

高等学校教材

大学信息技术导论

管会生 主编

管会生 张得太 任学惠 编

高等教育出版社

内 容 提 要

本书是依据教育部高等教育司非计算机专业计算机教学指导分委员会制订的高等院校计算机基础课程教学基本要求编写的。本书内容沿信息技术的历史、技术、应用三条主线展开,共包括信息技术基础、程序设计和应用三篇内容,重点介绍信息理论的发展历程与奠基人、中国特色的信息化道路、汉字识别技术、算法设计与基本数据结构、结构化程序设计、面向对象技术与可视化编程、软件方法学与软件工程学、Visual Basic 程序设计语言、中国“金”系列工程、信息技术与现代军事、生物计算机与生物信息学、计算机美术与计算机音乐、世界著名 IT 企业、校园卡打造数字化校园、智能化社区——数字家园、企业信息化与 ERP、软件工程与 CMM、密码技术与信息隐藏术等。通过大知识量地讲授,阐述这样一种理念:计算机的生命在于应用!

本书可作为高等学校计算机基础教育的教材,也可以作为培训教材和自学参考书。

中文在线出品

www.ChineseAll.com

序 言

——计算机的生命在于应用

我是 1956(猴)年生人,很幸运与计算机有缘,因为猴年也是中国计算机事业的本命年。1956年,是共和国计算机事业的开创起步年。这一年国家制定《1956—1967年科学技术发展远景规划》(简称12年科技规划)时,即将“计算技术的建立”列为四大紧急措施之首,开始筹建中国科学院计算技术研究所,并任命华罗庚教授为中国科学院计算技术研究所数字委员会主任。两年后,中国有了第一台电子管计算机——103机。1956年对于世界计算机界来说则更是具有里程碑意义的一年。首先,IBM公司创立了世界上第一个高级程序设计语言——FORTRAN语言,开创了计算机软件的新纪元,计算机从此真正走出了实验室,进入了应用普及大发展的新时代;其次,世界上第一台数字式电子计算机ENIAC这一年被送进了博物馆,1946年2月14日正式启用的ENIAC,仅运行了10年,人们就已经像看待出土文物一样对待它了,可见计算机这一新生事物发展速度是何等惊人;还有,这一年日本科学家刚刚造出了日本的第一台电子管计算机,这一年IBM公司制造出了世界上第一个磁盘存储系统,这一年计算机界首次提出了“人工智能(AI)”的概念,这一年推出了采用晶体管的商用计算机UNIVAC……

1995年尼葛洛庞帝(美)所写的《数字化生存》一书中强调:“计算不再只和计算机有关,它决定我们的生存。”的确如此,信息已是现代文明的血液,而信息化也已成为现代文明社会必不可少的内容。半个世纪以来,计算机像一个神奇的精灵,以人类历史上其他任何一种学科都无法与之比拟的发展速度改变着我们的生活方式和思维方式。计算机作为一种工具、一种手段正无孔不入地进入人们生产和生活的各个方面,并将成为一种知识、一种技能进入人们的智力结构,像读书、写字一样成为现代文明社会不可缺少的组成部分。同自然界一样,适者生存依然是信息社会必须遵循的法则。但古人云:“人之生也,与忧患俱来”。所以说,惟有惶者才能生存!天道酬勤,惟有强者才能生存!

面对计算机发展如此之快的技术和竞争如此激烈的市场,我们始终十分忙碌地工作着。的确太忙了,以至于顾不上为放弃曾经学过的一个个编程语言和操作软件而痛苦;也顾不上为丢弃曾经花费了大量人民币买来的一台台计算机设备而惋惜。我们已不能够说什么东西是“最新的”或“全新的”,也很难说什么东西是“最好的”或“最佳的”。结果是,我们一贯遵守的“循序渐进”的学习常规被彻底打破了。计算机技术的发展形态几十年来一直是“现在进行时”,正感慨“我怎么又落后了”的同时,新东西又来了。我们认了,这就是我们仅有的回答。

IT(信息技术)业的发展突飞猛进,但IT业作为朝阳产业的确还很年轻。企业的稚嫩,从业人员的成熟,加上高速成长的庞大市场和原本浮躁的媒体有意无意的炒作,这些因素共同把IT圈的浮躁推向了必然。回想起来,用“狂热”一词形容上世纪末.COM风潮似乎并

不为过。各种各样的“e道”高论曾使人眼花缭乱、目不暇接、无所适从。一些人为了商业目的，“恶炒”计算机热，什么来新闻效果就炒什么，其“炒”法与炒名人隐私和影星恋爱史的炒作心态是完全一样的。这些“喜新厌旧”的朋友，对于每一个新版的操作系统或编程工具都会热炒一番，全然不顾到底有多少中国人和中国的應用环境能使用和适应这些软件，忘却了面对如此广泛的计算机应用领域是根本不存在包打天下的编程语言的。更有许多人为能够用“猫”在因特网上浏览文字信息而欣喜若狂，欢呼“我们已经上了信息高速公路”，视而不见自己在提取图像信息（还不敢是动态的）和声音信息（还不敢是高保真的）时那难以忍耐的等待，忘却了在因特网上浏览时为大量的重复信息和无用信息（网络垃圾）所付出的宝贵时间。

不顾现实的跟风是IT业的另一通病。“解决方案”本来是从用户需求和系统角度出发，为用户提供全面服务的，不料却有了“解决方案是个筐，什么都可以往里装”的尴尬，就连扫描仪、打印机等也要打着解决方案的旗号，更有甚者干脆把多多少少相关的产品拼凑一下就可以贴上“XX解决方案”的标签。计算机界真的是到了该清醒的时候了。

在我们身边常常可以听到这样的话题：“我买个奔几合适啊？”；“快来看看，怎么又‘死’了？”；“这计算机怎么这么神，能放电影还能得病”。诚如斯言，虽然有趣，但听起来让人悲哀。人们可能忘记了，计算机的“神力”原本就是人赋予的，正如计算机界一句名言所说：“总有一些事情人可以比机器做得更好，更有一些事情只有人才可以做到，而且永远都是如此”。“9.11”美国遭袭击事件说明尽管美国是高科技武器及远程精确制导导弹拥有量最多的国家，高科技、高技术也并非是包打天下的利器，并非是争霸世界的惟一工具。所以说，正确地加以引导、普及计算机教育已是每一个计算机工作者不容推卸的责任，尽管对于一个有13亿人口的国家这是一个过于沉重的责任。

正是出于这种责任，在阅读收集和整理了信息技术领域大量资料的基础上，编者汇精聚萃，根据学习使用信息技术的特点和编者多年来的体会，将信息科学与信息技术领域研究的主流方向及主要应用进行了系统总结和概括，并以通俗易懂、简明扼要的方式介绍给读者。

本书是依据教育部对大学计算机基础课程的教学基本要求而编写的。全书沿信息技术的发展历史、技术、应用三条主线，共包括信息技术基础、程序设计和应用系统三篇内容，重点介绍信息理论的发展历程与奠基人、中国特色的信息化道路、汉字识别技术、算法设计与基本数据结构、结构化程序设计、面向对象技术与可视化编程、软件方法与软件工程学、Visual Basic 程序设计语言、中国“金”系列工程、信息技术与现代军事、生物计算机与生物信息学、计算机美术与计算机音乐、世界著名IT企业、校园卡打造数字化校园、智能化社区——数字家园、企业信息化与ERP、软件工程与CMM、密码技术与信息隐藏术等。同时也介绍信息资源与信息社会、多媒体与网络技术基础、无线接入技术、电子商务、电子政务、中国的计算机考试、数据仓库与数据挖掘、现代物流企业与物流中心建设、信息主权与知识产权管理、办公信息系统和移动办公、人工智能与神经网络、软件项目管理、客户关系管理（CRM）的理论和应用等前沿性内容，作为大学信息技术基础课程教学改革的尝试，以飨广大读者。

我们面临着“中国的计算机教育何时不再为Microsoft作义务操作培训和免费广告？”

“非零起点的大学本科计算机基础教育应如何改革?”等问题,而解决问题的方法也只有一个——“转变教育观念,创新教学理念”。观念的创新是最大的创新,观念的落后是最大的落后。

本书通过大知识量的讲授阐述这样一种理念:计算机的生命在于应用!我们不能再把对操作能力的评价作为对计算学科了解认知程度的尺度,就像我们不能把懂家电的人就认为他已懂得了电子学科一样。“兵无常法、药无常味、水无常形”。增兵减灶可打胜仗,减兵增灶也未必不赢。要汲取英语教学战线冗长(从小学到研究生要学19年)、投入时间和精力不少而效果甚微的教训,突出应用,讲求实效,这就是我们追求的目标。

本书遵照大学计算机应用基础课程教学改革“精细、深入、实用、简捷”的宗旨以及“简单是美、广度优先、授人以渔”的编写原则,率先在国内全面更新了传统的《计算机应用基础》课程以应用软件操作使用为主的教学体系结构,而以介绍信息系统和信息技术在各行各业应用中的新思想和新观念取而代之。在具体授课中,可以将第二部分(程序设计篇)作为要求掌握的基本内容重点讲授,其余两个部分(信息技术基础篇与应用篇)建议通过知识讲座的形式以扩大学生的知识面,要求学生了解信息技术的发展及应用状况就可以了。计算机教材或许是永远让编写者后悔的书,因为它无法时尚,更不可能新潮起来。但正因为如此,这种刺激和诱惑、挑战激励着我们一次又一次拿起手中的笔。

集25年的教学经验和体会,我所倡导的“计算机的生命在于应用!”、“素质不等于教育,教育不等于能力,能力不等于学历,学历不等于人才”以及“厚基础,宽专业;勤动手,强能力;重素质,善创新”等信息技术基础教育的新观念已得到业内同行的认可和响应。本书内容面广新颖、取材丰富实用、阐述深入浅出、结构合理清晰。我们建议:信息技术应用基础作为高校各本科专业必修课,可以统一教学内容,统一学时学分,统一组织考试。考虑到大学信息技术应用基础课程多学科的教学内容、多层次的教學对象、多媒体的表现形式以及多方面解决问题的需求,建议开课院校与我们一起来搭建立体化的教材体系,采用“教材+电子教案+试题库+多媒体网络课件+资料库+学科网站”的教学模式,授课、上机可同期集中安排。

本书由管会生主编,由张得太(5.7节和第6章)、任学惠(第4、第8章)和管会生(其余章节)三位老师共同编写完成,全书由管会生进行初审、修改和统稿。在本书编写过程中曾得到国内许多计算机公共课授课教师的大力支持,在此表示衷心感谢。由于时间仓促加上作者水平有限,书中粗浅疏漏或叙述欠严密之处在所难免,恳请读者给予批评指正。

管会生

2004年2月20日于兰州

目 录

第一部分 信息技术基础篇

第 1 章 信息资源与信息技术的发展	(3)	性能指标	(35)
1.1 信息资源与信息科学	(3)	2.2 计算机系统的基本硬件结构	(37)
1.1.1 信息及其主要特征	(3)	2.2.1 硬件的概念及基本结构	(37)
1.1.2 信息科学与信息技术	(5)	2.2.2 微处理器与并行处理系统	(39)
1.1.3 信息科学理论发展的百年 历史	(6)	2.2.3 存储器与磁盘阵列	(41)
1.2 信息理论的奠基人与 ACM 图 灵奖	(7)	2.2.4 输入设备与输出设备	(45)
1.2.1 信息论之父——香农	(7)	2.3 计算机操作系统	(48)
1.2.2 图灵与图灵机	(8)	2.3.1 操作系统的分类与并行系统 软件	(48)
1.2.3 冯·诺伊曼与冯氏结构计算机	(10)	2.3.2 Windows、Linux 和 UNIX	(49)
1.2.4 计算机界的诺贝尔奖——ACM 图 灵奖	(11)	2.4 计算机中数据的表示及转换	(52)
1.3 计算工具的演变和计算机的发展 ..	(13)	2.4.1 N 进制数的基数与权	(53)
1.3.1 古算史上的“哥德巴赫猜想”	(13)	2.4.2 四种进制数之间的相互转换 ..	(54)
1.3.2 信息技术及计算工具发展之大 事记	(14)	2.4.3 字符的二进制编码	(56)
1.3.3 ENIAC 的诞生与新一代 计算机	(17)	第 3 章 多媒体技术	(58)
1.3.4 生物计算机	(19)	3.1 多媒体技术的基本知识	(58)
1.3.5 信息电器与机器动物	(21)	3.1.1 多媒体技术的基本概念	(58)
1.4 中国信息技术的发展历程	(23)	3.1.2 信息压缩与信息压缩的国际 标准	(62)
1.4.1 中国计算机的发展初期	(23)	3.1.3 CD-ROM 与 DVD	(64)
1.4.2 近年来信息技术及软件业的发展 状况	(24)	3.2 多媒体技术的应用	(66)
1.4.3 有中国特色的信息化道路	(27)	3.2.1 多媒体技术在 Windows 中的 应用	(66)
第 2 章 计算机系统	(31)	3.2.2 音频与 MIDI	(68)
2.1 计算机的分类与微型计算机的 发展	(31)	第 4 章 计算机网络	(71)
2.1.1 计算机的分类与特点	(31)	4.1 计算机网络基础	(71)
2.1.2 微型计算机的发展	(33)	4.1.1 计算机网络概述	(71)
2.1.3 微型计算机的系统构成及主要		4.1.2 计算机网络的构成	(72)
		4.2 数据通信基础	(74)
		4.2.1 数据通信的基本知识	(74)
		4.2.2 数据通信方式	(75)
		4.2.3 数据传输技术	(76)
		4.2.4 多路复用技术	(77)
		4.2.5 数据交换技术	(78)

4.3 网络通信基础	(80)	4.5.3 城域网	(91)
4.3.1 网络传输介质和设备	(80)	4.5.4 光传输网络	(93)
4.3.2 OSI 参考模型网络及网络 协议	(81)	4.6 Internet	(95)
4.4 局域网	(83)	4.6.1 Internet 的基本工作原理	(95)
4.4.1 局域网的特点及构成	(83)	4.6.2 中国的 Internet 主干网络	(98)
4.4.2 局域网的拓扑结构	(84)	4.6.3 Internet 的接入方式	(99)
4.5 广域网与城域网	(85)	4.6.4 Internet 的接入技术	(101)
4.5.1 广域网的特征	(86)	4.6.5 无线接入技术的应用	(105)
4.5.2 广域网的公共传输平台	(86)	4.6.6 Internet 的应用	(108)
		第一部分 综合练习	(113)

第二部分 程序设计篇

第 5 章 程序设计基础	(121)	5.7.1 数据库管理系统的基本 概念	(166)
5.1 基本概念与程序设计语言的 发展	(121)	5.7.2 关系模型	(167)
5.1.1 基本概念	(121)	5.7.3 数据库设计	(168)
5.1.2 程序设计语言的发展	(127)	5.7.4 数据库语言 SQL	(169)
5.2 程序设计语言的基本成分与 使用	(134)	第 6 章 Visual Basic	(171)
5.2.1 数据及其运算	(134)	6.1 VB 的集成开发环境	(171)
5.2.2 函数与过程	(135)	6.1.1 Visual Basic 概述	(171)
5.2.3 程序设计语言的功能与 使用	(136)	6.1.2 VB 的集成开发 环境	(172)
5.2.4 程序运行的控制与环境	(137)	6.1.3 简单程序设计	(174)
5.3 算法设计与基本数据结构	(138)	6.2 对象的建立、编辑及属性、事件 与方法	(176)
5.3.1 问题描述与算法的概念	(139)	6.2.1 对象的建立和编辑	(176)
5.3.2 算法的表示方法	(141)	6.2.2 对象的属性、事件和方法	(177)
5.3.3 算法设计的基本方法	(143)	6.3 VB 程序设计语言 的语法	(182)
5.3.4 基本数据结构	(145)	6.3.1 数据类型	(182)
5.4 结构化程序设计	(153)	6.3.2 VB 中的常量与 变量	(184)
5.4.1 基本控制结构	(153)	6.3.3 基本语句	(187)
5.4.2 结构化分析与设计方法	(154)	6.3.4 VB 的运算符和表 达式	(192)
5.5 面向对象技术与可视化编程	(156)	6.3.5 VB 的控制结构 语句	(194)
5.5.1 面向对象程序设计的基本 概念	(157)	6.3.6 常用内部函数	(198)
5.5.2 可视化编程	(160)	6.3.7 数组	(200)
5.6 软件方法学与软件工程	(161)	6.3.8 自定义过程	(203)
5.6.1 软件方法学的分类和基本 内容	(161)	6.4 常用控件	(207)
5.6.2 软件工程	(163)		
5.7 数据库管理系统	(166)		

6.4.1 文本框和标签框	(207)	6.6.1 设计菜单	(219)
6.4.2 命令按钮、选择框、单选钮和 框架	(208)	6.6.2 建立弹出式菜单	(221)
6.4.3 列表框、组合框和滚动条	(210)	6.6.3 对话框	(222)
6.4.4 图片框、图像框和计时器	(213)	6.7 绘制图形	(226)
6.4.5 文件系统控件	(214)	6.7.1 Line 方法	(226)
6.5 多重窗体及多文档界面	(216)	6.7.2 Circle 方法	(227)
6.5.1 多重窗体	(216)	6.8 常用的文件操作语句与函数	(228)
6.5.2 多文档界面的设计	(218)	6.8.1 顺序文件	(228)
6.6 菜单与对话框	(219)	6.8.2 随机文件	(232)
		第二部分 综合练习	(235)

第三部分 应用篇

第7章 数字化生存	(243)	8.1.4 现代物流与电子商务	(270)
7.1 信息环境与信息社会	(243)	8.2 电子政务	(273)
7.1.1 因特网的发展与信息社会的 形成	(243)	8.2.1 电子政务概述	(273)
7.1.2 信息产业及信息技术的发展 趋势	(245)	8.2.2 电子政务的基本内容	(276)
7.1.3 中国“金”系列工程	(245)	8.2.3 电子政务的基础构成	(277)
7.2 信息技术与现代军事	(248)	第9章 信息技术与信息系统	(280)
7.2.1 科学技术是军事的第一战 斗力	(249)	9.1 办公信息系统	(280)
7.2.2 数字化部队与数字化装备	(250)	9.1.1 办公信息系统的基本概念、 基本内容和功能	(280)
7.2.3 信息技术在军事上的主要 功能	(251)	9.1.2 移动办公的展望与中国信 息港建设	(281)
7.3 计算机美术与计算机音乐	(253)	9.1.3 办公信息系统中软件系统 的结构设计	(282)
7.3.1 计算机美术	(253)	9.2 汉字信息处理	(284)
7.3.2 计算机音乐	(254)	9.2.1 汉字处理技术的发展概况	(284)
7.3.3 计算机音色转换	(256)	9.2.2 汉字的编码技术与字形 表示	(287)
7.4 机读身份证——中国信息化之 契机	(259)	9.2.3 汉字的输入和识别技术与 语音合成	(292)
7.4.1 机读身份证概况	(259)	9.3 软件工程与 CMM	(296)
7.4.2 机读身份证所引发的巨大 IT 应用市场	(260)	9.3.1 软件工程的发展与 CMM 模型	(297)
第8章 电子商务与电子政务	(262)	9.3.2 基于 CMM 的中国软件工程 改造	(299)
8.1 电子商务	(262)	9.3.3 软件项目管理	(300)
8.1.1 电子商务的概述	(262)	9.3.4 信息系统的开发	(302)
8.1.2 电子商务的功能及工作 模式	(264)	9.4 生物信息学与遗传算法	(304)
8.1.3 电子商务的技术支撑与运行 平台	(266)	9.4.1 生物信息学的概念与研究 动态	(304)

9.4.2 遗传算法的基本原理····· (307)	10.3.2 综合布线系统与家庭智能化····· (341)
9.4.3 遗传算法的基本步骤和应用····· (308)	10.4 数字校园····· (343)
9.5 人工智能与神经网络····· (310)	10.4.1 数字校园的概念与校园信息化建设····· (343)
9.5.1 “智能”与人工智能的概念····· (310)	10.4.2 大学资源计划····· (345)
9.5.2 人工智能的研究热点和实际应用····· (312)	10.4.3 校园卡打造数字化校园····· (346)
9.5.3 神经网络计算技术的发展····· (313)	10.4.4 计算机考试····· (352)
第10章 信息化与信息社会····· (316)	第11章 信息安全技术与信息社会的发展····· (356)
10.1 世界软件产业发展概况····· (316)	11.1 信息社会中的人文科学问题····· (356)
10.1.1 软件产业的崛起····· (316)	11.1.1 黑客犯罪与信息污染····· (356)
10.1.2 世界著名IT企业简介····· (318)	11.1.2 信息道德教育····· (358)
10.1.3 软件业市场前景与开发方向····· (323)	11.1.3 知识产权管理····· (359)
10.2 企业信息化和基于Web的ERP系统····· (326)	11.1.4 计算机神话——IT业批判····· (361)
10.2.1 信息时代的工业化····· (326)	11.1.5 文明的困惑——信息技术与伦理道德的“两难推理”····· (363)
10.2.2 企业信息化形成的新的管理模式及支持技术····· (328)	11.2 信息安全技术····· (368)
10.2.3 新型的企业管理模式和管理工具——ERP····· (329)	11.2.1 信息系统安全····· (368)
10.2.4 ERP项目实施与业务流程重组····· (330)	11.2.2 计算机安全技术····· (370)
10.2.5 ERP功能模块····· (332)	11.2.3 “9.11”事件与远程容灾系统····· (372)
10.2.6 J2EE平台与客户关系管理····· (333)	11.2.4 加密技术的基本概念····· (374)
10.2.7 数据挖掘与数据仓库——构筑企业竞争优势····· (335)	11.2.5 DES与RSA····· (377)
10.3 智能化社区——数字家园····· (339)	11.2.6 数字签名、数字时间戳和数字证书····· (379)
10.3.1 智能化社区的概念及发展状况····· (340)	11.2.7 生物特征识别技术的研究与应用····· (381)
	第三部分 综合练习····· (386)
	参考文献····· (388)

第一部分 信息技术基础篇

在漫长的人类社会发展的历史长河中,曾经历了四次信息技术革命:语言的形成,文字的创造,造纸术和印刷技术的发明,电报电话、广播电视的普及应用。以电子计算机的普及应用和计算机与现代通信技术有机结合而产生的因特网为主要特征的信息处理技术,则被人们称之为“第五次信息技术革命”。从1946年2月14日正式启用的第一台电子数字计算机ENIAC至今,谁也没有料到,时隔不到60年,这硕大无比的“怪物”竟经历了从主机、微机到网络三个时期的飞跃,并且开创了一个人类文明的新时代——信息时代。本篇将介绍以计算机技术、多媒体技术和网络通信技术为代表的信息技术的发展历程、计算机系统构成等内容。

第 1 章 信息资源与信息技术的发展

人们所说的“计算机(Computer)”并不仅仅是一台代替人工完成复杂计算的机器。确切地讲应该是“信息处理机”,它将人们听到的事实(Fact)和看到的景象(Phenomenon)等进入大脑的原始资料经过处理后变成有用的信息(Information)。信息同物质和能源一样,是人类社会赖以生存与发展的重要资源。信息技术是今天信息社会的“奠基石”,对当今社会经济的发展起着极大的促进作用。以信息技术为主导的信息时代在科技前沿也有其“三个代表”,那就是:以 Internet 为代表的网络技术,以纳米为代表的新型材料技术和以基因(基因工程、人类基因组图谱)为代表的生命科学和生物技术。

1.1 信息资源与信息科学

人类社会已经从以资源经济为主的农业社会和以资本经济为主的工业社会发展到了今天以知识经济为主的信息社会。信息社会是以信息活动为社会发展的基本活动,以信息技术为技术基础,以信息经济为主导经济,以信息产业为主导产业,以信息文化改变人类教育、生活和工作方式以及价值观念的新型社会形态。

也就是在 20 世纪 90 年代初,世界十大富翁的头衔还几乎被那些石油大王、船舶大王等所垄断,而 2000 年,Microsoft 总裁比尔·盖茨(Bill Gates)以 770 亿美元的资产赫然名列前茅。以计算机软件的生产和销售为其主业的 Microsoft 公司,从 1975 年以两千美元注册资金成立到今天已成为美国第一赢利公司,而具有 200 年发展历史的美国洛克菲勒公司才仅有 50 亿美元的资产。我们的邻邦印度,一个没有高速公路和地铁的国度,竟然演绎了一出现代无烟工业的奇迹,成为仅次于美国的软件大国,其软件产品出口 75 个国家,美国和日本是其最大的购买国,2000 年和 2001 年印度的软件出口额分别达到 63 亿和 79.9 亿美元,2003 年已超过 100 亿美元。

信息经济(或称知识经济)就是在充分知识化的社会中以信息智力资源的占有、投入和配置与知识产品的生产、分配(传播)和消费(使用)为最重要因素的经济。信息经济与工业社会的资本经济相比,除前者依赖于知识的程度高于后者以及知识在经济增长中的作用和价值大于后者外,最本质的不同是:信息和知识本身已成为知识经济中的一种最积极、最重要的投入要素。

1.1.1 信息及其主要特征

信息就像空气一样,虽然摸不到,但却不停地在我们身边流动,为人类服务。人们需要信息,研究信息,一时一刻也离不开信息。人类通过信息认识各种事物,借助信息的交流来沟通人与人之间的联系。

1. 信息的定义

信息研究者们对什么是信息科学的研究内容一直界定不清,它时而指计算机科学家所说的信息,时而指文献学家所说的信息,有时又代表通信科学家所说的信息。而真正的“信息科学”所关注的信息,应该不是指任何个别信息领域里的特定信息,它应该反映一切和信息有关领域里的信息现象,关心一切和信息有关领域里的信息问题。

著名的信息学家 Rafael Capurro 曾在他的著作中给出了一个关于信息的定义,他认为信息并不仅仅是交流的过程,也不是从一个主体转移到另一个主体的物质,信息是一种状态,人类的生产和生活以及其他一切活动都是在这种状态中进行的。的确如此,任何一个物种的存在,都依托于一个流通的信息环境,而当人类这个生灵在地球上出现的时候,信息就注定要以比其他物种更高的智能形态伴随着人类在艰难恶劣的自然环境中,从茹毛饮血的日子里一点一点走到今天。人类对信息的把握,成了他们认识世界和改造世界的有利工具。语言的产生,文字的出现,各种各样信息载体和媒介的不断拓展,使人们生活的信息状态越来越丰富和充盈。科学技术的发展,创造了一个又一个把想像变为现实的令人激动的时刻。人们惊异地发现,无处不在的信息不仅是我们离不开的生存环境,同时也是一种具有深厚挖掘潜力的资源。于是,对信息资源的开发和利用,就成了人们提升生活质量和工作效率的重要手段,在这个基础上,人们的想像力被牵引到了一个无穷无尽的空间。

迄今为止学术界尚没有对信息下一个准确完整的定义,对信息的解释也是众说纷纭,但可以大致归纳出信息的含义与特征:信息是用数据作为载体来描述和表示的客观现象;信息可以用数值、文字、声音、图像等多种形式表示;信息是对数据加工提炼的结果,是对人类有用的知识;信息是具有含义的符号或消息,而数据是计算机内信息的载体。显而易见,若想得到信息,必须要把客观世界中的现象和问题通过数据这种媒体记载下来。通常把对各种数据经过加工转换而得到信息的过程称为信息处理(或称数据处理)。

但是,信息和数据还是有区别的。对这里所说的数据也可以大致归纳出其含义与特征:数据应该是原始的、广义的、可鉴别的抽象符号;数据可以用来描述事物的属性、状态、程度、方式等;数据符号单独表示时没有任何含义,只有把它们放入特定的场合进行解释和加工,才能使其具有意义并升华为信息。

2. 信息的传递与应用

人类在古代就有了各种转载信息的手段:如结绳记事、烽火台、指南针、语言、文字、纸张、印刷术、风标、号角、邮传驿站、了望塔楼等。到了近代,又有了各种信息技术的产物:如望远镜、显微镜、算盘、手摇机械计算器。现代信息技术的产品就更多:电报电话、录音机、录像机及电子计算机、人造卫星、C⁴IRS 系统、机器人、无人驾驶系统等。人类在不断地改进信息处理的技术和工具,以适应社会发展中日益增多的信息处理的需要。

信息的传递目前主要有两种方式:通信和广播。通信包括交换、传输,广播包括发射、接收。信息的内容包括声音、数据、图像等。信息传递方式包括点到点(专线)、一点到多点(组播)、点到面(广播)和可选择的点到点(交换)等。交换包括电路交换和分组交换(包交换),两种交换方式还会有相当长的时间并存。传输包括有线传输和无线传输。有线传输主要是金属线和光纤系统;无线传输可按照波长来分,如长、中、短波和微波等;也可按照方式来分,如地面、卫星等。人和计算机之间的交流可能是信息传输式的,也可能是会话式的,

但这种会话并不真正发生在人与计算机之间。实际上人机会话是用户与程序员之间的对话,这种对话是间接的,对话中使用的符号是有限的,对话的主题也是事先选择好的,程序做出的各种应答也是预先“录好的”(用程序编写的)。所以,人机之间的会话目前实际上是通过指令和数据来表达的。

信息的应用非常广阔。认知、科学探索、知识传播、生产流程的控制、管理(宏观管理、微观管理)、娱乐(与声像设备结合)以及人与人之间的交流等,发展都很迅速。

3. 信息的主要特征

(1) 信息具有不灭性 物质是不灭的,能量也是不灭的,其形式可以转化。信息不像物体和能量,它不能单独存在,必须借助某种符号才能表现出来,必须寄载于某种物体之上,即信息必须依附于载体。信息是事物运动的状态和方式而不是事物本身。信息的不灭性是指一条信息产生后,其符号和物质载体可以变换,甚至载体可以被毁掉,如一本书、一张光盘,但信息本身并不会被消灭。另外,某些信息的使用有很强的时效性,如金融信息和战争信息等。

(2) 可传递性和共享性 信息在空间上的传递称为通信。在通信上,通常把信息的通路称为信道,如电话、微波、卫星等;把信息的发布者称为信源;把信息的接受者称为信宿。同一信源可供给多个信宿,而信源自身的信息可秋毫无损,这就是信息的共享性。信息在时间上的传递称为信息存储。利用存储介质(如纸张、磁带、磁盘、光盘等)可将用文字、声音和图像等符号表示的信息记录下来,使其永久为人类服务。信息可以廉价复制,广泛传播。信息的创造可能需要很大的投入,但复制只需要载体的成本。当然,无论是信息共享还是信息传播,都不能侵犯知识产权,这是信息社会必须遵守的信息道德。

(3) 信息的可处理性 可以通过压缩、存储、排序、转换形态等方法对信息进行加工处理。经过综合、分析等处理,原有的信息可以实现增值,能够更有效地服务于不同的人群或不同的领域。信息并不神秘,不断获取信息,不断提高信息识别和加工处理能力,自觉抵制污染社会环境的有害信息,这些都是信息时代应该倡导的信息意识。

1.1.2 信息科学与信息技术

信息科学是研究信息的获取、表述、存储、识别、编码、处理等传输过程和处理阶段中各种信息问题的科学。信息科学由计算机科学、电子与信息系统科学、半导体、光学和自动化科学等学科组成。

1. 关于信息科学的定义

“信息科学”这个名称早在50年前就出现了。从1948年《信息论》和《控制论》的问世到现在,尝试建立信息科学理论一直是许多人的梦想。而近年来,信息理论像一个多学科的讲坛,不同的信息问题研究者,把一些分布在不同领域里的信息概念加以阐述、系统化后提出了许多新颖的命题。传播学家说:公平的信息获取才是信息对称的基本保障;通信工程师说:当接收者对一个事件出现的概率估计值越小时,他获得的信息量就越大;计算机科学家则说:微处理器时钟频率的大小,决定着它处理信息的速度高低;生物学家又说:脱氧核糖核酸上的信息是控制子代和亲代相像的惟一源泉。

由美国普林斯顿大学马克鲁普1983年主编的《信息研究:学科之间的讯息》,德国科特布斯技术大学肯沃奇1994年主编的《信息:多学科概念中的新问题》和奥地利维也纳技术

大学霍夫克奇纳 1999 年主编的《探寻统一信息理论》这三个大型的信息科学交叉研究报告中,学者们提出:“信息科学”中的信息范围不应该再局限于某些特定领域,而应该是多元化的,只有统一的信息科学才是真正的信息科学。但建立统一信息科学理论是项艰巨复杂的工作,它代表着人类对纷繁信息世界的了解,并在理论上有了进一步的升华。首先对许多分支要有深刻的了解,信息科学涵盖着从电子学到哲学的许多基本问题,涉猎自然科学、技术科学、人文科学、社会科学。由于统一信息科学关注的信息对象分布在不同领域里,所以,要发现不同领域里信息现象的共同规律是很困难的。又由于信息是和通信或传播联系在一起,而分析通信与传播问题的首要条件是人们必须考虑两个以上的个体。所以,凡是使用信息的地方,就必然使人们产生出关于对象的联想或系统化思维,这就是信息概念的生命力和活力之所在。

所有部门信息科学所涉及的信息问题和人类的生产实践直接相关。而生产实践,也只有生产实践才是推动科学发展的真正动力。除了机械信息学、生物信息学、化学信息学、人类信息学等会继续有大的发展之外,了解低等动物之间通信的动物信息学也会有所进展。全(信)息学、信息物理学、土地信息学、地理信息学、心电信息学、医疗信息学、生命信息学等也正在吸引着许多学者的研究兴趣。

2. 信息技术

信息技术是指对信息的搜集、存储、传递、分析、使用等的处理技术和智能技术。具体地讲,包括软件开发技术、通信网络技术、微电子技术、信息处理技术和多媒体技术,而传感、自控和新材料技术等是信息技术的相关技术。信息技术是目前各领域高新技术的关键和核心,更是信息产业的基础。

现代信息技术包括四大类:① 电子信息技术:电子感测技术、电子通信技术、电子计算机、电子控制技术;② 量子信息技术:量子计算机;③ 激光信息技术:激光遥感、光导纤维通信、激光全息存储、激光控制技术、激光计算机;④ 生物信息技术:生物开关器件、生物存储器件、生物逻辑器件、生物计算机,等等。

信息技术产业的主要技术和产品范围包括:① 多媒体技术,其中包括多媒体计算机技术、PC 技术、液晶等高清晰度显示技术等;② 数据存储和处理技术,其中包括超巨型和超微型计算机技术、语言识别和神经网络等智能计算机技术、分子电子技术、计算机免疫系统技术等;③ 传输技术,包括光纤和卫星等通信技术、数字声像技术、各种调制和解调技术、各种传感器技术、交互式网络技术等。

1.1.3 信息科学理论发展的百年历史

信息科学理论的建立经历了一百多年的时间。

1854 年,英国数学家乔治·布尔(George Boole, 1815—1864)创建了逻辑代数理论,提前一个世纪为现代二进制计算机铺平了道路。

1924 年奈奎斯特(Nyquist)发表了题为“影响电报速率因素的确定”的文章。1928 年哈特莱(Hartley L. V. R)又发表了题为“信息传输”的文章,在这两篇文章里给出了信息度量的方法,研究了通信系统传输信息的能力。

1936 年,英国剑桥大学的数学家艾伦·图灵(Alan Mathison Turing, 1912—1954)发表了

题为“论可计算数及其在判定问题中的应用”的论文,提出了著名的理论计算机模型——图灵机。现在几乎所有的编程语言都建立在图灵机模型之上。图灵被誉为计算机科学的奠基人。

1945年,美籍数学家冯·诺伊曼(John von Neumann,1903—1957)等人首次发表题为“电子计算机逻辑结构初探”的报告,奠定了存储程序式计算机的理论基础,并开始研制相应的EDVAC计算机。

1947年,美国计算机协会(Association for Computing Machinery,ACM)在华盛顿哥伦比亚大学成立。

1950年,图灵在Mind杂志发表文章,提出了测试机器智能的Turing规范。

1951年,莫里斯·威尔克斯(Maurice Wilkes)提出微程序的设计概念,这是一种设计计算机系统控制部分的有效方法。

1952年,无线电工程师协会创办了电子计算机方面的I.R.E学报,这是IEEE计算机学报的前身。

1.2 信息理论的奠基人与ACM图灵奖

这里我们回顾计算机的发展史,介绍几位重要历史人物(如图1.1所示)。计算机史充满了智慧的火花,以史为鉴,前辈们刻苦钻研、锐意创新的精神,对今天的计算机工作者来说仍有很大的启迪作用。



图 1.1 计算机科学理论的奠基人

1.2.1 信息论之父——香农

1. 香农博士其人

克劳德·埃尔伍德·香农(Claude Elwood Shannon,1916—2001),美国数学家,信息论的创始人。在大学时他就表现出对数理问题的高度敏感,1940年获得麻省理工学院数学博士学位和电子工程硕士学位。他的硕士论文是关于布尔代数在逻辑开关理论中的应用,并证明布尔代数的逻辑运算,可以通过继电器电路来实现。1941年他加入了贝尔实验室数学

部,在此工作了 15 年。在这个世界上最著名的研究基地里,他的前辈最早研究了通信系统的信息传输能力,第一次提出了信息量的概念,并试图用数学公式予以描述。香农创造性地继承了他们的事业。

1948 年 6 月和 10 月,由贝尔实验室出版的《贝尔系统技术》杂志连载了香农博士的文章“通信的数学原理”。“通信的基本问题就是在一点重新准确地或近似地再现另一点所选择的消息”,这是香农在该文中的一句名言。正是沿着这一思路,香农应用数理统计方法来研究通信系统。他在文中用非常简洁的数学公式定义了信息时代的基本概念:熵和信道容量,指出了用降低传输速率来换取高保真通信的可能性。次年,香农又发表了“噪声下的通信”。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、信息度和信道容量与噪声的关系等有关精确地传送通信符号的基本技术问题。两篇文章成了现代信息论的奠基之作。这些贡献对今天的通信业具有重大影响。因此香农被誉为信息理论的奠基人。那时,他才不过刚刚 30 出头。

2. 香农极限及其应用

“熵”的概念起源于热力学,是度量分子不规则热运动的单位。香农的伟大贡献在于:利用概率分布的理论给出了“熵”的严格定义。在熵的基础上定义的信道容量也是通信中一个至关重要的概念。由此,香农推出了一个公式,明确表达了在不同噪声情况下传输速率与失真的定量关系。从这一个公式导出的为达到无失真通信的传输速率的极限,现被称为香农极限。

如今,这两个原理已广泛应用于信息处理和实际通信中。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。像现在人们在 PC 机上常用的 WinZip 就是一个例子。在数字通信中,各种语音信号在用数学模型提取参数(有损压缩)后也要对参数进行熵编码和无损压缩,这一无损压缩的算法就是根据熵的概念及其公式导出的。经过压缩后的数据还要进行扩张(信道编码),因为根据香农的第二个理论,要尽量做到无失真的传送,就要用类似重复的方法来纠正可能出现的错误。毫不夸张地说,没有这两个方法,就没有今天的手机通信。同样,在因特网上传递多媒体数据,无损压缩往往是必不可少的。比如现在所用的图像格式,除了 BMP 文件以外,其他格式都用到基于熵概念的无损压缩算法,还有当今广为人知的 MP3 音乐的压缩格式等。所以不论在电信业还是计算机业中,只要涉及信息的压缩与传递,就要用到香农的理论。香农的理论不仅在电子学的其他领域,如计算机、自动控制等方面大显身手,而且遍及物理学、化学、生物学、心理学、医学、经济学、人类学、语音学、统计学、管理学等学科。它已远远地突破了香农本人所研究和意料的范畴,即从香农的所谓“狭义信息论”发展到了“广义信息论”。

1.2.2 图灵与图灵机

1. 图灵其人

英国数学家艾伦·图灵,1912 年生于伦敦近郊的帕丁顿(Paddington)镇。1931 年图灵进了剑桥的“国王学院”(King's College)攻读数学。毕业时的数学学位考试拿了一等奖,取

得这个成绩的学生在剑桥大学有一个特别的荣誉称号,叫 Wrangler。图灵的学位论文是“概率论的中心极限定理”,1936年图灵因就同一课题所发表的论文“论可计算数及其在判定问题中的应用”而获得史密斯奖(Smith Prize)。

1935年,图灵开始对数理逻辑产生兴趣。数理逻辑(Mathematical Logic)又叫形式逻辑(Formal Logic)或符号逻辑(Symbolic Logic),是逻辑学的一个重要分支,其起源可追溯到17世纪德国的大数学家莱布尼茨(Gottfried Wilhelm Leibniz,1646—1716)。在莱布尼茨的思想中,数理逻辑、数学和计算机三者均出于一个统一的目的,即人的思维过程的演算化、计算机化,以至在计算机上实现。但是,“计算机”到底应该由哪些部分组成,如何进行计算和工作,在图灵之前的两个多世纪中,没有人能够说清楚。正是图灵上述1936年那篇论文第一次回答了这个问题,提出了著名的理论计算机的抽象模型,可以把推理化做一些简单的机械动作。图灵提出的计算模型被称做“图灵机”(Turing Machine)。几乎所有现代程序设计语言都建立在图灵机模型之上。图灵被誉为计算机科学的奠基人。1938年,图灵取得美国普林斯顿大学博士学位后回到英国剑桥大学,他的博士论文题目是“基于序数的逻辑系统”,主要研究了由乔治·布尔于1854年创建的逻辑代数,并且还自己动手用继电器搭逻辑门组成了一台乘法器。由于图灵对计算机的一系列杰出贡献和重大创造,1951年他被选为英国皇家学会院士。

图灵的一生充满神秘与传奇色彩。二次世界大战时,图灵领导了对德军密码的破译工作,他用继电器做成被叫做“霹雳弹”的译码机(Bombe,因为继电器工作时发出“霹雳啪啦”的声音而得名),破译了德军不少“恩尼格玛”密报(Enigma,德国密码机名称),为盟军战胜德国法西斯立了不少功劳。因此战后图灵被光荣授勋,被称为OBE(Officer Order of the British Empire)。“霹雳弹”后来改用电子管,命名为“巨人”(Colossus),被认为是第一台投入运行的电子数字计算机。但英国政府直到20世纪90年代初才解密“巨人”的资料。图灵身体健壮、酷爱长跑。1947年8月25日,The Times体育新闻报道他的马拉松成绩为2小时46分3秒。图灵加盟曼彻斯特大学后为计算机科学所做出的又一个杰出贡献是他在1950年10月发表的论文“计算机和智能(Computing Machinery and Intelligence)”。在这篇经典的论文中,图灵进一步阐明了他认为计算机可以有智能的思想,并提出了测试机器是否有智能的方法,他称之为“模仿游戏(Imitation Game)”,而大家现在称其为“图灵测试(Turing Test)”。

2. 图灵机的工作原理

图灵机是如何工作的?下面做一简单介绍^[2]。图灵机由三部分组成:一条双向都可无限延长的被分为一个个小方格的磁带、一个有限状态控制器和一个在带子上可以左右移动的读/写磁头,如图1.2所示。

图灵机有步骤地进行工作,工作情况取决于3个条件,即:机器的内部状态;读/写磁头扫描在磁带的哪个方格上;该方格上有什么信息。

机器执行一步工作的过程如下:

(1) 读/写磁头在所扫描的方格上写上符号,原有符号自然消除;



图1.2 图灵机的组成原理