

基础信息论

戴明远 编著

同济大学出版社

需要全本请访问 www.ertongbook.com

基础信息论

戴明远　编著

同济大学出版社

图书在版编目(CIP)数据

基础信息论/戴明远编著. —上海:同济大学出版社,2003.9

ISBN 7-5608-2681-4

Ⅰ. 基… Ⅱ. 戴… Ⅲ. 信息论—高等学校—教材 Ⅳ. G201

中国版本图书馆 CIP 数据核字(2003)第 056139 号

基础信息论

戴明远 编著

责任编辑 郁 峰 谢卫奋 责任校对 徐 栩 封面设计 潘向蓁

出版发行	同济大学出版社
	(上海四平路 1239 号 邮编 200092 电话 021-65985622)
经 销	全国各地新华书店
印 刷	常熟华顺印刷有限公司
开 本	850mm×1168mm 1/32
印 张	8.75
字 数	252000
印 数	1—3000
版 次	2003 年 9 月第一版 2003 年 9 月第一次印刷
书 号	ISBN 7-5608-2681-4/TN·6
定 价	14.00 元

本书若有印装质量问题,请向本社发行部调换

前　言

数十年来，科技发展日新月异，生化技术、生命科学、计算机技术、软件技术、通信技术以及信息论的发展尤为迅速。其中，信息论最早源于通信技术，但现在的应用领域已超越原通信技术领域，成为许多新兴学科共同的理论基础之一。

信息论的一些基本理论源于20世纪40年代，例如，信源熵的定义、量化，信源编码理论，信道编码理论，后来又发展了加密、解密编码理论，使得现代的通信既高效、可靠，又具有高度的保密性，从而在军事、工业、科研甚至经济管理、日常生活中都得到了广泛的应用。信息论不再是一门抽象的、陌生的学科，而已成为一门应用广泛、日益受到重视和欢迎的基础理论课。

在信息技术越来越热门的形势下，许多高等院校和高等技术专科院校都在相关专业中开设了或准备开设信息论课程。由于现在各院校学生课程都较重，因此希望有一本篇幅较小、所需学时数不多的信息论方面的教材，以适应少学时（约36学时）教学计划的要求。

本书共分六章：第一章概论，从通信的基本模型出发，讨论信息论的研究范围和方法；第二章基本信息论，研究信息测度、信源熵、熵速率、信道容量、信源优化编码；第三章理想通信系统，介绍理想通信系统的编码定理，实现的可能性与方法，简单的伪噪声编码通信基本原理和方法；第四章抗干扰二元编码的原理及方法，介绍了抗干扰编码的基本原理，并以简单的二元码为例，介绍了几种常见的抗干扰编码法；第五章信息论与密码学，以古典密码学为例介绍了加密方法和解密方法，并介绍了系统的完全保密性、理论保密性、实际保密性的概念和分析方法；第六章信息率失真函数，介绍信息率失真函数的概念，分析方法，保真度准则下的信源编码定理，与信道编码的对偶关系等。

本书在编写过程中得到了同济大学信息工程教研室各位老师的热情帮助和支持，作者深表感谢，并希望读者对书中的错误和不足之处给以批评和指教。

作　者

2003 年 6 月

需要全本请访问 www.ertongbook.com

目　录

前言

第 1 章　概论 …… (1)

1.1　信息论的概念 …… (1)
1.2　信息的分类 …… (5)
1.3　信息论研究的目的和范畴 …… (5)
1.3.1　通信系统的基本模型 …… (5)
1.3.2　信息论研究的目的 …… (8)
1.3.3　信息论研究的范畴 …… (8)

第 2 章　基本信息论 …… (12)

2.1　信源及信源的不确定性 …… (12)
2.1.1　不确定性的概念 …… (12)
2.1.2　信源不确定度的定义 …… (15)
2.1.3　信息度量 …… (16)
2.1.4　离散信源的熵 …… (23)
2.1.5　熵函数 $H(X)$ 的性质 …… (27)
2.1.6　加权熵的概念 …… (30)
2.2　平均互信息量 …… (31)
2.2.1　平均互信息量的定义 …… (32)
2.2.2　平均互信息量的物理意义 …… (33)
2.2.3　平均互信息量的性质 …… (37)
2.3　二元联合信源的联合熵(共熵)与条件熵 …… (39)
2.3.1　共熵 …… (40)

2.3.2 条件熵…… (41)
2.3.3 离散平稳信源的信源熵和极限熵…… (45)
2.4 马尔可夫信源…… (48)
2.5 信源冗余度…… (54)
2.5.1 冗余度的定义…… (55)
2.5.2 冗余的利用…… (58)
2.6 连续信源的熵…… (59)
2.6.1 三种特定连续信源的最大熵…… (64)
2.6.2 熵功率…… (70)
2.7 熵速率和信道容量…… (71)
2.7.1 信源熵速率…… (71)
2.7.2 信道容量的概念…… (72)
2.7.3 离散有噪声信道中的熵速率和信道容量…… (74)
2.7.4 可疑度的物理解释…… (77)
2.7.5 连续有噪信道的熵速率…… (81)
2.7.6 连续有噪信道的信道容量…… (82)
2.8 多用户信道…… (84)
2.8.1 多址接入信道…… (85)
2.8.2 广播信道…… (91)
2.8.3 相关信源的多用户信道…… (93)
2.9 信源编码…… (95)
2.9.1 等长编码定理…… (96)
2.9.2 变长编码…… (98)
2.9.3 信源优化的方法及编码原理…… (99)
2.9.4 非续长代码的构成方法 …… (104)
2.10 常用信源编码…… (105)
2.10.1 山农-费诺编码 …… (106)
2.10.2 哈夫曼编码…… (107)
2.10.3 冗余位编码…… (109)

第 3 章　理想通信系统及其分析方法 …… (118)
3.1　多维空间及信号的多维模型 …… (118)
3.1.1　高维球的体积及其特点 …… (119)
3.1.2　理想通信系统的最大传信率和错误概率 …… (121)
3.1.3　编码定理的证明 …… (122)
3.2　理想通信系统的信噪比,带宽和时间的互换…… (126)
3.3　理想通信系统的实现方法 …… (130)
3.4　理想接收机 …… (132)
3.4.1　相关接收机模型 …… (133)
3.4.2　匹配滤波器 …… (136)
3.5　伪噪声编码通信 …… (137)
3.5.1　等长二元码的相关系数 …… (138)
3.5.2　等长二元码的自相关运算 …… (138)
3.6　狭义伪噪声码举例 …… (139)
3.6.1　L 序列码 …… (139)
3.6.2　m 序列码 …… (140)
3.6.3　二元正交码 …… (143)
3.7　伪噪声编码通信的原理 …… (144)
第 4 章　抗干扰二元编码原理及方法 …… (150)
4.1　抗干扰编码的基本原理 …… (150)
4.1.1　编码和纠错能力的关系 …… (150)
4.1.2　抗干扰编码的基本原理 …… (153)
4.2　构造检错码的基本方法 …… (156)
4.2.1　奇偶校验码 …… (156)
4.2.2　定比码 …… (158)
4.2.3　ARQ 系统…… (160)
4.3　构造纠错码的基本方法 …… (161)

4.3.1 简单重复码 …… (162)
4.3.2 纠正一位错误的汉明码 …… (162)
4.3.3 循环码 …… (169)
4.4 编译码方法 …… (178)
4.4.1 多项式除法电路 …… (178)
4.4.2 循环码译码电路 …… (183)
4.5 卷积码 …… (184)
4.5.1 卷积码编码器原理 …… (184)
4.5.2 卷积码的生成矩阵 …… (188)
4.5.3 卷积码编码器 …… (190)
4.5.4 卷积码的译码方法 …… (193)
4.5.5 序列译码 …… (194)
4.6 纠正突发错误的编码 …… (199)
4.6.1 纠正突发错误的必要性 …… (199)
4.6.2 分组交织法 …… (199)
4.6.3 哈格尔伯格码 …… (200)
第5章 信息论与保密学 …… (204)
5.1 保密系统的数学模型 …… (205)
5.2 古典密码体制 …… (208)
5.2.1 单表密码 …… (208)
5.2.2 移位代换密码(加法密码) …… (209)
5.2.3 乘数密码 …… (209)
5.2.4 固定周期位移置换 …… (210)
5.2.5 多表代换密码 …… (211)
5.3 完全保密性 …… (212)
5.4 理论保密性 …… (219)
5.5 实际保密性 …… (224)

第 6 章　信息率失真函数 …… (227)
6.1　概述 …… (227)
6.1.1　信息率失真函数及平均失真度 …… (228)
6.1.2　信息率失真函数的定义 …… (233)
6.1.3　信息率失真函数的性质 …… (235)
6.2　离散信源的信息率失真函数 …… (243)
6.2.1　离散信源的信息率失真函数的参量表达式 …… (243)
6.2.2　二元离散等概率信源的信息率失真函数 …… (248)
6.3　连续信源的信息率失真函数 …… (258)
6.3.1　连续信源的信息率失真函数的参量表达式 …… (258)
6.3.2　高斯连续信源的信息率失真函数 …… (260)
参考文献 …… (268)

第1章 概 论

1.1 信息论的概念

物质不灭定律、能量不灭定律和爱因斯坦的质能公式展示了宇宙中存在物质、能量以及物质和能量之间的转换关系。宇宙中除物质和能量外，尚有第三个“要素”，即信息。三者都极为重要，如果没有物质，宇宙就会变得虚无缥缈；如果没有能量，宇宙就会失去演化的动力；而如果没有信息，宇宙就会变得杂乱无章，不可理喻。可见，信息的重要性绝不亚于物质和能量。

我们的宇宙是一次大爆炸的产物，大爆炸产生了物质和能量，同时也产生了信息。信息存在的历史与物质、能量一样久远，但人们对于信息的了解，却比对于物质和能量的了解晚了许多，甚至直到今天，信息还没有一个严格的、公认的定义，这反而激起了人类的好奇心——信息到底是什么？

信息的例子很多，蜜蜂的舞蹈是一种信息，其他蜜蜂获得此信息后，就会知道附近有花蜜；果子熟了变成红色是一种信息，具有颜色分辨能力的灵长目动物可以据此轻易地在茂密的绿叶中发现它；新鲜的口味表示食物正常的信息，而腐败的口味表示食物变质的信息。听课可以获得信息，看电视可以获得信息，盲人虽然看不见，却可以依赖触觉从盲文中获得信息。非生命物质中也存在信息，原始化石中存在有生物进化的信息，硬盘中存有程序的信息，甚至物质分子的特有光谱特性也是使这种物质区别于其他物质的信息，因此，信息无处不在。

信息到底是什么，归纳起来却很困难。信息无处不在，却又似乎看不见、摸不着，说不清它的形状，道不明它的大小，虽知道它无

需要全本请访问 www.ertongbook.com

处不在，却又觉得太抽象，如果要归纳信息的特性，可以有许多，诸如：

(1) 信息与物质、能量同属于宇宙，信息伴随于物质的运动而产生，但却可以独立于物质而存在。信息与能量密切相关，但同样可以独立于能量而存在。

(2) 信息源于宇宙世界，用于(生命)精神世界，就像工具源于物质却被智慧生物制造利用一样。

(3) 信息具有伴随物质和能量的普遍性，只要有运动物质和变化能量，就有信息，而绝对静止的物质和永恒不变的能量是不存在的，故信息具有普遍性。

(4) 信息具有可度量性。如前所述，信息很抽象，但现在已有合理的度量方法，并由此推导出许多有用的结论。

(5) 信息具有可创造性。与能量不能创造不同，信息却可以创造，例如文艺创作等。

(6) 信息具有无耗性。甲有一块铁，做成一把铁锤送给乙，甲就没有铁了。但是甲拥有的信息可以传输给乙，使乙拥有与甲一样的信息，而甲原有的信息不会因此而减少。

(7) 信息具有可传递性。信息可以从一方(信源)传输到另一方(信宿)，也可以从一方传输到多方。

(8) 信息具有可存储性。信息可以存储在某些媒介上长期保存。

(9) 信息具有不增性。信息可以被加工(压缩、处理、放大、变换等)，但信息量决不会因被加工而增加，一般只会减少，至多不变。

(10) 信息具有相对性。信息的相对性是指信息价值的相对性，即信息价值与精神世界有关，对一个人有价值的信息，对另一个人可以一文不值。

(11) 信息具有时效性。信息的产生依赖于物质和能量，但信息的价值或知识性可能因为时过境迁而失效。

(12) 信息具有催化性。信息可以使资金用在刀口上，使劳动力使用合理，使原材料、能量使用节省，从而大大提高劳动生产率，用同样的物质、能量和时间，可以制造出更多的社会财富，而信息在其中仅起催化剂作用，本身并不因此减少。

(13) 信息具有可加工性。信息可以被压缩(信源编码)，被赋予抗干扰能力(信道编码)，被加密(加密编码)，从而使得信息变得有效、坚强、保密而其内涵却保持不变。

(14) 信息具有排序性。用信息可以消除系统的不确定性，相当于对紊乱的系统进行有序排序。要使一个无序系统变得较为有序或有序，就必须从外部获得较多或足够的信息。

(15) 信息具有理想性。即使是在有噪信道，信源也可以与信道容量同样的传信率无差错地传输信息。

(16) 信息的体积无限小。分子可以分成原子，原子可以分成电子、质子、中子，分割无穷，信息无穷多，但体积无限小。

(17) 信息具有独立性。信息源于物质和能量，却能独立于二者而存在。

(18) 信息具有无限性。信息的无限性指无限的信息量可以基于有限的物质，或有限的物质一定包含无限的信息量。

信息具有如此复杂而丰富多彩的特性，难怪人类对其认识要明显地落后于对物质和能量的认识了。那么，科学家是如何认识信息的呢?

1924 年，奈奎斯特(H. Nyquist)发表了“影响电报速率因素的确定”一文，1928 年，哈特莱(R. V. Hartley)发表了“信息传输”的论文，最早研究了通信系统传输信息的能力问题，把信息看成是选择通信符号的方式，而选择的自由度的大小决定信息的大小，可选择的自由度越大，信息也越大，哈特莱还把具体的信息问题退去其物理意义，抽象成单纯的数学问题，认为，只要符号数目一定，发送码字长度一定，发信者发送的信息量也就确定了，他认为，信息是选择的自由度。

1948 年，山农(C. E. Shannon)在《贝尔系统电话杂志》上发表了题为“通信的数学理论”的长篇论文，以概率论为数学工具，深入讨论了通信的基本理论问题，给出了计算信源信息量和信道容量的方法和公式，次年，又发表了“噪声中通信”这篇论文，从而创立了信息论。但山农认为“信息就是一种消息”，显然，这种定义并不确切，后来的研究表明，消息中包含有信息，可以看成信息的载体，但它又不就是信息，因此，消息与信息有关系，但绝不是等同的关系。

1950 年，维纳(N. Wiener)在《控制论与社会》一书中，从社会学的角度来给信息下定义，认为“信息是人与外界相互作用的过程中所交换的内容的名称”，这是一种推广了的概念，是信息从物质(能量)到生命再到思维(认识)的推广，从而使信息变成了无所不包的“内容”。

信息论最先应用于通信工程，随着信息论的普及，也越来越广泛地应用于经济、管理和社会科学领域。于是，对信息论的定义趋于哲学化，我国学者钟义信认为：“信息就是事物运动的状态和方式，就是关于事物运动的千差万别的状态和方式的知识”，并进一步提出了语法信息、语义信息和语用信息三个不同的定义，从而避开了定义困难的所谓“全信息”。他认为，语法信息是事物运动状态与状态改变的方式本身，是研究事物运动各种可能的状态与状态间的关系，而不管是何种具体的事物(相当于宪法)；语义信息是事物运动状态和方式的具体含义，这就是研究各种状态与实物间的关系，也就是研究信息的具体的含义(相当于民法、刑法等)；语用信息是事物运动状态和方式及其含义对观察者的效用，是研究事物运动状态对确定观察者的作用效果，即信息的主观价值(相当于法庭审判)。

遗憾的是，国际上虽然已有不下百种对信息的定义，但至今尚未形成一个公认的、确切的定义。这说明人类对信息的认识尚不够深入，尚需进一步研究，这也正是学习信息论的魅力所在。

1.2 信息的分类

信息不但难于定义，要系统而简明地分类也很困难，于是，只能从许多角度进行分类，虽缺乏整体性，但对于认识、熟悉和理解信息还是有帮助的。

(1) 按层次分类，信息可分为语法信息、语义信息和语用信息。

(2) 按时域性质分类，信息可分为离散信息和连续信息。

(3) 按符号数分类，信息可分为二元信息和多元信息。

(4) 按专业分类，信息可分为自然信息、社会信息、科技信息、文艺信息和经济信息等。

(5) 按传递方向分类，信息可分为前馈信息和反馈信息。

(6) 按真伪分类，信息可分为真实信息和虚假信息。

(7) 按价值分类，信息可分为有用信息和无用信息。

(8) 按符号概率空间分类，信息可分为概率信息、突发信息、确定信息和模糊信息。

从应用、来源、载体等许多其他角度，也都可以对信息进行细致分类，因而信息的分类也很复杂。

1.3 信息论研究的目的和范畴

1.3.1 通信系统的基本模型

信息论最早在通信方面得到广泛的应用。古人类近距离用语言传递信息，远距离用手势或烽火传递信息；现代人类用现代通信工具如电话、电报、电视、网络传递信息。因此，传递信息就是通信的目的，为了便于分析，可以将通信系统抽象成如图 1.3.1 所示的通信系统模型。

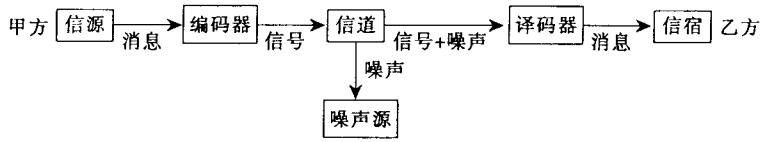


图 1.3.1　通信系统模型

图 1.3.1 所示的通信系统模型并不局限于电信系统，其他的信息流通系统，例如，生物的神经系统（听觉、视觉系统）和人类社会的管理系统等，也同样适用。信息论研究的对象已扩充到符合上述通信系统模型的任意系统，所以，信息论研究的对象，并不仅局限于通信领域，而是要比这更广泛得多。

通信系统模型共有六个部分：

（1）信源。信源是产生消息（或称符号）的源，可以是机械、电子设备、人、生物甚至是化石等。信源输出消息，消息携带有信息，故消息是信息的载体。信源输出的消息至少应为两种，何时输出何种消息是不确定的，随机的，但却是有规律的，因此，信源的特性可以用概率空间来描述。

（2）编码器。编码是把信源的原始信息按一定的理论或工程要求编成另一种消息形态，称之为信号。编码后的消息更紧凑（信源编码），更坚强（信道编码），更隐密（加密编码），在信道中传输，可以迅速而不易受到噪声的破坏，并使原始信息变得非法截取者所不可理解。不论是哪一编码，都应该使原始信息的内涵不变或基本不变。在实际通信系统中，编码器后面可能还包括换能、调制、发射等技术处理环节，以使信号能以合适的形态进入信道。

（3）信道。信道是信号从信源传输到信宿的通路，可以是各种有形的媒介，如同轴电缆、双绞线、波导、光纤等，也可以是星际真空（无线电波传播信道）。信道可以是星际真空，不局限于物质媒介，使得人类有可能利用整个宇宙的信息。电信系统是狭义的通信系统，而广义的通信系统用的信道可以是另外完全不同的媒

介，如邮政传递的书信、电报等。信道除了传送信号外，还有存储信号的作用，如电路中的延迟线等。

(4) 噪声源。噪声是无处不在的(温度高于绝对零度时)，分布于整个信道。为了研究方便，我们把信道中各处的噪声甚至系统中其他部分的噪声都等效地归纳为一个噪声源，而信道的输出就是信号与噪声的叠加，显然，这里把噪声看作是“加性噪声”。由于噪声干扰也具有随机性，因此，信道也可以用概率空间来描述，并且根据其统计特性来划分各种信道。

(5) 译码器。译码器是编码器的逆器件，译码器把信道输出的、包含干扰的信号进行反变换，以恢复信源消息。与编码器相对应，译码有信源译码、信道译码和解密译码。

(6) 信宿。信宿是消息传送的对象，信宿一般是合法的接收者。可以是机器，也可以是人。而非法的接收者，一般称为截取者或破译者，这是保密系统需要防范的对象。

图 1.3.1 的通信系统模型是一对一的单向通信(称为单工)，如果再建立一个同样的系统，用于信宿端(乙方)向信源端(甲方)发送消息，此时，甲、乙双方可同时互相传输信息(称为全双工，如果只能分时地互相传输信息，则称为半双工)。

再复杂一些的情况是一个信源对多个信宿传输信息(称为广播方式)；更复杂的情况是许多信源和信宿进行多输入、多输出，多向传输的通信，这就是网络通信系统了。但图 1.3.1 所示的通信系统模型却是最基本的模型，也是信息论研究的基本对象。

把实际复杂的通信系统归纳为一个简洁的通信系统模型，其目的就是为了便于找出信息传输过程中的内在规律和最佳处理方法，以提高信息传输的有效性、可靠性、保密性和认证性，从而使通信系统达到最优化。

所谓有效性，是指用最短的时间、最少的设备资源来传输同样的信息量。对于信源来说，是要压缩掉信源中的冗余，技术措施是

需要全本请访问 www.ertongbook.com