

高等学校教材

信息论基础

陈前斌 蒋 青 于秀兰 编著



高等教育出版社
Higher Education Press

高等学校教材

G201/24

2007

余韵谷夫

信息论基础

陈前斌 蒋 青 于秀兰 编著



高等教育出版社
Higher Education Press

内容简介

本书重点介绍由香农理论发展而来的信息论的基本理论以及编码理论,联系实际通信系统,用较多的例题和图示展示基本概念的运用方法,尽量减少繁杂的公式、定理证明;叙述上力求概念清楚、重点突出、深入浅出、通俗易懂;内容上力求科学性、先进性、系统性与实用性相统一。

全书共分 8 章,内容包括绪论、离散信源及其信息度量、离散信道及其信道容量、无失真信源编码、波形信源和波形信道、限失真信源编码、有噪信道编码和网络信息论初步。

本书既可作为高等学校通信工程、电子信息工程和其他相关专业本科学生教材,也可以作为信息、通信、电子工程相关专业技术人员的参考书。

图书在版编目(CIP)数据

信息论基础/陈前斌,蒋青,于秀兰编著. —北京:
高等教育出版社,2007.12

ISBN 978-7-04-022515-0

I. 信… II. ①陈… ②蒋… ③于… III. 信
息论—高等学校—教材 IV. G201

中国版本图书馆 CIP 数据核字 (2007) 第 172614 号

策划编辑	吴陈滨	责任编辑	关旭	封面设计	张楠
责任绘图	朱静	版式设计	马敬茹	责任校对	王超
责任印制	尤静				

出版发行 高等教育出版社
社址 北京市西城区德外大街 4 号
邮政编码 100011
总机 010-58581000

经销 蓝色畅想图书发行有限公司
印刷 北京四季青印刷厂

开本 787×960 1/16
印张 15.5
字数 280 000

购书热线 010-58581118
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landaco.com>
<http://www.landaco.com.cn>
畅想教育 <http://www.widedu.com>

版次 2007 年 12 月第 1 版
印次 2007 年 12 月第 1 次印刷
定价 19.70 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 22515-00

前 言

信息论是信息科学的主要理论基础之一,是一门具有高度概括性、综合性和应用广泛的学科。它利用数理统计方法研究信息的计量、发送、传递、交换、接收和存储中的一般规律和本质属性。随着以计算机技术、通信技术、网络技术为代表的现代信息技术的高速发展,信息技术成为当今社会渗透性最强、覆盖范围最宽、应用范围最广的高新技术领域之一。信息论的一些基本理论在通信、计算机网络、数字音像、信息处理等工程实践中得到广泛应用,并指导信息系统的最优化设计,包括在如何提高信息系统的可靠性、有效性、保密性等方面提供理论基础。随着信息革命和信息科学的发展,信息技术的研究和使用,目前信息论研究范畴远远超出了通信及相近学科,已应用到生物学、生理学、人类学、物理学、化学、电子学、经济学和管理科学等多个学科。

在这种形势下,信息论作为一门专业基础课程列入了许多高校的教学计划。同时,不同的高校和不同的专业,结合各自培养方案的特点,对信息论教学内容有不同的取舍和偏重。本书面向通信工程、电子信息工程、信息工程和电子信息科学与技术等专业教学需要,着重讲授以经典信息论内容为主的信息论基础,特别强调概念阐述清楚、理论体系相对完整,围绕信息传输系统这一工程应用背景建立信息论的数学分析方法。读者在学习时应结合信息传输系统的数学模型理解信息论中基本理论方法,培养抽象分析能力和系统工程概念。

本书参考学时为 32~48 学时,既可作为高等学校通信工程、电子信息工程和其他相近专业本科学生的教材,也可以作为信息、通信、电子工程相关专业科技人员的参考书。

本书由陈前斌担任主编,其中第 1 章和第 5 章由陈前斌编写;第 2 章、第 3 章和 8.1 节由蒋青编写;第 4 章、第 6 章、第 7 章和 8.2 节由于秀兰编写。全书由陈前斌统编定稿。

南京邮电大学徐澄圻教授担任本书的主审,对本书进行了仔细审阅,并提出了许多宝贵意见和修改建议,在此表示诚挚的谢意。

限于编者的水平,书中错误在所难免,敬请各位老师、学生、相关读者批评指正。

编者

2007 年 4 月

目 录

第 1 章 绪论	1
1.1 信息的概念	1
1.1.1 信息	1
1.1.2 香农“信息”的定义	3
1.1.3 “信息”的度量	4
1.2 信息论的形成和发展	7
1.3 信息论的研究范畴	10
1.3.1 通信系统模型	10
1.3.2 信息论的研究内容	14
第 2 章 离散信源及其信息度量	18
2.1 信源的分类及描述	18
2.2 信源的数学模型	18
2.2.1 单符号离散信源	19
2.2.2 简单的连续信源	20
2.2.3 多符号离散信源	20
2.2.4 多维连续平稳信源和波形信源	21
2.3 信息的度量和信源熵	22
2.3.1 自信息、联合自信息与条件自信息	22
2.3.2 互信息和条件互信息	25
2.3.3 信息熵	28
2.4 信源熵的基本性质和定理	32
2.5 离散无记忆扩展信源	36
2.6 离散平稳信源	38
2.6.1 离散平稳信源的数学模型	39
2.6.2 离散平稳信源的信息熵和极限熵	39
2.7 马尔可夫信源	46
2.7.1 马尔可夫信源的定义	47
2.7.2 马尔可夫信源的信息熵	50
2.8 信源的相关性和剩余度	54

习题	57
第3章 离散信道及其信道容量	61
3.1 信道的分类及其描述	61
3.2 单符号离散信道的数学模型	62
3.3 平均互信息及其特性	66
3.3.1 平均互信息的定义	66
3.3.2 损失熵和噪声熵	67
3.3.3 平均互信息的特性	69
3.4 信道容量及其计算方法	75
3.4.1 信道容量的定义	75
3.4.2 无噪信道的信道容量	76
3.4.3 对称离散信道的信道容量	77
3.4.4 准对称离散信道的信道容量	80
3.4.5 一般离散信道的信道容量	81
3.5 多符号离散信道的数学模型	84
3.6 离散无记忆扩展信道的信道容量	88
3.7 组合信道及其信道容量	90
3.7.1 独立并联信道及其信道容量	90
3.7.2 串联信道及其信道容量	91
3.8 信源与信道的匹配	93
习题	94
第4章 无失真信源编码	98
4.1 信源编码器	98
4.2 无失真离散信源编码定理	102
4.2.1 无失真定长信源编码定理	103
4.2.2 无失真变长信源编码定理	107
4.3 霍夫曼码和其他编码方法	110
4.3.1 唯一可译码和即时码	110
4.3.2 香农码	112
4.3.3 霍夫曼码	114
4.3.4 费诺码	121
4.4 算术码	122
4.4.1 香农-费诺-埃利斯码	122
4.4.2 算术码	124
习题	127

第 5 章 波形信源和波形信道	130
5.1 波形信源的统计特性和处理方法	130
5.1.1 波形信源的统计特性	130
5.1.2 随机波形信源的处理方法	131
5.2 连续信源的信息度量	132
5.2.1 连续信源的熵及其性质	132
5.2.2 连续信源的最大熵	135
5.2.3 熵功率	138
5.3 波形信源的信息度量	139
5.4 波形信道的分类和处理方法	140
5.4.1 波形信道的分类	140
5.4.2 波形信道的处理方法	142
5.5 连续信道的信道容量	143
5.5.1 连续信道的平均互信息	143
5.5.2 连续信道的信道容量	146
5.6 波形信道的信道容量	150
5.6.1 波形信道的信息传输速率	150
5.6.2 波形信道的信道容量	151
习题	154
第 6 章 限失真信源编码	156
6.1 失真测度	156
6.1.1 失真函数	157
6.1.2 平均失真	160
6.1.3 保真度准则	161
6.2 信息率失真函数	161
6.2.1 信息率失真函数的定义	163
6.2.2 $R(D)$ 函数的性质	163
6.2.3 $R(D)$ 函数的计算	166
6.3 限失真信源编码定理	169
习题	171
第 7 章 有噪信道编码	174
7.1 有噪信道编码定理	174
7.1.1 平均错误概率	174
7.1.2 信道编码的基本思想	178
7.1.3 有噪信道编码定理	180

7.1.4 信道编码的途径	182
7.2 联合信源信道编码定理	183
7.3 信道编码的基本原理	187
7.3.1 信道编码的一般方法	187
7.3.2 信道编码的检错和纠错能力	188
7.3.3 最小汉明距离译码	190
7.4 线性分组码	191
7.4.1 线性分组码的编码	192
7.4.2 线性分组码的译码	196
7.4.3 完备码和汉明码	198
7.4.4 循环码	198
7.5 卷积码	201
7.5.1 卷积码的解析表示	201
7.5.2 卷积码的图形描述	204
7.5.3 卷积码的译码方法	206
7.5.4 递归型系统卷积码	210
7.6 Turbo 码	211
7.6.1 Turbo 编码	211
7.6.2 Turbo 译码	212
习题	217
第 8 章 网络信息论初步	221
8.1 多用户信道	221
8.1.1 多址接入信道	222
8.1.2 广播信道	226
8.1.3 相关信源的多用户信道问题	227
8.2 无线信道	230
8.2.1 信道模型	230
8.2.2 圆对称复数高斯随机矢量	231
8.2.3 MIMO 高斯信道容量	232
附录 概率论基础知识概要	234
参考文献	235

第 1 章 绪 论

信息论是关于信息的理论,它的内涵和外延不断变化发展,因此有广义信息论和狭义信息论之分。在这一章中,讨论什么是信息并简介信息论的形成和发展,描述信息论研究的对象和内容。

1.1 信息的概念

1.1.1 信息

“信息(information)”一词大量充斥在网络、报刊、电视广播中,“信息时代”、“信息技术”、“信息产业”、“信息经济”、“信息资源”、“信息革命”、“信息报务”、“信息化”、“信息科学”、“信息处理”、“信息网络”、“信息安全”等有关信息的词层出不穷,不胜枚举,成为各种媒体上出现频率最高的词之一。信息伴随人类生活,并深刻影响了人们的日常生活,成为人类生活中不可或缺的部分。那么,什么是“信息”?

对于信息的含义,人们从不同的角度做出了多种描述:“信息就是谈论的事情、新闻和知识”(英国《牛津辞典》);“信息,就是在观察或研究过程中获得的数据、新闻和知识”(美国《韦氏字典》);“信息是所观察事物的知识”(日本《广辞苑》);“信息是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”(《辞海》1989年版)。可见,对于信息的含义众说纷纭。

在日常生活中,信息常常被认为就是“消息”、“情报”、“知识”、“情况”等。的确,信息与它们之间是有着密切联系的。但是,很明显我们日常生活中所说的信息的含义要更深刻、更广泛,它不能等同于消息、情报、知识和情况。例如,人们通过电视、电话、报刊等各种媒体,每时都在获取、加工、传递和利用着大量的信息。在行政工作中,看材料、学文件是获取信息,作决策、批文件是处理信息,作指示是传递信息。可见,信息来源于客观世界,范围广大,具有一定的利用价值,可以通过载体为人们所获知,用来指导人类认识世界、改造世界。

用文字、符号、数据、语言、音符、图片、图像等能够被人们感觉器官所感知的

形式,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(message)**。消息中包含信息,是信息的载体。得到消息,从而获得信息。同一则信息可用不同的消息形式来载荷,如球赛进展情况可用电视(图像)、广播(语言)、报纸(文字)等不同消息形式来表述。而一则消息也可载荷不同的信息,它可能包含非常丰富的信息,也可能只包含很少的信息。因此,信息与消息是既有区别又有联系的。

在各种实际通信系统中,往往为了克服时间或空间的限制而进行通信,必须对消息进行加工处理。把消息转换成适合信道传输的物理量,这种物理量称为**信号(signal)**,如电信号、光信号、声信号、生物信号等。信号携带着消息,它是消息的运载工具。信号携带信息,但不是信息本身。同样,同一信息可用不同的信号来表示。同一信号也可表示不同的信息。信息、消息和信号是既有区别又有联系的3个不同的概念。

近代控制论的创始人之一,美国科学家维纳(Norbert Wiener, 1894—1964) 1948年在《控制论——动物和机器中通信与控制问题》一书中论述:“信息就是信息,不是物质,也不是能量。”这句话听起来有点抽象,但指明了信息与物质和能量具有不同的属性。世界是由物质组成的,能量是一切物质运动的动力,信息是人类认识、了解自然和人类社会的凭据。信息、物质和能量,是人类社会赖以生存和发展的三大要素。

至此,我们简要阐述了信息的含义,但只是释义,还没有对“信息”进行准确定义。我们已经感觉到信息一词的含义模糊和难于捉摸,但人人都感觉到它的存在。每时每刻我们都在通过对周围世界的观察去获取它,并且通过一定的方法把它传送给别人、进行交换或把它存储起来留作以后使用。这种目前尚难明确定义的信息可以把它称为广义理解的信息,即**广义信息**。

对这种广义的信息有许多人研究它的定义。一般来说,广义信息是指与客观事物相联系,反映客观事物的运动状态,通过一定的物质载体被发出、传递和感受,对接收对象的思维产生影响并用来指导接收对象的行为的一种描述。其他定义有“信息是事物及其属性标识的集合”,试图从“属概念+种差”的界定来讨论信息定义的标准化模式;“信息是物质在相互作用中表征外部情况的一种普遍属性,它是一种物质系统的特性以一定形式在另一种物质系统中的再现”。但是,迄今为止,尚未形成被普遍接受的有关广义信息的定义。

尽管至今尚没有关于“信息”的确切定义,但随着生产力尤其是科学技术的发展,以及人们对“信息”理解的不断深入和对物质世界普遍属性理解的拓展,越来越多的人认同将信息与物质、能量一起,作为物质世界的三大支柱,支持“**信息是物质(事物)的一种普遍属性,是事物运动的状态、方式及其改变的反映**”的观点。

作为技术术语广泛使用的“信息”是指技术上可收集、识别、提取、转换、存

储、传递、处理、检索、检测、分析和利用的对象,它主要指信息的具体表达形式,虽然信息的形式总是与信息的内容有一定的联系,且不可能存在没有内容的形式,但作为技术术语的信息的确不考虑信息的内容。计算机所能处理的(特别是通信所能传送的)都是信息的载体或表达形式。计算机可把信息的一种形式转换成另一种形式,如把英语文本翻译成汉语文本,把数据库中的数据整理成所需形式的报表,或把气象数据进行处理后给出某一地区的气温等,而通信则把信息的具体载体或形式从甲地传送到乙地。因此,作为技术术语的“信息”实际上是指一切符号、记号、信号等表达信息所用的形式或载体,实际上把信息的形式或载体和它的具体内容区分开来。作为一个技术术语的信息,其意义当然要比前面广义信息的含义具体得多,但仍然是比较笼统的。

信息作为一个可以用严格的数学公式定义的科学名词,首先出现在统计学中,随后又出现在通信技术中。无论是在统计学中还是在通信技术中定义的信息都是一种统计意义上的信息,可以把它简称为统计信息。统计信息是一个抽象而明确的概念,它与内容无关,而且不随信息具体表达形式的变化(如把文字翻译成二进制码)而变化,因而也独立于形式。它反映了信息表达形式中统计方面的性质,是一个统计学上的抽象概念。其适用范围要比广义信息狭隘得多。我们在本书中讨论的信息论正是关于这种统计信息的理论。

1.1.2 香农“信息”的定义

香农(Claude Elwood Shannon, 1916—2001)在1948年发表了一篇著名的论文《通信的数学理论》。他从研究通信系统传输的实质出发,对信息作了科学的定义,并进行了定性和定量的描述。

各类通信系统——电报、电话、广播、电视、雷达、遥测等传送的是各种各样的消息。消息的形式可以不同,但它们都是能被传递的,能被人们感觉器官(眼、耳、触觉等)所感知的,而且消息表述的是客观物质和主观思维的运动状态或存在状态。香农将各种通信系统概括成如图1.1所示的框图。

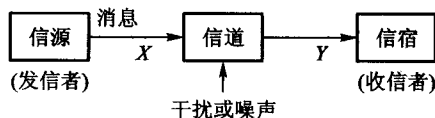


图 1.1 通信系统的简单模型

在各种通信系统中,其传输的形式是消息,但消息传递过程的一个最基本、最普通却又不十分引人注意的特点是:收信者在收到消息以前是不知道消息的具体内容的。在收到消息以前,收信者无法判断发信者将会发来描述何种事物

运动状态的具体消息;他更无法判断是描述这种状态还是那种状态。再者,即使收到消息,由于干扰的存在,他也不能断定所得到的消息是否正确和可靠。总之,收信者存在着“不知”、“不确定”或“疑问”。通过消息的传递,收信者知道了消息的具体内容,原先的“不知”、“不确定”和“疑问”消除或部分消除了。因此,对收信者来说,消息的传递过程是一个从不知到知的过程,或是从知之甚少到知之甚多的过程,或是从不确定到部分确定或全部确定的过程。如果不具备这样一个特点,那就根本不需要通信系统了。试想,如果收信者在收到电报或电话之前就已经知道报文或电话的内容,那还要电报、电话系统干什么呢?

由于主、客观事物运动状态或存在状态是千变万化的、不规则的、随机的,所以在通信以前,收信者存在“疑义”和“不知”。只要报文是清楚的,在传递过程中没有差错,那么,他收到报文以后,他原来所有的“不确定性”都没有了,他就获得了所有的信息。如果在传递过程中存在着干扰,使报文模糊不清,收信者收到报文以后,原先所具有的不确定性一点也没有减少,他就没有获得任何信息。如果干扰使报文发生部分差错,使收信者原先的不确定性减少了一些,但没有全部消除,他就获得了一部分信息。所以,通信过程是一种消除不确定性的过程。不确定性的消除就获得了信息。原先的不确定性消除得越多,获得的信息就越多,如果原先的不确定性全部消除了,就获得了全部的信息;若消除了部分不确定性,就获得了部分信息;若原先不确定性没有任何消除,就没有获得任何信息。由此可见,信息是事物运动状态或存在方式的不确定性的描述。这就是香农对信息的定义。

从以上分析可知,在通信系统中形式上传输的是消息,但实质上传输的是信息。消息只是表达信息的工具、载荷信息的客体。显然,在通信中被利用的(亦即携带信息的)实际客体是不重要的,而重要的是信息。信息较抽象,而消息是较具体的,但还不一定是物理性的。通信的结果是消除或部分消除不确定性从而获得信息。

1.1.3 “信息”的度量

根据香农有关信息的定义,信息如何测度呢?当人们收到一封电报或听了广播或看了电视,到底得到多少信息量呢?显然,信息量与不确定性消除的程度有关。消除多少不确定性,就得到多少信息量。那么,不确定性的大小能度量吗?

用数学的语言来讲,不确定就是随机性,具有不确定性的事件就是随机事件。因此,可运用研究随机事件的数学工具——概率论和随机过程来测度不确定性的大小。若从直接概念来讲,不确定性的大小可以直观地看成是事先猜测某随机事件是否发生的难易程度。

例如,假设有甲、乙两个布袋,各袋内装有大小均匀,对人手感觉完全一样的

球 100 个。甲袋内有红球、白球各 50 个,乙袋内有红、白、蓝、黑 4 种球各 25 个。现随意从甲袋或乙袋中取出一球,并猜测取出的是什么颜色的球,这一事件当然具有不确定性。显然,从甲袋中摸出红球要比从乙袋中摸出红球容易得多。这是因为,在甲袋中只在“红”与“白”两种颜色中选择一种,而且“红”与“白”机会均等,即摸取的概率各为 $1/2$ 。但在乙袋中,红球只占 $1/4$,摸出红球的可能性就小。自然,“从甲袋中摸出的是红球”比“从乙袋中摸出的是红球”的不确定性要小。从这个例子得出,不确定性的大小与可能发生的消息数目及各消息发生的概率有关。

再如气象预报,我们知道可能出现的气象状态有许多种。以十月份北京地区天气为例,经常出现的是“晴间多云”、“晴”或“多云”,其次是“多云转阴”、“阴”、“阴有小雨”等,而“小雪”这种天气状况出现的概率是极小的,“大雪”的可能性则更小。因此,在听气象预报前,我们大体上能猜测出天气的状况。由于出现“晴间多云”、“晴”或“多云”的可能性大,我们就可能确定这些天气状况的出现。因此,当预报明天白天“晴间多云”或“晴”时,我们并不觉得稀奇,因为和我们猜测的基本一致,所消除的不确定性较小,获得的信息量不大。而出现“小雪”的概率很小,我们很难猜测它是否会出现,所以这个事件的不确定性很大。如果预报是“阴有小雪”,我们就会大吃一惊,感到气候反常,这时就获得了很大的信息量。出现“大雪”的概率更小,几乎是不可能出现的现象,它的不确定性更大。如果一旦出现“大雪”的气象预报,我们将万分惊讶,这时将获得更大的信息量。由此可知,某一事物状态出现的概率越小,其不确定性越大;反之,某一事物状态出现的概率接近于 1,即预料中肯定会出现的事件,那它的不确定性就接近于零。

这两个例子告诉我们:某一事物状态的不确定性的该大小与该事物可能出现的不同状态的数目以及各状态出现的概率大小有关。既然不确定性的该大小能够度量,可见,信息是可以测度的。我们把某事物各种可能出现的不同状态,即所有可能选择的消息的集合,称为样本空间。每个可能选择的该消息是该样本空间的一个元素。对于离散消息的集合,概率测度就是对每一个可能选择的该消息指定一个该概率(该概率是非负的,且所有消息的该概率和为 1)。一个样本空间和它的该概率测度称为一个该概率空间。

一般该概率空间用 $[X, P(x)]$ 来表示。在离散情况下, X 的样本空间可写成 $\{a_1, a_2, \dots, a_q\}$ 。样本空间中该选择任一元素的该概率表示为 $P_X(a_i)$,其脚标 X 表示所考虑的该概率空间是 X 。在不会引起混淆的情况下,脚标可以略去,写成 $P(a_i)$ 。所以在离散情况下,该概率空间为

$$\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \cdots & a_q \\ P(a_1) & P(a_2) & \cdots & P(a_q) \end{bmatrix}$$

通常也记为

$$\begin{cases} X: a_1, a_2, \dots, a_q \\ P(x): P(a_1), P(a_2), \dots, P(a_q) \end{cases}$$

其中 $P(a_i)$ 就是选择符号 a_i 作为消息的概率, 称为先验概率。在接收端, 对是否选择消息(符号) a_i 的不确定性与 a_i 的先验概率成反比, 即对 a_i 的不确定性可表示为先验概率 $P(a_i)$ 的倒数的某一函数。我们取该函数为对数函数, 并把这样定义的不确定性称为该消息(符号) a_i 的自信息量, 即

$$I(a_i) \stackrel{\text{def}}{=} \log \frac{1}{P(a_i)} \quad (1.1)$$

由于信道中存在干扰, 假设接收端收到的消息(符号)为 b_j , 它可能与 a_i 相同, 也可能与 a_i 有差异。我们把条件概率 $P(a_i|b_j)$ 称为后验概率, 它是接收端收到消息(符号) b_j 后而发送端发送的是 a_i 的概率。那么, 接收端收到 b_j 后, 发送端发送的符号是否是 a_i 尚存在的不确定性, 应是后验概率的函数, 即是 $\log \frac{1}{P(a_i|b_j)}$ 。于是, 收信者在收到消息(符号) b_j 后, 已经消除的不确定性为: 先验的不确定性减去尚存在的不确定性, 这就是收信者获得的信息量, 定义为互信息, 即

$$I(a_i; b_j) = \log \frac{1}{P(a_i)} - \log \frac{1}{P(a_i|b_j)} \quad (1.2)$$

如果信道没有干扰, 信道的统计特性使 a_i 以概率 1 传送到接收端。这时, 收信者接到消息后, 尚存在的不确定性就等于零, 即 $P(a_i|b_j) = 1, \log \frac{1}{P(a_i|b_j)} = 0$, 不确定性全部消除。由此得互信息

$$I(a_i; b_j) = I(a_i) \quad (1.3)$$

以上就是香农关于信息的定义和度量。

1. 香农信息定义的优点

香农定义的信息概念在现有的各种理解中, 是比较深刻的, 它有许多优点:

- 它是一个科学的定义, 有明确的数学模型和定量计算。
- 它与日常用语中的信息的含义是一致的。例如, 设某一事件 a_i 发生的概率等于 1, 即 a_i 是预料中一定会发生的必然事件, 如果事件 a_i 果然发生了, 收信者将不会得到任何信息(日常含义), 因为他早知道 a_i 必定发生, 不存在任何不确定性。

根据公式(1.1), 因为 $P(a_i) = 1$, 所以得

$$I(a_i) = \log \frac{1}{P(a_i)} = 0$$

即自信息等于零。反之, 如果 a_i 发生概率很小, 即猜测它是否发生的不确定性很大, 一旦 a_i 果然发生了, 收信者就会觉得很意外和惊讶, 获得的信息量很大。

根据公式(1.1),因为 $P(a_i) \ll 1$,故得

$$I(a_i) = \log \frac{1}{P(a_i)} \gg 1$$

● 香农的信息定义排除了对信息一词某些主观上的含义。根据上述定义,同样一个消息对任何一个收信者来说,所得到的信息量(互信息)都是一样的。因此,信息的概念是纯粹形式化的概念。

2. 香农信息的局限性

香农定义的信息有其局限性,存在一些缺陷:

● 香农信息定义的出发点是假定事物状态可以用一个以经典集合论为基础的概率模型来描述。然而实际存在的某些事物运动状态,要寻找一个合适的概率模型往往是非常困难的。对某些情况来讲,是否存在这样一种模型还值得探讨。

● 香农信息的定义和度量没有考虑收信者的主观特性和主观意义,也撇开了信息的具体含义、具体用途、重要程度和引起后果等因素。这就与实际情况不完全一致。例如,当得到同一消息后,对不同的收信者来说常常会引起不同的感情、不同的关心程度、不同的价值,这些都应认为是获得了不同的信息。又例如,甲乙两人同去听一段音乐,若甲缺乏欣赏音乐的基本知识和必要训练,这种信息就不能发生什么作用。若乙是一位训练有素的音乐家,那么他将从这段音乐中获得大量信息。因此,信息有很强的主观性和实用性。

由此可见,香农信息的定义和度量是科学的,是能反映信息的某些本质的;但同时却是有缺陷的、有局限的,因此,它的适用范围会受到严重的限制。但由于人们对信息的本质认识还不够充分,所以,国际上尚未形成一个普遍公认的、完整的、确切的定义。为此,有关信息的定义和其测度的研究还在不断地深入。我们深信,随着人们对信息这一概念的不断深入研究,将会得出更合理、更确切的信息的定义和测度,彻底揭示信息的本质,全面和准确地把握信息。

1.2 信息论的形成和发展

人类的社会生活是不能离开信息的,人类的社会实践活动不仅需要对外界世界的情况有所了解并能做出正确的反应,而且还要与周围的人群进行沟通才能协调行动。这就是说,人类不仅时刻需要从自然界获得信息,而且人与人之间也需要进行通信,交流信息。人类需要随时获取、传递、加工、利用信息,否则就不能生存。

人们获得信息的方式有两种:一种是直接的,即通过自己的感觉器官,耳闻、目睹、鼻嗅、口尝、体触等直接了解外界情况;一种是间接的,即通过语言、文字、

信号……传递消息而获得信息。

通信是人与人之间交流信息的手段,语言是人类通信的最简单要素的基础。人类早期只是用语言和手势直接进行通信,交流信息。“仓颉造字”则使信息传递摆脱了直接形式,同时扩大了信息的存储形式,可算是一次信息技术的革命。印刷术的发明,扩大了信息的传播范围和容量,也是一次重大的信息技术变革。但真正的信息革命则是电报、电话、电视等现代通信技术的创造与发明,它们大大加快了信息的传播速度,增大了信息传播的容量。正是现代通信技术的发展导致了信息论的诞生。

1838年莫尔斯(F. B. Morse)发明电报使信息获得了电气的表现形式。1876年贝尔(A. G. Bell)发明了电话,使人类语言第一次获得电信号的形式。1895—1896年英国的马可尼(G. Marconi)和俄国的波波夫(A. C. Попов)发明无线电通信,使电报和电话可以通过电波加以传送。20世纪初(1907),德福雷斯特(L. De Forest)发明了能把电磁波放大的电子管。之后,很快出现了远距离无线电通信系统。大功率超高频电子管发明以后,电视系统就建立起来了(1925—1927)。接着,在20世纪30年代末和40年代初的第二次世界大战初期,微波通信系统、微波雷达系统等随着微波电子管的发明迅速发展起来。60年代初发明的激光技术,使人类进入了光纤通信的时代。因此,通信系统的最初发展可以概括为如何获得信息的电气表现形式以及如何将它们进行远距离传输。

进入20世纪后电信技术获得快速发展,如何提高信道利用率的问题开始提上日程。1917年坎培尔(G. A. Campbell)申请了第一个关于滤波器的专利,为频分复用信道提供了条件。1922年卡松(J. R. Carson)分析了振幅调制信号,开始明确上下边带的概念,提出边带理论,指明信号在调制(编码)与传送过程中与频谱宽度的关系。1924年奈奎斯特(H. Nyquist)开始分析电报信号传输中脉冲速率与信道带宽的关系。这一结果稍后又在其1928年的论文中得到发展,建立了限带信号的采样定理。哈特莱(R. V. L. Hartley)在1928年发表的论文《信息的传输》中首先提出消息是代码、符号而不是信息内容本身,使信息与消息区分开来,并提出用消息可能数目的对数来度量消息中所含有的信息量,为信息论的创立提供了思路。他研究接收机在估计接收脉冲幅度时只能分辨有限数目的脉冲幅度,假设这一数目是 M ,则 N 个脉冲所可能组成的不同序列的总数是 M^N ,哈莱特就把信息量 H 定义为 $H = N \log M$ 。这样,通过信道传输的信息量就与信道带宽和传输总时间的积成正比。从上面这些进展可以看出,在20世纪30年代以前通信的主要目标还集中在如何使发送信号并无失真地送到接收端,所用的分析方法还是分析确定性信号的方法。所以虽然哈特莱定义了信息量,但这还不是一个统计的概念,因此其意义还是相当有限的。

20世纪30年代,由于通信技术水平的提高以及随后第二次世界大战的爆发,

使通信中的噪声和抗干扰问题逐渐突出。1930年维纳(N. Wiener)开始把傅里叶(J. B. J. Fourier)分析方法全面引入到随机信号的研究中来,1936年兰登(V. D. Landon)发表了他第一篇有关噪声的论文。与此同时抗干扰的通信方法先后出现,1936年阿姆斯特朗(E. H. Armstrong)提出了频率调制,1939年达德莱(H. Dudley)发明了声码器,1939年瑞弗(H. Reeve)提出了具有强抗干扰能力的脉冲编码调制。这些对噪声的研究到1945年时由莱斯(S. O. Rice)作了全面的总结。

20世纪40年代初期,维纳在研究防空火炮的控制问题时,发表了“平稳时间序列的外推、内插与平滑及其工程应用”的论文。他把随机过程和数理统计的观点引入到通信和控制系统中来,揭示了信息传输和处理过程的统计本质。他还利用自己提出的“广义谐波分析理论”对信息系统中的随机过程进行了谱分析。这使得通信系统的理论研究上了一个新台阶。

1948年,香农在贝尔系统技术杂志上发表了《通信的数学理论》的文章。他利用概率测度和数理统计的方法系统地讨论了通信的基本问题,得出了几个重要而带有普遍意义的结论,并由此奠定了现代信息论的基础。

所以20世纪40年代中通信的理论已经全面走上统计分析的道路,抗干扰已经取代抗失真成为通信研究中的中心问题。在这样的背景下香农和维纳几乎同时提出了信息的统计定义,分别在香农的《通信的数学理论》和维纳的《控制论》进行了阐述,后来被公认为信息论的经典著作。但后者讨论的范围更广,它更主要的是控制论的经典著作。《通信的数学理论》一文,成为信息论诞生的标志。

在1948年以后的十余年中,香农对信息论的发展作出了巨大的贡献。在1973年出版的信息论经典论文集中,香农是49篇(总数)论文中12篇论文的作者。迄今为止,信息论的主要概念除通用编码外几乎都是香农首先提出的。除一系列基本的概念外,香农的贡献还在于证明了一系列编码定理,这些定理不但给出了某些性能的理论极限,而且实际上也是对香农所给基本概念的重大价值的证明。由于香农这一系列的贡献,他被认为是信息论的创始人。

香农信息理论的核心是:揭示了在通信系统中采用适当的编码后能够实现高效率和-high可靠性的传输信息,并得出了信源编码定理和信道编码定理。从数学观点看,这些定理是最优编码的存在定理;但从工程观点看,这些定理不是结构性的,不能从定理的结果直接得出实现最优编码的具体途径。然而,它们给出了编码的性能极限,在理论上阐明了通信系统中各种因素的相互关系,为人们寻找最佳通信系统提供了重要的理论依据。

在信息论的发展中,还有许多科学家做出了卓越的贡献。法国物理学家布里渊(L. Brillouin)1956年出版的专著《科学与信息理论》,从热力学和生命等许多方面探讨信息论,把热力学熵与信息熵直接联系起来,使热力学中争论了一个世纪之