

高等学校“十二五”规划教材
电子信息与通信工程系列

信息论与编码

Information Theory & Coding

江晓林 主 编

陈纯锴 苗雨 孙华东 副主编

高等学校“十二五”规划教材
电子信息与通信工程系列

信息论与编码

Information Theory & Coding

江晓林 主 编
陈纯锴 苗雨 孙华东 副主编

内容简介

本书是电子信息与通信工程系列教材之一,是在相关高等学校教师长期教学实践的基础上编写的。

全书共分6章,主要内容包括:绪论、信息量和信源的熵、信源编码、信道及信道编码、接近香农极限的编码和信息率失真函数,章节后面备有相关习题。

本书可以作为普通高等院校电子信息、通信类专业本科生的教材,也可供有关科技人员参考。

图书在版编目(CIP)数据

信息论与编码/江晓林主编. —哈尔滨:哈尔滨

工业大学出版社,2011.4

ISBN 978-7-5603-3205-5

I. ①信… II. ①江… III. ①信息论②信源编码-编
码理论③信道编码-编码理论 IV. ①TN911.2

中国版本图书馆 CIP 数据核字(2011)第 031096 号

策划编辑 许雅莹

责任编辑 李长波

封面设计 刘长友

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 哈尔滨市石桥印务有限公司

开 本 787mm×960mm 1/16 印张 11 字数 226千字

版 次 2011年5月第1版 2011年5月第1次印刷

书 号 ISBN 978-7-5603-3205-5

印 数 1~2 000册

定 价 26.00元

(如因印装质量问题影响阅读,我社负责调换)

总 序

电子信息与通信工程是当今世界发展最快的领域,该技术领域的新概念、新理论、新技术不断涌现,其知识更新速度也是令人吃惊。这就使得从事电子信息与通信工程技术的科技人员要不断学习,把握前沿动态,吸收最新知识。近年来,各高校通过教学改革,在引导学生将最新知识应用于社会实践和市场需求环境,解决实际问题,培养学生实践动手能力、探索性学习能力和创新思维能力等方面取得了可喜成就。

为了培养国家和社会急需的电子信息与通信工程领域的高级科技人才,配合高等院校电子信息与通信工程专业的教学改革和教材建设,哈尔滨工业大学出版社组织哈尔滨工业大学、哈尔滨理工大学、齐齐哈尔大学、佳木斯大学、黑龙江科技学院等多所高校联合编写了这套面向普通高等院校的“电子信息与通信工程系列”教材。

本系列教材的编写目标:结合新的专业规范,融合先进的教学思想、方法和手段,体现科学性、先进性和实用性,强调对学生实践能力的培养,以适应新世纪对通信、电子人才培养的需求。

本系列教材编写要求:专业基础课教材概念清晰、理论准确、深度合理、内容精练,并注意与专业课教学的衔接;专业课教材覆盖面广、深度适中,体现相关领域的新发展和新成果,注重理论联系实际。

本系列教材的编委会阵容强大,编者都是在教学工作第一线的骨干教师。他们具有多年丰富的教学和科研经历,掌握最新的理论知识,具有丰富的实践经验,是一支高水平的教材编写队伍。

本系列教材理论性与工程实践性紧密结合,旨在引导读者将电子信息与通信工程的理论、技术与应用有机结合,适合于高等学校电子、信息、通信和自动控制等专业的教材选取。我深信:这套教材的出版,对于推动电子信息与通信工程领域的教学改革、提高人才培养质量必将起到积极的推动作用,并以其内容的先进性、实用性和系统性为特色而获得成功。

吴群

哈尔滨工业大学教授

2010年4月

前 言

信息论与编码是研究信息的存储、传输、处理、控制和利用一般规律的科学,它主要研究如何提高信息系统的可靠性、有效性等问题,以使信息传输实现最优化。信息论是信息科学发展的源泉,也是信息科学的基础理论。随着电信网、广播电视网和计算机通信网三网合一进程的加快,多媒体技术、计算机技术、空间技术等信息通信网络技术出现了超出人们想象的、前所未有的发展速度,而在这些领域中,均涉及信息论与编码理论。现在人们对于信息论的基本理论已不再感到陌生、抽象和难以理解和掌握,也越来越感到学习和掌握信息理论的必要性和重要性。

信息论与编码是一门既具有广泛的数学理论知识,又有实际工程背景的课程,两者缺一不可。本书是作者多年在信息论与编码方面的教学和科研经验实践总结而加以编写的,全书注重基本概念、基本定理和基本分析方法的论述,结合实例建立概念和数学模型,给出详细的、必要的数学推演过程和证明,力求概念清晰,逻辑性、系统性强,力求做到既有实际应用背景,又有清晰的数学概念和数学思想。

全书内容由浅入深、循序渐进,共分6章。第1章绪论。第2章信息量和信源的熵。介绍信源的分类,信息量、信源的熵的计算方法;对离散信源和连续信源的熵进行了阐述,着重从联合熵、条件熵、平均互信息量(交互熵)三个角度加以展开。第3章信源编码。探讨了定长编码、变长编码、游程编码的原理及其实现方法,着重介绍了变长码中香农编码、霍夫曼编码和费诺编码的实现过程及其性能分析。第4章信道及信道编码。从信道的模型出发,探讨错误概率及其编、译码原则,论述了抗干扰信道编码定理及其逆定理;介绍信道编码的基本思想和几种常见的检错码,包括奇偶校验码、行列监督码和定比码;分析了线性分组码及其纠检错的实现;循环码的描述及编译码实现,并探讨了几种改进的循环码;分析了卷积码的编译码原理,并论述了限失真编码定理。第5章接近香农极限的编码。分析了Turbo码的编译码实现,并进行了性能分析;介绍了LDPC码的定义及描述方法、Tanner图的表示及其编译码的实现和性能分析。第6章信息率失真理论。从失真度、平均失真函数出发,介绍了信息率失真函数的定义及其性质,探讨了最小失真度和最大失真度。

分析了二元对称信源的率失真函数、保真度准则下的信源编码定理、离散无记忆信道的容量-代价函数。

本书内容丰富翔实,对基本概念和基础理论的阐述清晰明了;吸收了新技术和新方法。为了帮助读者掌握基本理论和分析方法,每章附有相应习题。本书可作为普通高等院校电子信息、通信类专业本科生的教材,也可供有关科学技术人员参考。

本书由江晓林任主编,黑龙江科技学院苗雨编写第4章,黑龙江科技学院陈纯锴编写第2,3章,哈尔滨商业大学孙华东编写第5章,黑龙江科技学院江晓林编写第1,6章,全书由江晓林统稿。

限于编者水平,疏漏和不足之处在所难免,殷切希望各界读者批评指正。

编 者

2011年2月

目 录

CONTENTS

第 1 章 绪论	1
1.1 信息	1
1.1.1 信息的概念	1
1.1.2 信息的性质	2
1.1.3 信息科学	2
1.2 通信系统	3
1.2.1 通信系统模型	3
1.2.2 通信系统的基本要求	4
1.3 信息论	6
1.3.1 信息论的发展历史	7
1.3.2 信息论分类	8
第 2 章 信息量和信源的熵	9
2.1 信源分类	9
2.2 自信息量与互信息量	10
2.2.1 自信息量	10
2.2.2 互信息量	11
2.3 平均自信息量——熵	12
2.4 熵的基本性质	13
2.5 熵的唯一性定理	15
2.6 离散集的平均互信息量	20
2.6.1 离散集的平均互信息量定义	20
2.6.2 平均互信息量的性质	20
2.7 马尔可夫信源	23
2.7.1 马尔可夫信源	23

2.7.2	马尔可夫信源的信息熵	25
2.7.3	随机过程的信息量和熵	26
2.8	信源的相关性和剩余度	28
2.9	连续随机变量的互信息和微分熵	29
2.9.1	连续随机变量的熵	30
2.9.2	连续随机变量的联合熵、条件熵以及平均互信息量	32
2.10	信道容量	38
2.10.1	单符号离散信道的信道容量	38
2.10.2	几种特殊信道的信道容量	39
2.10.3	对称信道的信道容量	41
2.10.4	强对称离散信道的信道容量	42
2.10.5	准对称离散信道的信道容量	43
2.10.6	一般离散信道的信道容量	45
	习题	47
第3章	信源编码	50
3.1	信源编码的相关概念	50
3.1.1	无失真信源编码的分类	50
3.1.2	编码器	51
3.1.3	码的分类	52
3.1.4	码树	53
3.2	定长编码定理	55
3.2.1	无失真编码条件	55
3.2.2	定长编码定理	56
3.3	变长码与变长信源编码定理	59
3.4	变长码的编码方法	61
3.4.1	香农编码	61
3.4.2	霍夫曼(Huffman)编码	63
3.4.3	费诺编码	68
3.5	游程编码	71
3.5.1	游程编码简介	71
3.5.2	MH 编码原理	72
	习题	73
第4章	信道及信道编码	75
4.1	信道及译码规则	75

4.1.1	信道的分类	75
4.1.2	信道模型	76
4.1.3	错误概率和译码规则	79
4.1.4	错误概率与编码方法	85
4.1.5	抗干扰信道编码定理及逆定理	91
4.2	信道编码的基本思想	91
4.2.1	差错控制系统分类	92
4.2.2	纠检错基本原理	93
4.3	检错码	96
4.3.1	奇偶校验码	96
4.3.2	行列监督码	98
4.3.3	定比码	98
4.4	线性分组码	99
4.4.1	汉明码的构成	99
4.4.2	监督矩阵 H 和生成矩阵 G	100
4.4.3	伴随式(校验子 S)	102
4.4.4	分组码的纠检错能力	102
4.5	循环码	103
4.5.1	循环码的描述	103
4.5.2	校验子与循环码的编、译码原理	106
4.5.3	译码方法和电路	108
4.5.4	几种改进循环码	110
4.6	卷积码	115
4.6.1	卷积码的基本原理	115
4.6.2	卷积码的译码	120
4.7	限失真编码定理	122
	习题	124
第5章	接近香农极限的编码	126
5.1	Turbo 码	126
5.1.1	Turbo 码的编码	127
5.1.2	Turbo 码的译码	130
5.1.3	译码性能分析	134
5.2	低密度校验码的定义	135
5.2.1	LDPC 码的概念与描述	135

5.2.2 LDPC 码的 Tanner 图表示	137
5.3 低密度校验码的编码和译码	139
5.3.1 LDPC 码的编码	139
5.3.2 LDPC 码的译码	141
5.4 低密度校验码的性能分析	144
习题	146
第 6 章 信息率失真理论	147
6.1 引言	147
6.2 信息率失真函数及其性质	149
6.2.1 失真度	149
6.2.2 信息率失真函数	151
6.2.3 率失真函数与信道容量的对偶性	152
6.2.4 最小失真度和最大失真度	152
6.3 二元对称信源的率失真函数	155
6.4 保真度准则下的信源编码定理	158
6.5 离散无记忆信道的容量-代价函数	160
6.5.1 容量-代价函数	160
6.5.2 容量-代价函数的理解	161
6.5.3 容量-代价函数的性质	162
习题	163
参考文献	165

第 1 章

绪 论

1.1 信 息

通信的目的,就是为了互通消息,实现信息的交换、传递。那么什么是信息呢?

1.1.1 信息的概念

“信息”是当代使用频率很高的一个概念,也是很难说清楚的一个概念。据不完全统计,信息的定义有 100 多种,它们都从不同的侧面、不同的层次揭示了信息的某些特征和性质,但至今仍没有统一的、能为各界普遍认同的定义。但信息的定义是信息论研究的一个基本内容。

20 世纪中叶,信息论、控制论、系统论等标新立异的新理论相继问世,有力地“晃动”着传统的科学框架。克劳德·香农是一位美国数学工程师,作为信息论的创始人,人们认为他是 20 世纪最伟大的科学家之一,人们尊崇香农为信息论及数字通信时代的奠基之父。1948 年克劳德·香农发表论文《A Mathematical Theory of Communication》,建立了信息论这一学科。论文中给出了通信系统的线性示意模型的新思想,首次引入“比特”(bit)一词。香农指出,信息是可以传递的,信息不是事物的本身,信息是抽象的;而消息、情报、指令、数据等本身不是信息。与此同时,美国另一位数学家诺伯特·维纳也发表了题为《时间序列的内插、外推和平滑化》的论文以及题为《控制论》的专著。他指出“通信的根本问题是报文的再生,在某一点与另外选择的一点上报文应该精确地或者近似地重现”。

近代人认为: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;信息就是知识等。

香农信息论认为:信息的多少等于不确定度的大小,信息是能使认识主体对某一事物的未知性和不确定性减少的有用知识。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。

信息是客观存在的,它是人类认识、改造客观世界的主要动力,是人类认识客观世界的更高层次。通信过程中经常会听到三个层次的描述:信号、消息、信息,人们容易混淆,区分不开。信息是指各个事物运动的状态及状态的变化方式,它是抽象的、非物理的,是哲学层表达。消息是具体的、非物理的,可描述为语言文字、符号、数据、图片,能够被感觉到,同时它也是信息的载荷体,是信息论中主要描述形式信息的数学层表达。信号是消息的物理体现,信号最具体,它是一物理量,可测量、可显示、可描述,同时它又是载荷信息的实体,是信息的物理层表达。

1.1.2 信息的性质

信息的基本概念在于它的不确定性,也就是说接收者在收到信息之前,对其内容是未知的,是新知识、新内容;信息可以产生,也可以消失,同时它可以被携带、被存储及处理;信息是可以度量的,信息量有多少的差别,可以归结为以下特性。

(1) 信息的可扩充性:相对物质和能量而言,信息资源没有限度,永远不会耗尽,而且会越来越多,导致信息爆炸,知识爆炸,能源危机。

(2) 信息的可压缩性:通过人脑的归纳和综合,信息可精练和压缩,产生专家系统、知识库。

(3) 信息的可替代性:信息可替代有形物质,信息出口,情报出口。

(4) 信息的可传递性:人与人之间,人与物之间,细胞之间,天体之间。

(5) 信息的可扩散性:总是以各种方式向外部扩散,绝对保密是无法实现的。

(6) 信息的可共享性:信息无法垄断,不能做转手交易。

(7) 信息的有效性:相同信息可有不同的有效性。

1.1.3 信息科学

信息不是静止的,它会产生也会消亡,人们需要获取它,并完成它的传输、交换、处理、检测、识别、存储、显示等功能,研究这方面内容的科学就是信息科学。信息科学是研究信息的概念、相关理论和应用的科学,它是一门新兴科学、边缘学科。

1. 信息科学的特点

(1) 多学科。它与许多基础科学和应用技术有关,互相渗透,如数学、逻辑学、心理学、语言文字学、生物学、控制论、计算机科学、通信技术、仿生学、人工智能技术。

(2) 产业化。它的应用服务于国民经济和社会生活的各个方面,从而形成一个新兴产业——信息产业。

2. 信息科学的研究范围

(1) 信息源:自然信息源(物理,化学,天体,地理,生物)、社会信息源(管理,金融,商

业)、知识信息源。

(2) 信息载体:第一载体(语言)、第二载体(文字)、第三载体(电磁波)。

(3) 信息的采集与转换:传感器,雷达,视、听、触、力,声、光、热、电、磁。

(4) 信息的传输:光、电磁波、神经、意念。

(5) 信息的存储与处理:计算机、视听系统。

1.2 通信系统

1.2.1 通信系统模型

通信就是互相传递和交换信息。

通信技术起源于人类社会初期,如语言、手势、文字、印刷术、电报(1840)、电话(1880)、无线电(1920)、广播、电视、计算机通信(1960)、通信卫星(1964)、Internet(1990)、全球信息网络。通信系统模型如图 1.1 所示。

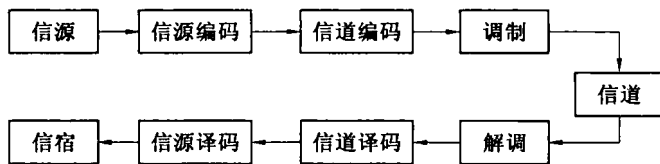


图 1.1 通信系统模型

(1) 信源与信宿

信源是信息的来源,是以符号的形式表现的具体消息。根据不同的分类方法有很多种类,比如数字信源/模拟信源、独立信源/相关信源、无记忆信源/有记忆信源。信宿是信息的接收者,即最终用户。

(2) 编码和译码

为了有效可靠地传递信息,对信息进行必要的加工,主要完成信号能量转换、编码和译码等。编码有信源编码和信道编码两种。信源编码器的作用是把信源发出的消息变换成由二进制码元(或多进制码元)组成的代码组,这种代码组就是基带信号;同时通过信源编码可以压缩信源的冗余度,以提高通信系统传输消息的效率。

信源编码器分为两类:一类是无失真信源编码,适用于离散信源或数字信号;另一类是限失真信源编码,用于连续信源或模拟信号,如语音、图像等信号的数字处理。信源编码器的主要指标是它的编码效率,即理论上能达到的码率与实际达到的码率之比。

信源译码器的作用是把信道译码器输出的代码组变换成信宿所需要的消息形式,它的作用相当于信源编码器的逆过程。

信道编码主要作用是提高信息传送的可靠性。在信源编码器输出的代码组上有目的地增加一些监督码元,使之具有检错或纠错的能力。信道编码的主要作用是增大码率或频带,即增大所需的信道容量。这恰与信源编码相反。

信道译码器的作用是进行译码,实现信号检错或纠错的功能,以提高传输消息的可靠性。

(3) 调制与解调

将各种基带信号转换成适于信道传输的调制信号(已调信号或频带信号)就是调制过程。调制实现的是用基带信号去控制载波信号的某个或几个参量的变化,将信息荷载在其上形成已调信号传输。而解调是调制的反过程,通过具体的方法从已调信号的参量变化中恢复原始的基带信号。

根据所控制的信号参量的不同,调制可分为调幅、调频、调相。调幅,是指使载波的幅度随着调制信号的大小变化而变化的调制方式;调频,是指使载波的瞬时频率随着调制信号的大小而变,而幅度保持不变的调制方式;调相,是利用原始信号控制载波信号的相位。

(4) 信道

通信过程中,信号都要通过信道才能从信源传送到信宿。根据分类的不同,信道的构成也有所不同。

① 按传输介质可将信道划分为有线信道和无线信道,有线信道包括明线、双绞线、同轴电缆、光缆等;无线信道包括长波信道、中波信道、短波信道、微波信道、卫星信道、短波电离层反射信道、散射信道。

② 按信道传输的信号类型可将信道划分为模拟信道和数字信道。

从研究消息传输的观点来说,信道的范围还可以扩大,除包括传输介质外,还可以包括有关的变换装置,比如发送设备、接收设备、调制器、解调器等。

1.2.2 通信系统的基本要求

衡量一个系统的好坏,必然要涉及通信系统的主要指标问题,否则无法衡量。人们对于通信系统的要求是多方面的,评价通信系统的优劣包括信息传输的有效性、可靠性、适应性、标准性、经济性,甚至还包括设备造型的优美性等。从消息的传输来说,有效性和可靠性是通信系统两个最重要指标,然而,它们却又相互矛盾,学习信息论与编码原理,最主要的也就是寻求最佳的方法来提高信息传输的有效性和可靠性。那么何为有效性和可靠性呢?有效性是通信系统传输信息速率的表征,而可靠性是通信系统传输信息质量的表征。传输速度越快,有效性越高;传输信号越准确,可靠性越高。在通信系统中,人们总是希望传输信息既快又准,既有效又可靠。然而有效性和可靠性两者却是矛盾的,这一点,通过以后的进一步学习,将会有更深的体会。一般情况下,要增加系统的有效性,就要降低可靠性,反之亦然。在实际中,常常依据实际系统要求采取相对统一的办法,即在满足一定可靠性指标下,尽量

提高消息的传输速率,即有效性;或者,在维持一定有效性条件下,尽可能提高系统的可靠性。二者如何进行选择,这需要根据通信系统具体要求来进行。

1. 有效性

通信系统的有效性可用传输速率来衡量,传输速率越高,则系统的有效性越好。对应于信号,传输速率快慢通常用以下参量来衡量。

(1) 码元传输速率 R_B

通信是用有限个波形来代表信息的,那么每秒钟所传输的数字波形的数目就是传输速率,即码元传输速率,通常又称为码元速率或波形速率,用符号 R_B 来表示,单位为波特(Baud),常用符号“B”表示(注意,不能用小写)。例如,某系统在 2 s 内共传送 4 800 个码元,则系统的传码率为 2 400 B。码元速率与码元宽度 T_b 有关,具体计算公式为

$$R_B = \frac{1}{T_b} \quad (1.1)$$

然而,仅用码元速率来表征有效性是不够的,因为码元速率仅仅表征的是单位时间内传送数字波形的个数,而没有限定这种波形采用的是多少进制。在实际通信中,由于传输的信息量通常比较大,一般采用多进制来进行信息的传输。码元速率 R_B 表征的是单位时间内传送的码元数,若通信系统能够不变,码元速率应该只与码元宽度 T_b 有关,与信号的进制数无关。根据这点我们可以知道多进制(N)码元速率 R_{BN} 与二进制码元速率 R_{B2} 之间,在保证系统信息速率不变的情况下,相互可转换,转换关系式为

$$R_{B2}/B = R_{BN} \cdot \log_2 N \quad (1.2)$$

式中 $N = 2^k$, 其中 $k = 2, 3, 4, \dots$ 。

(2) 信息传输速率 R_b

信息传输速率简称信息速率,又称为传信率、比特率等。信息传输速率用符号 R_b 表示。 R_b 是指单位时间(每秒钟)内传送的信息量,单位为比特/秒(bit/s),简记为 b/s 或 bps。例如,若某信源在 1 s 内传送 1 200 个符号,且每一个符号的平均信息量为 1 bit,则该信源的 $R_b = 1\,200$ b/s 或 1 200 bps, R_b 也与 N 有关。

(3) 消息传输速率 R_m

消息传输速率也称消息速率,它被定义为单位时间(每秒钟)内传输的消息数,用 R_m 表示。因消息的衡量单位不同,有各种不同的含义。例如,当消息的单位是汉字时,则 R_m 的单位为字/秒。消息速率在实际中应用不多。

(4) R_b 与 R_B 之间的互换

在二进制中,码元速率 R_{B2} 同信息速率 R_{b2} 的关系在数值上相等,但单位不同。在多进制中, R_{BN} 与 R_{bN} 之间数值不同,单位也不同。它们在数值上有如下关系式

$$R_{bN} = R_{BN} \cdot \log_2 N \quad (1.3)$$

在码元速率保持不变条件下,二进制信息速率 R_{b2} 与多进制信息速率 R_{bN} 之间的关系为

$$R_{b2} = \frac{R_{bN}}{\log_2 N} \quad (1.4)$$

(5) 频带利用率

频带利用率指的是传输效率问题,也就是说,我们不仅关心通信系统的传输速率,还要看在这样的传输速率下所占用的信道频带宽度是多少。如果频带利用率高,说明通信系统的传输效率高,否则相反。

频带利用率是指单位频带内码元传输速率的大小,单位为 B/Hz,即

$$\eta = \frac{R_B}{B} \quad (1.5)$$

频带宽度 B 的大小取决于码元速率 R_B ,而码元速率 R_B 与信息速率有确定的关系。因此,频带利用率还可用信息速率 R_b 的形式来定义,以便比较不同系统的传输效率,即

$$\eta / (\text{b} \cdot (\text{s} \cdot \text{Hz})^{-1}) = \frac{R_b}{B} \quad (1.6)$$

2. 可靠性

衡量通信系统可靠性的指标,具体可用信号在传输过程中出错的概率来表述,即用差错率来衡量。差错率越大,表明系统可靠性越差。差错率通常有两种表示方法。

(1) 码元差错率 P_e 。

码元差错率 P_e 简称误码率,它是指接收错误的码元数在传送总码元数中所占的比例,更确切地说,误码率就是码元在传输系统中被传错的概率。其表达式为

$$P_e = \frac{\text{单位时间内接收的错误码元数}}{\text{单位时间内系统传输的总码元数(正确码元数 + 错误码元数)}} \quad (1.7)$$

(2) 信息差错率 P_{eb} 。

信息差错率 P_{eb} 简称误信率,或误比特率,它是指接收错误的信息量在传送信息总量中所占的比例,或者说,它是码元的信息量在传输系统中被丢失的概率。其表达式为

$$P_{eb} = \frac{\text{单位时间内系统传输中出错(丢失)的比特数(信息量)}}{\text{单位时间内系统传输的总比特数(总信息量)}} \quad (1.8)$$

误码率表达式(1.7)和误信率表达式(1.8)常常用来计算系统的误码率和误信率。

这里需要注意的问题是,一定要把码元速率 R_B 和信息速率 R_b 的条件搞清楚。另外还需强调的是,如果已知条件给出码元速率和收端出现错误的信息量,则同样需注意速率转换问题。

1.3 信 息 论

信息论是信息科学的主要理论基础之一,通过利用数学方法研究信息的度量、传递、交换和存储的一门研究信息的基本理论的科学,主要研究可能性和存在问题,为具体实

现更好通信提供理论基础。

1.3.1 信息论的发展历史

1832 年莫尔斯电码对香农编码理论的启发;

1885 年凯尔文研究了一条电缆的极限传信速率;

1922 年卡逊对调幅信号的频谱结构进行研究;

1924 年奈奎斯特证明了信号传输速率和带宽成正比;

1928 年 Hartley 提出信息应该由随机事件样本空间中元素个数的对数值来衡量;

1939 年 Dudley 发明声码器;

1940 年维纳将随机过程和数理统计引入通信与控制系统;

1948 年香农发表《通信中的数学模型》;

1949 年香农发表《保密系统中的通信理论》。

香农信息理论的基本思路,大致可归结为以下三个基本观点。

1. 非决定论观点

我们知道,在科学史上,直到 20 世纪初,拉普拉斯的决定论观点始终处于统治地位。这种观点认为,世界上一切事物的运动都严格地遵从一定的机械规律。因此,只要知道了它的原因,就可以唯一地决定它的结果;反过来,只要知道了它的结果,也就可以唯一地确定它的原因。或者,只要知道了某个事物的初始条件和运动规律,就可以唯一地确定它在各个时刻的运动状态。这种观点只承认必然性,排斥、否认偶然性。

根据通信问题研究对象的特点,信息理论按照非决定论的观点,采用了概率统计的方法,作为分析通信问题的数学工具,因而比以往的研究更切合实际、更科学、更有吸引力。

2. 形式化假说

可提出如下假设:虽然信息的语义因素和语用因素对于广义信息来说并不是次要因素,但对于作为“通信的消息”来理解的狭义信息来说是次要因素。因此,在描述和度量作为“通信的消息”来理解的狭义信息时,可以先把语义、语用因素搁置起来,假定各种信息的语义信息量和语用信息量恒定不变,而只单纯考虑信息的形式因素。

3. 不确定性

对通信过程作进一步分析就可发现,人们要进行通信,不外有两种情形:第一种情况是自己有某种形式的信息要告诉对方,同时估计对方既会对这种信息感兴趣,而又尚不知道这个信息。也就是说,对方在关于这个信息的信息上存在着不确定性。另一种情况是,自己有某种疑问要向对方询问,而且估计对方能够解答自己的疑问。在前一种情况下,如果估计对方已经了解了所欲告之的消息,自然就没有必要通信了;在后一种情况,如果自己没有疑问,当然就不必询问了。