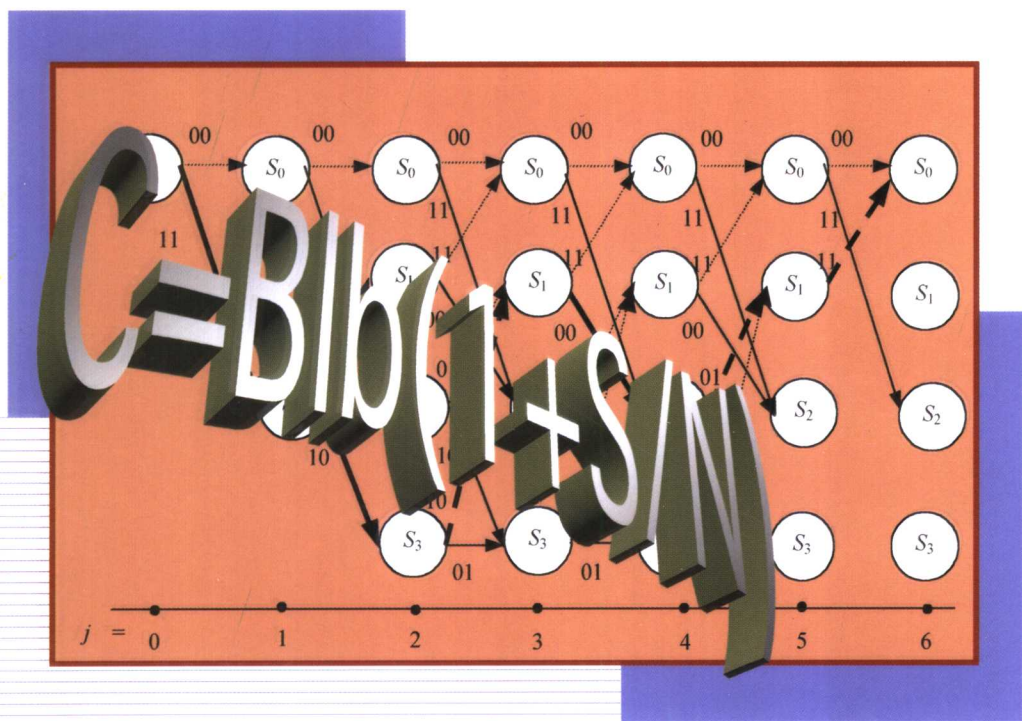


信息与通信工程研究生系列教材

信息论与编码

沈连丰 叶芝慧 编著
杨千里 主审

信息与通信工程研究生系列教材

信息论与编码

沈连丰 叶芝慧 编著

杨千里 主审

科学出版社

北京

内 容 简 介

本书系统、深入地论述了信息论和编码的基本原理和主要技术,内容包括:信息科学及其发展,信息的度量,离散信源及其编码,离散信道及其编码,连续消息和连续信道,网络信息理论初步,信息率失真理论及其应用,差错控制的基本概念,分组码,卷积码,纠突发错误码,保密通信的理论基础及信息理论的广泛应用等。全书把严密的教学语言、合理的物理解释和丰富的应用实例相结合,将经典理论的介绍和最新的研究成果相结合。每章后面都附有习题。

本书可作为高等院校信息与通信工程以及电子类、计算机科学与工程类、自动控制类等学科硕士研究生或高年级本科生的教材,也可作为有关技术、科研和管理人员的参考书。

图书在版编目(CIP)数据

信息论与编码 / 刘连丰,叶芝慧编著. —北京: 科学出版社, 2004

(信息与通信工程研究生系列教材)

ISBN 7-03-013177-0

I. 信… II. ①刘… ②叶… III. ①信息论-研究生-教材 ②信源编码-编码理论-研究生-教材 ③信道编码-编码理论-研究生-教材 IV. TN911.2

中国版本图书馆 CIP 数据核字(2004)第 033524 号

责任编辑:匡 敏 贾瑞娜 / 责任校对:包志虹

责任印制:钱玉芬 / 封面设计:陈 敬

科 学 出 版 社 出 版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

源海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2004年9月第一版 开本:B5(720×1000)

2004年9月第一次印刷 印张:27 3/4

印数:1—3 000 字数:542 000

定价:39.00元

(如有印装质量问题,我社负责调换〈环伟〉)

序

今天,信息论已经成为高校电子、通信以及计算机等专业的必修课,关于信息论,外文的、中文的书籍不少。但是信息科学的快速发展,常常使人不易选到一本适合的信息论教材,而本书恰是我近年看到的值得推荐的一本。教材建设是高等学校建设的主要内容之一。编写、出版一本好的教材,对提高相关课程的教学质量,促进相关专业的良性发展,甚至推动崭新学科的诞生都有着十分重要的作用。东南大学无线电工程系的师生在信息科学方面做了很多研究,数十年锲而不舍,取得了大量成果,2003年将本书列为学校“十五”规划重点教材,体现了学校十分重视教材建设。

自香农的著名论文 *The Mathematical Theory of Communication* 发表以来,信息理论在通信领域发挥着越来越重要的作用,并且迅速渗透到其他相关的自然科学以及社会科学领域,显示出强大的生命力和不可估量的应用前景,产生了信息科学和信息产业。在工程学科领域,信息论是通信理论的重要基石,它强调通过编码来提高通信的有效性和可靠性。正如 R. W. Hamming 所言,从逻辑上说编码理论导致了信息理论,而信息理论又给出了编码的性能极限。因此常将信息理论与编码理论放在一起讲授,本书正是按照这一思路撰写的。

据我所知,本书作者及其带领的青年学者近年来在科研和教学两方面都做了大量的努力。在科研上,他们承担着国家、教育部和省、市以及企业委托的众多研究开发课题;在教学和教材建设上,他们与科学出版社等单位合作,编著和翻译了多部优秀的电子与通信领域的教材,还研制了多个系列的有助于提高教学质量和创新能力的教学实验产品。因此,尽管国内外在信息论与编码方面已出版了一些专著和教材,但是本书能够汲取众家之长,融进自己的科研教学成果,有其显著特色。在内容上,理论分析与工程应用并重,既有必要的数学推导和定理证明,又有大量的最新成果和实际应用的介绍;在语言上,既严谨又通俗风趣;在编排上,注重循序渐进和深入浅出,从浩瀚的素材中作了恰当的取舍,一方面为读者展现经典信息理论的丰硕成果,另一方面引导读者进一步探索更深层次的问题。给我印象最深的是,特别注意对物理概念的透彻剖析和说明。本书内容新颖、系统,反映了信息论与编码的最新理论与实践成果,相信它的出版对相关课程和专业的普及和提高会起到积极的推动作用。

本书的出版,有科学出版社的大力支持,有作者几十年教学经验的总结,有众多青年学者、博士生和研究生的心血,是集体智慧的结晶。在此向为本书的出版付

出辛勤劳动的人表示由衷的祝贺,预祝他们在今后的工作中为我国的教育事业和信息产业的发展做出更大的贡献。

A handwritten signature in black ink, appearing to be '李锐' (Li Rui), written in a cursive style.

2004年3月于北京

前 言

“信息论与编码”是国内外众多大学为研究生或高年级本科生开设的专业理论课之一,更是通信、电子、信息等学科学生的必修课。东南大学无线电工程系从1979年开始为硕士研究生开设本课程,先后使用过多种国内外教材和参考书,但随着信息技术的迅猛发展和研究生数量的大量增长,本课程可用的教科书和参考书无论从数量上还是质量上均不能满足教学要求,作者在教学实践中感到,出版以面向硕士研究生为主并可供其他层次学生和科研人员参考的“信息论与编码”教科书,变得更为迫切。鉴于此,作者在1998年将多年的讲稿整理成讲义并付诸教学实践,本书就是在此基础上经过数次修改完善而成的。它可以作为“信息与通信工程”一级学科硕士研究生“信息论与编码”课程的教科书或相关学科类似课程的参考书,建议授课60学时左右。课程已有完备的多媒体课件,使用本教材的老师可直接向科学出版社免费索取,也可直接与作者联系(电子邮件:lfshen@seu.edu.cn)。

本书共有15章。第1章简单介绍信息科学及其发展,给出了本书的主要内容和课程要求;第2~8章是Shannon信息论的基本内容,分别讨论了信息的度量、离散信源、离散信源的信源编码、离散信道及其信道编码、连续消息和连续信道、网络信息理论初步和信息率失真理论及其应用等,重点论述了各种熵的概念、Shannon的3个定理和信道容量公式等基础知识;第9~13章给出了目前常用的各种差错控制方法,可以说都是Shannon第2定理的应用,首先讨论了纠错编码的基本概念及其本质,接着分别讨论了几种常用的分组码、卷积码、纠突发错误码;第14章讨论保密通信的理论基础,给出了保密通信的技术体制及数学模型、传输线路保密技术的信息论基础、信息保密技术的基础知识等,并对几种密码体制作了介绍;第15章简单介绍了信息理论的广泛应用,包括在多媒体通信中、计算机通信网中、移动通信系统和其他学科的应用等。每章后面都附有习题。

和已出版的国内外同类著作相比,本书力图把严密的数学语言、合理的物理解释和丰富的应用实例相结合,有利于自学;将经典理论的介绍和最新的研究成果相结合,借以增强读者学习的主动性;许多例题和习题取自东南大学无线电工程系历年的“信息论与编码”课程考试试题和博士生“通信原理(含信息论与编码)”入学考试试题,绝大部分来源于教师的思考和科研实践,在语言上努力做到严谨和通俗风趣相结合,以此提高读者对本门课程的学习兴趣。

作者衷心地感谢已故导师吴伯修先生和其他许多在信息论和编码领域做出贡献

献的国内外先师们,是他们将作者带进了这一无限宽广的科学领域,在教学和写作中参考了他们已公开的成果并将其列于本书参考文献中,因此本书首先是学习前人论文著作的体会和在前人成果基础上所做工作的汇报。

本书第1~12章及附录以沈连丰为主编写,第13~15章和9.4节及12.5节以叶芝慧为主编写;瞿慧、董丽君对全书初稿补充了许多理论证明,提出了不少修改意见,选编了大部分习题,完成了索引,承担了全书的图表绘制、公式验算校对和计算机输录等工作;全书的定稿由沈连丰负责,后期的大量具体工作由叶芝慧负责。

作者诚挚感谢对本书的写作、出版给予各种帮助的领导、同仁和学生。著名通信专家杨千里将军在百忙中抽出时间担任本书的主审,这位备受敬重的前辈学长以他渊博的学识、丰富的经验和严谨的学风,给予了作者无私指导;东南大学将本书列为“十五”规划重点教材,校研究生院和教务处的老师给予多方面的支持;作者在东南大学无线电工程系及移动通信国家重点实验室的师长、同事程时昕教授、毕光国教授、尤肖虎教授、张锡昌教授、吴镇扬教授、曹秀英教授、陈明教授、宋铁成教授、徐平平教授等对本课程一直给予关心、指导和帮助,对本书的大纲和初稿提出许多中肯的修改意见。祝宗泰教授将自己多年的讲稿、习题及解答、多届学生的考题等宝贵资料无偿地奉献给作者;博士后刘云、张磊,博士生刘彤、许国军、钱晓勇、傅永生、谷金山、徐智勇、邱晓华、顾燕、李文峰、黄忠虎、鲍淑娣、胡静等分别阅读了书稿的部分章节并提出许多修改意见;东南大学无线电工程系1994~2003级的硕士研究生,在课程讲义和本书初稿的使用过程中提出了许多问题并帮助修改了不少错误,特别是瞿慧和董丽君,差不多牺牲了一年的业余时间为本书做了大量细致的工作;最后,作者要特别感谢科学出版社的钟谊编辑和匡敏编辑,是她们的鼓励、支持和指导,使本书得以出版。

本书内容虽经多届学生使用并得到众多同仁的指正,但是信息理论和编码技术的发展日新月异,限于作者水平,书中一定存在不少疏漏之处,敬请同行专家和读者不吝指正。

沈连丰

2004年3月于东南大学
移动通信国家重点实验室

常用符号

- [D] 失真矩阵
- A 符号集合
- B 符号序列消息信源编码后的代码组长度,带宽
- B_D 失真许可试验信道集合
- \bar{B} 符号序列消息源编码后的代码组平均长度
- b 单符号消息信源编码的代码组长度,突发错误长度
- \bar{b} 单符号消息信源编码的代码组平均长度
- C 信道容量,码字,密文
- D 信道基本符号数,失真限定值
- D_K 解密变换
- d 距离
- d_i 自由距离
- E 错误图样
- E_b/N_0 信噪比
- E_K 加密变换
- $E_r(R)$ 可靠性函数
- F, f 函数
- F 域
- G 群,生成矩阵
- $GF(q)$ Galois 域
- G 增益
- g 译码函数,生成多项式
- H 一致校验矩阵
- $H(x_0)$ 绝对熵
- $I(X, Y)$ 互收息量
- $J(\cdot)$ 辅助函数
- \bar{N} 熵功率
- N_0 噪声功率谱密度
- n_c 编码约束长度
- P_e 错误概率
- p 概率,密度函数
- R, R_i 信息传输速率
- $R(D)$ 信息率失真函数
- r 接收消息序列
- S 状态空间
- S/N 信噪比
- W 码字集合
- X 信源空间,消息集合
- Y 信宿空间,消息集合
- H 熵,疑义度
- $h(x)$ 相对熵
- $H(X)$ 熵
- H_t 时间熵
- \notin 不属于
- σ^2 方差
- λ $GF(q)$ 域的特征
- η 编码效率
- $P(x)$ 概率分布
- β 代码的折算因子
- ϵ 任意小的正数,交叉传输概率
- \in 属于
- Δ 增量,步长
- ρ 相关系数
- \oplus 模运算

目 录

序

前言

常用符号

第 1 章 信息科学及其发展	1
1.1 通信系统的基本概念	1
1.2 信息科学的有关概念	4
1.3 信息理论的研究内容	8
1.4 香农信息论概述.....	10
1.5 主要内容及课程要求.....	10
习题	11
第 2 章 信息的度量	12
2.1 度量信息的基本思路.....	12
2.2 信源熵和条件熵.....	15
2.3 互信息量和平均互信息量.....	17
2.4 多维随机变量的熵.....	26
本章小结	28
习题	28
第 3 章 离散信源	30
3.1 离散信源的分类及其描述.....	30
3.2 离散信源的熵.....	36
3.3 信源的冗余度.....	42
3.4 信源符号序列分组定理.....	45
3.5 平稳离散信源及其性质.....	48
本章小结	50
习题	51
第 4 章 离散信源的信源编码	53
4.1 信源编码的模型.....	53
4.2 信息传输速率和编码效率.....	56
4.3 单义可译定理.....	59
4.4 无失真信源编码定理.....	65

4.5	几种典型的信源编码方法	68
4.6	汉字编码方法及其讨论	80
4.7	图像的信源编码	88
4.8	误码对信源译码的影响	92
	本章小结	92
	习题	93
第5章	离散信道及其信道编码	94
5.1	信道的分类及其描述	94
5.2	无扰离散信道的传输特性	95
5.3	有扰离散信道的传输特性	99
5.4	译码准则	109
5.5	有扰离散信道的信道编码定理	113
5.6	信道编码定理的应用	123
	本章小结	125
	附录 Fano 不等式的证明	125
	习题	128
第6章	连续消息和连续信道	131
6.1	连续消息的信息度量	131
6.2	连续消息在信道上的传输问题	146
6.3	香农信道容量公式	147
6.4	连续消息的识别和理想接收机	150
6.5	连续信源的数字处理及其编码	153
	本章小结	161
	习题	161
第7章	网络信息理论简介	164
7.1	多址信道及其传输特性	164
7.2	广播信道	173
	本章小结	176
	习题	176
第8章	信息率失真理论及其应用	177
8.1	失真函数和平均失真度	177
8.2	信息率失真函数	182
8.3	信息率失真函数 $R(D)$ 的计算	191
8.4	保真度准则下的信源编码定理	196
8.5	信息率失真函数与信息价值	205

本章小结	209
习题	210
第 9 章 差错控制的基本概念	212
9.1 差错控制系统及其理论基础	212
9.2 纠错编码的基本概念及其本质	216
9.3 纠错编码方法的性能评价	221
9.4 基于图形的编码	223
本章小结	231
习题	231
第 10 章 线性分组码	232
10.1 近世代数的基础知识	232
10.2 线性分组码的编码	250
10.3 线性分组码的译码	254
10.4 线性分组码举例	258
本章小结	259
习题	260
第 11 章 循环码	262
11.1 循环码的描述	262
11.2 循环码的编码和译码	269
11.3 二元 BCH 码	277
11.4 多元 BCH 码和 R-S 码	295
11.5 其他 BCH 码	299
本章小结	299
习题	300
第 12 章 卷积码	302
12.1 卷积码的编码及描述	302
12.2 卷积码的最大似然译码	310
12.3 卷积码的序列译码	314
12.4 多进制卷积码	320
12.5 Turbo 码	321
12.6 恶性卷积码及译码的错误扩展	328
12.7 卷积码的汉明距离及各种好码的概念	329
本章小结	330
习题	331
第 13 章 纠突发错误码	333

13.1	纠突发错误码的定义及基本性质	333
13.2	法尔码	335
13.3	交错码	338
13.4	伯顿码	340
13.5	纠突发错误卷积码	341
13.6	岩垂码	342
13.7	纠突发错误循环码的译码	345
13.8	纠突发和随机错误码	347
	本章小结	350
	习题	350
第 14 章	保密通信的理论基础	352
14.1	保密通信的技术体制及数学模型	352
14.2	信息保密技术的基础知识	357
14.3	数据加密标准 DES	363
14.4	流加密技术	368
14.5	公共密钥密码系统	369
14.6	PGP 标准	371
14.7	通信网络安全的加密方案	375
	本章小结	376
	习题	376
第 15 章	信息理论的广泛应用	378
15.1	多媒体通信中的编译码技术	378
15.2	计算机网络通信中的编译码技术	387
15.3	信息论在数字移动通信系统中的应用	394
15.4	信息论在宽带无线接入系统中的应用	406
15.5	信息论在其他领域的应用	411
	本章小结	413
	习题	413
附录 A	思考题	414
附录 B	综合练习题	417
	参考文献	422
	索引	426

第 1 章 信息科学及其发展

1.1 通信系统的基本概念

1.1.1 从通信谈起

人们常说通信的目的是传输信息,也说是传递消息。可以这么说,自从有了人类,就有了伴随着人类的通信。手势、书信、电话、电视等都是传递消息的方式,因此都是通信,但是近代的通信特指“电通信”。通信系统的主要任务是借助电信号这种载体来传输消息。这样,信息(information)、消息(message)、信号(signal)就成为人们最常使用的 3 个基本术语,那么,它们的确切定义及基本含义是什么?下面通过如图 1.1 所示的最简单的通信系统模型来说明。

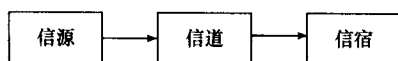


图 1.1 最简单的通信系统模型

图 1.1 中给出了 3 个基本单元:信源(source)、信道(channel)和信宿(destination)。

一般来说,信源是提供消息的人、设备或事物。例如,打电话时,人发出声音或电信设备发出电信号;因特网(Internet)上计算机系统中制作好的可供浏览的网页;广播、电视发送的节目信号;天体无线电信号等都是信源。因此信源本身颇具复杂性。但从传输的角度,或被人所感知的角度,人们最感兴趣的是信源给出的消息。

消息是由符号、文字、数字、语声、图像等组成的序列。消息通常具有如下特征:

- (1) 消息可以产生、传递和获取

事物的运动产生消息,这说明它是客观的、可感知的,但它看不见、摸不着。

传递消息要借助于载体(或媒质),用载体传送,从载体中取出。因此,载体是消息的物理体现,如在纸上写字,用磁带、光碟、存储器等记录语音、图像、数据等,而最值得研究的载体是电的、磁的、光的载体,这就是各类信号。也就是说,信号是一种载体,是消息的物理体现,它使消息能够看得见,变得具体化。人们通过时域、频域及各种变换域来研究信号的波形、频谱等结构,构成

比较完善的信号理论。

(2) 消息有具体内容

消息的内容是客观存在的。但对消息的接受者来说,消息有新、旧或已知、未知之分,还有有用、无用及有大用、小用之别。这就是说,客观存在的消息,对不同的接收者来说,其性质、作用有所不同,甚至截然不同,这就促使人们去研究消息的内涵以及人的判断等主观因素,从而产生了信息理论及其各种分支。

从图 1.1 中可以看出,信源产生了消息,信道传递了消息,而信宿则是接收了消息。消息从信源通过信道到信宿,如何有效、可靠地传输,是通信系统要解决的根本问题。

1) 有效性(validity)。消息若能在信源中先进行去粗取精的处理,则必能提高通信的有效性。对于数字通信来说,这就是信源编码(source coding)要解决的主要问题。

2) 可靠性(reliability)。信宿对接收到的“消息”若能够进行判断、评估,进行去伪存真的处理,则必能提高通信的可靠性。对于数字通信来说,信道编码(channel coding)的主要任务就是解决这个问题。

信息理论最初就是从解决通信系统的有效性和可靠性而发展起来的,从信源、信道、信宿的特性入手,逐步发展成应用极为广泛的新兴学科——信息科学。为此,本章首先给出信源、信道、信宿的比较确切的定义,然后再讨论信息的涵义。

信源:消息的源,通常指向通信系统提供消息的人、设备或事物。信源输出的是以符号形式出现的具体消息,它载荷信息。信源输出的消息有多种形式,亦有不同的分类方法。例如,从信源发出的消息在时间上和幅度上的分布情况这一角度,可分为离散信源(discrete source)和连续信源(continuous source)。前者是指该信源发出的消息在时间和幅度上都是离散分布的,如字母、文字、数字等符号组成的符号序列或者单个符号;后者则指在时间或幅度上是连续分布的,如话音、图像等。

信道:传递消息的通道,广义上是指从信源到信宿间传递物理信号的媒质和设施。信道的种类很多,如电信中常用的架空明线、同轴电缆、波导、光纤、传输电磁波的空间等都是信道。也可以从信道的性质或其传送的信号情况来分类,如无干扰信道和有干扰信道、恒参信道和变参信道、离散信道(discrete channel)和连续信道(continuous channel)、单用户信道和多用户信道等。

信宿:消息传递的对象,即接收消息的人、设备或事物。根据实际需要,信宿接收的消息,其形式可以与信源发出的相同,也可以不同。当两者形式不同时,存在一个映射关系。

1.1.2 信息的一般含义及定义

从通信的实质意义来讲,如果信宿收到的消息是已知的,则等于没有收到任何

消息。因此,人们更感兴趣的是消息中所包含的未知的成分,用概率论的术语来说,就是具有不确定性的成分。消息中这种未知的或不确定的成分,通常被称为消息中所包含的信息。

由此看来,通信中的信息、消息和信号是紧密相连的:通信系统传输的是信号,信号是消息的载体,消息中的未知成分是信息。

人们对消息和信号有直观的认识,教科书对它们也有明确的定义,但对信息到底是什么并不是那么清楚。事实上,如何给信息下定义始终是困扰大家的一个问题。我国的《辞海》对信息作了如下解释:

信息: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。通常须通过处理和分析来提取。信息、物质和能量被称为系统的三大要素。信息和物质的量值与信息的随机性有关,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。

上述解释首先肯定了信息、消息和信号的密不可分性。为此可将这三者合在一起说明如下:

信息·消息·信号: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。信息存在于自然界,也存在于人类社会。可以这样说:哪里有事物,哪里就有事物的运动,同时哪里就产生了信息。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。信息的特征为:①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。②信息是能使认识某一事物的未知性或不确定性减少的有用知识。③信息可以产生,也可以消失,同时信息可以被携带、存储及处理。④信息是可以量度的,信息量有多少的差别。消息是指包含有信息的语言、文字和图像等。消息中的不确定成分是信息。在通信中,消息是指担负着传送信息任务的单个符号或符号序列。这里的符号包括文本、数据、语言、图形和图像等。消息是具体的,它载荷信息,但它不是物理性的。信号是消息的物理体现,为了在信道上传输消息,就必须把消息加载到具有某种物理特征的信号上去。信号是信息的载荷子或载体,它是物理性的。在近代通信中信号往往是电信号和光信号。

应该说,以上两种解释都力图对信息做出正确及全面的定义。然而,信息科学还在飞速发展,人们对信息也不断赋予新的内涵,对信息的理解不但要考虑信息理论发展的历史,也要考虑它的将来,因此,正确及全面地理解它的涵义远比记住它的定义重要。为此,先讨论信息科学的有关概念及信息理论研究的主要内容,再回过头来看看这些定义,也许会有更深的理解。

1.2 信息科学的有关概念

1.2.1 信息科学的创立

1. 信息论的创始人

信息科学的创立是以信息理论(简称信息论)的建立为标志的。众所周知,美国科学家香农(C. E. Shannon)在1948年和1949年先后发表的两篇文章 *the Mathematical Theory of Communication* 和 *Communication in the Presence of Noise*, 奠定了信息论的基础。香农在这两篇论文中,讨论了信息的度量、信息特征、信息传输速率、信道容量(channel capacity)、干扰对信息传输的影响等问题,全部理论建立在信息是可以度量的基础上,但他没有给出信息的定义,只是提出问题,分析问题,得出结论。至于结论的意义,他没作多少解释。

同一时期,美国另一科学家维纳(N. Wiener)出版了 *Extrapolation, Interpolation and Smoothing of Stationary Time Series* 和 *Control Theory* 两本名著。维纳是控制论领域的专家,重点讨论微弱信号的检测理论,并形成了信息论的一个分支。他对信息作了如下定义: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内容的名称。看来维纳力图从信息的本质来下定义,但人们同外部世界交换的内容实在太多,而最本质、最基础的内容是物质和能量。那么,信息是不是物质?有没有能量的属性?与物质、能量的关系是什么?

Shannon 和 Wiener 没有给出上述问题的解答,随着信息理论向通信以外的学科特别是向社会科学领域的渗透,人们为了弄清这些问题进行了很多研究,但直到现在也还没有定论。

因此,人们一般认为信息论的创始人是香农和维纳,但由于香农的贡献更大,所以更多的人认为香农更合适。

2. 信息的特征

前面的讨论已经涉及信息的一般特征,但从更一般的意义来说,信息具有如下特征:

- 1) 未知性或不确定性。这是信息的最基本属性,否则就不是信息。
- 2) 由不知到知,等效为不确定性的集合的元素的减少。这样,就可以基于集合论和概率论来描述信息的传输。
- 3) 可以度量。这是信息理论的基础,使得信息可以比较大小、价值等。
- 4) 可以产生、消失,可以被携带、存储、处理。由此可见信息既有物质的某些属性,又不同于物质。信息和物质都可以被携带、存储、处理,但物质不灭,它只能从一种形态变化成另一种形态,而信息可以产生和消失。

5) 可以产生动作。说明信息能够发挥作用,获得信息后可能产生结果。可见信息既有能量的某些属性,又不同于能量。能量产生的动作是客观的,而信息的影响含有客观和主观的双重因素。

下面通过一个具体的例子来进一步理解信息的上述特征。

某人于某日得知如下两条消息:①2021年9月11日上午9时,一颗小行星将与地球相撞。②2022年9月11日晚上9:00将发生月食。试比较它们所包含的信息。

这两条消息都属于天文学的范畴,如果预先不知道,则他(她)都会从中获得信息;但这两个事件发生的概率有很大的不同,显然前者的出现概率极小,如果它真的出现了,必将带来极大的信息;若真正度量它们,还应该知道天文学范畴的事件集合及它们的出现概率;产生的影响和引起的动作肯定不同,前者可能使得相当多科学家倾其全部智慧及精力去研究,后者则可能不会有上述反响,但并不排除有人并不理会前者而仅对后者感兴趣。因此,一般来说,前者所包含的信息远远大于后者,但还会因人而异。

3. 信息理论要解决的基本问题

从信息的特征可以看出,信息理论要回答的问题很多,由于它最初是由通信发展而来的,因此首先将回答通信中遇到的问题。图 1.2 是通信系统的一般框图,与图 1.1 相比,加上了信源编码、信道编码等 9 个部分,其目的是体现“有效、可靠地传送信息”。结合信息的特征,下面一些最基本的理论问题等待着信息论的回答:

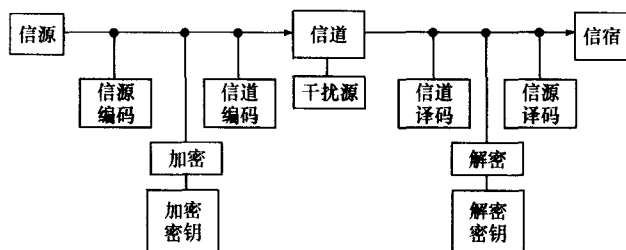


图 1.2 通信系统的一般框图

- 1) 什么是信息? 如何度量?
- 2) 在给定的信道中,信息传输有没有极限?
- 3) 信息能否被压缩和恢复? 极限条件是什么?
- 4) 从实际环境(如干扰、噪声)中抽取信息,极限条件是什么?
- 5) 在允许一定失真(distortion)的条件下,信息能否被更大程度地压缩? 极限条件是什么?