

信息 基础理论解析

李迎春 李长青 等 编著

An Analysis of the Basic Theory
of Information

本书以香农信息论的基本理论（香农三大定理）为主线，系统地解析了信息的度量、信源与信道、信源与信道编码、网络信息论、信息论应用等理论及应用问题



中国发展出版社
CHINA DEVELOPMENT PRESS

信息 基础理论解析

An Analysis of the Basic Theory
of Information

李迎春 李长青 等 编著



中国发展出版社
CHINA DEVELOPMENT PRESS

图书在版编目 (CIP) 数据

信息基础理论解析 / 李迎春等编著. —北京: 中国发展出版社, 2017. 4
ISBN 978-7-5177-0618-2

I. ①信… II. ①李… III. ①信息论 IV. ①G201

中国版本图书馆CIP数据核字 (2016) 第290163号

书 名: 信息基础理论解析

著作责任者: 李迎春 李长青 等

出版发行: 中国发展出版社

(北京市西城区百万庄大街16号8层 100037)

标准书号: ISBN 978-7-5177-0618-2

经 销 者: 各地新华书店

印 刷 者: 北京明恒达印务有限公司

开 本: 720mm × 960mm 1/16

印 张: 20.5

字 数: 292千字

版 次: 2017 年 4 月第 1 版

印 次: 2017 年 4 月第 1 次印刷

定 价: 69.80元

联系电话: (010) 68990646 68990692

购书热线: (010) 68990682 68990686

网络订购: <http://zgfcbs.tmall.com>

网购电话: (010) 68990639 88333349

本社网址: <http://www.developress.com.cn>

电子邮件: cheerfulreading@sina.com

版权所有 · 翻印必究

本社图书若有缺页、倒页, 请向发行部调换

“人类曾经以采集食物为生，而如今他们重新要以采集信息为生，尽管这看上去有点不和谐。但在这个角色上，电子时代的人类与他们旧石器时代的祖先一样，也是游牧族。”

——科学家马歇尔·麦克卢汉

“生活中最成功的人通常是那些拥有最佳信息的人。”

——政治家本杰明·迪斯雷利

香农1948年发表的著名论文《通信的数学理论》(Shannon, 1948)标志着信息论的诞生，从那以后，信息研究便成为一门专门学问，人类社会开始逐渐走向自觉地认识和高水平地利用信息的阶段。

信息论是现代信息科学的主要基础理论之一。信息论主要研究对象是香农所提出的一般通信系统模型，主要研究内容是消息如何在通信系统中进行传输，主要研究目的是提高通信系统的有效性、可靠性和安全性，以便达到通信系统的最优化。

信息论的基础理论围绕香农三大定理展开，即无失真信源编码定理、有噪信道编码定理、保真度准则下的限失真信源编码定理。除此以外，信息论的基础理论还包括与之相关的信源、信道、信息度量、信道容量等基本概念及理论。

本书共分八章。第一章主要介绍信息论中重要的基本概念和相关的知识

信息基础理论解析

背景。第二章重点分析了信息的度量问题。第三章、第四章分别讨论了信源和信道问题。前四章是后续章节讨论香农三大定理的基础。第五章重点介绍了香农第一定理（无失真变长信源编码定理）以及香农第三定理（保真度准则下的限失真信源编码定理），由于都是针对信源编码，故将这两部分内容融入一章中。第六章重点讨论有噪信道编码定理及几种典型的信道编码方法。第七章介绍了信息传输网络化理论。第八章介绍了信息基本理论在各个领域内的应用。

本书作者长期从事“信息论基础理论及应用”的教学工作，结合已有的教学经验和相关领域的研究成果编写了此书。本书由李迎春、李长青、李海俊、范有臣共同完成。徐华正、吕登龙、张记瑞、胡一博参与了编写工作。

本书突出了信息理论的基本概念、基本理论及其相关数学表达式的物理意义的解析，尽量避免繁琐的公式证明及推导。全书深入浅出、概念清晰、内容系统性强。本书适合作为高等教育信息类专业本科生、研究生的教材或教学参考书，亦可作为有关科技人员拓展知识领域的参考书籍。

由于作者水平所限，书中难免纰漏之处，期望读者及时斧正。

第1章	信息论概述	/ 1
1.1	信息论与信息科学	/ 3
1.2	信息的概念与性质	/ 5
1.2.1	信息的概念	/ 5
1.2.2	信息的性质	/ 9
1.2.3	信息、消息及信号	/ 9
1.3	信息论研究的对象、目的和内容	/ 11
1.3.1	信息论研究的对象	/ 11
1.3.2	信息论研究的目的	/ 13
1.3.3	信息论研究的内容	/ 14
1.4	信息论的发展历程	/ 16
1.4.1	信息论的酝酿	/ 16
1.4.2	信息论的建立	/ 18
1.4.3	信息论的发展	/ 19
1.4.4	香农及其贡献	/ 19
第2章	信息的统一度量	/ 21
2.1	自信息	/ 24
2.1.1	自信息	/ 24
2.1.2	条件自信息	/ 26

信息基础理论解析

2.2	信息熵	/ 26
2.2.1	熵的定义	/ 27
2.2.2	熵的性质	/ 30
2.2.3	联合熵	/ 34
2.2.4	条件熵	/ 35
2.3	互信息	/ 36
2.3.1	互信息的定义	/ 36
2.3.2	互信息的性质	/ 37
2.4	平均互信息	/ 38
2.4.1	平均互信息的定义	/ 39
2.4.2	平均互信息的性质	/ 40
2.5	各种信息度量之间的关系	/ 45
2.6	扩展信源的信息度量	/ 46
2.7	连续随机变量的信息度量	/ 48

第3章 信源 / 57

3.1	信源的分类	/ 60
3.2	信源的数学描述	/ 61
3.2.1	用随机变量描述离散消息	/ 61
3.2.2	用随机矢量描述离散序列	/ 62
3.2.3	用随机过程描述波形信源	/ 64
3.3	马尔可夫信源	/ 65
3.3.1	马尔可夫信源定义	/ 65
3.3.2	状态转移概率及描述	/ 65
3.3.3	切普曼—柯尔莫戈洛夫方程	/ 68

3.3.4	马尔可夫链的状态分类	/ 69
3.3.5	马尔可夫链的遍历性	/ 70
3.3.6	离散马尔可夫信源的熵	/ 70
3.4	信源的相关性和剩余度	/ 74
第4章	信道	/ 79
4.1	信道的分类	/ 82
4.2	信道的数学描述	/ 85
4.2.1	离散信道模型	/ 86
4.2.2	连续信道模型	/ 89
4.2.3	波形信道模型	/ 90
4.3	信道疑义度和噪声熵	/ 91
4.4	信道的组合	/ 92
4.4.1	串联信道	/ 92
4.4.2	并联信道	/ 95
4.5	信道容量	/ 96
4.5.1	信道容量的定义	/ 97
4.5.2	简单离散信道的信道容量	/ 98
4.5.3	对称离散信道的信道容量	/ 101
4.5.4	组合信道的信道容量	/ 104
4.5.5	一般离散信道的信道容量	/ 106
4.5.6	离散无记忆信道容量的迭代算法	/ 118
4.5.7	连续型信道的信道容量	/ 126
4.6	信源与信道的匹配	/ 136

第5章 信源编码 / 143

- 5.1 信源编码的基本概念 / 146
- 5.2 唯一可译性的判决问题 / 151
 - 5.2.1 唯一可译性的存在性判决 / 151
 - 5.2.2 唯一可译码判断准则 / 152
- 5.3 无失真信源编码 / 154
 - 5.3.1 等长码的编码长度 / 154
 - 5.3.2 等长无失真信源编码定理 / 156
 - 5.3.3 变长无失真信源编码定理(香农第一定理) / 160
 - 5.3.4 变长码的常用编码方法 / 168
- 5.4 限失真信源编码 / 176
 - 5.4.1 失真测度 / 176
 - 5.4.2 信息率失真函数 / 181
 - 5.4.3 信息率失真函数的计算 / 187
 - 5.4.4 限失真信源编码定理(香农第三定理) / 188

第6章 信道编码 / 199

- 6.1 信道编码的基本概念 / 202
 - 6.1.1 编码信道 / 202
 - 6.1.2 译码准则 / 202
 - 6.1.3 错误概率 / 203
 - 6.1.4 码字空间 / 204
 - 6.1.5 汉明距离 / 204
 - 6.1.6 纠检错能力 / 205

6.2	两种典型的译码准则	/ 206
6.2.1	最大后验概率准则	/ 206
6.2.2	最大似然译码准则	/ 207
6.3	有噪信道编码定理 (香农第二定理)	/ 210
6.4	信源信道联合编码定理	/ 211
6.5	信道编码	/ 212
6.5.1	奇偶校验码	/ 213
6.5.2	简单重复码	/ 216
6.5.3	线性分组码	/ 218
6.5.4	循环码	/ 220
6.5.5	卷积码	/ 223

第7章 网络信息论 / 235

7.1	网络信道的分类	/ 239
7.1.1	多源接入信道	/ 239
7.1.2	广播信道	/ 240
7.1.3	中继信道	/ 240
7.1.4	串扰信道	/ 241
7.1.5	反馈信道	/ 241
7.2	网络信道的容量	/ 242
7.3	网络编码的基本原理	/ 250
7.3.1	网络编码相关概念	/ 251
7.3.2	网络编码的优点	/ 252
7.3.3	网络编码的构造算法及发展方向	/ 258
7.3.4	网络编码的适用条件	/ 265

第8章 信息论的应用 / 267

8.1 生物信息学 / 269

8.1.1 核酸序列分析 / 271

8.1.2 蛋白质序列分析 / 271

8.1.3 生物医学成像 / 272

8.2 信息光学 / 279

8.2.1 光学全息 / 281

8.2.2 光学信息处理 / 283

8.2.3 光通信 / 286

8.3 信息论与密码学 / 288

8.3.1 保密系统的数学模型 / 290

8.3.2 传统加密技术 / 291

8.3.3 现代加密技术 / 295

8.4 经济信息学 / 298

8.4.1 香农的投资之路 / 299

8.4.2 凯利公式 / 302

8.4.3 优化投资组合——马科维茨理论 / 304

8.4.4 优化投资组合的熵理论 / 306

8.5 信息论与管理 / 307

8.5.1 管理的最简信息模型 / 308

8.5.2 复杂管理信息模型构造 / 312

第 1 章

信息论概述

“信息就是信息，既不是物质，也不是能量。信息是人们在适应客观世界，并使这种适应被客观世界感受的过程中与客观世界进行交换的内容的名称。”

——控制论创始人诺伯特·维纳

“信息是用来减少随机不确定性的东西。”

——信息论创始人克劳德·艾尔伍德·香农

人类经历了六千年的农业社会和近三百年的工业社会，现在已步入第三个文明社会——信息社会。在信息社会里，信息就像空气一样充斥在我们这个时代政治、经济、文化和社会生活的方方面面。

社会信息化的一个典型技术标志就是信息网络的出现：计算机和通信设备结合起来而构成的全球性通信网络，从根本上消除了距离对通信的限制。社会信息化的另一个重要标志就是大多数人都从事或参与信息的生产、应用和保障工作。



图 1-1 信息网络连接的社会

在19世纪以前,对信息的认识和利用一直处在原始和经验的状况。人类真正自觉地认识信息问题并开始触及到信息的本质,是在20世纪40年代以后。其中,理论上最主要的标志是香农的论文《通信的数学理论》(Shannon, 1948)和维纳的专著《控制论——动物和机器中的通信与控制问题》(Wiener, 1948)。正因为如此,一部关于人类认识和利用信息的历史可以粗略地以1948年为界进行划分:在此之前,人类不自觉地认识和低水平地利用信息;在此之后,人类才开始逐渐走向自觉地认识和高水平地利用信息的阶段。

1.1 信息论与信息科学

信息论产生于上世纪40年代末,最早仅限于通信领域。由于现代自然科学各门学科的相互联系、相互渗透,信息概念及理论超越通信领域运用于其他学科。在上世纪60年代末至70年代初出现了信息科学。信息科学是在信息论的基础上发展起来的,它涉及数学、通信理论、控制论、计算机科学、人工智能、电子学和自动化技术以及物理学、生物学等多个领域,与哲学关系也很密切。信息科学是一门多边缘的学科,亦称之为横断学科。

信息科学是以信息作为主要研究对象、以信息过程的运动规律作为主要研究内容、以信息科学方法论作为主要研究方法、以扩展人的信息功能(全部信息功能形成的有机整体就是智力功能)作为主要研究目标的一门科学。

信息科学研究内容包括:阐明信息的概念和本质(哲学信息论);探讨信息的度量和变换(基本信息论);研究信息的提取方法(识别信息论);澄清信息的传递规律(通信理论);探明信息的处理机制(智能理论);探究信息的再生理论(决策理论);阐明信息的调节原则(控制理论);完善信息的组织理论(系统理论)等。

信息科学现在主要的研究课题集中在以下六个方面:

第一,信源理论和信息的获取,研究自然信息源和社会信息源,以及从

信息基础理论解析

信息源提取信息的方法和技术。

第二,信息的传输、存储、检索、变换和处理。

第三,信号的测量、分析、处理和显示。

第四,模式信息处理,研究对文字、图像、声音等信息的处理、分类和识别研制机器视觉系统和语音识别装置。

第五,知识信息处理,研究知识的表示、获取和利用,建立具有推理和自动解决问题能力的知识信息处理系统即专家系统。

第六,决策和控制,在对信息的采集、分析、处理、识别和理解的基础上做出判断、决策或控制,从而建立各种控制系统、管理信息系统和决策支持系统。

信息科学的支柱理论包括信息论、系统论和控制论(信息科学的三论)。

1. 信息论

信息论是信息科学的前导,是一门用数理统计方法研究信息的度量、传递和交换规律的科学,主要研究通信和控制系统中普遍存在着的信息传递的共同规律,以及建立最佳地解决信息的获取、度量、变换、存储、传递等问题的基础理论。

2. 控制论

控制论的创立者是美国科学家维纳,1948年他发表《控制论》一书,明确提出控制论的两个基本概念——信息和反馈,揭示了信息与控制规律。控制论是关于动物和机器中的控制和通信的科学,它研究各种系统共同控制规律。在控制论中广泛采用功能模拟和黑箱方法。控制系统实质上是反馈控制系统。负反馈是实现控制和使系统稳定工作的重要手段。控制论中,对系统控制调节通过信息的反馈来实现。

3. 系统论

系统论的基本思想是把系统内各要素综合起来进行全面考察统筹,以求整体最优化。整体性原则是其出发点,层次结构和动态原则是其研究核心。

综合化、有序化是其精髓。

1.2 信息的概念与性质

信息论的出发点是认识信息的本质及其运动规律，它的归宿则是利用信息来解决现实问题。对信息的定义是信息论研究的一个基本内容，信息是一个看似简单实则非常抽象的概念，对它进行准确定义相当困难。到目前为止，关于信息的不同定义已在百种以上，这本身就足以说明，信息概念还没有得到最后的澄清。

1.2.1 信息的概念

1. 广义信息

信息的本质是什么？信息是物质的？信息是精神的？信息既是物质的也是精神的？信息既不是物质的也不是精神的，而是独立于物质和精神的第三存在？这是信息科学对于传统的哲学研究提出的最直接的挑战。这个问题的一切逻辑上可能的上述4种答案，都有坚定的哲学家阵营支持。

目前，大多数哲学家和科学家普遍认为，信息不是物质，信息就是对物质存在和运动形式的一般描述。

物质、能量和信息是构成客观世界的三大要素。三要素中物质是基础，是实体；能量是物质运动的形式，物质可转换成能量，而能量又是改造客观世界的主要动力；信息则表征了物质的存在状态和运动形式。美国安东尼·欧廷教授对物质、能量和信息三者作了深刻的描述：没有物质，就什么东西也不存在；没有能量，就什么事情也不发生；没有信息，就什么东西也无意义。三者的关系如图1-2所示。



图 1-2 构成客观世界的三大要素

可以看出，信息有如下基本特征：

第一特征：信息来源于物质，又不是物质本身。它从物质的运动中产生出来，又可以脱离源物质而寄生于媒体，相对独立地存在。

第二特征：信息也来源于精神世界。既然信息是事物运动的状态和状态变化的方式，而事物运动既可以是物质的运动也可以是精神的运动（思维的过程，比如信息可以被提炼成为知识），那么，精神领域的事物运动当然也就成为信息的一个来源。

第三特征：信息与能量息息相关，但是又与能量有本质的区别。信息与能量都与事物的运动相关联。在一定的意义上可以说，信息与能量两者都是事物运动状态的函数。不仅如此，传输信息或处理信息总需要一定的能量来支持，而控制和利用能量则总需要有信息来引导。

2. 香农信息（狭义信息）

香农信息论面向通信工程，决定了它在理论上具有强烈的通信特色。通信系统的基本问题是：在（统计）噪声背景下，在信息接收端近似地或精确地复制发送端发出的信号波形。

1948年香农在《贝尔系统技术杂志》发表了一篇题为《通信的数学理论》的论文中对信息进行了定义：信息是用来减少随机不确定性的东西，即