信息论

写明公式！！！

一、离散信息的度量+离散信源及信源熵

根据所编函数计算如下信源的自信息量及信源熵。

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \end{bmatrix}$$

17. 给定随机变量 X 和 Y 的联合概率：

	$x=0$	$x=1$
$y=0$	$1/3$	$1/3$
$y=1$	0	$1/3$

计算：（1）熵 $H(X)$、$H(Y)$；（2）条件熵 $H(X|Y)$、$H(Y|X)$；（3）联合熵 $H(XY)$；（4）$H(Y)-H(Y|X)$；（5）互信息量 $I(X;Y)$。

4、有两个二元随机变量 X 和 Y，它们的联合概率分布函数如下表：

Y / X	0	1
0	1/8	3/8
1	3/8	1/8

条件概率公式：$p(x\mid y)=\dfrac{p(xy)}{p(y)}$，试求：

a、熵 H(X),联合熵 H(X,Y);

b、条件熵 H(X|Y);

c、平均互信息 I(X;Y);

维拉图

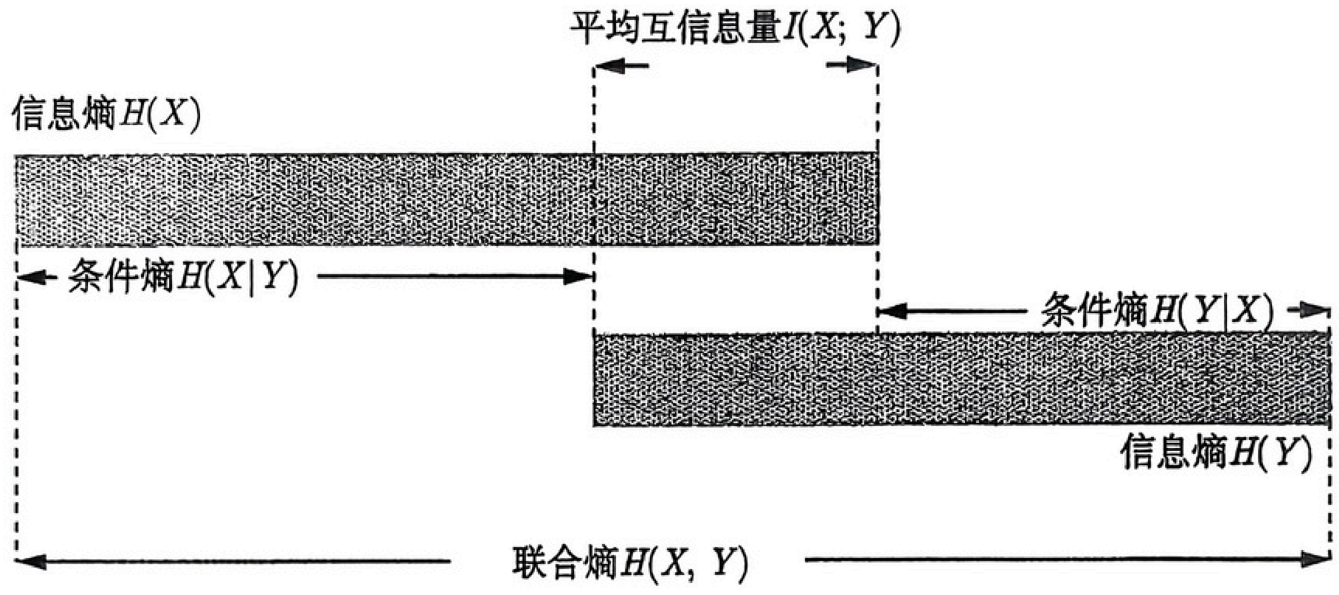

图 2.11 平均互信息量与熵关系的维拉图

各种熵之间的关系

二、信源编码（哈夫曼编码）

对给定信源进行二进制霍夫曼编码

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \end{bmatrix}$$

2、给出下面信源的霍夫曼编码（手写出每一步的编码过程，并用程序验证结果的准确性）

$$\begin{pmatrix} U \\ P \end{pmatrix} = \begin{bmatrix} u_1 & u_2 & u_3 & u_4 & u_5 & u_6 & u_7 & u_8 \\ 0.1 & 0.18 & 0.4 & 0.05 & 0.06 & 0.10 & 0.07 & 0.04 \end{bmatrix}$$

分类图

信源编码分类，如图 5.1所示。

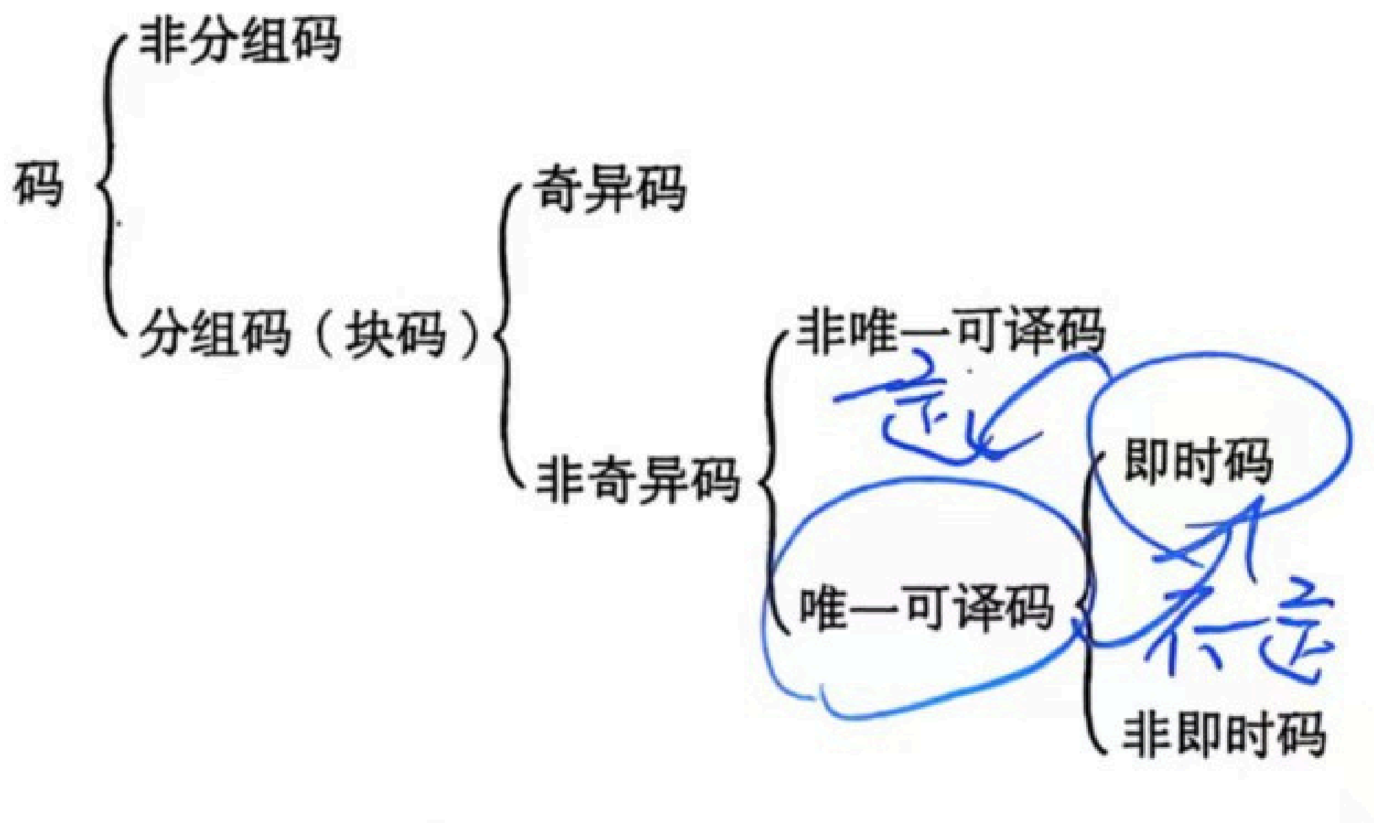

图 5.1　信源编码分类

唯一可译码 不一定 是即时码

即时码 一定是 唯一可译码

三、信道编码

①给定监督矩阵求生成矩阵

②给定信息序列做编码

③收到码字做校验

已知（7，4）汉明码的监督矩阵为 $[H]=\begin{bmatrix}0&1&1&1&1&0&0\\1&0&1&1&0&1&0\\1&1&0&1&0&0&1\end{bmatrix}$

（1）给出其生成矩阵；

（2）当有一序列 110101010001……送入时，写出编码器编出的汉明码序列；

（3）若收到码字为 11101010101010，请给出译码过程和结果。

关系：
汉明监督次数与码字长度
最小码距与纠检错能力

✅ 一、汉明监督次数（校验位数 r）与码字长度 n 的关系

结论（必须背）：

$$2^r \geq n+1$$

其中

- r：汉明监督位（校验位）个数
- n：码字长度（信息位 k + 校验位 r）

含义：

监督位必须足够多，能够**唯一标识所有 n 个可能出错的位置 + 表示无错的 1 个状态。**

码的最小距离越大 平均错误概率越小

✅ 二、最小码距 d 与纠错 / 检错能力的关系

（1）最大可检错误数：

$$\text{可检错能力} = d-1$$

（2）最大可纠错误数：

$$\text{可纠错能力} = t = \left\lfloor \frac{d-1}{2} \right\rfloor$$

解释：

- 最小码距越大，码字越“分散”，错误不易跨越到另一个合法码字。
- 所以能纠的越多，能检的越多。