信息论

息论

论

Xin Xi Lun

仲振玉 编著

北京理工大学出版社

信 息 论

金振玉 编著

北京理工大学出版社

（京）新登字149号

内 容 提 要

本书共分八章，比较全面系统地介绍了香农信息论的基本内容。第二、三、四、八章主要讨论了离散无记忆信源、有记忆信源（马尔可夫信源）以及连续信源。在这几章里不仅讨论了它们的熵和编码定理，还讨论了与信源压缩编码有关的信息率失真理论。第五、六、七章主要讨论了各类典型信道容量的计算方法。特别是给出了计算信道容量的迭代公式，并对连续信道也做了比较详细的论述。同时，用适当篇幅讨论了实用信道编码。此外，在附录中分别介绍了凸函数的基本特性、某些重要的不等式、概率空间的基本概念以及计算熵用的数值表。

本书可作为高年级大学生和研究生的教材，亦可用作有关工程技术人员的参考书。

信 息 论

金振玉 编著

●

北京理工大学出版社出版

新华书店北京发行所发行 各地新华书店经售

北京密云华都印刷厂印刷

*

850×1168毫米 32开本 11印张 285千字

1991年12月第一版 1991年12月第一次印刷

ISBN 7-81013-441-8/TN·28

印数：1—3200册 定价：3.65元

前　言

随着新技术革命的蓬勃兴起，物质需求和生产增长将让位于信息，它的使用价值在一定意义上将超过物质和能。人类进入信息社会时，就不能不高度重视信息。在浩如烟海的信息中，使用者能否迅速可靠地得到他所需要的部分，就要进行信息处理、存贮和传输。例如正在兴起的综合业务数字网（ISDN），它是将语音、数据、图像等通信业务综合到一个通信网上，把信息可靠地传送给各个用户。这些信息的处理、存贮和传输所依赖的基本规律就是信息论所要讨论的基本内容。

随着信息在社会中重要性的增加，信息论将被更多的人所重视。信息论是从统计学的观点研究信息系统。它的主要任务是研究信息系统的可靠性和有效性，以及两者的关系。目前关于信息论的研究主要沿着两个方向进行，其一是理论研究，其二是应用研究。当然，应用研究不能撇开理论，而是二者相结合的研究方法。因此，认为信息论是脱离实际的抽象理论的看法是一种误解。况且，由于现代科学技术互相渗透，信息论又以“横断学科”的姿态伸向各个领域，预计它将有较大的发展。

本书大部分内容经过多次教学实践，注重应用，不追求数学的严密性，重视直观的理解。为此，多通过图、例来加深物理概念的理解，尽量避免烦杂的数学推导。

本书除包括香农信息论基本内容外，对更具有实用价值的马尔可夫信源作为一章进行了较深入的论述。此外，考虑当今信息量越来越大的特点，对信源压缩编码的基础理论“信息率失真理论”给予相当的重视。在信道方面，不仅分析了连续信道，特别对离散有扰信道给予较为重点的分析，并对信道编码也用一定的

篇幅加以叙述。因此，本书较为全面地论述了香农信息论，并对在香农理论基础上得以发展的信道编码，给予了充分注意。

本书由北京理工大学电子工程系王中教授担任主审，她对本书提出了宝贵的意见，作者在此向她表示诚挚的感谢。

作者限于水平，又兼时间仓促，书中错误与不足之处敬希读者指正。

编者

1989年6月

目　　录

第一章　绪　　论

§1-1　信息论的形成和发展……（1）
§1-2　通信系统模型……（3）
§1-3　早期关于信息的度量……（7）

第二章　信　息　量

§2-1　离散无记忆信源的信息量……（11）
§2-2　熵函数的代数性质……（14）
§2-3　熵函数的解析性质……（20）
§2-4　互信息(mutual Information)……（24）
§2-5　数据处理中信息的变化（信息不增性原理）……（27）
§2-6　平稳随机序列的熵……（29）
§2-7　冗余度……（33）
§2-8　连续信源的信息量……（34）
§2-9　最大熵定理……（40）

第三章　马尔可夫信源

§3-1　马尔可夫信源……（48）
§3-2　正规马尔可夫信源……（52）
§3-3　各态历经的马尔可夫信源……（58）
§3-4　一般马尔可夫信源与其次数……（62）
§3-5　马尔可夫信源的熵……（68）
§3-6　伴随信源……（70）
§3-7　马尔可夫信源的扩展……（74）
§3-8　语言模型……（80）

第四章　信源编码

§4-1　信源编码简介……(85)
§4-2　编码的定义……(87)
§4-3　单义码存在定理……(89)
§4-4　最佳编码……(96)
§4-5　最佳编码与译码的技术设备……(100)
§4-6　香农与费诺编码法……(103)
§4-7　离散信源变长编码定理……(106)
§4-8　平稳信源和马尔可夫信源的变长编码定理……(112)
§4-9　离散无记忆信源定长编码定理……(114)

第五章　信道与信道容量

§5-1　信道的分类……(121)
§5-2　信道模型……(125)
§5-3　信道与互信息……(132)
§5-4　对称信道……(133)
§5-5　延长信道……(136)
§5-6　无损信道和确定信道的互信息……(137)
§5-7　有扰离散无记忆信道的信道容量……(139)
§5-8　一些特殊信道的信道容量的计算……(142)
§5-9　任意有扰离散无记忆信道容量的计算——迭代计算法……(151)

第六章　信 道 编 码

§6-1　信道编码……(162)
§6-2　信道编码定理……(164)
§6-3　在二元对称信道上信道编码定理的证明……(168)
§6-4　可靠性函数……(172)
§6-5　信道编码的基本概念……(174)
§6-6　单一错误的检测与纠正……(176)
§6-7　码的纠错能力……(188)
§6-8　循环码……(199)
§6-9　有限域（伽罗瓦域Galois *field*）……(211)

§6-10 BCH码与Fire码 …… (216)

第七章 连续信道

§7-1 连续信道与其互信息量 …… (219)
§7-2 信道中的噪音 …… (221)
§7-3 高斯信道 …… (225)
§7-4 时间连续信道 …… (229)
§7-5 连续信道的信道容量 …… (236)
§7-6 具有任意噪音功率谱密度的信道容量 …… (238)
§7-7 信号空间与连续信道编码定理 …… (245)
§7-8 相关检波 …… (247)
§7-9 连续信道上离散信息的传输 …… (252)
§7-10 连续信道的二元信息传输 …… (254)
§7-11 连续信道的M元信息传输 …… (256)
§7-12 连续信道的连续信息传输——线性调制理论 …… (262)
§7-13 连续信道的连续信息传输——非线性调制 …… (267)
§7-14 时间连续信号的传输 …… (269)

第八章 信息率失真理论

§8-1 失真函数的基本概念 …… (273)
§8-2 信息率失真函数$R(D)$ …… (277)
§8-3 信息率失真函数$R(D)$的性质 …… (283)
§8-4 $R(D)$的参量表示法 …… (288)
§8-5 二元信源的信息率失真函数$R(D)$ …… (295)
§8-6 $R(D)$函数的迭代计算公式 …… (299)
§8-7 限失真信源编码定理 …… (305)
§8-8 模拟信源的信息率失真函数 …… (306)
附录A 凸函数 …… (316)
附录B 某些重要不等式 …… (322)
附录C 概率空间 …… (326)
附录D 计算熵用的数值表 …… (332)
主要参考文献 …… (341)

第一章 绪 论

信息、物质、能量是科学技术的三大支柱。科学技术的发展使人类正在进入一个新的时代——信息时代。在这个时代里，计算机技术与通信技术紧密结合，完成信息的产生、获取、传输、交换、处理、检测、识别、存贮、显示等功能，研究这方面的科学就是信息科学。信息论是信息科学的主要理论基础之一，无疑它将在信息时代中起着越来越大的作用，并将与控制论、系统论以及人工智能、生物学等互相渗透互相促进，预计将会得到更大的发展。

§1-1 信息论的形成和发展

信息论自诞生到现在不过40多年，这在人类科学史上是相当短暂的，但它的发展和对学术界及工程界的影响是相当深刻的。信息作为一种资源，如何开发、利用、共享是人们普遍关心的问题。

在人类历史的长河中，信息经历了五次重大变革：第一次变革是语言的产生。人类用语言能准确的传递感情和意图，使语言成为传递信息的重要工具。第二次变革是文字的产生。不久又发明了纸，此后，便开始用书信的方式交换信息。公元前500年左右，波斯帝国的邮政路线长达2500公里。第三次变革是印刷术的发明。它使信息能大量存贮和大量的流通，并显著地扩大了信息传递的范围。第四次变革是电话、电报的发明。它开始了人类电信时代。本世纪30年代新的调制方法，如调频、调相、单边带调制、脉冲编码调制和增量调制等的出现，使人们对信号能量、带

宽和干扰的关系有了进一步认识。1939年出现了带通声码器，这是最早的语音数据压缩系统。在这一时期无线电广播和电视广播诞生了。通信技术的进步使人们更加深入地考虑一些问题：究竟如何定量地研究通信系统中的信息、怎样才能更有效和更可靠地传递信息，现有的各种通信体制如何改进等。

1924年Ngquist解释了信号带宽和信息速率之间的关系。1928年，Hartley引入了非统计（等概率事件）信息量概念。Armstrang在1936年提出增大带宽可以使抗干扰能力加强。Hartley的工作给Shannon很大的影响，他在1941～1944年对通信和密码进行深入研究，用概率论的方法研究通信系统，揭示了通信系统传递的对象就是信息，并对信息给以科学的定量描述，提出了信息熵的概念。指出通信系统的中心问题是在噪音下如何有效而可靠地传送信息，以及实现这一目标的主要方法是编码等。这一成果1948年才以"通信的数学理论(*"A mathematical theory of communication"*)为题公开发表。该文的发表标志着信息论的诞生。与此同时，Wiener在研究火控系统和人体神经系统时，提出了熵和在干扰作用下的信号最佳滤波理论，成为通信理论的一个重要分支。

50年代信息论在学术界引起了巨大的反响。1951年美国IRE成立了信息论组，并于1955年正式出版了信息论汇刊。60年代信道编码技术有较大进展，使它成为信息论的又一重要分支。它把代数方法引入到纠错码的研究，使分组码技术发展达到了高峰，找到了大量可纠正多个错误的码，而且提出了可实现的译码方法。其次是卷积码和概率译码有了重大突破；提出了序列译码和Viterbi译码方法。

信源编码的研究落后于信道编码。Shannon1959年的文章(*"Coding theorems for adiscrete source aith a fidelty criterion"*)系统地提出了信息率失真理论（*hate-Distortion Theory*)。它是数据压缩的数学基础。为各种信源压缩编码的研

究奠定了基础。

第五次变革是计算技术与通信技术相结合，促进了通信网的发展。所谓“综合业务数字网（ISDN：*Integrated Service Digital Network*），给人们提供了不仅仅是电话业务的多种服务，还包括电报、传真、图像、数据等信息的传送，使人类社会进入了信息时代。从通信来看，是由单向通信扩展到多用户双向通信。70年代以来从经典Shannon的单向通信的信息论推广到多用户信息理论。

此外，广义信息论也相应地发展着，如语义信息（即通信符号所表达的含义）。在情报检索系统中就要考虑信息的语义，才能从浩瀚的科技文献中，根据专业术语检索出所需的文章。语用信息是涉及信息的价值和效用的一种理论。同语义信息相比，它对收信者的依赖性更强，而且与时间有着密切的关系。例如一个管理系统，一份信息过时后，就几乎没有什么价值了。建立在模糊数学基础上的模糊信息也取得了一定的进展。信息论不仅在通信、广播、电视、雷达、导航、计算机、电子对抗等电子学领域得到了直接应用，而且还广泛地渗透到诸如医学、生物学、生理学、心理学、神经学、语言学、社会学、经济学、美学等各个方面。

总之，从20世纪末期开始，“现代化”在科学技术上的主要标志，就是信息科学和信息技术的高度发展。在信息时代里，由于信息量激增，人们需要提高处理信息的能力和利用信息的能力。为此要掌握先进的信息处理技术。当然，对于信息科学基础理论之一的信息论的了解和掌握是不可缺少的，只有这样才能迎接新时代对我们提出的挑战。

§1-2 通信系统模型

实际通信系统形式虽然很多，但可用以下的基本模型加以概括。根据分析问题的需要可分为3种。

一、Shannon模型

这是1948年Shannon最初的论文所采用的模型，如图1-1所示。这个模型不仅适用各类通信系统，还适用于一切有信息传递

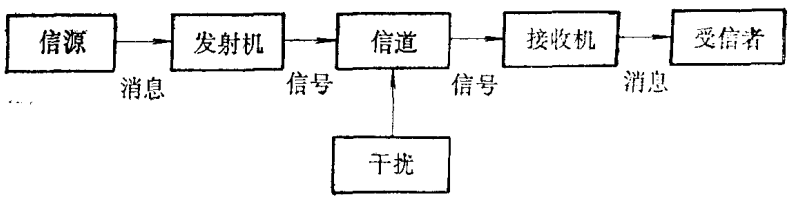

图1-1 通信系统模型之一

和变换的过程。Shannon信息论就是从这个基本模型进行研究的。这个模型由六个部分组成，其中：

(1) 信源 (*information Source*)：它是通信的起点，是产生消息的源泉；也可以说，是有效行动或者是控制动作的司令源。从信源输出的消息载荷着一定的信息。

(2) 消息 (*Message*)：它是信息的载荷者，是信息的表现形式。例如，由音素构成的语言；由字母构成的文本；由点、划构成的电报；由图像、语音和音乐构成的电视节目等。

(3) 发射机 (*Transmitter*)：把信源发出的消息变成适合信道传输的装置称为发射机。信息论不是具体研究发射技术，而是研究用适当的编码方式将消息通过信道传送出去。因此，把这一部分称为编码器更为合适 (*Encoder*)。

(4) 信号 (*Signal*)：消息的载荷者。对电信号而言，它往往是时间的函数。例如编码组、调幅波、调频波、单边带等都是信号。

(5) 信道 (*Channel*)：连接发射机与接收机的媒介。传递的

信号通过这种媒介从一个系统进入另一个系统。如载有信息的电磁波、光波、红外线、水下超声波、电缆和光纤等。图1-1中噪音源产生的干扰加到信道中去，这是一种等效看法，这里的干扰是把整个通信系统的干扰都集中起来从信道加入。例如，大气干扰、火花干扰、太阳黑子的干扰、邻近电台干扰、接收机内部噪音等。在信息论中主要研究加性干扰。

(6) 接收机 (*Receiver*)：它将信道输出的信号与噪音接收下来，还原成原来的消息。从信息论角度来看这是个译码过程，往往称为译码器 (*Decoder*)。

(7) 受信者 (*Destination*)：它是通信的目的地，消息的归宿，亦称为信宿。接收者可能是人或者是机器。

那么，究竟在信息论中所讨论的信息是什么呢？由上述通信基本模型可知，通信之所以成为需要，是因为接收者，在未收到消息以前，不知道消息的内容是什么。如果事先知道了就不需要通信了。例如某甲发了一封电报，内容是“平安到达北京”。电报未发之前某乙不知道某甲是否“平安到达北京”。对某乙来说，收到电报以后消除了一些原先不能肯定的情况。把这种原先不能肯定的情况称为先验不肯定性。假定电报内容只有“到达北京”几个字。对某乙来说还是不能肯定是否平安。只是部分的消除了不肯定性。因此，对这种不肯定性应该用一个量去度量它。在信息论中把这种先验不肯定性用一个专用名词“信息”来表示。因此，可以说：信息存在于一切事物中，是一种不可度量的抽象量，是一种先验不肯定性。而信息量，不是研究信息的抽象量，而是研究消息的可能的量。例如，某人去发电报，首先在大脑中形成电文，然后把电文写在电报稿纸上，写在纸上的电文就是消息。因此，在大脑中形成的电文是无法定量研究的，而只能研究写出的电文，即消息的信息量。在后面章节中可以知道，信息量就是不肯定性平均减小的量。

在结束关于信息的讨论时应该指出，在学术界关于信息的定

义众说纷云。信息是从*INFORMATION*这个词翻译过来的。有消息、情报知识、资料等概念。这里介绍几种典型的定义。

英国牛津英语辞典给它下了这样的定义：通过各种方式可以被传递、传播、传达、感受的，以声音、图像、文件所表征，并与某些特定的事实、主题或事件相联系的消息、情报、知识都可以泛称为信息。可见，这是一个非常广义的范畴。它与任何人的工作、学习乃至日常生活，与任何事物的运动、发展都不可分割地紧密联系在一起。信息存在于一切事物的发展过程中，而每一事物的发展过程存在着自始至终的信息运动。

美国的*Webster's New Collegiate Dictionary*对信息定义为：(1) 由调查、研究和学习所得到的资料和知识，(2) 知识、学问和新闻，(3) 事实、情况和数据，(4)表示数据的记号、信号和标识。这也是一种广义的说法。

日本广辞林定义为：关于事物的内容和情况的通报。

在哲学界对信息一词也十分感兴趣，他们提出信息究竟是物质还是精神。想回答这个问题最好还是用维纳说法：信息就是信息，即不是物质也不是能量。

总之，还能列举出许多说法和定义，但在信息论中信息的含义还是明确的。它是一个抽象量，存在于一切事物中，是一种先验不肯定性。一般人所说的信息多半指消息和情报而言，其实，它们的内含才是信息。这种混淆，在实践和应用中并无重要妨碍。

二、Fano的模型

从本质上讲，它与Shannon模型没有多大变化，只是表现方法不同而已。只要把图1-1中的发射机变成二个编码器，把接收机变成二个译码器。因此，在图1-2中的信道应该包括调制器和解调器，并且，噪音包括在接收信号中。

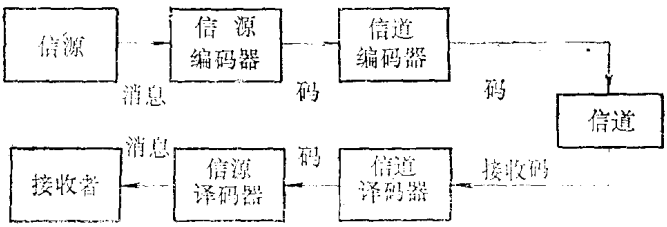

图1-2 通信系统的模型之二

三、Marko的模型

这个模型在生物和社会生活方面更为适合。其特点是，连接部件是输入码和输出码的处理点，如图1-3所示。

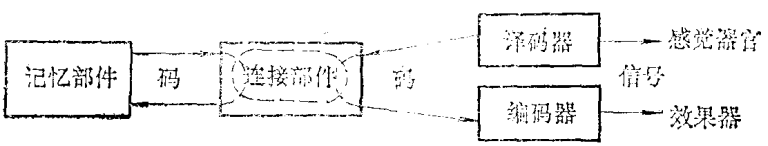

图1-3 通信系统模型之三

综上所述，在本书中如不作特别声明，分析问题的依据都是根据Shannon模型进行的。

§1-3 早期关于信息的度量

关于信息如何度量，最早是由Hartley提出的。他说：从逻辑上讲应该选择对数单位来度量信息。为了解释Hartley的定义，我们假定通信系统中不存在干扰，该系统只传送“接通”和“断开”两种消息，这样只要用两个符号表示这两种状态就可以了，例如，“接通”用“1”表示，“断开”用“0”表示。并假定“断开”

或“接通”的机会相等。也就是说，系统中出现“0”或“1”的概率相等。如果要传送的是四种消息，若仍用“0”、“1”表示四种状态，在接收端为了能区分不同的状态，每种状态最少要用两个二进制符号表示（00，11，01，10）。如果系统要传送8种状态，最少要用三个二进制符号表示一种状态（000，111，110，001，101，010，011，100），才能区分出不同的状态。依此类推，如表1-1所示。

表1.1 信息量与二进制符号关系表

信息量	消息的数量	二进制符号的位数
$\log_2 2 = 1$	2个等概率出现的消息	(0,1)用一位二进制符号分别表示2个消息
$\log_2 4 = 2$	4个等概率出现的消息	(00,11,10,01)用二位二进制符号分别表示4个消息
$\log_2 8 = 3$	8个等概率出现的消息	(000,111,001,011,010,110,101,100)用三位二进制符号分别表示8个消息
⋮	⋮	⋮
$\log_2 n = m$	n个等概率出现的消息	用m位二进制符号分别表示n个消息

由表1-1可知，取以2为底的消息数量的对数，恰恰是用以表示消息的二进制符号位数（或是单元符号数）。还可以看到，消息出现的概率越小，不肯定性就越大，即信息量越大。亦即消息数量的对数可以表示不肯定性的大小。n个等概率出现的消息的信息量，应该等于对n个消息数取对数，即

$$\text{信息量} = \log_2 n = -\log_2 p \qquad (1\text{-}1)$$

这就是哈特莱信息量的由来，即等概率出现的消息之一的信息

量，等于表示消息的二进制码所需的最小位数，也就是消息数的对数。

以上讨论，是消息出现为等概率的情况，如果信源输出消息的概率不等时，信息量如何表示呢？如果将信源输出的众多消息看成一个事件集，集合中某一事件 x 的出现概率为 $P(x)$，该事件的信息量用 $I(x)$ 表示，我们仿照等概率事件的情况，写出该事件的信息量为

$$I(x)=\log\frac{1}{P(x)} \tag{1-2}$$

把这种信息量称为某事件的自信息量（*Self information*）。

这样做是否能表达 x 事件出现的信息量呢？回答是肯定的，原因如下：

（1）当概率 $P(x)=1$ 时，$I(x)=0$。$P(x)=1$ 表明是必然事件，而必然事件是没有不肯定性的，因而，信息量为零，这与直观看法相符。

（2）$I(x)$ 是事件 x 发生概率 $P(x)$ 的单调递减函数，即随事件 x 发生概率 $P(x)$ 的增加，信息量减小，如图 1-4 所示。

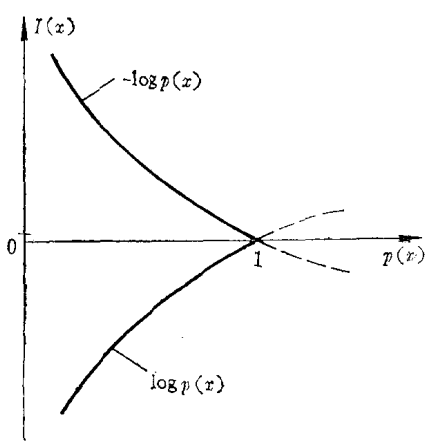

图1-4 $I(X)$ 与 $P(X)$ 关系曲线