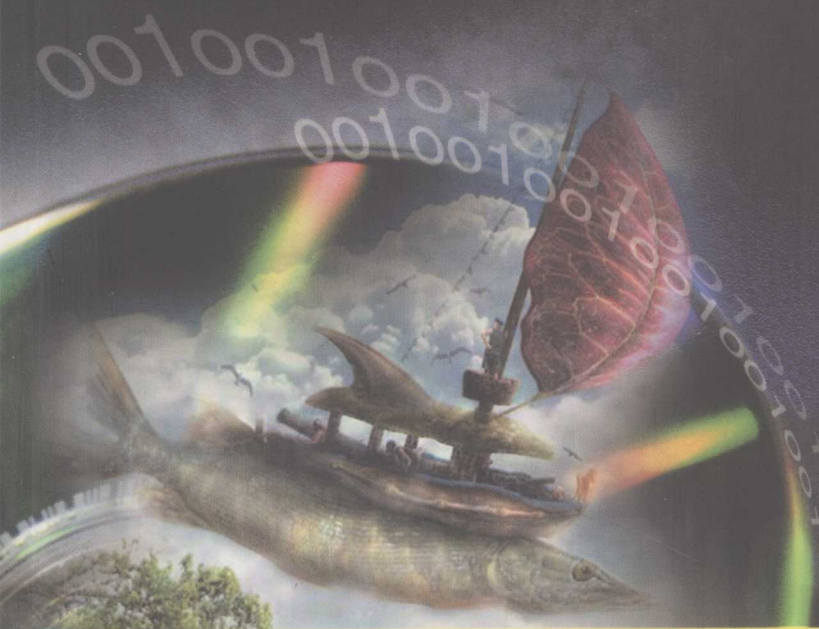


福建省高校非计算机专业教材编写委员会 组织编写
主 编 \ 鄂大伟



大学 信息技术 基础

福建省高校非计算机专业教材编写委员会

组织编写

大学 信息技术 基础

主 编：鄂大伟
副主编：宁正元
俞建家
陈锻生

TP3

387

ASIMO

HONDA

28

图书在版编目 (CIP) 数据

大学信息技术基础/鄂大伟主编. —厦门: 厦门大学出版社, 2005. 6

ISBN 7-5615-2351-1

I. 大… II. 鄂… III. 电子计算机—高等学校—教材 IV. TP3

中国版本图书馆CIP数据核字 (2005) 第064656号

厦门大学出版社出版发行

(地址: 厦门大学 邮编: 361005)

<http://www.xmupress.com>

xmup@public.xm.fj.cn

福建省沙县方圆印刷有限公司印刷

2005年7月第1版 2005年7月第1次印刷

开本: 787×1092 1/16 印张: 17

字数: 432千字 印数: 00001~17000

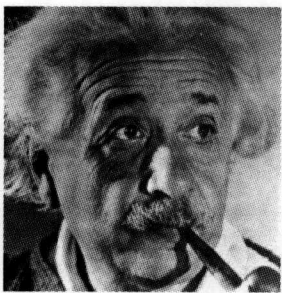
定价: 23.80元

本书如有印装质量问题请直接寄承印厂调换

致 科 学

想像力比知识更为重要。重要的是不要停止问问题。

——爱因斯坦——



阿尔伯特·爱因斯坦

科学,是一个散发着崇高人性光辉的字眼。科学是一个漫长而渐进的历程,科学的金字塔由许多一生为追求科学真理的科学家和科学工作者所构筑。科学技术是推动人类社会前进的巨大动力,每个民族和国家的发展,都离不开科学技术的进步。科学技术的发展和创新,决定着人类文明的进程,决定着民族和国家的盛衰兴亡。

科学的本质在于创新。无论是科学规律的发现,还是技术的发明,都离不开创新。科学认识是求知、求真、求是的过程,贯穿着批判思维和创新思维的交互作用,体现着科学工作者永不衰竭的创新意识和进取精神。科学的创造常常以扎实的专业知识为基础,以艰苦卓绝的精神劳动为途径,以敏锐的观察能力、丰富的想象能力、深刻的洞察能力、见微知著的直觉能力和灵感为向导,以构建关于自然界的本质和规律的理论体系为目标。

整个科学史可以说就是以天才的名字来点缀的灿烂银河,而有几颗特别明亮的星辰,它们所发射的光芒穿越了整个宇宙。在信息科学和计算机科学与技术的发展史上,也涌现了许多天才般的人物和先驱者,他们以自己的孜孜不倦的追求和探索精神,浇灌和培育了信息科学之花。人类在走向信息文明时,每一步前进与超越,都闪耀着他们睿智的思想光辉。他们的精神令人钦佩,他们的经历给人以启迪。在信息化浪潮铺面而来的今天,有志于学习或从事信息科学与技术的人们,在信息科学殿堂的入口处,应以勤奋的学习,艰苦的探索,开启信息科学的大门,并准备为科学的进步而奉献一生。

今年,全球以接力传递光信号形式纪念伟大而孤独的科学家爱因斯坦逝世 50 周年,同时纪念相对论发表 100 周年(1905—2005)。值此之际,请允许我引用牛顿的一句名言:有些人一生在伟大真理海洋的沙滩上拾集晶莹的卵石。他们日复一日地注视着那虽然缓慢,但却确定无疑地上涨的气势磅礴的海潮,这股海潮的胸怀包藏着无数能把人类生活装点得更高尚美好的珍宝。

让我们向本书提及的以及未提及的那些为人类文明进步和科学发展做出贡献的科学家和科学工作者们致敬!

作者

2005 年 6 月

前言

——对大学信息技术教育的思考

自文明肇始以来,人类就生活在信息的海洋中。人类社会的生存和发展,时刻都离不开接收信息、传递信息、存储信息和利用信息。原始人的“结绳记事”,古人的“烽火报警”无疑是信息的表示、存储和传送的方法之一。如果说文字的出现、印刷术的产生,是人类文明发展进程中的两个重要里程碑,那么,以计算机技术为代表的信息技术的发展和普及,将成为人类文明发展史上的第三个里程碑。

以计算机技术、网络与通信技术和微电子技术为代表的现代信息技术,正在改变人们传统的生活、学习和工作方式,同时也影响着教育的内容与方法。信息技术教育成为全世界课程改革研究的热点研究课题。作为前言,这里我们想谈谈编写本书的动机与目的。

信息技术教育是指学习、利用信息技术,培养信息素养,促进学与教优化的理论与实践。信息技术教育的本质是利用信息技术培养信息素养。信息素养是指人所具有的对信息进行获取、识别、加工、利用、评价和管理的知识、能力等各方面基本品质的总和。

为了迎接 21 世纪的挑战,国家教育部已决定在中小学基本普及信息技术教育,将信息技术课纳入中小学必修课程,把学生培养成有信息素质的终身学习者。这是我国面向 21 世纪国际竞争、提高全民素质、培养具有信息素质和创新素质的新型人才的一项重要举措。随着中小学信息技术教育的实施,中小學生将掌握计算机和网络的基本知识技能,学生的信息素养将会大幅度的提高。目前,经过基本信息技术教育的高中毕业生将陆续升入大学,他们对计算机基本操作技能的掌握程度会越来越好,这是不争的事实。

从学生受教育的系统性和整体性来看,大学的信息技术教育应充分考虑学生在中小学时期所打下的信息技术的基础。在规划大学信息技术教育的课程时,应首先了解中小学信息技术课程的开设情况,并结合学生对信息技术教育内容的掌握程度,有的放矢地进行教学,以避免教育的重复、资源的浪费。使那些在中小学接受了基本信息技术教育的学生,在步入大学殿堂后,能够受到更高层次的、与培养目标相适应的信息技术教育。

从目前的情况看,由于各地方教育发展的不平衡性,似乎还没有一个统一的解决办法。有些高校基本上假定学生的信息技术基础为“零起点”,其教学内容仍定位在计算机文化基础教学的层次上,教学要求单纯地体现为对计算机操作技能的培养上。从教学内容来看,与中小学的信息技术课程的内容大多是重复的。如果大学的计算机信息教育课程仍以传授计算机操作技能为主,则将偏离大学信息技术教育的目标,也将使在小学、初中、高中分阶段开设信息技术课程的作法变得毫无意义。

在由原先大学作为计算机教育的起点开始过渡到以中小学作为普及信息技术教育的起点的背景下,大学信息技术教育的内容是什么;如何体现大学信息技术教育的特点;如何与中小学衔接;如何紧跟迅速发展的信息技术,构建适合我国国情的大学信息技术课程与教材体系;如何进行信息技术教育与其他学科教育的整合以及课堂教学和实践环节的整合;这些都是高

校计算机基础教学课程改革需要认真思考的问题。

当然,对以上问题的认识,人们有着不同的观点。我们以为:大学信息技术教育其实质是对信息技术教育根本目的的认识,对学习目标的_{理解},对课程性质的定位。信息技术教育的根本目的是培养创新人才,不是培养机器的操作者。从更深层的意义上讲,信息技术教育是一种素质教育,它不是以某种技能掌握为目标的技能培训。不能以计算机基础知识与操作技能的学习与掌握来替代信息技术的学习与掌握,只能将计算机作为信息获取、分析、处理的工具进行学习。对计算机技术的学习应从信息技术的高度、信息技术的角度进行学习,而不是脱离这个前提进行学习。

基于以上认识,我们在2001年就开始进行大学信息技术教学改革的尝试。当时的初衷是准备给大学生开一门类似于“信息技术基础”的课程。可是寻遍书店也未找到合用的教材。因为目前关于大学计算机文化基础的教材比比皆是,而真正介绍大学信息技术的教材甚鲜,即使有些挂以“信息技术”之名的大学教材仍未脱离“计算机文化基础”的窠臼。这才使我们萌生了编写一本全面介绍大学信息技术的相关概念与知识,内容丰富,可读性强的教材,通过教材与内容的改革来贯彻我们对大学信息技术教学改革的思想与理念。现在,作为《大学信息技术》教学研究项目的重要成果,本教材终于与读者见面了,同时,我们的教学改革项目获得了所在学校的教学成果特等奖,2005年福建省高校教育教学成果二等奖,我们感到十分欣慰。

本书冠以《大学信息技术基础》之名,意即大而学之,信息先导;论述基础,注重原理;阐明要义,把握整体。本书内容组织参照国家教育部的新大纲编写,从信息科学与计算机技术的最新成果中汲取知识,力求与相关学科相互融合,使课程内容超越原“计算机文化基础”的局限,不涉及所述对象使用或操作的繁文缛节,最大限度地回归和体现大学“信息技术”的“整体”面目,成为培养大学生信息能力、信息素养和创新意识的载体。这正是我们所期待的。

在本书的编写过程中,我们力求能在结构上和_{内容上}有所创新,并有较鲜明的特色。首先是取材新颖,许多内容取自于国内外最新的资料和Internet资源,以反应信息技术当前的发展;其次,为使学生了解信息技术的全貌,开阔视野,对在信息技术发展历史上出现的重要事件、人物或概念以提示方式介绍,旨在增加知识性;再次,全书包含的信息量较大,深浅程度不一,目的是让具有不同需求的读者都有收益。在教学中可根据教学对象的专业背景和需求对内容加以调整,使“深者得其深,浅者得其浅”。

为克服传统书本媒体的先天局限性,充分发挥网络教学的优势。我们准备提供教师教学和学生自学使用的网络教学平台和网络版课程(已得到福建省高等学校网络课程立项课题的资助,正在建设中)。网络课程的设计是对教材内容的进一步的补充和说明,可以使学习者在任何时间,任何地点进行学习。学习者还可以通过网络教材提供的网络链接查阅到众多相关的学习资源,实现现实世界与课堂教学之间的交互功能。

在信息化社会的今天,信息素养已成为科学素养的重要构成部分。迅速地筛选和获取信息、准确地鉴别信息、创造性地加工和处理信息,将是所有社会成员应具备的、终生有用的基础能力之一。在本课程的学习中,我们希望能够营造一种信息素质教育的环境。学生应能关注学习过程,积极参与自主的学习活动,提出与自己学过的学科内容有关的问题、使用环境、资源、工具等进行学习,从而实现教学相长。“博学而笃志,切问而近思”,学则固矣。

参加本书编写的教师有(按编写章节顺序)集美大学鄂大伟教授、福建农林大学宁正元教授、福州大学俞建家教授、华侨大学陈锻生教授。全书由鄂大伟负责策划和统稿。

在“大学信息技术基础”的教改实践中和在本书的编写过程中,福建省教育厅高教处、福建

省高校计算机水平考试指导委员会的领导和专家们对我们的教学改革与探索给予了肯定和大力支持,并予以专门立项资助。本项目同时还得到福建省高等学校网络课程立项课题的资助,集美大学的余元辉老师为网络课程平台的建设付出了辛勤的、创造性的劳动。厦门大学出版社为教材的出版也给予了大力帮助,在此一并深致谢忱。恭疏短语,难竭鄙诚。

信息技术的发展一日千里,相关的每个学科都变得越来越精微和深奥。囿于作者的水平及篇幅所限,俾本书内容难以准确反映和把握信息的整体与全貌,疏漏、欠妥、悖谬之处,恳请读者指正。冀收博见,嘉惠来学。

作者

2005年6月于厦门集美学村

(74)	主 要 参 考 文 献	2.2.3
(74)	器 材 容 内 容 提 要	2.2.4
目 录	器 材 容 内 容 提 要	2.2.5
(70)	香 农 对 信 息 的 定 义	2.2.6
(72)	香 农 信 息 论 的 局 限 性	2.2.7
(72)	数 据 、 消 息 、 信 号 与 信 息 的 区 别	2.2.8
前 言	数 据 、 消 息 、 信 号 与 信 息 的 区 别	2.2.9
第 1 章 信息与计算科学		(1)
(72) 1.1 探索信息的真谛		(2)
(72) 1.1.1 什么是信息		(2)
(72) 1.1.2 从信息论到信息科学		(3)
(74) 1.1.3 香农对信息的定义		(4)
(72) 1.1.4 信息的度量		(5)
(72) 1.1.5 香农信息论的局限性		(7)
(72) 1.1.6 数据、消息、信号与信息区别		(8)
(70) 1.2 信息技术与信息科学		(9)
(70) 1.2.1 信息技术的发展与定义		(9)
(72) 1.2.2 信息技术的核心		(10)
(72) 1.2.3 信息科学		(13)
(72) 1.3 计算与计算科学		(14)
(74) 1.3.1 探索计算之源		(14)
(74) 1.3.2 计算模型与图灵机		(15)
(72) 1.3.3 计算的困惑——如何认识计算科学		(17)
(72) 1.3.4 计算机科学的研究领域		(18)
(72) 1.4 计算机的信息表示与编码		(20)
(70) 1.4.1 信息在计算机中的表示		(20)
(72) 1.4.2 信息的编码		(22)
(72) 1.4.3 数制及其转换		(24)
(74) 1.5 计算机的逻辑运算与逻辑门电路		(26)
第 2 章 计算机系统:硬件与软件		(31)
(72) 2.1 从历史走向未来——计算机的发展史		(32)
(72) 2.1.1 现代计算机的“史前”时代(—1946)		(32)
(70) 2.1.2 冯·诺依曼型计算机的基本结构		(34)
(70) 2.1.3 第一台现代电子数字计算机的诞生		(35)
(72) 2.1.4 现代计算机发展的四个阶段		(36)
(72) 2.1.5 无“心”的机器——计算机有智能吗?		(40)
(72) 2.1.6 巨型计算机与矢量计算		(42)
(72) 2.2 计算机硬件系统		(43)
(70) 2.2.1 微型计算机		(43)
(72) 2.2.2 微处理器——给你一颗奔腾的“芯”		(44)

2.2.3 主板	(47)
2.2.4 内存储器	(48)
2.2.5 外存储器	(49)
2.2.6 输入设备	(50)
2.2.7 输出设备	(53)
2.2.8 微型计算机的总线及标准	(55)
2.2.9 计算机与外部设备的接口及标准	(55)
(1) 2.3 计算机软件系统	(56)
(3) 2.3.1 软件的性质	(57)
(3) 2.3.2 软件技术的进化史	(58)
(3) 2.3.3 软件系统的分层结构	(63)
(4) 2.3.4 操作系统	(64)
(2) 2.3.5 应用软件	(75)
第3章 多媒体技术基础	(79)
(8) 3.1 多媒体的概念	(80)
(9) 3.1.1 媒体的分类	(80)
(9) 3.1.2 多媒体与多媒体技术	(80)
(10) 3.1.3 多媒体计算机系统	(81)
(13) 3.2 光盘存储系统	(82)
(11) 3.2.1 光盘及其特点	(82)
(11) 3.2.2 光盘的类型	(84)
(12) 3.2.3 光盘的标准	(84)
(12) 3.2.4 光盘系统记录与读取信息的原理	(87)
(18) 3.3 多媒体音频信号处理	(88)
(20) 3.3.1 音频信号的形式	(88)
(20) 3.3.2 音频的数字化过程	(90)
(22) 3.3.3 数字音频的文件格式	(92)
(24) 3.4 多媒体图像信息处理	(93)
(25) 3.4.1 计算机图像处理的概念	(94)
(31) 3.4.2 图像的数字化过程	(95)
(32) 3.4.3 图像的压缩与编码	(97)
(32) 3.4.4 矢量图与位图	(98)
(34) 3.4.5 图像文件格式	(100)
(35) 3.4.6 计算机图形处理与图像处理的区别与联系	(101)
(36) 3.5 多媒体视频信息处理	(102)
(40) 3.5.1 视频的定义	(102)
(42) 3.5.2 视频的分类	(103)
(43) 3.5.3 电视信号制式	(103)
(44) 3.5.4 YUV 与 RGB 彩色模型	(104)
(44) 3.5.5 视频的数字化过程	(105)

3.5.6 视频信号的压缩与编码	(106)
第4章 程序数据的数据、组织和管理	(109)
4.1 程序设计与算法	(110)
4.1.1 算法与程序	(110)
4.1.2 程序设计的语言	(116)
4.1.3 程序设计的方法	(122)
4.1.4 程序的调试与纠错	(131)
4.2 程序数据的组织与结构	(133)
4.2.1 数据与数据类型	(133)
4.2.2 简单数据结构的应用	(137)
4.2.3 较复杂的数据结构的组织	(143)
4.2.4 数据结构在程序设计语言中的实现	(144)
第5章 数据库技术基础	(147)
5.1 关系数据模型	(148)
5.1.1 数据模型	(148)
5.1.2 什么是关系模型	(152)
5.1.3 关系数据库的基本概念	(154)
5.1.4 关系数据库的标准语言——SQL	(157)
5.2 数据库应用系统设计	(167)
5.2.1 数据库设计内容	(167)
5.2.2 需求分析	(169)
5.2.3 建立 E-R 模型	(172)
5.2.4 建立关系模型	(173)
5.2.5 关系范式	(174)
5.2.6 数据库应用系统的体系结构及开发工具	(175)
5.3 数据库技术的新发展	(177)
5.3.1 数据仓库、数据挖掘和联机分析处理技术	(177)
5.3.2 面向对象数据库	(178)
5.3.3 基于 XML 的数据库	(178)
第6章 计算机网络与应用	(183)
6.1 计算机网络概述	(184)
6.1.1 什么是计算机网络	(184)
6.1.2 网络的形成与发展	(184)
6.1.3 网络的功能	(185)
6.1.4 网络的基本组成	(185)
6.1.5 计算机网络的分类	(186)
6.2 数据通信基础	(186)
6.2.1 通信系统模型	(186)
6.2.2 数据通信方式	(187)
6.2.3 数据交换技术	(189)

6.2.4 基带传输与宽带传输	(190)
6.2.5 网络传输比特速率与带宽	(191)
6.2.6 网络传输介质	(192)
6.2.7 网络通信协议	(196)
6.3 计算机局域网	(198)
6.3.1 局域网的基本概念	(198)
6.3.2 局域网的拓扑结构	(198)
6.3.3 局域网的工作模式	(200)
6.3.4 局域网的硬件组成	(201)
6.3.5 局域网的软件	(202)
6.3.6 局域网互联技术	(204)
6.4 Internet 基础	(209)
6.4.1 Internet 的发展	(209)
6.4.2 Internet 的 IP 地址与域名	(211)
6.4.3 Internet 提供的信息服务	(215)
6.5 Internet 的接入方式	(217)
6.5.1 电话拨号接入	(217)
6.5.2 宽带接入方式	(218)
6.6 Internet 的信息检索与发布	(219)
6.6.1 Internet 信息检索	(219)
6.6.2 网页制作与工具	(222)
6.6.3 网站的建设与管理	(223)
第 7 章 信息系统安全与社会责任	(229)
7.1 信息系统的安全构架	(230)
7.1.1 信息安全的含义	(230)
7.1.2 信息安全的保障体系	(231)
7.1.3 信息安全的主要目标	(232)
7.1.4 信息安全研究的四个层面	(232)
7.2 数据备份	(233)
7.2.1 数据备份的基本概念	(233)
7.2.2 如何制定和规划备份策略	(234)
7.2.3 备份策略的实施时的注意事项	(234)
7.2.4 完善的备份软件应该具备的功能	(235)
7.3 计算机病毒与防治	(236)
7.3.1 什么是计算机病毒	(236)
7.3.2 计算机病毒的历史及产生原因	(236)
7.3.3 计算机病毒的传播途径	(237)
7.3.4 计算机病毒的特征	(238)
7.3.5 病毒的工作原理	(239)
7.3.6 网络病毒	(240)

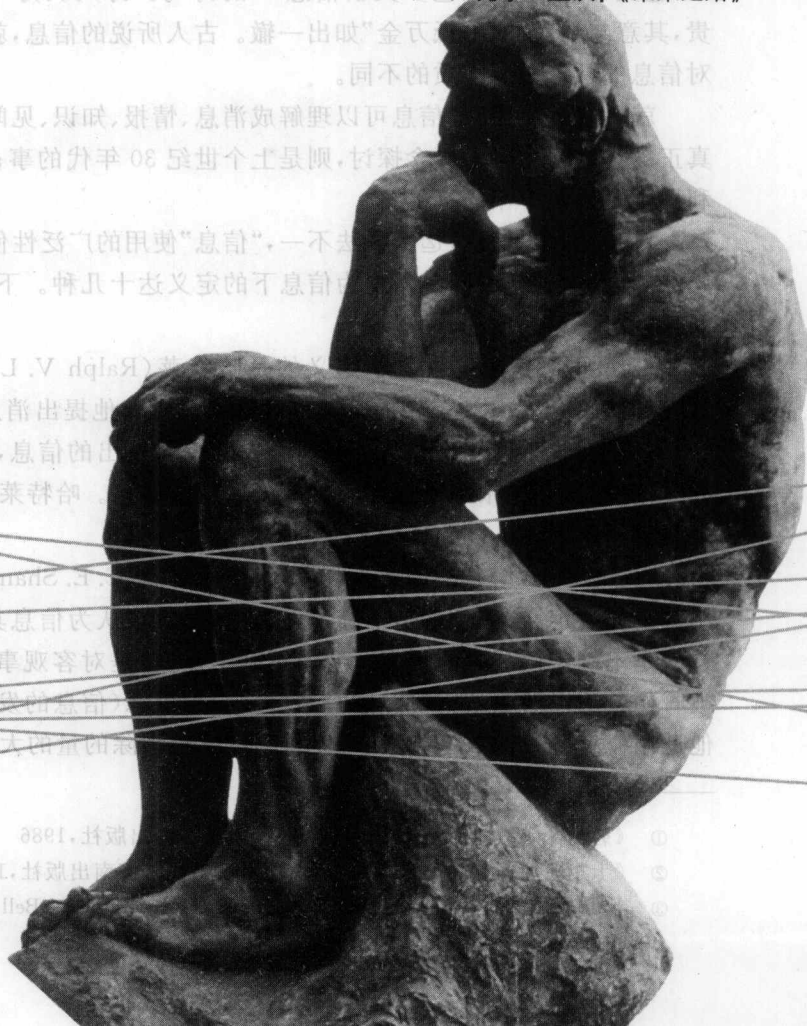
7.3.7 计算机病毒的预防措施	(242)
7.3.8 常见杀毒软件	(243)
7.4 防火墙	(243)
7.4.1 什么是防火墙	(244)
7.4.2 防火墙的功能	(244)
7.4.3 防火墙的基本准则	(245)
7.4.4 防火墙的种类	(245)
7.4.5 防火墙的部署	(247)
7.4.6 部署个人防火墙	(247)
7.5 数据加密技术	(248)
7.5.1 加密与加密系统	(248)
7.5.2 密码体制	(249)
7.5.3 数字签名	(251)
7.6 信息政策与法规	(252)
7.6.1 信息时代的法律	(252)
7.6.2 计算机犯罪的定义与分类	(253)
7.6.3 计算机犯罪的立法	(254)
7.6.4 信息安全教育	(255)
7.6.5 信息道德	(256)

第 1 章

信息与计算机科学

当我第一次听说“信息时代”这个儿时，就感到心痒难熬。之后，我读到有关学术界预言各国将为控制信息，而不是控制资源而战。这听起来挺玄乎，但他们听说的信息究竟是什么意思呢？

——比尔·盖茨：《未来之路》



信息、物质和能量是现代社会主要依赖的三种资源。信息作为一种客观存在,从远古直到当今的文明社会,一直都在积极发挥着人类意识到或没意识到的重大作用。信息犹如空气一样普遍存在于人类社会时空之中。也许正是因为我们整天都淹没在信息的海洋中,我们对信息并没有给予太多的关注。什么是信息?它的实质是什么?它有什么特征?它怎样度量?对这些问题的透彻理解,是收集、处理和利用信息的前提,让我们就从这里开始探索信息资源的宝库,迈向信息科学的大门。

1.1 探索信息的真谛

1.1.1 什么是信息

信息的本质究竟是什么?人类始终在不断追问自己。今天,人类已经跨入信息时代。对于信息的本质,我们应当做出什么样的诠释呢?

信息一词来源于拉丁文“Information”,且在英文、法语、德语、西班牙语中同字,以及俄语、南斯拉夫语中同音,表明了它在世界范围内使用的广泛性。我国古人很早就已经知道信息的重要性,认为信息是可以寻觅、获取的。信息一词也多见于我国古代的诗词中,唐朝诗人李中在诗中就留下了“梦断美人沉信息,目穿长路倚楼台”的佳句^①。宋代女词人李清照也赋有“不乞随珠与和璧,只乞乡关新信息”^②的诗句。诗人认为“乡关新信息”比随珠、和璧更为宝贵,其意与杜甫“家书抵万金”如出一辙。古人所说的信息,就是指“音信”、“消息”,与今天人们对信息的认识没有本质的不同。

就一般意义而言,信息可以理解成消息、情报、知识、见闻、通知、报告、事实、数据等等。但真正被作为一个科学概念探讨,则是上个世纪30年代的事;而被作为科学为人们普遍认识和利用则是近几十年的事情。

对于什么叫信息,迄今说法不一,“信息”使用的广泛性使得我们难以给它下一个确切的定义。专家、学者从不同的角度为信息下的定义达十几种。下面所叙述的几种定义是人们从不同角度对信息的理解:

1. 最早对信息进行科学定义的是哈特莱(Ralph V. L. Hartley)。他在1928年发表的《信息传输》^③一文中,首先提出“信息”这一概念。他提出消息是代码、符号而不是信息内容本身,使信息与消息区分开来。他认为,发信者所发出的信息,就是他在通信符号表中选择符号的具体方式,并主张用所选择的自由度来度量信息。哈特莱的思想和研究成果,为信息论的创立奠定了基础。

2. 1948年,信息论创始人,美国科学家香农(C. E. Shannon)从研究通信理论出发,第一次用数学方法定义“信息就是不确定性的消除量”。认为信息具有使不确定性减少的能力,信息量就是不确定性减少的程度。所谓不确定性,就是对客观事物的不了解、不肯定。因此,信息被看作是用以消除信宿(信息的接收者)对于信源(信息的发出者)发出哪些消息的不确定性。他还用概率统计的数学方法,来度量不定性被消除的量的大小。

① 《碧云集·暮春怀故人》——全唐诗,上海古籍出版社,1986

② 《上枢密韩肖胄诗》,徐北文《李清照全集译注》,济南出版社,1998

③ Ralph V. L. Hartley: “Transmission of Information”, in Bell System Tech. Journal, vol. 7, 1928

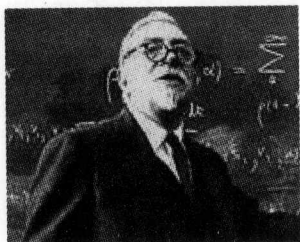


图 1-1 N. Wiener

3. 控制论创始人之一,美国科学家维纳(N. Wiener)(图 1-1)在 1948 年发表的名著《控制论——动物和机器中的通讯与控制问题》^①一书中曾经指出“信息就是信息,不是物质,也不是能量。”维纳在《控制论》一书的导言中曾明确地指出:“必须发展一个关于信息量的统计理论,在这个理论中,单位信息就是对二择一的事物作单一选择时所传递出去的信息”。后来,维纳在《人有人的用处——控制论与社会》^②一书中写到:“信息是在人们适应外部世界,并且使这种适应反作用于外部世界的过程中,同外部世界进行互相交换的内容的名称。”“要有效地生活,就必须有足够的信息。”在这里,维纳把人们与外界环境交换信息的过程看成是一种广义的通信过程,试图从信息自身具有的内容属性给信息定义。这两本著作标志着控制论这门新兴学科的兴起。

4. 关于信息的定义,有人提出用变量来度量,认为“信息就是差异”。持这种观点的典型代表是意大利学者朗格(G. Longe)。他提出:“信息是反映事物的形式、关系和差别的东西。信息是包含于客体间的差别中,而不是在客体本身中。”按照这种观点,自然界和人类社会普遍存在着可传递的差异性。差异越大,信息量就越大,没有差异就没有信息,不可传递的东西也不是信息。所谓信息量就是对事物差异度的量度或测度。

5. 我国信息论学者钟义信教授认为:信息是“事物运动状态和方式,也就是事物内部结构和外部联系的状态和方式。”^③

6. 权威性工具书《辞源》对信息定义为:“信息就是收信者事先所不知道的报导。”^④《韦氏字典》(中国)对信息的描述是:“信息是指对消息接受者来说预先不知道的报导。”

对于信息的含义,至今仍是众说纷纭,莫衷一是,人们出于不同的研究目的,从不同的角度出发,对信息作用的不同理解和解释而对信息做出了定义。各种信息定义都反映了信息的某些特征。这样,难免就存在差异性和多样化。

随着时间的推移,时代将赋予信息以新的含义,这将是一个动态的概念。现代“信息”的概念,已经与半导体技术、微电子技术、计算机技术、通信技术、网络技术、多媒体技术、信息服务业、信息产业、信息经济、信息化社会、信息管理、信息论等含义紧密地联系在一起。但信息的本质是什么,这仍然是需要进一步探讨的问题。

1.1.2 从信息论到信息科学

20 世纪初以来,特别是 20 世纪 40 年代,通信技术的迅速发展,迫切需要解决一系列信息理论问题,例如如何从接收的信号中滤除各种噪声,怎样解决火炮自动控制系统跟踪目标问题等。这就促使科学家在各自研究领域对信息问题进行认真的研究,以便揭示通信过程的规律和重要概念的本质。

信息论作为一门严密的科学,主要应归功于美国数学家香农(C. E. Shannon, 1916—2001)

① N. Wiener.《控制论——动物和机器中的通讯与控制问题》. 科学出版社,1963

② N. Wiener.《人有人的用处》. 商务印书馆,1978

③ 钟义信.《信息科学原理》. 北京邮电大学出版社,1996

④ 《辞源》. 商务印书馆修订本,1996



图 1-2 Shannon

(图 1-2)。1948 年,香农在《贝尔系统技术杂志》上发表重要论文《通信的数学理论》^①,1949 年,香农又发表另一重要论文《在噪声中的通信》^②。在这两篇文章中,他解决了过去许多悬而未决的问题:经典地阐明了通信的基本问题,提出了通信系统的模型,给出了信息量的数学表达式,解决了信道容量、信源统计特性、信源编码、信道编码等有关精确地传送通信符号的基本技术问题。两篇文章成了现在信息论的奠基著作。而香农也一鸣惊人,成了这门新兴学科的奠基人。由于香农提出的信息论是关于通信技术的理论,它是以数学方法研究通信技术中关于信息的传输和变换规律的一门科学。所以,人们又将其称为狭义信息论,或经典信息论。

信息论发展的第二个阶段是一般信息论。这种信息论虽然主要还是研究通信问题,但是新增加了噪声理论、信号的滤波、检测、信号的编码与译码、信号的调制与解调,以及信息的处理等问题。通信的目的是要使接收者获得可靠的信息,以便做出正确的判断与决策。为此,一般信息论特别关心信号被噪声干扰时的处理问题。

信息论发展的第三个阶段是广义信息论。它是随着现代科学技术的纵横交叉的发展而逐渐形成的。一般地说,在对信息的研究中,仅考虑其形式的方面而不考虑其内容和用途,即是狭义信息。如果考虑信息的语义和有效性问题,则是广义信息。广义信息论,远远超出了通信技术的范围来研究信息问题,它以各种系统、各门科学中的信息为对象,广泛地研究信息的本质和特点,以及信息的获取、计量、传输、储存、处理、控制和利用的一般规律。广义信息论的研究与很多学科密切相关,例如,数学、物理学、控制论、计算机科学、逻辑学、心理学、语言学、生物学、仿生学、管理科学等。信息论在各个方面得到了广泛的应用。主要研究以计算机处理为中心的信息处理的基本理论,包括语言、文字的处理、图像识别、学习理论及其各种应用。

显然,广义信息论包括了狭义信息论和一般信息论的内容,但其研究范围却比通信领域广泛得多,从而拓宽了信息论的研究方向,使得人类对信息现象的认识与揭示不断丰富和完善。是狭义信息论和一般信息论在各个领域的应用和推广,因此,它的规律也更一般化,适用于各个领域。

1.1.3 香农对信息的定义

什么是信息?信息能否度量?怎样度量?这些是信息论所必须回答的问题。

香农在他发表的著名论文《通信的数学理论》中,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。

香农认为:信息是有秩序的量度,是人们对事物了解的不确定性的消除或减少。信息是对组织程度的一种测度,信息能使物质系统有序性增强,减少破坏、混乱和噪音。

香农提出:信息的传播过程是“信源”(信息的发送者)把要提供的信息经过“信道”传递给“信宿”(信息的接收者),信宿接收这些经过“译码”(即解释符号)的信息符号的过程。并由此建立了通信系统模型。

① 《通信的数学理论》(A mathematical Theory of Communication) Bell System Technical Journal, 1948 年 6 月和 10 月,连载

② 《在噪声中的通信》, Bell System Technical Journal, 1949 年

什么是信道呢？信道是在物理线路上划分的逻辑通道。由于物理上的限制，信道都只有有限的带宽，而且存在噪声，因此信道能够传递的最大数据速率是受信道带宽制约的。对于这个问题，奈奎斯特(H. Nyquist, 1889—1976)和香农先后展开了研究。由此，香农推出了受噪声(所谓噪声是指“外加于信号之上，而非属信息源本身的信号”)干扰的信道情况下传输速率与信噪比(信号功率与噪声功率之比)之间的关系，指出了用降低传输速率来换取高保真通信的可能性。该公式已广泛用于有噪声情况下的信道最大传输速率的计算(见公式 6-2)。

在香农确定信息量名称时，将热力学中的“熵”的概念应用到信息领域。一个系统的熵就是它的无组织程度的度量。而一个系统中的信息量是它的组织化程度的度量，这说明信息与熵恰好是一个相反的量，信息是负熵，所以在信息熵的公式中有负号。它表示系统获得后无序状态的减少或消除，即消除不定性的多少。

由于熵表达了事物所含的信息量，我们不可能用少于熵的比特数来确切表达这一事物。所以这一概念已成为所有无损压缩的标准和极限。同时，它也是导出无损压缩算法做到或接近“熵”的编码的源泉。



什么是熵

“熵”的概念起源于热力学，是度量分子不规则热运动的单位，即“不确定性”在热力学里用“熵”来度量。熵表示系统的无组织或混乱程度，熵愈大意味着该系统愈混乱无序，熵愈小表明该系统的组织程度愈高。香农的伟大贡献在于，将热力学中的“熵”的概念应用到信息领域，并利用概率分布的理论给出信息量——熵的概念，即信息是以它对事物不确定性的减少或消除来度量。

1.1.4 信息的度量

根据香农的有关信息的定义，信息源所发出的消息带有不确定性。用数学的语言来讲，不确定就是随机性。那么信息如何测度呢？显然，信息量与不确定性消除程度有关。消除多少不确定性，就获得多少信息量。不确定性的多少可以直观地看成是事先猜测某随机事件是否发生的可能程度。



著名的科学家史蒂芬·霍金在他所著的《时间简史》序言中风趣地写道：“有人告诉我，放在书中的每一个方程都会使本书的销售量减半，吓跑一半我的潜在读者。”

这里要介绍的内容当然要比《时间简史》容易理解的多，即使这样，为了说明信息是如何度量的，这里不得不给出两个公式。希望不要吓跑全部的读者。

由于信息源发出的消息是随机的，可以用随机变量来表示。如果用符号表示信息源发出的信息，假设事件的基本空间包含个元素，即，且每一个可能值的概率为：

$$P(X=x_i)=p \quad i=1,2,\dots,m$$

那么定义一个随机事件 x 所含的信息量称为 x 的自信息量，即：

$$I(x)=\log \frac{1}{p(x)}=-\log p(x) \quad (1-1)$$