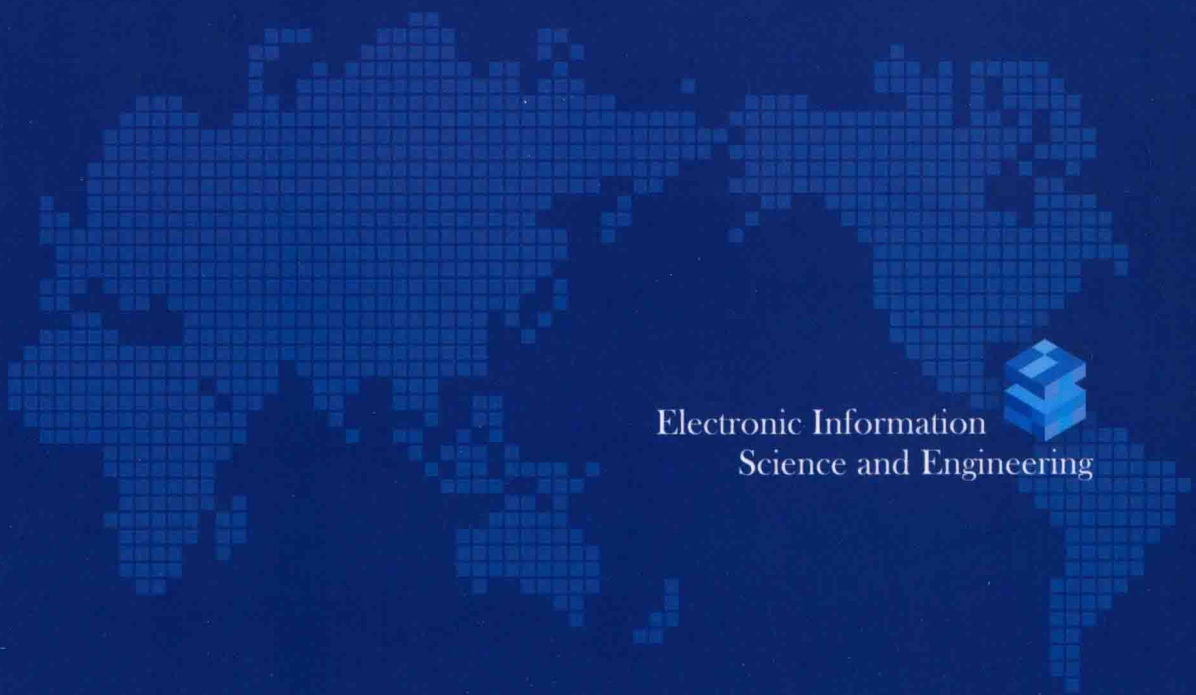
全国电子信息类和财经类优秀教材
普通高等教育“十三五”规划教材

信息论基础与应用

◆李 梅 编著



Electronic Information
Science and Engineering

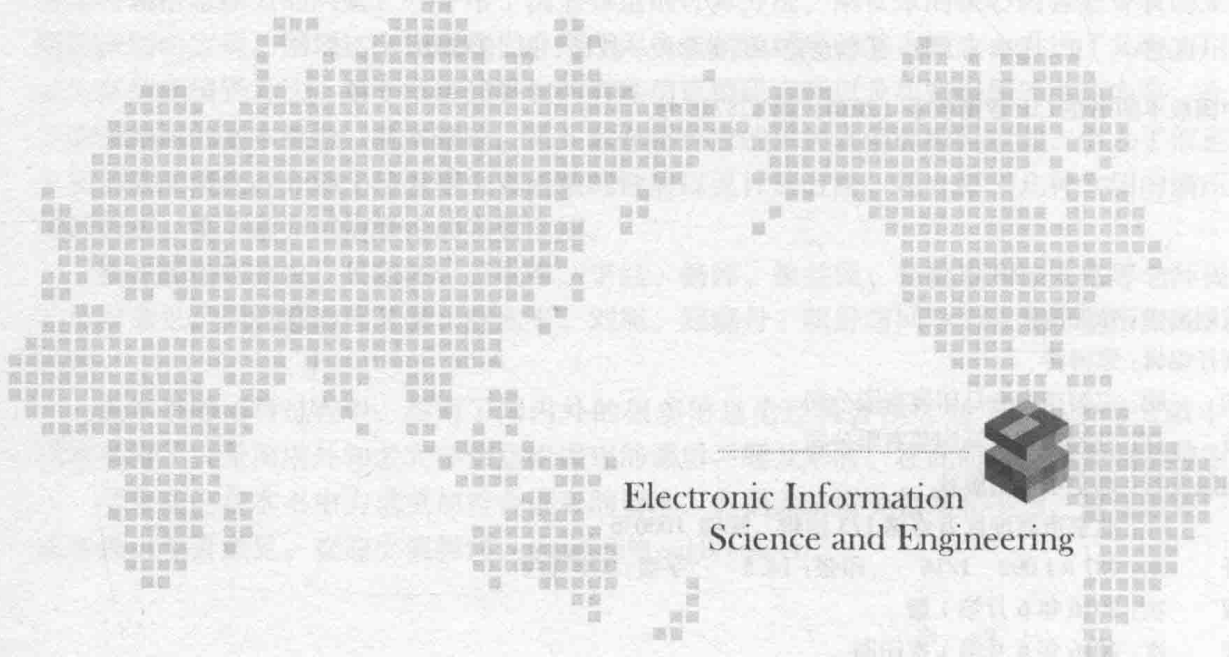
中国工信出版集团
電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

本书含二维码应用

CESE
全国电子信息类和财经类优秀教材
普通高等教育“十三五”规划教材

信息论基础与应用

李 梅 编著



電子工業出版社
Publishing House of Electronics Industry
北京 · BEIJING

内容简介

信息论是现代信息通信领域的基础理论，是研究信息传输和信息处理的一般规律的科学。我们在借鉴了国内外众多的信息论优秀教材和参考资料之后编写了本书。本书以香农的三个编码定理为中心，重点讲述了相关的基本概念、原理、方法和应用。本书介绍经典信息论的内容，不涉及过多的分支。全书选择了很多与日常生活密切相关，具有一定趣味性的习题，同时增加了动手编程实践。

本书可作为通信及电子信息类相关专业高年级本科生和研究生的教材，也可作为相关专业科研人员的参考书。

图书在版编目（CIP）数据

信息论基础与应用/李梅编著. —北京：电子工业出版社，2016.6
ISBN 978-7-121-29026-8

Ⅰ. ①信… Ⅱ. ①李… Ⅲ. ①信息论—高等学校—教材 Ⅳ. ①G201

中国版本图书馆 CIP 数据核字（2016）第 128738 号

策划编辑：章海涛
责任编辑：章海涛
印　　刷：三河市鑫金马印装有限公司
装　　订：三河市鑫金马印装有限公司
出版发行：电子工业出版社
　　北京市海淀区万寿路 173 信箱　邮编 100036
开　　本：787×1 092　1/16　　印张：14.5　　字数：370 千字
版　　次：2016 年 6 月第 1 版
印　　次：2016 年 6 月第 1 次印刷
定　　价：36.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：192910558（QQ 群）。

前　言

信息论是现代信息通信领域的基础理论，是研究信息传输和信息处理的一般规律的科学，因此目前各高等院校的相关专业的本科生、研究生都开设了这门课。

在借鉴了国内外众多的信息论优秀教材和参考资料之后，作者根据多年的教学实践经验编写了本书。本书以使读者掌握基本概念和方法为目的，力图以读者最易接受的方式介绍信息论的基本内容及应用。本书可作为通信及电子信息类相关专业高年级本科生和研究生的教材，也可作为相关专业科研人员的参考书。

教材以香农的三个编码定理为中心，重点讲述了相关的基本概念、基本原理和基本方法。鉴于目前各大高校都在削减学时，教材只介绍经典信息论的内容，没有涉及过多的分支。同时，我们认为，“信息论”是一门理论和实践紧密结合的课程，因此选择了很多与日常生活密切相关，具有一定趣味性的习题，同时增加了动手编程实践。

除绪论外，本书还包括7章内容。绪论主要介绍香农信息论的研究对象、目的和内容。第1章介绍信息度量的几个重要概念：自信息、互信息、信息熵、平均互信息以及数据处理定理。第2章研究定量度量信源产生信息的能力和信源冗余度的问题。第3章研究定量描述信道传递信息能力的问题，并介绍了信道容量的计算方法。第4章的核心内容是香农的无失真信源编码定理。围绕这个定理我们介绍了无失真信源编码的基本概念，讲述了几种实用的无失真信源编码方法。第5章讲述香农的有噪信道编码定理以及纠错编码的主要内容，介绍了信道编码的基本概念、基本理论。第6章介绍香农的限失真信源编码定理，引入了信息率失真函数的概念并介绍了信息率失真函数的性质以及计算方法，还介绍了几种常用的熵压缩编码算法。第7章介绍信息论的主要应用。

在此感谢李亦农、吴韶波、王永峰、李红、杨师、徐益民、杨世忠和徐安庭等老师提出的宝贵意见，同时感谢邢开颜、赵永平、刘旭、甄晓丹、明舒晴同学为收集习题所做的大量工作。

在本书的编写过程中，参阅了国内外的很多信息论经典著作（列于本书参考文献中），同时参考了大量国内外知名大学信息论课程的课后习题及解答，在此向有关作者表示感谢！

尽管我们在本书中力求更加符合读者的要求，但仍无法避免错漏和不当，恳切希望广大读者提出宝贵意见。欢迎交流探讨：maggieli@ cugb. edu. cn。

李　梅

2016年5月于中国地质大学（北京）

目　录

绪　论

本章首先介绍信息理论中最重要的概念——信息，澄清与之容易混淆的几个概念，然后介绍“信息论”课程的学习目的和内容，希望读者对信息论的研究对象、研究内容以及学习信息论的意义有整体的了解。

0.1　信息的概念

信息论是通信的数学基础，它是随着通信技术的发展而形成和发展起来的一门新兴的横断学科。信息论创立的标志是 1948 年 Claude Shannon（香农）发表的论文 *A Mathematical Theory of Communication*。为了解决在噪声信道中有效传输信息的问题，香农在这篇文章中创造性地采用概率论的方法来研究通信中的问题，并对信息给予了科学的定量描述，第一次提出了“信息熵”的概念。

在日常生活中，人们往往对“消息”和“信息”不加区分，消息被认为就是信息。例如，收到一封电报或者听了天气预报，人们就说得到了信息。

收到消息后，如果消息告诉了人们很多原来不知道的新内容，人们会感到获得了很多的信息，而如果消息是人们基本已经知道的内容，那么所获到的信息并不多。所以，信息应该是可以度量的。那么，怎样度量信息呢？人们需要一个可以用数学模型来表示的信息概念。

1928 年，哈特莱（Hartley）首先提出了用对数度量信息的概念。一个消息包含的信息量用其所有可能取值的个数的对数来表示。比如，抛掷一枚硬币可能有两种结果：正面和反面，所以得知抛掷结果后获得的信息量是$\log_2 2 = 1$ 比特。而一个十进制数字可以表示 0 ~ 9 中的任意一个符号，所以一个十进制数字包含$\log_2 10 = 3.3219$ 比特的信息量。这里的对数取以 2 为底，信息量的单位为比特（bit）。

哈特莱的工作给了香农很大的启示，香农进一步注意到，消息的信息量不仅与它的可能值的个数有关，还与消息本身的不确定性有关。例如，抛掷一枚偏畸硬币，如果正面向上的可能性是 90%，当人们得知抛掷结果是反面时得到的信息量，会比得知抛掷结果是正面时得到的信息量大。

一个消息之所以会包含信息，正是因为它具有不确定性，一个不具有不确定性的消息是不会包含任何信息的。通信的目的就是为了消除或部分消除这种不确定性。比如，在得知硬币的抛掷结果前，人们对于结果是出现正面还是出现反面是不确定的，通过通信，人们得知了硬币的抛掷结果，消除了不确定性，从而获得了信息。因此，信息是对事物运动状态或存在方式的不确定性的描述。这就是香农信息的定义。

用数学的语言来讲，不确定就是随机性，具有不确定性的事件就是随机事件。因此，可运用研究随机事件的数学工具——概率来测度不确定性的大小。在信息论中，人们把消息用

随机事件表示，发出这些消息的信源则用随机变量来表示。比如，抛掷一枚硬币的试验可以用一个随机变量来表示，而抛掷结果可以是正面或反面，这个具体的消息则用随机事件表示。

某个消息 x_i 出现的不确定性的大小被定义为自信息，用这个消息出现的概率的对数的负值来表示，即

$$I(x_i) = -\log p(x_i) \tag{0.1}$$

自信息同时表示这个消息包含的信息量，也就是能够给予收信者的最大信息量。如果消息能够正确传输，收信者就能够获得这样多的信息量。

信源包含的信息量定义为信源发出的所有可能消息的平均不确定性。香农把信源包含的信息量称为信息熵。自信息的统计平均定义为信源熵，即

$$H(X) = -\sum_{i=1}^{q} p(x_i)\log p(x_i) \tag{0.2}$$

式中，q 表示信源消息的个数。信息熵表示信源的平均不确定性的大小，同时表示信源输出的消息所含的平均信息量。因此，虽然信源产生的消息可能包含不同的信息量，如抛掷一枚偏畸硬币的结果“是正面”和“是反面”这两个消息所含的信息量不同，但是可以用它们的平均值来表示这个信源（抛掷一枚偏畸硬币的试验）的平均不确定性。

在接收端，信源的不确定性得到了部分或全部消除，接收者就得到了信息。信息在数量上等于通信前后“不确定性”的消除量（减少量）。这种建立在概率模型上的信息概念排除了日常生活中“信息”一词主观上的含义和作用，只是对消息的统计特性的定量描述，所以信息可以度量，而且与日常生活中“信息”的概念并不矛盾，因此是一个科学的定义。根据这样的定义，同样一个消息对于任何接收者来说，包含的信息量都是一样的。事实上，信息具有很强的主观性和实用性，同样一个消息对不同的人常常有不同的主观价值或主观意义。例如，同一则气象预报对在室外工作的人和在室内工作的人，可能会有不同的意义和价值，因此所提供的信息量也应该不同。所以，香农信息在某些情况下也具有一定的局限性。

0.2 信息论的研究对象、目的和内容

从诞生到现在，信息论虽然只有短短的几十年，但它的发展对学术界及人类社会的影响是相当广泛和深刻的。如今，信息论的研究内容不仅包括通信，还包括所有与信息有关的自然和社会领域，如模式识别、计算机翻译、心理学、遗传学、神经生理学、语言学、语义学甚至社会学中有关信息的问题。香农信息论迅速发展成为涉及范围极广的广义信息论——信息科学。

信息论的研究对象是广义的通信系统，它把所有的信息流通系统都抽象成如图 0.1 所示的模型。这个模型不仅包括电话、电报、传真、电视、雷达等狭义的通信系统，还包括生物有机体的遗传系统、神经系统、视觉系统甚至人类社会的管理系统。信息是以消息的形式在这个通信系统中传递的。通过研究通信系统中消息的传输和处理，人们得到信息传输与处理的规律，以提高通信的可靠性和有效性。

任何信息流通系统中都有一个发出信息的发送端（信源）、一个接收信息的接收端（信宿）以及信息流通的通道（信道）。在信息传递的过程中不可避免地会有噪声，所以会有一

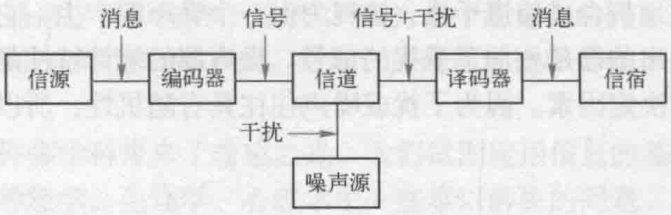


图 0.1　通信系统模型

个噪声源。为了把信源发出的消息变成适合在信道中传输的信号，还需要加入编码器，在送到信宿之前要进行反变换，所以要加入译码器。

这个通信系统主要包括如下 5 部分。

（1）信源

顾名思义，信源是产生消息和消息序列的源。信源可以是人、生物、机器或其他事物。比如，“各种气象状态”是信源，能够产生独特气味吸引蜜蜂采蜜的花朵是信源，人的大脑活动也是一种信源。信源的输出是消息（或消息序列）。

消息有不同的形式，如文字、符号、语言、图片、图像、气味等。消息以能被通信双方所理解的形式，通过通信进行传递和交换。消息携带着信息，是信息的载体。

信源输出的消息是随机的、不确定的，但有一定的规律性，因此用随机变量或随机矢量等数学模型来表示信源。

（2）编码器

编码就是把消息变成适合在信道中传输的物理量，这种物理量称为信号（如电信号、光信号、声信号、生物信号等）。信号携带着消息，它是消息的载体。

编码器可分为信源编码器和信道编码器。

信源编码的目的是压缩信源的冗余度（即多余度），提高信息传输的效率，这是为了提高通信系统的有效性。信源编码又可分为无失真信源编码和限失真信源编码。

信道编码是为了提高信息传输的可靠性而有目的地对信源编码器输出的代码组添加一些监督码元，使之具有纠错、检错能力。比如，老师讲课需要把知识进行加工和提炼，以提高信息传输的有效性，而为了让学生听得明白，有时需要适当重复，这是为了提高信息传输的可靠性。

在实际的通信系统中，可靠性和有效性常常是矛盾的，提高有效性必须去掉信源符号的冗余部分，但是这会导致可靠性的下降；而提高可靠性需要增加监督码元，这又降低了有效性。有时为了兼顾有效性，不一定要求绝对准确地在接收端再现原来的消息，而是可以允许一定的误差或失真，也就是说，允许近似地再现原来的消息。

（3）信道

信道是指通信系统把载荷消息的信号从发送端送到接收端的媒介或通道，是包括收发设备在内的物理设施。除了传播信号外，信道还有存储信号的作用。在狭义的通信系统中，实际信道有明线、电缆、波导、光纤、无线电波传播空间、磁盘、光盘等，这些都属于传输电磁波能量的信道。对于广义的通信系统来说，信道还可以是其他传输媒介。

在信道中引入噪声和干扰是一种简化的表达方式。为了分析方便，系统其他部分产生的

干扰和噪声都被等效地折合成信道干扰，被视为由一个噪声源产生，它将作用于所传输的信号上。这样，信道输出的已是叠加了干扰的信号。噪声源的统计特性是划分信道的依据，并且是信道传输能力的决定因素。因为干扰或噪声往往具有随机性，所以信道用输入和输出之间的条件概率来描述。

(4) 译码器

译码就是把信道输出的已迭加了干扰的编码信号进行反变换，变成信宿能够理解的消息。译码器也可分为信源译码器和信道译码器。译码器需要尽可能地再现信源输出的消息。

(5) 信宿

信宿是消息传送的对象，即接收消息的人或机器。

以上考虑的是收发两端单向通信的情况，只有一个信源和一个信宿，信息传输也是单向的。在组网通信的情况下，如电话网、计算机网络等，可能有很多单独的信源、信道和信宿同时进行信息交换。例如，广播信道是一个输入、多个输出的单向信道，卫星通信则是多个输入、多个输出的多向传输的通信。这就需要把两端单向通信的模型进行适当修正，得出多用户通信系统的模型，即把两端单向通信的信息理论发展成为多用户通信信息理论。

信息论研究的是关于这个通信系统的最根本、最本质的问题。例如：

① 什么是信息？如何度量信息？

② 怎样确定信源的输出中含有多少信息量？

③ 对于一个信道，它传输信息量的最高极限（信道容量）是多少？

④ 为了能够无失真地传输信源信息，对信源编码时所需的最少的码符号数是多少？（无失真信源编码即香农第一定理）

⑤ 在有噪信道中有没有可能以接近信道容量的信息传输率传输信息而错误概率几乎为零？（有噪信道编码即香农第二定理）

⑥ 如果对信源编码时允许一定量的失真，所需的最少的码符号数又是多少？（限失真信源编码即香农第三定理）

毫无疑问，如果对这些问题都有了确定的答案，那么在设计通信系统时就有了目标和指导方向，也有了评价通信系统优劣的标准。

下面举几个成功地应用信息论的概念和方法指导通信系统设计的例子。

(1) 无失真信源编码的应用之一：计算机文件的压缩

由于数据库的广泛应用，存储计算机文件所需的存储量问题日益突出。目前，对计算机文件的压缩已发展了至少 20 多种算法，其中较好的算法能使文件压缩后所需的存储量只为原文件的 30% 左右。

(2) 有噪信道编码的应用之一：模拟话路中数据传输速率的提高

最早的调制解调器的速率只有 300 bps，此后调制解调器的速率为 4800 bps、9600 bps、14.4 kbps、19.2 kbps、28.8 kbps，如今的实际速率已达到 33.6 kbps，非常接近于理论极限。

(3) 限失真信源编码的应用之一：语音信号压缩

按照信息理论的分析，话音信号所需的编码速率可以远远低于按奈奎斯特采样定理和量化噪声理论所确定的编码速率。几十年来，人们在这方面的工作取得了巨大进展。CCITT 关于长途电话网的语音编码速率标准已从 1972 年 G.711 标准中的 64 kbps 降低到 1992 年标准

中的16 kbps。在移动通信中，1988 年，欧洲 GSM 标准中的语音编码速率为 13. 2 kbps。1989 年，美国 CTIA 标准中的速率为 7. 95 kbps。目前，声码器的速率可低于 100 bps，已接近信息论指出的极限。

信息论的成就给许多学科带来了希望之光，人们试图应用信息的基本理论解决诸如组织化、语义化、听觉、神经学、生理学、心理学中一些难以解决的问题。目前，人们已将信息论广泛应用于物理、化学、生物学、心理学、管理学等学科，信息的概念和方法已广泛渗透到各个科学领域，迫切要求突破香农信息论的狭隘范围，以便推动许多新兴学科的进一步发展。一门研究信息的科学——广义信息论正在形成。对于信息论的研究范围，一般有如下 3 种理解：

① 狭义信息论：又称为香农信息论，主要通过数学描述和定量分析，研究通信系统从信源到信宿的全过程，包括信息的测度、信道容量以及信源和信道编码理论等问题，强调通过编码和译码使收、发两端联合最优化，并且以定理的形式证明极限的存在。这部分内容是信息论的基础理论。

② 一般信息论：也称为工程信息论，主要研究信息传输和处理问题，除香农信息论的内容外，还包括噪声理论、信号滤波和预测、统计检测和估计、调制理论、信息处理理论、保密理论等。

③ 广义信息论：不仅包括上述两方面的内容，还包括所有与信息有关的自然和社会科学领域，如模式识别、机器翻译、心理学、遗传学、神经生理学、语言学、语义学甚至社会学中有关信息的问题。

扩展阅读：信息论的形成和发展

香农

克劳德·艾尔伍德·香农（Claude Elwood Shannon，1916—2001 年），美国数学家、信息论创始人。1948 年和 1949 年，香农在《贝尔系统技术杂志》（*Bell System Technical Journal*）上分别发表了著名论文《通信的数学原理》和《噪声下的通信》。香农阐明了通信的基本问题，给出了通信系统的模型，提出了信息量的数学表达式，并解决了信道容量、信源统计特性、信源编码、信道编码等一系列基本问题。这两篇论文成为信息论的奠基性著作。

从有人类的那一天开始，人类就生活在信息的海洋里。人类日常生活、工农业生产、科学研究和战争等，一切都离不开消息传递和信息流动。为了突破语言、文字的局限性，人类不断地创造出许许多多的信息传递方法。1844 年，美国人莫尔斯发明了高效率编码电报法；1876 年，贝尔发明了世界上第一台可用的电话机；1887 年，马可尼发明了无线电报，加上调幅广播、电视、调频广播、数字通信系统、声码器、扩频通信等，通信系统已日益成为人类社会的神经系统。

信息论是在长期的通信工程实践和理论研究的基础上发展起来的，它研究如何认识信息，利用信息，以改变自己的生存条件、创造更好的生活环境等。尽管人类对信息的认识、利用源远流长，但真正对信息理论的研究只有半个多世纪。

1922 年，卡松提出了边带理论，指明了信号在调制（编码）和传送的过程中与频谱宽

度的关系。1924 年，美国的奈奎斯特证实了卡松的理论。1928 年，哈特莱发表了《信息的传输》，首先提出了消息是代码、符号、序列，而不是内容本身。他第一次提出了“信息量”的概念，并试图用数字公式加以描述，为信息论的创立提供了思路。1945 年，莱斯对噪声的研究做了全面的总结，通信理论已经全面走上统计分析之路。1946 年计算机和 1947 年晶体管的诞生与相应技术的发展，则是信息论产生的物质基础。

第二次世界大战中，由于通信在军事上的重要意义，香农开始从事信息论的研究。1948 年，香农发表了《通信的数学理论》，主要内容是研究信源、信宿、信道及编码问题。战后，由于通信事业的需要和电子技术的飞速发展，促进了信息论的进一步发展，许多国家的学者对此进行了大量的研究工作，并卓有成就。1951 年，美国无线电工程学会承认了信息论这门新学科，建立了信息论学组。

扩展阅读：量子通信与量子信息论

20 世纪 80 年代，量子力学与信息科学相结合，诞生了一门新型的交叉学科——量子信息学（Quantum Information Theory），主要包括量子通信和量子计算，为确保信息安全和提高计算速度提供了全新的方案。

量子通信主要基于量子纠缠态的理论，使用量子隐形传态（传输）的方式实现信息传递。根据实验验证，具有纠缠态的两个粒子无论相距多远，只要一个发生变化，另外一个也会瞬间发生变化。利用这个特性实现量子通信的过程如下：事先构建一对具有纠缠态的粒子，将两个粒子分别放在通信双方，将具有未知量子态的粒子与发送方的粒子进行联合测量（一种操作），则接收方的粒子瞬间发生坍塌（变化），坍塌（变化）为某种状态，这个状态与发送方的粒子坍塌（变化）后的状态是对称的，然后将联合测量的信息通过经典信道传送给接收方，接收方根据接收到的信息，对坍塌的粒子进行幺正变换（相当于逆转变换），可得到与发送方完全相同的未知量子态。

量子纠缠可以用“薛定谔猫”来帮助理解：把一只猫放到一个有毒的盒子中，然后将盒子盖上，然后问这只猫现在是死了还是活着。量子物理学的答案是：它既是死的，也是活的。有人会说，打开盒子看一下不就知道了。是的，打开盒子确实能知道猫是死是活，但按量子物理的解释：这种死或活的状态是人为观察的结果，也是人的宏观干扰使得猫变成了死的或活的，并不是盒子盖着时的真实状态。同样，微观粒子在不被“干扰”之前就一直处于“死”和“活”两种状态的叠加，也可以说它既是“0”也是“1”。

以笛卡儿、伽利略、牛顿为代表的主流物理学家认为，宇宙的组成部分相互独立，它们之间的相互作用受到时空的限制。而量子纠缠效应脱离了时空，证实了任何两种物质之间，不管距离多远，都有可能相互影响，不受四维时空的约束，是非局域的（nonlocal），不仅宇宙证实了“爱因斯坦的幽灵”——超距作用（spooky action in a distance）的存在，也证实了中国人一直强调的因果报应等观点——任何两种物质在冥冥之中存在深层次的内在联系。

与量子通信相比，经典通信的安全性和高效性都无法与之相提并论。量子通信绝不会“泄密”，确保了通信的安全性。其一，量子加密的密钥是随机的，即使被窃取者截获，也无法得到正确的密钥，因此无法破解信息；其二，通信双方分别有两个处于纠缠态的粒子，其中一个粒子的量子态发生变化，另一方的粒子量子态就会立刻随之变化。根据量子理论，

宏观的任何观察和干扰都会立刻改变量子态，导致其坍塌，因此窃取者由于干扰而得到的信息已经破坏，并非原有信息。高效性体现在被传输的未知量子态在被测量之前会处于纠缠态，即可以同时代表多个状态。例如，一个量子态可以同时表示0和1两个数字，7个这样的量子态就可以同时表示128个状态或128个数字：0~127。量子通信的这样一次传输就相当于经典通信方式的128次。可以想象，如果传输带宽是64位或更高，那么效率之差将是惊人的。

量子通信是国际量子物理和信息科学的研究热点。我国从20世纪80年代开始从事量子光学领域的研究，并且在量子通信领域开始领跑世界。

1997年，在奥地利留学的中国青年学者潘建伟与荷兰学者波密斯特等人合作，首次实现了未知量子态的远程传输。这是国际上首次实验成功将一个量子态从甲地的光子传输到乙地的光子上。实验中传输的只是表达量子信息的“状态”，作为信息载体的光子本身并不被传输。

2006年夏，中国科学技术大学教授潘建伟小组、美国洛斯·阿拉莫斯国家实验室、欧洲慕尼黑大学－维也纳大学联合研究小组，各自独立地实现了诱骗态方案，同时实现了超过100 km的诱骗态量子密钥分发实验，由此打开了量子通信走向应用的大门。

2008年底，潘建伟的科研团队成功研制了基于诱骗态的光纤量子通信原型系统，在合肥成功组建了世界上首个3节点链状光量子电话网，成为实用化量子通信网络实验研究的两个团队之一（另一个团队为欧洲联合实验团队）。

2009年9月，潘建伟的科研团队在3节点链状光量子电话网的基础上，建成了世界上首个全通型量子通信网络，首次实现了实时语音量子保密通信，标志着中国在城域量子网络关键技术方面已经达到了产业化要求。

2012年，中国科学家潘建伟等人在国际上首次成功实现百千米量级的自由空间量子隐形传态和纠缠分发。

2016年8月16日，由潘建伟担任首席科学家的世界第一颗量子科学实验卫星“墨子号”顺利发射升空。“墨子号”是中国科学院空间科学先导专项首批实验卫星之一，主要科学目标是星地高速量子密钥分发实验，在此基础上实验广域量子密钥网络，以期空间量子通信实用化；在太空中分发纠缠光子，实验量子隐形传态，并检验空间尺度的量子力学完备性。2016年年底，“京沪干线”将建成，全长超过2000 km，将成为连接北京、济南、合肥、上海等城域网络的量子保密通信线路，也将是全球首个远距离广域光纤量子保密通信骨干线路。

第1章 信息的度量

在信息论尚未作为一门学科建立起来之前，人们对于信息的度量并没有一个定量的概念，自香农开始，才将信息量的定量描述确定下来。对信息的定量描述有助于人们更方便地研究通信系统的可靠性和有效性。

在最简单的离散随机变量的情况下，下面引入关于信息度量的几个最重要的概念。

- 自信息：一个事件（消息）本身所包含的信息量，它是由事件的不确定性决定的。比如，“抛掷一枚硬币的结果是正面”这个消息包含的信息量。
- 互信息：一个事件所给出的关于另一个事件的信息量。比如，今天下雨所给出的关于明天下雨的信息量。
- 平均自信息（信息熵）：事件集（用随机变量表示）所包含的平均信息量，表示信源的平均不确定性。比如，抛掷一枚硬币的试验所包含的信息量。
- 平均互信息：一个事件集给出的关于另一个事件集的平均信息量。比如，今天的天气给出的关于明天的天气的信息量。

1.1 自信息和互信息

1.1.1 自信息

在绪论中讲过，信源发出的消息（事件）具有不确定性，而事件发生的不确定性与事件发生的概率大小有关，概率越小，不确定性越大，事件发生以后包含的信息量就越大。小概率事件的不确定性大，一旦出现必然使人感到意外，因此产生的信息量就大，特别是几乎不可能出现的事件一旦出现，必然产生极大的信息量。大概率事件是预料之中的事件，不确定性小，即使发生，也没什么信息量，特别是概率为1的确定事件发生以后，不会给人以任何信息量。因此，随机事件的自信息量$I(x_i)$是该事件发生概率$p(x_i)$的函数，并且$I(x_i)$应该满足以下公理化条件：

① $I(x_i)$是$p(x_i)$的严格递减函数。当$p(x_1)<p(x_2)$时，$I(x_1)>I(x_2)$，概率越小，事件发生的不确定性越大，事件发生以后包含的自信息量越大。

② 极限情况下，当$p(x_i)=0$时，$I(x_i)\to\infty$；当$p(x_i)=1$时，$I(x_i)=0$。

③ 从直观概念上讲，由两个相对独立的不同消息所提供的信息量，应等于它们分别提供的信息量之和。

可以证明，满足以上公理化条件的函数形式是对数形式。

【定义1-1】随机事件的自信息量定义为该事件发生概率的对数的负值。设事件x_i的概率为$p(x_i)$，则它的自信息定义为

$$I(x_i) \overset{\text{def}}{=\!=} -\log p(x_i) = \log \frac{1}{p(x_i)} \tag{1.1}$$

自信息量的函数曲线如图 1.1 所示。可以看到，$I(x_i)$的这种定义正是满足上述公理性条件的函数形式。在它的定义域[0, 1]内，自信息是非负的。

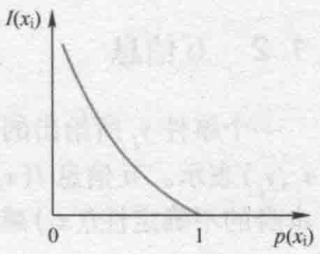


图 1.1　自信息量

$I(x_i)$代表两种含义：事件 x_i 发生前，等于事件 x_i 发生的不确定性的大小；事件 x_i 发生后，表示事件 x_i 包含或所能提供的信息量。在无噪信道中，事件 x_i 发生后，能准确无误地传输给接收者，所以 $I(x_i)$等于接收者接收到 x_i 后所获得的信息量。这是因为消除了 $I(x_i)$大小的不确定性后，才获得了这样大小的信息量。

自信息量的单位与所用对数的底有关。

通常取对数的底为 2，信息量的单位为比特（bit，binary unit）。比特是信息论中最常用的信息量单位。$p(x_i) = \frac{1}{2}$时，$I(x_i) = 1$ 比特，即概率等于$\frac{1}{2}$的事件具有 1 比特的自信息量。例如，一枚均匀硬币的任何一种抛掷结果均含有 1 比特的信息量。当取对数的底为 2 时，2 常省略。注意：计算机术语 bit 是位的单位（bit，binary digit），与信息量的单位不同，但有联系，1 位的二进制数字最大能提供 1 比特的信息量。

若取自然对数（对数以 e 为底），则自信息量的单位为奈特（nat，natural unit）。理论推导中或用于连续信源时，用以 e 为底的对数比较方便。

$$1 \text{ 奈特} = \log_2 e \text{ 比特} = 1.443 \text{ 比特}$$

工程上用以 10 为底的对数较为方便。若以 10 为对数的底，则自信息量的单位为哈特莱（Hartley）（因为哈特莱最先提出用对数来度量信息）。

$$1 \text{ 哈特莱} = \log_2 10 \text{ 比特} = 3.322 \text{ 比特}$$

如果取以 r 为底的对数($r > 1$)，则 $I(x_i) = -\log_r p(x_i)$（r 进制单位）。

$$1 \quad r \text{ 进制单位} = \log_2 r \text{ 比特}$$

【例 1.1】

（1）英文字母中“a”出现的概率为 0.064，“c”出现的概率为 0.022，分别计算它们的自信息量。

（2）假定前后字母出现是互相独立的，计算“ac”的自信息量。

（3）假定前后字母出现不是互相独立的，当“a”出现以后，“c”出现的概率为 0.04，计算“a”出现以后，“c”出现的自信息量。

【解】

（1）

$$I(a) = -\log 0.064 = 3.96 \text{ 比特}$$

$$I(c) = -\log 0.022 = 5.51 \text{ 比特}$$

（2）由于前后字母出现是互相独立的，“ac”出现的概率为 0.064×0.022。

$$\begin{aligned} I(ac) &= -\log(0.064 \times 0.022) \\ &= -(\log 0.064 + \log 0.022) \\ &= I(a) + I(c) = 9.47 \text{ 比特} \end{aligned}$$