1、消息、信号及信息

2、信息量、平均信息量

# 消息、信号及信息

# 消息、信号及信息

**消息(message)**

**通信系统传输的对象，具有不同形式和内容。**

如：气象中的温度、天气的变化，语音，

图像画面，文字及计算机数据等。

**消息是对事件的物理状态变化进行描述的一种具体形式，具有人们能感知的物理特性。**

**消息**通常不能直接传输。

为传输消息，应把消息通过不同的传感器转换为**信号**。

## 信号( signal )

**信号是由消息转换来的可以被传输和处理的具体形式，是消息的运载工具（载体）。**

## 电信号

**随某参数（通常为t）变化的电量（电压、电流、电荷量、磁通量、电场强度、磁场强度...）。**

## 信息(information)

信息是消息或信号随机变化中的"不确定性"，是消息中所含的待知的本质内容。

**确定的消息中不含有信息。**

信息的主要特征：

（1）信息无体积和重量；（2）信息易扩散和传递；

（3）信息具有依附性；（4）信息可共享。

# 信息量

# 信息量

**信息量(information content)：信息的度量。**

**信息量 I 与消息出现的概率 P(x) 之间的关系、规律：**

**（1）信息量是概率的函数。**

**（2）P(x)越小，I越大；反之，I越小。**

**“9.11”事件信息量极大。**---“竟有如此事情发生，大为震惊！”

# 信息量

**信息量(information content)：信息的度量。**

**信息量 I 与消息出现的概率 P(x) 之间的关系、规律：**

**（1）信息量是概率的函数。**

**（2）P(x)越小，I越大；反之，I越小。**

P(x)=1，I为零； P(x)=0，I为无穷大。

“11.9” 特朗普赢得美国总统大选，出人意料，情理之中。

# 信息量

**信息量(information content)：信息的度量。**

**信息量 I 与消息出现的概率 P(x) 之间的关系、规律：**

**（1）信息量是概率的函数。**

**（2）P(x)越小，I越大；反之，I越小。**

**P(x)=1，I为零；P(x)=0，I为无穷大。**

**（3）若干个互相独立事件构成的消息，所含信息量等于各独立事件信息量之和，也就是说，信息具有可加性。**

# 信息量

**信息量I与消息出现的概率P(x)之间的关系为：**

$$I = \log_a \frac{1}{P(x)} = -\log_a P(x)$$

**即：消息包含的信息量为消息出现概率的倒数的对数。**

# 信息量

**信息量单位：**

a=2时，信息量单位为比特（bit---binary unit）；

a=e时，信息量的单位为奈特（nat ---nature unit ）；

a=10时，信息量的单位为哈特莱（Hartley or decit）。

# 平均信息量

# 平均信息量

一般地说，离散信源产生多个独立的消息（符号），每个消息发生的概率不同，故包含的信息量也不相同。

$$
\begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ p(x_1) & p(x_2) & p(x_3) & \cdots & p(x_n) \end{bmatrix}
$$

描述离散信源的方法 --- 概率场

$$\sum_i p(x_i) = 1 \quad \text{(完备性)}$$

这时，通常考虑信源的**统计平均信息量**（信源的**平均信息量**或**信源熵**）。

# 平均信息量

西安交通大学 XI'AN JIAOTONG UNIVERSITY

单个独立离散消息的信息量为

$$I_{x_i} = \log_2 \frac{1}{p(x_i)} = -\log_2 p(x_i)$$

独立等概时信息量

若有***M***个独立等概出现的消息，每个消息出现的概率为**1/M**，则每个消息的信息量为

$$I = \log_2\left(\frac{1}{P}\right) = \log_2 1 \Big/ \frac{1}{M} = \log_2 M$$

# 平均信息量

$M$=2时，每个消息出现的概率为1/2，
则消息的信息量为

$$I = -\log_2 1 \Big/ \frac{1}{2} = \log_2 2 = 1bit$$

$$\begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$$

**结论：**一个等概二进制消息（符号）包含的信息量为1 *bit* 。

工程上定义一个二进制符号包含的信息量为1比特。

# 平均信息量

若有M个独立等概的消息之一要传送，

且满足 $M=2^k$ 时，此消息可用 $k$ 个二进制

符号（脉冲）传递，即此消息包含的信息量

为 $k$ 比特。

$$\boxed{I=\log_2 M=\log_2 2^k=k}$$ 比特

# 平均信息量

独立等概时，信源平均信息量与单个离散消息的信息量相同。

# 平均信息量



## 独立不等概时平均信息量

$$\begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ p(x_1) & p(x_2) & p(x_3) & \dots & p(x_n) \end{bmatrix}$$

$$p(x_1) \neq p(x_2) \neq \dots \neq p(x_n) \qquad \sum_i p(x_i) = 1$$

消息 $x_i$ 包含的信息量为： $I_i = -\log_2 p(x_i)$

平均信息量（信源熵 ***Entropy***）由每个消息的信息量按概率加权求和得到，即：

$$\boxed{H(x) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i)} \quad \text{bit/符号}$$

# 平均信息量

例1.1：某信源的符号集由*A*、*B*、*C*、*D*和*E* 组成，设每一个符号独立出现，出现的概率分别为1/4、1/8、 1/8、3/16和5/16。试求该信源的平均信息量。

解：信源的平均信息量，即信源熵，为

$$H(x) = -\sum_{i=1}^{n} P(x_i)\log_2 P(x_i)$$

$$= -\left[\frac{1}{4}\log_2\frac{1}{4} + \frac{1}{8}\log_2\frac{1}{8} + \frac{1}{8}\log_2\frac{1}{8} + \frac{3}{16}\log_2\frac{3}{16} + \frac{5}{16}\log_2\frac{5}{16}\right]$$

$$=2.23 \text{ 比特/符号}$$

# 平均信息量

西安交通大学 XI'AN JIAOTONG UNIVERSITY

## 条件平均信息量

当信源各符号出现不独立（统计相关）时，必须用条件概率来计算信源的平均信息量，称为条件平均信息量。

转移概率矩阵

$$\begin{bmatrix} P(x_1/x_1) & P(x_1/x_2) & \dots & P(x_1/x_n) \\ P(x_2/x_1) & P(x_2/x_2) & \dots & P(x_2/x_n) \\ & & \dots & \\ P(x_n/x_1) & P(x_n/x_2) & \dots & P(x_n/x_n) \end{bmatrix}$$

其中：$P(x_j/x_i)$ 为前一符号为 $x_i$，后一符号为 $x_j$ 的条件概率。

# 平均信息量

XI'AN JIAOTONG UNIVERSITY

## 条件平均信息量为：

$$H\left(x_j/x_i\right)=\sum_{i=1}^{n}P\left(x_i\right)\sum_{j=1}^{n}\left[-P\left(x_j/x_i\right)\log_2 P\left(x_j/x_i\right)\right]$$

$$=-\sum_{i=1}^{n}\sum_{j=1}^{n}\left[P\left(x_i\right)P\left(x_j/x_i\right)\log_2 P\left(x_j/x_i\right)\right]$$

其中： $P(x_i)$ 为符号 $x_i$ 出现概率

$P(x_j/x_i)$ 为前一符号为 $x_i$, 后一符号为 $x_j$ 的条件概率

# 平均信息量

**例1.2** 某离散信源由A、B两种符号组成，其转移概率矩阵为

$$\begin{bmatrix} P(A/A) & P(A/B) \\ P(B/A) & P(B/B) \end{bmatrix} = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix}$$

已知$P(A)=1/4$，$P(B)=3/4$，试求该信源的平均信息量。

**解：** 由条件平均信息量计算式可得

$$\begin{aligned} H(x_j/x_i) &= -\sum_{i=1}^{2}\sum_{j=1}^{2}\left[P(x_i)P(x_j/x_i)\log_2 P(x_j/x_i)\right] \\ &= -P(A)\left[P(A/A)\log_2 P(A/A)+P(B/A)\log_2 P(B/A)\right] \\ &\quad -P(B)\left[P(A/B)\log_2 P(A/B)+P(B/B)\log_2 P(B/B)\right] \\ &= 0.532 \quad \text{比特/符号} \end{aligned}$$

# 平均信息量



当A、B两个符号独立出现时，信源的平均信息量为

（P(A)=1/4，P(B)=3/4）

$$
\begin{aligned}
H(x) &= -\sum_{i=1}^{2} P(x_i)\log_2 P(x_i) \\
&= -\frac{1}{4}\log_2\frac{1}{4} - \frac{3}{4}\log_2\frac{3}{4} \\
&= 0.81 \quad \text{比特/符号}
\end{aligned}
$$

**结论：** 符号间统计独立时的信源熵大于符号间统计相关时的信源熵。

# 平均信息量

独立且等概时平均信息量

$$H(x) = -\sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{1}{n} = \log_2 n$$

结论：信源符号独立等概出现时的平均信息量最大。