信息论及其应用

仇佩亮　编著

浙江大学出版社

信息论及其应用

仇佩亮　编著

浙江大学出版社

图书在版编目（CIP）数据

信息论及其应用 / 仇佩亮编著. —杭州：浙江大学出版社，1999.7
ISBN 7-308-02124-6

Ⅰ.信… Ⅱ.仇… Ⅲ.信息论 Ⅳ.G201

中国版本图书馆 CIP 数据核字（1999）第 26282 号

责任编辑 王文文
封面设计 宋纪浔
出版发行 浙江大学出版社
（杭州浙大路 38 号 邮政编码 310027）
（网址：http://www.zjupress.com）
（E-mail：zupress@mail.hz.zj.cn）
经　　销 浙江省新华书店
排　　版 浙江大学出版社电脑排版中心
印　　刷 杭州富阳何云印刷有限公司印刷
开　　本 850mm×1168mm 1/32
印　　张 9.75
字　　数 260 千
版　　次 1999 年 7 月第 1 版
印　　次 2000 年 8 月第 2 次
印　　数 1001—3000
定　　价 11.00 元

前　言

信息论是研究信息传输和信息处理过程中一般规律的一门学科，也是现代信息科学和技术的一门基础理论。因此目前各高等院校的电子信息类专业的本科生、研究生都把信息论作为一门重要的专业基础理论课。同时，由于信息论的思想和方法已广泛地渗透到许多其他的学科，信息论的许多结果也有相当的普遍意义，因此信息论在许多其他领域，如在计算机科学、系统科学、统计学、经济学甚至在社会学中都获得了成功的应用。信息论对于从事这些相关领域工作和学习的人员来说也是极有参考价值的。

本书系作者根据多年教学实践和经验编著而成。教材主要介绍了香农(Shannon)信息论的基本概念、基本分析方法和主要结果；同时介绍了香农理论在密码学、计算理论、系统科学和统计学方面的一些应用。本书避免使用过多的抽象数学理论，用基本概率论工具阐明香农信息论中的精华——典型列理论和随机编码方法。本书吸取了 R. G. Gallager 和 T. Cover 所著的两本信息论经典著作的优点，结合编著者的教学经验，使得本书写得深入浅出，既保持理论的完整性和系统性，又概念清楚，易读好懂。

本书分十一章，前七章是信息论的基础理论，后四章是应用篇。除第一章绪论外，第二章介绍信息量的定义和性质；第三章介绍离散信源的无错压缩编码；第四章介绍信道容量和信道编码定理；第五章介绍常见的连续信道——高斯信道；第六章介绍了失真受到限制的信源压缩编码理论——率失真理论；第七章介绍有多个发射机和多个收信机的信源编码和信道编码理论，也称为多用户信息论；第八章介绍密码学，特别介绍香农信息论对密码学的应

用;第九章介绍在系统理论和信号处理中极为有用的最大信息原则和最大熵谱估计;第十章介绍类型理论及其在统计学和通用信源编码方面的应用;第十一章介绍 Kolmogorov 复杂性理论。

本书的编著是在浙江大学姚庆栋教授的推荐和支持下进行的,在此表示深切的感谢。同时作者在编著过程中得到了浙江大学信息与通信工程研究所同仁们的关心,他们对本书提出了许多宝贵的意见,在此一并表示感谢。

编著者

1999 年 4 月

目　录

第一章　绪论

信息论是应用近代概率统计方法研究信息传输、交换、存贮和处理的一门学科，也是源于通信实践发展起来的一门新兴应用科学。

信息是系统传输、交换、存贮和处理的对象，信息载荷在语言、文字、数据、图象等消息之中。在信息论中，信息和消息是紧密相关联的两个不同概念。同样一个消息，比如一张当日报纸，对于不同的人从中可获得的信息是不一样的；同样的天气预报"明天有雨"，对于干旱地区和雨量充沛地区来说其信息含量也不一样。一张纸写上几个字成为一封家信，对于收信者是家书抵万金，但对旁人可能是废纸一张。因此信息是一种奇妙的东西，它是有别于物质和能量的一种存在。信息的本质和它的科学定义是当前科学界，乃至哲学界热衷研究的课题，但信息的重要性是毋庸置疑的。控制论创始人维纳说过："要有效地生活，就要有足够多的信息"。目前社会上流行一些提法，如"信息、材料、能源是现代科学的三大支柱"、"信息、物质、能量是构成一切系统的三大要素"，这些提法充分说明了人们对信息的重要性的认识。

信息的度量是信息论研究的基本问题。从目前的研究来看，要对通常意义下的信息给出一个统一的度量是困难的。至今最为成功，也是最为普及的信息度量是由信息论创始人香农(Shannon)在他的光辉著作《通信的数学理论》[44]中提出的，是建立在概率模型上的信息度量。他把信息定义为"用来消除不确定性的东西"。用概率的某种函数来描述不确定性是自然的，所以香农用

$$I(A) = -\log P(A)$$

来度量事件 A 发生所提供的信息，其中 $P(A)$ 为事件 A 的概率。这个定义与人们的直觉经验相吻合。如果一个随机试验有 N 个可能结果或者说一个随机消息有 N 个可能值，若它们出现的概率分别为 $p_1, p_2, \cdots, p_N$，则这些事件的自信息的平均值

$$H = -\sum_{i=1}^{N} p_i \log p_i,$$

作为这个随机试验或随机消息所提供的平均信息也是合理的。H 也称为熵，这是借用于统计物理学中的一个名词。在物理学中熵是描述系统的不规则性或不确定性程度的一个物理量。

信息论所研究的通信系统基本模型如图 1.1.1 所示。

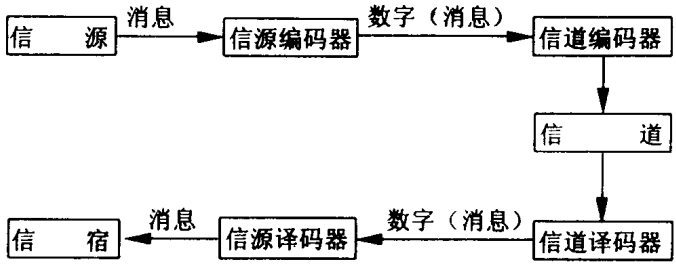

图 1.1.1　通信系统的基本模型

信源是产生消息(或消息序列)的源。消息通常是符号序列或时间函数。消息取值服从一定的统计规律，所以信源的数学模型可以是一个离散的随机序列或连续的随机过程。

信源编码器把信源产生的消息变换成数字序列。在不允许编码失真的情况下，信源编码器的目的是在保证能从其输出数字序列无错误地恢复输入消息序列的前提下，减少输出数字序列的速率；也就是保证在不失真的条件下对输入消息序列进行压缩。或者在允许失真的情况下，信源编码的目的是对给定信源，在保证消息平均失真不超过某给定允许值 D 的条件下，尽量减少输出数字序

列的速率。

信道在实际通信系统中是指传输信号的媒介或通道，如架空明线、电缆、电离层、人造卫星等。在信息论的模型中也把发送端和接收端的调制、解调器等归入到信道，并把系统中各部分的噪声和干扰都归入信道中。在信道的输入、输出模型中，根据噪声和干扰的统计特性，用输入、输出的条件概率（或称转移概率）来描述信道特性。

信道编码器把信源编码输出的数字序列变换成适合于信道传输的，由信道入口符号组成的序列。信道编码器的最主要作用是要对其输出序列提供保护，以抵抗信道噪声和干扰。

信道译码器和信源译码器分别是信道编码和信源编码的反变换，信宿是消息的接收者。

信息论解决了通信中的两个基本问题。首先对于信源编码，信息论回答了："达到不失真信源压缩编码的极限（最低）编码速率是多少？"这一问题。香农的答复是这个极限速率等于该信源的熵 H。事实上香农认为每个随机过程，不管是音乐、语言、图象，都有一个固有的复杂性，该随机过程不能被无失真地压缩到该固有复杂性之下，这个固有复杂性就等于该随机过程的熵。信息论对通信解决的第二个问题是关于信道编码方面的。在香农以前，人们都认为增加信道的信息传输速率总要引起错误概率的增加，但香农却出人意料地证明，只要信息传输速率小于信道容量 C，传输的错误概率可以任意地小，如果超过信道容量，则传输的错误是不可避免的。对每个信道可以根据它的噪声干扰特征计算出它的容量 C。

信息论除了指出通信中信源编码和信道编码的极限速率外，还提供了为达到这些极限速率的具体编码方法。然而这些方法虽然在理论上极为漂亮，但在实际上却无法实现。因为这些方法在计算上是不可实现的。随着许多通信专家的努力，以及现代电子科学技术的发展，无论对信源编码还是信道编码，目前都已有许多具体

有实用价值的编译码方案，它们的性能都逐步地向香农指出的极限性能逼近。

香农不仅解决了上述两个通信中的基本问题，而且也为通信中的一些现代课题，例如率失真理论、网络信息论、保密理论，做了奠基性工作。他提出的随机编码理论，典型列方法已成为现代信息论研究的最有效工具。

香农的信息论源于通信实践。它对通信领域的成功应用使得香农的信息论被称为通信的数学理论。但是香农理论的思想、方法甚至某些结论也已渗透到许多其他学科。

在计算机科学中，Kolmogorov 等人把数据串的复杂性定义为利用通用计算机计算出这个数据串所需的最短二元程序的长度，也就是把数据串复杂性定义为最小描述长度。Kolmogorov 证明，一个随机源按概率分布所输出的数据序列的这种复杂性近似于该随机源的熵。因而 Kolmogorov 复杂性理论与香农信息论建立了密切的联系。

最大熵准则或最大信息原则是许多科学研究中所常用的准则，实际证明这个准则是有效的、合理的。信息论赋于最大熵准则以明确的内涵。最大熵准则和最小描述长度准则都是一种科学的方法论，在信息论中可找到它们的联系。这给相信"最简单的解释是最好的"信条的人的一个佐证。

类型理论是典型列理论的翻版。利用类型理论成功地建立了通用的信源压缩算法。同时利用类型理论使我们能估计出罕见事件的概率（所谓大偏离理论）和估计假设检验的最佳错差指数。

信息论思想和方法还在经济学，生物科学等方面获得应用。

第二章　熵和互信息

§2.1　离散信源的熵和信息量

通常的信息处理系统如图 2.1.1 所示。

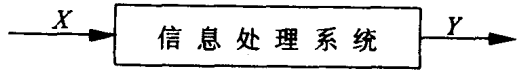

图 2.1.1　信息处理系统

其中 X 和 Y 表示输入输出随机变量，中间方框代表对输入作某种变换。令 $\mathscr{X}$ 和 $\mathscr{Y}$ 表示 X 和 Y 的取值集合。当 $\mathscr{X}$ 和 $\mathscr{Y}$ 是离散事件集合时，可用

$$\mathscr{X} = \{x_k; k = 1,2,\cdots\cdots,K\}$$

$$\mathscr{Y} = \{y_j; j = 1,2,\cdots\cdots,J\}$$

表示。对每个 $x_k \in \mathscr{X}$，相应概率为 $q(x_k) = q_k$。由概率定义：

$$\begin{cases} q_k \geqslant 0 \\ \sum\limits_{k=1}^{K} q_k = 1 \end{cases} \qquad k = 1,2,\cdots\cdots,K \tag{2.1.1}$$

我们称 $\{X,\mathscr{X},q(x)\}$ 为输入概率空间，它完全描述了随机变量 X。同样有输出概率空间 $\{Y,\mathscr{Y},\omega(y)\}$，其中

$$\begin{cases} \omega(y_j) = \omega_j \geqslant 0 \qquad j = 1,2,\cdots\cdots,J \\ \sum\limits_{j=1}^{J} \omega_j = 1 \end{cases} \tag{2.1.2}$$

随机变量 X 和 Y 的联合空间为 $\{XY, \mathscr{X} \times \mathscr{Y}, p(x,y)\}$。对每对 $(x_k, y_j) \in \mathscr{X} \times \mathscr{Y}$ 有相应的概率 $p(x_k, y_j)$，且

$$\begin{cases} \sum_k \sum_j p(x_k, y_j) = 1 \\ \sum_j p(x_k, y_j) = q(x_k) \\ \sum_k p(x_k, y_j) = \omega(y_j) \end{cases} \tag{2.1.3}$$

相应的条件概率为

$$\begin{cases} p(y_j | x_k) = \dfrac{p(x_k, y_j)}{q(x_k)} \\ p(x_k | y_j) = \dfrac{p(x_k, y_j)}{\omega(y_j)} \end{cases} \tag{2.1.1}$$

2.1.1 事件的互信息

定义 2.1.1 两个离散概率空间 $\{X, \mathscr{X}, q(x)\}$ 和 $\{Y, \mathscr{Y}, \omega(y)\}$，事件 $Y = y_j \in \mathscr{Y}$ 的出现给出事件 $X = x_k \in \mathscr{X}$ 的互信息量 $I(x_k; y_j)$，并定义为

$$I(x_k; y_j) = \log_a \frac{p(x_k | y_j)}{q(x_k)} \tag{2.1.5}$$

当 a 取 2 时，信息量单位为比特(bit)；当 a 取 e 时，单位为奈特(nat)。

由定义显然

$$\begin{aligned} I(x_k; y_j) &= \log_a \frac{p(x_k | y_j)}{q(x_k)} \\ &= \log_a \frac{p(x_k, y_j)}{q(x_k)\omega(y_j)} \\ &= \log_a \frac{p(y_j | x_k)}{\omega(y_j)} \\ &= I(y_j; x_k) \end{aligned} \tag{2.1.6}$$

因此事件的互信息具有对称性。若 y_j 的出现有助于肯定 x_k 的出