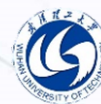


# 平均互信息量及其性质

武汉理工大学

Information theory  
and  
coding



武汉理工大学

## »»» 内容提要

---

### 第2章 信息的度量

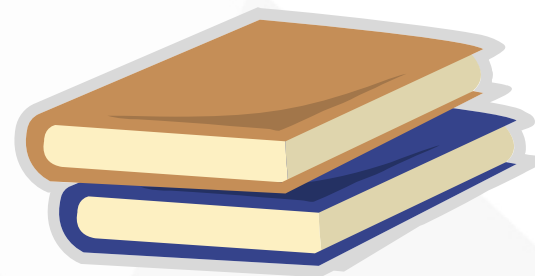
#### 2.6 平均互信息量及其性质

**知识要点：** 定义

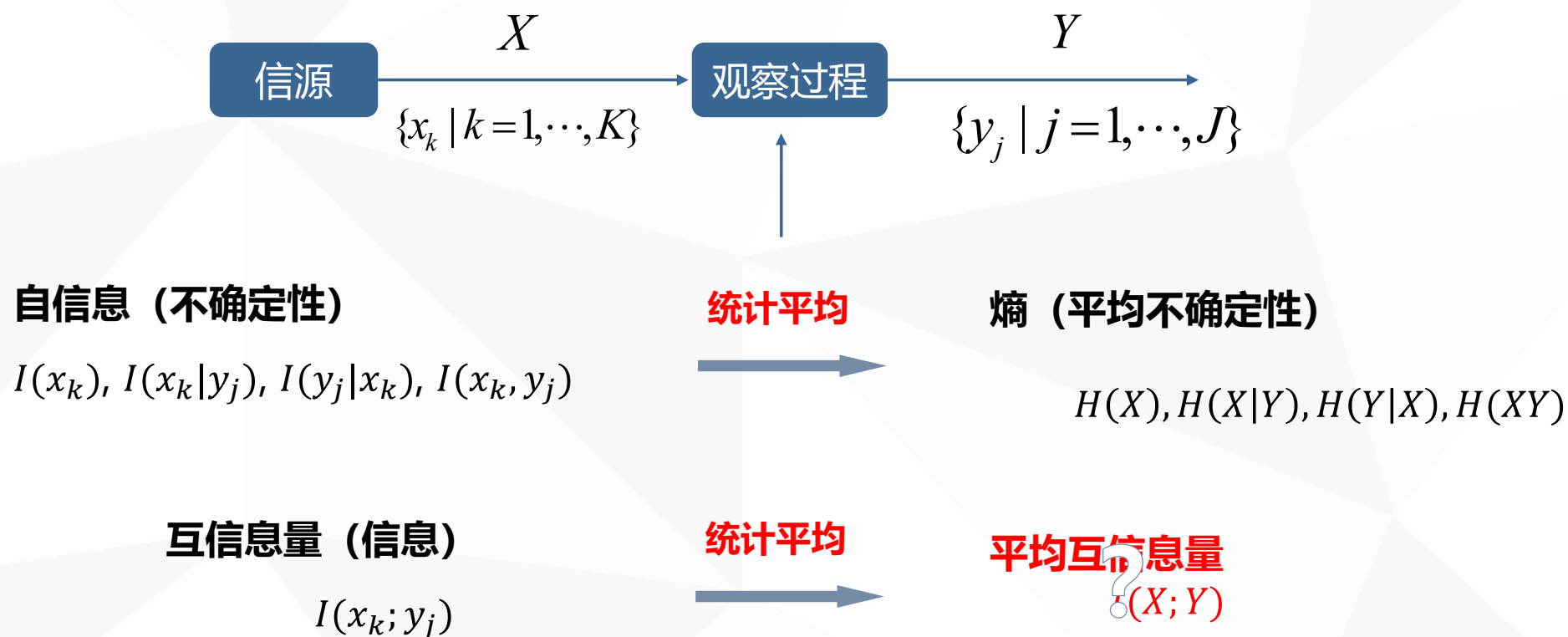
物理意义

性质

各类熵与平均互信息量的关系



## 回顾：不确定性与信息



## ➤➤➤ 平均互信息量的定义

联合概率空间:  $[XY, P_{XY}] = [(x_k, y_j), P(x_k, y_j) \mid k = 1, 2, \dots, K; j = 1, 2, \dots, J]$

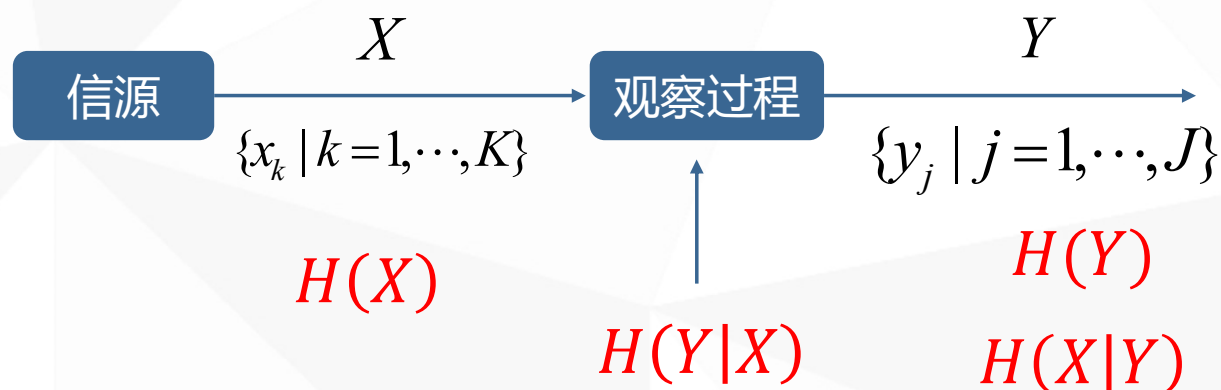
互信息量:  $I(x_k; y_j) = \log \frac{P(x_k, y_j)}{P(x_k)P(y_j)} \quad k = 1, 2, \dots, K; j = 1, 2, \dots, J$



统计平均

平均互信息量: 
$$I(X; Y) = \sum_{k=1}^K \sum_{j=1}^J P(x_k, y_j) I(x_k; y_j)$$
$$= \sum_{k=1}^K \sum_{j=1}^J P(x_k, y_j) \log \frac{P(x_k, y_j)}{P(x_k)P(y_j)} \quad \text{bit/符号}$$

## 平均互信息量的物理解释



### 1. $I(X; Y) = H(X) - H(X|Y)$ ——从输入端看

物理意义：从Y中获得的关于X的信息 $I(X; Y)$ 等于X的先验(平均)不确定性 $H(X)$ 减去X的后验(平均)不确定性 $H(X|Y)$ 。

### 2. $I(X; Y) = H(Y) - H(Y|X)$ ——从输出端看

物理意义：输出Y包含的全部信息 $H(Y)$ 包含了来自输入X的有用信息 $I(X; Y)$ 和噪声引入的无用信息 $H(Y|X)$ 。