



“十二五”国家重点图书出版规划项目
电子与信息工程系列

INFORMATION AND CODING THEORY

信息与编码理论

● 顾学迈 石硕 贾敏 编著



哈尔滨工业大学出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS



“十二五”国家重点图书出版规划项目
电子与信息工程系列

INFORMATION AND CODING THEORY

信息与编码理论

● 顾学迈 石硕 贾敏 编著



哈尔滨工业大学出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

内容简介

信息与编码理论是信息技术领域的基础理论,本书根据作者多年教学实践经验和体会系统地介绍了信息与编码理论的相关内容。全书共 10 章,包括信息概念及信息度量、信源熵和信道容量、信源编码定理、信道编码定理、率失真编码定理、线性分组码、循环码和卷积码等基本内容,同时也介绍了 Turbo 码、LDPC 码、数字喷泉码及多用户信息论等扩展内容。本书力求叙述清晰、深入浅出,着重介绍信息与编码理论的基本概念和基本方法。

本书适合作为高等院校电子信息类学科的研究生及高年级本科生教材,也可作为信息科学及系统工程专业领域教学科研人员的参考书。

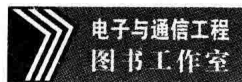
图书在版编目(CIP)数据

信息与编码理论/顾学迈,石硕,贾敏编著. —哈尔滨:哈尔滨工业大学出版社,2014. 10

ISBN 978 - 7 - 5603 - 4911 - 4

I. ①信… II. ①顾…②石…③贾… III. ①信息论—高等学校—教材②信源编码—通信理论—高等学校—教材 IV. ①TN911.2

中国版本图书馆 CIP 数据核字(2014)第 196836 号



责任编辑 李长波

封面设计 刘洪涛

出版发行 哈尔滨工业大学出版社

社 址 哈尔滨市南岗区复华四道街 10 号 邮编 150006

传 真 0451 - 86414749

网 址 <http://hitpress.hit.edu.cn>

印 刷 黑龙江省地质测绘印制中心印刷厂

开 本 787mm×1092mm 1/16 印张 21.75 字数 520 千字

版 次 2014 年 10 月第 1 版 2014 年 10 月第 1 次印刷

书 号 ISBN 978 - 7 - 5603 - 4911 - 4

定 价 42.00 元

(如因印装质量问题影响阅读,我社负责调换)

“十二五”国家重点图书 电子与信息工程系列

编 审 委 员 会

顾 问 张乃通

主 任 顾学迈

副 主 任 张 晔

秘 书 长 赵雅琴

编 委 (按姓氏笔画排序)

王 钢 邓维波 任广辉 沙学军

张钧萍 吴芝路 吴 群 谷延锋

孟维晓 赵洪林 赵雅琴 姜义成

郭 庆 宿富林 谢俊好 冀振元

序

FOREWORD

教材建设一直是高校教学建设和教学改革的主要内容之一。针对目前高校电子与信息工程教材存在的基础课教材偏重数学理论,而数学模型和物理模型脱节,专业课教材对最新知识增长点和研究成果跟踪较少等问题,及创新型人才的培养目标和各学科、专业课程建设全面需求,哈尔滨工业大学出版社与哈尔滨工业大学电子与信息工程学院的各位老师策划出版了电子与信息工程系列精品教材。

该系列教材是以“寓军于民,军民并举”为需求前提,以信息与通信工程学科发展为背景,以电子线路和信号处理知识为平台,以培养基础理论扎实、实践动手能力强的创新型人才为主线,将基础理论、电信技术实际发展趋势、相关科研开发的实际经验密切结合,注重理论联系实际,将学科前沿技术渗透其中,反映电子信息领域最新知识增长点和研究成果,因材施教,重点加强学生的理论基础水平及分析问题、解决问题的能力。

本系列教材具有以下特色:

(1) 强调平台化完整的知识体系。该系列教材涵盖电子与信息工程专业技术理论基础课程,对现有课程及教学体系不断优化,形成以电子线路、信号处理、电波传播为平台课程,与专业应用课程的四个知识脉络有机结合,构成了一个通识教育和专业教育的完整教学课程体系。

(2) 物理模型和数学模型有机结合。该系列教材侧重在经典理论与技术的基础上,将实际工程实践中的物理系统模型和算法理论模型紧密结合,加强物理概念和物理模型的建立、分析、应用,在此基础上总结牵引出相应的数学模型,以加强学生对算法理论的理解,提高实践能力。

(3) 宽口径培养需求与专业特色兼备。结合多年来有关科研项目的科研经验及丰硕成果,以及紧缺专业教学中的丰富经验,在专业课教材编写过程中,在兼顾电子与信息工程毕业生宽口径培养需求的基础上,突出军民兼用特色,在

满足一般重点院校相关专业理论技术需求的基础上,也满足军民并举特色的要求。

电子与信息工程系列教材是哈尔滨工业大学多年来从事教学科研工作的各位教授、专家们集体智慧的结晶,也是他们长期教学经验、工作成果的总结与展示。同时该系列教材的出版也得到了兄弟院校的支持,提出了许多建设性的意见。

我相信:这套教材的出版,对于推动电子与信息工程领域的教学改革、提高人才培养质量必将起到重要推动作用。

哈尔滨工业大学教授 张乃通
中国工程院院士



2010年11月于哈工大

前言

PREFACE

信息技术高速发展与广泛应用已经极大地改变了人们的生活方式和社会的管理方式。随着信息化社会的不断进步,信息的概念早已超越了狭义的通信系统范畴。信息与编码理论不仅是通信系统、信号处理和计算机网络等领域的基础理论,而且在智能控制、管理科学、生物医学以及社会学、语言学等领域也有越来越广泛的应用。

信息论是用概率论和数理统计方法,定量研究信息的获取、加工、处理、传输和控制的一门科学。信息是消息中所包含的新内容与新知识,是用来减少和消除人们对于事物认识的不确定性,是系统保持一定结构、实现其功能的基础。狭义信息论是研究在通信系统中普遍存在的信息传递的共同规律,以及如何提高信息传输系统的有效性和可靠性的一门基础理论。广义信息论被理解为运用狭义信息论的观点来研究相关问题的理论。信息论认为,系统正是通过获取、传递、加工与处理信息而实现其有目的的运动。信息论能够揭示人类认识活动和产生飞跃的实质,掌握信息论有助于探索与研究人们的思维规律和推动与进化人们的思维活动。

通信系统的基本要求是可靠而有效地传递信息。数字化技术是实现信息有效传输的基础,信源编码的目的是在保证信源信息不损失或有限损失的基础上实现信源压缩,进而提高信息传输的有效性。实际通信系统的传输环境通常会受到干扰,包括系统内部和外部噪声以及电波传播时变特性。由于这些干扰的存在,必须采用适当的信道编码方法来保证信息的可靠传输,信道编码理论提供了适合信道环境的信道编码方法。

本书作为高等学校信息类专业基础理论课教材,系统地介绍了狭义信息论和编码理论的主要内容。本书根据作者多年教学实践的经验和体会,参考了国内外多部信息与编码理论的相关教材,结合近年来高等学校专业基础课教学手段和授课对象的特点,按照问题提出、解决思路、数学建模、物理实现和实际应用的思维逻辑,系统地介绍了信息与编码理论的相关内容。

本书共分10章。第1章为绪论。主要介绍通信系统的基本模型,信息的基本概念,通信系统的基本要求,信息论的研究范畴以及信息论与信息科学的内涵。第2章为离散信源熵与交互信息量。主要介绍信息的度量,信息熵的概念及最大熵定理,交互信息量的概念与性质,同时介绍了离散单符号信源、离散多符号信源和马尔柯夫信源的信源熵和交互信息量的计算方法。第3章为信道容量与高斯信道。主要介绍了离散信道的熵速率与信道容量概念,典型离散信道容量的计算方法,连续信源的熵,连续有噪声信道的信道容量等问题;重点介绍了通信理论中最著名的香农公式的形成过程。第4章为信源编码与率失真函数。主要介绍了信源编码的概念和基本思想,重点介绍了香农第一编码定理(无失真信源编码定理)

和香农第三编码定理(限失真信源编码定理),同时介绍了 Huffman 编码的基本方法。第 5 章为信道编码原理。主要介绍了信道编码和译码准则的概念,重点介绍了香农第二编码定理(有噪声信道编码定理),同时介绍了差错控制编码的基本知识。本章的 5.5 节经典序列与信道编码定理作为选读内容介绍了信道编码定理的证明方法。第 6 章为代数编码基础。本章作为编码理论的预备知识介绍了近世代数和线性代数的相关内容,包括集合与映射、群、域、域上多项式和向量空间的相关概念和定理,并通过较多例题来帮助读者掌握编码理论中常用的数学工具。第 7 章为线性分组码。主要介绍线性分组码及汉明码的概念,线性分组码的矩阵描述,循环码和 BCH 码的概念及其多项式描述,同时介绍了循环码和 BCH 码的基本编译码方法。第 8 章为卷积码。主要介绍了卷积码的概念和描述方法,卷积码的维特比译码和序列译码的基本原理和基本方法。本章的 8.4 节介绍了卷积码的软判决译码和 BCJR 译码方法,可作为选读内容。第 9 章为信道编码的其他方法。主要介绍了 Turbo 码、LDPC 码和喷泉码等几种信道编码新方法,主要用于选读,目的是希望读者能够了解和掌握信道编码的最新概念和方法,同时也可以进一步加深理解信道编码理论的基本思想。第 10 章为多用户信息论(有些教材称为网络信息论)。主要介绍了多用户信道模型,重点介绍了几种典型的多址信道、广播信道、中继信道的信道容量计算方法,简单介绍了相关信源编码的基本概念。本章作为经典信息论内容的拓展,可作为相关读者的选读内容。

本书主要作为信息与通信工程学科研究生及高年级本科生的专业基础课教材,也可作为其他信息类相关专业学生了解和掌握信息与编码理论的教学参考书,可根据授课学时数选择部分章节内容。

本书第 1~5 章由顾学迈编写,第 6~8 章由石硕编写,第 9~10 章及全书习题由贾敏编写,顾学迈负责全书策划、修改和统编。博士研究生顾术实参加了第 9 章部分内容编写,硕士研究生孙博雅、博士研究生许恩玮、刘通、马健欣、王雪参加了本书的修改和校对工作,哈尔滨工业大学通信技术研究所多位老师和学生对本书的编写工作给予了关心和指导,在此一并表示诚挚感谢。

本书内容虽经多次授课实践并经过不断补充修改,但由于信息与编码理论相关内容不断发展,限于作者水平,书中必然还有不少疏漏和不足之处,敬请同行专家和读者不吝指教。

顾学迈

2014 年 9 月于哈尔滨工业大学

目 录

CONTENTS

第 1 章 绪论	1
1.1 通信系统与信息	1
1.1.1 通信系统的基本模型	1
1.1.2 信息的概念	2
1.2 信息论的研究范畴	4
1.2.1 狭义信息论	4
1.2.2 一般信息论	5
1.2.3 广义信息论	5
1.3 信息论与信息科学	6
1.3.1 信息论的发展历史	6
1.3.2 信息科学	7
第 2 章 离散信源熵与交互信息量	9
2.1 离散信源的熵	9
2.1.1 信息的度量	9
2.1.2 单符号离散无记忆信源的熵	11
2.1.3 熵函数的性质	13
2.1.4 离散信源的最大熵	14
2.1.5 联合熵与条件熵	16
2.2 离散信道的平均交互信息量	20
2.2.1 离散信道的数学模型	20
2.2.2 交互信息量	22
2.2.3 平均交互信息量	24
2.3 平均交互信息量的特性	27
2.3.1 凸函数和 Jensen 不等式	27
2.3.2 平均交互信息量的简单性质	29
2.3.3 平均交互信息量的凸函数性	33
2.4 离散随机序列信源	37
2.4.1 N 维扩展信源的熵	37

2.4.2 马尔柯夫信源·····	43
习 题 ·····	49
第3章 信道容量与高斯信道 ·····	53
3.1 离散信道的信道容量·····	53
3.1.1 熵速率与信道容量·····	53
3.1.2 几种简单信道的信道容量计算·····	54
3.1.3 信道容量的一般计算方法·····	56
3.1.4 对称信道和准对称信道的信道容量·····	60
3.2 串联信道的交互信息量·····	62
3.3 连续信源的熵·····	67
3.3.1 连续信源熵的定义·····	67
3.3.2 几种连续信源的熵·····	69
3.4 连续信源的最大熵·····	71
3.4.1 连续信源的最大熵·····	71
3.4.2 连续信源的熵功率·····	75
3.4.3 连续信源熵的变换·····	75
3.5 连续有噪声信道的信道容量·····	78
3.5.1 连续有噪声信道·····	78
3.5.2 连续信道的平均交互信息量·····	82
3.5.3 连续信道的信道容量·····	84
3.6* Fano 不等式 ·····	88
习 题 ·····	91
第4章 信源编码与率失真函数 ·····	94
4.1 离散信源编码·····	94
4.1.1 编码器·····	94
4.1.2 单义可译码·····	95
4.1.3 平均码字长度·····	97
4.2 无失真信源编码定理·····	99
4.2.1 编码效率·····	99
4.2.2 无失真信源编码定理 ·····	100
4.3 Huffman 编码 ·····	103
4.3.1 Shannon-Fano 算法 ·····	103
4.3.2 Shannon-Fano 算法的最佳条件 ·····	105
4.3.3 Huffman 算法 ·····	105
4.4 率失真函数 ·····	108
4.4.1 失真度与平均失真度 ·····	108
4.4.2 率失真函数的定义 ·····	113

4.4.3	率失真函数的值域和定义域	114
4.4.4	率失真函数的性质	119
4.4.5	率失真函数的计算	120
4.4.6	限失真离散无记忆信源编码定理	125
习 题		127
第 5 章	信道编码原理	130
5.1	信道编码的基本概念	130
5.1.1	信道编码的含义	130
5.1.2	汉明距离	132
5.2	译码准则	133
5.2.1	译码准则的含义	133
5.2.2	错误译码概率	134
5.2.3	最大后验概率准则	134
5.2.4	最大似然准则	136
5.3	有噪声信道编码定理	138
5.4	信息传输的差错控制方法	142
5.4.1	差错控制方法的分类	142
5.4.2	信道编码的性能评价	145
5.4.3	简单的检错码举例	147
5.5*	经典序列与信道编码定理	149
5.5.1	渐近等分割性和 ϵ 经典序列	149
5.5.2	联合 ϵ 经典序列	153
5.5.3	信道编码定理的证明	157
习 题		161
第 6 章	代数编码基础	163
6.1	集合与映射	163
6.1.1	集合	163
6.1.2	映射	164
6.2	群	164
6.3	域	166
6.4	二元域上多项式	169
6.4.1	二元域上多项式的计算	169
6.4.2	二元扩展域的构造	171
6.4.3	二元扩展域的性质	173
6.5	向量空间	178
6.5.1	向量空间概念	178
6.5.2	矩阵及其变换	181

习 题	184
第 7 章 线性分组码	186
7.1 汉明码	186
7.1.1 线性分组码的描述	186
7.1.2 校验子与标准阵译码	192
7.1.3 线性分组码的最小码距	196
7.1.4 汉明码与完备码	198
7.2 循环码	200
7.2.1 循环码的描述	200
7.2.2 循环码的编译码方法	202
7.2.3 循环码的编码电路	205
7.3 循环码的译码	206
7.3.1 梅吉特译码器	206
7.3.2 捕错译码	209
7.3.3 大数逻辑译码	211
7.4 BCH 码	216
7.4.1 本原 BCH 码	216
7.4.2 非本原 BCH 码	218
7.4.3 RS 码	220
7.4.4 BCH 码的译码	222
习 题	228
第 8 章 卷积码	230
8.1 卷积码的编码	230
8.1.1 卷积码编码器	230
8.1.2 卷积码的监督矩阵	232
8.1.3 卷积码的生成矩阵	234
8.1.4 卷积码的编码举例	237
8.2 卷积码的维特比译码	242
8.2.1 卷积码的图形表示法	242
8.2.2 译码量度	245
8.2.3 维特比算法	247
8.2.4* 卷积码的性能	249
8.3 卷积码的序列译码	253
8.3.1 序列译码算法的基本思路	253
8.3.2 序列译码的堆栈算法	254
8.4* 卷积码的其他译码方法	259
8.4.1 软判决译码	259

8.4.2 卷积码的软输出维特比算法	261
8.4.3 卷积码的 BCJR 译码算法	265
习 题	271
第 9 章 信道编码的其他方法	273
9.1 级联码	273
9.1.1 单级级联码	273
9.1.2 交织码	275
9.2 Turbo 码	276
9.2.1 Turbo 码编码原理	277
9.2.2 Turbo 码的译码	280
9.2.3 Turbo 乘积码	283
9.3 LDPC 码	286
9.3.1 LDPC 码的描述	286
9.3.2 LDPC 码的 Tanner 图	288
9.3.3 LDPC 码的译码	290
9.4 喷泉码	296
9.4.1 喷泉码概念	297
9.4.2 LT 码	299
9.4.3 Raptor 码	302
习 题	304
第 10 章 多用户信息论	305
10.1 多用户信道模型	305
10.2 多址接入信道	308
10.2.1 离散多址接入信道	308
10.2.2 高斯多址接入信道	314
10.3 广播信道	316
10.4 中继信道	324
10.5 相关信源编码	326
习 题	330
参考文献	332

信息论是通信系统工程的基础理论,是人们在日常生活中进行信息交流与传输过程中通过不断研究与实践总结和发展起来的一门应用基础科学。信息论是用概率与统计的数学方法来研究信息的传输、变换、存储、处理和应用问题。1948 年,克劳德·艾尔伍德·香农(Claude E. Shannon)发表了经典论文“通信中的数学理论”,随着研究的不断深入和发展,信息论的应用已逐渐超越了通信工程的范畴,被广泛应用到越来越多的科学领域,成为现代信息科学的基础。本章主要介绍信息的基本概念,信息论的基本范畴,信息论与通信理论的关系,信息论在现代信息科学发展中的地位 and 作用。

1.1 通信系统与信息

1.1.1 通信系统的基本模型

在现代社会中,通信与通信系统的概念已经广泛地被人们所了解。那么准确地讲,什么是通信呢?香农在《通信中的数学理论》中描述,“通信的基本问题就是在某一地点精确或者近似地再生另一地点被选择的信息”。在通信理论教科书中的描述是“通信就是相互传递和交换信息”或“通信就是迅速而准确地传递信息”,等等。

广义上讲,通信是信息或消息的传递,手势、书信、电话、电视及网络等都是消息传递的方式。而狭义上讲,现代通信是指利用光、电能量作为传输媒介的“电子通信”。通信系统的主要任务就是利用电信号作为载体来传递消息。这样,信息、消息、信号就是通信问题中的三个常用术语,它们的准确含义和区别是什么呢?图 1.1 给出了通信系统的基本模型。

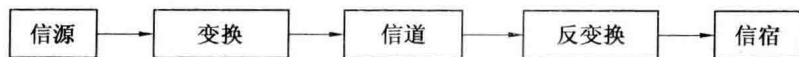


图 1.1 通信系统的基本模型

信源就是信息的源,由信源产生信息,信源可能是人、机器、生物及自然界,因此信源本身是复杂的,而经典信息论研究的信源只是人或机器这类最简单的信源。

信源发出的信息是以消息的形式出现的,消息是由符号、文字、数字、语音、图像等组成的序列。消息的传递一般要借助于载体,因此,载体是消息的物理体现,例如用纸张、磁带、光碟等。在传统的通信理论中,语言为第一载体,文字为第二载体,电磁波为第三载体。在现代信息技术领域,电磁波是一种最广泛应用的消息载体,这种载体即被称为信号。也就是说,信号是一种载体,信号是消息的物理体现,信号是消息具体化。人们通过时域、频域及各



种变换域来研究信号的波形、频谱等特征,从而形成了系统的通信信号处理理论。

信道是传递消息的通道,广义上讲是指信源与信宿之间传递物理信号的媒介和设备,包括声、光、电,甚至包括神经、脉络、意念等。在电信领域中,导线、电缆、光纤、波导、电磁空间都是典型的信道。通常的传递电信号的信道划分包括有线信道和无线信道,恒参信道和变参信道,数字信道和模拟信道,离散信道和连续信道等。

信宿就是消息传递的对象,是信息的接收者,它可以是人、机器、生物及自然界。

在通常情况下,为了有效地传递消息,信源发出的消息要经过必要的加工或处理,这一过程称为变换。在接收端,为了还原消息要进行反变换。变换与反变换通常是由通信设备完成的,典型的数字通信系统的变换包括能量变换、模数转换、信源编码、信道编码和调制解调等。图 1.2 给出了一个数字通信系统的模型。

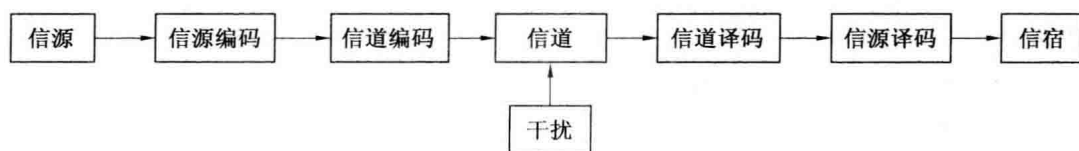


图 1.2 数字通信系统的模型

研究通信系统的目的就是要找到信息传输过程的基本规律,提高信息传输的可靠性、有效性、保密性和安全性,以达到信息传输系统最优化。评价通信系统性能的最重要的两个指标是有效性和可靠性。

通信系统的有效性是指用尽可能少的时间、频带、功率和成本传送一定数量的信息。例如,用信息传输速率描述时间有效性,用频带利用率描述频率有效性等,对于数字通信系统来说,信源编码是提高系统有效性的主要手段。通信系统的可靠性是指要使信源发出的消息经过信道传输以后,尽可能准确地、不失真地再现在接收端。在数字通信系统中通常用误码率来表示,误码率是接收到错误码元在传送总码元数中的比率,通常用信道编码方法来降低系统误码率,误码率越低,可靠性越高。

从信息论的角度讲,提高可靠性和提高有效性常常会发生矛盾,这就需要通信系统的设计者根据具体应用来统筹兼顾。信息论最早就是从解决通信系统有效性和可靠性问题而发展起来的,香农信息论最大的贡献是在理论上证明了提高通信系统有效性和可靠性的统一。

1.1.2 信息的概念

当今社会被称为信息化社会,信息已经融入人们日常生活的各个方面,因此,人们对信息的概念并不陌生,然而,如何给信息下一个准确的定义却是一个困难的问题。现代科学认为,物质、能量和信息是构成客观世界的三大要素;信息是物质和能量在空间和时间上分布的不均匀程度;信息是关于事物运动的状态和规律;物质、能量和信息三者是相辅相成,缺一不可的。

从广义信息论也就是信息科学的角度来讲,信息有以下三个不同层次上的定义。

语法信息:事物运动状态和运动规律本身就是信息。它只研究事物运动可能出现的各种状态,以及这些状态之间的关系,不研究状态的含义和状态变化的影响。香农信息论定义



的信息属于这一层面上的一种描述,即从概率统计的角度来研究事物运动状态及状态变化的关系。

语义信息:事物运动状态和运动规律的具体含义和影响是信息。它研究事物运动状态及运动规律与客观实体的关系,即研究信息的具体含义,语义信息比语法信息更为复杂,更为具体,针对性更强。

语用信息:事物运动状态和运动规律及其含义对观察者的效用才是信息。它研究事物运动状态和规律与使用者的关系,研究信息的主观价值。语用信息比语义信息有更强的主观性和相对性。

语法信息是最抽象的、最基础的层次,狭义信息论和通信系统中讨论的信息主要是语法信息。

信息通常被理解为消息,但严格来说,消息和信号一样只是信息的载体,信息是消息中的内涵。从通信系统的角度讲,如果信宿收到的消息是已知的,就等于没有收到任何信息。人们更关心的是消息中所包含的未知的部分,用概率论的术语来说,就是具有不确定的部分。消息中未知的或不确定的成分通常被称为消息中包含的信息。

由此看来,通信系统中的信息、消息和信号是三个不同层面的概念,通信系统传输的是信号,信号是消息的载体,而消息中的未知成分是信息。信号、消息和信息特征如下:

信号:是信息的物理表达层,是三个层次中最具体的层次。它是一个物理量,是载荷信息的实体,可测量、可描述、可显示,如电信号、光信号、声信号等。

消息:或称为符号,是信息的数学表达层,它虽不是一个物理量,但是可以定量地加以描述,它是具体物理信号的进一步数学抽象,例如文字、数据、代码、图形等。

信息:信息是指各个事物运动的状态及状态变化的方式。人们从来自对周围世界的观察得到的数据中获得信息,信息是信号和消息的更高表达层次。

尽管我们可以对信息的特征给予基本描述,但是如果得到信息的准确定义仍然是比较困难的,这就像我们对物质和能量也很难给出精确定义一样。

我国的《辞海》对信息的描述:① 音讯:“消息”。② 通信系统传输和处理的对象,泛指消息和信号的具体内容和意义,通常需通过处理和分析来提取。信息的量值与其随机性有关,如在接收端无法预估消息或信号中所蕴含的内容或意义,即预估的可能性越小,信息量就越大。我国的《新华词典》(2001年修订版)对“信息”的注释:① 音信:消息。② 信息论中指用符号传送的报道,报道的内容是接收符号者预先不知道的。③ 事物的运动状态和关于事物运动状态的陈述。美国的《韦氏字典》对信息的解释:信息是用以通信的事实,是在观察中得到的数据、新闻和知识。

控制论创始人维纳认为:信息是适应客观世界过程中与客观世界交换的内容。

我国的信息论教材对信息的描述还包括:信息是人、生物、机器等认知主体所感受或所表达的事物运动状态和运动变化规律的方式;信息是人们在适应外部世界和控制外部世界的过程中,同外部世界进行交换的内容。

信息论创始人香农认为:信息是可以传递的,具有不确定性的消息(情报、数据、信号等)中所包含的表示事物特性的部分。香农信息论认为信息的多少等于无知度的大小,因此,香农信息也称为概率信息。

综上所述,信息的特征主要包括:① 接收者在收到信息之前,对它的内容是未知的,所



以信息是新内容、新知识;②信息是能够使认识某一事物的未知程度和不确定性减少的有用知识;③信息可以产生,也可以消失,可以传递、交换、携带、存储和处理;④信息是可以度量的,信息量有大小或多少的差别。

理解信息的概念有几个要点需要说明,①狭义来讲,信息不是事物的本身,信息是抽象的,而消息、文字、指令、数据等是具体的,它们本身不是信息,只是信息的载体。②香农信息论认为,信息的多少等于无知度的大小,人们已知的消息不是信息,而好像、大概、可能之类的不确切的内容包含着信息。因此,有时说“信息冗余”与“信息压缩”是不严格的。③广义来讲,信息是具有新内容的消息,是对于决策有价值的情报,是一切所感知的信号,信息就是新知识等。

应该说明,以上的各种解释都试图对信息做出正确及全面的定义;然而,信息科学仍在不断发展,人们对信息的内涵仍在不断丰富,目前在信息科学领域还没有形成一个被普遍认可的、完整的、确切的定义。也就是说,信息的定义本身也是信息科学研究的内容之一。

1.2 信息论的研究范畴

信息论是源于通信实践和发展的一门新兴学科的基础理论,信息论是应用近代概率与数理统计的数学方法研究信息的基本性质和度量方法,研究信息的获取、传输、存储和处理的一般规律的理论。信息论是信息科学的基础,它强调用数学语言来描述信息科学中的共性问题和解决方法。目前,普遍的观点是把信息论划分为狭义信息论、一般信息论和广义信息论三个层面。

1.2.1 狭义信息论

狭义信息论也就是香农信息论,主要是以概率信息的度量为基础,研究通信系统的有效性和可靠性两大基本问题。通信系统的有效性和可靠性涉及信源、信道和信宿各个部分,其重点是信源编码和信道编码。美国数学家香农是控制论创始人维纳的学生,1948年6月到10月,香农在《贝尔系统技术杂志》上连载发表了影响深远的论文“通信中的数学理论”,1949年又发表了论文“噪声下的通信”。两篇论文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的数学描述,奠定了香农信息论的基础。

狭义信息论讨论的问题主要包括:①信息的度量,即通信系统中的信息是否可以度量,如何度量;②信道容量,典型的通信信道如何模型化,信息的传输速率可以提高到什么程度,极限是多少;③信源编码,信源的效率如何描述,信源的效率如何提高,极限是多少;④信道编码,信道的干扰如何克服,在有干扰的信道上如何提高信息传输速率;⑤率失真函数,在允许一定失真的条件下,信源是否可以进一步压缩,压缩到什么程度;⑥多用户信道容量,在多用户及网络环境下,信道容量如何描述。

香农信息论的主要贡献就是明确回答了上述问题,分别给出了哈特莱(Hartly)公式、信息熵、信道容量、三个编码定理及香农公式等,这些问题也就是本书的主要内容。

香农信息论的基本观点包括三个方面:①非决定论观点,承认偶然性,同时也承认必然性,分析时利用概率论研究事物的统计规律;②形式化假说,通信的目的只是在接收端恢复