

高等学校“十一五”规划教材

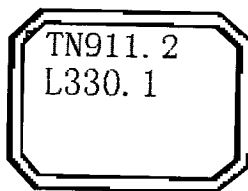
信息论与编码

Xinxilun Yu Bianma

李子臣 主编

中国矿业大学出版社

China University of Mining and Technology Press



高等学校“十一五”规划教材

信息论与编码

主 编 李子臣

副主编 吴 冰 张卷美

中国矿业大学出版社

内容简介

本书以香农信息论为基础,系统地介绍了通信系统中两大编码理论。重点阐述了香农信息论的基本理论、信源压缩编码及信道编码的原理与具体方法,力图将信息理论及编码理论与实际应用紧密结合。

本书可作为高等理工院校电子信息工程、通信工程、信号与信息处理学科的本科生教材,也可作为相关专业研究生、科技和工程技术人员的学习及参考用书。

图书在版编目(CIP)数据

信息论与编码/李子臣主编. —徐州:中国矿业大学出版社,2006.8

ISBN 7-81107-362-5

I. 信… II. 李… III. ①信息论—高等学校—教材
②信源编码—编码理论—高等学校—教材③信道编码—编码理论—高等学校—教材 IV. TN911.2

中国版本图书馆 CIP 数据核字(2006)第 077915 号

书 名 信息论与编码

主 编 李子臣

责任编辑 何 戈

责任校对 杜锦芝

出版发行 中国矿业大学出版社

(江苏省徐州市中国矿业大学内 邮编 221008)

网 址 <http://www.cumtp.com> E-mail: cumtpvip@cumtp.com

排 版 中国矿业大学出版社排版中心

印 刷 徐州中矿大印发科技有限公司

经 销 新华书店

开 本 787×1092 1/16 印张 21.75 字数 540 千字

版次印次 2006 年 8 月第 1 版 2006 年 8 月第 1 次印刷

定 价 45.00 元

(图书出现印装质量问题,本社负责调换)

前 言

随着科学技术,特别是信息技术的发展,信息理论在通信领域中发挥着越来越重要的作用,《信息论与编码》正是为适应这一发展要求而编写的。“信息论与编码”是高等理工院校电子信息工程、通信工程、信号与信息处理学科的一门重要的技术基础课,它的应用领域非常广泛,几乎遍及各个工程技术学科。

随着传统的物质型经济在全世界范围内大规模地向信息型经济转变,人们在思想上也不断地接受着信息化的洗礼。自从美国科学家 C. E. Shannon 于 1948 年创立信息论以来,经过众多通信科技工作者的努力,信息论和编码理论的研究取得了丰硕的成果。在信息的度量、信息传输特性、纠错编码与压缩编码性能极限等理论问题及各种纠错编码和信源压缩编码方法、信道传输容量的研究方面,都取得了重大突破,有力地促进了通信科技的不断发展。特别是进入 20 世纪 90 年代以来,“信息高速公路”、国家信息基础结构等一个又一个崭新的概念,一个又一个全新的事物,接踵而至,越来越多的人渴望学习和掌握信息科学技术知识。为了适应高等学校教学改革和教材改革的需求,依据我国当前电子信息工程、通信工程学科课程设置与教学改革的实际情况,对传统的“信息论基础”、“信息处理与编码”、“数字图像处理与压缩编码技术”课程的教学内容,经适当选择和裁剪组合,以信息与通信系统的优化为主线,力求简明、扼要、适用地编写了这本以介绍信息工程中的基本理论、基本技术、基本方法为主的教材,以达到加强基础、拓宽专业、压缩专业课程的目的。

本书是作者在多年讲授“信息论与编码基础”课程的基础上整理编写而成的,书中包括香农信息论的基本内容及主要结论,压缩编码的基本原理,纠错编码原理、方法以及基本应用等章节。内容涵盖了通信中有关信息处理的基本原理和方法。在编写过程中我们强调基本原理的理解,取材注意循序渐进、难度适中,并注重理论对实际应用的指导作用,在写法上力求条理清楚、逻辑性强,每章还附有相应的习题。因此,本书可作为信息类各专业的高年级本科生教材,也可作为其他专业学生及通信科技人员的参考书。

全书共分九章,主编为河南理工大学的李子臣教授,副主编是吴冰、张卷美。其中第一章、第二章由河南理工大学张卷美编写,第三章由河南理工大学高娜编写,第四章第一节、第三节~第八节由河南科技大学孙士保编写,第四章的第二节由河南理工大学王新良编写,第五章由河南理工大学侯占伟编写,第六章由河南理工大学许焱平编写,第七章由河南理工大学刘娟编写,第八章由河南理工大学吴冰编写,第九章由河南理工大学翟海霞编写。

本书在编写过程中得到了有关领导的大力支持和帮助,许多兄弟院校的同行为本书的编写提出了许多宝贵意见并提供了帮助,在此一并表示衷心的感谢。

编 者

2006 年 5 月

目 录

第一章 信息的概念和特征	1
第一节 信息的定义与性质	1
第二节 信息的分类	3
第三节 信息论与编码研究的主要内容	4
一、信息的度量问题	4
二、通信信道的基本模型	5
三、通信系统中信息的传递过程	5
第四节 信息论发展概况	7
一、信息论的产生与发展	7
二、信息论的早期酝酿	8
三、信息论的建立与发展	8
四、信息论的近期发展	10
五、信息论在 IT 领域中的应用	10
本章小结	12
习题一	12
第二章 信息的度量方法	13
第一节 信息度量概述	13
一、预备知识	13
二、信息量	16
三、互信息量与条件互信息量	22
第二节 熵的概念	26
一、信源熵	26
二、条件熵	30
三、联合熵	31
第三节 熵的基本性质	32
一、信息熵的代数性质	32
二、熵的解析性质	39
第四节 加权熵	44
一、加权熵的定义	44
二、加权熵的基本性质	45
第五节 交互熵	53
一、交互熵的定义	53
二、交互熵的物理意义	54

三、交互熵的性质	56
四、各种熵之间的关系	63
第六节 连续型随机变量的信息量	64
一、连续型随机变量的熵	64
二、几种特殊连续型随机变量的熵	67
三、连续型随机变量的熵的性质	69
第七节 最大熵定理	72
一、限峰值功率的最大熵定理	72
二、限平均功率的最大熵定理	73
三、均值受限条件下的最大熵定理	74
四、熵功率	75
本章小结	76
习题二	76
第三章 信源与信息熵	79
第一节 信源的数学模型及其分类	79
一、信源的数学模型	79
二、信源的分类	81
第二节 离散无记忆信源	82
第三节 离散无记忆信源的扩展信源	84
一、最简单的离散信源	84
二、 N 次扩展信源	85
三、 N 次扩展信源的熵	86
第四节 离散平稳信源	87
一、平稳信源	87
二、平稳信源的熵	88
三、极限熵	89
第五节 马尔可夫信源	91
一、有限状态马尔可夫链	91
二、马尔可夫信源	95
第六节 信源的相关性和剩余度	98
第七节 连续信源及其熵	99
一、连续信源	99
二、连续信源熵的定义	99
三、连续信源的最大熵	101
四、联合熵、条件熵和平均交互信息量	105
本章小结	106
习题三	106

第四章 信道及其容量	108
第一节 信道的分类与描述	108
一、信道的分类	108
二、信道的描述	110
第二节 离散无记忆信道	111
一、离散信道的数学模型	111
二、单符号离散信道	112
三、信道疑义度及平均互信息	115
第三节 离散无记忆的扩展信道	120
一、 N 次扩展信道数学模型	121
二、 N 次扩展信道平均互信息	123
第四节 信道的组合	127
一、串联信道	127
二、并联信道	131
三、和信道	132
第五节 信道容量	132
一、几个基本定义	132
二、离散无记忆信道的 N 次扩展信道	142
三、组合信道容量	142
第六节 信源与信道的匹配	143
第七节 连续信道及其容量	144
一、时间离散信道及其容量	144
二、时间连续信道的容量	146
第八节 多用户信道	147
一、引言	147
二、多址信道	148
三、广播信道	152
四、相关信源的多用户信道	154
本章小结	155
习题四	155
第五章 信源编码	158
第一节 无失真信源编码	158
一、编码器	158
二、分组码	160
三、等长码和等长信源编码定理	161
四、变长编码定理	166
第二节 限失真信源编码定理	168
一、失真测度	169

二、信息率失真函数	170
三、离散信源和连续信源的 $R(D)$ 计算	174
四、限失真信源编码定理	176
第三节 常用信源编码方法	177
一、香农编码	177
二、哈夫曼编码	178
三、游程编码	179
四、算术编码	180
五、预测编码	183
六、变换编码	185
第四节 语音压缩编码	188
一、波形编码 ADPCM 基本原理	188
二、参量编码的线性预测编码器 LPC	190
三、混合编码的各类方法	191
第五节 图像编码	192
一、静止图像压缩编码及其技术标准 JPEG	193
二、活动图像压缩编码及其技术标准 MPEG	194
本章小结	196
习题五	196
 第六章 信道编码	 198
第一节 信道编码的概念	198
一、信道错误图样	198
二、信道编码基本思想	199
三、信道编码的分类	199
四、码距与纠错能力	200
第二节 信道的编码与译码	201
一、错误概率与译码规则	201
二、译码规则的选择准则	203
三、错误概率和编码方法	206
第三节 信道的正、反编码定理	210
一、有噪信道编码定理	210
二、有噪信道编码逆定理	213
三、联合信源信道编码定理	215
第四节 线性分组码	216
一、线性分组码的基本概念	216
二、线性分组码代数结构描述	220
三、线性分组码的译码方法	221
四、线性分组码的纠错能力与码限	225

第五节 循环码	227
一、循环码的基本概念及基本性质	227
二、循环码的矩阵描述及对偶码	229
三、循环码的性质及其应用	232
本章小结	233
习题六	234
 第七章 信息的保密问题	 237
第一节 信息论与密码学	237
一、信息安全面临的威胁	237
二、信息安全的模型	239
三、密码学概述	241
第二节 传统密码体制	243
一、单表代换密码	243
二、多表代换密码	246
三、多字母代换密码	249
四、转置密码	250
第三节 分组密码	252
一、分组加密的基本概念	252
二、数据加密标准	254
三、高级加密标准	261
四、AES 算法过程	264
第四节 公开密钥密码体制	268
一、公开密钥的一般原理	268
二、RSA 体制	269
三、ELGamal 密码	270
四、椭圆曲线密码	272
本章小结	275
习题七	276
 第八章 网络信息理论	 277
第一节 双输入单输出信道的信道容量	277
第二节 离散二址接入信道的容量界限	280
一、计算容量 C_1	283
二、计算容量 C_2	284
三、计算容量 C_{12}	286
第三节 高斯加性二址接入信道的容量界限	288
第四节 单输入双输出信道的信道容量	294
第五节 高斯链式接续信道的容量界限	296

第六节 相关信源的边信息和公信息.....	302
本章小结.....	305
习题八.....	305
第九章 若干应用问题.....	307
第一节 数据压缩问题.....	307
一、数据压缩概述.....	307
二、静态图像数据压缩的基本运算.....	308
三、数据压缩技术分析的主要指标.....	311
四、数据压缩中的其他问题.....	312
第二节 卷积码.....	312
一、卷积码的代数结构.....	313
二、卷积码的图形描述和重量计数.....	318
三、卷积码的 Viterbi 译码算法.....	323
第三节 几种实用的编码问题.....	331
一、汉字编码问题.....	331
二、图形码概述.....	332
本章小结.....	333
习题九.....	333
参考文献.....	335

第一章 信息的概念和特征

[本章重点] 信息的概念;信息的分类;信息论与编码所研究的主要内容。

[本章难点] 信息的度量。

第一节 信息的定义与性质

今天,我们正生活在信息时代,有关信息的新名词、新术语层出不穷,信息产业在社会经济中所占份额越来越大,信息基础设施建设速度之快成了我们这个社会的重要特征之一,物质、能源、信息构成了现代社会生存发展的三大基本支柱。信息是信息论中最基本、最重要的概念,是一个既复杂又抽象的概念。那么什么是信息呢?

科学来源于实践,科学上许多术语往往来自日常生活,信息这个术语同样如此,它是在人类社会互通情报的实践中产生的。例如,当人们收到一封电报、接到一个电话,或早晨从收音机里听到了气象预报,或看了电视里的新闻之后,人们就说得到了“信息”,但到底得到了多少信息呢?对于这个问题,有时仅仅有一些模糊的似乎可以比较的概念而无确切的答案。比如,当打电话的对方告诉的消息是早就知道的,就觉得这个消息没有什么意义,或者说没有带来什么消息。相反,如果电话传来的消息,不但是原来不知道的,而且完全出乎意料,就会觉得这个消息印象很深,给了很多消息。又如,甲乙两人同去听某一学者讲课,由于两人的业务基础不同,他们听到的虽然是同一内容,但听后所得到的新知识,或者说所得到的有用的信息往往是不一样的。科学上所说的信息正是从这个原始的、含糊不清的概念中加以概括、提炼、提高和开拓而得到的,它有严格、确切的含义,有一定的数学模式,并能定量地度量。

在日常生活中,我们收到一封电报、接到一个电话或听到气象预报后,所得到的具体内容是对某一事务状态的描述。例如,气象预报说:“晴转多云”,这是对气象状态的具体描述;电报上的报文“母病愈”,这是对母亲身体健康情况的描述。这些都是人脑以外的客观物质世界表现出来的各种不同的运动状态,而电话中说“他想去北京”,这是存在于他头脑里的思维活动,反映了人的主观世界——大脑物质的思维活动所表现出来的思维状态。

在信息社会中,“信息”这个名词应用十分广泛(如报纸上和电视节目中的经济信息、商品信息、人才交流信息等等),但由于信息科学比起其他学科,如物理学、化学、天文学、生物学等还显得很年轻,人类对信息的认识还很不够。迄今为止,信息还没有形成一个完整的、确切的定义,不同的研究学派对信息的本质及其定义还没有形成统一的意见和认识。综合信息的各种定义,比较一致的定义是:信息是认识主体(人、生物、机器)所感受的或所表达的事物运动的状态和运动状态变化的方式。可以这样认为:

(1) 信息是普遍存在的,它存在于自然界,存在于人类社会,也存在于思维领域,哪里有事物,哪里就有事物的运动状态和方式。

(2) 信息与物质是既有联系又有区别的两个概念:物质是信息的载体,物质的运动是信

息的源泉,但信息只是事物运动的状态和方式,不是事物的本身,信息与物质不能等同。

(3) 信息与能量也是既有联系又有区别的两个概念:传递信息需要能量,驾驭能量则需要信息;然而,信息是事物运动的状态和方式,能量是物体做功的本领,两者之间有质的区别。

(4) 既然信息是事物运动的状态和方式,是认识主体所感知、所表述的事物运动和方式,那么人类要认识事物就必须取得信息。同样,要变革事物也必须要有信息。

由此可知,信息是个十分抽象的概念,信息本身看不见、摸不着,它必须依附于一定的物质形式,如文字、声波、电磁波等,这种运载信息的物质称为信息载体,一切物质都有可能成为信息的载体。

以文字、语音、图像等这些能够为人们的感觉器官所感知的物理现象,把客观物质运动和主观思维活动的状态表达出来就成为消息。从通信的观点出发,构成消息的各种物理现象要具有两个条件:一是能为通信双方所理解,二是可以传递。因此从电报、电话或电视中得到的是一些具体的消息,是描述主、客观各种事物运动状态的消息,而电报、电话、电视等都是这些消息的传递系统。

由此可知,信息是一个十分抽象而又复杂的概念,是人们对客观事物感触到的新知识;消息是信息的载荷者,是描述信息的一种表现形式(如语言、文字等),同一种信息可用不同的消息来载荷(如某一事物,既可用语言来表达,也可用文字来描述),既是对某个事物的描述和反映,也是表示事物内涵的一种形式;而信号则是运载或携带消息的任何物理量,达到迅速有效地传送和交换信息的目的。目前,在信息传输中,主要使用的信号是电信号、光信号和无线电信号。信息包含在消息之中,是通信系统中传送的对象,信息作为现代社会生存发展的第三支柱,具有以下特征:

(1) 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相对独立地存在。

(2) 信息来源于精神世界,但是又不局限于精神领域。

(3) 信息与能量息息相关,但是又与能量有本质的区别。

(4) 信息具有知识的本性,但是又比知识的内涵更广泛。

(5) 信息是具体的,并且可以被(生物、机器等)所感知、提取、识别,可以被传递、存储、交换、处理、显示、检索和利用。

(6) 信息可以被众多用户所共享。

(7) 语法信息在传递和处理过程中永不增值,并且在封闭系统中,语法信息的最大可能值不变。

(8) 接收者在收到信息之前,对它的内容是不知道的,所以信息是新知识、新内容。

(9) 信息是能使认识主体对某一事物的未知性或不确定性的减少的有用的知识。

(10) 信息可以产生,也可以消失,同时信息还可以被携带、储存及处理。

(11) 信息是可以度量的,信息量有多少的差别。

根据以上特征和信息的定义,可以得到信息的一些重要的性质。

(1) 普遍性:信息是普遍存在的。

(2) 无限性:在整个宇宙时空中,信息是无限的;即使在有限的空间(时间有限或无限)中,信息也是无限的。

- (3) 相对性:对于同一个事物,不同的观察者所获得的信息量可能不同。
- (4) 转换性:信息可以在时间上或空间中从一点转换到另一点。
- (5) 变化性:信息是可变的,它可以用不同的载体和不同的方法来载荷。
- (6) 有序性:信息可以用来消除系统的不定性,增加系统的有序性。
- (7) 动态性:信息具有动态性质,一切活的信息都随时间而变化。因此,信息也是有时效、有“寿命”的。
- (8) 转化性:从潜在的意义上讲,信息是可以转化的,它在一定条件下可以转化为物质、能量、时间及其他。
- (9) 共享性:即同一信息可以被无限的人所获得。信息是可以共享的,信息的交流不但不会使交流者失去原有的信息,而且可以获得新的消息。
- (10) 可度量性:信息的数量与质量是可度量的,即信息量。

第二节 信息的分类

上一节介绍了关于信息的概念、特征和性质,使我们对信息有了定性的认识。但要全面、准确地掌握信息的概念,必须对信息有定量的认识。这就要求首先能够确切描述信息。

由前面可知,信息是一个十分复杂的研究对象,要找到一种通用的方法来恰如其分地描述各种各样的信息,即使不是不可能,至少也是很困难的。就算能够找到一种方法可以用来描述一切可能的信息,这种方法也必然是非常笼统、非常一般化,不可能具体细致地刻画各种不同类型的信息。为了具体地描述信息,一定要把信息进行分类,分门别类地进行描述,建立分门别类的描述方法。

按照信息的性质,可以分成语法信息、语义信息、语用信息。

按照观察的过程,可以分成实在信息、先验信息和实得信息。

按照信息的地位,可以分成客观信息和主观信息。

按照信息的作用,可以分成有用信息、无用信息和干扰信息。

按照信息的逻辑意义,可以分成真实信息、虚假信息 and 不定信息。

按照信息的传递方向,可以分成前馈信息和反馈信息。

按照信息的生成领域,可以分成宇宙信息、自然信息、社会信息和思维信息。

按照信息的应用部门,可以分成工业信息、农业信息、军事信息、政治信息、科技信息、文化信息、经验信息、市场信息和管理信息等。

按照信息的来源,可以分成语音信息、图像信息、文字信息、数据信息、计算信息等。

按照信息载体的性质,可以分成语声信息、图像信息、文字信息、电磁信息、光学信息和生物信息等。

按照携带信息的信号性质,还可以分成连续信息、离散信息和半连续信息等。

.....

我们研究信息的目的,就是要准确地把握信息的本质和特点,以便更有效地利用信息。因此,在众多的分类原则和方法中,最重要的是按照信息性质的分类。

按照性质的不同可以把信息分成语法信息、语义信息、语用信息三个基本的类型,其中最重要的也是最抽象的类型是语法信息,它是迄今为止在理论上研究得最多的类型。

语法信息考虑的是事物运动状态和变化方式的外在形式。首先,根据事物运动状态和方式在形式上的不同,语法信息还可以进一步分成有限状态和无限状态;其次,事物运动状态可能是连续的,也可能是离散的,于是,又可以分成连续状态语法信息和离散状态语法信息;再者,事物运动状态还可能是清晰的或者是模糊的,这样,又可以分成状态明晰的语法信息和状态模糊的语法信息。

当然,按照事物运动的方式,还可以把信息进一步细分为概率信息、偶发信息、确定信息和模糊信息。香农信息论主要讨论的是语法信息中的概率信息,本书也是以概率信息为主要研究对象。

上述分类可以用图 1-1 直观地表示。

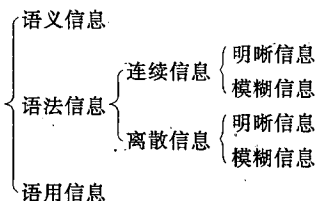


图 1-1 不同性质的信息分类

第三节 信息论与编码研究的主要内容

信息与编码理论的主要内容有信息的度量问题、通信系统与编码问题、信源与信道编码、代数码理论等,我们对此作简要介绍。

一、信息的度量问题

什么是信息?信息能否度量?怎么度量?这些是信息论所必须回答的问题。在回答这些问题之前必须解决两个前提,其一就是如何引进一种新的概念或新的度量,从科学发展的历史来看,这种新概念或新的度量的引进必须符合两个条件:第一个条件是一种新概念或度量的引进必须大体合理,这种新概念所反映的问题与我们以往所了解的特征大体相等;第二个条件是这种新的概念最后可以解决什么问题。这两个条件缺一不可,它们相互呼应与补充,因此我们既要从这种度量的引进来理解它的含义,也要从它最终所解决的问题来加深并确定它的本质含义。香农熵的引进正是从这两个条件出发,它首先把信息的度量看作事物不肯定性的度量来引进,同时在编码理论中发挥了重大作用,只有在信息度量确定后,通信的编码理论才得到了定量化的描述。正因为香农在 1948 年所发表的论文中同时解决这两个问题,所以人们才会把这篇论文作为信息论的奠基性论文。

需要解决的第二个问题是:既然信息是具有广泛含义的概念,那么它的度量与物理量一样,不可能通过一种量来确定它的全部度量,在物理量中有速度、加速度、作用力、电量、电磁强度等不同的量,它们从各种不同的角度来刻画这些物体的运动变化的定量特征。因此在信息量中也有同样的情形,在信息论中我们将介绍它们的多种度量形式,对这种不同类型的度量我们不仅要从它们的引进定义来理解它们的意义,而且也要从它们能解决的问题上进一步加深对它们的理解。

香农熵是香农信息论中信息的度量基础,它与概率分布相联系,以“不肯定性”作为它度量的基础,对此我们在下面还要详细介绍。由香农熵可引进联合熵、条件熵、互信息、信道容量、互熵等,我们都把它们作为信息度量的各种不同形式。这里需要特别指出的一点是香农熵虽以概率分布的不肯定性作为度量的基础,但是随着信息科学的不断发展,人们对香农熵的理解也在日益加深与扩大,香农熵将作为事物复杂度的普遍意义来理解,并随着信息科学的发展,新的信息度量与新的学科分支一定还会产生。

二、通信信道的基本模型

香农信息论的研究对象如图 1-2 所示的通信系统模型,这个模型主要包括以下五个部分:

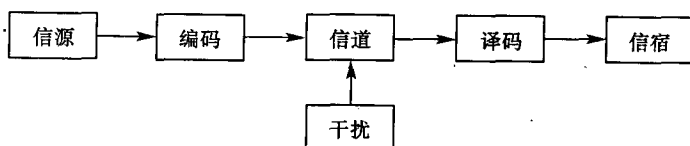


图 1-2 通信系统模型

(1) 信源。信源是产生消息的源头,是消息运动的出发点。信源消息有多种形式,可以是离散的或连续的,也可以是时间序列,它们分别可用离散型随机变量、连续型随机变量及随机过程等数学模型来表示。

(2) 编码。编码是对消息符号进行编码处理的过程。编码包括信源编码、保密编码、信道编码三大类,香农信息论研究得较为深入的是信源编码和信道编码两类。信源编码是对信源输出的消息进行适当的变换和处理,以尽可能提高信息传输的效率;信道编码是为了提高信息传输的可靠性而对信息进行的变换和处理。香农信息论分别用几个重要定理给出了编码的理论性极限,几十年来鼓舞着一批又一批通信理论工作者为达到这些极限而殚精竭虑、苦苦求索,从而推动了编码技术研究的空前繁荣。

(3) 信道。信道是信息的传递媒介。实际的信道有明线、电缆、波导、光纤、无线电波传播空间等。信息的传输不可避免地会引入噪声和干扰,为了分析方便,我们把系统所有其他部分的干扰和噪声都等效地折合成信道干扰,看成是一个噪声源产生的,并叠加于所传输的信号上,这样,信道输出的是已经叠加了干扰的信号。由于干扰和噪声均具有随机性,所以信道的特性同样可以用概率空间来描述,而噪声源的统计特性又是划分信道类型的主要依据。

(4) 译码。译码是把信道输出的编码信号进行反变换,以尽可能准确地恢复原始的信源符号。相应的译码器有信源译码器和信道译码器。

(5) 信宿。信宿即信息传输的目的地。

香农信息论在解决了信息的度量问题之后,主要致力于研究如何提高图 1-2 所示的通信系统中传输的可靠性和有效性。

三、通信系统中信息的传递过程

信息的传递过程可归结如下:首先由信源发出消息,在信息论中把这种消息又称为原始信息,由编码将原始信息变为信号,并进入信道成为信道的输入信号(简称输入信号或入口

信号);输入信号经信道的传送,达到通信的另一端,形成输出信号(或出口信号);输出信号经译码运算把输出信号变为消息,这种消息是原始消息的还原,所以又称为还原消息;还原消息最终由接收者接收,实现通信的目标。

在信道中,信号的产生、发射、传送与接收实际上是一个很复杂的过程,如果输入是脉冲信号,那么它还要变成具有不同振幅的电磁波才能发射。因此,信道终端最初接收到的信号是电磁振荡信号,再由电磁振荡信号变为信道输出的脉冲信号,由译码运算把输出信号还原成消息。

由脉冲信号变为电磁振荡信号,再由电磁振荡信号变为信道输出的脉冲信号,在通信工程中称为调制解调,它涉及信号处理与电磁波理论,在本书中不作讨论。

由于干扰的存在,信道的输出信号可能与输入信号不容,从而形成还原消息与原始消息的不同,这种现象称为通信误差,是通信系统中需要克服的现象。

通信误差的克服一般通过硬件与软件两个途径来解决。所谓硬件就是改造通信系统的物理条件,如元器件的改进、加大信号功率与频带、降低噪声干扰等,但这种改进是有限度的,而且会加大设备成本。软件的改进就是编码方式的改进,这就是在外部硬件环境基本不变的条件下,通过编码与译码函数的设计来实现提高通信速度、降低通信误差的目标,这正是信息论所要讨论的基本内容,因此信息论又称为信息编码理论。而在代数编码理论中,编码问题则归结为纠错码的构造与设计问题,同时也要考虑编码的算法问题,尤其是译码的算法的可计算性问题,所构造的译码算法必须在通信工程中同步实现。

关于信息论研究的具体内容,是一个有争议的问题,有人认为信息论只是概率论的一个分支,这是数学家的观点,当然,这种看法有一定的根据,因为香农信息论确实为概率论开拓了一个新的分支,但如果把信息论限制在数学范围内,这就太狭隘了。也有人认为信息论只是熵的理论,这是某些物理学家的观点。他们对“熵”特别感兴趣,熵的概念确实是香农信息论的基本概念之一,但信息论的全部内容要比熵的概念广泛得多。

归纳起来,信息论研究的主要内容包括以下几个方面:

(1) 通信的统计理论的研究。主要研究利用统计数学工具分析信息和信息传输的统计规律,其具体内容有:① 信息的度量;② 信息速率与熵;③ 信道传输能力——信道容量。

(2) 信源的统计特性。主要包括:① 文字(如汉字)、字母(如英文)统计特性;② 语音的参数分析和统计特性;③ 图片及活动图像(如电视)的统计特性;④ 其他信源的统计特性。

(3) 受信者接收器的研究。主要包括:① 人的听觉和视觉器官的特性;② 人的大脑感受和记忆能力的模拟。这些问题的研究与生物学、生理学、心理学的研究密切相关。

(4) 编码理论与技术的研究。主要包括:① 有效性编码:用来提高信息传输效率,主要是针对信源的统计特性进行编码,所以有时也称为信源编码;② 抗干扰编码:用来提交信息传输的可靠性,主要是针对信道统计特性进行编码,所以有时也称为信道编码。

(5) 提交信息传输效率的研究。主要包括:① 功率的节约;② 频带的压缩;③ 传输时间的缩短,即快速传输问题。

(6) 抗干扰理论与技术的研究。主要包括:① 各种调制制度的抗干扰性;② 理想接收机的实践。

(7) 噪声中信号检测理论与技术的研究。主要包括:① 信号检测的最佳准则;② 信号最佳检测的实践。

由此可以看出,信息论的研究内容极为广泛,是一门新兴的边缘学科,是当代信息科学基本和重要的理论基础,因此,信息论的研究范畴可以概括为以下三方面的内容:

(1) 信息论基础(狭义信息论)。主要研究信息的测度、信道容量以及信源和信道编码理论等问题,即为香农基本理论。

(2) 一般信息论。主要研究通信的一般理论,除了香农理论以外,还包括噪声理论、信号滤波和预测、统计检测与估计理论、调制理论以及信息处理理论等。后一部分是以美国科学家维纳(N. Wiener)为代表,其中最有贡献的是维纳和苏联科学家柯莫各洛夫(А. КОМОПОВ),主要是研究信息传输和处理问题。

虽然维纳和香农等人都是运用概率和统计数学的方法来研究准确地或近似地再现消息的问题,都是为了使消息传送和接收最优化,但它们之间却有一个重要的区别。

维纳研究的重点是在接收端,研究一个信号(消息)如果在传输过程中被某些因素(如噪声、非线性失真等)所干扰时,在接收端怎样把它恢复、再现,从干扰中提取出来。在此基础上,创立了最佳线性过滤理论(维纳滤波器)、统计检测与估计理论、噪声理论等。

而香农研究的对象则是从信源到信宿之间的全过程,是收、发端联合最优的问题,其重点是放在编码。他指出,只要在传输前后对消息进行适当的编码,就能保证在干扰的存在下,最佳地传送和准确或近似地再现消息,为此发展了信息测度理论、信道容量理论和编码理论等。

(3) 广义信息论。广义信息论不仅包括上述两方面的内容,而且包括所有与信息有关的领域,如心理学、遗传学、神经生理学、语言学、语义学甚至于包括社会学中有关信息的问题。

综上所述,信息论是一门应用概率论、随机过程、数理统计和近代代数的方法来研究广义的信息传输、提取和处理系统中一般规律的工程科学;它的主要目的是提高信息系统的可靠性和有效性,以便达到系统的最优化;它的主要内容(分支)包括香农理论、编码理论、维纳理论、检测与估计理论、信号设计和处理理论、调制理论和随机噪声理论等。

由于信息论研究的内容极为广泛,而各分支又有一定的相对独立性,因此本书仅论述信息论的基础理论。

第四节 信息论发展概况

信息论是信息科学的主要理论基础之一,它是在长期通信工程实践和理论研究的基础上发展起来的。信息论从诞生到今天,已有 50 多年的历史,现已成为一门独立的理论科学。本节介绍信息论与编码理论的产生、发展、形成过程,重要的应用领域等,使我们知道理论是如何从实践中经过抽象、概括、提高而逐步形成的。

一、信息论的产生与发展

信息的概念是一个普遍的概念,从细胞分裂到植物、动物的遗传与进化、从语言文字到动作表情都有信息表达与传递,但是信息论的产生、发展与通信、计算机技术的产生、发展密切相关。回顾它的历史,大体可以分为早期酝酿、理论建立、理论发展、理论应用与近代发展等几个阶段,我们先对此作一个简单的回顾。