad X_k \in \{S : S_i, i = 1, 2, \dots, L\}$$

编码输出码组(码字)

$$\bar{Y}_{n_J} = (Y_1 Y_2 \dots Y_{n_J}) \quad Y_k \in \{C : C_i, i = 1, 2, \dots, D\}$$

译码后，存在虽然经过的是无扰信道，但仍存在 $\bar{X}'_J \neq \bar{X}_J$ 的可能性，是因为编码后的码字，有可能不是唯一可译的。

唯一可译码

若待编码的符号序列的不同组合个数为

$$|\bar{X}_J| = L^J$$

码字集中不同的码字个数

$$|\bar{Y}_{n_J}| = D^{n_J}$$

编码输出为唯一可译码的(必要)条件

$$D^{n_J} \geq L^J$$

对于码元取自 $\{C : C_i, i = 1, 2, \dots, D\}$, 码字长度为 n_J 一个码字,
其最大可携带的信息量(由最大熵定理)为 $n_J \log_2 D$

编码速率与编码效率

编码表示一个信源符号所需的平均信息量的定义为**编码速率** R

等长编码的编码速率

$$R = \frac{1}{J} (n_J \log_2 D)$$

不等长编码的编码速率

$$R = \frac{1}{J} (\bar{n}_J \log_2 D)$$

其中 \bar{n}_J 为不等长编码的平均码长。

编码效率为信源熵与编码速率的比值。

离散无记忆信源的等长编码

单个符号独立编码 采用 D 进制码元编码

若信源符号集有 L 种符号，要保证译码的惟一性，由

$$D^{n_J} \geq L^J \quad J=1$$

码字长度应取 $n_J = n_1 \geq \frac{\log_2 L}{\log_2 D}$

取整得 $n_1 = \begin{cases} \log_2 L / \log_2 D & \text{若 } \log_2 L / \log_2 D \text{ 为整数} \\ \lfloor \log_2 L / \log_2 D \rfloor + 1 & \text{若 } \log_2 L / \log_2 D \text{ 不为整数} \end{cases}$

编码效率： $\eta_c = \frac{H(S)}{R} = \frac{H(S)}{n_1 \log_2 D}$

离散无记忆信源的等长编码

扩展编码(联合编码): 将 J 个信源符号进行联合编码

由译码唯一性要求

$$D^{n_J} \geq L^J \quad \rightarrow \quad n_J \geq \frac{J \log_2 L}{\log_2 D}$$

取整得

$$n_J = \begin{cases} J \log_2 L / \log_2 D & \text{若 } J \log_2 L / \log_2 D \text{ 为整数} \\ \lfloor J \log_2 L / \log_2 D \rfloor + 1 & \text{若 } J \log_2 L / \log_2 D \text{ 不为整数} \end{cases}$$

平均每个符号所需的码元数

$$n_1 = \frac{N_J}{J} = \begin{cases} J \log_2 L / \log_2 D & \text{若 } J \log_2 L / \log_2 D \text{ 为整数} \\ \frac{\lfloor J \log_2 L / \log_2 D \rfloor}{J} + \frac{1}{J} & \text{若 } J \log_2 L / \log_2 D \text{ 不为整数} \end{cases}$$

$J \uparrow \longrightarrow \frac{1}{J} \rightarrow 0$ J 取值的增大有利于效率的提高。

离散无记忆信源的等长编码

采用二进制码元分别对 $J=4$ 的两信源符号序列进行编码。

$$\begin{array}{lcccccccc} S: & S_1 & S_2 & S_3 & S_4 & S_5 & S_6 & S_7 & S_8 \\ p(S): & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \end{array}$$

$$\begin{array}{lcccccccc} S': & S'_1 & S'_2 & S'_3 & S'_4 & S'_5 & S'_6 & S'_7 & S'_8 \\ P(S'): & 1/16 & 1/16 & 1/16 & 1/16 & 1/8 & 1/8 & 1/4 & 1/4 \end{array}$$

离散无记忆信源的等长编码

信源1:

$$\begin{array}{lcccccccc} S: & S_1 & S_2 & S_3 & S_4 & S_5 & S_6 & S_7 & S_8 \\ p(S): & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \end{array}$$

$$H(S) = -\sum_{i=1}^8 P(S_i) \log_2 P(S_i) = -8 \times \left(\frac{1}{8} \log_2 \frac{1}{8} \right) = 3$$

$$\text{因为 } n_J = \frac{J \log_2 L}{\log_2 D} = \frac{4 \log_2 8}{\log_2 2} = 12 \quad \text{故可取 } n_J = 12$$

$$\text{平均每个符号的码元数 } n_1 = n_J / 4 = 12 / 4 = 3$$

$$\text{编码效率 } \eta_c = \frac{H(S)}{n_1 \log_2 D} = \frac{3}{3 \times 1} = 1$$

离散无记忆信源的等长编码

信源2:

S' :	S'_1	S'_2	S'_3	S'_4	S'_5	S'_6	S'_7	S'_8
$P(S')$:	1/16	1/16	1/16	1/16	1/8	1/8	1/4	1/4

$$\begin{aligned} H(S') &= -\sum_{i=1}^8 P(S'_i) \log_2 P(S'_i) \\ &= -4 \times \left(\frac{1}{16} \log_2 \frac{1}{16} \right) - 2 \times \left(\frac{1}{8} \log_2 \frac{1}{8} \right) - 2 \times \left(\frac{1}{4} \log_2 \frac{1}{4} \right) = \frac{11}{4} \end{aligned}$$

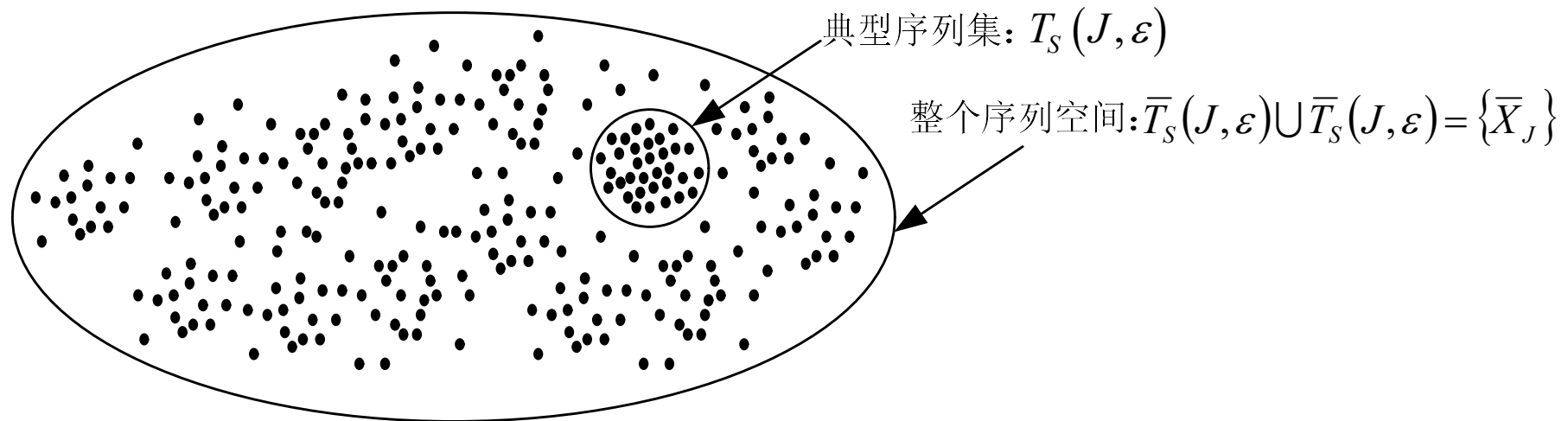
同理有 $\frac{J \log_2 L}{\log_2 D} = \frac{4 \log_2 8}{\log_2 2} = 12$ 取 $n_J = 12$

平均每个符号的码元数 $n_1 = n_J / 4 = 12 / 4 = 3$

编码效率

$$\eta'_c = \frac{H(S')}{n_1 \log_2 D} = \frac{11/4}{3 \times 1} = \frac{11}{12}$$

离散无记忆信源的有损等长编码



研究典型序列集特性的意义

如果容许一定的失真，只对典型序列编码，对非典型序列不予编码传输，码字长度可大大缩短，从而有效提高传输效率。

什么是典型序列？什么是非典型序列？

离散无记忆信源的有损等长编码

序列集	序列出现概率 $P(S_i S_j S_k S_l)$	序列包含信息量 $I(S_i S_j S_k S_l)$	序列包含平均信息量 $I(S_i S_j S_k S_l) / 4$
$S_0 S_0 S_0 S_0$	0.4096	1.2877	0.3219
$S_0 S_0 S_0 S_1$	0.1024	3.2877	0.8219
$S_0 S_0 S_1 S_0$	0.1024	3.2877	0.8219
$S_0 S_0 S_1 S_1$	0.0256	5.2876	1.3219
$S_0 S_1 S_0 S_0$	0.1024	3.2877	0.8219
$S_0 S_1 S_0 S_1$	0.0256	5.2876	1.3219
$S_0 S_1 S_1 S_0$	0.0256	5.2876	1.3219
$S_0 S_1 S_1 S_1$	0.0064	7.2877	1.8219
$S_1 S_0 S_0 S_0$	0.1024	3.2877	0.8219
$S_1 S_0 S_0 S_1$	0.0256	5.2876	1.3219
$S_1 S_0 S_1 S_0$	0.0256	5.2876	1.3219
$S_1 S_0 S_1 S_1$	0.0064	7.2877	1.8219
$S_1 S_1 S_0 S_0$	0.0256	5.2876	1.3219
$S_1 S_1 S_0 S_1$	0.0064	7.2877	1.8219
$S_1 S_1 S_1 S_0$	0.0064	7.2877	1.8219
$S_1 S_1 S_1 S_1$	0.0016	9.2877	2.3219

离散无记忆信源的不等长编码

等长编码 VS 不等长编码
优缺点？

等长容易切分（须展开）

不等长能从统计意义提升效率（须展开）

不等长编码的平均码长怎么算？

两种不等长编码方式

[逗点码]

每个码字都含有一个特定的符号用于标识

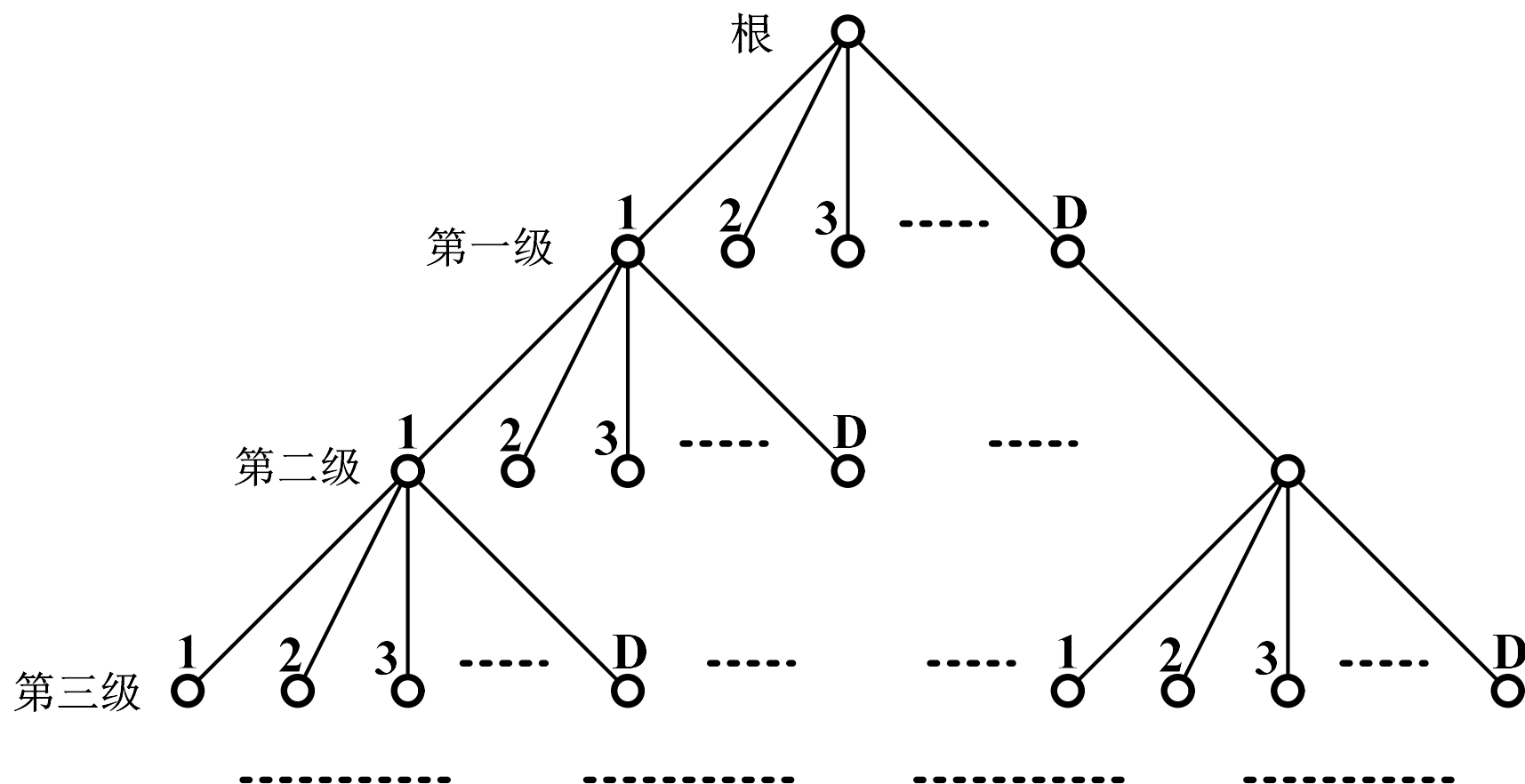
如： 0, 01, 011, 0111

[异字头码]/[异前缀码]

若码字集中任一码字都不是另一码字的字头

两者的优缺点？

树编码实现异字头码



动手试一下~~~/////如何让树编码的效率最大?

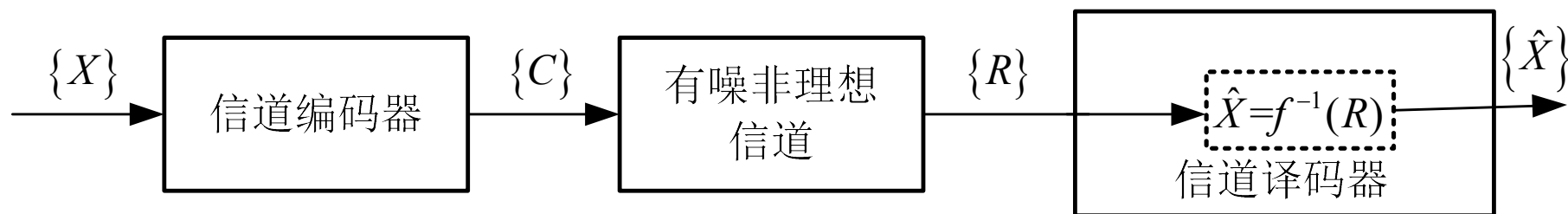
异头字码的示例

$S:$	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
$P(S):$	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$

消息符号	符号概率 $p(S_i)$	第一次划分 (编码)	第二次划分 (编码)	第三次划分 (编码)	码字
S_1	$1/3$	$1/3 (0)$			0
S_2	$1/9$	$1/3 (1)$	$1/9$	$1/9 (0)$	10
S_3	$1/9$		$1/9$	$1/9 (1)$	11
S_4	$1/9$		$1/9$	$1/9 (2)$	12
S_5	$1/9$	$1/3 (2)$	$1/9$	$1/9 (0)$	20
S_6	$1/9$		$1/9$	$1/9 (1)$	21
S_7	$1/27$		$1/27$	$1/27 (0)$	220
S_8	$1/27$		$1/27$	$1/9 (2)$ $1/27 (1)$	221
S_9	$1/27$		$1/27$	$1/27 (2)$	222

信道编码

理想信道模型



信道相关概念回顾

$$P(R) = \begin{pmatrix} p(r_1) & p(r_2) & \dots & p(r_N) \end{pmatrix}$$

$$[P(R/C)] = \begin{bmatrix} p(r_1/c_1) & p(r_2/c_1) & \dots & p(r_N/c_1) \\ p(r_1/c_2) & p(r_2/c_2) & \dots & p(r_N/c_2) \\ \dots & \dots & \dots & \dots \\ p(r_1/c_M) & p(r_2/c_M) & \dots & p(r_N/c_M) \end{bmatrix}$$

$$[P(C/R)] = \begin{bmatrix} p(c_1/r_1) & p(c_1/r_2) & \dots & p(c_1/r_N) \\ p(c_2/r_1) & p(c_2/r_2) & \dots & p(c_2/r_N) \\ \dots & \dots & \dots & \dots \\ p(c_M/r_1) & p(c_M/r_2) & \dots & p(c_M/r_N) \end{bmatrix}$$

两种信道译码准则

- 最大似然译码准则

$$p(\hat{c}_i = c_i / r_j) = \max \left(p(c_1 / r_j), p(c_2 / r_j), \dots, p(c_M / r_j) \right), \quad j = 1, 2, \dots, N$$

- 最大后验概率译码准则

$$p(c_i / r_j) = \max_{k=1,2,\dots,M} p(c_k / r_j)$$

两者有什么**优劣**和**关联**？

End