



普通高等教育“十一五”国家级规划教材

计算机基础教程

湖北省教育厅 组编



 华中师范大学出版社



普通高等教育“十一五”国家级规划教材

计算机基础教程

(第四版)

主 编 胡金柱 何婷婷 杨 青 刘宝忠

副主编 何统洲 纪 鹏 王 诚 朱 丽

主 审 孙俊逸(湖北省高教学会高校计算机教育专业委员会)

编 委 (以姓氏笔画为序)

王 诚(湖北美术学院)

朱 丽(武汉音乐学院)

朱 玲(郧阳师范高等专科学校)

朱晓钢(湖北大学)

纪 鹏(湖北理工学院)

祁文青(湖北理工学院)

杨 静(郧阳师范高等专科学校)

何统洲(郧阳师范高等专科学校)

汪 毅(武汉设计工程学院)

赵甫哲(华中师范大学)

袁 涌(湖北理工学院)

曹加恒(武汉大学珞珈学院)

华中师范大学出版社

内 容 提 要

本书是为大专院校非计算机专业的学生编写的计算机基础教材。

全书分为五章:第1章介绍计算机的产生、发展和新技术,以及计算机系统的结构和信息安全知识;第2章介绍 Windows 7 的主要功能和基本使用方法;第3章介绍计算机网络的基本概念和 Internet 的使用方法,包括网络设置、浏览器的使用、电子邮件等;第4章介绍 Office 2010 的使用方法,包括 Word 2010、Excel 2010 和 PowerPoint 2010 的使用方法;第5章是多媒体及其常用软件的使用方法,介绍了多媒体的概念和几种常用软件的使用,包括当前主流的图形图像处理软件 Photoshop、声音处理软件 Audition、动画创作与编辑工具软件 Flash。每章都有案例分析和习题,以便读者复习和练习。

本书所讲授的内容丰富、知识点层次清晰、示例图文并茂、叙述通俗易懂,适合作为高等院校各类本科生的计算机教材,也可供计算机初学者学习参考。

新出图证(鄂)字 10 号

图书在版编目(CIP)数据

计算机基础教程/胡金柱等主编. —4 版. —武汉:华中师范大学出版社,2015. 8
ISBN 978-7-5622-7046-1

I. ①计… II. ①胡… 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(2015)第 148961 号

计算机基础教程(第四版)

© 湖北省教育厅 组编

胡金柱 何婷婷 杨 青 刘宝忠 主编

责任编辑:罗 挺 胡小忠

编辑室:第二编辑室

出版发行:华中师范大学出版社

社址:湖北省武汉市珞喻路 152 号

销售电话:027-67863426/67863280

邮购电话:027-67861321

网址:<http://press.ccnu.edu.cn/>

印刷:武汉兴和彩色印务有限公司

开本:787mm×1092mm 1/16

版次:2015 年 8 月第 4 版

定价:32.00 元

责任校对:刘 峥

电话:027-67867362

邮编:430079

传真:027-67863291

电子信箱:hscbs@public.wh.hb.cn

督印:王兴平

印张:17

印次:2015 年 8 月第 1 次印刷

封面设计:胡 灿

字数:420 千字

欢迎上网查询、购书

敬告读者:欢迎举报盗版,请打举报电话 027-67861321。

第一版编写人员

主 编：胡金柱 苏光奎 杨先麟 汪白云
编 委：（以姓氏笔画为序）

王舜燕 朱少林 任卓梅 刘冠蓉
杨先麟 金汉均 袁志勇 顾保磊
游小明

第二版编写人员

主 编：胡金柱 庞丽萍 杨健霭 刘大革
编 委：（以姓氏笔画为序）

刘大革 刘卫国 刘春燕 肖 明
杨 青 金汉均 胡必鑫 夏玉勤
袁国刚 彭绪富

第三版编写人员

主 编：胡金柱 兰顺碧 刘宝忠 王建勇
主 审：孙俊逸 殷 朴 何婷婷
编 委：（按编写章节排序）

刘宝忠 王建勇 何统洲 李成海
兰顺碧 马学军 胡金柱 赵甫哲
朱晓钢 唐霁虹 李梁军

第四版前言

随着计算机的日益普及和网络技术的迅速发展,计算机技术的应用能力已经成为当今大学生必备的能力,因此,高等院校计算机基础教育显得尤为重要。计算机基础教育不仅能让学生掌握计算机的基本知识和基本使用方法,而且能启发学生对先进科学技术的追求,使学生思维敏捷、思路开阔。

本教材由湖北省教育厅组织编写,被选为普通高等教育“十一五”国家级规划教材,是高等院校非计算机专业计算机基础公共课教材的第四版。本次修订仍然遵循前三版的基本原则,即相对先进性、实用性、科学性和简单易学性,同时照顾零起点的读者。第四版在第三版的基础上加入了目前的新技术和新应用,如大数据,更新了软件版本,介绍的操作系统由 Windows XP 改为 Windows 7, Office 2003 更新为 Office 2010,学生通过本教材的学习,可以掌握计算机基本的知识,培养学习计算机技术的兴趣,同时为进一步学习计算机技术打下基础。

教材以计算机知识为基础,以实际应用为主导,参考当前国内外最新资料,理论结合实践。全书分为五章,第 1 章介绍了计算机的产生、发展和新技术,介绍了计算机系统的结构以及信息安全知识,该章内容是广大读者必须学习和了解的基本原理和常识;第 2 章介绍操作系统,主要介绍了 Windows 7 的主要功能和基本使用方法,包括资源管理器的使用、控制面板的使用等;第 3 章介绍计算机网络,介绍了计算机网络的基本概念和 Internet 的使用方法,包括网络设置、浏览器的使用、电子邮件等;第 4 章介绍 Office 的使用方法,包括 Word 2010、Excel 2010 和 PowerPoint 2010 的使用方法,同时还简单介绍了国产办公应用软件 WPS;第 5 章介绍多媒体及其常用软件的使用方法,介绍了多媒体的概念和几种常用软件的使用方法,包括当前主流的图形图像处理软件 Photoshop、声音处理软件 Audition、动画创作与编辑工具软件 Flash。

为了便于教师组织教学,也使学习者更好地掌握理论知识和实际应用,本教材沿用了第三版的模式,将教材分为主教材《计算机基础教程》和配套的实验教材《计算机基础教程上机指导》。主教材重点讲解计算机基本知识、基本原理和一般操作,每一章都有案例分析,帮助学习者深入理解计算机知识,实验教材为学习者提供实验方法和实验具体指导,主教材和实验教材相辅相成,为学习者提供帮助。

本教材的作者都是长期从事计算机基础教学工作的教师,有着丰富的教学经验。第 1 章由武汉工程大学的刘宝忠编写;第 2 章由湖北美术学院的王诚编写;第 3 章由鄱阳师范高等专科学校的朱玲、杨静共同编写,鄱阳师范高等专科学校的何统洲负责本章的审阅和修改工作;第 4 章由湖北理工学院的袁涌、祁文青和武汉音乐学院的朱丽共同编写,湖北理工学院的纪鹏负责本章的审阅和修改工作;第 5 章由朱丽和王诚共同编写。全书由华中师范大

学的胡金柱、何婷婷、杨青统稿、定稿。

本书在编写过程中,得到了湖北省各院校同仁的热情帮助和支持,特别是得到了华中师范大学、武汉工程大学、湖北美术学院、武汉音乐学院、湖北理工学院、鄱阳师范高等专科学校的大力支持,得到了华中师范大学出版社老师的热情帮助,在此表示最诚挚的谢意。

目前,计算机技术的发展日新月异,计算机技术的应用越来越广,由于编者水平有限,书中难免存在诸多不妥之处,恳请广大读者批评指正。

— 编 者

2015 年 6 月

目 录

第 1 章 计算机基础知识概论	1
1.1 计算机系统概述	1
1.1.1 计算机的发展、特点与分类	1
1.1.2 计算机的应用	3
1.2 计算机中的数制与编码	6
1.2.1 二进制与其他数制	6
1.2.2 常用信息编码	9
1.3 计算机系统的基本组成	11
1.3.1 冯·诺依曼计算机的基本原理	11
1.3.2 计算机系统的基本结构	12
1.3.3 计算机硬件系统	13
1.3.4 计算机软件系统	14
1.3.5 计算机系统的主要技术指标	14
1.3.6 微型计算机系统的组成	15
1.4 计算机病毒及其防治	19
1.4.1 计算机病毒简介	19
1.4.2 常用杀毒软件简介	22
1.5 信息安全与知识产权	23
1.5.1 信息安全	24
1.5.2 职业道德与软件知识产权	26
1.5.3 计算机犯罪与防范	30
习题	32
第 2 章 操作系统应用	33
2.1 操作系统概述	33
2.1.1 操作系统的类型	33
2.1.2 操作系统的基本功能	34
2.1.3 典型的操作系统简介	35
2.2 Windows 7 操作系统的安装与恢复	37
2.2.1 Windows 操作系统的安装	37
2.2.2 Windows 操作系统的快速恢复(系统还原)	38
2.3 Windows 7 的基本操作	39

2.3.1	Windows 操作系统的窗口	39
2.3.2	快捷键的使用	43
2.4	Windows 7 文件管理	43
2.4.1	文件及文件夹的基本概念	43
2.4.2	Windows 7 系统文件夹	45
2.4.3	Windows 7 文件与文件夹操作	45
2.5	Windows 7 的应用程序管理	49
2.5.1	应用程序的安装、运行与删除	49
2.5.2	任务管理器	52
2.6	Windows 7 的系统管理	53
2.6.1	Windows 7 控制面板的使用	53
2.6.2	系统安全设置	59
2.6.3	磁盘管理	65
2.7	案例分析	67
	习题	68
第3章	计算机网络及其应用	70
3.1	计算机网络概述	70
3.1.1	计算机网络的定义与功能	70
3.1.2	计算机网络的分类	71
3.1.3	计算机网络的体系结构	72
3.2	计算机网络的组成	73
3.2.1	网络硬件系统	74
3.2.2	网络软件系统	75
3.3	Internet 基础	76
3.3.1	Internet 概述	76
3.3.2	IP 地址和域名地址	77
3.3.3	Internet 接入方法	80
3.3.4	移动网络简介	82
3.4	常用网络工具软件	85
3.4.1	浏览器的使用	85
3.4.2	收发电子邮件	92
3.4.3	搜索引擎	95
3.4.4	知网的使用	98
3.4.5	文件的上传与下载	101
3.4.6	云盘的使用	103

3.5 无线局域网与物联网及其应用	106
3.5.1 无线局域网简介.....	106
3.5.2 无线局域网的应用.....	107
3.5.3 物联网的概念及其应用.....	108
3.6 案例分析	111
习题.....	118
第4章 办公自动化软件及其应用.....	120
4.1 Word 2010 文字处理系统	120
4.1.1 Word 2010 基本操作.....	120
4.1.2 Word 2010 表格处理.....	133
4.1.3 Word 2010 图文混排.....	137
4.1.4 科学论文排版.....	138
4.1.5 Word 2010 审阅.....	145
4.1.6 Word 2010 打印输出.....	148
4.2 电子表格系统	149
4.2.1 Excel 2010 的基本操作.....	149
4.2.2 Excel 2010 公式与函数应用.....	153
4.2.3 Excel 2010 数据图表的操作.....	158
4.2.4 数据统计.....	163
4.2.5 数据透视表.....	173
4.2.6 Excel 2010 多个工作表的联动操作.....	175
4.3 演示文稿软件	176
4.3.1 PowerPoint 2010 的基本操作	176
4.3.2 PowerPoint 2010 演示文稿的多媒体效果	184
4.3.3 PowerPoint 2010 演示文稿的动画效果	192
4.3.4 PowerPoint 2010 演示文稿的放映设置与发布	198
4.4 WPS 中文办公系统.....	201
4.4.1 WPS 简介	201
4.4.2 WPS 协同办公	202
4.5 案例分析	205
4.5.1 图文混排.....	205
4.5.2 利用邮件合并功能批量制作名片.....	215
4.5.3 Excel 实例	217
4.5.4 PowerPoint 超链接交互功能	222
习题.....	223

第5章 多媒体应用基础	225
5.1 多媒体技术基础知识	225
5.1.1 计算机与多媒体技术.....	225
5.1.2 多媒体信息的基本要素.....	226
5.1.3 多媒体技术的特点与应用.....	228
5.1.4 多媒体系统的组成.....	229
5.2 图形图像信息处理	229
5.2.1 图形图像处理基础知识.....	229
5.2.2 Photoshop 工作界面	231
5.2.3 Photoshop 基本操作	232
5.2.4 Photoshop 图像处理	233
5.2.5 Photoshop 图像合成输出	235
5.3 声音信息处理	236
5.3.1 声音概述.....	236
5.3.2 音频处理技术.....	237
5.3.3 常用的音频处理软件.....	238
5.3.4 Adobe Audition 音频编辑软件简介	239
5.4 多媒体动画制作	248
5.4.1 动画制作的基本概念与文件格式.....	248
5.4.2 二维动画制作软件 Flash 工作界面	250
5.4.3 Flash 动画制作的基本方法	252
5.5 视频信息处理	253
5.5.1 视频概述.....	253
5.5.2 视频处理技术.....	254
5.5.3 常用的视频编辑软件.....	255
5.5.4 Corel Video Studio 视频编辑	256
习题.....	260
参考文献.....	262

第 1 章 计算机基础知识概论

电子计算机(Electronic Computer)又称电脑(Computer),诞生于 20 世纪 40 年代。几十年来,计算机的发展突飞猛进,其应用已深入到社会的各个领域以及人们生活的各个方面。学习和了解计算机的基础知识,已经成为人们的必需。

本章主要介绍计算机的发展、特点与分类,介绍计算机系统的基本结构、数制及数制间的转换方法,还介绍微型计算机系统的基本组成与计算机病毒的相关知识。

1.1 计算机系统概述

1946 年世界上第一台电子计算机问世以来,尤其是微型计算机的出现和计算机网络的发展,有力地推动了信息社会的发展。计算机的发明及应用发挥和激发了人类的创造力,标志着人类文明的发展进入了一个崭新的阶段,同样,计算机技术也随着社会生产和科学技术的进一步发展而不断发展。

1.1.1 计算机的发展、特点与分类

1. 计算机的发展历史

(1) 电子计算机的发展

半个多世纪以来,根据电子计算机所采用的电子器件,一般将电子计算机的发展历史分成四个阶段,其主要特点如表 1-1 所示。

表 1-1 各代电子计算机比较

	第一代 (1946—1956 年)	第二代 (1957—1964 年)	第三代 (1965—1970 年)	第四代 (1971 年至今)
电子器件	电子管	晶体管	中、小规模集成电路	大规模、超大规模集成电路
主存储器	磁芯、磁鼓	磁芯、磁鼓	磁芯、磁鼓、半导体存储器	半导体存储器
辅助存储器	磁带、磁鼓	磁带、磁鼓、磁盘	磁带、磁鼓、磁盘	磁盘、磁带、光盘
处理方式	机器语言,汇编语言	作业连续处理,编译语言	多道程序,实时处理	网络结构,实时、分时处理
运算速度 (次·s ⁻¹)	5000~4 万	几十万~一百万	一百万~几百万	几百万~千万亿
代表机种	ENIAC、EDVAC、IBM 705	IBM 7090、CDC 6600	IBM 360、PDP 11、NOVA 1200	IBM370、CRAY II、VAX11、IBM PC

(2) 计算机的发展趋势

进入 21 世纪以来,计算机技术的发展非常迅速,产品不断升级换代,融入了各项新技术,使得计算机功能越来越强。计算机在各个领域的广泛应用,也积极地推动了社会的发展和科学技术的进步,促进了计算机技术的更新和发展,因而产生了新一代计算机,可称为第五代计算机,主要是用于支持知识库的智能计算机、神经网络计算机和生物计算机等。

新一代计算机主要是将信息采集、存储、加工、通信和人工智能结合在一起的智能计算机,将突破传统计算机的结构模式,注重智能化,即对数据进行处理的同时还具备模拟的功能。因此,未来计算机的发展趋势是微型化、巨型化、网络化、智能化。

①微型化。随着计算机技术的不断发展,计算机的体积越来越小,功能越来越强,如笔记本计算机、掌上计算机等便携式计算机。

②巨型化。巨型化是指最大、最快、最贵的计算机的运算速度越来越快。例如,目前世界上运行最快的计算机速度达到 5 亿亿次每秒浮点运算。生产巨型计算机的公司有美国的 Cray 公司和 TMC 公司、日本的富士通公司和日立公司等。我国研制的天河二号计算机是截至 2014 年底世界运算速度最快的计算机。

③智能化。计算机依据不确定的输入作出决定,它模仿人脑的工作方式,具有直观判断和处理不完整的模糊信息的能力,甚至有接近人的审美和情感的能力。也就是说,计算机工作时只需要告诉它“做什么”,而不必“手把手”教它“怎么做”。目前科学家们正在采取人工智能和神经网络的方法开发智能计算机。

④网络化。随着计算机技术与通信技术的飞速发展,全球网络化时代已经到来,可以说有了计算机就拥有了网络资源。

2. 计算机的特点

(1)运算速度快:计算机的运算速度已从早期的每秒几千次发展到现在的每秒亿亿次。如此快的计算速度,不仅极大地提高了工作效率,而且使许多极复杂的科学问题得以解决。例如,外国的一位数学家花了 15 年的时间把圆周率 π 的值算到小数点后 707 位,而这项工作现代计算机只用不到 1 小时就完成了。

(2)计算精度高:尖端科学技术的发展往往需要高度准确的计算能力,只要电子计算机内用以表示数值的位数足够多,就能提高运算精度。一般的计算工具只有几位有效数字,而计算机的有效数字可以精确到十几位、几十位甚至数百位,这样就能精确地进行数据计算并且表示数据的计算结果。

(3)存储功能强:计算机具有存储信息的存储装置,可以存储大量的数据,当需要时可准确无误地取出来。计算机这种存储信息的“记忆”能力,使它成为信息处理的有力工具。

(4)具有逻辑判断能力:计算机既可以进行数值运算,也可以进行逻辑运算。计算机可以对文字或符号进行判断和比较,也可以进行逻辑推理和证明,这是其他任何计算工具无法相比的。

(5)具有自动运行能力:计算机不仅能存储数据,还能存储程序。计算机内部操作是按照人们事先编制的程序一步一步自动进行的,不需要人工操作和干预,这是计算机与其他计算工具最本质的区别。

3. 计算机的分类

通常,人们用“分代”来表示计算机在纵向的历史中的发展情况,而用“分类”来表示计算

机在横向的不同领域的发展、分布和使用情况。根据美国电气和电子工程师协会(IEEE)1989年提出的标准来划分,可以将计算机分成巨型机、小巨型机、大型主机、小型主机、工作站和个人计算机6类。

(1)巨型机(Supercomputer)。也称为超级计算机,它占地最大、价格最贵、功能最强、浮点运算速度最快(2004年IBM的蓝色基因达到135.5万亿次每秒)。目前,只有少数国家的几家公司能够生产,多用于战略武器(如核武器和反导武器)的设计、空间技术、石油勘探、天气预报以及社会模拟等领域。巨型机的研制水平、生产能力及其应用程度,已成为衡量一个国家经济实力和科技水平的重要标志。

(2)小巨型机(Minisupercomputer)。又称为小型超级电脑或桌上型超级计算机,出现于20世纪80年代中期。它的功能低于巨型机,运算速度能达到1TELOPS(即每秒10亿次),价格只有巨型机的十分之一。

(3)大型主机(Mainframecomputer)。也称为大型电脑,包含国内通常所说的大、中型机。其特点是大型、通用、可靠、安全、向后兼容,具有很强的处理和管理能力,主要用于大银行、大公司、规模较大的高校和科研院所。在计算机向网络化发展的今天,大型主机仍有其生存空间。

(4)小型机(Minicomputer或Minis)。小型机结构简单,可靠性高,成本较低,不需要经过长期培训即可维护和使用,对于广大中、小用户较为适用。

(5)工作站(Workstation)。工作站是功能介于PC机和小型机之间的一种高档微型计算机,运算速度快,具有较强的联网功能,用于特殊领域,如图像处理、计算机辅助设计等。它与网络系统中的工作站在用词上相同,而含义不同。网络上的工作站泛指联网用户的节点,以区别于网络服务器,常常由一般的PC充当。

(6)个人计算机(Personal Computer,PC)。个人计算机即通常所说的电脑、微机或计算机,它出现于20世纪70年代,以其设计先进(总是率先采用高性能的微处理器MPU)、软件丰富、功能齐全、价格便宜等优势而拥有广大的用户,大大推动了计算机的普及应用。可以说,PC机无所不在、无所不用,除了台式型的,还有膝上型、笔记本、掌上型、手表型等类型。

1.1.2 计算机的应用

计算机的应用已渗透到社会的各行各业,正在改变着传统的工作、学习和生活方式,推动着社会的发展。

1. 科学计算(数值计算)

科学计算是指利用计算机来完成科学研究和工程技术中提出的数学问题的计算。在现代科学技术工作中,科学计算问题是大量的和复杂的。利用计算机的高速计算、大存储容量和连续运算的能力,可以解决人工无法解决的各种科学计算问题。

例如,建筑设计中为了确定构件尺寸,通过弹性力学导出了一系列复杂方程,长期以来,由于计算方法跟不上而一直无法求解。而计算机不但能求解这类方程,并且引起弹性理论研究的一次突破,出现了有限单元法。

2. 数据处理(信息处理)

数据处理是对各种数据进行收集、存储、整理、分类、统计、加工、利用、传播等一系列活

动的统称。据统计,80%以上的计算机主要用于数据处理,这类工作量大面广,是计算机应用的主要方向。

数据处理从简单到复杂已经历了三个发展阶段。

(1)电子数据处理(Electronic Data Processing,EDP),它以文件系统为手段,实现一个部门内的单项管理。

(2)管理信息系统(Management Information System,MIS),它以数据库技术为工具,实现一个部门的全面管理,以提高工作效率。

(3)决策支持系统(Decision Support System,DSS),它以数据库、模型库和方法库为基础,帮助决策者提高决策水平,改善运营策略的正确性与有效性。

目前,数据处理已广泛地应用于办公自动化、企事业计算机辅助管理与决策、情报检索、图书管理、电影电视动画设计、会计电算化等各行各业。信息正在形成独立的产业,多媒体技术使信息展现在人们面前的形式不仅是数字和文字,也有“声情并茂”的声音和图像信息。

3. 辅助技术(计算机辅助设计与制造)

计算机辅助技术包括 CAD、CAM 和 CAI 等。

(1)计算机辅助设计(Computer Aided Design,CAD)

(2)计算机辅助制造(Computer Aided Manufacturing,CAM)

计算机辅助制造是利用计算机系统进行生产设备的管理、控制和操作的过程。例如,在产品的制造过程中,用计算机控制机器的运行,处理生产过程中所需的数据,控制和处理材料的流动以及对产品进行检测等。使用 CAM 技术可以提高产品质量,降低成本,缩短生产周期,提高生产效率和改善劳动条件。将 CAD 和 CAM 技术集成,实现设计生产自动化,这种技术称为计算机辅助制造系统(CIMS),它的实现将真正做到无人化工厂(或车间)。

(3)计算机辅助教学(Computer Aided Instruction,CAI)

计算机辅助教学是利用计算机系统使用课件来进行教学。课件可以用制作工具或高级语言来开发制作,它能引导学生循环渐进地学习,使学生轻松自如地从课件中学到所需知识。CAI 的主要特色是交互教育、个别指导和因人施教。

4. 过程控制(实时控制)

过程控制是指利用计算机及时采集检测数据,按最优值迅速地对控制对象进行自动调节或自动控制。采用计算机进行过程控制,不仅可以大大提高控制的自动化水平,而且可以提高控制的及时性和准确性,从而改善劳动条件,提高产品质量及合格率。因此,计算机过程控制已在机械、冶金、石油、化工、纺织、水电、航天等部门得到广泛的应用。

例如,在汽车工业方面,利用计算机控制机床和整个装配流水线,不仅可以实现精度要求高、形状复杂的零件加工自动化,而且可以使整个车间或工厂实现自动化。

5. 人工智能(智能模拟)

人工智能(Artificial Intelligence,AI)是指计算机模拟人类的智能活动,诸如感知、判断、理解、学习、问题求解和图像识别等。现在人工智能的研究已取得不少成果,有些已开始走向实用阶段。例如,能模拟高水平医学专家进行疾病诊疗的专家系统,具有一定思维能力的智能机器人等。

6. 网络应用

计算机技术与现代通信技术的结合构成了计算机网络。计算机网络的建立,不仅解决

了一个单位、一个地区、一个国家中计算机与计算机之间的通信,各种软件和硬件资源的共享,也大大促进了国际间的文字、图像、视频和声音等各类数据的传输与处理。

7. 大数据技术

大数据(Big Data)是指那些超过传统数据库系统处理能力的数据,它的数据规模很大,对传输速度要求很高,或者其结构不适合传统的数据库系统。为了获取大数据中的价值,我们必须选择另一种方式来处理它。数据中隐藏着有价值的模式和信息,在以往需要相当的时间和成本才能提取这些信息。如沃尔玛和谷歌这类领先企业都要付高昂的代价才能从大数据中挖掘信息。而当今的各种资源,如硬件、云架构和开源软件使得大数据的处理更为方便和廉价,即使是在车库中创业的公司也可以用较低的价格租用云服务时间了。对于企业组织来讲,大数据的价值体现在两个方面:分析使用和二次开发。对大数据进行分析能揭示隐藏其中的信息。例如零售业中对门店销售、地理和社会信息的分析能提升对客户理解。对大数据的二次开发则是那些成功的网络公司的长项。例如 Facebook 通过结合大量用户信息,定制出高度个性化的用户体验,并创造出一种新的广告模式。这种通过大数据创造出新产品和服务的商业行为并非巧合,谷歌、雅虎、亚马逊和 Facebook 都是大数据时代的创新者。

(1) 大数据的 4V 特征

大量化(Volume)。企业面临着数据量的大规模增长。IDC 最近的报告预测称,到 2020 年,全球数据量将扩大 50 倍。目前,大数据的规模尚是一个不断变化的指标,单一数据集的规模范围从几十 TB 到数 PB 不等。简而言之,存储 1PB 数据需要两万台配备 50GB 硬盘的个人电脑。此外,各种意想不到的来源都能产生数据。

多样化(Variety)。一个普遍的观点认为,人们使用互联网搜索是形成数据多样性的主要原因,这一看法并不完全正确。数据多样性的增加主要是由于新型多结构数据,以及包括网络日志、社交媒体、互联网搜索、手机通话记录及传感器网络等数据类型造成。其中,部分传感器安装在火车、汽车和飞机上,每个传感器都增加了数据的多样性。

快速化(Velocity)。数据创建和移动的速度将更快。在高速网络时代,通过基于实现软件性能优化的高速电脑处理器和服务器的,创建实时数据流已成为流行趋势。企业不仅需要了解如何快速创建数据,还必须知道如何快速处理、分析并返回给用户,以满足他们的实时需求。根据 IMS Research 关于数据创建速度的预测,到 2020 年全球将拥有 220 亿部互联网连接设备。

价值(Value)。大量的不相关信息,浪里淘沙却又弥足珍贵,可用于对未来趋势与模式的可预测分析等。

(2) 大数据时代对生活、工作的影响

大数据除了经济方面的影响,同时也在政治、文化等方面产生深远的影响。大数据可以帮助人们开启循“数”管理的模式,也是我们当下“大社会”的集中体现,三分技术,七分数据,得数据者得天下。

大数据增加了对信息管理专家的需求。大数据的影响并不仅仅限于信息通信产业,也正在“吞噬”和重构很多传统行业,广泛运用数据分析手段管理和优化运营的公司其实质都是一个数据公司。麦当劳、肯德基以及苹果公司的旗舰专卖店都是在数据分析基础之上进行精准选址。而在零售业中,数据分析的技术与手段更是得到广泛的应用,传统企业如沃尔

玛通过数据挖掘重塑并优化供应链,新崛起的电商如卓越亚马逊、淘宝等则通过对海量数据的掌握和分析,为用户提供更加专业化和个性化的服务。

大数据经常含有一些详细的潜在的能够展示有关个人用户的信息,逐渐引起了人们对个人隐私的担忧,一些处理大数据的公司需要认真对待这个问题。美国天睿资讯的一个科学家提出,我们不应该简单地服从法律方面的隐私保护问题,这些远远不够,公司都应该遵从谷歌不作恶的原则,甚至更应该作出更积极的努力。

(3) 大数据时代的发展方向、趋势

ESM 国际电子商情针对 2013 年大数据应用现状和趋势的调查显示,被调查者最关注的大数据技术中,排在前五位的分别是大数据分析(12.91%)、云数据库(11.82%)、Hadoop(11.73%)、内存数据库(11.64%)以及数据安全(9.21%)。Hadoop 已不再是人们心目中仅有的大数据技术,而大数据分析成为最被关注的技术。从中可以看出,人们对大数据的了解已经逐渐深入,关注的技术点也越来越多。既然大数据分析是最被关注的技术趋势,那么大数据分析中的哪项功能是最重要的呢?排在前三位的功能分别是实时分析(21.32%)、丰富的挖掘模型(17.97%)和可视化界面(15.91%)。2012 年也曾做过类似的调查,当时选择丰富的挖掘模型(27.22%)比实时分析(19.88%)多 7.34%。短短一年时间内,企业对实时分析的需求激增,成就了很多以实时分析为创新技术的大数据厂商。从调查结果可以看出:企业在未来一两年中有迫切部署大数据的需求,并且已经从一开始的基础设施建设,逐渐发展为对大数据分析和整体大数据解决方案的需求。与此同时,大数据还面临人才缺乏的挑战,需要企业和高校联合起来,培养数据领域的复合型人才,帮助企业打赢这场“数据战”。

1.2 计算机中的数制与编码

数制也称计数制,是指用一组固定的符号和统一的规则来表示数值的方法。编码是采用少量的基本符号,选用一定的组合原则,以表示大量复杂多样的信息的技术。计算机处理的任何信息必须转换成二进制形式数据后才能由计算机处理、存储和传输。

1.2.1 二进制与其他数制

1. 数制的概念

对于不同的数制,它们的共同特点如下:

(1) 每种数制都有固定的符号集,如十进制数制的符号有 10 个:0,1,2,...,9;二进制数制的符号有两个:0 和 1。

(2) 数制都使用位置表示法,即处于不同位置的数符所代表的值不同,与它所在位置的权值有关。例如,十进制数 5555.555 可表示为:

$$5555.555 = 5 \times 10^3 + 5 \times 10^2 + 5 \times 10^1 + 5 \times 10^0 + 5 \times 10^{-1} + 5 \times 10^{-2} + 5 \times 10^{-3}$$

可以看出,各种进位记数制中的权的值恰好是基数的某次幂。因此,对任何一种进位记数制表示的数都可以写出按其权展开的多项式之和,任意一个 m 位整数, k 位小数的 r 进制数 N 可表示为:

$$N = \sum_{i=m-1}^{-k} D_i \times r^i$$

式中的 D_i 为该数制采用的基本数符, r^i 是权, r 是基数, 不同的基数表示不同的进制数。

2. 常用的几种数制

人们常用的是十进制数, 而计算机一般只能处理表二进制数, 编码中也采用八进制或十六进制数。表 1-2 给出了几种常用的进位数制。

表 1-2 几种常用的数制

	二 进 制	八 进 制	十 进 制	十 六 进 制
规则	逢二进一	逢八进一	逢十进一	逢十六进一
基数	$r=2$	$r=8$	$r=10$	$r=16$
数符	0, 1	0, 1, ..., 7	0, 1, ..., 9	0, 1, ..., 9, A, B, C, D, E, F
权	2^i	8^i	10^i	16^i
形式表示	B	O	D	H

3. 数制转换

(1) r 进制与十进制

将 r 进制数值转换为十进制数值, 其转换公式为:

$$N = \sum_{i=m-1}^{-k} D_i \times r^i$$

公式本身就提供了将 r 进制数转换为十进制数的方法。比如, 把二进制数转换为相应的十进制数, 只要将二进制数只出现 1 的数位权相加即可。

例如, 把二进制数 11010 转换成相应的十进制数:

$$(11010)_B = 1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 = (26)_D$$

例如, 把二进制数 100110.101 转换成相应的十进制数:

$$(100110.101)_B = 1 \times 2^5 + 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^{-1} + 1 \times 2^{-3} = (38.625)_D$$

其中, 下标 B 和 D 分别表示括号内是二进制数和十进制数。

(2) 十进制与 r 进制

将十进制数值转换为 r 进制数值, 整数部分和小数部分的转换方法是不相同的, 下面分别加以介绍。

①整数部分的转换。把一个十进制的整数不断除以基数 r , 取其余数(除 r 取余法), 就能够转换成以 r 为基数的数。例如, 为了把十进制的数转换成相应的二进制数, 只要把十进制数不断除以 2, 并记下每次所得余数(余数总是 1 或 0), 所有余数连起来即为相应的二进制数, 这种方法称为除 2 取余法。

例如, 把十进制数 25 转换成二进制数, 计算过程如图 1-1 所示。

所以 $(25)_D = (11001)_B$ 。

注意: 第一位余数是低位, 最后一位余数是高位。

②小数部分转换。要将一个十进制小数转换成 r 进制小数时, 可将十进制小数不断地乘以 r , 取其整

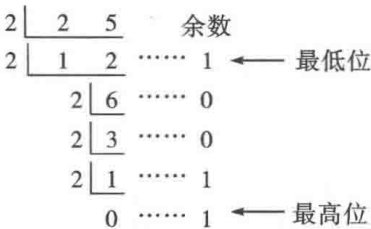


图 1-1 十进制整数转换成二进制数