

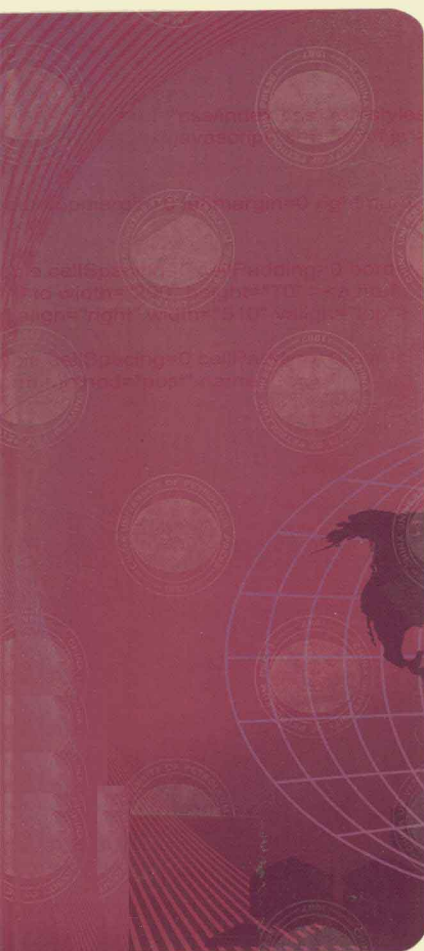


普通高等教育“十一五”国家级规划教材
山东省高校统编教材

■ 第五版

大学IT

山东省教育厅组编



☒ Windows XP
☒ Office 2003

中国石油大学出版社



普通高等教育“十一五”国家级规划教材
山东省高校统编教材

大学 IT

(第五版)

山东省教育厅组编

主 编 刘法胜

副主编 曹宝香 李成江

周应兵 张庆德

编 委 刘庆华 刘法胜 刘培玉

宋秀芹

张 磊 张庆德 李 丽 李成江

李盛恩 周应兵 郑永果 赵铭建

耿玉水 曹宝香 尉永清

中国石油大学出版社

图书在版编目 (CIP) 数据

大学 IT / 山东省教育厅组编. — 5 版. — 东营:
中国石油大学出版社, 2010.7

ISBN 978-7-5636-3130-8

I. ①大... II. ①山... III. ①电子计算机—高等学校
—教材 IV. ①TP3

中国版本图书馆 CIP 数据核字 (2010) 第 116790 号

大 学 IT (第五版)

山东省教育厅组编

主 编: 刘法胜

责任编辑: 刘玉兰

出 版 者: 中国石油大学出版社 (山东 东营, 邮编 257061)

网 址: <http://www.uppbook.com.cn>

电子信箱: eyi0213@163.com

印 刷 者: 沂南县汇丰印刷有限公司

发 行 者: 中国石油大学出版社 (电话 0546-8392062)

开 本: 185 × 260 印张: 16.75 字数: 429 千字

版 次: 2010 年 7 月第 5 版第 1 次印刷

定 价: 27.60 元

版权专有, 翻印必究。举报电话: 0546-8392062

本书封面覆有带中国石油大学出版社标志的激光防伪膜。

本书封面贴有带中国石油大学出版社标志的电码防伪标签, 无标签者不得销售。

前言 Preface

欢迎选用《大学IT》！编写这本教材的灵感和思路来自《大学数学》、《大学英语》、《中学信息技术》和“Fluency with Information Technology: Skills, Concepts, and Capabilities”。

人类社会已进入信息时代，计算机文化、计算机应用和计算机技术已渗透到社会的各个领域。社会信息化和信息社会化已经成为社会演化、发展和进步的车轮。信息社会的主要特征之一是，谁能获得信息技术（IT）的优势，谁就能在一日千里的竞争中占据优势。每一个民族和国家都在努力争取占领信息技术的制高点。高等学校中的信息技术教育，在科教兴国现代化战略的实施中起着重要的作用。

这是本书第五版修订。至此，简要回顾一下信息技术教育的历程或许是有益的。信息技术的主体是计算机技术。我国高等院校计算机公共课教学始于20世纪80年代，大体经历了三个阶段：第一阶段是以程序语言为主要教学内容的阶段，由于历史因素的限制和计算机学科发展过程的惯性影响，将计算机理解为单纯的计算工具，高等院校计算机基础教育也就留下了“偏离”的烙印；进入20世纪90年代初，以直接应用为主体的操作技术教育形成了第二阶段的计算机教育特点；自20世纪90年代中期以来，计算机文化教育的开展在真正意义上形成了高等院校计算机公共基础教育的气候。

1995年春，原山东省教育委员会在《关于加强普通高校计算机基础教学的意见》中明确规定，高校非计算机专业的计算机基础教育公共课分为三个层次：计算机文化基础、计算机应用基础和计算机技术基础。山东省高等院校计算机公共课程教学在此基础上得到了大的发展和提高。我们在1994年全国第一本《大学计算机文化》初版的基础上，组织出版了《计算机文化基础》教学用书，至今已经连续八次修订再版。

进入新世纪，为适应信息技术发展和教学需要，我们又编写了全国第一本信息技术教材《大学IT》，以期逐步更替原来的《计算机文化基础》教学用书。经过试用，先后修订了四次，2010年初又在12所高校征求使用和修订意见，为适应知识更新和教学改革的需要，吸收国际上大学IT的有效做法和风格，完善《大学IT》的相对独立性，在好教、好学和好用上下功夫，对《大学IT》进行了系统修改。

建立信息的采集、加工和输出的思维习惯，具备使用计算机处理信息的观念和能力，是信息社会的时代文化。20世纪80年代，美国计算机文化协会组织人、

计算机教育先驱 Luehrmann 博士称之为 Computer Literacy; 20 世纪 90 年代, 美国计算机文化教育权威 O' Leary 夫妇称之为 Computer Competency。

《大学 IT》努力体现基础理论和先进技术的统一。计算机不仅是一个具体的物理对象, 还是一个抽象的计算模型。如果说中学信息技术教材体现直观和实用性, 那么, 《大学 IT》在此基础上还要体现抽象性和系统性。《大学 IT》内容包括信息社会、信息科学、信息技术、信息工程、信息管理和信息经济等。

《大学 IT》一书, 参考最新《普通高等院校计算机基础教育大纲》及美国计算机科学与技术专业教学改革 CC2001 方案, 借鉴计算机专业教学改革思路, 与中小学 IT 教育大纲接轨, 在山东省高校教学改革立项课题《大学 IT 教学体系建设与改革研究》工作基础上编写而成。

全书共分九章, 包括信息技术中的信息来源、信息传输、信息处理和信息应用的诸多方面。每章分五节, 内容包括该章概述、理论方法、技术要求、系统和应用。《大学 IT》新版努力做到使其内容丰富, 形式简约。

“知之莫如欲知, 欲知莫如乐知”。将知识的灌输和积累过程转变为知识的指导与求知过程是本书的指导思想之一。当然, 将获取知识的过程变成享受过程目前只是一种奢望而已。不过, 理论指导、上机实践和网络平台学习三个环节, 为上述目标的实现奠定了基础。

“知者行之始, 行者知之成”。《大学 IT》主要供课堂教学使用, 与本书配套使用的《大学 IT 实验教程》, 包含了相应内容的实验指导模块, 使大学 IT 知识体系的认知与实践相得益彰。《大学 IT》电子版本、教学课件与相应的网络平台建设也在同步进行, 以实现资源共享, 优势互补。

信息学科的综合性和交叉性, 基础性和先进性等决定了该课程的特点。《大学 IT》的目的是给学生一个体系构架, 变知识灌输为理论指导。《大学 IT》给教师和学生一个开放的学习平台。

《大学 IT》作为山东省高校计算机基础教学改革教材是集体智慧和汗水的结晶。参加本书编写工作的院校有山东科技大学、山东师范大学、曲阜师范大学、聊城大学、山东建筑大学、山东警察学院、山东交通学院、潍坊学院、德州学院、山东轻工业学院等。各章执笔人员为: 第 1 章, 郭爱章、徐鑫; 第 2 章, 鲁燃; 第 3 章, 史士英; 第 4 章, 曹宝香; 第 5 章, 杜忠友; 第 6 章, 张庆德; 第 7 章, 尉永清、张辉; 第 8 章, 张磊、王鑫; 第 9 章, 宋秀芹。山东科技大学的连剑和任国强两位老师为本书的初稿校读、整理和课件制作做了大量的工作。全书由刘法胜、李成江统稿。

由于编者水平所限, 不妥甚至谬误之处在所难免, 希望同行和读者批评指正, 以期再版时改进。

编者

2010 年 5 月

目 录 Contents

第1章 IT概论	1	2.1.5 计算机的发展趋势	35
1.1 信息与信息社会	1	2.2 计算机的工作原理	36
1.1.1 信息定义	1	2.2.1 冯·诺依曼计算机结构	36
1.1.2 信息的特征	2	2.2.2 计算机的心脏——CPU	39
1.1.3 信息社会	4	2.2.3 存储器	43
1.2 信息的数据表示	5	2.2.4 计算机的工作过程	50
1.2.1 数据的有效表示	5	2.3 微机系统及其主要指标	51
1.2.2 常见数据的表示	6	2.3.1 微机分类	51
1.2.3 数据压缩	9	2.3.2 微机的主要性能指标	51
1.3 信息科技	10	2.3.3 微机的常见外部设备	52
1.3.1 信息科学	10	2.4 嵌入式计算机系统	55
1.3.2 信息技术	11	2.4.1 嵌入式系统基础知识	55
1.4 信息经济	14	2.4.2 嵌入式系统的组成	57
1.4.1 电子商务	14	2.4.3 嵌入式系统的应用	60
1.4.2 IT产业	17	2.5 计算机应用	62
1.5 信息科技应用	18	2.5.1 计算机应用的社会意义	62
1.5.1 信息科技在通信领域中的应用	18	2.5.2 计算机内部结构、输入与输出	62
1.5.2 信息科技在计算领域中的应用	19	2.5.3 计算机硬件的安装	65
1.5.3 信息科技在管理领域中的应用	20	小 结	67
1.5.4 信息科技其他日常应用	21	信息网	67
小 结	23	习 题	67
信息网	23	阅读材料	69
习 题	25	第3章 计算机软件	73
阅读材料	25	3.1 软件概述	73
第2章 计算机硬件系统	29	3.1.1 计算机软件	73
2.1 信息工具——计算机	29	3.1.2 软件与硬件	74
2.1.1 计算机的起源与发展	30	3.1.3 软件兼容性	74
2.1.2 计算机的特点	32	3.1.4 软件版本	75
2.1.3 计算机的分类	32	3.1.5 软件版权	76
2.1.4 计算机的应用	33	3.2 系统软件	76
		3.2.1 操作系统	76
		3.2.2 几种典型的计算机操作系统	77

3.2.3 数据库管理系统	78	4.5.2 软件开发工具	110
3.2.4 程序设计语言处理程序	79	4.5.3 VB 计算机语言	112
3.2.5 实用工具程序	79	小 结	114
3.3 应用软件	80	信息网	114
3.3.1 通用应用软件	80	习 题	115
3.3.2 专用应用软件	81	阅读材料	116
3.3.3 中间件	81		
3.4 软件工程	82	第 5 章 网络与通信	119
3.4.1 软件危机	82	5.1 通信基础	120
3.4.2 软件工程	83	5.1.1 通信发展历程	120
3.4.3 软件企业	83	5.1.2 通信原理	120
3.4.4 软件产业	83	5.1.3 通信技术	121
3.5 计算机软件应用	84	5.1.4 通信系统	123
3.5.1 计算机软件的社会应用	85	5.2 网络世界	125
3.5.2 计算机软件的技术应用	85	5.2.1 走进网络	125
3.5.3 计算机软件的日常应用	86	5.2.2 计算机网络基础	127
小 结	91	5.3 网络技术	129
信息网	91	5.3.1 计算机网络的基本组成	129
习 题	92	5.3.2 网络拓扑结构	131
阅读材料	93	5.3.3 网络协议	133
		5.3.4 网络操作系统	135
第 4 章 计算机程序设计与语言	96	5.4 Internet	136
4.1 概 述	96	5.4.1 Internet 基础	136
4.1.1 计算机程序	97	5.4.2 Internet 的接入方式	138
4.1.2 程序设计步骤	97	5.4.3 Internet 在中国	139
4.1.3 程序设计语言	97	5.5 通信与网络应用	140
4.2 模型与算法	98	5.5.1 通信与网络的社会应用	140
4.2.1 建立模型	98	5.5.2 通信与网络的技术应用	142
4.2.2 算法设计	98	5.5.3 通信与网络的日常应用	142
4.2.3 算法表达	100	小 结	145
4.3 程序设计实现	101	信息网	145
4.3.1 程序举例	101	习 题	146
4.3.2 编写程序	102	阅读材料	147
4.4 程序的测试、调试与维护	104		
4.4.1 程序测试	105	第 6 章 网络信息安全	149
4.4.2 程序调试	106	6.1 网络信息安全概述	149
4.4.3 程序维护	106	6.1.1 网络信息安全的含义	149
4.5 程序设计语言应用	107	6.1.2 网络信息安全的结构层次	150
4.5.1 程序设计语言的分类与选择	107	6.1.3 网络信息安全面临的威胁	150

6.1.4 网络信息安全对策	151	7.3.3 数据挖掘	182
6.2 密码理论简介	152	7.4 常用的数据库管理系统	184
6.2.1 基本概念	152	7.4.1 Access 数据库	184
6.2.2 网络通信中的加密方式	153	7.4.2 Visual FoxPro 数据库	185
6.2.3 著名密码算法举例	154	7.4.3 SQL Server 2005 数据库	185
6.3 计算机病毒	155	7.5 数据库应用	186
6.3.1 病毒的原理、特点与传播途径	155	7.5.1 Access 2003 的启动与退出	186
6.3.2 病毒的类型	155	7.5.2 Access 2003 环境	187
6.3.3 病毒的预防	156	7.5.3 Access 数据库	188
6.3.4 病毒的清除	157	小 结	197
6.4 防火墙与入侵检测	158	信息网	197
6.4.1 防火墙的概念	158	习 题	198
6.4.2 防火墙体系结构	158	阅读材料	199
6.4.3 入侵检测的概念	160	第 8 章 多媒体	203
6.4.4 入侵检测系统的工作原理和分类	161	8.1 多媒体世界	203
6.5 信息安全应用	162	8.1.1 多媒体的发展过程	204
6.5.1 社会信息安全	162	8.1.2 多媒体的未来	204
6.5.2 信息安全技术的应用	163	8.2 多媒体中的基本概念	205
6.5.3 信息安全技术的操作与实践	163	8.2.1 媒体及其五种形式	205
小 结	163	8.2.2 多媒体信息类型	205
信息网	163	8.2.3 超文本与超媒体	206
习 题	164	8.3 多媒体技术	206
阅读材料	166	8.3.1 多媒体技术的特点	206
第 7 章 数据管理	168	8.3.2 音频信息处理技术	207
7.1 数据管理概述	168	8.3.3 图像信息处理技术	209
7.1.1 数据	168	8.3.4 视频信息处理技术	211
7.1.2 数据组织的基本概念	168	8.4 多媒体计算机系统	212
7.1.3 数据管理技术的发展	169	8.4.1 多媒体计算机的硬件系统	212
7.2 数据模型	171	8.4.2 多媒体计算机的软件系统	213
7.2.1 数据模型的概念	171	8.5 多媒体应用	214
7.2.2 概念数据模型	171	8.5.1 多媒体社会应用	214
7.2.3 结构数据模型	173	8.5.2 多媒体技术应用	215
7.3 数据管理技术	176	8.5.3 实用多媒体软件	218
7.3.1 数据库	176	小 结	220
7.3.2 数据仓库	178	信息网	220
		习 题	220
		阅读材料	221

第9章 IT新应用展望	224	9.4 IT应用大视角	240
9.1 信息采集技术	224	9.4.1 导航与卫星定位技术	240
9.1.1 传感器	225	9.4.2 3S技术	241
9.1.2 传感器的应用	226	9.4.3 智能机器人	244
9.1.3 传感器的发展方向	227	9.4.4 智能大厦	245
9.1.4 物联网世界	228	9.5 IT发展展望	248
9.2 信息传输技术	229	9.5.1 3G与三网合一	248
9.2.1 光纤通信原理	230	9.5.2 自动驾驶	250
9.2.2 红外和蓝牙	231	9.5.3 量子计算机	251
9.2.3 卫星通信	233	9.5.4 生物计算机	252
9.3 信息处理技术	234	9.5.5 社会计算	254
9.3.1 数字图像处理	234	9.5.6 云计算	256
9.3.2 声音信息处理	236	小 结	258
9.3.3 电子指纹	236	信息网	258
9.3.4 自动翻译	237	习 题	259
9.3.5 模式识别	238	阅读材料	259

第 1 章

Chapter One

IT 概 论

不出户，知天下；不窥牖，见天道。

——老子《道德经》

“The time has come,” the Walrus said, “to talk of many things.”

——L.Carroll

本章概要

作为信息时代的主角，信息越来越受到人们的重视。那么，什么是信息，它有什么特征，又有什么作用呢？本章将围绕这些问题介绍信息、信息科学和信息技术及其相关知识，从不同层次给出信息论的总体框架。学完本章，你将能够：

- ◎描述信息的概念和特征，了解信息社会；
- ◎了解数据的表示和常见信息的表示；
- ◎说明信息科学和信息技术概况；
- ◎了解信息经济中几种业务模式；
- ◎了解信息系统及其作用。

1.1 信息与信息社会

信息犹如空气和水一样存在于人类社会。从远古到当今的文明社会，信息一直在发挥着重要作用，它是人类生存和社会发展所必需的基本资源。在日常生活中，人们时时刻刻都在与信息打交道，听到、看到、感受到各种信息。社会信息化和信息社会化已经成为当今社会的一个主要特征。

1.1.1 信息定义

信息能够反映物质和能量的形态、结构、状态等特征，可以用人的感官感觉到，也可以用仪器、仪表和各种传感器进行探测。信息是一个动态的概念，随着时间的推移，信息不断地被赋予新的含义。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、信息管理和信息论等紧密地联系在一起。

信息是一个不断变化，至今尚未有公认的、确切的、统一定义的概念，先后出现了多种定义。

◎1948年,美国数学家、信息论的创始人香农(Shannon)在题为《通信的数学理论》的论文中指出:“信息是用来消除随机不确定性的东西。”

◎1948年,美国著名数学家、控制论的创始人维纳(Wiener)在《控制论》一书中指出:“信息就是信息,既非物质,也非能量。”

◎钟义信教授提出:“信息是关于事物运动的状态和规律,或者说是关于事物运动的知识。”李衍达院士则认为,信息反映的是事物的状态、特性和变化。

◎逆香农信息定义:信息是确定性的增加。

◎逆维纳信息定义:信息就是信息,是事物属性标识的集合。

这些说法虽然不一致,但都从不同侧面、不同角度反映了信息的本质。一般认为:信息是在自然界、人类社会和人类思维活动中普遍存在的一切物质和事物的属性。

1.1.2 信息的特征

1. 信息的特征

信息虽无确切定义,但是却具有两个明显的特征:广泛性与抽象性。

所谓广泛性可从以下三个方面来理解:

(1)客观世界充满着信息:天上的星体、地下的矿藏,一切客观物质无不具有自己的信息。

(2)人类离不开信息:人的五官在不停地感知、接收信息;人的神经系统在不停地传递信息;人的大脑在不停地处理与决策信息;人与人之间又在不停地交流信息;人活在世上百分之百的时间都在自觉或不自觉地与信息打交道。

(3)知识、书本是有用信息的积累:人类依靠知识改造自然、适应自然,靠知识促进社会的发展与进步。

所谓抽象性是指,信息是组成客观世界并促进社会发展的最基本的三大要素之一。

物质、能量和信息是构成世界的三大要素。物质是一种资源,它提供各种各样的材料;能量是一种资源,它提供各种形式的动力;信息也是一种资源,它提供知识。没有物质,宇宙将成为虚无;没有能量,世界将变得死寂;没有信息,人生将失去意义。物质运动是由物质、能量和信息构成的,信息源自于物质运动。信息是对事物运动状态或存在方式的描述和反映,它通过某种物质形式存在、传播和表现。信息依附于物质和能量,但又不等同于物质和能量。没有信息就不能更好地利用物质和能量,人类利用信息和知识改造物质,创造新物质,提高能量利用效率,发现新能量形式。信息也是客观存在的,它是人类认识、改造客观世界的主要动力,是人类认识客观世界的更高层次。

2. 信息独特的性质

信息不同于物质和能量,有其独特的性质。信息具有共享性、独立性、形式多样性、本质性、普遍性和无限性。

1) 共享性

信息的传播是一个完整的系统过程,由信源、信宿和信道等组成。其中,信源是信息的发布者,信宿是信息的接收者,信源和信宿之间的信息传输的途径与设备称为信道。信息的共享性是指同一信源可供多个信宿,信源发出信息后,其自身的信息并不减少。也就是说,信息可以被多次使用而不会损耗和消失。信息的这种特性是信息不同于物质和能量的重要特征之一。

一般来说,物质和能量使用之后,就消耗或转化了,而信息则不然。广播、电视不会因

为听的人或看的人多了,其他人听或看的内容就少了。信息随着每一次交换,都会产生一个新的副本,而不会消失或损耗。你有一个苹果,我有一个苹果,交换之后,每一人还是只有一个苹果;你有一条信息,我有一条信息,交流后则每人拥有两条信息。这体现了物质与信息的一种本质差别。

2) 独立性

信息可以脱离它所反映的事物而被独立保存和传播。我们现在可以看到《道德经》、《逍遥游》、《论语》、《中庸》等两千多年前的著作,这些书里保存着我们祖先深刻的思想精髓,反映的是当时的信息。时光流逝,沧海桑田,这些书产生的背景早已消失不见,然而书中的信息却还能让我们感受到先人的睿智。我们能从资料上看到火星的地貌、木星的光环、太阳的日冕。人们并未到达这些地方,之所以能了解这些星体的情况,是因为我们能通过各种仪器接收它们所发出的光线等信息。对于我们居住的地球,人们并未进入地壳以下的地方,但人们对地球的内部并非一无所知。人们通过考察重力、地磁、地震波等情况就能探寻出这个星球的许多秘密。信息可以与它所反映的事物分离而独立存在。我们也可以保存信息,使它摆脱产生它的时间和空间的限制,这样我们就能与先哲对话、看遥远的宇宙、探察神秘的地核运动。当然,信息虽然具有独立性,但并不意味着信息可以凭空产生。

3) 形式多样性

信息的保存、表现与传播必须有一定的载体,载体的形式可以是多样的,信息的形式可用其载体形式表征。以书籍作例子,若想保存《道德经》,我们可以把它印成书,不管是简体的还是繁体的,是横排还是竖排,它的内容——信息都是一样的。用磁带、光盘作为它的载体也行,反映的信息仍然相同。信息的载体可以多种多样,但是不存在没有载体的信息。信息是物质运动的反映,它也必须依附于物质运动而存在。这决定了信息必须借助于某种载体而存在、传播和表现。

4) 本质性

信息能够反映事物的运动、状态、规律和变化。相对于物质和能量,信息更能表现事物的内在规律和本质内涵。物质不灭、能量守恒本身就是关于事物本质的信息。信息的获取、传播与保存更有利于人们认识和掌握客观事物规律。信息的交流对人类的认知、对人类依据客观条件做出科学决策等起着至关重要的作用。

5) 普遍性

宇宙中所有事物都是物质,都在运动着。信息是物质运动的反映,因而所有事物都包含信息。信息世界包括了所有事物,这就是信息的普遍性,也是我们整个社会都离不开信息的根本原因。

6) 无限性

信息是可以无限增长的资源。随着社会进步、科技进步,能被人类所感知的信息越来越多。我们可以看到原子的世界,可以考察千百万光年以外的宇宙,人类获取信息的范围空前扩大。新的科学手段不断地把更多的信息纳入人们可以利用的范围。物质是多样的,物质运动是多样且永不停息的,因而信息也是无穷无尽的,这与哲学上认识的无限性是一致的。

3. 信息与数据、知识的联系

信息与数据、知识等有密切的联系,但它们不是等价的概念。

1) 数据与信息

信息是决策者进行成功决策的基础,是经过加工处理并对人类客观行为产生影响的数据

表现形式。如果没有合理的数据与信息,决策只能是一种猜想。

信息与数据之间存在着固有的联系:数据(data)是反映客观事物属性的记录,是信息的载体;信息则是数据的内涵,是对数据语义的解释。信息常常通过数据形式表示出来,被人们理解和接受。从计算机的角度看,数据泛指可以被计算机接受并能够被计算机处理的符号。

数据与信息是有区别的。信息是有用的数据,数据是信息的一种表现形式。信息是通过数据符号来传播的,数据如不具有知识性和有用性则不能称其为信息。

从信息处理角度看,任何事物的属性都是通过数据来表示的。数据经过加工处理后,可以具有知识性并对人类活动产生决策作用,从而形成信息。用数据符号表示信息,其形式通常有三种:数值型数据,即对客观事物进行定量记录的符号,如体重、年龄、价格等;字符型数据,即对客观事物进行定性记录的符号,如姓名、单位和地址等;特殊型数据,如声音、视频、图像等。

2) 知识与信息

信息不一定是知识,但知识肯定是信息,信息是所有知识的来源。信息作为资源,未必都有用,只有把各种有用的信息进行系统的组织才能成为知识,而这些知识经过反复实践、完善和提高,才形成科学知识。

知识反映的是这个世界的状况、变化和规律,从这个角度上讲,知识就是信息。知识是从常识信息中提炼出来的,反映的是比常识信息本身更本质的东西。例如,每天太阳升起,普照大地,泽被万物,这只是常识。天文学正是从太阳每天都升起这类常识中汲取营养并发展起来的,天文学表述的是天体的运行规律。所以知识是信息,但不能说信息就是知识。

信息与知识两者具有极其密切的联系,又有着重要的区别。信息是知识的原材料,知识是由信息提炼出来的抽象产物。这既反映了知识与信息之间的本质联系,也彰显了它们之间的区别。

1.1.3 信息社会

从哲学的角度来看,世界的本原是物质,物质的存在形式是运动。物质是无所不在的,物质的运动也是无所不在的;物质是多样的,物质的运动也是多样的。信息是物质运动的反映,这决定了信息也是无所不在的,也是多种多样的。这个世界是信息的,我们的世界是信息世界。

科学技术的发展使人类正在进入一个新的时代,这个时代的主要特征之一是对信息的需求和利用变得空前重要,因此有人称它为信息时代。我们现在所处的社会则被称为信息社会或信息化社会。

信息社会是信息起主要作用的社会。在现代社会,信息的重要性前所未有地增加了。信息成了重要的资源和关键的因素。在商业领域,竞争对手们争相扩大生产、提高效率、节省成本、加快资本循环、采用新的技术。在这一过程中,经理们发现信息成了他们制胜的法宝,企业如果掌握了相应的信息,就能比对手领先一步,采取恰当的行动,获取更大的利润,在竞争中取得更有利的地位。在竞争中,如何掌握信息、掌握先机,成为获胜的重要因素。过去,在军事领域中,一般决定战争胜负的是双方的军事实力,但在现代社会,军事力量已经不能仅仅用拥有多少坦克、飞机、士兵来衡量,还要看制信息权如何,也就是军队是否能及时、准确、全面地获取战场信息,并及时地传递信息、处理信息和利用信息。信息科技的发展是信息社会发展的主线。

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。

信息作为一种资源自古就有,人类也是自古以来就在利用信息。语言是思维的工具,也是传播的载体,人类使用大脑存储和处理信息,使用语言交流和传播信息,人类的信息活动也从具体走向抽象。

信息处理技术被人们称为“第五次信息技术革命”,它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。从1946年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 60 多年,这硕大无比的“怪物”竟经历了从大型机、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。

由上可见,在各种生产、科学研究和社会活动中,无处不涉及信息的交换和利用。迅速获取信息,正确处理信息,充分利用信息……信息的重要性不言而喻。因而人们把我们的社会叫做信息社会,把我们的时代叫做信息时代。

我们虽然已经处于信息社会,但世界各国的社会信息化程度还不尽相同。社会信息化和信息社会化是一个迅速而又长远的过程。信息化社会是以信息为社会发展的基本动力,以信息技术为实现信息化社会的手段,以信息经济为社会存在和发展的主导经济,以信息文化改变着人类的教育、生活和工作方式以及价值观念和时空观念的新兴社会形态;信息社会化主要从公共信息公开和共享以及社会文明程度方面进行表征。

1.2 信息的数据表示

1.2.1 数据的有效表示

信息可以用各种合适的方式加以表示和传输。信息表示与传输首先需要对信息进行编码,编码就是有目的的信息变换。编码的历史可以追溯到人类文明的启蒙时代。“结绳记事”其实就是最简单、最原始的一种编码。在绳子上打一个结或做一个记号,就表示记录了某一特定事件,包罗万象的大千世界,从此有了分类、识别、标记、定名……继而诞生了文字。今天,电话号码、汽车编号、身份证号码、商品号码以及这本《大学 IT》的书号等铺天盖地的编码形成了信息世界的直观形态。

大家都知道烽火戏诸侯的典故,敌人发动攻击的信息通过点燃烽火表示,诸侯得到消息就派兵救援;抗日战争中,若敌人偷袭,群众岗哨可通过放倒消息树来通知大家准备抗击。这告诉我们一个简单的事实——信息可以有不同的表示。在计算机中,信息是用“0”和“1”进行编码和存储的。

1. 进制

所谓进制也称计数制,是用一组固定符号和统一规则来表示数值的方法。人们在生产实践中用到了很多种进制,比如常用的十进制、钟表的六十进制等。所谓 r 进制,即逢 r 进位。一个数制所包含的数字符号的个数称为基数。如十进制含 0~9 十个数字符号,其基数为 10;二进制基数为 2,只有 0、1 两个数字;十六进制基数为 16,即 0~9 和 A~F(分别表示 10~15)。为区分不同数制的数,约定:对于任一 r 进制的数 n ,记作: $(n)_r$ 。例如, $(1001)_2$ 、 $(70)_8$ 和 $(15)_{16}$ 分别表示二进制数、八进制数和十六进制数。不用括号及下标的数,默认为十进制数,如 15。人们也习惯在数后面加上对应进制英文的首字母 D(十进制)、B(二进制)、O(八进

制)、H(十六进制)来表示其前面的数所采用的数制。

除“基数”概念外,“位权”也是进制中的一个重要概念,它是指数码在不同位置上的权值。在进位计数制中,处于不同数位的数码代表的数值不同。例如,十进制数111,个位数上的1权值为 10^0 ,十位数上的1权值为 10^1 ,百位数上的1权值为 10^2 。

2. 比特

比特来自英文bit,是二进制位(Binary Digit)的缩略写法。比特代表一个位,只有两个可能状态,常用“0”或“1”表示。二进制在物理上容易实现,目前计算机中一般使用二进制。比特是计算机能处理的数据的最小单位,计算机里各种各样的数据都是用比特表示的。计算机在存储数据时,常常把8位二进制数视为一个存储单元或称做一个字节。用 2^n 来计算存储容量,把 2^{10} (即1 024)个存储单元称为1 KB;把 2^{10} K(即1 024 K)个存储单元称为1 MB;把 2^{10} M(即1 024 M)个存储单元称为1 GB。

1.2.2 常见数据的表示

1. 数值数据的编码表示

可以用二进制表示数值数据。烽火台点燃我们可以用1表示,反之用0表示,那么,4座烽火台的排列组合可以表示16种敌军情,0000表示没有敌人,0001表示一队敌人,0010表示两队敌人,……1111表示15队敌人。

在计算机内用于存储数据的逻辑部件有两种可能状态,即高电位和低电位两种状态,我们不妨把它们看成是1状态和0状态,那么十进制数字在计算机中存储,实际上也是用二进制存储。但二进制无法表示负数,人们就找了替代方案,用0和1的序列来表示负数。这种替代方案也使用0和1,因而称为二进制编码。

2. 字符数据的编码表示

在计算机中,数值数据用二进制表示,但人们使用计算机还涉及各种字符,这些字符也无法直接在计算机中表示和存储。人们同样用一系列0和1串表示字符,这称为字符编码。计算机的字符编码是公开的,这有利于计算机之间共享和通信。目前采用的字符编码主要是ASCII码,它是American Standard Code for Information Interchange(美国标准信息交换代码)的缩写,已被国际标准化组织采纳,作为国际通用的信息交换标准代码。ASCII码是一种西文机内码,有7位ASCII码和8位ASCII码两种,7位ASCII码称为标准ASCII码,8位ASCII码称为扩展ASCII码。7位标准ASCII码用一个字节(8位)表示一个字符,并规定其最高位为0,实际只用到7位,因此可表示128个不同字符。Unicode是16个比特的字符编码,可以表示65 000多个字符,理论上可以表示任何在使用的语言字符,这对于通信和商业非常有用。

3. 汉字的编码表示

(1) 汉字交换码:由于汉字数量极多,一般用连续的两个字节(16个二进制位)来表示一个汉字。1980年,我国颁布了第一个汉字编码字符集标准,即GB2312-80《信息交换用汉字编码字符集基本集》,该标准编码简称国标码,是我国大陆地区及新加坡等海外华语区通用的汉字交换码。GB2312-80收录了6 763个汉字,以及682个符号,共7 445个字符,这奠定了中文信息处理的基础。

(2) 汉字机内码:国标码GB2312不能直接在计算机中使用,因为它没有考虑与基本的信息交换代码ASCII码的冲突。为了能区分汉字与ASCII码,在计算机内部表示汉字时把交换码(国标码)两个字节最高位改为1,称为机内码。

(3) 汉字字形码：所谓汉字字形码实际上就是用来将汉字显示到屏幕上或打印到纸上所需要的图形数据。汉字字形码记录汉字的外形，是汉字的输出形式。通常有两种方法——点阵法和矢量法，分别对应两种字形编码——点阵码和矢量码。点阵码是一种用点阵表示汉字字形的编码，它把汉字按字形排列成点阵，一个 16×16 点阵的汉字要占用 32 个字节，一个 32×32 点阵的汉字则要占用 128 个字节，而且点阵码缩放困难并容易失真。矢量码效率高，可靠性强。

(4) 汉字输入码：将汉字通过键盘输入到计算机中采用的代码称为汉字输入码，也称为汉字外部码（外码）。汉字输入码的编码原则是应该易于接受、学习、记忆和掌握，码长尽可能短。目前我国的汉字输入码编码方案已有上千种，但是在计算机上常用的只有几种。根据编码规则，这些汉字输入码可分为流水码、音码、形码和音形结合码四种。

4. 图形和视频表示

图形不同于数字和字符，它在计算机中也必须表示成“0”和“1”串，才能存储和传输。图形分为两种：位图和矢量图。我们知道，计算机屏幕是由一个个像素组成的，像素的黑白不同则组成了一幅幅计算机黑白画面。因而我们存储黑白位图时，只需存储计算机屏幕上所有像素的黑白即可，可以用“1”代表白，“0”代表黑。人们肉眼分辨不出屏幕是由一个个像素组成的，是因为像素极小而数目极多。屏幕的水平分辨率即表示屏幕有多少像素列，垂直分辨率表示有多少像素行，将屏幕的垂直和水平分辨率相乘就得出像素总数。分辨率 800×600 的屏幕有 480 000 个像素，即需要 480 000 个比特来存储一个黑白位图。

某大学有名大学生利用上述信息表示原理，独特地表达了他的爱的信息。在学生公寓大楼里，他到每个宿舍送了糖果和可口可乐，希望大家按他的要求在固定的时间打开或关闭照明灯。奇迹出现了：在约定的时间，偌大宿舍楼的窗户灯光组成了醒目的“I♥U”。

黑白位图只由黑白像素组成，看起来不太真实，因此人们用不同程度的灰色来表示位图，这种图像看起来像黑白照片，比黑白位图真实。灰度的级别越多，图像越真实。计算机常用 256 级灰度显示图像，每个像素可以选择 256 级灰度中的任何一种，但是这也增加了存储所需的容量。计算机存储 256 级灰度需要 8 比特，而黑白图像仅需一个比特，因此，计算机存储 256 级灰度的图像所需容量是同样大小黑白图像的 8 倍。现在图像更多是彩色的，有 16 色、256 色和 24 位图像（真彩色图像）等，像素分别可以选择 16 种、256 种和 2^{24} 种色彩，每个像素需要 4 个比特、8 个比特和 24 个比特来存储，因而所需存储容量分别是黑白图像的 4 倍、8 倍和 24 倍。这些彩色图像的存储空间更大，而传输时间也更长。

人们采用许多技术来减少图像的传输时间和存储容量。一种方法是压缩文件大小；另一种方法是采用“抖动”技术。位图使用一系列比特表示，因而可以修改和编辑单个像素。图 1-1 所示是利用微软的 Paint 编辑位图的窗口界面。其他常用的工具有 Photoshop、Paintbrush 等，它们都提供了一系列工具用于位图的编辑处理，如剪切、复制和粘贴等。

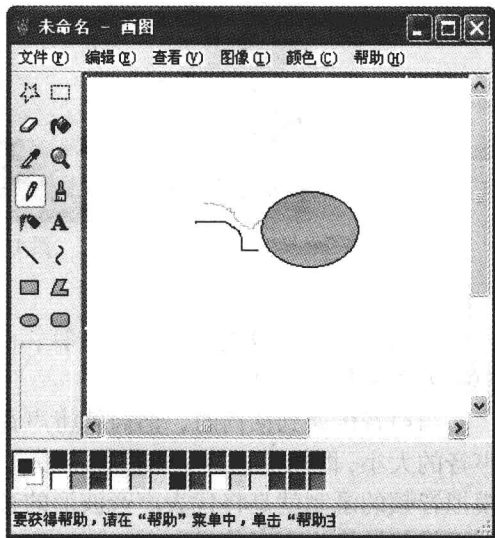


图 1-1 位图表示示例

矢量图由一系列可重构图像的指令组成。在创建矢量图时，计算机把不同颜色的线和图形转换成能够重构的指令，计算机存储的是指令而非真正的图像。矢量图的结构决定了它的优点：矢量图的存储空间比位图小，其大小取决于图像的复杂度。图像越复杂，指令越多，占用的存储空间就越大。使用矢量图像软件可以方便地修改图像。矢量图实际上是由多个“对象”构成的，每个对象都可以单独地拉伸、拖动、变形和删除。另外，位图放大后，细节就会令人感觉很粗糙，而矢量图放大后细节仍然精细，如图1-2所示的卡通狗，中图为位图放大，右图为矢量图放大。包含矢量图的文件扩展名常为 .wmf、.dxf、.mgx 和 .cgm。



图 1-2 矢量图示例

数字视频是由一系列的帧组成的，每一帧都是一幅静止的位图图像，视频每秒钟显示30帧图像。位图是静态的，而视频是动态的，这是由于人眼看到每一幅图像时会产生视觉暂留。如果在视觉暂留退去之前显示下一幅图像，人们就会感觉到动态的效果，而不是静止的图像。要保证这种效果，两帧间切换的时间要小于视觉暂留时间，这与播放电影的原理相同。图1-3所示的是一架飞机在穿越音障的视频截图。



图 1-3 数字视频示例

5. 声音表示

人们不仅将计算机用于科学计算，也用于听歌曲等娱乐。我们处理声音的方式有两种：模拟方式和数字方式。

当声音传递到话筒时，话筒就将声音引起的振动转换为电信号，并用不同的电压值表示声音的大小，即用电压波动来模拟声波，这种表示声音的电信号称为模拟音频，如图1-4所示。模拟音频的录制就是将代表声音波形的音频信号存储到适当的媒体上，如磁带或唱片，播放时再将其转换为声音信号。

音频信号是一种模拟量，而在计算机中，所有的信息都是以二进制的数字形式表示的，因